Some lower bounds in parameterized AC^0

Yijia Chen
School of Computer Science
Fudan University
yijiachen@fudan.edu.cn

Jörg Flum
Mathematisches Institut
Universität Freiburg
joerg.flum@math.uni-freiburg.de

Abstract

We demonstrate some lower bounds for parameterized problems via parameterized classes corresponding to the classical AC^0. Among others, we derive such a lower bound for all fpt-approximations of the parameterized clique problem and for a parameterized halting problem, which recently turned out to link problems of computational complexity, descriptive complexity, and proof theory. To show the first lower bound, we prove a strong AC^0 version of the planted clique conjecture: AC^0-circuits asymptotically almost surely can not distinguish between a random graph and this graph with a randomly planted clique of any size $\leq n^\xi$ (where $0 \leq \xi < 1$).

1. Introduction

For $k \in \mathbb{N}$ the k-clique problem asks, given a graph G, whether it contains a clique of size k. In [21], Rossman showed that the k-clique problem has no bounded-depth and unbounded-fan-in circuits of size $O(n^{k/4})$, where n is the number of vertices in an input graph. Therefore, there doesn't exist a family $\left(\mathsf{C}_{\binom{n}{2},k}\right)_{n,k\in\mathbb{N}}$ of circuits such that for some functions $d, f : \mathbb{N} \to \mathbb{N}$,

- every $\mathsf{C}_{\binom{n}{2},k}$ has depth at most $d(k)$ and size bounded by $f(k) \cdot n^{k/4}$,
- an n-vertex graph G has a k-clique if and only if $\mathsf{C}_{\binom{n}{2},k}(G) = 1$. Here $\mathsf{C}_{\binom{n}{2}}$ has an input node for every potential edge.

If the constraint on the depth of the circuits could be removed, then we would immediately obtain that the *parameterized clique problem*

p-CLIQUE	
Instance:	A graph G and $k \in \mathbb{N}$.
Parameter:	k.
Question:	Does G contain a clique of size k?

cannot be solved in time $f(k) \cdot n^{O(1)}$. Thus, p-CLIQUE would not be fixed-parameter tractable (FPT) and hence, FPT $\neq$ W[1], since p-CLIQUE is in the parameterized class W[1]. Therefore, Rossman's result may be viewed as an AC^0 version of FPT $\neq$ W[1], an inequality conjectured by most experts of the field (recall that the complexity class AC^0 contains all problems that can be computed by bounded-depth and unbounded fan-in circuits of polynomial size).

In [11] Elberfeld et al. introduced the parameterized class para-AC^0 as the AC^0 analog of the class FPT: A problem is in para-AC^0 if it can be computed by *dlogtime-uniform* AC^0-circuits after

an (arbitrarily complex) *precomputation* [12] on the parameter. Later in [3] it was shown that para-AC^0 contains the *parameterized vertex cover problem* (p-VERTEX-COVER), one of the archetypal fixed-parameter tractable problems. For various other problems the authors of [3] also proved their membership in para-AC^0. Concerning nonmembership, a result in [6] shows that the parameterized st-connectivity problem (p-STCONN), i.e., the problem of deciding whether there is a path of length at most k between vertices s and t in a graph G, parameterized by k, is not in para-AC^0. It is worth noting that st-connectivity is solvable in polynomial time, and hence p-STCONN $\in$ FPT.

The class AC^0 is one of the best understood classical complexity classes. Already in [1, 14] it was shown that PARITY, the problem of deciding whether a binary string contains an even number of 1's, is not in AC^0. Since PARITY has a very low complexity, for many other problems, including VERTEX-COVER and CLIQUE, the AC^0-lower bound can be easily derived by reductions from PARITY. Similarly, as p-CLIQUE $\notin$ para-AC^0, it is not very hard to see, using some appropriate weak parameterized reductions, that many other parameterized problems, including the dominating set problem, are not in para-AC^0.

It is well known that the class AC^0 is intimately connected to first-order logic (FO). In fact, the problems decidable by dlogtime-uniform AC^0-circuits are precisely those definable in $\mathrm{FO}(<,+,\times)$, that is, in first-order logic for ordered structures with built-in predicates of addition and multiplication.

Now we can also study various parameterized classes based on fragments of $\mathrm{FO}(<,+,\times)$. Let us emphasize that this is not merely an academic exercise. Logic and parameterized complexity are surprisingly intertwined with each other, which, among others, is witnessed by various algorithmic meta-theorems (see e.g. [16]). Moreover, the problem whether there is a logic for PTIME, a central problem of descriptive complexity, turned out (see [9] for a thorough discussion) to be related to the complexity of the parameterized halting problem

p-HALT	
Instance:	$n \in \mathbb{N}$ in *unary* and a nondeterministic Turing machine (NTM) $\mathbb{M}$.
Parameter:	$\lvert\mathbb{M}\rvert$, the size of he machine $\mathbb{M}$.
Question:	Does $\mathbb{M}$ accept the empty input tape in at most n steps?

In fact, already in [20] it was shown that PTIME has a logic if p-HALT has an algorithm with running time $n^{f(|\mathbb{M}|)}$ for some function f. We get a family $(\mathsf{C}_{n,k})_{n,k\in\mathbb{N}}$ of circuits such that

- every $\mathsf{C}_{n,k}$ has depth 2 and size $g(k)\cdot n$ for some function $g:\mathbb{N}\to\mathbb{N}$,
- an NTM $\mathbb{M}$ accepts the empty input tape in at most n steps if and only if $\mathsf{C}_{n,|\mathbb{M}|}(n,\mathbb{M})=1$

by hard-wiring into $\mathsf{C}_{n,k}$ the NTMs of size k which halt on empty input in $\leq n$ steps.

Therefore, p-HALT is in a *nonuniform* version of para-AC^0. So the question arises whether p-HALT $\in$ para-AC^0. Note that a positive answer will yield that p-HALT $\in$ FPT, which is considered to be highly unlikely [9]. Hence, the goal is to show *unconditionally* that p-HALT $\notin$ para-AC^0. To the best of our knowledge, all existing AC^0 lower bounds apply to both uniform and nonuniform circuits. Perhaps, in order to settle the complexity of p-HALT with respect to para-AC^0, a better understanding of the uniformity conditions of circuits is really required.

Our work. In this paper, we systematically investigate lower bounds in terms of para-AC^0. We show that a number of problems are not in this class or in some of its proper subclasses. To some extent, our results appear rather separated and our proofs are often built on known results and techniques. Nevertheless, as *unconditional lower bounds* are still rare in parameterized complexity, para-AC^0 is in our opinion the best starting point for this line of research.

Following the framework proposed in [12], we first compare two possible definitions of para-AC^0 depending on different ways to obtain parameterized classes from classical ones. We have already mentioned the first one, in which an arbitrary precomputation can be performed on the parameter before a standard computation according to the corresponding classical class. The second approach requires the parameterized problem to be in the classical class if we restrict to instances were the parameter is far smaller than the size of the input. We show that both views lead to the same para-AC^0.

