First-Order Logic with Counting
At Least, *Weak* Hanf Normal Forms Always Exist and Can Be Computed!

Dietrich Kuske
TU Ilmenau

Nicole Schweikardt
Humboldt-Universität zu Berlin

***Abstract*—We introduce the logic FOCN($\mathbb{P}$) which extends first-order logic by counting and by numerical predicates from a set $\mathbb{P}$, and which can be viewed as a natural generalisation of various counting logics that have been studied in the literature.**

We obtain a locality result showing that every FOCN($\mathbb{P}$)-formula can be transformed into a formula in Hanf normal form that is equivalent on all finite structures of degree at most d. A formula is in Hanf normal form if it is a Boolean combination of formulas describing the neighbourhood around its tuple of free variables and arithmetic sentences with predicates from $\mathbb{P}$ over atomic statements describing the number of realisations of a type with a single centre. The transformation into Hanf normal form can be achieved in time elementary in d and the size of the input formula. From this locality result, we infer the following applications:

(1) The Hanf-locality rank of first-order formulas of bounded quantifier alternation depth only grows polynomially with the formula size.

(2) The model checking problem for the fragment FOC($\mathbb{P}$) of FOCN($\mathbb{P}$) on structures of bounded degree is fixed-parameter tractable (with elementary parameter dependence).

(3) The query evaluation problem for fixed queries from FOC($\mathbb{P}$) over fully dynamic databases of degree at most d can be solved efficiently: there is a dynamic algorithm that can enumerate the tuples in the query result with constant delay, and that allows to compute the size of the query result and to test if a given tuple belongs to the query result within constant time after every database update.

1. Introduction

The counting ability of first-order logic is very limited: it can only make statements of the form "there are at least k witnesses x for $\varphi(x)$" for a constant $k \in \mathbb{N}$. To overcome this problem, one can add number variables κ to first-order logic and means to express that κ equals the number of witnesses for the formula $\varphi(x)$. In order to make use of these number variables, one also adds numerical functions like addition and numerical predicates like $\kappa \leqslant \kappa'$ or "κ is a prime" to the logic. These and similar ideas led to the extensions of first-order logic by the Rescher quantifier, the Härtig quantifier, or arbitrary unary counting quantifiers [13], [24], [27], to logics like FO(D_p) from [23], FO(Cnt) from [21], and FO+C from [11]. In this paper, we introduce an extension FOCN($\mathbb{P}$) of first-order logic by counting, number variables, and numerical predicates from a set $\mathbb{P}$. By choosing $\mathbb{P}$ appropriately, we use this extension as a general framework for counting-extensions of first-order logic (it subsumes all the logics mentioned above).

Clearly, two isomorphic graphs cannot be distinguished by logical sentences. Even more: suppose there is a bijection between the vertex sets of two undirected graphs $\mathcal{A}$ and $\mathcal{B}$ such that, for every node of $\mathcal{A}$, the neighbourhood of radius $2^{\mathcal{O}(q)}$ of that node is isomorphic to the neighbourhood of its image. Then, first-order logic cannot distinguish the two graphs by first-order sentences of quantifier rank at most q (this goes back to [12]; the actual bound $2^{q-1} - 1$ was obtained in [20]). Consequently, to determine whether a sentence φ of quantifier rank q holds in an undirected graph $\mathcal{A}$, it suffices to count how often each neighbourhood type of radius $2^{\mathcal{O}(q)}$ is realised in $\mathcal{A}$.

It actually suffices to count these realisations up to a certain threshold (that depends on q and the degree d of the graph $\mathcal{A}$) [8]. Bounding the degree of $\mathcal{A}$ by d, there are only finitely many neighbourhood types of radius $2^{\mathcal{O}(q)}$ that can be realised. Consequently, this condition can be expressed as a first-order sentence; i.e., as a first-order sentence in *Hanf normal form*.

A similar story can be told, e.g., for the extension FO(D_2) of first-order logic by the ability to express that the number of witnesses for $\varphi(x)$ is even. To determine whether such a sentence holds in an undirected graph $\mathcal{A}$, one has to count the number of realisations up to a certain threshold and one has to determine the parity of this number [23]. Again, this leads to an FO(D_2)-sentence in Hanf normal form that expresses the said condition in the graph $\mathcal{A}$.

We say that a logic can only express local properties if validity of a sentence in a structure can be determined by solely counting the number of realisations of neighbourhood types. This property has traditionally been proven by suitable notions of games. Often, the *existence* of a Hanf normal form follows from this directly. But there is no obvious way to extract an *algorithm* for the construction of it. On the other hand, these Hanf normal forms have also found various applications in algorithms and complexity (cf., e.g., [2], [3], [6], [9], [14], [16], [17], [21], [25], [26]). In particular, there are very general algorithmic meta-theorems stating that model checking is fixed-parameter tractable for various classes of structures, and that the results of queries

978-1-5090-3018-7/17/$31.00

against various classes of databases can be enumerated with constant delay after a linear-time preprocessing phase. In this context, questions about the efficiency of the normal forms have attracted interest (cf. e.g., [3], [5], [14], [22]).

The main result of this paper is the effective construction of a Hanf normal form from an arbitrary formula of our logic FOCN($\mathbb{P}$). This construction extends the constructions from [3], [14] and can be carried out in 5-fold exponential time. We also provide a 4-fold exponential lower bound for the fragment FO($\mathbb{P}$) and a restricted form of Hanf normal forms. From the existence and the computability of Hanf normal forms, we infer four applications:

(1) The model checking problem for the (large) fragment FOC($\mathbb{P}$) of the logic FOCN($\mathbb{P}$) on structures of bounded degree is fixed-parameter tractable (with elementary parameter dependence) where we assume an oracle for the numerical predicates from $\mathbb{P}$.

(2) The Hanf-locality rank of first-order formulas of bounded quantifier alternation depth only grows polynomially with the formula size. This complements Libkin's bound $2^{q-1}-1$ for q the quantifier rank of the formula [20] and (partly) proves a conjecture from [19].

(3) For a sentence φ in FOCN($\mathbb{P}$), we can compute a first-order description of the numerical condition that is equivalent to validity of φ. This description is expressed in an extension of integer arithmetic with predicates from $\mathbb{P}$.

(4) The query evaluation problem for fixed queries from FOC($\mathbb{P}$) over fully dynamic databases of degree $\leqslant d$ can be solved efficiently: there is a dynamic algorithm that can enumerate the tuples in the query result with constant delay, and that allows to compute the size of the query result and to test if a given tuple belongs to the query result within constant time after every database update.

Above, we said that the existence of a Hanf normal form follows "often". A counterexample to this is the fragment FO($\mathbb{P}$) of FOCN($\mathbb{P}$) that we consider in [14]. The problem there is that, in general, FO($\mathbb{P}$) does not allow to formulate the necessary numerical condition. In Corollary 3.3 we present a weakening of the notion of a Hanf normal form that also works in this case.

The rest of the paper is structured as follows. Sections 2 and 3 introduce the logic FOCN($\mathbb{P}$) and the according notion of Hanf normal form. Theorem 3.2 summarises the paper's technical main result, the proof of which is given in Sections 4 and 6. Section 5 describes the mentioned applications. Due to space restrictions, many details had to be omitted; they will appear in the paper's full version (a preprint is available at https://arxiv.org/abs/1703.01122).

Acknowledgements. The second author would like to acknowledge the financial support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under grant SCHW 837/5-1.

2. First-order logic with counting and numerical predicates

We write $\mathbb{Z}$, $\mathbb{N}$, and $\mathbb{N}_{\geqslant 1}$ for the sets of integers, non-negative integers, and positive integers, resp. For $S \subseteq \mathbb{Z}$ and $k \in \mathbb{Z}$ we let $S+k = \{s+k : s \in S\}$. For all $m,n \in \mathbb{N}$, we write $[m,n]$ for the set $\{k \in \mathbb{N} : m \leqslant k \leqslant n\}$ and $[m] = [1,m]$. For a k-tuple $\overline{x} = (x_1,\dots,x_k)$ we write $|\overline{x}|$ to denote its *length* k. The exponential functions $\exp_k \colon \mathbb{N} \to \mathbb{N}$ are defined by induction on k via $\exp_0(n) = n$ and $\exp_{k+1}(n) = 2^{\exp_k(n)}$ for all $k \in \mathbb{N}$. We write $\operatorname{poly}(n)$ for the set of functions $\bigcup_{k\in\mathbb{N}} \mathcal{O}(n^k)$.

A *signature* σ is a *finite* set of relation and constant symbols. Associated with every relation symbol $R \in \sigma$ is a positive integer $\operatorname{ar}(R)$ called the *arity* of R. The *size* $\|\sigma\|$ of a signature σ is the number of its constant symbols plus the sum of the arities of its relation symbols. We call a signature *relational* if it does not contain any constant symbol. A σ*-structure* $\mathcal{A}$ consists of a *finite* non-empty set A called the *universe* of $\mathcal{A}$, a relation $R^{\mathcal{A}} \subseteq A^{\operatorname{ar}(R)}$ 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 these definitions, all signatures and all structures considered in this paper are *finite*. To indicate that two σ-structures $\mathcal{A}$ and $\mathcal{B}$ are isomorphic, we write $\mathcal{A} \cong \mathcal{B}$.

In the following, we define the extension FOCN($\mathbb{P}$) of first-order logic FO by counting and by numerical predicates from a set $\mathbb{P}$. Our notation extends standard notation concerning first-order logic, cf. [7], [21].

Let vars and nvars be fixed disjoint countably infinite sets of *structure variables* and *number variables*, respectively. In our logic, structure variables from vars will always denote elements of the structure, and number variables from nvars will denote integers. Typical structure variables are x and y, typical number variables are λ and κ. Often, we use z as an arbitrary variable from $\mathsf{vars} \cup \mathsf{nvars}$.

A σ*-interpretation* $\mathcal{I} = (\mathcal{A}, \beta)$ consists of a σ-structure $\mathcal{A}$ and an *assignment* β *in* $\mathcal{A}$, i.e., $\beta \colon \mathsf{vars} \cup \mathsf{nvars} \to A \cup \mathbb{Z}$ with $\beta(x) \in A$ for $x \in \mathsf{vars}$ and $\beta(\kappa) \in \mathbb{Z}$ for $\kappa \in \mathsf{nvars}$. For $k,\ell \in \mathbb{N}$, for $a_1,\dots,a_k \in A$, $n_1,\dots,n_\ell \in \mathbb{Z}$, and for pairwise distinct $y_1,\dots,y_k \in \mathsf{vars}$ and $\kappa_1,\dots,\kappa_\ell \in \mathsf{nvars}$, we write $\beta\frac{a_1,\dots,a_k}{y_1,\dots,y_k}\frac{n_1,\dots,n_\ell}{\kappa_1,\dots,\kappa_\ell}$ for the assignment β' in $\mathcal{A}$ with $\beta'(y_j) = a_j$ for all $j \in [k]$, $\beta(\kappa_j) = n_j$ for all $j \in [\ell]$, and $\beta'(z) = \beta(z)$ for all $z \in (\mathsf{vars} \cup \mathsf{nvars}) \setminus \{y_1,\dots,y_k,\kappa_1,\dots,\kappa_\ell\}$. For $\mathcal{I} = (\mathcal{A},\beta)$ we let $\mathcal{I}\frac{a_1,\dots,a_k}{y_1,\dots,y_k}\frac{n_1,\dots,n_\ell}{\kappa_1,\dots,\kappa_\ell} = \big(\mathcal{A}, \beta\frac{a_1,\dots,a_k}{y_1,\dots,y_k}\frac{n_1,\dots,n_\ell}{\kappa_1,\dots,\kappa_\ell}\big)$.

Definition 2.1 Let σ be a signature. The set of FO[σ]-*formulas* is built according to the following rules:

(1) $x_1{=}x_2$ and $R(x_1,\dots,x_{\operatorname{ar}(R)})$ are *formulas*, where $R \in \sigma$ and $x_1,x_2,\dots,x_{\operatorname{ar}(R)}$ are structure variables or constant symbols in σ
(2) if φ and ψ are *formulas*, then so are $\neg\varphi$ and $(\varphi \vee \psi)$
(3) if φ is a *formula* and $y \in \mathsf{vars}$, then $\exists y\, \varphi$ is a *formula*

The semantics $[\![\varphi]\!]^{\mathcal{I}} \in \{0,1\}$ for a σ-interpretation $\mathcal{I} = (\mathcal{A},\beta)$ and a FO[σ]-formula φ is defined as usual.

In a first step, we extend first-order logic such that numerical statements on the number of witnesses for a formula are possible. These numerical statements are based on numerical predicates that we define first.

Definition 2.2 A *numerical predicate collection* is a triple $(\mathbb{P}, \mathrm{ar}, [\![.]\!])$ where $\mathbb{P}$ is a countable set of *predicate names*, $\mathrm{ar}\colon \mathbb{P} \to \mathbb{N}_{\geqslant 1}$ assigns the *arity* to every predicate name, and $[\![\mathsf{P}]\!] \subseteq \mathbb{Z}^{\mathrm{ar}(\mathsf{P})}$ is the *semantics* of the predicate name $\mathsf{P} \in \mathbb{P}$.

