Hanf normal form for first-order logic with unary counting quantifiers

Lucas Heimberg
Humboldt-Universität zu Berlin
heimberl@informatik.hu-berlin.de

Dietrich Kuske
Technische Universität Ilmenau
dietrich.kuske@tu-ilmenau.de

Nicole Schweikardt
Humboldt-Universität zu Berlin
schweika@informatik.hu-berlin.de

Abstract

We study the existence of Hanf normal forms for extensions $\mathrm{FO}(\mathbf{Q})$ of first-order logic by sets $\mathbf{Q} \subseteq \mathcal{P}(\mathbb{N})$ of unary counting quantifiers. A formula is in Hanf normal form if it is a Boolean combination of statements of the form "the number of witnesses y of $\psi(\overline{x}, y)$ belongs to Q" where $\mathsf{Q} \subseteq \mathbb{N}$ is a unary counting quantifer and ψ describes the isomorphism type of a local neighbourhood around its free variables $\overline{x}, y$.

We show that a formula from $\mathrm{FO}(\mathbf{Q})$ can be transformed into a formula in Hanf normal form that is equivalent on all structures of degree $\leqslant d$ if, and only if, all counting quantifiers occurring in the formula are ultimately periodic. This transformation can be carried out in worst-case optimal 3-fold exponential time.

In particular, this yields an algorithmic version of Nurmonen's extension of Hanf's theorem for first-order logic with modulo counting quantifiers.

1. Introduction

The intuition that first-order logic can only express local properties is formalised by the theorems by Hanf, by Gaifman, and by Schwentick and Barthelmann [5, 7, 8, 17]. All these results give rise to normal forms for first-order formulas.

Hanf's and Gaifman's theorem have found various applications in algorithms and complexity (cf., e.g., [3, 6, 10–12, 18, 19]). In particular, there are very general algorithmic meta-theorems stating that first-order model checking is fixed-parameter tractable for various classes of structures, and that the results of first-order queries against various classes of databases can be enumerated with constant delay after a linear-time preprocessing phase. In the context of such algorithms, questions about the efficiency of the normal forms have recently attracted interest (cf. e.g., [1, 2, 13]).

Notions of locality have also been developed for extensions of first-order logic, and they have found application in proving inexpressibility results for these logics (cf., e.g., [9, 12, 16]). When restricting attention to classes of finite structures of bounded degree, these locality notions also give rise to normal forms for the respective logics. Let us focus on the particular case of Hanf-locality:

Hanf's locality theorem for first-order logic implies that for every first-order sentence φ over a finite relational signature σ, and for every degree bound $d \in \mathbb{N}$, there exists a first-order sentence ψ that is equivalent to φ on all finite σ-structures of degree $\leqslant d$, such that ψ is a Boolean combination of statements of the form "the number of elements x whose r-neighbourhood has isomorphism type τ is $\geqslant k$". Such a formula ψ is said to be in *Hanf normal form*. A worst-case optimal algorithm for constructing ψ when given φ and d has been developed in [1].

In [16], Nurmonen extended Hanf's locality theorem to the extension of first-order logic by modulo counting quantifiers D_p (for positive integers p), where a formula of the form $\mathsf{D}_p\, y\, \psi(\overline{x}, y)$ states that the number of witnesses y for $\psi(\overline{x}, y)$ is divisible by p. As an easy consequence of Nurmonen's theorem, one obtains that for every sentence φ of first-order logic with modulo counting quantifiers, and for every degree bound $d \in \mathbb{N}$ there exists a first-order sentence with modulo counting quantifiers ψ that is equivalent to φ on all finite structures of degree $\leqslant d$, such that ψ is a Boolean combination of statements of the form "the number of elements x whose r-neighbourhood has isomorphism type τ is congruent k modulo p" and statements of the form "the number of elements x whose r-neighbourhood has isomorphism type τ is $\geqslant k$". Again, we say that ψ is in *Hanf normal form*.

For algorithmic applications, an effective procedure for computing ψ when given φ and d, would be highly desirable (cf., e.g., the use of Nurmonen's theorem in the full version of [15]). The proof of [16], however, does *not* lead to such an effective procedure. The two main questions which started the research whose results are presented in this paper are

(1) Is there an algorithmic version of Nurmonen's result?

(2) For which classes of unary counting quantifiers does an analogue of Nurmonen's result hold?

Answering question (2), our first main result provides a precise characterisation: A class $\mathbf{Q}$ of unary counting quantifiers permits "*Hanf normal forms*" (analogous to the ones obtained from Nurmonen's result) if, and only if, all counting quantifiers in $\mathbf{Q}$ are ultimately periodic.

Answering question (1), our second main result provides an algorithm which, when given a degree bound d and a formula φ of the extension of first-order logic with ultimately periodic unary counting quantifiers, transforms φ into a corresponding "Hanf normal form" ψ which is equivalent to φ on all structures of degree $\leqslant d$. This algorithm uses 3-fold exponential time and is worst-case optimal.

The rest of the paper is structured as follows. Section 2 fixes basic notations used throughout the paper. Section 3 gives precise statements of our two main results. Sections 4 and 5 are devoted to the proof of the "only if"-direction and the "if"-direction, respectively, of our characterisation of the sets of unary counting quantifiers that permit Hanf normal forms. Section 6 is devoted to the runtime analysis of our algorithm for transforming a given formula into Hanf normal form. Section 7 concludes the paper and points out directions for future work.

2. Preliminaries

We write $\mathcal{P}(S)$ to denote the power set of a set S. We write $\mathbb{N}$ for the set of non-negative integers, and we let $\mathbb{N}_{\geqslant 1} := \mathbb{N}\setminus\{0\}$. For all $m, n \in \mathbb{N}$ with $m \leqslant n$, we write $[m, n]$ for the set $\{i \in \mathbb{N} : m \leqslant i \leqslant n\}$, and we let $[m, n) := [m, n] \setminus \{n\}$. For a real number $r > 0$, we write $\log(r)$ to denote the logarithm of r with respect to base 2.

We say that a function f from $\mathbb{N}$ to the set $\mathbb{R}_{\geqslant 0}$ of non-negative reals is *at most k-fold exponential*, for some $k \in \mathbb{N}$, if there exists a number $c > 0$ such that for all sufficiently large $n \in \mathbb{N}$ we have $f(n) \leqslant T(k, n^c)$, where $T(k, m)$ is a tower of 2s of height k with an m on top (i.e., $T(0, m) = m$ and $T(k{+}1, m) = 2^{T(k,m)}$ for all $k, m \geqslant 0$).

For an ω-word $w = w_0w_1w_2 \cdots \in \{0, 1\}^\omega$ and a number $n \in \mathbb{N}$, we write $w[n]$ to denote the letter w_n in w at position n. For numbers $i, j \in \mathbb{N}$ with $i \leqslant j$, we write $w[i, j]$ for the (finite) word $w_iw_{i+1} \cdots w_j$. Similarly, we write $w(i, j]$ for the (finite) word $w_{i+1} \cdots w_j$. In particular, $w(i, i]$ is the empty word ϵ, and $w(j{-}1, j] = w[j]$.

Structures and formulas. A *signature* σ is a *finite* set of relation symbols and constant symbols. Associated with every relation symbol R is a positive integer $\mathrm{ar}(R)$ called the *arity* of R. A *σ-structure* $\mathcal{A}$ consists of a *finite* non-empty set A called the *universe* of $\mathcal{A}$, a relation $R^{\mathcal{A}} \subseteq A^{\mathrm{ar}(R)}$ for each relation symbol $R \in \sigma$, and an element $c^{\mathcal{A}} \in A$ for each constant symbol $c \in \sigma$. Note that according to this definition, all signatures and all structures considered in this paper are *finite*. We call a signature *relational*, if it only contains relation symbols.

We use the standard notation concerning first-order logic and extensions thereof, cf. [4, 12]. By $\mathrm{FO}[\sigma]$ we denote the class of all first-order formulas of signature σ, and by FO we denote the union of all $\mathrm{FO}[\sigma]$ for arbitrary signatures σ.

By $\mathrm{free}(\varphi)$ we denote the set of all *free variables* of φ. A *sentence* is a formula φ with $\mathrm{free}(\varphi) = \emptyset$.

We write $\varphi(\overline{x})$, for $\overline{x} = (x_1, \ldots, x_n)$ with $n \geqslant 0$, to indicate that $\mathrm{free}(\varphi) \subseteq \{x_1, \ldots, x_n\}$. If $\mathcal{A}$ is a σ-structure and $\overline{a} = (a_1, \ldots, a_n) \in A^n$, we write $\mathcal{A} \models \varphi[\overline{a}]$ to indicate that the formula $\varphi(\overline{x})$ is satisfied in $\mathcal{A}$ when interpreting the free occurrences of the variables $x_1, \ldots, x_n$ with the elements $a_1, \ldots, a_n$.

For a class $\mathfrak{C}$ of σ-structures, two formulas $\varphi(\overline{x})$ and $\psi(\overline{x})$ of signature σ are called *equivalent on* $\mathfrak{C}$ (for short: $\mathfrak{C}$-equivalent) if for all σ-structures $\mathcal{A} \in \mathfrak{C}$ and for all $\overline{a} \in A^n$ we have $\mathcal{A} \models \varphi[\overline{a}] \iff \mathcal{A} \models \psi[\overline{a}]$.

Unary counting quantifiers. All quantifiers considered in this article are *unary counting quantifiers* (for short: quantifiers), i.e., subsets of $\mathbb{N}$. We will use the terms "set (of natural numbers)" and "quantifier" interchangeably.

For a quantifier $\mathsf{Q} \subseteq \mathbb{N}$ and a formula $\varphi(\overline{x}, y)$ over a signature σ, the formula $\mathsf{Q}y\,\varphi(\overline{x}, y)$ is satisfied by a σ-structure $\mathcal{A}$ and an interpretation $\overline{a}$ of the variables $\overline{x}$ if

$$|\{b \in A : \mathcal{A} \models \varphi[\overline{a}, b]\}| \in \mathsf{Q}.$$

For a *set* $\mathbf{Q} \subseteq \mathcal{P}(\mathbb{N})$ *of quantifiers* we write $\mathrm{FO}(\mathbf{Q})$ to denote the *extension of* FO *with the quantifiers from* $\mathbf{Q}$. To avoid unnecessary special cases in proofs, we will often tacitly assume that the existential quantifier $\exists = \mathbb{N}_{\geqslant 1}$ is contained in the sets $\mathbf{Q}$ considered.[1]

The quantifier rank $\mathrm{qr}(\varphi)$ of an $\mathrm{FO}(\mathbf{Q})$-formula φ is defined as the maximal nesting depth of *all* quantifiers.

For a number $k \geqslant 0$ we write $(\mathsf{Q}{+}k)y\,\varphi(\overline{x}, y)$ as a short hand for a formula expressing in a σ-structure $\mathcal{A}$ and for an interpretation $\overline{a}$ of the variables $\overline{x}$ that the number of elements $b \in A$ such that $\mathcal{A} \models \varphi[\overline{a}, b]$ belongs to the set $(\mathsf{Q}{+}k) := \{n{+}k : n \in \mathsf{Q}\}$. Clearly, $(\mathsf{Q}{+}k)y\,\varphi(\overline{x}, y)$ can be expressed by the formula

$$\begin{aligned} \exists y_1 \cdots \exists y_k \Big(& \bigwedge_{1 \leqslant i < j \leqslant k} \neg\, y_i{=}y_j \\ & \wedge\ \forall y \,\big(\bigvee_{1 \leqslant i \leqslant k} y{=}y_i \ \rightarrow\ \varphi(\overline{x}, y) \,\big) \\ & \wedge\ \mathsf{Q}y \,\big(\varphi(\overline{x}, y) \wedge \bigwedge_{1 \leqslant i \leqslant k} \neg\, y{=}y_i \,\big)\Big). \end{aligned} \tag{1}$$

For every $k \geqslant 1$, we will write $\exists^{\geqslant k}y\,\varphi$ and $\exists^{=k}y\,\varphi$ for the formulas $(\exists{+}(k{-}1))y\,\varphi$ and $\exists^{\geqslant k}y\,\varphi \wedge \neg\exists^{\geqslant k+1}y\,\varphi$, resp.