Then we derive a first set of lower bound results: We show that many natural $\mathrm{W}[1]$-hard problems are not in para-AC^0 by arguing that the corresponding reductions from p-CLIQUE can be made in AC^0. Among others, they include the weighted satisfiability problems for classes of propositional formulas, which define the W-hierarchy.

We present a modeltheoretic tool, based on the color-coding method, which allows to show membership in AC^0 (similarly as done in [3] via circuits).

We generalize Rossman's result mentioned at the beginning of this introduction and show that any fpt-approximation of p-CLIQUE is not in para-AC^0. To get this result we prove that AC^0-circuits asymptotically almost surely can not distinguish between a random graph and this graph with a randomly planted clique of any size $\leq n^\xi$ with $0 \leq \xi < 1$. Our first proof of the last two results used the sophisticated machinery in [21]. Here we outline a proof, suggested to us anonymously, which is directly built on Beame's *Clique Switching Lemma* [5]. The fpt-approximation lower bound of p-CLIQUE again can be transferred to the weighted satisfiability problems, provided the propositional formulas are of odd depth.

Finally we turn to p-HALT. We are not able to show p-HALT $\notin$ para-AC^0, however, using the decidability of Presburger's arithmetic we prove that p-HALT is not in para-$\mathrm{FO}(<,+)$, not even in $\mathrm{XFO}(<,+)$. On the other hand, p-HALT $\in$ nonuniform-para-$\mathrm{FO}(<,+)$.

2. Preliminaries

By $\mathbb{N}$ we denote the set of nonnegative integers. For every $n \in \mathbb{N}$ we let $[n] := \{1, \dots, n\}$. Moreover, let $\mathbb{R}$ be the set of real numbers, $\mathbb{R}_+ := \big\{r \in \mathbb{R} \bigm| r > 0\big\}$, and $\mathbb{R}_{\geq 1} := \big\{r \in \mathbb{R} \bigm| r \geq 1\big\}$. For any set A and $k \in \mathbb{N}$ we define $\binom{A}{k}$ as the class of k-element subsets of A, i.e., $\big\{S \subseteq A \bigm| |S| = k\big\}$.

A (simple) graph $G = (V(G), E(G))$ $\big($for short, $G = (V, E)\big)$ is undirected and has no loops and multiple edges. Here, $V(G)$ is the vertex set and $E(G)$ the edge set, respectively. A subset $C \subseteq V(G)$ is a *clique* of G if for every $u, v \in C$ either $u = v$ or $\{u, v\} \in E(G)$. And $D \subseteq V(G)$ is a *dominating set* of G if for every $v \in V(G)$ either $v \in D$ or there exists $u \in D$ with $\{u, v\} \in E(G)$.

Relational structures and first-order logic. A *vocabulary* τ is a finite set of relation symbols. Each relation symbol has an *arity*. A *structure* $\mathcal{A}$ of vocabulary τ, or simply structure, consists of a finite set A called the *universe*, and an interpretation $R^{\mathcal{A}} \subseteq A^r$ of each r-ary relation symbol $R \in \tau$. For example, a graph G can be identified with a structure $\mathcal{A}(G)$ of vocabulary $\{E\}$ with binary relation symbol E such that $A(G) := V(G)$ and $E^{\mathcal{A}(G)} := \{(u, v) \mid \{u, v\} \in E(G)\}$.

Formulas of first-order logic of vocabulary τ are built up from atomic formulas $x = y$ and $Rx_1 \dots x_r$, where $x, y, x_1, \dots, x_r$ are variables and $R \in \tau$ is of arity r, using the boolean connectives and existential and universal quantification. For example, for every $k \geq 1$ let

$$\mathit{clique}_k := \exists x_1 \dots \exists x_k \Big(\bigwedge_{1 \leq i < j \leq k} (\neg x_i = x_j \wedge E x_i x_j) \Big).$$

Then a graph G has a k-clique if and only if $\mathcal{A}(G) \models \mathit{clique}_k$.

Parameterized complexity. We fix an alphabet $\Sigma := \{0,1\}$. A *parameterized problem* (Q,κ) consists of a classical problem $Q \subseteq \Sigma^*$ and a function $\kappa : \Sigma^* \to \mathbb{N}$, the *parameterization*, computable in polynomial time. As an example, we have already seen p-CLIQUE in the Introduction. A similar problem is the *parameterized dominating set problem*.

p-DOMINATING-SET	
Instance:	A graph G and $k \in \mathbb{N}$.
Parameter:	k.
Question:	Does G contain a dominating set of size k?

Both, p-CLIQUE and p-DOMINATING-SET, play an important role in parameterized complexity, mainly because they are complete for the classes $\mathrm{W}[1]$ and $\mathrm{W}[2]$, respectively. Recall that the classes of the W-hierarchy are defined by taking the closure under fpt-reductions of the following weighted satisfiability problem for suitable classes Γ of propositional formulas or circuits.

p-WSAT(Γ)	
Instance:	$\gamma \in \Gamma$ and $k \in \mathbb{N}$.
Parameter:	k.
Question:	Does γ have a satisfying assignment of Hamming weight k?

Definition 2.1. Let (Q,κ) and (Q',κ') be two parameterized problems. An *fpt-reduction* from (Q,κ) to (Q',κ') is a mapping $R : \Sigma^* \to \Sigma^*$ such that:

- For all $x \in \Sigma^*$ we have $\big(x \in Q \iff R(x) \in Q'\big)$.
- For all $x \in \Sigma^*$, the image $R(x)$ is computable in time
$$f(\kappa(x)) \cdot |x|^{O(1)}$$
for a computable $f : \mathbb{N} \to \mathbb{N}$.
- There is a computable function $g : \mathbb{N} \to \mathbb{N}$ such that $\kappa'(R(x)) \leq g(\kappa(x))$ for all $x \in \Sigma^*$.

If there is an fpt-reduction from (Q,κ) to (Q',κ'), then we write $(Q,\kappa) \leq^{\text{fpt}} (Q',\kappa')$.

For $t \geq 0$ and $d \geq 1$ we inductively define the following classes $\Gamma_{t,d}$ and $\Delta_{t,d}$ of propositional formulas:

$$\begin{aligned}
\Gamma_{0,d} &:= \{\lambda_1 \wedge \ldots \wedge \lambda_c \mid c \leq d, \lambda_1, \ldots, \lambda_c \text{ literals}\},\\
\Delta_{0,d} &:= \{\lambda_1 \vee \ldots \vee \lambda_c \mid c \leq d, \lambda_1, \ldots, \lambda_c \text{ literals}\},\\
\Gamma_{t+1,d} &:= \Big\{ \bigwedge_{i \in I} \delta_i \mid I \text{ finite}, \delta_i \in \Delta_{t,d} \text{ for all } i \in I \Big\},\\
\Delta_{t+1,d} &:= \Big\{ \bigvee_{i \in I} \gamma_i \mid I \text{ finite}, \gamma_i \in \Gamma_{t,d} \text{ for all } i \in I \Big\}.
\end{aligned}$$

Now we are ready to define the classes of the *W-hierarchy*.