Basic examples of numerical predicates are P_+, $\mathsf{P}_\cdot$, $\mathsf{P}_=$, $\mathsf{P}_\leqslant$, Prime with $[\![\mathsf{P}_+]\!] = \{(m,n,m+n) : m,n \in \mathbb{Z}\}$, $[\![\mathsf{P}_\cdot]\!] = \{(m,n,m\cdot n) : m,n \in \mathbb{Z}\}$, $[\![\mathsf{P}_=]\!] = \{(m,m) : m \in \mathbb{Z}\}$, $[\![\mathsf{P}_\leqslant]\!] = \{(m,n) \in \mathbb{Z}^2 : m \leqslant n\}$, and $[\![\mathsf{Prime}]\!] = \{n \in \mathbb{N} : n \text{ is a prime number}\}$. Also, D_p with $[\![\mathsf{D}_p]\!] = p\mathbb{Z}$ (for each fixed $p \in \mathbb{N}_{\geqslant 1}$) and the halting problem (i.e., the set of indices of Turing machines that halt with empty input) are possible numerical predicates.

Definition 2.3 Let σ be a signature and $(\mathbb{P}, \mathrm{ar}, [\![.]\!])$ be a numerical predicate collection. The sets of *formulas* and *counting terms for* $\mathrm{FO}(\mathbb{P})[\sigma]$ are built according to the rules (1)–(3) and the following rules:

(4) if $\mathsf{P} \in \mathbb{P}$, $m = \mathrm{ar}(\mathsf{P})$, and $t_1, \ldots, t_m$ are *counting terms*, then $\mathsf{P}(t_1, \ldots, t_m)$ is a *formula*

(5') if φ is a formula, $y \in \mathsf{vars}$, and $k \in \mathbb{N}$, then $\#(y).\varphi - k$ is a *counting term*.

We will write $\#(y).\varphi$ as a shorthand for the counting term $\#(y).\varphi - 0$.

Let $\mathcal{I} = (\mathcal{A}, \beta)$ be a σ-interpretation. For every formula φ and every counting term t from $\mathrm{FO}(\mathbb{P})[\sigma]$, the semantics $[\![\varphi]\!]^{\mathcal{I}} \in \{0,1\}$ of φ in $\mathcal{I}$ and the semantics $[\![t]\!]^{\mathcal{I}} \in \mathbb{Z}$ of t in $\mathcal{I}$ extend the definition for $\mathrm{FO}[\sigma]$-formulas as follows:

(4) $[\![\mathsf{P}(t_1, \ldots, t_m)]\!]^{\mathcal{I}} = 1$ iff $\big([\![t_1]\!]^{\mathcal{I}}, \ldots, [\![t_m]\!]^{\mathcal{I}}\big) \in [\![\mathsf{P}]\!]$

(5') $[\![\#(y).\varphi - k]\!]^{\mathcal{I}} = \big|\big\{a \in A \,:\, [\![\varphi]\!]^{\mathcal{I}\frac{a}{y}} = 1\big\}\big| - k$

Remark 2.4 For $\mathrm{FO}(\mathbb{P})$ and the following logics $\mathrm{FOC}(\mathbb{P})$ and $\mathrm{FOCN}(\mathbb{P})$, *expressions* are formulas or counting terms.

Example 2.5 For structure variables $y \in \mathsf{vars}$, the quantifier $\exists y$ can be replaced by using a suitable numerical predicate: Let $\mathsf{P}_\exists$ be the numerical predicate with $\mathrm{ar}(\mathsf{P}_\exists) = 1$ and $[\![\mathsf{P}_\exists]\!] = \mathbb{N}_{\geqslant 1}$. Consider a σ-interpretation $\mathcal{I} = (\mathcal{A}, \beta)$. Since $\mathcal{A}$ is finite, we have $\mathcal{I} \models \exists y\, \varphi \iff \mathcal{I} \models \mathsf{P}_\exists(\#(y).\varphi)$.

The following examples provide choices of $\mathbb{P}$ for which the logic $\mathrm{FO}(\mathbb{P})$ has been studied in the literature.

Example 2.6

(a) Let $\mathbb{E} = \{\exists^{\geqslant k} : k \in \mathbb{N}_{\geqslant 1}\}$ with $\mathrm{ar}(\exists^{\geqslant k}) = 1$ and $[\![\exists^{\geqslant k}]\!] = \{k, k{+}1, \ldots\}$ for every $k \geqslant 1$. The logic $\mathrm{FO}(\mathbb{E})$ is equivalent to the logic $\mathrm{FO}(\mathrm{C})$ of [7].

(b) The logic $\mathrm{FO}(\{\mathsf{D}_p\})$ is equivalent to the extension of first-order logic by the divisibility quantifier D_p, considered by Nurmonen in [23].

(c) Let $(\mathbb{P}, \mathrm{ar}, [\![.]\!])$ be a numerical predicate collection with $\mathrm{ar}(\mathsf{P}) = 1$ for all $\mathsf{P} \in \mathbb{P}$. Then $\mathrm{FO}(\mathbb{P})$ is equivalent to the logic considered in [14].

(d) The logics $\mathrm{FO}(\{\mathsf{P}_\leqslant\})$ and $\mathrm{FO}(\{\mathsf{P}_=\})$ are equivalent to the extension of first-order logic by the *Rescher quantifier* and the *Härtig quantifier*, resp. [13], [24].

(e) Let $\mathbb{U} = \{0,1\}^*\$\{0,1\}^+$ and, for $u\$v \in \mathbb{U}$, let $i \in [\![u\$v]\!]$ if, and only if, the ω-word uv^ω carries a 1 at position $i \in \mathbb{N}$. Then $\{[\![u\$v]\!] : u\$v \in \mathbb{U}\}$ is the set of ultimately periodic (or semilinear) subsets of $\mathbb{N}$. The logic $\mathrm{FO}(\mathbb{U})$ is equivalent to the extension of first-order logic by ultimately periodic unary counting quantifiers, considered in [14].

(f) In [21, Sect. 8.1], Libkin considers the extension FO(unary) of first-order logic by the class of all unary generalised quantifiers. It is not difficult to see that every FO(unary)-formula is equivalent to an $\mathrm{FO}(\mathbb{P})$-formula, for a suitable numerical predicate collection $(\mathbb{P}, \mathrm{ar}, [\![.]\!])$.

Our next logic $\mathrm{FOC}(\mathbb{P})$ allows not only numerical statements on numbers given by counting terms of the form $\#(y).\varphi - k$, but on polynomials over such terms. In addition, the logic $\mathrm{FOC}(\mathbb{P})$ allows to count tuples.

Definition 2.7 Let σ be a signature and let $(\mathbb{P}, \mathrm{ar}, [\![.]\!])$ be a numerical predicate collection. The set of *expressions for* $\mathrm{FOC}(\mathbb{P})[\sigma]$ is built according to the rules (1)–(4) and the following rules:

(5) if φ is a *formula*, $k \in \mathbb{N}_{\geqslant 1}$, and $\overline{y} = (y_1, \ldots, y_k)$ is a tuple of pairwise distinct structure variables (i.e., variables in vars), then $\#\overline{y}.\varphi$ is a *counting term*

(6) every integer $i \in \mathbb{Z}$ is a *counting term*

(7) if t_1 and t_2 are *counting terms*, then so are $(t_1 + t_2)$ and $(t_1 \cdot t_2)$

Let $\mathcal{I} = (\mathcal{A}, \beta)$ be a σ-interpretation. For every expression ξ of $\mathrm{FOC}(\mathbb{P})[\sigma]$, the semantics $[\![\xi]\!]^{\mathcal{I}}$ is given by the semantics for the rules (1)–(4) and the following:

(5) $[\![\#\overline{y}.\varphi]\!]^{\mathcal{I}} = \big|\big\{(a_1, \ldots, a_k) \in A^k \,:\, [\![\varphi]\!]^{\mathcal{I}\frac{a_1,\ldots,a_k}{y_1,\ldots,y_k}} = 1\big\}\big|$, where $\overline{y} = (y_1, \ldots, y_k)$

(6) $[\![i]\!]^{\mathcal{I}} = i$

(7) $[\![(t_1{+}t_2)]\!]^{\mathcal{I}} = [\![t_1]\!]^{\mathcal{I}} + [\![t_2]\!]^{\mathcal{I}}$, $[\![(t_1 \cdot t_2)]\!]^{\mathcal{I}} = [\![t_1]\!]^{\mathcal{I}} \cdot [\![t_2]\!]^{\mathcal{I}}$

If s and t are counting terms, then we write $s - t$ for the counting term $(s + ((-1) \cdot t))$. With this convention, we can understand $\mathrm{FO}(\mathbb{P})$ as a fragment of $\mathrm{FOC}(\mathbb{P})$. Note that counting terms of $\mathrm{FOC}(\mathbb{P})$ are polynomials while counting terms of $\mathrm{FO}(\mathbb{P})$ are special linear polynomials. In addition, counting terms of $\mathrm{FOC}(\mathbb{P})$ can count *tuples* of elements of the universe while counting terms of $\mathrm{FO}(\mathbb{P})$ only count single *elements* of the universe.

Example 2.8 The following $\mathrm{FOC}(\{\mathsf{Prime}\})$-formula (expressing that the sum of the numbers of nodes and edges of a graph is a prime) is not an $\mathrm{FO}(\{\mathsf{Prime}\})$-formula:

$$\mathsf{Prime}\,\big(\,(\,\#(x).x{=}x \;+\; \#(x,y).E(x,y)\,)\,\big)\,.$$

Our final extension of the logic allows besides structure variables also number variables, and it allows to quantify over "small" numbers, i.e., over numbers in $\{0, 1, \ldots, |A|\}$, when evaluated in a σ-structure $\mathcal{A}$:

Definition 2.9 Let σ be a signature and let $(\mathbb{P}, \mathrm{ar}, [\![.]\!])$ be a numerical predicate collection. The set of *expressions for* $\mathrm{FOCN}(\mathbb{P})[\sigma]$ is built according to the rules (1)–(7) and the following rules:

(8) every variable from nvars is a *counting term*
(9) if φ is a *formula* and $\kappa \in \mathsf{nvars}$, then $\exists\kappa\, \varphi$ is a *formula*

Let $\mathcal{I} = (\mathcal{A}, \beta)$ be a σ-interpretation. For every expression ξ of FOCN($\mathbb{P}$)[σ], the semantics $[\![\xi]\!]^{\mathcal{I}}$ is given by the semantics for the rules (1)–(7) and the following:

(8) $[\![\kappa]\!]^{\mathcal{I}} = \beta(\kappa)$ for $\kappa \in \mathsf{nvars}$
(9) $[\![\exists\kappa\, \varphi]\!]^{\mathcal{I}} = \max\{[\![\varphi]\!]^{\mathcal{I}\frac{k}{\kappa}} : k \in \{0, 1, \ldots, |A|\}\}$

By FOCN($\mathbb{P}$), we denote the union of all FOCN($\mathbb{P}$)[σ] for arbitrary signatures σ, and similarly for FOC($\mathbb{P}$), FO($\mathbb{P}$), and FO.

Example 2.10 Consider the FOCN($\{\mathsf{Prime}, \mathsf{P}_=\}$)-formula

$$\exists\kappa\ \mathsf{Prime}\Big(\#(y).\mathsf{P}_=\big(\kappa,\ \#(z).E(y,z)\big)\Big)\,.$$

The counting term $\#(z).E(y,z)$ denotes the out-degree of y, hence the formula $\mathsf{P}_=\big(\kappa,\ \#(z).E(y,z)\big)$ expresses that κ is the out-degree of y. Thus, the whole formula says that there is some degree κ such that the number of nodes of out-degree κ is a prime. Since 0 is not a prime, this FOCN($\mathbb{P}$) formula is equivalent to the following FO($\mathbb{P}$)-formula

$$\exists x\ \mathsf{Prime}\bigg(\#(y).\mathsf{P}_=\big(\,\#(z).E(x,z),\ \#(z).E(y,z)\,\big)\bigg)\,.$$

Remark 2.11 The logics FO(Cnt) from [21] and FO+C from [11] can be viewed as fragments of FOCN($\mathbb{P}$) where $\mathbb{P}$ contains the predicates P_+, $\mathsf{P}_\cdot$, $\mathsf{P}_=$ and $\mathsf{P}_\leqslant$. But these two logics have no mechanism for counting *tuples*. E.g., it is not clear how to express in FO(Cnt) or FO+C that the number of edges of a graph is a square number, while this is FOCN($\mathbb{P}$)-expressible by $\exists\kappa\ \mathsf{P}_=\big(\,\#(x,y).E(x,y)\,,\ (\kappa{\cdot}\kappa)\,\big)$.

Note that we restrict the quantification over numbers to the size of the universe of the structure $\mathcal{A}$. This is analogous to the semantics of the logics FO(Cnt) and FO+C from [11], [21]. As a consequence, the logic FOCN($\mathbb{P}$)[σ] does not have the full power of integer arithmetic. Let us mention that our main result Theorem 3.2 also holds for the variant of FOCN($\mathbb{P}$) where quantifications of number variables range over arbitrary integers (rather than just numbers in $\{0, 1, \ldots, |A|\}$); the model-checking algorithm described in Section 5, however, does not carry over to this variant of FOCN($\mathbb{P}$).