[1] Since we only consider finite structures, the classical quantifier $\exists$ and the unary counting quantifier $\mathbb{N}_{\geqslant 1}$ are equivalent.

The *displacement* of a formula ψ is the smallest $K \geqslant 0$ such that for every subformula of ψ of shape $(\mathsf{Q}{+}k)y\,\varphi$ with $\mathsf{Q} \subseteq \mathbb{N}$ and $k \geqslant 0$ it holds that $k \leqslant K$.

It is the aim of this paper to study the locality of the logics FO($\mathbf{Q}$) in the sense of Hanf's theorem [4, 5, 8]. To define the according locality notion for this logic, we need the concepts introduced in the remainder of this section.

Gaifman graph. Let $\mathcal{A}$ be a σ-structure. Its *Gaifman graph* $G_{\mathcal{A}}$ is the undirected, loop-free graph with vertex set A and an edge between two distinct vertices $a, b \in A$ iff there exists $R \in \sigma$ and a tuple $(a_1, \ldots, a_{\mathrm{ar}(R)}) \in R^{\mathcal{A}}$ such that $a, b \in \{a_1, \ldots, a_{\mathrm{ar}(R)}\}$.

The *distance* $\mathit{dist}^{\mathcal{A}}(a, b)$ between two elements $a, b \in A$ is the minimal length (i.e., the number of edges) of a path from a to b in $G_{\mathcal{A}}$ (if no such path exists, we set $\mathit{dist}^{\mathcal{A}}(a, b) = \infty$).

For $r \geqslant 0$ and $a \in A$, the *r-neighbourhood of a in $\mathcal{A}$* is the set $N_r^{\mathcal{A}}(a) := \{b \in A : \mathit{dist}^{\mathcal{A}}(a, b) \leqslant r\}$. For a tuple $\overline{a} = (a_1, \ldots, a_n) \in A^n$, we write $N_r^{\mathcal{A}}(\overline{a})$ for the union of the sets $N_r^{\mathcal{A}}(a_i)$ for all $i \in [1, n]$.

Types and spheres. Let σ be a relational signature and let $c_1, c_2, \ldots$ be a sequence of pairwise distinct constant symbols. For every $n \geqslant 1$ we write σ_n for the signature $\sigma \cup \{c_1, \ldots, c_n\}$.

For every $r \geqslant 0$ and $n \geqslant 1$, a *type with n centres and radius r* (for short: *r-type with n centres*) is a σ_n-structure $(\mathcal{A}, a_1, \ldots, a_n)$, where $\mathcal{A}$ is a σ-structure, the constant symbols $c_1, \ldots, c_n$ are interpreted by the elements $a_1, \ldots, a_n \in A$, and $A = N_r^{\mathcal{A}}(a_1, \ldots, a_n)$. The elements $a_1, \ldots, a_n$ are called the *centres* of the r-type.

For a σ-structure $\mathcal{A}$ and a tuple $\overline{a} = (a_1, \ldots, a_n) \in A^n$, the *$r$-sphere of $\overline{a}$ in $\mathcal{A}$* is the r-type $\mathcal{N}_r^{\mathcal{A}}(\overline{a}) := (\mathcal{A}[N_r^{\mathcal{A}}(\overline{a})], \overline{a})$ (where $\mathcal{A}[B]$ denotes the restriction of the structure $\mathcal{A}$ to the universe $B \subseteq A$).

Bounded structures. The *degree* of a σ-structure $\mathcal{A}$ is the degree of its Gaifman graph $G_{\mathcal{A}}$. If this degree is $\leqslant d$, then we call $\mathcal{A}$ *d-bounded*.

Two formulas $\varphi(\overline{x})$ and $\psi(\overline{x})$ of signature σ are called *d-equivalent* if they are $\mathfrak{C}_d$-equivalent where $\mathfrak{C}_d$ is the class of all d-bounded σ-structures.

Note that for all $d \geqslant 2$, $r \geqslant 0$ and $n \geqslant 1$, every d-bounded r-type τ with n centres contains at most $n \cdot d^{r+1}$ elements. Thus, there is an FO[σ]-formula $\mathrm{sph}_\tau(\overline{x})$ of size $(n \cdot d^{r+1})^{\mathcal{O}(\|\sigma\|)}$ such that for every σ-structure $\mathcal{A}$ and every tuple $\overline{a} \in A^n$ we have that $\mathcal{A} \models \mathrm{sph}_\tau[\overline{a}] \iff \mathcal{N}_r^{\mathcal{A}}(\overline{a}) \cong \tau$.

Hanf normal form for first-order logic. An FO-sentence over a relational signature σ is said to be in *Hanf normal form* (cf. [1]) if it is a Boolean combination[2] of so-called *Hanf-sentences*. A Hanf-sentence is a sentence of the form

$$\exists^{\geqslant k} y \ \mathrm{sph}_\tau(y),$$

expressing that there exist at least k elements whose r-sphere is isomorphic to τ, for a given r-type τ. More generally, a *Hanf-formula* is a formula of the form

$$\exists^{\geqslant k} y \ \mathrm{sph}_\tau(\overline{x}, y),$$

where $k \geqslant 1$, $\overline{x}$ is a tuple of $n \geqslant 0$ variables, and τ is a type with $n{+}1$ centres and radius $r \geqslant 0$, such that for every σ-structure $\mathcal{A}$ and every tuple $\overline{a} \in A^n$,

$$\begin{aligned} & (\mathcal{A}, \overline{a}) \models \exists^{\geqslant k} y \ \mathrm{sph}_\tau(\overline{x}, y) \\ \iff \ & |\{b \in A : \mathcal{N}_r^{\mathcal{A}}(\overline{a}, b) \cong \tau\}| \ \geqslant \ k. \end{aligned}$$

An FO[σ]-formula is said to be in Hanf normal form (HNF, for short) if it is a Boolean combination of Hanf-formulas.

It is known that every first-order formula is d-equivalent to some formula in Hanf normal form [4] that can be computed in 3-fold exponential time [1]. A main result of the present paper is that this generalizes to formulas from FO($\mathbf{Q}$) only in case that all quantifiers in $\mathbf{Q}$ are ultimately periodic — which motivates the last definitions of this section.

Ultimately periodic sets. Let $\mathsf{Q} \subseteq \mathbb{N}$ be a unary quantifier. It is *ultimately periodic* if there exist numbers $p, n_0 \in \mathbb{N}$ with $p \geqslant 1$, such that

$$\text{for all } n \geqslant n_0 \text{ we have} \quad n \in \mathsf{Q} \iff n{+}p \in \mathsf{Q}. \qquad (2)$$

The minimal number $p \geqslant 1$ for which there exists an n_0 such that statement (2) is true is called the *period* of Q, and n_0 is called an *offset* of Q. Examples of ultimately periodic sets are the existential quantifier $\exists := \mathbb{N}_{\geqslant 1}$ (with period 1 and offset 1) and the divisibility quantifier $\mathsf{D}_p := \{\, p \cdot m \ : \ m \in \mathbb{N} \,\}$ for every $p \in \mathbb{N}$ with $p \geqslant 2$ (with period p and offset 0).

The *characteristic sequence* χ_{Q} of $\mathsf{Q} \subseteq \mathbb{N}$ is the ω-word $w = w_0 w_1 \cdots \in \{0,1\}^\omega$ with $\mathsf{Q} = \{n \in \mathbb{N} : w_n = 1\}$.

Fact 2.1. Let $\mathsf{Q} \subseteq \mathbb{N}$. If Q is ultimately periodic with offset n_0 and period p, then there exist finite words $\alpha \in \{0,1\}^*$ and $\pi \in \{0,1\}^+$ such that $\chi_{\mathsf{Q}} = \alpha \cdot \pi^\omega$, $|\alpha| = n_0$ and $|\pi| = p$.

If, conversely, $\xi_{\mathsf{Q}} = \alpha \cdot \pi^\omega$, then Q is ultimately periodic, its period divides $|\pi|$ and $|\alpha|$ is an offset.

We can thus represent an ultimately periodic set Q by the finite word $\mathit{rep}(\mathsf{Q}) := \alpha\#\pi$, where $\chi_{\mathsf{Q}} = \alpha \cdot \pi^\omega$. To make this definition unambiguous, we demand that $p := |\pi|$ is the period of Q, and $n_0 := |\alpha|$ is the *smallest* offset of Q for the period p. The *size* $\|\mathsf{Q}\|$ of Q is defined as the length of $\mathit{rep}(\mathsf{Q})$.

Let $\mathbf{Q} \subseteq \mathcal{P}(\mathbb{N})$ be a set of ultimately periodic sets. The size $\|\varphi\|$ of an FO($\mathbf{Q}$)-formula[3] φ of signature σ is its length when viewed as a word over the alphabet $\sigma \cup \mathrm{Var} \cup \{=, \exists, \neg, \vee, (,)\} \cup \{,\} \cup \{0, 1, \#\}$, where Var is a countable set of variable symbols, and where each quantifier $\mathsf{Q} \in \mathbf{Q}$ is represented by the word $\mathit{rep}(\mathsf{Q})$.

[2] Throughout this paper, whenever we speak of *Boolean combinations*, we mean *finite* Boolean combinations.

[3] as usual, $\exists$, $\neg$, $\vee$ belong to the official syntax, whereas $\forall$, $\wedge$, $\rightarrow$, $\leftrightarrow$ will be used as abbreviations when constructing formulas

3. Main results

In the following, we denote by σ a relational signature.

We generalise the notion of Hanf normal form to extensions of first-order logic by unary counting quantifiers in the following way. A *Hanf-formula* is a formula of the form

$$(\mathsf{Q}{+}k)y\ \mathrm{sph}_\tau(\overline{x},y),$$

where $\mathsf{Q} \subseteq \mathbb{N}$, $k \geqslant 0$, and τ is an r-type with $n{+}1$ centres. The free variables of this formula are $\overline{x}$, and the formula expresses that the number of interpretations for y such that the r-sphere of $\overline{x}, y$ is isomorphic to τ, belongs to the set $(\mathsf{Q}{+}k)$. I.e., for every σ-structure $\mathcal{A}$ and every tuple $\overline{a} \in A^n$,

$$\begin{aligned} &(\mathcal{A},\overline{a}) \models (\mathsf{Q}{+}k)y\ \mathrm{sph}_\tau(\overline{x},y) \\ \iff\ & |\{b \in A : \mathcal{N}_r^{\mathcal{A}}(\overline{a},b) \cong \tau\}| \in (\mathsf{Q}{+}k). \end{aligned}$$

The *locality radius of the Hanf-formula* $(\mathsf{Q}{+}k)y\ \mathrm{sph}_\tau(\overline{x},y)$ is the radius of the type τ. A *Hanf-sentence* is a Hanf-formula that does not have any free variable. A formula in *Hanf normal form* (for short: HNF) is a Boolean combination of Hanf-formulas. The *locality radius of a Hanf normal form* is the maximum of the locality radii of its Hanf-formulas.

Definition 3.1. Let $\mathbf{Q} \subseteq \mathcal{P}(\mathbb{N})$ be a set of unary counting quantifiers. We say that $\mathbf{Q}$ *permits Hanf normal forms* if for every relational signature σ and every degree bound $d \geqslant 0$, every $\mathrm{FO}(\mathbf{Q})[\sigma]$-formula is d-equivalent to an $\mathrm{FO}(\mathbf{Q})[\sigma]$-formula in HNF.