Definition 2.2. Let $t \geq 1$. Then
$$\mathrm{W}[t] := \bigcup_{d \geq 1} \big\{(Q,\kappa) \mid (Q,\kappa) \leq^{\text{fpt}} p\text{-WSAT}(\Gamma_{t,d})\big\}.$$

Circuit Complexity. A circuit C with n input gates is a directed acyclic graph in which every node (i.e., gate) is labelled by $\bigwedge$, $\bigvee$, $\neg$, or by one of the variables, or by 0 or 1. All $\bigwedge$ and $\bigvee$ gates may have arbitrarily many inputs, i.e., C is of *unbounded fan-in*. The *depth* of C is the length of a longest directed path in C. The *size* of C, denoted by $|\mathsf{C}|$, is the number of gates in C. We often tacitly identify C with the function $\mathsf{C} : \{0,1\}^n \to \{0,1\}^m$ it computes. Here, n is the number of variables of C and m the number of its *output gates*.

AC^0 is the class of problems that can be computed by circuits of bounded-depth and polynomial size. More precisely:

Definition 2.3. Let $Q \subseteq \Sigma^*$. We say that $Q \in \mathrm{AC}^0$ if there exists a family of boolean circuits $(\mathsf{C}_n)_{n\in\mathbb{N}}$ such that:

(A1) The depth of every C_n is bounded by a fixed constant.

(A2) $|\mathsf{C}_n| = n^{O(1)}$.

(A3) Let $x \in \Sigma^*$. Then ($x \in Q$ if and only if $C_{|x|}(x) = 1$). In particular, C_n has n input gates.

(A4) $(\mathsf{C}_n)_{n\in\mathbb{N}}$ is *dlogtime-uniform*, that is: there is a *deterministic logtime Turing machine* $\mathbb{M}$ which on input 1^n outputs the circuit C_n. More precisely, $\mathbb{M}$ recognizes the language

$$\big\{(b, i, 1^n) \;\big|\; \text{the } i\text{th bit of the binary encoding of } C_n \text{ is } b\big\}$$

$\big($cf. Section 6 of [4]$\big)$.

Often, $(\mathsf{C}_n)_{n\in\mathbb{N}}$ are called AC^0*-circuits*.

We remark that most lower bounds in our paper still hold without the requirement (A4). Therefore, (A4) is irrelevant for most of our results. However, with this uniformity condition, AC^0 characterizes precisely the class of problems that are definable in $\mathrm{FO}(<, +, \times)$ [4].

3. para-AC^0 and Some Natural Examples

Definition 3.1 ([3])**.** Let (Q, κ) be a parameterized problem. Then (Q, κ) is in para-AC^0 if there exists a family $\big(\mathsf{C}_{n,k}\big)_{n,k\in\mathbb{N}}$ circuits such that:

(P1) The depth of every $\mathsf{C}_{n,k}$ is bounded by a fixed constant.

(P2) $|\mathsf{C}_{n,k}| \leq f(k) \cdot n^{O(1)}$ for every $n, k \in \mathbb{N}$, where $f : \mathbb{N} \to \mathbb{N}$ is a computable function.

(P3) Let $x \in \Sigma^*$. Then $\big(x \in Q$ if and only if $\mathsf{C}_{|x|,\kappa(x)}(x) = 1\big)$.

(P4) There is a deterministic Turing machine that on input $(1^n, 1^k)$ computes the circuit $\mathsf{C}_{n,k}$ in time $g(k) + O(\log n)$, where $g : \mathbb{N} \to \mathbb{N}$ is a computable function.

For future reference, we restate a para-AC^0 version of Rossman's main result [21] as follows.

Theorem 3.2. *Let $k \in \mathbb{N}$. Then there is no family $\Big(\mathsf{C}_{\binom{n}{2}}\Big)_{n\in\mathbb{N}}$ of circuits such that the following conditions are all satisfied.*

- *The depth of every $\mathsf{C}_{\binom{n}{2}}$ is bounded by a fixed constant $d \in \mathbb{N}$.*
- *The size of $\mathsf{C}_{\binom{n}{2}}$ is $O(n^{k/4})$.*

– *Let G be a graph and $n := |V(G)|$. Then G has a k-clique if and only if $\mathsf{C}_{\binom{n}{2}}(G) = 1$. Here, $\mathsf{C}_{\binom{n}{2}}$ has an input node for every potential edge.*

In particular, p-CLIQUE $\notin$ para-AC^0.

Remark 3.3. Recall that Chen et al. [7] showed that p-CLIQUE has no algorithms of running time $f(k) \cdot |n|^{o(k)}$ unless the Exponential Time Hypothesis (ETH) fails. Theorem 3.2 in fact establishes an AC^0 version of this result without using ETH.

Next, we give two equivalent characterizations of para-AC^0. The first one (i.e., between (i) and (ii)) was already mentioned in [11]. Note that in [11] it is required that a problem in para-AC^0 has an AC^0 computable parameterization.

Proposition 3.4. *Let (Q, κ) be a parameterized problem. Consider the following statements.*

(i) $(Q, \kappa) \in$ para-AC^0.

(ii) There is a precomputation*, that is, a computable function pre* $: \mathbb{N} \to \Sigma^*$ *and* AC^0*-circuits* $\left(\mathsf{C}_n\right)_{n\in\mathbb{N}}$ *such that for every* $x \in \Sigma^*$*,*

$$x \in Q \iff \mathsf{C}_{|(x,\mathit{pre}(\kappa(x)))|}(x, \mathit{pre}(\kappa(x))) = 1.$$

(iii) Q is decidable, and there is a computable function $h : \mathbb{N} \to \mathbb{N}$ and AC^0*-circuits* $\left(\mathsf{C}_n\right)_{n\in\mathbb{N}}$ *such that for every $x \in \Sigma^*$ with $|x| \geq h(\kappa(x))$,*

$$x \in Q \iff \mathsf{C}_{|x|}(x) = 1.$$

Then (iii) $\Rightarrow$ (i) and (1) $\Leftrightarrow$ (ii). If, in addition, the parameterization κ can be computed by AC^0*-circuits, then (i) $\Rightarrow$ (iii), i.e., all three statements are equivalent.*

Proof : (i) $\Rightarrow$ (ii) Let $(Q, \kappa) \in$ para-AC^0 be witnessed by a family $\left(\mathsf{C}_{n,k}\right)_{n,k\in\mathbb{N}}$ of circuits. Moreover, let $f, g : \mathbb{N} \to \mathbb{N}$ be the corresponding computable functions in (P2) and (P4). Without loss of generality, we assume that g is increasing and $f(k) = 2^{g(k)}$.