The construct $\exists z$ binds the variable $z \in \mathsf{vars} \cup \mathsf{nvars}$, and the construct $\#\overline{y}$ in a counting term binds the structure variables from the tuple $\overline{y}$; all other occurrences of variables are free. We denote the set of free variables of the expression ξ by $\mathrm{free}(\xi)$ and write $\xi(\overline{z})$, for $\overline{z} = (z_1, \ldots, z_n)$ with $n \geqslant 0$, to indicate that at most the variables from $\{z_1, \ldots, z_n\}$ are free in the expression ξ. A *sentence* is a formula without free variables, a *ground term* is a counting term without free variables. Furthermore, a *number formula* is a formula whose free variables all belong to nvars. For instance, $\mathsf{P}(\kappa, \#(y).\varphi(y, \kappa))$ is a number formula, but not a sentence since $\kappa \in \mathsf{nvars}$ is free in this formula.

Note that the semantics $[\![\xi]\!]^{\mathcal{I}}$ for an expression $\xi(\overline{x}, \overline{\kappa})$ and a σ-interpretation $\mathcal{I} = (\mathcal{A}, \beta)$ only depends on $\mathcal{A}$ and on $\beta(z)$ for variables z in $\overline{x}\,\overline{\kappa}$.

Let $\overline{x} = (x_1, \ldots, x_m)$ and $\overline{\kappa} = (\kappa_1, \ldots, \kappa_n)$ be tuples of structure and number variables, resp., let $\mathcal{I} = (\mathcal{A}, \beta)$ be a σ-interpretation, and let $\overline{a} = (\beta(x_i))_{i \in [m]}$ and $\overline{\kappa} = (\beta(\kappa_j))_{j \in [n]}$. If $t(\overline{x}, \overline{\kappa})$ is a counting term, then we write $t^{(\mathcal{A}, \overline{a}, \overline{k})}$ or $t^{\mathcal{A}}[\overline{a}, \overline{\kappa}]$ for the integer $[\![t]\!]^{\mathcal{I}}$. Furthermore, for a formula $\varphi(\overline{x}, \overline{\kappa})$ we write $(\mathcal{A}, \overline{a}, \overline{k}) \models \varphi$ or $\mathcal{A} \models \varphi[\overline{a}, \overline{k}]$ to indicate that $[\![\varphi]\!]^{\mathcal{I}} = 1$, i.e., the formula $\varphi(\overline{x}, \overline{\kappa})$ is satisfied in $\mathcal{A}$ when interpreting the free occurrences of the structure variables x_i with $\beta(x_i)$ and the free occurrences of the number variables κ_i with $\beta(\kappa_i)$. In case that $m = n = 0$ (i.e., φ is a sentence and t is a ground term), we simply write $t^{\mathcal{A}}$ instead of $t^{\mathcal{A}}[\overline{a}, \overline{k}]$, and $\mathcal{A} \models \varphi$ instead of $\mathcal{A} \models \varphi[\overline{a}, \overline{k}]$.

Two expressions ξ and ξ' are *equivalent* (for short, $\xi \equiv \xi'$), if $[\![\xi]\!]^{\mathcal{I}} = [\![\xi']\!]^{\mathcal{I}}$ for every σ-interpretation $\mathcal{I}$.

The size $\|\xi\|$ of an expression is its length when viewed as a word over the alphabet $\sigma \cup \mathsf{vars} \cup \mathsf{nvars} \cup \mathbb{P} \cup \{,\} \cup \{=, \exists, \neg, \vee, (,)\} \cup \{\#, .\}$. The *number quantifier rank* $\mathrm{nqr}(\xi)$ of an FOCN($\mathbb{P}$)-expression ξ is the maximal nesting depth of quantifiers of the form $\exists\kappa$ with $\kappa \in \mathsf{nvars}$. The *binding rank* $\mathrm{br}(\xi)$ of ξ is the maximal nesting depth of constructs of the form $\exists y$ with $y \in \mathsf{vars}$ and $\#\overline{y}$ with $\overline{y}$ a tuple in vars. The *binding width* $\mathrm{bw}(\xi)$ is the maximal length $|\overline{y}|$ for $\#\overline{y}.\psi$ occurring in ξ; if ξ contains no such term, then $\mathrm{bw}(\xi) = 1$ if ξ contains an existential quantifier $\exists y$ with $y \in \mathsf{vars}$, and $\mathrm{bw}(\xi) = 0$ otherwise. Note that quantification over number variables does not influence the binding rank nor the binding width and, conversely, quantification over structure variables does not influence the number quantifier rank.

Example 2.12 The sentence $\exists x\, \mathsf{Prime}\big(\#(y).E(x,y)\big)$ has number quantifier rank 0, binding rank 2, binding width 1, and size 16. When evaluated in a directed graph $\mathcal{A} = (A, E^{\mathcal{A}})$, the sentence states that $\mathcal{A}$ contains a node whose out-degree is a prime number.

3. Hanf Normal Form

Gaifman graph and bounded structures. Let σ be a signature. The *Gaifman graph* $G_{\mathcal{A}}$ of a σ-structure $\mathcal{A}$ is the undirected 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 structure $\mathcal{A}$ is called *connected* if its Gaifman graph $G_{\mathcal{A}}$ is connected; the *connected components* of $\mathcal{A}$ are the connected components of $G_{\mathcal{A}}$. The *degree* of $\mathcal{A}$ is the degree of its Gaifman graph, i.e., the maximum number of neighbours of a node of $G_{\mathcal{A}}$. For $d \in \mathbb{N}$, a σ-structure $\mathcal{A}$ is *d-bounded* if its degree is at most d. Two formulas or two counting terms ξ and ξ' over a signature σ are *d-equivalent* ($\xi \equiv_d \xi'$, for short), if $[\![\xi]\!]^{\mathcal{I}} = [\![\xi']\!]^{\mathcal{I}}$ for every σ-interpretation $\mathcal{I} = (\mathcal{A}, \beta)$ with $\mathcal{A}$ *d-bounded*.

Let $\mathcal{A}$ be a σ-structure, $\overline{a} \in A^n$ for some $n \geqslant 1$, and $b \in A$. The *distance* $\mathrm{dist}^{\mathcal{A}}(\overline{a}, b)$ between $\overline{a}$ and b is the minimal number of edges of a path from some element of the tuple $\overline{a}$

to b in $G_{\mathcal{A}}$ (if no such path exists, we let $\operatorname{dist}^{\mathcal{A}}(\overline{a}, b) = \infty$). For every $r \geqslant 0$, the *r-neighbourhood of $\overline{a}$ in $\mathcal{A}$* is the set $N_r^{\mathcal{A}}(\overline{a}) = \{b \in A : \operatorname{dist}^{\mathcal{A}}(\overline{a}, b) \leqslant r\}$.

Types, spheres, and sphere-formulas. Let σ be a relational signature and let $c_1, c_2, \ldots$ be a sequence of pairwise distinct constant symbols. 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 structure $\tau = (\mathcal{A}, a_1, \ldots, a_n)$ over the signature $\sigma \cup \{c_1, \ldots, c_n\}$, where $\mathcal{A}$ is a σ-structure and $(a_1, \ldots, a_n) \in A^n$ with $A = N_r^{\mathcal{A}}(a_1, \ldots, a_n)$, i.e., each element of $\mathcal{A}$ is "close" to some element from $\{a_1, \ldots, a_n\}$. The elements $a_1, \ldots, a_n$ are the *centres* of τ.

Let $\mathcal{A}$ be a σ-structure. For every non-empty set $B \subseteq A$, we write $\mathcal{A}[B]$ to denote the restriction of the structure $\mathcal{A}$ to the universe $B \subseteq A$, i.e., the σ-structure with universe B, where $R^{\mathcal{A}[B]} = R^{\mathcal{A}} \cap B^{\operatorname{ar}(R)}$ for each symbol $R \in \sigma$.

For each tuple $\overline{a} = (a_1, \ldots, a_n) \in A^n$, the *$r$-sphere of $\overline{a}$ in $\mathcal{A}$* is defined as the r-type with n centres

$$\mathcal{N}_r^{\mathcal{A}}(\overline{a}) \;=\; \big(\mathcal{A}[N_r^{\mathcal{A}}(\overline{a})], \overline{a}\big)$$

over the signature $\sigma \cup \{c_1, \ldots, c_n\}$. We say that $\overline{a}$ *is of (or, realises the) type τ in $\mathcal{A}$* iff $\mathcal{N}_r^{\mathcal{A}}(\overline{a}) \cong \tau$.

For any d-bounded structure $\mathcal{A}$, any node $a \in A$, and any $r \in \mathbb{N}$, we have $|N_r^{\mathcal{A}}(a)| \leqslant \nu_d(r) := 1 + d \cdot \sum_{0 \leqslant i < r} (d-1)^i$. Observe that for all $r \geqslant 0$ we have $\nu_0(r) = 1$, $\nu_1(r) \leqslant 2$, $\nu_2(r) = 2r+1$, and $(d-1)^r \leqslant \nu_d(r) \leqslant d^{r+1}$ for $d \geqslant 3$, i.e., ν_d grows linearly for $d \leqslant 2$ and exponentially for $d \geqslant 3$.

For every $d, r \geqslant 0$ and $n \geqslant 1$, the universe of every d-bounded r-type τ with n centres contains at most $n \cdot \nu_d(r)$ elements. Thus, given τ and r, one can construct a *sphere-formula* $\operatorname{sph}_\tau(\overline{x})$, i.e., an $\mathrm{FO}[\sigma]$-formula such that for every σ-structure $\mathcal{A}$ and every tuple $\overline{a} \in A^n$ we have

$$\mathcal{A} \models \operatorname{sph}_\tau[\overline{a}] \quad\Longleftrightarrow\quad \mathcal{N}_r^{\mathcal{A}}(\overline{a}) \;\cong\; \tau\,.$$

The formula $\operatorname{sph}_\tau(\overline{x})$ can be constructed in time $\mathcal{O}(\|\sigma\|)$ if $n \cdot \nu_d(r) = 1$, and otherwise in time $(n \cdot \nu_d(r))^{\mathcal{O}(\|\sigma\|)}$.

3.1. Formulas in Hanf normal form for FO($\mathbb{P}$)

In this subsection, we fix a relational signature σ and a *unary* numerical predicate collection, i.e., a numerical predicate collection $(\mathbb{P}, \operatorname{ar}, [\![.]\!])$ with $\operatorname{ar}(\mathsf{P}) = 1$ for all $\mathsf{P} \in \mathbb{P}$. We recall the notion of formulas in Hanf normal form for the logic $\mathrm{FO}(\mathbb{P})$ from [14] (it extends the classical notion for first-order logic FO, see, e.g., [3]).

A *numerical condition on occurrences of types with one centre* (or *numerical oc-type condition*) for $\mathrm{FO}(\mathbb{P})[\sigma]$ is a sentence of the form

$$\mathsf{P}\big(\#(y).\operatorname{sph}_\tau(y) - k\big)\,,$$

where $\mathsf{P} \in \mathbb{P} \cup \{\mathsf{P}_\exists\}$, $k \in \mathbb{N}$, and τ is an r-type with 1 centre, for some $r \in \mathbb{N}$ (in [14], such sentences are called *Hanf-sentences*). We call r the *locality radius* of the numerical oc-type condition. The condition expresses that the number of interpretations for y such that the r-sphere around y is isomorphic to τ belongs to the set $[\![\mathsf{P}]\!] + k$.

A formula $\varphi(\overline{x})$ is in *Hanf normal form for* $\mathrm{FO}(\mathbb{P})[\sigma]$ if it is a Boolean combination of numerical oc-type conditions for $\mathrm{FO}(\mathbb{P})[\sigma]$ and sphere-formulas from $\mathrm{FO}[\sigma]$; in particular, this means that $\varphi \in \mathrm{FO}(\mathbb{P} \cup \{\mathsf{P}_\exists\})[\sigma]$. Accordingly, a *sentence* is in Hanf normal form if it is a Boolean combination of numerical oc-type conditions. We will speak of *hnf-formulas* (for $\mathrm{FO}(\mathbb{P})[\sigma]$) when we mean "formulas in Hanf normal form" (for $\mathrm{FO}(\mathbb{P})[\sigma]$), and similarly for hnf-sentences. The *locality radius* of an hnf-formula is the maximum of the locality radii of its numerical oc-type conditions and its sphere-formulas. The following theorem summarises the main results of [14] and was the starting point of the work to be reported in the present paper.