Our main result characterises the sets $\mathbf{Q}$ that permit Hanf normal forms in terms of ultimately periodic sets:

Theorem 3.2. *A set $\mathbf{Q} \subseteq \mathcal{P}(\mathbb{N})$ of unary counting quantifiers permits Hanf normal forms if, and only if, every quantifier $\mathsf{Q} \in \mathbf{Q}$ is ultimately periodic.*

For the "only if" direction of Theorem 3.2, we consider a unary counting quantifier $\mathsf{S} \in \mathbf{Q}$ that is *not* ultimately periodic and show that already for the signature $\sigma_P := \{P\}$ with P unary, no sentence in HNF can express "$|A| \in \mathsf{S}$". I.e., we show that the formula $\mathsf{S}y\, y{=}y$ is not equivalent to any $\mathrm{FO}(\mathbf{Q})[\sigma_P]$-sentence in HNF. For achieving this, we utilize that no finite factor of χ_{S} determines the remainder of the characteristic word. The proof details are given in Section 4.

For the "if" direction, consider a set $\mathbf{Q} \subseteq \mathcal{P}(\mathbb{N})$ of ultimately periodic sets. Furthermore, let $\mathbf{D} \subseteq \mathcal{P}(\mathbb{N})$ be the set which contains for each $\mathsf{Q} \in \mathbf{Q}$ with period $p \geqslant 2$ the divisibility quantifier D_p.[4] The proof of the "if" direction proceeds by showing the following:

(1) Any $\mathrm{FO}(\mathbf{Q})$-formula φ can be translated into an $\mathrm{FO}(\mathbf{D})$-formula that is equivalent to φ on all finite structures (irrespective of their degree).

(2) For $t \geqslant 1$ let $\sigma_{[1,t]}$ be the signature consisting of the unary relation symbols $P_1, \ldots, P_t$. The properties "$|A| \in (\exists{+}k)$" and "$|A| \in (\mathsf{D}_p{+}k)$" can be expressed by sentences in HNF from $\mathrm{FO}[\sigma_{[1,t]}]$ and $\mathrm{FO}(\{\mathsf{D}_p\})[\sigma_{[1,t]}]$, resp. (for all $k \in \mathbb{N}$ and all $p \geqslant 2$).

(3) The set $\mathbf{D}$ permits Hanf normal forms.

(4) If $\mathsf{Q} \subseteq \mathbb{N}$ is ultimately periodic with period $p \geqslant 2$ and $\psi = (\mathsf{D}_p{+}k)y\ \varphi$, then there is a Boolean combination of formulas $(\mathsf{Q}{+}\ell)y\ \varphi$ and $(\exists{+}\ell)y\ \varphi$ for suitable numbers $\ell \in \mathbb{N}$ that is equivalent to ψ on all finite structures (irrespective of their degree).

Step (1) is straightforward. Steps (2) and (3) are the crucial steps (and (2) is used for proving (3)). Step (4) is obtained by an application of a basic result on word combinatorics. Note that the "if" direction of Theorem 3.2 is an immediate consequence of steps (1), (3), and (4). Proof details for the steps (1)–(4) can be found in Section 5.

Our second main result provides a worst-case optimal algorithm for transforming formulas into Hanf normal form:

Theorem 3.3. *There is an algorithm which receives as input a degree bound $d \in \mathbb{N}$ and a $\varphi \in \mathrm{FO}(\mathbf{Q})$, where $\mathbf{Q} \subseteq \mathcal{P}(\mathbb{N})$ is a set of ultimately periodic sets, and constructs a d-equivalent HNF $\psi \in \mathrm{FO}(\mathbf{Q})$ of the same signature as φ. For $d \geqslant 3$, the formula ψ has locality radius $\leqslant 4^{qr(\varphi)}$ and displacement in $d^{2^{\mathcal{O}(\|\varphi\|)}}$, and the algorithm's runtime is in*

$$2^{d^{2^{\mathcal{O}(\|\varphi\|)}}}.$$

We will present the proof of the "if" direction of Theorem 3.2 in a way which allows it to be read as the algorithm of Theorem 3.3. An upper bound on the algorithm's runtime is obtained by a careful analysis of the time required for performing each of the steps of that proof; see Section 6 for details. For obtaining the 3-fold exponential upper bound, it is crucial that the main construction within our proof of Step (2) is done via a divide and conquer approach (a more straightforward brute-force approach only yields a 4-fold exponential upper bound).

A 3-fold exponential lower bound (for $d = 3$) was shown in [1] already for plain first-oder logic, i.e., for the special case where $\mathbf{Q} = \{\exists\}$.

4. Proof of the "only if" direction of Theorem 3.2

In this section we show that, whenever a set $\mathbf{Q}$ of quantifiers contains some quantifier that is *not* ultimately periodic, then $\mathbf{Q}$ does *not* permit Hanf normal forms. For this, it suffices to consider a signature consisting of a single unary relation symbol.

Lemma 4.1. *Let $\sigma_P := \{P\}$ be the signature consisting of a unary relation symbol P. Let $\mathbf{Q} \subseteq \mathcal{P}(\mathbb{N})$ be a set of unary counting quantifiers which contains a quantifier $\mathsf{S} \subseteq \mathbb{N}$ that is not ultimately periodic. There is no $\mathrm{FO}(\mathbf{Q})[\sigma_P]$-sentence*

[4] I.e., for each $\mathsf{Q} \in \mathbf{Q}$ of period $p \geqslant 2$ we have $\mathsf{D}_p \in \mathbf{D}$, and for each $\mathsf{D}_p \in \mathbf{D}$ there exists a $\mathsf{Q} \in \mathbf{Q}$ of period p.

δ in Hanf normal form, such that for all σ_P-structures $\mathcal{A}$ we have $\mathcal{A} \models \delta \iff |A| \in \mathsf{S}$.

Proof. For contradiction, assume that δ is an $\mathrm{FO}(\mathbf{Q})[\sigma_P]$-sentence in HNF expressing "$|A| \in \mathsf{S}$".

Since P is unary, any r-neighbourhood of an element a in a σ_P-structure $\mathcal{A}$ consists of its center a, only. Consequently, $P(y)$ and $\neg P(y)$ are the only formulas $\mathrm{sph}_\tau(y)$, where τ is a type with one center. Hence, there are a finite set $\mathbf{Q}' \subseteq \mathbf{Q}$ and a natural number $k \geqslant 1$ such that δ is a Boolean combination of sentences of the following form, with $\mathsf{Q} \in \mathbf{Q}'$ and $\ell \in [0, k{-}1]$:

$$(\mathsf{Q}{+}\ell)y\ P(y) \qquad \text{or} \qquad (\mathsf{Q}{+}\ell)y\ \neg P(y).$$

Let $\mathsf{Q}_1, \ldots, \mathsf{Q}_j$ be a list of all $\mathsf{Q} \in \mathbf{Q}'$. For each $a \in \mathbb{N}$ with $a \geqslant k$ consider the word w_a of length $k \cdot j$ defined as the concatenation of the bitstrings $\chi_{\mathsf{Q}_i}(a{-}k, a]$ for $i = 1, \ldots, j$.

Clearly, there exist natural numbers $b > a > k$ with $w_a = w_b$. Consequently, $\chi_{\mathsf{Q}}(a{-}k, a] = \chi_{\mathsf{Q}}(b{-}k, b]$ for all $\mathsf{Q} \in \mathbf{Q}'$.

If, for all $c \geqslant 0$, we have $a + c \in \mathsf{S} \iff b + c \in \mathsf{S}$, then S is ultimately periodic (with period dividing $b{-}a$ and offset a). Since this is not the case, there is $c \geqslant 0$ with $a + c \in \mathsf{S} \iff b + c \notin \mathsf{S}$.

Now consider σ_P-structures $\mathcal{A}$ and $\mathcal{B}$ with $|A| = a{+}c$, $|B| = b{+}c$, and $|P^{\mathcal{A}}| = |P^{\mathcal{B}}| = c$. By choice of a, b, c, we have $|A| \in \mathsf{S} \iff |B| \notin \mathsf{S}$, and thus, $\mathcal{A} \models \delta \iff \mathcal{B} \not\models \delta$.

Nevertheless, $\mathcal{A}$ and $\mathcal{B}$ cannot be distinguished by any of the Hanf-sentences that occur in δ: To this end, let $\mathsf{Q} \in \mathbf{Q}'$ and $\ell \in [0, k{-}1]$. Then $\mathcal{A} \models (\mathsf{Q}{+}\ell)y\ P(y)$ iff $|P^{\mathcal{A}}| \in (\mathsf{Q}{+}\ell)$. This is equivalent to $|P^{\mathcal{B}}| \in (\mathsf{Q}{+}\ell)$ since $|P^{\mathcal{A}}| = |P^{\mathcal{B}}|$, and therefore to $\mathcal{B} \models (\mathsf{Q}{+}\ell)y\ P(y)$. On the other hand, $\mathcal{A} \models (\mathsf{Q}{+}\ell)y\ \neg P(y)$ iff $a - \ell \in \mathsf{Q}$, since $|A \setminus P^{\mathcal{A}}| = a$. This is equivalent to $b - \ell \in \mathsf{Q}$, since $\chi_{\mathsf{Q}}(a{-}k, a] = \chi_{\mathsf{Q}}(b{-}k, b]$. Finally, $b{-}\ell \in \mathsf{Q}$ is equivalent to $\mathcal{B} \models (\mathsf{Q}{+}\ell)y\ \neg P(y)$, since $|B \setminus P^{\mathcal{B}}| = b$. In summary, the structures $\mathcal{A}$ and $\mathcal{B}$ satisfy the same Hanf-sentences that occur in δ. As δ is a Boolean combination of these Hanf-sentences, we obtain that $\mathcal{A} \models \delta$ iff $\mathcal{B} \models \delta$. This is a contradiction, completing the proof of Lemma 4.1. ☐

The "only if" direction of Theorem 3.2 is an immediate consequence: For $\mathsf{S} \in \mathbf{Q}$, the $\mathrm{FO}(\mathbf{Q})[\sigma_P]$-sentence $\mathsf{S}y\ y{=}y$ expresses "$|A| \in \mathsf{S}$". But if S is not ultimately periodic, this cannot be expressed by any $\mathrm{FO}(\mathbf{Q})[\sigma_P]$-sentence in HNF by Lemma 4.1.

5. Proof of the "if" direction of Theorem 3.2

For the proof of the "if" direction of Theorem 3.2, we have to show that for every set $\mathbf{Q} \subseteq \mathcal{P}(\mathbb{N})$ of ultimately periodic sets, for every relational signature σ, for every degree bound $d \geqslant 0$, and for every $\mathrm{FO}(\mathbf{Q})[\sigma]$-formula φ, there is a d-equivalent $\mathrm{FO}(\mathbf{Q})[\sigma]$-formula ψ in Hanf normal form.

For this, we introduce the notion of the *generalised quantifier rank gqr(φ) of a formula φ*, which is defined in the same way as the quantifier rank, with the only exception that a formula ψ of the shape $(\mathsf{Q}{+}k)y\ \varphi$ has generalised quantifier rank $\mathrm{gqr}(\psi) = \mathrm{gqr}(\varphi) + 1$ (in contrast, ψ has quantifier rank $\mathrm{qr}(\psi) = \mathrm{qr}(\varphi) + k + 1$).