Then, we define the precomputation as $\mathit{pre}(k) := (k, f(k))$. We need to construct a family of circuits $\left(\mathsf{D}_m\right)_{m\in\mathbb{N}}$ such that for every $x \in \Sigma^*$, $y := \left(x, \mathit{pre}(\kappa(x))\right)$, and $m := |y|$

$$x \in Q \iff \mathsf{D}_m(y) = 1. \tag{1}$$

The circuit D_m is basically an $\bigvee$-gate on all possible $\mathsf{C}_{n,k}$'s with $n + f(k) \leq m$. On input $y = \left(x, \mathit{pre}(\kappa(x))\right) = \left(x, (\kappa(x), f(\kappa(x)))\right)$, it detects the input x and the parameter $k = \kappa(x)$, and then uses $\mathsf{C}_{n,k}$ to evaluate on x. Clearly (1) holds. Note the size of D_m can be bounded as

$$|\mathsf{D}_m| \leq O\left(\sum_{n+f(k)\leq m} |\mathsf{C}_{n,k}|\right) = O\left(\sum_{n+f(k)\leq m} f(k) \cdot n^{O(1)}\right) \leq m^{O(1)},$$

where the last equality is by (P2) and $f(k) \leq m$. The dlogtime-uniformity of D_m is also easy to see by (P4) and $f(k) = 2^{g(k)}$.

(ii) $\Rightarrow$ (i) Given AC^0-circuits $\left(\mathsf{C}_m\right)_{m\in\mathbb{N}}$ and a precomputation *pre* $: \mathbb{N} \to \Sigma^*$ as in (2), it is our goal to construct a family $\left(\mathsf{D}_{m,k}\right)_{m,k\in\mathbb{N}}$ of circuits which satisfies (P1) – (P4) in Definition 3.1. For every

$m, k \in \mathbb{N}$ let $\mathsf{D}_{m,k}$ simulate the circuit $\mathsf{C}_{m+|\mathit{pre}(k)|}(\_, \mathit{pre}(k))$, i,e, we fix the second part of the input of $\mathsf{C}_{m+|\mathit{pre}(k)|}$ as $\mathit{pre}(k)$. Then for every $x \in \Sigma^m$

$$\mathsf{D}_{m,\kappa(x)}(x) = 1 \iff \mathsf{C}_{m+|\mathrm{pre}(\kappa(x))|}(x, \mathit{pre}(k)) = 1 \iff x \in Q.$$

This establishes (P3). The conditions on the depth, the size, and the uniformity of $\mathsf{D}_{m,k}$ are routine.

(iii) $\Rightarrow$ (i) Let the AC^0-circuits $(\mathsf{C}_n)_{n\in\mathbb{N}}$ be as in (iii) and let $n, k \in \mathbb{N}$. By assumption, if $n \geq h(k)$, then the circuit C_n satisfies that $\big(x \in Q \iff \mathsf{C}_n(x) = 1\big)$ for every $x \in \Sigma^n$ with $\kappa(x) = k$. So we can take $\mathsf{D}_{n,k} := \mathsf{C}_n$. Otherwise, $n < h(k)$, then we define

$$\mathsf{D}_{n,k}(x) := \bigvee_{y \in Q \cap \Sigma^n} x = y,$$

Here $x = y$ is the abbreviation of the circuit $\bigwedge_{i\in[n]} x_i = y_i$, where every x_i (y_i) is the ith bit of x (y, respectively).

Now assume that there are AC^0-circuits $\big(\mathsf{PC}_n\big)_{n\in\mathbb{N}}$ such that for every $x \in \Sigma^*$ we have $\mathsf{PC}_{|x|}(x) = \kappa(x)$. We show the direction from (i) to (iii). Let $\big(\mathsf{C}_{n,k}\big)_{n,k\in\mathbb{N}}$, $f, g : \mathbb{N} \to \mathbb{N}$ be as stated in Definition 3.1 for (i). Again, we assume that g is increasing and $f(k) = 2^{g(k)}$. Now for every $n \in \mathbb{N}$ and $x \in \Sigma^n$ we define

$$\mathsf{D}_n(x) = \bigvee_{\substack{k\in\mathbb{N} \text{ with} \\ f(k)\leq n}} \big((\mathsf{PC}_n(x) = k) \wedge \mathsf{C}_{n,k}(x)\big).$$

Then for every $x \in \Sigma^n$ with $k := \kappa(x)$ and $|x| \geq f(k)$ it holds

$$x \in Q \iff \mathsf{C}_{n,k}(x) = 1 \iff \mathsf{D}_n(x) = 1.$$

It is easy to verify that $\big(\mathsf{D}_n\big)_{n\in\mathbb{N}}$ are AC^0-circuits. □

In order to use Theorem 3.2 to show para-AC^0 lower bounds for other problems, we introduce a more restricted form of fpt-reductions.

Definition 3.5. Let (Q, κ) and (Q', κ') be two parameterized problems. A *para-*AC^0*-reduction* from (Q, κ) to (Q', κ') is a mapping $R : \Sigma^* \to \Sigma^*$ such that:

(R1) For all $x \in \Sigma^*$ we have $\big(x \in Q \iff R(x) \in Q'\big)$.

(R2) There is a family of circuits $\big(\mathsf{C}_{n,k}\big)_{n,k\in\mathbb{N}}$, whose depth is bounded by a fixed constant, such that

- (a) for all $x \in \Sigma^*$, $\mathsf{C}_{|x|,\kappa(x)}(x)$ outputs $R(x)$;
- (b) $\big|\mathsf{C}_{n,k}\big| \leq f(k) \cdot |x|^{O(1)}$ for a computable function $f : \mathbb{N} \to \mathbb{N}$;
- (c) there is a deterministic Turing machine that on input $(1^n, 1^k)$ computes the circuit $\mathsf{C}_{n,k}$ in time $g(k) + O(\log n)$, where $g : \mathbb{N} \to \mathbb{N}$ is a computable function.

(R3) There is a computable function $h : \mathbb{N} \to \mathbb{N}$ such that $\kappa'(R(x)) \leq h(\kappa(x))$ for all $x \in \Sigma^*$.

If there is a para-AC^0-reduction from (Q, κ) to (Q', κ'), then we write $(Q, \kappa) \leq^{\mathrm{pac}} (Q', \kappa')$.

However, in general para-AC^0 is *not* closed under para-AC^0-reductions as witnessed by the following example.

Example 3.6. Define

$$Q := \left\{ (x,b) \,\middle|\, x \in \{0,1\}^* \text{ and } b = \sum_{i \in [|x|]} x_i \mod 2 \right\}.$$

Clearly, Q is equivalent to the classical PARITY problem of deciding whether there is an even number of 1's in x. Thus $Q \notin \text{AC}^0$.

We define two polynomial time computable parameterizations of Q by

$$\kappa_1(x,b) := 0 \quad \text{and} \quad \kappa_2(x,b) := \sum_{i \in [|x|]} x_i \mod 2.$$

Then it is easy to see that $(Q,\kappa_1) \notin$ para-AC^0 and $(Q,\kappa_2) \in$ para-AC^0; yet $(Q,\kappa_1) \leq^{\text{pac}} (Q,\kappa_2)$ by the identity mapping $R(x,b) = (x,b)$.