Theorem 3.1 ([3], [14])

(a) *Let $(\mathbb{P}, \operatorname{ar}, [\![.]\!])$ be a numerical predicate collection with $\operatorname{ar}(\mathsf{P}) = 1$ for all $\mathsf{P} \in \mathbb{P}$. The following are equivalent:*
- *For any relational signature σ, any degree bound $d \in \mathbb{N}$, and any formula $\varphi \in \mathrm{FO}(\mathbb{P})[\sigma]$, there exists a d-equivalent hnf-formula for $\mathrm{FO}(\mathbb{P})[\sigma]$.*
- *For all $\mathsf{P} \in \mathbb{P}$, the set $[\![\mathsf{P}]\!]$ is ultimately periodic.*

(b) *Let $(\mathbb{U}, \operatorname{ar}, [\![.]\!])$ be the numerical predicate collection from Example 2.6*(e). *There is an algorithm which receives as input a degree bound $d \geqslant 2$, a relational signature σ, and a formula $\varphi \in \mathrm{FO}(\mathbb{U})[\sigma]$, and constructs a d-equivalent hnf-formula ψ for $\mathrm{FO}(\mathbb{U})[\sigma]$. The algorithm's running time is in*

$$\exp_3\big(\mathcal{O}(\|\varphi\| + \|\sigma\|) + \log\log(d)\big)\,.$$

(c) *There exists a relational signature σ and a sequence of $\mathrm{FO}[\sigma]$-sentences φ_n of size $\mathcal{O}(n)$ such that every 3-equivalent hnf-sentence $\psi_n \in \mathrm{FO}(\{\mathsf{P}_\exists\})[\sigma]$ has at least $\exp_3(n)$ subformulas.*

The first claim above implies in particular the existence of d-equivalent hnf-formulas for first-order logic (cf. e.g. [3]). For $\mathbb{P} = \{\mathsf{D}_p\}$ (cf. Example 2.6(b)), the existence of d-equivalent hnf-formulas for $\mathrm{FO}(\mathbb{P})$ also follows from Nurmonen's work [23], and the second claim provides an algorithmic version of Nurmonen's theorem. The second claim also implies the main result from [3]. Finally, the third claim was already shown in [3].

3.2. Hanf normal form for FOCN($\mathbb{P}$)

To also allow some kind of "Hanf normal form" for numerical predicates that are not ultimately periodic, we introduce the notion of a formula in "Hanf normal form", where the "numerical oc-type conditions for $\mathrm{FO}(\mathbb{P})$" are replaced by more general "numerical oc-type conditions for $\mathrm{FOCN}(\mathbb{P})$" and, in addition to Boolean combinations, we also allow quantification over number variables (but not over structure variables). Recall that the numerical oc-type condition for $\mathrm{FO}(\mathbb{P})$ is of the form $\mathsf{P}(\#(y).\operatorname{sph}_\tau(y) - k)$ and expresses that the number of realisations of the type τ with a single centre belongs to the set $[\![\mathsf{P}]\!] + k$. In numerical oc-type conditions for $\mathrm{FOCN}(\mathbb{P})$, the "difference between the number of realisations of τ and k" is replaced by an arbitrary multivariate integer polynomial whose variables are the number of realisations of one-centred types $\tau_1, \ldots, \tau_n$ and number variables from nvars. In addition, we give up the restriction to unary numerical predicate collections. The precise definition is as follows.

A *basic counting term* for $\mathrm{FOCN}(\mathbb{P})[\sigma]$ is a counting term t of the form $\#(y).\mathrm{sph}_\tau(y)$, where $y \in \mathsf{vars}$, $r \in \mathbb{N}$ and τ is an r-type with one centre (over σ). The number r is called the *locality radius* of t. In a σ-structure $\mathcal{A}$, the basic counting term t specifies the number $t^{\mathcal{A}}$ of elements $a \in A$ with $\mathcal{N}_r^{\mathcal{A}}(a) \cong \tau$.

A *simple counting term* for $\mathrm{FOCN}(\mathbb{P})[\sigma]$ is a polynomial over basic counting terms, number variables, and integers.

An *atomic numerical oc-type condition* for $\mathrm{FOCN}(\mathbb{P})[\sigma]$ is a formula of the form $\mathsf{P}(t_1, \ldots, t_m)$ with $\mathsf{P} \in \mathbb{P} \cup \{\mathsf{P}_\exists\}$, $m = \mathrm{ar}(\mathsf{P})$, and simple counting terms $t_1, \ldots, t_m$.

A *numerical oc-type condition* for $\mathrm{FOCN}(\mathbb{P})[\sigma]$ is built from atomic numerical oc-type conditions by Boolean combinations and quantification over number variables. The *locality radius* of any of these terms and conditions is the maximal locality radius of the involved basic counting terms.

A formula $\varphi(\overline{x}, \overline{\kappa})$ is in *Hanf normal form for* $\mathrm{FOCN}(\mathbb{P})[\sigma]$ if it is a Boolean combination of numerical oc-type conditions for $\mathrm{FOCN}(\mathbb{P})[\sigma]$ and sphere-formulas from $\mathrm{FO}[\sigma]$; in particular, this means that $\varphi \in \mathrm{FOCN}(\mathbb{P} \cup \{\mathsf{P}_\exists\})[\sigma]$. The locality radius of such a formula is the maximal locality radius of the involved numerical oc-type conditions and sphere-formulas. We abbreviate "formula in Hanf normal form" by *hnf-formula*. Accordingly, an *hnf-sentence* is a sentence in Hanf normal form, i.e., a Boolean combination of numerical oc-type conditions without free number variables.

When speaking of the *number of distinct numerical oc-type conditions* in an hnf-formula ψ we mean the minimal number s of numerical oc-type conditions $\chi_1, \ldots, \chi_s$ such that each χ_i is either an atomic numerical oc-type condition or starts with a number quantifier (i.e., is of the form $\exists\kappa\, \chi_i'$ with $\kappa \in \mathsf{nvars}$), and ψ is a Boolean combination of $\chi_1, \ldots, \chi_s$ and of sphere-formulas from $\mathrm{FO}[\sigma]$.

In analogy to the first two statements in Theorem 3.1, the following is our main result regarding the existence and computability of hnf-formulas for $\mathrm{FOCN}(\mathbb{P})$.

Theorem 3.2 *Let* $(\mathbb{P}, \mathrm{ar}, [\![.]\!])$ *be a numerical predicate collection.*

(a) For any relational signature σ, any degree bound $d \in \mathbb{N}$, and any $\mathrm{FOCN}(\mathbb{P})[\sigma]$-formula φ, there exists a d-equivalent hnf-formula ψ for $\mathrm{FOCN}(\mathbb{P})[\sigma]$ of locality radius less than $(2\,\mathrm{bw}(\varphi) + 1)^{\mathrm{br}(\varphi)}$. Moreover, $\mathrm{free}(\psi) = \mathrm{free}(\varphi)$, $\mathrm{nqr}(\psi) \leqslant \mathrm{nqr}(\varphi)$, and the number of distinct numerical oc-type conditions in ψ is at most

$$\exp_4\big(\mathrm{poly}(\|\varphi\| + \|\sigma\|) + \log\log(d)\big)\,.$$

(b) There is an algorithm which receives as input a degree bound $d \geqslant 2$, a relational signature σ, and an $\mathrm{FOCN}(\mathbb{P})[\sigma]$-formula φ, and constructs such an hnf-formula ψ in time

$$\exp_5\big(\mathrm{poly}(\|\varphi\| + \|\sigma\|) + \log\log(d)\big)\,.$$

The proofs of these two statements can be found in Section 4. Concerning the numerical predicates, our proofs are purely syntactical and do not rely on the particular semantics $[\![\mathsf{P}]\!]$ of the numerical predicates $\mathsf{P} \in \mathbb{P}$.

From statement (a), we infer in Section 5 a polynomial bound for the locality rank of first-order formulas of bounded quantifier alternation depth, as well as a connection between our logic and bounded arithmetic; algorithmic applications of statement (b) for model checking and query evaluation are also discussed in Section 5.

Note that Theorem 3.2(a) implies that if φ is in $\mathrm{FOC}(\mathbb{P})$ (i.e., contains no number variables), then the hnf-formula ψ is in $\mathrm{FOC}(\mathbb{P} \cup \{\mathsf{P}_\exists\})$. For φ in $\mathrm{FO}(\mathbb{P})$, however, Theorem 3.1(a) ensures that ψ is not always in $\mathrm{FO}(\mathbb{P} \cup \{\mathsf{P}_\exists\})$.

Suppose $\mathrm{ar}(\mathsf{P}) = 1$ for all $\mathsf{P} \in \mathbb{P}$ (i.e., $\mathbb{P}$ is *unary*) and let $\varphi \in \mathrm{FO}(\mathbb{P})$. Since $\mathrm{FO}(\mathbb{P}) \subseteq \mathrm{FOC}(\mathbb{P})$, there is a d-equivalent hnf-formula ψ for $\mathrm{FOCN}(\mathbb{P})$, and we even know that $\psi \in \mathrm{FOC}(\mathbb{P} \cup \{\mathsf{P}_\exists\})$. An analysis of the proof of Theorem 3.2(a) yields that all counting terms that appear in ψ have the form i with $i \in \mathbb{N}$ or

$$\Big(\sum_{\tau \in T} \#(y).\mathrm{sph}_\tau(y)\Big) \; - \; k$$

for some set T of types of radius r and some $k \in \mathbb{N}$. Since all predicates from $\mathbb{P}$ are unary, we can eliminate the constant counting terms $i \in \mathbb{N}$ by replacing $\mathsf{P}(i)$ with *true* or *false* depending on whether $i \in [\![\mathsf{P}]\!]$ or not. Clearly,

$$\sum_{\tau \in T} \#(y).\mathrm{sph}_\tau(y) \quad \equiv \quad \#(y).\bigvee_{\tau \in T} \mathrm{sph}_\tau(y)$$

since the types from T all have the same radius and no element can satisfy the sphere-formulas for two different spheres of the same radius. Hence, ψ can be transformed into a Boolean combination of sphere-formulas and of sentences of the form

$$\mathsf{P}\big(\#(y).\bigvee_{\tau \in T} \mathrm{sph}_\tau(y) \; - \; k\big) \tag{1}$$

with $\mathsf{P} \in \mathbb{P} \cup \{\mathsf{P}_\exists\}$. We call such a Boolean combination a *formula in weak Hanf normal form for* $\mathrm{FO}(\mathbb{P})$ or *whnf-formula*, since it weakens the condition on numerical oc-type conditions in hnf-formulas for $\mathrm{FO}(\mathbb{P})$ (that requires $|T| = 1$). As a result, we obtain

Corollary 3.3 *Let* $(\mathbb{P}, \mathrm{ar}, [\![.]\!])$ *be a unary numerical predicate collection. For any relational signature σ, any degree bound $d \in \mathbb{N}$, and any formula $\varphi \in \mathrm{FO}(\mathbb{P})[\sigma]$, there exists a d-equivalent whnf-formula ψ.*

As can be seen from the proof above, the formula ψ can be constructed effectively and the bounds from Theorem 3.2 apply here as well; in particular, the number of subformulas of the form (1) is at most 4-fold exponential in the size of φ.

For this setting, we get a matching lower bound in analogy to Theorem 3.1(c); the proof can be found in Section 6:

Theorem 3.2 (continued)

(c) There exists a unary numerical predicate collection $(\mathbb{P}, \mathrm{ar}, [\![.]\!])$*, a relational signature* σ*, and a sequence of* $\mathrm{FO}(\mathbb{P})[\sigma]$*-sentences* φ_n *of size* $\mathcal{O}(n)$ *such that for every 3-equivalent whnf-sentence* ψ_n*, the number of distinct subformulas of the form (1) in* ψ_n *is at least* $\exp_4(n)$*, for every* $n \geqslant 1$*.*

4. Construction of hnf-formulas for $\mathrm{FOCN}(\mathbb{P})$

For our algorithms it will be convenient to work with a fixed list of representatives of d-bounded r-types, provided by the following lemma (see [2], [14] for a proof).

Lemma 4.1 *There is an algorithm which upon input of a relational signature* σ*, a degree bound* $d \geqslant 2$*, a radius* $r \geqslant 0$*, and a number* $k \geqslant 1$*, computes a list* $\mathcal{L}_r^{\sigma,d}(k) = \tau_1, \ldots, \tau_\ell$ *(for a suitable* $\ell \geqslant 1$*) of* d*-bounded* r*-types with* k *centres over* σ*, such that for every* d*-bounded* r*-type* τ *with* k *centres over* σ *there is exactly one* $i \in [\ell]$ *such that* $\tau \cong \tau_i$*. The algorithm's runtime is* $2^{(k\nu_d(r))^{\mathcal{O}(\|\sigma\|)}}$*. Furthermore, upon input of a* d*-bounded* r*-type* τ *with* k *centres over* σ*, the particular* $i \in [\ell]$ *with* $\tau \cong \tau_i$ *can be computed in time* $2^{(k\nu_d(r))^{\mathcal{O}(\|\sigma\|)}}$*.*

Throughout the remainder of this paper, $\mathcal{L}_r^{\sigma,d}(k)$ will always denote the list provided by Lemma 4.1. We will write $\tau \in \mathcal{L}_r^{\sigma,d}(k)$ to express that τ is one of the types $\tau_1, \ldots, \tau_\ell$ of the list $\mathcal{L}_r^{\sigma,d}(k)$. Our proof of parts (a) and (b) of Theorem 3.2 proceeds by induction on the construction of the input formula. A major technical step for the construction is provided by the following lemma.