For the remainder of this section we denote by $\mathbf{Q} \subseteq \mathcal{P}(\mathbb{N})$ a set of ultimately periodic quantifiers. Furthermore, we let $\mathbf{D} \subseteq \mathcal{P}(\mathbb{N})$ be the set which contains for each $\mathsf{Q} \in \mathbf{Q}$ with period $p \geqslant 2$ the divisibility quantifier D_p.

The proof follows the four steps outlined in Section 3.

5.1 Step (1)

Step (1) is established by the following lemma.

Lemma 5.1. *Let $\mathsf{Q} \subseteq \mathbb{N}$ be ultimately periodic with period p and offset n_0. Every formula of the shape $\mathsf{Q}y\ \varphi$ is equivalent to a Boolean combination of formulas of the form $(\mathsf{D}_p{+}\ell)y\ \varphi$ and $(\exists{+}\ell)y\ \varphi$, for $\ell < n_0{+}p$.*
This Boolean combination has size $\mathcal{O}((n_0{+}p)^3 \cdot \|\varphi\|)$, and it has the same generalised quantifier rank as $\mathsf{Q}y\ \varphi$.

Proof. Let $n_1 \in \mathbb{N}$ be the (unique) number in $[n_0, n_0{+}p)$ that is divisible by p. Clearly, Q is also ultimately periodic with period p and offset n_1. Let $\mathsf{Q}_1 := \mathsf{Q} \cap [0, n_1)$ and $R := \{r \in [0, p) : n_1{+}r \in \mathsf{Q}\}$. It is straightforward to verify that $\mathsf{Q}y\ \varphi$ is equivalent to the formula

$$\Big(\bigvee_{\ell \in \mathsf{Q}_1} \exists^{=\ell} y\ \varphi \Big) \ \vee\ \Big(\exists^{\geqslant n_1} y\ \varphi\ \wedge\ \bigvee_{r \in R} (\mathsf{D}_p{+}r)y\ \varphi \Big). \quad (3)$$

Clearly, this formula has the same generalised quantifier rank as $\mathsf{Q}y\ \varphi$, it has displacement $< n_0{+}p$, and it is of size $\mathcal{O}((n_0{+}p)^3 \cdot \|\varphi\|)$. ☐

5.2 Step (2)

For every $t \in \mathbb{N}_{\geqslant 1}$ let $\sigma_{[1,t]}$ be the signature consisting of t unary relation symbols $P_1, \ldots, P_t$. Step (2) consists of proving the following lemma.

Lemma 5.2. *There is an algorithm which receives as input numbers $i, j, t, k \in \mathbb{N}$ with $1 \leqslant i \leqslant j < t$, and a quantifier $\mathsf{Q} \in \{\exists\} \cup \boldsymbol{D}$ with period $p \geqslant 1$, and constructs a formula $\delta^{(\mathsf{Q}+k)}_{[i,j]} \in \mathrm{FO}(\boldsymbol{D})[\sigma_{[1,t]}]$ in HNF such that for every $\sigma_{[1,t]}$-structure $\mathcal{B}$ where the relations $P_1^{\mathcal{B}}, \ldots, P_t^{\mathcal{B}}$ are mutually disjoint,*

$$\mathcal{B} \models \delta^{(\mathsf{Q}+k)}_{[i,j]} \quad \iff \quad \Big| \bigcup_{s=i}^{j} P_s^{\mathcal{B}} \Big| \in (\mathsf{Q}{+}k).$$

The displacement of $\delta^{(\mathsf{Q}+k)}_{[i,j]}$ is $\leqslant \max\{k, p\}$.

From this lemma, one can even construct a formula in HNF that works without the assumption on the relations $P_i^{\mathcal{B}}$ to be mutually disjoint. But this formula would be larger and we later need the small formula (see Cor. 6.2).

Proof. In the following, let $\mathfrak{C}$ denote the class of all $\sigma_{[1,t]}$-structures $\mathcal{B}$ whose relations $P_s^{\mathcal{B}}$ are mutually disjoint.

Observe that for every $\mathcal{B} \in \mathfrak{C}$, we have

$$\left|\bigcup_{s=i}^{j} P_s^{\mathcal{B}}\right| \in (\mathsf{Q}{+}k) \iff \mathcal{B} \models (\mathsf{Q}{+}k)y \bigvee_{s=i}^{j} P_s(y).$$

Thus, if $i = j$, we are done.

The algorithm proceeds by a recursive subdivision of the interval $[i, j]$. For this, let $h := \lfloor \frac{j-i}{2} \rfloor$.

Case 1: $\mathsf{Q} = \exists$. The formula $(\exists{+}k)y \bigvee_{s=i}^{j} P_s(y)$ is $\mathfrak{C}$-equivalent (but not equivalent on all structures) to

$$(\exists{+}k)y \bigvee_{s=i}^{h} P_s(y) \quad\vee\quad (\exists{+}k)y \bigvee_{s=h+1}^{j} P_s(y) \quad\vee$$
$$\bigvee_{\ell=1}^{k} \Big((\exists{+}(\ell{-}1))y \bigvee_{s=i}^{h} P_s(y) \;\wedge\; (\exists{+}(k{-}\ell))y \bigvee_{s=h+1}^{j} P_s(y) \Big).$$

The algorithm proceeds recursively by decomposing the quantified subformulas in the same manner, and arrives at a Boolean combination $\delta_{[i,j]}^{(\exists+k)}$ of Hanf-formulas of the shape $(\exists{+}\ell)y\, P_s(y)$ with $\ell \leqslant k$.

Case 2: $\mathsf{Q} = \mathsf{D}_p$ *with* $p \geqslant 2$ *and* $k < p$. The formula $(\mathsf{D}_p{+}k)y \bigvee_{s=i}^{j} P_s(y)$ is $\mathfrak{C}$-equivalent (but not equivalent on all structures) to

$$\bigvee_{\substack{k_1,k_2 \in [0,p),\\ k_1+k_2 \equiv k \bmod p}} \Big((\mathsf{D}_p{+}k_1)y \bigvee_{s=i}^{h} P_s(y) \;\wedge\; (\mathsf{D}_p{+}k_2)y \bigvee_{s=h+1}^{j} P_s(y) \Big).$$

In the same way as in Case 1, the algorithm proceeds recursively and arrives at a Boolean combination $\delta_{[i,j]}^{(\mathsf{D}_p+k)}$ of Hanf-formulas of the shape $(\mathsf{D}_p{+}\ell)y\, P_s(y)$ for $\ell < p$.

Case 3: $\mathsf{Q} = \mathsf{D}_p$ *with* $p \geqslant 2$ *and* $k \geqslant p$. Let $k' \in [0, p{-}1]$ with $k' \equiv k \mod p$. For every $\mathcal{B} \in \mathfrak{C}$, we have

$$\left|\bigcup_{s=i}^{j} P_s^{\mathcal{B}}\right| \in (\mathsf{Q}{+}k) \iff \left|\bigcup_{s=i}^{j} P_s^{\mathcal{B}}\right| \in (\exists{+}(k{-}1)) \cap (\mathsf{Q}{+}k')$$

Therefore, the algorithm can output the HNF

$$\delta_{[i,j]}^{(\mathsf{D}_p+k)} \;:=\; \delta_{[i,j]}^{(\exists+(k-1))} \;\wedge\; \delta_{[i,j]}^{(\mathsf{D}_p+k')},$$

where $\delta_{[i,j]}^{(\exists+(k-1))}$ and $\delta_{[i,j]}^{(\mathsf{D}_p+k)}$ are HNFs, constructed according to Case 1 and Case 2, respectively. Note that in this case, the HNF $\delta_{[i,j]}^{(\mathsf{D}_p+k)}$ has displacement $\leqslant \max\{k, p\}$. □

5.3 Step (3)

The following lemma states this subsection's main result. In particular, it implies that the set **D** permits Hanf normal forms.

Lemma 5.3. *There is an algorithm which receives as input a degree bound $d \geqslant 3$ and a $\varphi \in \mathrm{FO}(\boldsymbol{D})[\sigma]$ (where σ is a relational signature), and constructs a HNF $\psi \in \mathrm{FO}(\boldsymbol{D})[\sigma]$ that is d-equivalent to φ.*

Furthermore, ψ has locality radius $\leqslant 4^q$ and displacement $\leqslant (q{+}1) \cdot \|\varphi\| \cdot d^{4^q+1} + \max\{K, P\}$, where $q \geqslant 0$ is the generalised quantifier rank of φ, $K \geqslant 0$ is the displacement of φ, and $P \geqslant 2$ is an upper bound on the periods of the quantifiers occurring in φ.

In the following, we let σ be a finite relational signature. For every $d \geqslant 3$, we write $\mathfrak{C}_d$ for the *class of all d-bounded σ-structures*. For each $n \geqslant 1$ and each $r \geqslant 0$ we write $T_r^d(n)$ for a *set of all (up to isomorphism) d-bounded r-types with n centres*. I.e., for every d-bounded r-type τ with n centres, there is precisely one $\tau' \in T_r^d(n)$ such that $\tau \cong \tau'$.

Furthermore, we write $\top$ for a fixed *tautological* $\mathrm{FO}[\sigma]$-sentence in HNF; e.g., we can choose $\top := \exists y\, \mathrm{sph}_\tau(y) \vee \neg\exists y\, \mathrm{sph}_\tau(y)$, where τ is an arbitrary, fixed type of radius 0 with one centre. We let $\bot := \neg\top$ be the corresponding *unsatisfiable* sentence in HNF.

For the proof of Lemma 5.3, we generalise a construction by Bollig and Kuske [1], proceeding by induction on the shape of $\mathrm{FO}(\mathbf{D})[\sigma]$-formulas. While the case of quantifier-free formulas is straightforward, much more work is needed to transform a formula $\varphi(\overline{x})$ of the shape $(\mathsf{Q}{+}k)y\; \psi'(\overline{x}, y)$ into a d-equivalent HNF. Suppose that $\psi'(\overline{x}, y)$ is already in HNF. Our construction is carried out along the following steps:

- For each structure $\mathcal{A} \in \mathfrak{C}_d$ and each tuple $\overline{a} \in A^n$, we let
$$B^{\mathcal{A}}(\overline{a}) \;:=\; \{b \in A \,:\, \mathcal{A} \models \psi'[\overline{a}, b]\}.$$
Clearly, $\mathcal{A} \models \varphi[\overline{a}] \iff |B^{\mathcal{A}}(\overline{a})| \in (\mathsf{Q}{+}k)$.
Suppose that $\varphi(\overline{x})$ has $n \geqslant 0$ free variables and that the HNF $\psi'(\overline{x}, y)$ has locality radius $r \geqslant 0$. Observe that for every $b \in B^{\mathcal{A}}(\overline{a})$, there is exactly one $\tau \in T_{4r}^d(n{+}1)$ such that $\mathcal{N}_{4r}^{\mathcal{A}}(\overline{a}, b) \cong \tau$.
Let $\tau_1, \ldots, \tau_t$, for $t := |T_{4r}^d(n{+}1)|$, be an enumeration of $T_{4r}^d(n{+}1)$. Let $\mathcal{B}$ be the $\sigma_{[1,t]}$-structure $(B, P_1^{\mathcal{B}}, \ldots, P_t^{\mathcal{B}})$ with universe $B := B^{\mathcal{A}}(\overline{a})$ and, for each $s \in [1, t]$, the set $P_s^{\mathcal{B}} := \{b \in B \,:\, \mathcal{N}_{4r}^{\mathcal{A}}(\overline{a}, b) \cong \tau_s\}$. In the following we will denote the structure $\mathcal{B}$ also by $\mathcal{B}_{4r}^{\mathcal{A}}(\overline{a})$.
Clearly, B is the disjoint union of the sets $P_1^{\mathcal{B}}, \ldots, P_t^{\mathcal{B}}$. Hence, for the HNF sentence $\delta_{[1,t]}^{(\mathsf{Q}+k)} \in \mathrm{FO}(\mathbf{D})[\sigma_{[1,t]}]$, obtained from Lemma 5.2 in Step (2), we have that
$$|B| \in (\mathsf{Q}{+}k) \iff \mathcal{B} \models \delta_{[1,t]}^{(\mathsf{Q}+k)}.$$
- For every quantifier $\mathsf{R} \in \{\exists\} \cup \mathbf{D}$, each $\ell \geqslant 0$, and each $s \in [1, t]$ we construct a HNF $\psi_s^{(\mathsf{R}+\ell)}(\overline{x}) \in \mathrm{FO}(\mathbf{D})[\sigma]$, such that for every σ-structure $\mathcal{A}$, for every tuple $\overline{a} \in A^n$, and for the $\sigma_{[1,t]}$-structure $\mathcal{B} := \mathcal{B}_{4r}^{\mathcal{A}}(\overline{a})$, we have that
$$\mathcal{A} \models \psi_s^{(\mathsf{R}+\ell)}[\overline{a}] \iff |P_s^{\mathcal{B}}| \in (\mathsf{R}{+}\ell).$$
I.e., we interpret the $\sigma_{[1,t]}$-structure $\mathcal{B}_{4r}^{\mathcal{A}}(\overline{a})$ in $(\mathcal{A}, \overline{a})$.