Note (Q,κ_2) also serves as a counterexample for the direction from (i) to (iii) in Proposition 3.4.

Therefore we need a further requirement on pac-reductions. The previous example suggests to require the AC^0-computability of the parameterization (as done in [11]). In fact, para-AC^0 is closed under those reductions. However, we choose another requirement, which is simpler to verify and is satisfied by almost all natural reductions.

Definition 3.7. Let (Q,κ) and (Q',κ') be two parameterized problems. A *weak para-*AC^0*-reduction* from (Q,κ) to (Q',κ') is a para-AC^0-reduction which satisfies:

(R3') There is a computable function $h : \mathbb{N} \to \mathbb{N}$ such that $\kappa'(R(x)) = h(\kappa(x))$ for all $x \in \Sigma^*$.

$(Q,\kappa) \leq^{\text{pwac}} (Q',\kappa')$ means that there is a weak para-AC^0-reduction from (Q,κ) to (Q',κ').

It is straightforward to verify that para-AC^0 is closed under weak para-AC^0-reductions.

Lemma 3.8. *Let (Q,κ) and (Q',κ') be two parameterized problems with $(Q,\kappa) \leq^{\text{pwac}} (Q',\kappa')$. If $(Q',\kappa') \in$* para-AC^0*, then $(Q,\kappa) \in$* para-AC^0*, too.*

It is well known that p-CLIQUE is fpt-reducible to p-DOMINATING-SET. The reduction presented in the next proof is even a weak para-AC^0-reduction and thus, by Theorem 3.2 and the previous lemma yields:

Proposition 3.9. p-DOMINATING-SET $\notin$ para-AC^0.

Proof : By the previous remark it suffices to present a weak para-AC^0-reduction from p-CLIQUE to p-DOMINATING-SET. Let (G,k) be an instance of p-CLIQUE with $G = (V,E)$. We may assume that E is not empty. Let $k \geq 2$. We construct a graph $H = (W,F)$ with

$$G \text{ has a } k\text{-clique} \quad \Longleftrightarrow \quad H \text{ has a dominating set of size } k + \binom{k}{2}. \tag{2}$$

Let new(i) and new(i,j) with $i,j \in [k]$ and $i < j$ be new vertices. The vertex set W of H is the disjoint union of three types of sets:

(a) $\{\text{new}(i)\} \cup V(i)$ for $i \in [k]$, where each $V(i)$ is a (disjoint) copy of V;

(b) $\{i\} \times \{j\} \times V(i) \times V(j)$ for $i, j \in [k]$ with $i < j$;

(c) $\{\text{new}(i,j)\} \cup E(i,j)$ for $i, j \in [k]$ with $i < j$, where each $E(i,j)$ is a (disjoint) copy of the edge set E.

We denote by $v(i)$ the copy of $v \in V$ in $V(i)$ and by $e(i,j)$ the copy of $e \in E$ in $E(i,j)$. The set F consists of the following edges:

(d) for $i \in [k]$ edges that make $\{\text{new}(i)\} \cup V(i)$ a clique;

(e) for $i, j \in [k]$ with $i < j$ edges that make $\{\text{new}(i,j)\} \cup E(i,j)$ a clique;

(f) for $i, j \in [k]$ with $i < j$ and every $(i, j, u(i), v(j)) \in \{i\} \times \{j\} \times V(i) \times V(j)$ an edge from this vertex to every $u'(i) \in V(i)$ with $u \neq u'$ and an edge to every $v'(j) \in V(j)$ with $v \neq v'$; furthermore, if $\{u, v\} \in E$, then an edge from $(i, j, u(i), v(j))$ to the vertex $\{u, v\}(i, j)$ (in $E(i, j)$).

Then the equivalence (2) holds. In fact, first assume that $C := \{u_1, \ldots, u_k\}$ is a k-clique in G. Then the set

$$D(C) := \big\{u_1(1), \ldots, u_k(k)\big\} \cup \big\{\{u_i, v_j\}(i,j) \mid i, j \in [k] \text{ and } i < j\big\}$$

is a dominating set in H.

Conversely, assume D is a dominating set in H of size $k + \binom{k}{2}$. In view of the elements of the form new(i) and new(i,j), we see that D must contain elements of each $\{\text{new}(i)\} \cup V(i)$ and of each $\{i\} \times \{j\} \times V(i) \times V(j)$. Thus, D consists of exactly one element of each of these sets. Note that the element from $\{\text{new}(i)\} \cup V(i)$ must be distinct from new(i), as otherwise at most one element from every $\{i\} \times \{j\} \times V(i) \times V(j)$ can be dominated by D (but $|V| \geq 2$ as, by assumption, $E \neq \emptyset$). So let $u_i(i)$ with $u_i \in V$ be the element of D in $V(i)$. As D dominates the element $(i, j, u_i(i), u_j(j))$, we see that it has to be dominated by some element of $E(i,j)$; that is, $\{u_i, u_j\} \in E$. Thus $\{u_1, \ldots, u_k\}$ is a clique. □

Corollary 3.10. *Let $t, d \geq 1$ with $t + d \geq 3$. Then p-WSAT$(\Gamma_{t,d}) \notin$ para-AC^0.*

Proof : For every graph $G = (V, E)$ we define a propositional formula

$$\delta_G := \bigwedge_{\substack{u, v \in V \text{ with} \\ u \neq v \text{ and } \{u,v\} \notin E}} \neg X_u \vee \neg X_v.$$

Clearly, for every $k \in \mathbb{N}$,

$$G \text{ has a } k\text{-clique} \iff \delta_G \text{ has a satisfying assignment of weight } k. \tag{3}$$

This gives a weak para-AC^0-reduction from p-CLIQUE to p-WSAT$(\Gamma_{1,2})$, or p-WSAT$(\Gamma_{t,1})$ in case $t \geq 2$. □

Similarly, one can show that basic problems like p-HOM, p-EMB, p-SUBGRAPH-ISOMORPHISM, and p-MC(Σ_1^1) are not in para-AC^0 (we use the notations of [12]).