Lemma 4.2 *Let* σ *be a relational signature, let* $r \geqslant 0$*,* $n \geqslant 0$*,* $k \geqslant 1$*, let* $(x_1, \ldots, x_n, y_1, \ldots, y_k)$ *be a tuple of* $n{+}k$ *pairwise distinct variables in* vars*, let* $\overline{x} = (x_1, \ldots, x_n)$*, let* $\overline{y} = (y_1, \ldots, y_k)$*, let* $\tau \in \mathcal{L}_r^{\sigma,d}(n{+}k)$*, and let*

$$t(\overline{x}) \quad := \quad \#\overline{y}.\mathrm{sph}_\tau(\overline{x}, \overline{y}).$$

- *If* $n = 0$*, then there is a simple counting term* $\hat{t}$ *without number variables such that* $t^{\mathcal{A}} = \hat{t}^{\mathcal{A}}$*, for any* d*-bounded* σ*-structure* $\mathcal{A}$*.*
- *If* $n \neq 0$*, then for every* $R' \geqslant R := r + k{\cdot}(2r{+}1)$ *and every* $\rho \in \mathcal{L}_{R'}^{\sigma,d}(n)$*, there is a simple counting term* $\hat{t}_\rho$ *without number variables, such that* $t^{\mathcal{A}}[\overline{a}] = \hat{t}_\rho^{\mathcal{A}}$ *holds for any* d*-bounded* σ*-structure* $\mathcal{A}$ *and any tuple* $\overline{a} \in A^n$ *of type* ρ *(i.e., any* $\overline{a} \in A^n$ *with* $\mathcal{A} \models \mathrm{sph}_\rho[\overline{a}]$*).*

Furthermore, $\hat{t}$ *and* $\hat{t}_\rho$ *have locality radius at most* $\hat{R} := r + (k{-}1)(2r{+}1)$*. Moreover, there is an algorithm which constructs* $\hat{t}$ *and* $\hat{t}_\rho$*, resp., in time* $\exp_1\big(((n{+}k){\cdot}\nu_d(\hat{R}))^{\mathcal{O}(\|\sigma\|)}\big)$*.*

Proof idea: The proof relies on a similar analysis of neighbourhood types as the proof of Lemma 4.7 in [18] and proceeds by an induction on the number of *components of* τ *w.r.t.* $\overline{x}$. These components are defined as follows. Let $\tau = (\mathcal{T}, e_1, \ldots, e_n, f_1, \ldots, f_k)$ and let $G = (V, E)$ be the Gaifman graph of τ. Decompose G into its connected components $V_1, \ldots, V_s$. In case that $n = 0$, the tuple $\overline{x}$ is the empty tuple, and the *components of* τ *w.r.t.* $\overline{x}$ are defined as the connected components $V_1, \ldots, V_s$ of G. In case that $n \neq 0$, we can assume w.l.o.g. that there is an $i \in \{1, \ldots, s\}$ such that each of the sets $V_1, \ldots, V_i$ and none of the sets $V_{i+1}, \ldots, V_s$ contains an element of $\{e_1, \ldots, e_n\}$. The *components of* τ *w.r.t.* $\overline{x}$ are defined as the sets $\cup_{j=1}^{i} V_j$, $V_{i+1}, \ldots, V_s$. □

We also use the following lemma from [2].

Lemma 4.3 *Let* σ *be a relational signature. Let* $s \geqslant 0$ *and let* $\chi_1(\overline{\kappa}), \ldots, \chi_s(\overline{\kappa})$ *be arbitrary number formulas from* $\mathrm{FOCN}(\mathbb{P})[\sigma]$.[1] *Let* $r \geqslant 0$*,* $k \geqslant 1$*,* $d \geqslant 2$*, and let* $\mathcal{L}_r^{\sigma,d}(k) = \tau_1, \ldots, \tau_\ell$*. Let* $\overline{x} = x_1, \ldots, x_k$ *be a list of* k *pairwise distinct variables in* vars*. For every Boolean combination* $\psi(\overline{x}, \overline{\kappa})$ *of the formulas* $\chi_1(\overline{\kappa}), \ldots, \chi_s(\overline{\kappa})$ *and of* d*-bounded sphere-formulas of radius at most* r *over* σ*, and for every* $J \subseteq [s]$ *there is a set* $I \subseteq [\ell]$ *such that*

$$\psi_J(\overline{x}) \quad \equiv_d \quad \bigvee_{i \in I} \mathrm{sph}_{\tau_i}(\overline{x}),$$

where $\psi_J(\overline{x})$ *is the formula obtained from* $\psi(\overline{x}, \overline{\kappa})$ *by replacing every occurrence of a formula* $\chi_j(\overline{\kappa})$ *with true if* $j \in J$ *and with false if* $j \notin J$ *(for every* $j \in [s]$*). Given* ψ *and* J*, the set* I *can be computed in time* $\mathrm{poly}(\|\psi\|){\cdot}2^{(k\nu_d(r))^{\mathcal{O}(\|\sigma\|)}}$*.*

By combining the Lemmas 4.3 and 4.2 we obtain:

Lemma 4.4 *Let* σ *be a relational signature. Let* $s \geqslant 0$ *and let* $\chi_1(\overline{\kappa}), \ldots, \chi_s(\overline{\kappa})$ *be arbitrary number formulas from* $\mathrm{FOCN}(\mathbb{P})[\sigma]$*. Let* $r \geqslant 0$*,* $n \geqslant 0$*,* $k \geqslant 1$*, and* $d \geqslant 2$*. Let* $(x_1, \ldots, x_n, y_1, \ldots, y_k)$ *be a tuple of* $n{+}k$ *pairwise distinct variables in* vars*, let* $\overline{x} = (x_1, \ldots, x_n)$*, and let* $\overline{y} = (y_1, \ldots, y_k)$*. Let* $\psi(\overline{x}, \overline{y}, \overline{\kappa})$ *be a Boolean combination of the formulas* $\chi_1(\overline{\kappa}), \ldots, \chi_s(\overline{\kappa})$ *and of* d*-bounded sphere-formulas of radius at most* r *over* σ*, and let*

$$t(\overline{x}, \overline{\kappa}) \quad := \quad \#\overline{y}.\psi(\overline{x}, \overline{y}, \overline{\kappa})\,.$$

Let $\mathcal{L}_r^{\sigma,d}(n{+}k) = \tau_1, \ldots, \tau_\ell$*. For every* $i \in [\ell]$ *let*

$$t_i(\overline{x}) \quad := \quad \#\overline{y}.\mathrm{sph}_{\tau_i}(\overline{x}, \overline{y})\,.$$

For every $J \subseteq [s]$ *there is an* $I \subseteq [\ell]$ *such that the following is true for every* d*-bounded* σ*-structure* $\mathcal{A}$ *and every tuple* $\overline{k} \in \mathbb{Z}^{|\overline{\kappa}|}$ *with*

$$(\mathcal{A}, \overline{k}) \quad \models \quad \chi_J(\overline{\kappa}) \quad := \quad \bigwedge_{j \in J} \chi_j(\overline{\kappa}) \;\wedge\; \bigwedge_{j \in [s] \setminus J} \neg\, \chi_j(\overline{\kappa})\,.$$

- *If* $n = 0$*, then*

$$t^{\mathcal{A}}[\overline{k}] \quad = \quad \sum_{i \in I} \hat{t}_i^{\mathcal{A}}$$

1. The lemma's statement in [2] was formulated for sentences $\chi_1, \ldots, \chi_s$ of first-order logic with modulo-counting quantifiers; the proof, however, is independent of the particular kind of formulas and also applies for number formulas in $\mathrm{FOCN}(\mathbb{P})$ (or any other logic).

where $\hat{t}_i$ is the simple counting term (without number variables) provided by Lemma 4.2 for the term t_i. We let $\hat{t}_J := \sum_{i\in I} \hat{t}_i$.

- *If $n \neq 0$, then for every $R' \geqslant R := r + k\cdot(2r+1)$, every $\rho \in \mathcal{L}^{\sigma,d}_{R'}(n)$ and every tuple $\overline{a} \in A^n$ of type ρ we have*

$$t^{\mathcal{A}}[\overline{a},\overline{k}] \quad = \quad \sum_{i\in I} \hat{t}^{\mathcal{A}}_{i,\rho}\,,$$

where $\hat{t}_{i,\rho}$ is the simple counting term (without number variables) provided by Lemma 4.2 for the term $t_i(\overline{x})$ and the type ρ. We let $\hat{t}_{J,\rho} := \sum_{i\in I} \hat{t}_{i,\rho}$.

The locality radii of $\hat{t}_J$ and $\hat{t}_{J,\rho}$ are at most $\hat{R} := r + (k-1)(2r+1)$. Moreover, there is an algorithm which upon input of $\psi(\overline{x},\overline{y},\overline{\kappa})$ and J (and ρ, in case that $n \neq 0$), constructs $\hat{t}_J$ (resp, $\hat{t}_{J,\rho}$) in time $\operatorname{poly}(\|\psi\|)\cdot 2^{((n+k)\cdot\nu_d(\hat{R}))^{\mathcal{O}(\|\sigma\|)}}$. Furthermore, $|\{\hat{t}_J : J \subseteq [s]\}|$ and $|\{\hat{t}_{J,\rho} : J \subseteq [s]\}|$ is at most 2^{ℓ} for $\ell \in 2^{((n+k)\nu_d(r))^{\mathcal{O}(\|\sigma\|)}}$.

We are now ready to prove Theorem 3.2(a)+(b):

Theorem 4.5 *Let $(\mathbb{P}, \operatorname{ar}, [\![.]\!])$ be a numerical predicate collection. There is an algorithm which upon input of a degree bound $d \geqslant 2$, a relational signature σ, and an $\mathrm{FOCN}(\mathbb{P})[\sigma]$-formula φ, constructs an hnf-formula ψ for $\mathrm{FOCN}(\mathbb{P})[\sigma]$ with $\psi \equiv_d \varphi$.*

Furthermore, $\operatorname{free}(\psi) = \operatorname{free}(\varphi)$, $\operatorname{nqr}(\psi) \leqslant \operatorname{nqr}(\varphi)$, *and the locality radius of ψ is $< (2\operatorname{bw}(\varphi){+}1)^{\operatorname{br}(\varphi)}$. The number of distinct numerical oc-type conditions in ψ is at most*

$$\exp_3\big(\operatorname{poly}(\|\varphi\| + \|\sigma\|)\big) \ \ \textit{for} \ \ d = 2 \quad \textit{and}$$
$$\exp_4\big(\operatorname{poly}(\|\varphi\| + \|\sigma\|) + \log\log(d)\big) \ \ \textit{for} \ \ d \geqslant 3.$$

The construction of ψ takes time at most

$$\exp_4\big(\operatorname{poly}(\|\varphi\| + \|\sigma\|)\big) \ \ \textit{for} \ \ d = 2 \quad \textit{and}$$
$$\exp_5\big(\operatorname{poly}(\|\varphi\| + \|\sigma\|) + \log\log(d)\big) \ \ \textit{for} \ \ d \geqslant 3.$$

Proof: Here, we only prove the theorem's first statement; proofs of the theorem's further statements can be found in the paper's full version. We assume that $\mathsf{P}_\exists \in \mathbb{P}$. Using Example 2.5, we can w.l.o.g. assume that φ does not contain any existential quantification $\exists y$ with $y \in \mathsf{vars}$. We proceed by induction on the shape of φ. Throughout the proof, we let $\overline{x} = (x_1, \ldots, x_n)$ be the free structure variables, and $\overline{\kappa}$ be the free number variables of φ.

Case 1: Suppose that φ is an atomic formula of the form $x_1{=}x_2$ or $R(x_1, \ldots, x_{\operatorname{ar}(R)})$ with $R \in \sigma$. Clearly, φ is equivalent to the formula $\psi := \bigvee_{\tau\in J} \operatorname{sph}_\tau(\overline{x})$ where J is the set of all types $\tau \in \mathcal{L}^{\sigma,d}_0(n)$ that satisfy φ.

Case 2: Suppose that φ is of the form $\neg\varphi'$ or $(\varphi' \vee \varphi'')$. Since the class of hnf-formulas is closed under Boolean combinations, it suffices to apply the induction hypothesis.

Case 3: Suppose that φ is of the form $\mathsf{P}(t_1, \ldots, t_m)$ with $\mathsf{P} \in \mathbb{P} \cup \{\mathsf{P}_\exists\}$, $m = \operatorname{ar}(\mathsf{P})$, and where $t_1, \ldots, t_m$ are counting terms.

According to Definition 2.9, for every $j \in [m]$, the counting term t_j is built by using addition and multiplication based on integers, on number variables from $\overline{\kappa}$, and on counting terms θ' of the form $\#\overline{y}.\theta$. Let Θ' be the set of all these counting terms θ' and let Θ be the set of all the according formulas θ. By the induction hypothesis, for each θ in Θ there is a d-equivalent hnf-formula $\psi^{(\theta)}$. Let Ψ be the set of all these $\psi^{(\theta)}$. Each ψ in Ψ is a Boolean combination of d-bounded sphere-formulas and of numerical oc-type conditions. Let $\chi_1(\overline{\kappa}), \ldots, \chi_s(\overline{\kappa})$ be a list of numerical oc-type conditions such that any of the $\psi \in \Psi$ is a Boolean combination of sphere-formulas and of formulas in $\{\chi_1, \ldots, \chi_s\}$. Let r be the maximum locality radius of any of the sphere-formulas that occur in any $\psi \in \Psi$, and let k be the maximal length $|\overline{y}|$ for $\#\overline{y}.\theta$ occurring in Θ'.

For each $\theta'(\overline{x},\overline{\kappa}) = \#\overline{y}.\theta(\overline{x},\overline{y},\overline{\kappa})$ in Θ', we apply Lemma 4.4 to the term

$$t^{(\theta')}(\overline{x},\overline{\kappa}) \quad := \quad \#\overline{y}.\psi^{(\theta)}(\overline{x},\overline{y},\overline{\kappa})$$

and obtain for every $J \subseteq [s]$

- a simple counting term $\hat{t}^{(\theta')}_J$ without number variables, in case that $n = 0$,
- and for every $\rho \in \mathcal{L}^{\sigma,d}_R(n)$, with $R := r + k(2r{+}1)$, a simple counting term $\hat{t}^{(\theta')}_{J,\rho}$ without number variables, in case that $n \neq 0$.