- In the HNF $\delta^{(\mathsf{Q}+k)}_{[1,t]}$, we replace every Hanf-formula of the shape $(\mathsf{R}+\ell)y\, P_s(y)$ by the HNF $\psi_s^{(\mathsf{R}+\ell)}(\overline{x})$.

Clearly, the resulting $\mathrm{FO}(\mathbf{D})[\sigma]$-formula $\psi(\overline{x})$ is in HNF and, furthermore, d-equivalent to $\varphi(\overline{x})$.

The details of the above construction are carried out as follows.

Proof of Lemma 5.3. We describe the algorithm on input of a degree bound $d \geqslant 3$ and an $\mathrm{FO}(\mathbf{D})[\sigma]$-formula $\varphi(\overline{x})$ over a finite relational signature σ. Let $q \geqslant 0$ be the generalised quantifier rank of φ. Furthermore, let $K \geqslant 0$ be the displacement of φ and let $P \geqslant 2$ be an upper bound on the periods of the quantifiers occurring in φ.

The algorithm proceeds by induction on the shape of $\varphi(\overline{x})$. We will show that for the HNF $\psi(\overline{x}) \in \mathrm{FO}(\mathbf{D})[\sigma]$, which the algorithm constructs, the following claim holds:

Claim 5.4. *(a) $\psi(\overline{x})$ has locality radius $\leqslant 4^q$.*

(b) $\psi(\overline{x})$ is d-equivalent to $\varphi(\overline{x})$.

(c) $\psi(\overline{x})$ has displacement $\leqslant (q+1)\cdot N + \max\{K,P\}$, for $N := \|\varphi\| \cdot d^{4^q+1}$.

In the following, the main steps of the algorithm are numbered. In Appendix D, an analysis of the time complexity of each of these steps can be found by the same number.

Suppose that φ is quantifier-free, i.e., $q = 0$. Here, φ is d-equivalent to a disjunction over all types in $T_0^d(n)$ that satisfy $\varphi(\overline{x})$. The only minor technical difficulty here is to express this disjunction as a proper HNF. The algorithm proceeds in the following steps:

(1) Compute the set $T_0^d(n{+}1)$.

(2) Let $\overline{c} = (c_1,\ldots,c_n)$ and compute the set $T \subseteq T_0^d(n{+}1)$ that contains precisely the types $\tau = (\mathcal{T},\overline{c},c_{n+1})$ from $T_0^d(n{+}1)$ with $\mathcal{T} \models \varphi[\overline{c}]$.

(3) It T is the empty set, then φ is unsatisfiable and, hence, equivalent to the HNF $\bot$. Otherwise, let

$$\psi(\overline{x}) \;:=\; \bigvee_{\tau\in T} \exists^{\geqslant 1} y\; \mathrm{sph}_\tau(\overline{x},y).$$

It is clear that in both cases, $\psi(\overline{x})$ is a HNF which in particular satisfies *Claim 5.4 (a) and (c).*

For a proof of Claim 5.4 (b), consider a structure $\mathcal{A} \in \mathfrak{C}_d$ and a tuple $\overline{a} = (a_1,\ldots,a_n) \in A^n$.

"$\Longrightarrow$": If $\mathcal{A} \models \varphi[\overline{a}]$ then, by definition of T, there is a type $\tau \in T$ such that $\mathcal{N}_0^{\mathcal{A}}(\overline{a},a_1) \cong \tau$. By interpreting the variable y with a_1 it follows that $(\mathcal{A},\overline{a}) \models \exists^{\geqslant 1} y\; \mathrm{sph}_\tau(\overline{x},y)$ and therefore $\mathcal{A} \models \psi[\overline{a}]$.

"$\Longleftarrow$": If $\mathcal{A} \models \psi[\overline{a}]$ then, by construction of ψ, there is a type $(\mathcal{T},\overline{c},c_{n+1}) \in T$ such that $(\mathcal{A},\overline{a}) \models \exists^{\geqslant 1} y\; \mathrm{sph}_\tau(\overline{x},y)$. In particular, there is an element $a_{n+1} \in A$ such that $\mathcal{N}_0^{\mathcal{A}}(\overline{a},a_{n+1}) \cong \tau$. Since, by construction of T we know that $\mathcal{T} \models \varphi[\overline{c}]$, also $\mathcal{A} \models \varphi[\overline{a}]$.

Suppose that φ is a Boolean combination with generalised quantifier rank $q \geqslant 1$. If $\varphi = \neg\varphi'$ then let ψ' be an $\mathrm{FO}(\mathbf{D})[\sigma]$-formula in HNF that is d-equivalent to φ' and let $\psi := \neg\psi'$. If $\varphi = (\varphi_1 \vee \varphi_2)$, then let ψ_1 and ψ_2 be $\mathrm{FO}(\mathbf{D})[\sigma]$-formulas in HNF that are d-equivalent to φ_1 and φ_2, respectively, and let $\psi := (\psi_1 \vee \psi_2)$.
In both cases, Claim 5.4 (a)–(c) is obviously satisfied.

Suppose that $\varphi(\overline{x}) = (\mathsf{Q}+k)y\,\varphi'(\overline{x},y)$. Let $p \in [1,P]$ be the period of Q. Recall that $k \leqslant K$. The algorithm proceeds as follows:

(4) By Claim 5.4 (a)–(c), there is a HNF $\psi'(\overline{x},y) \in \mathrm{FO}(\mathbf{D})[\sigma]$ that is d-equivalent to $\varphi'(\overline{x},y)$ and which has locality radius $\leqslant r := 4^{q-1}$ and displacement $\leqslant k'$ for

$$k' \;:=\; q\cdot(N{-}1) + \max\{K,P\}.$$

(5) Let $t := |T_{4r}^d(n{+}1)|$ and let $\tau_1,\ldots,\tau_t$ be an enumeration of the set $T_{4r}^d(n{+}1)$. Recall that for every $\mathcal{A} \in \mathfrak{C}_d$ and every tuple $\overline{a},b \in A^{n+1}$, there is exactly one $s \in [1,t]$ such that $\mathcal{N}_{4r}^{\mathcal{A}}(\overline{a},b) \cong \tau_s$.

(6) By Lemma 5.2, there is an $\mathrm{FO}(\mathbf{D})[\sigma_{[1,t]}]$-sentence $\delta^{(\mathsf{Q}+k)}_{[1,t]}$ such that for every $\sigma_{[1,t]}$-structure $\mathcal{B}$ where the universe B is a disjoint union of the sets $P_1^{\mathcal{B}},\ldots,P_t^{\mathcal{B}}$, we have that

$$\mathcal{B} \models \delta^{(\mathsf{Q}+k)}_{[1,t]} \quad\Longleftrightarrow\quad |B| \;\in\; (\mathsf{Q}+k).$$

Furthermore, $\delta^{(\mathsf{Q}+k)}_{[1,t]}$ is a Boolean combination of Hanf-formulas of the shape $(\mathsf{R}+\ell)y\,P_s(y)$ with $\mathsf{R} \in \{\exists\}\cup\mathbf{D}$, $s \in [1,t]$, and $\ell < \max\{k,p\} \leqslant \max\{K,P\}$.

(7) For every $s \in [1,t]$, we apply the following Lemma 5.5 to every Hanf-formula in ψ' that is not already a sentence, to obtain an $\mathrm{FO}(\mathbf{D})[\sigma]$-*sentence* ψ_s' in HNF that is *equivalent to $\psi'(\overline{x},y)$ in respect to* τ_s, i.e., for every σ-structure $\mathcal{A}$ and for every tuple $\overline{a},b \in A^{n+1}$ with $\mathcal{N}_{4r}^{\mathcal{A}}(\overline{a},b) \cong \tau_s$,

$$\mathcal{A} \models \psi'[\overline{a},b] \quad\Longleftrightarrow\quad \mathcal{A} \models \psi_s'.$$

Lemma 5.5. *There is an algorithm which receives as input a degree bound $d \geqslant 3$, a Hanf-formula $\alpha(\overline{x}) \in \mathrm{FO}(\boldsymbol{D})[\sigma]$ (where σ is a relational signature) with $n \geqslant 1$ free variables and locality radius $r \geqslant 0$, and a type $\tau \in T_{4r}^d(n)$, and constructs a HNF sentence $\alpha_\tau \in \mathrm{FO}(\boldsymbol{D})[\sigma]$ that is equivalent to $\alpha(\overline{x})$ in respect to τ. Furthermore, α_τ has locality radius $\leqslant r$ and displacement $\leqslant k + n\cdot d^{(2r+1)+1}$, where k is the displacement of α.*

We will prove Lemma 5.5 after having completed the proof of Lemma 5.3.
It follows from Lemma 5.5 that for every $s \in [1,t]$, the HNF ψ_s' has locality radius $\leqslant r$. Furthermore, since $(n{+}1)\cdot d^{(2r+1)+1} \leqslant N$, the formula ψ_s' has displacement

$$\begin{aligned} &\leqslant\; q\cdot(N{-}1) + \max\{K,P\} + N \\ &\leqslant\; (q{+}1)\cdot N + \max\{K,P\}. \end{aligned}$$

(8) For every Hanf-formula of the shape $(\mathsf{R}{+}\ell)y\,P_s(y)$ in the HNF $\delta^{(\mathsf{Q}+k)}_{[1,t]}$, we let

$$\psi^{(\mathsf{R}+\ell)}_s(\overline{x}) \;:=\; \psi'_s \;\wedge\; (\mathsf{R}{+}\ell)y\,\mathrm{sph}_{\tau_s}(\overline{x},y).$$

Note that $\psi^{(\mathsf{R}+\ell)}_s(\overline{x})$ has locality radius $\leqslant 4r = 4^q$. Since $\ell < \max\{k,p\} \leqslant \max\{K,P\}$, $\psi^{(\mathsf{R}+\ell)}_s(\overline{x})$ has displacement $\leqslant (q{+}1)\cdot N + \max\{K,P\}$. Furthermore, for every σ-structure $\mathcal{A}$ and every tuple $\overline{a} \in A^n$, we have

$$\mathcal{A} \models \psi^{(\mathsf{R}+\ell)}_s[\overline{a}] \iff \mathcal{B}^{\mathcal{A}}_{4r}(\overline{a}) \models (\mathsf{R}{+}\ell)y\,P_s(y). \quad (4)$$

Let $\psi(\overline{x})$ be the $\mathrm{FO}(\mathbf{D})[\sigma]$-formula which we obtain from $\delta^{(\mathsf{Q}+k)}_{[1,t]}$ when we replace every Hanf-formula of the shape $(\mathsf{R}{+}\ell)y\,P_s(y)$ by the HNF $\psi^{(\mathsf{R}+\ell)}_s(\overline{x})$. It is easy to verify that $\psi(\overline{x})$ satisfies *Claim 5.4 (a) and (c)*.