In view of Corollary 3.10 the reader might wonder about the status of p-WSAT$(\Gamma_{1,1})$. Using the color-coding technique as in [3], one can show that the problem is in fact solvable in para-AC^0. We present a more logic-oriented technique for such proofs. It uses FO$(<, +, \times)$ instead of dlogtime-uniform AC^0. First, we recall the following lemma from [13, page 349]:

Lemma 3.11. *For every sufficiently large $n \in \mathbb{N}$, it holds that for all $k \leq n$ and for every k-element subset X of $[n]$, there exists a prime $p < k^2 \cdot \log_2 n$ and $q < p$ such that the function $h_{p,q} : [n] \to \{0, \ldots, k^2 - 1\}$ given by $h_{p,q}(m) := (q \cdot m \mod p) \mod k^2$ is injective on X.*

For $n \in \mathbb{N}$ denote by $<^{[n]}$ the natural ordering on $[n]$. Clearly, if $\mathcal{A}$ is any ordered structure, then $(A, <^{\mathcal{A}})$ is isomorphic to $([|A|], <^{[|A|]})$ and the isomorphism is unique. Furthermore, for ternary relation symbols $+$ and $\times$ we consider the ternary relations $+^{[n]}$ and $\times^{[n]}$ on $[n]$ that are the relations underlying the addition and the multiplication of $\mathbb{N}$ restricted to $[n]$. That is,

$$\begin{aligned} +^{[n]} &:= \big\{(a,b,c) \;\big|\; a,b,c \in [n] \text{ with } c = a + b\big\}, \\ \times^{[n]} &:= \big\{(a,b,c) \;\big|\; a,b,c \in [n] \text{ with } c = a \cdot b\big\}. \end{aligned} \tag{4}$$

Let τ be a vocabulary which does not contain the relation symbols $<, +, \times$ and set $\tau_{<,+,\times} := \tau \cup \{<, +, \times\}$. We say that a $\tau_{<,+,\times}$-structure $\mathcal{A}$ *has built-in addition and built-in multiplication* if $(A, <^{\mathcal{A}}, +^{\mathcal{A}}, \times^{\mathcal{A}})$ is isomorphic to $([|A|], <^{[|A|]}, +^{[|A|]}, \times^{[|A|]})$. Sometimes we write $\varphi \in \mathrm{FO}(<, +, \times)$ to emphasize that φ is a first-order formula in a vocabulary containing the symbols $<, +, \times$.

Corollary 3.12. *There is a computable function which associates every $k \in \mathbb{N}$ with a structure $\mathcal{C}(k)$ and every* FO*-formula $\varphi(x)$ with an* FO$(<, +, \times)$*-sentence χ_φ such that for every structure $\mathcal{A}$,*

$$\begin{aligned} &[\mathcal{A} : \mathcal{C}(k)] \models \chi_\varphi \\ &\quad \iff \textit{there are pairwise distinct } x_1, \ldots, x_k \in A \textit{ with } \mathcal{A} \models \varphi(x_i) \textit{ for every } i \in [k]. \end{aligned} \tag{5}$$

Here, $[\mathcal{A} : \mathcal{C}(k)] := \mathcal{B} = \big(\mathcal{A} \dot\cup \mathcal{C}(k), U^{\mathcal{B}}, <^{\mathcal{B}}, +^{\mathcal{B}}, \times^{\mathcal{B}}\big)$ is defined as follows.

- *$\mathcal{A} \dot\cup \mathcal{C}(k)$ is the disjoint union of $\mathcal{A}$ and $\mathcal{C}(k)$.*
- *$U^{\mathcal{B}} := A$.*
- *$<^{\mathcal{B}}$ is an ordering of B and every element of A precedes all elements of $C(k)$. Furthermore $<^{\mathcal{B}}$ extends the ordering $\prec^{\mathcal{C}(k)}$ given in $\mathcal{C}(k)$.*
- *$\mathcal{B}$ has built-in addition $+^{\mathcal{B}}$ and multiplication $\times^{\mathcal{B}}$.*

Proof : Let $\tau_0 := \{K, \prec, F\}$ with unary K, binary $\prec$, and ternary F. We first define $\mathcal{C} = \mathcal{C}(k)$, a τ_0-structure, which basically embodies all functions from $\big\{0, \ldots, k^2 - 1\big\}$ to $\{0, \ldots, k - 1\}$.

- $C = \big\{0, \ldots, k^{k^2} - 1\big\}$.
- $K^{\mathcal{C}} := \{k - 1\}$ is the singleton set containing the k-th element in $\mathcal{C}$.
- Let $\prec^{\mathcal{C}}$ is the natural ordering on C.
- Let $g_0, \ldots, g_{k^{k^2}-1}$ be an enumeration of all functions from $\big\{0, \ldots, k^2 - 1\big\}$ to $\{0, \ldots, k - 1\}$. Then we define a ternary relation

$$F^{\mathcal{C}} := \big\{(i, x, y) \;\big|\; g_i(x) = y\big\}.$$

Now let $\mathcal{A}$ be any structure in a vocabulary τ. We may assume that $\tau \cap \tau_0 = \emptyset$ by renaming symbols in τ if necessary. Then, by the disjoint union $\mathcal{A} \dot\cup \mathcal{C}(k)$ of $\mathcal{A}$ and $\mathcal{C}(k)$ we mean the structure

$$\big(A \dot\cup C(k), (R^{\mathcal{A}})_{R\in\tau}, (S^{\mathcal{C}(k)})_{S\in\tau_0}\big),$$

where $A \dot\cup C(k)$ is the disjoint union of the sets A and $C(k)$.

We view the universe of $\mathcal{B} := \mathcal{A} \dot\cup \mathcal{C}(k)$ as

$$B = \Big\{0, \ldots, |A| + k^{k^2} - 1\Big\}.$$

In order to make formulas more readable, we introduce some abbreviations. We freely use terms as in $x + y + z = w$ $\big($which is equivalent to $\exists u(+xyu \wedge +uzw)\big)$. Note that $k-1$ can be defined in $\mathcal{B}$ as the unique x satisfying the formula

$$\exists u \exists v\big(\neg Uu \wedge \forall z(z < u \to Uz) \wedge Kv \wedge u + x = v\big).$$

Thus, the number k, i.e., the $(k+1)$th element in $\mathcal{B}$, can also be easily defined. Clearly, $x = (y \mod z)$ is an abbreviation for

$$\exists u\big(x = u \times z + y \wedge y < z\big).$$

Moreover, $g_i(x) = y$ is a shorthand for the formula

$$\begin{aligned}\gamma(i,x,y) := \exists u\big(&\neg Uu \wedge \forall z(z < u \to Uz)\\ &\wedge \exists i'\exists x'\exists y'(i' = u + i \wedge x' = u + x \wedge y' = u + y \wedge Fi'x'y')\big).\end{aligned}$$

Now let

$$\chi_\varphi := \exists p \exists q \exists i\ \psi_\varphi(p,q,i),$$

where

$$\psi_\varphi(p,q,i) := \forall j\big(j < k \to \exists u(Uu \wedge \varphi^U(u) \wedge \rho(p,q,i,u,j)\big)$$

$\big($here $\varphi^U(u)$ is obtained from $\varphi(u)$ by relativizing all quantifiers to $U\big)$, and

$$\rho(p,q,i,u,j) := g_i\big((q \times (u \mod p) \mod p) \mod k^2) = j\big).\ ^{1}$$

Note that $\rho(p,q,i,u,j)$ is equivalent to

$$g_i\big(h_{p,q}(u)\big) = j,$$

where $h_{p,q}$ is defined in Lemma 3.11. We replaced $(q \times u \mod p)$ by $(q \times (u \mod p) \mod p)$, since $q \times u$ might exceed the size of $\mathcal{B}$, i.e., $|A| + k^{k^2}$.