By Lemma 4.4, the following is true for every $J \subseteq [s]$ and $\chi_J := \bigwedge_{j\in J} \chi_j \wedge \bigwedge_{j\in[s]\setminus J} \neg\chi_j$.

If $n = 0$, then

$$\begin{aligned} &\big(\ \chi_J(\overline{\kappa}) \ \wedge \ \mathsf{P}(t_1(\overline{\kappa}), \ldots, t_m(\overline{\kappa}))\ \big) \\ \equiv_d \ &\big(\ \chi_J(\overline{\kappa}) \ \wedge \ \mathsf{P}(t_{1,J}(\overline{\kappa}), \ldots, t_{m,J}(\overline{\kappa}))\ \big) \end{aligned}$$

where, for every $i \in [m]$, we let $t_{i,J}(\overline{\kappa})$ be the simple counting term obtained from $t_i(\overline{\kappa})$ by replacing each occurrence of a term $\theta' \in \Theta'$ by the ground term $\hat{t}^{(\theta')}_J$.

If $n \neq 0$, then for every $\rho \in \mathcal{L}^{\sigma,d}_R(n)$ we have

$$\begin{aligned} &\big(\ \operatorname{sph}_\rho(\overline{x}) \ \wedge \ \chi_J(\overline{\kappa}) \ \wedge \ \mathsf{P}(t_1(\overline{x},\overline{\kappa}), \ldots, t_m(\overline{x},\overline{\kappa}))\,\big) \\ \equiv_d \ &\big(\ \operatorname{sph}_\rho(\overline{x}) \ \wedge \ \chi_J(\overline{\kappa}) \ \wedge \ \mathsf{P}(t_{1,J,\rho}(\overline{\kappa}), \ldots, t_{m,J,\rho}(\overline{\kappa}))\,\big) \end{aligned}$$

where, for every $i \in [m]$, we let $t_{i,J,\rho}(\overline{\kappa})$ be the simple counting term obtained from $t_i(\overline{x},\overline{\kappa})$ by replacing each occurrence of a term $\theta' \in \Theta'$ by the ground term $\hat{t}^{(\theta')}_{J,\rho}$.

In summary, we obtain the following:

If $n = 0$, then $\ \varphi(\overline{\kappa}) \ = \ \mathsf{P}(t_1, \ldots, t_m)$

$$\begin{aligned} &\equiv_d \bigvee_{J\subseteq[s]} \big(\ \chi_J \ \wedge \ \mathsf{P}(t_1, \ldots, t_m)\ \big) \\ &\equiv_d \bigvee_{J\subseteq[s]} \big(\ \chi_J \ \wedge \ \mathsf{P}(t_{1,J}, \ldots, t_{m,J})\ \big) \quad =: \quad \psi(\overline{\kappa})\,. \end{aligned}$$

The formula χ_J is a Boolean combination of the numerical oc-type conditions $\chi_1, \ldots, \chi_s$. The terms $t_{i,J}$ are polynomials over the simple counting terms $\hat{t}^{(\theta')}_J$ and number variables from $\overline{\kappa}$, i.e., they are simple counting terms. Hence ψ is a Boolean combination of numerical oc-type conditions and therefore an hnf-formula without free structure variables.

If $n \neq 0$, then for $L := \mathcal{L}_R^{\sigma,d}(n)$ we have $\varphi(\overline{x},\overline{\kappa})$

$$
\begin{aligned}
&= \mathsf{P}(t_1,\ldots,t_m)\\
&\equiv_d \bigvee_{\rho\in L}\Big(\mathrm{sph}_\rho(\overline{x}) \wedge \bigvee_{J\subseteq[s]}\big(\chi_J \wedge \mathsf{P}(t_1,\ldots,t_m)\big)\Big)\\
&\equiv_d \bigvee_{\rho\in L}\Big(\mathrm{sph}_\rho(\overline{x}) \wedge \bigvee_{J\subseteq[s]}\big(\chi_J \wedge \mathsf{P}(t_{1,J,\rho},\ldots,t_{m,J,\rho})\big)\Big)\\
&=: \psi(\overline{x},\overline{\kappa})\,.
\end{aligned}
$$

As above, the formula χ_J is a Boolean combination of numerical oc-type conditions. The terms $t_{i,J,\rho}$ are polynomials over the simple counting terms $\hat{t}_{J,\rho}^{(\theta')}$ and number variables from $\overline{\kappa}$, i.e., they are simple counting terms. Hence $\psi(\overline{x},\overline{\kappa})$ is a Boolean combination of sphere-formulas and of numerical oc-type conditions, i.e., an hnf-formula.

Case 4: Suppose that φ is of the form $\exists\lambda\,\varphi'$ with $\lambda \in \mathsf{nvars}$. By the induction hypothesis, $\varphi' \equiv_d \psi'$ for an hnf-formula $\psi'(\overline{x},\overline{\kappa},\lambda)$. Let R be the locality radius of ψ'. From every $\tau = (\mathcal{T},\overline{c}) \in \mathcal{L}_R^{\sigma,d}(n)$, we now construct a numerical oc-type condition $\psi'_\tau(\overline{\kappa},\lambda)$ as follows: Consider a type $\rho = (\mathcal{S},\overline{d})$ such that the sphere-formula $\mathrm{sph}_\rho(\overline{x})$ occurs in ψ', and let r be the locality radius of this sphere-formula. If $\mathcal{N}_r^{\mathcal{T}}(\overline{c}) \cong \rho$, then we replace every occurrence of the sphere-formula $\mathrm{sph}_\rho(\overline{x})$ in ψ' by *true*, otherwise we replace it by *false*. As a result, we get $\varphi = \exists\lambda\,\varphi' \equiv_d \exists\lambda\,\psi'$

$$
\begin{aligned}
&\equiv_d \exists\lambda \bigvee_{\tau\in\mathcal{L}_R^{\sigma,d}(n)} \big(\mathrm{sph}_\tau(\overline{x}) \wedge \psi'\big)\\
&\equiv_d \exists\lambda \bigvee_{\tau\in\mathcal{L}_R^{\sigma,d}(n)} \big(\mathrm{sph}_\tau(\overline{x}) \wedge \psi'_\tau(\overline{\kappa},\lambda)\big)\\
&\equiv \bigvee_{\tau\in\mathcal{L}_R^{\sigma,d}(n)} \big(\mathrm{sph}_\tau(\overline{x}) \wedge \exists\lambda\,\psi'_\tau(\overline{\kappa},\lambda)\big) \quad =: \quad \psi
\end{aligned}
$$

which is an hnf-formula. □

5. Applications

Fixed-parameter model-checking. As a straightforward application of Theorem 3.2(a)+(b), we obtain that Seese's FO model-checking algorithm for classes of structures of bounded degree from [25] can be generalised to the logic FOCN($\mathbb{P}$) for arbitrary numerical predicate collections $(\mathbb{P}, \mathrm{ar}, [\![.]\!])$:

Theorem 5.1 *Let $(\mathbb{P}, \mathrm{ar}, [\![.]\!])$ be a numerical predicate collection. There is an algorithm with oracle $\{(\mathsf{P},\overline{n}) : \mathsf{P} \in \mathbb{P}, \overline{n} \in [\![\mathsf{P}]\!]\}$ which receives as input a formula $\varphi(\overline{x},\overline{\kappa}) \in$ FOCN($\mathbb{P}$), a σ-structure $\mathcal{A}$ (where σ consists of precisely the relation symbols that occur in φ), a tuple $\overline{a} \in A^{|\overline{x}|}$, and a tuple $\overline{k} \in \mathbb{Z}^{|\overline{\kappa}|}$, and decides whether $\mathcal{A} \models \varphi[\overline{a},\overline{k}]$.*

If $d \geqslant 2$ is an upper bound on the degree of $\mathcal{A}$, then the algorithm runs in time

$$f(\varphi,d) \;+\; g(\varphi,d)\cdot|A| \;+\; f(\varphi,d)\cdot|A|^{\mathrm{nqr}(\varphi)}$$

where $f(\varphi,d) \in \exp_5\big(\mathrm{poly}(\|\varphi\|) + \log\log(d)\big)$ *and* $g(\varphi,d) \in \exp_3\big(\mathrm{poly}(\|\varphi\|) + \log\log(d)\big)$.

Since $\mathrm{nqr}(\varphi) = 0$ for all $\varphi \in$ FOC($\mathbb{P}$) this, in particular, implies that on classes of structures of bounded degree, model-checking of FOC($\mathbb{P}$) is fixed-parameter tractable (even fixed-parameter linear) when using oracles for the predicates in $\mathbb{P}$.

Hanf-locality of FOCN($\mathbb{P}$) and the locality rank of FO. The following notion is taken from [15], see also the textbook [21]. Let $\mathcal{A}$ and $\mathcal{B}$ be structures over a relational signature σ, let $k \in \mathbb{N}$ and $\overline{a} \in A^k$ and $\overline{b} \in B^k$. Let furthermore $r \in \mathbb{N}$. Then $(\mathcal{A},\overline{a})$ and $(\mathcal{B},\overline{b})$ are *r-equivalent* (denoted $(\mathcal{A},\overline{a}) \leftrightarrows_r (\mathcal{B},\overline{b})$) if there exists a bijection $f\colon A \to B$ such that, for all $c \in A$, we have

$$\mathcal{N}_r^{\mathcal{A}}(\overline{a},c) \;\cong\; \mathcal{N}_r^{\mathcal{B}}(\overline{b},f(c))\,.$$

Now let $\varphi(\overline{x})$ be an FOCN($\mathbb{P}$)-formula with k free structure variables and without free number variables. The formula $\varphi(\overline{x})$ is *Hanf-local* if there exists $r \geqslant 0$ such that for all structures $\mathcal{A}$ and $\mathcal{B}$ and all $\overline{a} \in A^k$ and $\overline{b} \in B^k$ with $(\mathcal{A},\overline{a}) \leftrightarrows_r (\mathcal{B},\overline{b})$, we have

$$\mathcal{A} \models \varphi[\overline{a}] \quad\Longleftrightarrow\quad \mathcal{B} \models \varphi[\overline{b}]\,.$$

The minimal such r is called the *Hanf-locality rank of* φ and is denoted by $\mathrm{hlr}(\varphi)$.

Let τ be a type with a single centre and let $\mathcal{A}$ be a σ-structure. By $\mathrm{real}_\tau^{\mathcal{A}}$, we denote the number of realisations of the type τ in $\mathcal{A}$. For $r,d \in \mathbb{N}$, the *Hanf-tuple for radius r and degree d* for a structure $\mathcal{A}$ is the tuple

$$\mathsf{HT}_r^d(\mathcal{A}) \quad=\quad \big(\mathrm{real}_\tau^{\mathcal{A}}\big)_{\tau\in\mathcal{L}_r^{\sigma,d}(1)}\,.$$

By Theorem 3.2(a), every formula φ has a d-equivalent hnf-formula ψ of locality radius $r < (2\,\mathrm{bw}(\varphi)+1)^{\mathrm{br}(\varphi)}$. Furthermore, $(\mathcal{A},\overline{a}) \leftrightarrows_r (\mathcal{B},\overline{b})$ implies $\mathsf{HT}_r^d(\mathcal{A}) = \mathsf{HT}_r^d(\mathcal{B})$ and $\mathcal{N}_r^{\mathcal{A}}(\overline{a}) \cong \mathcal{N}_r^{\mathcal{B}}(\overline{b})$. Since the validity of ψ only depends on this information, we get the following:

Corollary 5.2 *Every* FOCN($\mathbb{P}$)*-formula φ without free number variables is Hanf-local with Hanf-locality rank* $\mathrm{hlr}(\varphi) < (2\,\mathrm{bw}(\varphi)+1)^{\mathrm{br}(\varphi)}$.

For first-order formulas φ, we have $\mathrm{hlr}(\varphi) \in \exp_1(\mathcal{O}(\|\varphi\|))$ (actually, $\mathrm{hlr}(\varphi) \leqslant 2^{q-1}-1$ where q is the quantifier depth of φ [20]). Our results allow us to bound the Hanf-locality rank of $\varphi \in$ FO by a polynomial in $\|\varphi\|$ whose degree is the quantifier alternation depth of φ.

As usual, we write Σ_n to denote the set of all FO-formulas of quantifier alternation depth $\leqslant n$ whose outermost quantifier block is existential.

Theorem 5.3 *Every* FO*-formula $\varphi \in \Sigma_n$ is Hanf-local with Hanf-locality rank* $\mathrm{hlr}(\varphi) < (2\|\varphi\|+1)^n$.

Proof: The formula φ is equivalent to a formula of the form $\exists\overline{x_1}\,\neg\exists\overline{x_2}\cdots\neg\exists\overline{x_n}\,\neg\psi$ where ψ is quantifier-free and $\overline{x_1}, \overline{x_2}, \ldots, \overline{x_n}$ are tuples of variables of length $\leqslant \|\varphi\|$.