To verify that $\psi(\overline{x})$ also satisfies *Claim 5.4 (b)*, consider a structure $\mathcal{A} \in \mathfrak{C}_d$ and a tuple $\overline{a} \in A^n$. Let $\mathcal{B}$ denote the $\sigma_{[1,t]}$-structure $\mathcal{B}^{\mathcal{A}}_{4r}(\overline{a})$. The following equivalence holds:

$$\begin{aligned}
\mathcal{A} \models \varphi[\overline{a}] &\iff (\mathcal{A},\overline{a}) \models (\mathsf{Q}{+}k)y\,\varphi'(\overline{x},y)\\
&\iff (\mathcal{A},\overline{a}) \models (\mathsf{Q}{+}k)y\,\psi'(\overline{x},y)\\
&\iff |B| \in (\mathsf{Q}{+}k)\\
&\iff \mathcal{B} \models \delta^{(\mathsf{Q}+k)}_{[1,t]}\\
&\iff \mathcal{A} \models \psi[\overline{a}].
\end{aligned}$$

The last equivalence follows from the construction of $\psi(\overline{x})$ and equivalence (4). ☐

It remains to prove Lemma 5.5.

Proof of Lemma 5.5. We describe the algorithm on input of a degree bound $d \geqslant 3$, a Hanf-formula $\alpha(\overline{x}) \in \mathrm{FO}(\mathbf{D})[\sigma]$ (where σ is a finite relational signature) with locality radius $r \geqslant 0$ and $n \geqslant 1$ free variables, and for a type $\tau \in T^d_{4r}(n)$.

Suppose that $\overline{x} = (x_1,\ldots,x_n)$ are the free variables of $\alpha(\overline{x})$ and that $\alpha(\overline{x}) := (\mathsf{Q}{+}k)y\,\mathrm{sph}_\rho(\overline{x},y)$ with $\mathsf{Q} \in \{\exists\} \cup \mathbf{D}$, displacement $k \geqslant 0$, and $\rho \in T^d_r(n{+}1)$. Furthermore, suppose that $\rho = (\mathcal{R},\overline{c},d)$ for the centres $\overline{c} = (c_1,\ldots,c_n)$ and that $\tau = (\mathcal{T},\overline{e})$ for the centres $\overline{e} = (e_1,\ldots,e_n)$.

The algorithm proceeds by the following case distinction:

Suppose that $\mathit{dist}^\rho(\overline{c},d) \leqslant 2r+1$. Here, it holds that

$$\begin{aligned}
&|\{b \in A : \mathcal{N}^{\mathcal{A}}_r(\overline{a},b) \cong \rho\}|\\
=\;&|\{b \in N^{\mathcal{A}}_{2r+1}(\overline{a}) : \mathcal{N}^{\mathcal{A}}_r(\overline{a},b) \cong \rho\}|.
\end{aligned} \quad (5)$$

Furthermore, we have that $N^\rho_r(d) \subseteq N^\rho_{3r+1}(\overline{e})$. Note that, since $2r{+}1{+}r \leqslant 3r{+}1 \leqslant 4r$, this implies that for every σ-structure $\mathcal{A}$ and every tuple $\overline{a} \in A^n$ with $\mathcal{N}^{\mathcal{A}}_{4r}(\overline{a}) \cong \tau$,

$$\begin{aligned}
&|\{b \in N^{\mathcal{A}}_{2r+1}(\overline{a}) : \mathcal{N}^{\mathcal{A}}_r(\overline{a},b) \cong \rho\}|\\
=\;&\underbrace{|\{f \in N^{\tau}_{2r+1}(\overline{e}) : \mathcal{N}^{\tau}_r(\overline{e},f) \cong \rho\}|}_{=:\,k_\tau}
\end{aligned} \quad (6)$$

Hence, putting (5) and (6) together, the algorithm outputs $\alpha_\tau := \bot$ if $k_\tau \notin (\mathsf{Q}{+}k)$ and $\alpha_\tau := \top$ if $k_\tau \in (\mathsf{Q}{+}k)$.

Suppose that $\mathit{dist}^\rho(\overline{c},d) > 2r+1$. In this case, the sets $N^\rho_r(\overline{c})$ and $N^\rho_r(d)$ are disjoint and there are no edges in the Gaifman graph of ρ between the nodes from $N^\rho_r(\overline{c})$ and the nodes from $N^\rho_r(d)$.

If $\mathcal{N}^\rho_r(\overline{c})$ and $\mathcal{N}^\tau_r(\overline{e})$ are *not isomorphic* then we have for every σ-structure $\mathcal{A}$ and for each $\overline{a} \in A^n$ with $\mathcal{N}^{\mathcal{A}}_{4r}(\overline{a}) \cong \tau$, that $|\{b \in A : \mathcal{N}^{\mathcal{A}}_r(\overline{a},b) \cong \rho\}| = 0$. Hence, if $0 \in (\mathsf{Q}{+}k)$ we let $\alpha_\tau := \top$ and if $0 \notin (\mathsf{Q}{+}k)$ we let $\alpha_\tau := \bot$.

In the following we assume that $\mathcal{N}^\rho_r(\overline{c})$ and $\mathcal{N}^\tau_r(\overline{e})$ are isomorphic. In this case we have for each σ-structure $\mathcal{A}$ and every tuple $\overline{a} \in A^n$, that

$$\begin{aligned}
&|\{b \in A : \mathcal{N}^{\mathcal{A}}_r(\overline{a},b) \cong \rho\}|\\
=\;&|\{b \in A \setminus N^{\mathcal{A}}_{2r+1}(\overline{a}) : \mathcal{N}^{\mathcal{A}}_r(b) \cong \mathcal{N}^\rho_r(d)\}|.
\end{aligned} \quad (7)$$

Since $2r{+}1{+}r \leqslant 3r{+}1 \leqslant 4r$, the following equivalence holds for every σ-structure $\mathcal{A}$ and every tuple $\overline{a} \in A^n$ with $\mathcal{N}^{\mathcal{A}}_{4r}(\overline{a}) \cong \tau$:

$$\begin{aligned}
&|\{b \in N^{\mathcal{A}}_{2r+1}(\overline{a}) : \mathcal{N}^{\mathcal{A}}_r(b) \cong \mathcal{N}^\rho_r(d)\}|\\
=\;&\underbrace{|\{f \in N^{\tau}_{2r+1}(\overline{e}) : \mathcal{N}^{\tau}_r(f) \cong \mathcal{N}^\rho_r(d)\}|}_{=:\,\ell_\tau}.
\end{aligned} \quad (8)$$

Hence, putting (7) and (8) together, we know that

$$\begin{aligned}
&|\{b \in A : \mathcal{N}^{\mathcal{A}}_r(\overline{a},b) \cong \rho\}|\\
=\;&|\{b \in A : \mathcal{N}^{\mathcal{A}}_r(b) \cong \mathcal{N}^\rho_r(d)\}| \;-\; \ell_\tau.
\end{aligned}$$

It follows, that the algorithm can output the the Hanf-formula

$$\alpha_\tau \;:=\; (\mathsf{Q}{+}(k{+}\ell_\tau))y\,\mathrm{sph}_{\mathcal{N}^\rho_r(d)}(y).$$

Finally, observe that $\ell_\tau \leqslant n \cdot d^{(2r+1)+1}$ and hence, α_τ has displacement $\leqslant k + n \cdot d^{(2r+1)+1}$. ☐

5.4 Step (4)

Step (4) is established by the following lemma.

Lemma 5.6. *Let $\mathsf{Q} \subseteq \mathbb{N}$ be ultimately periodic with period $p \geqslant 2$ and offset n_0. Every formula of the shape $(\mathsf{D}_p{+}k)y\,\varphi$, for a $k \geqslant 0$, is equivalent to a Boolean combination of formulas of the shape $(\mathsf{Q}{+}\ell)y\,\varphi$ and $(\exists{+}\ell)y\,\varphi$ with $\ell < n_0{+}k{+}2p$. Furthermore, such a Boolean combination can be computed in time $\mathcal{O}((\|\mathsf{Q}\|{+}k)^3 \cdot \|\varphi\|)$.*

Proof. Let $n_1 \in \mathbb{N}_{\geqslant 1}$ be the smallest number $\geqslant \max\{n_0,k\}$ such that $n_1{-}1 \equiv k \bmod p$. Clearly, Q is ultimately periodic with period p also for the offset n_1. From Fact 2.1 we obtain bitstrings α and π of length $|\alpha| = n_1$ and $|\pi| = p$ such that $\chi_{\mathsf{Q}} = \alpha \cdot \pi^\omega$. I.e., $\chi_{\mathsf{Q}}[0,n_1{-}1] = \alpha$, and

$$\chi_{\mathsf{Q}}[n_1,n_1{+}p{-}1] = \pi = \chi_{\mathsf{Q}}[n_1{+}(s{-}1)p,\; n_1{+}sp{-}1], \quad (9)$$

for every $s \geqslant 1$.

Claim 5.7. *For all $n \in \mathbb{N}$ with $n \geqslant n_1{+}p{-}1$ we have $n \in (\mathsf{D}_p{+}k) \iff \chi_{\mathsf{Q}}(n{-}p,n] = \pi$.*

Before presenting the proof of the claim, let us first show how the claim can be used to prove Lemma 5.6.

Letting $\pi_0, \pi_1, \dots, \pi_{p-1} \in \{0,1\}$ such that $\pi = \pi_{p-1} \cdots \pi_1 \pi_0$, it is straightforward to see that for all $n \in \mathbb{N}$ with $n \geqslant n_1{+}p{-}1$ we have

$$\begin{aligned} & \chi_{\mathsf{Q}}(n{-}p, n] \;=\; \pi \\ \iff \quad & \bigwedge_{\substack{i \in [0,p):\\ \pi_i = 1}} n \in (\mathsf{Q}+i) \;\;\wedge\;\; \bigwedge_{\substack{j \in [0,p):\\ \pi_j = 0}} n \notin (\mathsf{Q}+j). \end{aligned}$$

Thus, the formula $(\mathsf{D}_p{+}k)y\,\varphi$ is equivalent to the formula

$$\begin{aligned} & \bigvee_{\ell \in S} \exists^{=\ell} y\,\varphi \quad \vee \quad \Big(\exists^{\geqslant n_1+p-1} y\,\varphi \\ & \quad \wedge \bigwedge_{\substack{i \in [0,p):\\ \pi_i = 1}} (\mathsf{Q}{+}i)y\,\varphi \;\;\wedge \bigwedge_{\substack{j \in [0,p):\\ \pi_j = 0}} \neg\,(\mathsf{Q}{+}j)y\,\varphi \Big), \end{aligned} \tag{10}$$

where S is the set of all $n \in (\mathsf{D}_p{+}k)$ with $n < n_1{+}p{-}1$. Since $n_1 \leqslant \max\{n_0, k\}{+}p$, each of the quantifiers that explicitly occur in the formula (10) has displacement $< n_0{+}k{+}2p \leqslant 2\|\mathsf{Q}\| + k$. Using this, it is straightforward to see that the formula (10) has size $\mathcal{O}((\|\mathsf{Q}\|{+}k)^3 \cdot \|\varphi\|)$ and can easily be computed within the same time bound.

To complete the proof of Lemma 5.6, it only remains to prove Claim 5.7.