We still need to show the equivalence (5). The direction from right to left is easy, since $\mathcal{B} \models \chi_\varphi$ means that for some p, q, i there exist $u_0, u_1, \ldots, u_{k-1} \in A$ with

$$\mathcal{A} \models \varphi(u_j) \quad \text{and} \quad g_i(h_{p,q}(u_j)) = j$$

for every $0 \le j < k$. The second condition implies that all u_j's are distinct.

[1] Let us emphasize that $\rho(p,q,i,u,j)$ does not depend on k which is defined from the unary relation $K^{\mathcal{C}}$. Hence neither does χ_φ.

For the other direction, assume that there are k elements $u_0, u_1, \ldots, u_{k-1} \in A$ with $\mathcal{A} \models \varphi(u_j)$ for all j. By Lemma 3.11 there exist $p < k^2 \cdot \log_2 n$ and $q < p$ such that $h_{p,q}$ is injective on $\{u_0, \ldots, u_{k-1}\}$. Since the range of $h_{p,q}$ is $\{0, \ldots, k^2 - 1\}$, we can choose a function $g_i : \{0, \ldots, k^2 - 1\} \to \{0, \ldots, k-1\}$ such that $g_i(h_{p,q}(u_j)) = j$ for every $0 \leq j < k$. Since $q < p < k^2 \cdot \log_2 n$, we can guarantee that

$$(q \times (u \mod p)) < k^4 (\log_2 |A|)^2 \leq |A| + k^{k^2}.$$

Hence $\rho(p, q, u, j)$ gives the correct answer. □

Let χ_φ^{-1} be the formula obtained by defining $\psi_\varphi(p, q, i)$ by

$$\psi_\varphi(p, q, i) := \forall j \big(j < (k-1) \to \exists u (Uu \wedge \varphi^U(u) \wedge \rho(p, q, u, j)) \big),$$

where the formula $\rho(p, q, u, j)$ remains unchanged. Then the last part of the previous proof shows that

$$\begin{aligned} &[\mathcal{A} : \mathcal{C}(k)] \models \chi_\varphi^{-1} \\ &\quad \iff \text{there are pairwise distinct } x_1, \ldots, x_{k-1} \in A \text{ with } \mathcal{A} \models \varphi(x_i) \text{ for every } i \in [k-1]. \end{aligned} \tag{6}$$

Proposition 3.13. p-WSAT$(\Gamma_{1,1}) \in$ para-AC^0

Proof : Let δ be a propositional formula in $\Gamma_{1,1}$. Clearly δ has a satisfying assignment of Hamming weight k if and only if in δ

(i) no propositional variable occurs both positively and negatively,

(ii) there are at least k propositional variables,

(iii) there are *no* $k+1$ variables which occur positively.

Without loss of generality, we assume that (i) always holds. It is easy to view δ as a structure $\mathcal{A}(\delta)$ such that for some formulas $\varphi_{\text{var}}(x)$ and $\varphi_{\text{pos}}(x)$ we have for every propositional variable Y,

$$\begin{aligned} \mathcal{A}(\delta) \models \varphi_{\text{var}}(Z) &\iff Z \text{ occurs in } \delta \\ \mathcal{A}(\delta) \models \varphi_{\text{pos}}(Z) &\iff Z \text{ occurs positively in } \delta. \end{aligned}$$

By Corollary 3.12 and (6) we see that

$$\delta \text{ has a satisfying assignment of Hamming weight } k \iff [\mathcal{A}(\delta) : \mathcal{C}(k+1)] \models \big(\neg \chi_{\varphi_{\text{pos}}} \wedge \chi_{\varphi_{\text{var}}}^{-1} \big).$$

This equivalence shows that the problem can be decided by FO$(<, +, \times)$ after a precomputation on the parameter k. The result then follows from Proposition 3.4. □

4. Inapproximability of p-CLIQUE by para-AC^0

We recall the notion of fpt approximation introduced in [10]. We present the definition for p-CLIQUE, the problem which interests us. It can easily be generalized to any maximization problem.

If not stated otherwise, $\rho : \mathbb{N} \to \mathbb{R}_{\geq 1}$ is always a computable function such that the mapping $k \mapsto k/\rho(k)$ is nondecreasing and unbounded.

Definition 4.1. An algorithm $\mathbb{A}$ is a *parameterized approximation for* p-CLIQUE *with approximation ratio* ρ if for every graph G and $k \in \mathbb{N}$ with $\omega(G) \geq k$ the algorithm $\mathbb{A}$ computes a clique C of G such that

$$|C| \geq \frac{k}{\rho(k)}.$$

Here the clique number $\omega(G)$ is the size of a maximum clique of G. If the running time of $\mathbb{A}$ is bounded by $f(k) \cdot |G|^{O(1)}$ where $f : \mathbb{N} \to \mathbb{N}$ is computable, then $\mathbb{A}$ is an *fpt approximation algorithm*.

We tend to believe that p-CLIQUE has no fpt approximation algorithm for any ratio ρ. Since para-AC^0 is a class of decision problems, in order to prove a lower bound it is more convenient to deal with decision algorithms instead of algorithms computing a clique.

Definition 4.2 ([10])**.** A decision algorithm $\mathbb{A}$ is a *parameterized cost approximation for* p-CLIQUE *with approximation ratio* ρ if for every graph G and $k \in \mathbb{N}$,

- if $k \leq \omega(G)/\rho(\omega(G))$, then $\mathbb{A}$ accepts (G, k);
- if $k > \omega(G)$, then $\mathbb{A}$ rejects (G, k).

In other words, $\mathbb{A}$ decides the *promise* problem:

p-GAP$_\rho$-CLIQUE	
Instance:	A graph G and $k \in \mathbb{N}$ such that either $k \leq \omega(G)/\rho(\omega(G))$ or $k > \omega(G)$.
Parameter:	k.
Question:	Is $k \leq \omega(G)/\rho(\omega(G))$?

The intuition behind this definition: If G contains a clique far bigger than k, detecting a k-clique might become easier. It is straightforward to verify that if p-CLIQUE has no parameterized fpt cost approximation of ratio ρ, then it has no parameterized fpt approximation of ratio ρ either [10].

Theorem 4.3. *Let $\rho : \mathbb{N} \to \mathbb{R}_{\geq 1}$ be a computable function such that the mapping $k \mapsto k/\rho(k)$ is nondecreasing and unbounded. Then*

$$p\text{-GAP}_\rho\text{-CLIQUE} \notin \text{para-AC}^0.$$

Our original proof of this result was based on a generalization of the machinery developed in [21], a generalization we first used to prove that AC^0 circuits are not sensitive to planted cliques of a reasonable size, see Theorem 4.6. The much simpler proof of Theorem 4.6 we present here is based on Beame's Clique Switching Lemma [5] and was suggested to us anonymously.