By induction, set $\psi_n = \psi$ and $\psi_{i-1} = \mathsf{P}_\exists(\#\overline{x_i}.\neg\psi_i)$. Clearly, $\psi_0 \equiv \varphi$ has binding rank n and binding width

$\max\{|\overline{x_i}| : 1 \leqslant i \leqslant n\} \leqslant \|\varphi\|$. Then $\mathrm{hlr}(\varphi) = \mathrm{hlr}(\psi_0) < (2\|\varphi\|+1)^n$ holds by Corollary 5.2. □

Remark 5.4 If, in the definition of Hanf-locality, we allow also *infinite* structures, then the statement "every FOCN($\mathbb{P}$)-formula φ without free number variables is Hanf-local" becomes false — and it becomes false already for pure first-order logic, i.e., for $\varphi \in$ FO: Let $\mathcal{A}$ be a countably infinite complete graph and $\mathcal{B}$ the disjoint union of two copies of $\mathcal{A}$. Then $\mathcal{A} \models \varphi$ and $\mathcal{B} \not\models \varphi$, for $\varphi := \forall x \forall y\, \big(E(x,y) \vee x{=}y\big)$. But $\mathcal{A} \leftrightarrows_r \mathcal{B}$, for every $r \in \mathbb{N}$.

As shown in [3], [7], Corollary 5.2 holds for first-order logic if the definition of the Hanf-locality rank is based on possibly infinite structures of bounded degree d. In [19] it is conjectured that, for every fixed $n \in \mathbb{N}$, this locality rank of formulas $\varphi \in \Sigma_n$ is polynomial in the size of φ. The above theorem confirms this conjecture at least for finite structures.[2]

Hanf-locality and bounded arithmetic. For a numerical predicate collection $(\mathbb{P}, \mathrm{ar}, [\![.]\!])$ consider the extension $\mathcal{Z}_{\mathbb{P}} = (\mathbb{Z}, +, \cdot, 0, \leqslant, ([\![\mathsf{P}]\!])_{\mathsf{P} \in \mathbb{P} \cup \{\mathsf{P}_\exists\}})$ of integer arithmetic with the predicates from $\mathbb{P} \cup \{\mathsf{P}_\exists\}$. A first-order formula $\Phi(\overline{v})$ in the signature of this structure is *bounded* if every quantification $\exists v$ is of the form $\exists v\, (0 \leqslant v \leqslant \sum_{1 \leqslant i \leqslant |\overline{v}|} v_i \wedge \ldots)$, i.e., quantification is restricted to numbers between 0 and the sum of the free variables of Φ.

Let $\varphi \in$ FOCN($\mathbb{P}$) be a sentence, let $r = (2\,\mathrm{bw}(\varphi)+1)^{\mathrm{br}(\varphi)} - 1$, and let $d \in \mathbb{N}$. By Theorem 3.2(a), validity of φ in a d-bounded structure $\mathcal{A}$ only depends on the tuple $\mathsf{HT}^d_r(\mathcal{A}) \in \mathbb{N}^{\mathcal{L}^{\sigma,d}_r(1)}$. Since hnf-sentences are Boolean combinations of numerical oc-type conditions, Theorem 3.2(a) ensures that $\mathcal{A} \models \varphi$ is a first-order property of the tuple $\mathsf{HT}^d_r(\mathcal{A})$ in $\mathcal{Z}_{\mathbb{P}}$. More generally, we obtain the following:

Theorem 5.5 *Let $(\mathbb{P}, \mathrm{ar}, [\![.]\!])$ be a numerical predicate collection. There is an algorithm which receives as input a degree bound d, a relational signature σ, a formula $\varphi(\overline{x}) \in$ FOCN($\mathbb{P}$)$[\sigma]$ with k free structure variables and without free number variables, and a type $\rho \in \mathcal{L}^{\sigma,d}_r(k)$, for $r = (2\,\mathrm{bw}(\varphi)+1)^{\mathrm{br}(\varphi)} - 1$, and constructs a bounded first-order formula Ψ_ρ in the signature of $\mathcal{Z}_{\mathbb{P}}$ and with free variables $(v_\tau)_{\tau \in \mathcal{L}^{\sigma,d}_r(1)}$, such that the following holds for all d-bounded σ-structures $\mathcal{A}$ and all $\overline{a} \in A^k$ with $\rho \cong \mathcal{N}^{\mathcal{A}}_r(\overline{a})$:*

$$\mathcal{A} \models \varphi[\overline{a}] \quad \Longleftrightarrow \quad \mathcal{Z}_{\mathbb{P}} \models \Psi_\rho[\mathsf{HT}^d_r(\mathcal{A})]\,.$$

If the formula φ belongs to FOC($\mathbb{P}$), i.e., contains no number variables, then the formula Ψ_ρ is quantifier-free. With $\mathbb{U}$ the numerical predicate collection from Example 2.6(e), the formula Ψ_ρ can be rewritten into a formula in the signature of $(\mathbb{Z}, +, 0, \leqslant)$ (for $\varphi \in$ FOC($\mathbb{U}$)). Furthermore, using [8] and in particular [3], a similar proof for $\varphi \in$ FO$[\sigma]$ yields a quantifier-free formula Ψ_ρ in the signature of $(\mathbb{Z}, \leqslant)$. Recall that the counting logics FO(Cnt) from [21] and FO+C from [11] are fragments of FOCN($\mathbb{P}$) where $\mathbb{P}$ contains only arithmetical predicates. Consequently, the d-bounded models of any formula from these logics are determined by some set definable in bounded arithmetic.

2. [19, Conjecture 6.2] expected the bound $\|\varphi\| \cdot 2^n$ as opposed to our result $(2\|\varphi\|{+}1)^n$.

Query-evaluation on dynamic databases. In [2], Berkholz, Keppeler, and Schweikardt used the Hanf normal form result of [14] to design efficient algorithms for evaluating queries of first-order logic with modulo-counting quantifiers on dynamic databases. It turns out that the methods of [2] can easily be adapted to generalise to FOC($\mathbb{P}$)-queries, when using the Hanf normal form for FOC($\mathbb{P}$) obtained from Theorem 3.2(a)+(b).

To give a precise statement of the results, we need to provide some notation from [2]. We fix a countably infinite set **dom**, the *domain* of potential database entries. Consider a relational signature $\sigma = \{R_1, \ldots, R_{|\sigma|}\}$. A *$\sigma$-database* D (σ-db, for short) is of the form $D = (R_1^D, \ldots, R_{|\sigma|}^D)$, where each R_i^D is a finite subset of $\mathbf{dom}^{\mathrm{ar}(R_i)}$. The *active domain* adom(D) of D is the smallest subset A of **dom** such that $R_i^D \subseteq A^{ar(R_i)}$ for each R_i in σ. As usual in database theory, we identify a σ-db D with the σ-structure $\mathcal{A}_D$ with universe adom(D) and relations R_i^D for each $R_i \in \sigma$. The *degree* of D is the degree of $\mathcal{A}_D$. The *cardinality* $|D|$ of D is defined as the number of tuples stored in D, i.e., $|D| := \sum_{R \in \sigma} |R^D|$. The *size* $\|D\|$ of D is defined as $\|\sigma\| + |\mathrm{adom}(D)| + \sum_{R \in \sigma} \mathrm{ar}(R) \cdot |R^D|$ and corresponds to the size of a reasonable encoding of D.

For an FOCN($\mathbb{P}$)-formula φ with free$(\varphi) \subseteq$ vars and for any tuple $\overline{x} = (x_1, \ldots, x_k)$ of pairwise distinct structure variables such that free$(\varphi) \subseteq \{x_1, \ldots, x_k\}$, the query result $[\![\varphi(\overline{x})]\!]^D$ of $\varphi(\overline{x})$ on D is defined via

$$[\![\varphi(\overline{x})]\!]^D \quad = \quad \{\overline{a} \in \mathrm{adom}(D)^k : \mathcal{A}_D \models \varphi[\overline{a}]\}\,.$$

If φ is a sentence, then the *answer* $[\![\varphi]\!]^D$ of φ on D is defined as $[\![\varphi]\!]^D = [\![\varphi]\!]^{\mathcal{A}_D} \in \{0,1\}$.

We allow to update a given σ-database by inserting or deleting tuples as follows (note that both types of commands may change the database's active domain and degree). A *deletion* command is of the form $\mathsf{delete}\, R(a_1, \ldots, a_m)$ for $R \in \sigma$, $m = \mathrm{ar}(R)$, and $a_1, \ldots, a_m \in$ **dom**. When applied to a σ-db D, it results in the updated σ-db D' with $R^{D'} = R^D \setminus \{(a_1, \ldots, a_m)\}$ and $S^{D'} = S^D$ for all $S \in \sigma \setminus \{R\}$. An *insertion* command is of the form $\mathsf{insert}\, R(a_1, \ldots, a_m)$ for $R \in \sigma$, $m = \mathrm{ar}(R)$, and $a_1, \ldots, a_m \in$ **dom**. When applied to a σ-db D in the unrestricted setting, it results in the updated σ-db D' with $R^{D'} = R^D \cup \{(a_1, \ldots, a_r)\}$ and $S^{D'} = S^D$ for all $S \in \sigma \setminus \{R\}$. Here, we restrict attention to databases of degree at most d. Therefore, when applying an insertion command to a σ-db D of degree $\leqslant d$, the command is carried out only if the resulting database D' still has degree $\leqslant d$; otherwise D remains unchanged and instead of carrying out the insertion command, an error message is returned.

As in [2], we adopt the framework for dynamic algorithms for query evaluation of [1]. These algorithms are based on Random Access Machines (RAMs) with $\mathcal{O}(\log n)$

word-size and a uniform cost measure (cf., e.g., [4]). We assume that the RAM's memory is initialised to 0. Our algorithms will take as input an FOC($\mathbb{P}$)-formula $\varphi(\overline{x})$ with $\overline{x} = (x_1, \ldots, x_k) \in \mathsf{vars}^k$ and a σ-db D_0 of degree $\leqslant d$. For all query evaluation problems considered here, we aim at routines **preprocess** and **update** which achieve the following.

Upon input of $\varphi(\overline{x})$ and D_0, **preprocess** builds a data structure D which represents D_0 (and which is designed in such a way that it supports the evaluation of $\varphi(\overline{x})$ on D_0). Upon input of a command $\mathsf{update}\ R(a_1, \ldots, a_m)$ (with $\mathsf{update} \in \{\mathsf{insert}, \mathsf{delete}\}$), calling **update** modifies the data structure D such that it represents the updated database D. The *preprocessing time* t_p is the time used for performing **preprocess**; the *update time* t_u is the time used for performing an **update**. By **init** we denote the particular case of the routine **preprocess** upon input of a formula $\varphi(\overline{x})$ and the *empty* database $D_\emptyset$ (where $R^{D_\emptyset} = \emptyset$ for all $R \in \sigma$). The *initialisation time* t_i is the time used for performing **init**. In the dynamic algorithms presented here, the **preprocess** routine for input of $\varphi(\overline{x})$ and D_0 carries out the **init** routine for $\varphi(\overline{x})$ and then performs a sequence of $|D_0|$ update operations to insert all the tuples of D_0 into the data structure. Consequently, $t_p = t_i + |D_0| \cdot t_u$.

Whenever speaking of a *dynamic algorithm* we mean an algorithm that has at least the routines **preprocess** and **update**. In the following, D will always denote the database that is currently represented by the data structure D. To *answer* a *sentence* φ under updates, apart from the routines **preprocess** and **update**, we aim at a routine **answer** that outputs $[\![\varphi]\!]^D$. The *answer time* t_a is the time used for performing **answer**.

The following corollary is obtained by a straightforward adaptation of the proof of Theorem 4.1 of [2], where all uses of the Hanf normal form result for first-order logic with modulo-counting quantifiers of [14] are replaced by uses of Theorem 3.2(a)+(b).

Corollary 5.6 *Let* $(\mathbb{P}, \mathrm{ar}, [\![.]\!])$ *be a numerical predicate collection. There is a dynamic algorithm with oracle* $\{(\mathsf{P}, \overline{n}) : \mathsf{P} \in \mathbb{P}, \overline{n} \in [\![\mathsf{P}]\!]\}$ *which receives as input a relational signature* σ*, a degree bound* $d \geqslant 2$*, an* FOC($\mathbb{P}$)[σ]*-sentence* φ*, and a* σ*-db* D_0 *of degree* $\leqslant d$*, and computes within* $t_p = f(\varphi, d) \cdot \|D_0\|$ *preprocessing time a data structure that can be updated in time* $t_u = f(\varphi, d)$ *and allows to return the query result* $[\![\varphi]\!]^D$ *with answer time* $t_a = \mathcal{O}(1)$.
The function $f(\varphi, d)$ *is of the form* $\exp_5(\mathrm{poly}(\|\varphi\|) + \log\log d)$.

Regarding the evaluation of queries $\varphi(\overline{x})$ where $\overline{x} = (x_1, \ldots, x_k)$ is a tuple of arity $k > 0$, the framework of [2] considers the following problems. To *test* if a given tuple belongs to the query result, we aim at a routine **test** which upon input of a tuple $\overline{a} \in \mathbf{dom}^k$ checks whether $\overline{a} \in [\![\varphi(\overline{x})]\!]^D$. The *testing time* t_t is the time used for performing a **test**. To solve the *counting problem under updates*, we aim at a routine **count** which outputs the cardinality $|[\![\varphi(\overline{x})]\!]^D|$ of the query result. The *counting time* t_c is the time used for performing a **count**. To solve the *enumeration problem under updates*, we aim at a routine **enumerate** such that calling **enumerate** invokes an enumeration of all tuples (without repetition) that belong to the query result $[\![\varphi(\overline{x})]\!]^D$. The *delay* t_d is the maximum time used during a call of **enumerate**

- until the output of the first tuple (or the end-of-enumeration message EOE, if $[\![\varphi(\overline{x})]\!]^D = \emptyset$),
- between the output of two consecutive tuples, and
- between the output of the last tuple and the end-of-enumeration message EOE.