Proof of Claim 5.7.
Fix an arbitrary $n \in \mathbb{N}$ with $n \geqslant n_1{+}p{-}1$.
The direction "$\Longrightarrow$" is an immediate consequence of equation (9).
Note that for proving the direction "$\Longleftarrow$", it suffices to prove the following: If $\chi_{\mathsf{Q}} = \beta \cdot \pi^\omega$ for some word $\beta \in \{0,1\}^*$ with $|\alpha| \leqslant |\beta|$, then $|\alpha| \equiv |\beta| \bmod p$ (in our case, $\beta = \chi_{\mathsf{Q}}[0, n{-}1]$). Since α is not longer than β, there exist $i \in \mathbb{N}$, a word $u \in \{0,1\}^*$ with $|u| < |\pi|$, and a word $v \in \{0,1\}^+$ with $\alpha\pi^i u = \beta\pi$ and $\beta\pi v = \alpha\pi^{i+1}$. Hence $\alpha\pi^i uv = \alpha\pi^{i+1}$ implies $\pi = uv$. Consequently, $\beta\pi vu = \alpha\pi^{i+1}u$ is a prefix of $\alpha\pi^\omega = \beta\pi^\omega$ which implies that vu is a prefix of π^ω of length $|vu| = |uv| = |\pi|$, i.e., $vu = \pi$. Since u and v commute (i.e., $uv = vu$), a basic result in word combinatorics [14, Proposition 1.3.2] implies the existence of some word $w \in \{0,1\}^+$ such that $u, v \in w^*$ and therefore $\pi = uv \in w^*$. But then $\chi_{\mathsf{Q}} = \alpha\pi^\omega = \alpha w^\omega$ implies $|\pi| = |w|$ since $p = |\pi|$ was the period of χ_{Q}, and hence minimal. Hence, $w = \pi = uv$. Consequently, $u = \varepsilon$ and therefore $\beta = \alpha\pi^i$. This ensures $|\alpha| \equiv |\beta| \bmod p$. This completes the proof of Claim 5.7 and therefore the proof of Lemma 5.6. ☐

Note that the "if" direction of Theorem 3.2 is an immediate consequence of the Lemmas 5.1, 5.3, and 5.6.

6. Complexity

This section is devoted to the proof of Theorem 3.3. We proceed along the steps outlined in Section 3 and provide a runtime analysis of the algorithm obtained from the proof presented in Section 5. Due to space restrictions, most proofs had to be deferred to an appendix.

The following Corollary 6.1 shows how to use Lemma 5.1 to compute, for a given FO($\mathbf{Q}$)-formula, an equivalent FO($\mathbf{D}$)-formula. Here, we let $w(\varphi) := 1$ if φ is quantifier-free. Otherwise, we denote by $w(\varphi)$ the largest number $\|\mathsf{Q}\|$ for all quantifiers Q occurring in φ.

Corollary 6.1. *There is an algorithm which receives as input a $\varphi \in$ FO($\mathbf{Q}$) of quantifier rank q, and constructs in time*

$$\|\varphi\| \cdot w(\varphi)^{\mathcal{O}(q)}$$

an equivalent $\psi \in$ FO($\boldsymbol{D}$) of the same signature as φ and with displacement $\leqslant w(\varphi)$ and generalised quantifier rank q.

Corollary 6.2. *The algorithm from Lemma 5.2 runs in time*

$$\max\{k, p\}^{\mathcal{O}(\log(j-i+1))}.$$

We proceed with a runtime analysis of the algorithms constructed in Step (3). We start with the algorithm provided by Lemma 5.5, and afterwards proceed to the algorithm provided by Lemma 5.3.

A task needed within both algorithms is to check whether two structures of size $\leqslant N$ (for some number $N \geqslant 1$) are isomorphic to each other. By using a brute-force algorithm, this can be done in time at most $N^{c_I} \cdot N^N$, where $c_I \geqslant 1$ is a suitable number of size $\mathcal{O}(\|\sigma\|)$: for each of the at most N^N bijections, it can be checked in time N^{c_I} whether the bijection is indeed an isomorphism.

Corollary 6.3. *The algorithm from Lemma 5.5 runs in time*

$$k \cdot 2^{(n \cdot d^{4r+1})^{\mathcal{O}(\|\sigma\|)}}.$$

Corollary 6.4. *The algorithm from Lemma 5.3 runs in time*

$$\max\{K, P\}^{(\|\varphi\| \cdot d^{4^q+1})^{\mathcal{O}(\|\sigma\|)}}.$$

We are now ready to prove Theorem 3.3.

Proof of Theorem 3.3.
We analyse the algorithm's runtime when receiving as input a degree bound $d \geqslant 3$ and a formula $\varphi \in$ FO($\mathbf{Q}$). Let σ denote the relational signature consisting of all relation symbols that occur in φ, and let $q \geqslant 0$ be the quantifier rank of φ.

- In *Step (1)*, the algorithm of Corollary 6.1 computes an FO($\mathbf{D}$)[σ]-formula $\tilde{\varphi}$ that is equivalent to φ and that has displacement $\leqslant w(\varphi)$ and generalised quantifier rank q. This takes time in

$$\|\varphi\| \cdot w(\varphi)^{\mathcal{O}(q)} \quad \subseteq \quad 2^{\|\varphi\|^{\mathcal{O}(1)}}$$

For the latter inclusion, recall that $q, w(\varphi), \|\sigma\| < \|\varphi\|$.

- In *Step (3)*, the algorithm of Lemma 5.3 computes a HNF $\tilde{\psi} \in \text{FO}(\mathbf{D})[\sigma]$ that is d-equivalent to $\tilde{\varphi}$. According to Corollary 6.4, for $K, P \leqslant \|\varphi\|$ this takes time

$$\max\{K,P\}^{\left(\|\tilde{\varphi}\|\cdot d^{4^q+1}\right)^{\mathcal{O}(\|\sigma\|)}} \quad\subseteq\quad 2^{d^{2^{\mathcal{O}(\|\varphi\|)}}}.$$

Note that $\tilde{\psi}$ has locality radius $\leqslant 4^q$ and displacement $\leqslant$

$$(q{+}1)\cdot\left(\|\tilde{\varphi}\|\cdot d^{4^q+1}+\max\{K,P\}\right) \quad\subseteq\quad d^{2^{\mathcal{O}(\|\varphi\|)}}.$$

- In *Step (4)*, the algorithm of Lemma 5.6 computes for each Hanf-formula in $\tilde{\psi}$ an equivalent $\text{FO}(\mathbf{Q})[\sigma]$-formula in HNF. For each Hanf-formula in $\tilde{\psi}$ of the shape $(\mathsf{Q}{+}k)y\,\gamma$, this takes time $\mathcal{O}((\|\mathsf{Q}\|{+}k)^3\cdot\|\gamma\|)$. Hence, to compute a HNF $\psi \in \text{FO}(\mathbf{Q})[\sigma]$ that is equivalent to $\tilde{\psi}$, takes time

$$\|\tilde{\psi}\|\cdot\left(\|\varphi\|+d^{2^{\mathcal{O}(\|\varphi\|)}}\right)^3 \quad\subseteq\quad 2^{d^{2^{\mathcal{O}(\|\varphi\|)}}}.$$

ψ has locality radius $\leqslant 4^q$ and displacement in $d^{2^{\mathcal{O}(\|\varphi\|)}}$.

Altogether, the proof of Theorem 3.3 is complete, since ψ is d-equivalent to φ and can be computed in time

$$2^{d^{2^{\mathcal{O}(\|\varphi\|)}}}. \qquad \square$$

7. Conclusion

We have generalised the notion of Hanf normal forms (HNF, for short) from first-order logic FO to first-order logic with unary counting qantifiers $\text{FO}(\mathbf{Q})$.

Our first main result (see Theorem 3.2) completely characterizes those sets Q of unary counting quantifiers that permit HNF: the logic $\text{FO}(\mathbf{Q})$ permits HNF if, and only if, all sets in $\mathbf{Q}$ are ultimately periodic.

Our second main result (see Theorem 3.3) provides an algorithm which, for any set $\mathbf{Q}$ of ultimately periodic sets and any degree bound $d \in \mathbb{N}$, transforms an input $\text{FO}(\mathbf{Q})$-formula φ into an $\text{FO}(\mathbf{Q})$-formula in HNF that is equivalent to φ on all (finite) structures of degree at most d. We showed that this algorithm uses time at most 3-fold exponential in the size of φ. A lower bound of [1] shows that already for $d = 3$ and plain first-order logic FO, this is worst-case optimal. A more refined runtime analysis (that will be included in the paper's full version) shows that for the degree bound $d = 2$, our algorithm can be implemented in such a way that its runtime is only 2-fold exponential in the size of $\varphi \in \text{FO}(\mathbf{Q})$.

For future work, we plan to consider relaxed variants of HNF, where instead of the Hanf-formulas defined in the current paper, more general formulas are allowed. E.g., formulas of the form $(\mathsf{Q}{+}k)\overline{y}\ \text{sph}_\tau(\overline{x},\overline{y})$, stating that "the *number of tuples* $\overline{y}$ satisfying $\mathcal{N}_r(\overline{x},\overline{y}) \cong \tau$, belongs to the set $(\mathsf{Q}{+}k)$". Or, formulas of the form $(\mathsf{Q}{+}k)y\ \psi(\overline{x},y)$, where $\psi(\overline{x},y)$ is local around its free variables, but is not required to describe the isomorphism type of an r-neighbourhood around $\overline{x},y$. Finally, it is not clear whether the results of this paper can be extended to infinite structures.

References

[1] B. Bollig and D. Kuske. An optimal construction of Hanf sentences. *J. Applied Logic*, 10(2):179–186, 2012.

[2] A. Dawar, M. Grohe, S. Kreutzer, and N. Schweikardt. Model Theory Makes Formulas Large. In *Proc. ICALP'07*, pages 1076–1088, 2007. Full version available as preprint NI07003-LAA, Isaac Newton Institute of Mathematical Sciences (2007).

[3] A. Durand and E. Grandjean. First-order queries on structures of bounded degree are computable with constant delay. *ACM Trans. Comput. Log.*, 8(4), 2007.

[4] H.-D. Ebbinghaus and J. Flum. *Finite Model Theory*. Springer, 1999.

[5] R. Fagin, L. Stockmeyer, and M. Vardi. On monadic NP vs. monadic co-NP. *Information and Computation*, 120(1):78–92, 1995.

[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 J. Stern, editor, *Proceedings of the Herbrand Symposium, Logic Colloquium '81*, pages 105–135. North Holland, 1982.

[8] W. Hanf. Model-theoretic methods in the study of elementary logic. In J. Addison, L. Henkin, and A. Tarski, editors, *The Theory of Models*, pages 132–145. North Holland, 1965.

[9] L. Hella, L. Libkin, and J. Nurmonen. Notions of locality and their logical characterizations over finite models. *J. Symb. Log.*, 64(4):1751–1773, 1999.

[10] W. Kazana and L. Segoufin. First-order query evaluation on structures of bounded degree. *Logical Methods in Computer Science*, 7(2), 2011.

[11] S. Kreutzer. Algorithmic meta-theorems. In *Finite and Algorithmic Model Theory*. Cambridge University Press, 2011. London Mathematical Society Lecture Notes, No. 379.

[12] L. Libkin. *Elements of Finite Model Theory*. Springer, 2004.

[13] S. Lindell. A normal form for first-order logic over doubly-linked data structures. *Int. J. Found. Comp. Sci.*, 19(1):205–217, 2008.

[14] M. Lothaire. *Combinatorics on words*. Cambridge University Press, 1984.

[15] F. Neven, N. Schweikardt, F. Servais, and T. Tan. Distributed streaming with finite memory. In *18th International Conference on Database Theory (ICDT 2015)*, pages 324–341, 2015.