The proof of the following corollary is obtained from the proofs of Theorem 6.1, Theorem 8.1, and Theorem 9.4 of [2] by replacing all uses of Theorem 4.1 of [2] by Corollary 5.6.

Corollary 5.7 *Let* $(\mathbb{P}, \mathrm{ar}, [\![.]\!])$ *be a numerical predicate collection. There is a dynamic algorithm with oracle* $\{(\mathsf{P}, \overline{n}) : \mathsf{P} \in \mathbb{P}, \overline{n} \in [\![\mathsf{P}]\!]\}$ *which receives as input a relational signature* σ*, a degree bound* $d \geqslant 2$*, an* FOC($\mathbb{P}$)[σ]*-formula* $\varphi(\overline{x})$ *with free variables* $\overline{x} = (x_1, \ldots, x_k) \in \mathsf{vars}^k$ *(for some* $k \in \mathbb{N}$*), and a* σ*-db* D_0 *of degree* $\leqslant d$*, and computes within* $t_p = f(\varphi, d) \cdot \|D_0\|$ *preprocessing time a data structure that can be updated in time* $t_u = f(\varphi, d)$ *and allows to*

- *test for any input tuple* $\overline{a} \in \mathbf{dom}^k$ *whether* $\overline{a} \in [\![\varphi(\overline{x})]\!]^D$ *within testing time* $t_t = \mathcal{O}(k^2)$
- *return the cardinality* $|[\![\varphi(\overline{x})]\!]^D|$ *of the query result within time* $\mathcal{O}(1)$
- *enumerate* $[\![\varphi(\overline{x})]\!]^D$ *with* $\mathcal{O}(k^3)$ *delay.*

The function $f(\varphi, d)$ *is of the form* $\exp_5(\mathrm{poly}(\|\varphi\|) + \log\log d)$.

6. The lower bound

By Corollary 3.3, any FO($\mathbb{P}$)-formula φ has a d-equivalent formula ψ in weak Hanf normal form. As the bounds from Theorem 3.2(a) apply here as well, it follows that the number of sentences $\mathsf{P}(\#(y).\bigvee_{\tau \in T} \mathrm{sph}_\tau(y) - k)$ in ψ is at most 4-fold exponential in the size of φ. In this section, we show a matching lower bound for certain numerical predicate collections. For two tuples $\overline{a}, \overline{b} \in \mathbb{N}^s$, we write $\overline{a}^\mathsf{T}\, \overline{b}$ for the usual inner product $\sum_{1 \leqslant i \leqslant s} a_i \cdot b_i$.

Definition 6.1 A set $R \subseteq \mathbb{N}$ of natural numbers is *rich* if, for all $s, u, v \in \mathbb{N}$, all $\overline{a_0} \in \{0,1\}^s \setminus \{\overline{0}\}$, $\overline{a_1}, \ldots, \overline{a_u} \in \mathbb{N}^s$, and all $c_1, \ldots, c_u \in \mathbb{N}$ with $(\overline{a_0}, 0) \neq (\overline{a_i}, c_i)$ for all $i \in [u]$, there exist $\overline{x}, \overline{y} \in (v + \mathbb{N})^s$ such that

- $\overline{a_0}^\mathsf{T}\, \overline{x} \in R \iff \overline{a_0}^\mathsf{T}\, \overline{y} \notin R$ and
- $\overline{a_i}^\mathsf{T}\, \overline{x} - c_i \in R \iff \overline{a_i}^\mathsf{T}\, \overline{y} - c_i \in R$ for all $i \in [u]$.

Example 6.2 In the paper's full version we show the following. We can produce probabilistically a set R of natural numbers as follows: for every $n \in \mathbb{N}$, toss a fair coin and place n into R iff the outcome is tail. Then, with probability 1, we get a rich set. Furthermore, a set $R \subseteq \mathbb{N}$ is rich whenever it has "large gaps", i.e., R is infinite, $0 \notin R$, and for all $d \in \mathbb{N}_{\geqslant 1}$ there exists $q \in R$ such that $[\lfloor q/d \rfloor, d \cdot q] \cap R = \{q\}$. Examples of such sets are $\{n^n : n \in \mathbb{N}\}$, $\{\lfloor 2^{n^c} \rfloor : n \in \mathbb{N}\}$ for all reals $c > 1$, $\{n! : n \in \mathbb{N}\}$, as well as all infinite subsets of these sets. But note that neither $\{2^n : n \in \mathbb{N}\}$ nor (by Bertrand's postulate) the set of all primes has large gaps.

For $n \in \mathbb{N}$ let $\mathfrak{T}_n$ denote the set of all *(complete labeled ordered binary) trees of height* 2^n, i.e., of all structures $\mathcal{A} = (A, E_0^{\mathcal{A}}, E_1^{\mathcal{A}}, X^{\mathcal{A}})$ with A the set of binary words of length at most 2^n, $E_b^{\mathcal{A}} = \{(u, ub) : ub \in A\}$ for $b \in \{0, 1\}$, and $X^{\mathcal{A}} \subseteq A$. We denote by σ_{tree} the signature of these trees.

A tree $\mathcal{A}$ is *marked* if the root ε belongs to $X^{\mathcal{A}}$, i.e., is labeled. Otherwise, $\mathcal{A}$ is *unmarked*. For a tree $\mathcal{A} = (A, E_0^{\mathcal{A}}, E_1^{\mathcal{A}}, X^{\mathcal{A}})$, let $\mu(\mathcal{A}) = (A, E_0^{\mathcal{A}}, E_1^{\mathcal{A}}, X^{\mathcal{A}} \cup \{\varepsilon\})$ denote the marked tree that is obtained by labelling the root; the unmarked tree $\bar{\mu}(\mathcal{A}) = (A, E_0^{\mathcal{A}}, E_1^{\mathcal{A}}, X^{\mathcal{A}} \setminus \{\varepsilon\})$ is defined analogously.

For a finite set S of trees, we let $\mu(S) := \{\mu(\mathcal{A}) : \mathcal{A} \in S\}$. A *forest over* S (or, an S*-forest*) is a disjoint union of finitely many copies of trees from S. Since every tree is finite, the same applies to every S-forest. For a $\mathfrak{T}_n$-forest $\mathcal{F}$ and a set $R \subseteq \mathbb{N}$, let $P_R(\mathcal{F})$ be the following property:

> The number of unmarked trees $\mathcal{A}$ in $\mathcal{F}$ such that $\mu(\mathcal{A})$ also appears in $\mathcal{F}$ belongs to R.

The following lemma is the technical core of the proof of Theorem 6.4, which implies Theorem 3.2(c).

Lemma 6.3 *Let $\mathbb{P} = \{\mathsf{R}\}$ with $\mathrm{ar}(\mathsf{R}) = 1$ and $[\![\mathsf{R}]\!] \subseteq \mathbb{N}$ rich. Let $n \in \mathbb{N}$ and let $\psi \in \mathrm{FO}(\mathbb{P} \cup \{\mathsf{P}_\exists\})[\sigma_{\mathrm{tree}}]$ be a whnf-sentence such that for all $\mathfrak{T}_n$-forests $\mathcal{F}$ we have*

$$\mathcal{F} \models \psi \quad \iff \quad P_{[\![\mathsf{R}]\!]}(\mathcal{F}) \,.$$

Then, the number of distinct subformulas of the form $\mathsf{R}(\#(y).\bigvee_{\tau \in T} \mathrm{sph}_\tau(y) - k)$ in ψ is at least as big as the number of non-empty sets $B \subseteq \mu(\mathfrak{T}_n)$ of marked trees of height 2^n.

Theorem 6.4 *Let $\mathbb{P} = \{\mathsf{R}\}$ with $\mathrm{ar}(\mathsf{R}) = 1$ and $[\![\mathsf{R}]\!] \subseteq \mathbb{N}$ rich. There is a sequence $(\varphi_n)_{n \geqslant 1}$ of $\mathrm{FO}(\mathbb{P})[\sigma_{\mathrm{tree}}]$-sentences of size $\mathcal{O}(n)$ such that, for all $n \geqslant 1$, every whnf-sentence ψ_n from $\mathrm{FO}(\mathbb{P} \cup \{\mathsf{P}_\exists\})[\sigma_{\mathrm{tree}}]$ that is 3-equivalent to φ_n contains at least $\exp_4(n)$ distinct subformulas of the form $\mathsf{R}(\#(y).\bigvee_{\tau \in T} \mathrm{sph}_\tau(y) - k)$.*

Proof: A construction by Frick & Grohe [10, Lemma 25] provides us with a sequence of formulas $\varphi_n \in \mathrm{FO}(\mathbb{P})[\sigma_{\mathrm{tree}}]$ of size $\mathcal{O}(n)$ such that, for all $n \in \mathbb{N}$ and all $\mathfrak{T}_n$-forests $\mathcal{F}$, we have $P_{[\![\mathsf{R}]\!]}(\mathcal{F}) \iff \mathcal{F} \models \varphi_n$.

Let $n \geqslant 1$ and let $\psi \in \mathrm{FO}(\mathbb{P} \cup \{\mathsf{P}_\exists\})[\sigma_{\mathrm{tree}}]$ be a whnf-sentence that is 3-equivalent to φ_n. Then, by Lemma 6.3, ψ contains at least $2^{\frac{1}{2}|\mathfrak{T}_n|} - 1$ distinct subformulas of the form $\mathsf{R}(\#(y).\bigvee_{\tau \in T} \mathrm{sph}_\tau(y) - k)$. Since $|\mathfrak{T}_n| = 2^{(2^{2^n+1}-1)}$, we obtain that $2^{\frac{1}{2}|\mathfrak{T}_n|} - 1 \geqslant \exp_4(n)$, and we are done. $\square$

References

[1] C. Berkholz, J. Keppeler, and N. Schweikardt. Answering conjunctive queries under updates. In *Proc. PODS'17*, 2017. Preprint available at https://arxiv.org/abs/1702.06370.

[2] C. Berkholz, J. Keppeler, and N. Schweikardt. Answering FO+MOD queries under updates on bounded degree databases. In *Proc. ICDT'17*, pages 8:1–8:18, 2017. Preprint available at https://arxiv.org/abs/1702.08764.

[3] B. Bollig and D. Kuske. An optimal construction of Hanf sentences. *Journal of Applied Logic*, 10:179–186, 2012.

[4] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms (3. ed.)*. MIT Press, 2009.

[5] A. Dawar, M. Grohe, S. Kreutzer, and N. Schweikardt. Model Theory Makes Formulas Large. In *Proc. ICALP'07*, pages 1076–1088, 2007.

[6] 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.

[7] H.-D. Ebbinghaus and J. Flum. *Finite Model Theory*. Springer, 1995.

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

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

[10] M. Frick and M. Grohe. The complexity of first-order and monadic second-order logic revisited. *APAL*, 130(1-3):3–31, 2004.

[11] Martin Grohe. *Descriptive Complexity, Canonisation, and Definable Graph Structure Theory*. Cambridge Univ. Press, 2017. To appear. Preprint available at https://www.lii.rwth-aachen.de/de/13-mitarbeiter/professoren/39-book-descriptive-complexity.html.

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

[13] K. Härtig. Über einen Quantifikator mit zwei Wirkungsbereichen. Colloquium on the foundations of mathematics, mathematical machines and their applications, Tihany 1962, pp. 31–36, 1965.

[14] L. Heimberg, D. Kuske, and N. Schweikardt. Hanf normal form for first-order logic with unary counting quantifiers. In *Proc. LICS'16*, pages 277–286, 2016.

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

[16] W. Kazana and L. Segoufin. First-order query evaluation on structures of bounded degree. *LMCS*, 7(2), 2011.

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

[18] S. Kreutzer and N. Schweikardt. On Hanf-equivalence and the number of embeddings of small induced subgraphs. In *Proc. CSL-LICS'14*, pages 60:1–60:10, 2014.

[19] D. Kuske and M. Lohrey. Automatic structures of bounded degree revisited. *Journal of Symbolic Logic*, 76(4):1352–1380, 2011.

[20] L. Libkin. Logics with counting and local properties. *ACM Transactions on Computational Logic*, 1:33–59, 2000.

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

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

[23] J. Nurmonen. Counting modulo quantifiers on finite structures. *Information and Computation*, 160:62–87, 2000.

[24] N. Rescher. Plurality quantification. *JSL*, 27:373–374, 1962.

[25] D. Seese. Linear time computable problems and first-order descriptions. *Math. Struct. in Comput. Sci.*, 6:506–526, 1996.

[26] L. Segoufin. A glimpse on constant delay enumeration (invited talk). In *Proc. STACS'14*, pages 13–27, 2014.

[27] J. A. Väänänen. Unary quantifiers on finite models. *Journal of Logic, Language and Information*, 6(3):275–304, 1997.