[16] J. Nurmonen. Counting modulo quantifiers on finite structures. *Inf. Comput.*, 160(1-2):62–87, 2000.

[17] T. Schwentick and K. Bartelmann. Local normal forms for first-order logic with applications to games and automata. *Discrete Mathematics and Computer Science*, 3:109–124, 1999.

[18] D. Seese. Linear time computable problems and first-order descriptions. *Math. Struc. in Comp. Sci.*, 6(6):505–526, 1996.

[19] L. Segoufin. A glimpse on constant delay enumeration (invited talk). In *31st International Symposium on Theoretical Aspects of Computer Science (STACS 2014)*, pages 13–27, 2014.

APPENDIX

This appendix contains proofs omitted in Section 6.

A. Proof of Corollary 6.1

Let φ be an $\mathrm{FO}(\mathbf{Q})$-formula of quantifier rank $q \geqslant 0$. The algorithm proceeds by induction on the shape of φ, and uses the formulas constructed in the proof of Lemma 5.1.

The only interesting step is the case of an $\mathrm{FO}(\mathbf{Q})$-formula φ of shape $\mathsf{Q}y\,\varphi'$ for a $\mathsf{Q} \in \mathbf{Q}$ and with quantifier rank $q \geqslant 1$. As the first step, the algorithm is called recursively to compute an $\mathrm{FO}(\mathbf{D})[\sigma]$-formula ψ' that is equivalent to φ' and that has displacement $\leqslant w(\varphi')$ and generalised quantifier rank $q{-}1$.

If Q is the existential quantifier, the algorithm can output $\psi := \exists y\ \psi'$. Otherwise, the algorithm uses formula (3) from the proof of Lemma 5.1 to obtain an $\mathrm{FO}(\mathbf{D})$-formula ψ that is equivalent to $\mathsf{Q}y\,\psi'$. Recall that $w(\varphi) \geqslant \|\mathsf{Q}\| = n_0{+}p{+}1$, where p and n_0 are the minimal period and the minimal offset of Q. Hence, ψ has size $\mathcal{O}(w(\varphi)^3 \cdot \|\psi'\|)$, displacement $\leqslant w(\varphi)$ and generalised quantifier rank q.

For the algorithm's runtime analysis note that the only step which increases the formula size is the one for the quantifiers in $\mathbf{Q}$. Hence, $\|\psi\|$ is at most $\|\varphi\| \cdot w(\varphi)^{\mathcal{O}(q)}$. It is easy to see that ψ can also be constructed within the same time bound. □

B. Proof of Corollary 6.2

We proceed along the three cases provided in the proof of Lemma 5.2. Note that in any case, the algorithm has recursion depth $\mathcal{O}(\log(j{-}i{+}1))$.

Observe that the formula used in Case 1 is a Boolean combination of $2k$ quantified subformulas. Hence, the number of Hanf-formulas in $\delta^{(\exists+k)}$ is in $k^{\mathcal{O}(\log(j-i+1))}$. Furthermore, each of these Hanf-formulas has size $\mathcal{O}(k^2)$.

In the same way it can be shown for Case 2 that the number of Hanf-formulas in $\delta^{(\mathsf{D}_p+k)}$ (for $k < p$) is $p^{\mathcal{O}(\log(j-i+1))}$, and that each of these Hanf-formulas has size in $\mathcal{O}(p^2)$.

It follows, that the number of Hanf-formulas in the HNF $\delta^{(\mathsf{D}_p+k)}$ (for $k \geqslant p$), constructed in Case 3, is $\max\{k,p\}^{\mathcal{O}(\log(j-i+1))}$. Furthermore, each of these Hanf-formulas has size $\mathcal{O}(\max\{k,p\}^2)$.

Altogether, it follows that the HNF $\delta^{(\mathsf{Q}+k)}$ can be constructed in time $\max\{k,p\}^{\mathcal{O}(\log(j-i+1))}$. □

C. Proof of Corollary 6.3

In the following, we let $N := n \cdot d^{4r+1}$, and we use the same notation as in the proof of Lemma 5.5.

Recall that ρ has at most $(n{+}1)\cdot d^{r+1}$ elements. Therefore, the time needed do decide whether $\mathit{dist}^{\rho}(\overline{c},d) \leqslant 2r{+}1$ or $\mathit{dist}^{\rho}(\overline{c},d) > 2r{+}1$ is in $N^{\mathcal{O}(\|\sigma\|)}$.

If $\mathit{dist}^{\rho}(\overline{c},d) \leqslant 2r{+}1$, we have to compute the number k_τ defined in equation (6). This requires us to check at most $n \cdot d^{(2r+1)+1} \leqslant N$ d-bounded σ-structures with each at most $(n{+}1) \cdot d^{r+1} \leqslant N$ elements for isomorphism. Hence, in this case, the algorithm uses time $2^{N^{\mathcal{O}(\|\sigma\|)}}$.

If $\mathit{dist}^{\rho}(\overline{c},d) > 2r{+}1$, we have to check whether the two structures $\mathcal{N}_r^{\rho}(\overline{c})$ and $\mathcal{N}_r^{\tau}(\overline{e})$ (each with $\leqslant N$ elements) are isomorphic. If this is the case, we have to compute the number ℓ_τ defined in equation (8). This requires us to check at most $n \cdot d^{(2r+1)+1}$ d-bounded σ-structures, each with at most $n \cdot d^{r+1} \leqslant N$ elements, for isomorphism. This can be done in time $2^{N^{\mathcal{O}(\|\sigma\|)}}$. Since $\ell_\tau \leqslant N$, the formula α_τ can be computed in time $\mathcal{O}((k{+}N)^2) + N^{\mathcal{O}(\|\sigma\|)}$.

We can conclude that altogether, the algorithm of Lemma 5.5 can be carried out in time

$$(k{+}N)^{\mathcal{O}(1)} + 2^{N^{\mathcal{O}(\|\sigma\|)}} \quad\subseteq\quad k \cdot 2^{(n\cdot d^{4r+1})^{\mathcal{O}(\|\sigma\|)}}.$$

□

D. Proof of Corollary 6.4

For proving Corollary 6.4, it suffices to extend the statement of *Claim 5.4* by

(d) *There is a number* $c \geqslant 3$ *of size* $\mathcal{O}(\|\sigma\|)$ *such that the algorithm terminates after at most* $\max\{K,P\}^{N^c}$ *time steps, where* $N := \|\varphi\| \cdot d^{4^q+1}$.

In the following, the steps of the computation are numbered in the same way as in the description of the algorithm provided in the proof of Lemma 5.3.

If φ *is quantifier-free, i.e.,* $q = 0$, the algorithm proceeds along the following steps:

(1) Compute the set $T_0^d(n{+}1)$. Note that each type in $T_0^d(n{+}1)$ has size $\leqslant n{+}1 < N$. Furthermore, there is a number $c_1 \geqslant 1$ of size $\mathcal{O}(\|\sigma\|)$ such that $T_0^d(n{+}1)$ contains at most $2^{(n+1)^{c_1}} < 2^{N^{c_1}}$ types. Hence, there is a number $c_2 \geqslant c_1$ of size $\mathcal{O}(\|\sigma\|)$ such that $T_0^d(n{+}1)$ can be enumerated in $\leqslant 2^{N^{c_2}}$ time steps.

(2) Compute the set $T \subseteq T_0^d(n{+}1)$ of all types $\tau = (\mathcal{T},\overline{c},d)$ in $T_0^d(n{+}1)$ where $\mathcal{T} \models \varphi[\overline{c}]$. Since φ is quantifier-free, this can be done in at most $2^{N^{c_3}}$ time steps for a number $c_3 \geqslant c_1$ of size $\mathcal{O}(\|\sigma\|)$.

(3) The construction of the formula $\psi(\overline{x})$ takes at most $2^{N^{c_4}}$ time steps, for a number $c_4 \geqslant c_1$ of size $\mathcal{O}(\|\sigma\|)$.

Hence, for a suitable $c \geqslant 1$ of size $\mathcal{O}(\|\sigma\|)$, the algorithm can be carried out in at most 2^{N^c} time steps.

For φ *with generalised quantifier rank* $q \geqslant 1$, the case of Boolean combinations is trivial. For φ of the shape $(\mathsf{Q}{+}k)y\ \varphi'(\overline{x},y)$, where $\mathsf{Q} \in \{\exists\} \cup \mathbf{D}$ has period $p \geqslant 1$ and where $k \leqslant K$, the algorithm proceeds along the following steps:

(4) By the induction hypothesis (Claim 5.4 (d)), the algorithm computes the HNF $\psi'(\overline{x},y)$ in $\leqslant \max\{K,P\}^{(N-1)^c}$ time steps.

(5) Since the set $T^d_{4^q}(n{+}1)$ contains $t \leqslant 2^{N^{c_1}}$ types of size $< N$, it can be enumerated in $\leqslant 2^{N^{c_2}}$ time steps.

(6) By Corollary 6.2, there is a $c_4 \geqslant c_1$ of size $\mathcal{O}(\|\sigma\|)$, such that it takes $\leqslant \max\{k, p\}^{N^{c_4}} \leqslant \max\{K, P\}^{N^{c_4}}$ time steps to construct the HNF $\delta^{(\mathsf{Q}+k)}_{[1,t]} \in$ FO(**D**)$[\sigma_{[1,t]}]$.

(7) For each $s \in [1, t]$, compute the FO(**D**)$[\sigma]$-sentence ψ'_s. Since we know that $(n{+}1) \cdot d^{4r+1} \leqslant N$ and that furthermore $k' := q \cdot (N{-}1) + \max\{K, P\}$, we obtain from Corollary 6.3 that this takes time $k' \cdot 2^{N^{\mathcal{O}(\|\sigma\|)}} \subseteq \max\{K, P\}^{N^{\mathcal{O}(\|\sigma\|)}}$ for each Hanf-formula in ψ'. Since $\|\psi'\| \leqslant 2^{(N-1)^c}$, there is a number $c_5 \geqslant 1$ of size $\mathcal{O}(\|\sigma\|)$ such that it takes $\leqslant \max\{K, P\}^{(N-1)^c} \cdot \max\{K, P\}^{N^{c_5}}$ time steps to construct ψ'_s.

(8) Replace each Hanf-formula of the shape $(\mathsf{R}{+}\ell)y\, P_s(y)$ in $\delta^{(\mathsf{Q}+k)}_{[1,t]}$ by the HNF $\psi^{(\mathsf{R}+\ell)}_s(\overline{x})$. By the size of $\delta^{(\mathsf{Q}+k)}_{[1,t]}$, this takes at most $\max\{K, P\}^{N^{c_5}}$ time steps, for a number $c_6 \geqslant c_4$ of size $\mathcal{O}(\|\sigma\|)$.

Altogether, the algorithm takes $\leqslant \max\{K, P\}^{(N-1)^c + N^{c_6}}$ time steps, for a number $c_7 \geqslant \max\{c_2, c_4, c_5, c_6\}$ of size $\mathcal{O}(\|\sigma\|)$.

Choosing $c := c_7{+}1$ enforces that $(N{-}1)^c{+}N^{c_7} \leqslant N^c$, since

$$\tfrac{(N-1)^c + N^{c-1}}{N^c} \;=\; \left(\tfrac{N-1}{N}\right)^c + \tfrac{1}{N} \;\leqslant\; \tfrac{N-1}{N} + \tfrac{1}{N} \;=\; 1.$$

It follows, that, the algorithm terminates after at most $\max\{K, P\}^{N^c} = \max\{K, P\}^{(n \cdot d^{4^q+1})^c}$ time steps.

This completes the proof of *Claim 5.4 (d)* and hence, the proof of Corollary 6.4. □