4.1. Beame's Clique Switching Lemma. Let $n \in \mathbb{N}$. We consider graphs with vertex set $[n]$. To represent functions on those graphs, every potential edge $e \in \binom{[n]}{2}$ is encoded by a Boolean variable X_e. We set

$$\mathcal{X}_n := \left\{ X_e \;\middle|\; e \in \binom{[n]}{2} \right\}.$$

In particular, $X_e = 1$ means that e is present in the given graph, otherwise $X_e = 0$. Sometimes, it is convenient to understand e as a natural number with $e \in \left[\binom{n}{2}\right]$. Then, e is the eth potential edge in an n-vertex graph, and X_e is the eth variable in $\mathcal{X}_n$.

For every $\ell \in [n]$ and $q \in \mathbb{R}$ with $0 \leq q \leq 1$ let $\mu \in \mathscr{C}_n^{\ell,q}$ be a *random restriction*, $\mu : \mathcal{X}_n \to \{0, 1, \star\}$, generated as follows:

– Choose $U \in \binom{[n]}{\ell}$ uniformly at random and then set $\mu(X_e) := \star$ for every $e \in \binom{U}{2}$.

– For $e \notin \binom{U}{2}$ we set $\mu(X_e) := 1$ with probability q and $\mu(X_e) := 0$ with probability $1 - q$.

Let F be a boolean function defined on the set of assignments from $\mathcal{X}_n$ to $\{0,1\}$ and $\mu \in \mathscr{C}_n^{\ell,q}$. The function $F\restriction_\mu$ is defined on the set of assignments from $\mu^{-1}(\star)$ to $\{0,1\}$ by: for any assignment $S : \mu^{-1}(\star) \to \{0,1\}$

$$F\restriction_\mu (S) := F(S \cup \mu),$$

where $S \cup \mu : \mathcal{X}_n \to \{0,1\}$ is the assignment given by

$$(S \cup \mu)(X_e) := \begin{cases} S(X_e) & \text{if } X_e \in \mu^{-1}(\star) \\ \mu(X_e) & \text{otherwise.} \end{cases}$$

Recall that a rooted binary tree is a *decision tree* on some variable set $\mathcal{X} \subseteq \mathcal{X}_n$ if every leaf is labeled either 0 or 1, every internal node is labelled by a variable of $\mathcal{X}$, and the edges between an internal node and its two children are labelled 0 and 1. The *vertex height* of a path P in T is the number of distinct vertices occurring in edges e such that the corresponding X_e appears in P. The *vertex height* $|T|_v$ of T is the maximum vertex height of a path in T.

For any boolean function F as above, we set

$$\text{DTdepth}_{\text{vertex}}(F) = \min\{|T|_v \mid T \text{ a decision tree computing } F\}.$$

Lemma 4.4 (Beame's Clique Switching Lemma [5][2])**.** *Let $n, r \in \mathbb{N}$ and $0 \le q \le 1/2$. Moreover, let F be a* DNF*-formula of variable set $\mathcal{X}_n$ with conjunctive clauses of vertex length at most r. For $s, \ell \in \mathbb{N}$ with $\ell := pn$, where $s \ge 0$ and $\ell := pn$ with $p \le 1/(r(2/q)^{(r+s)/2})$, we have*

$$\Pr_{\mu \in \mathscr{C}_n^{\ell,q}} \left[\text{DTdepth}_{\text{vertex}}\big(F\restriction_\mu\big) > s \right] < \frac{8\big((2/q)^{(s+r-1)/2} pr\big)^s}{3}.$$

Here, the vertex length *of a clause is the number of distinct vertices in edges e with X_e appearing in this clause.*

We apply Lemma 4.4 inductively on bounded-depth circuits.

Lemma 4.5. *Assume*

– $k : \mathbb{N} \to \mathbb{R}_+$ *with $k(n) \le \log_2 n$ for all sufficiently large n and $\lim_{n\to\infty} k(n) = \infty$,*

– $S : \mathbb{N} \to \mathbb{N}$ *with $S(n) \ge n$,*

– $d : \mathbb{N} \to \mathbb{N}$.

Define $q : \mathbb{N} \to \mathbb{R}_+$ and $s : \mathbb{N} \to \mathbb{N}$ by

$$q(n) := n^{-1/k(n)} \quad \textit{and} \quad s(n) := \left\lfloor \sqrt{k(n)(\log_n S(n) d(n))} \right\rfloor, \tag{7}$$

and $\ell_i : \mathbb{N} \to \mathbb{N}$ inductively by

$$\ell_0(n) := n \quad \textit{and} \quad \ell_{i+1}(n) := \left\lfloor \frac{\ell_i(n)}{n^{5s(n)/k(n)}} \right\rfloor. \tag{8}$$

[2] See the imbalanced version of [5, Lemma 3] mentioned in the first paragraph of page 12 of that paper.

Then, $\ell_{d(n)}(n) = n^{1-\Theta\left(5d(n)\sqrt{(\log_n S(n)d(n))/k(n)}\right)}$ *and for every circuit* C *with variable set* $\mathcal{X}_n$*, size bounded by* $S(n)$*, and depth bounded by* $d(n)$*,*

$$\Pr_{\mu\in\mathscr{C}_n^{\ell_{d(n)}(n),q(n)}}\Big[\mathsf{C}\restriction_\mu \textit{ is constant}\Big] = 1 - o(1).$$

Moreover, the convergence rate can be bounded in terms of S*,* d*, and* k*.*

Proof : We fix an $n \in \mathbb{N}$ and let $k := k(n)$, $q := q(n)$, $S := S(n)$, $d := d(n)$, $s := s(n)$, and $\ell_i := \ell_i(n)$. It is easy to see that for every $i \in [d]$

$$\frac{\ell_i}{\ell_{i+1}} \geq n^{5s/k} \tag{9}$$

and

$$\ell_d = n^{1-\Theta\left(5d\sqrt{(\log_n Sd)/k}\right)}.$$

Let μ_0 be the empty restriction, i.e., $\mu_0(X_e) = \star$ for every $X_e \in \mathcal{X}_n$. For every $i \in [d]$ we let π_i be a random restriction from $\mathscr{C}_{\ell_{i-1}}^{\ell_i,q}$. We set

$$\mu_i := \mu_{i-1} \circ \pi_i,$$

where $\circ$ is defined in such a way that for every $X_e \in \mathcal{X}_n$,

$$\mu_i(X_e) = \begin{cases} \mu_{i-1}(X_e) & \text{if } \mu_{i-1}(X_e) \neq \star, \\ \pi_i(X_{e'}) & \text{if } X_e \text{ is the } e'\text{th variable in } \mathcal{X}_{\ell_{i-1}} \text{ with } \mu_{i-1}(X_e) = \star\,. \end{cases}$$

It is easy to see that $\mu := \mu_d$ has the distribution of $\mathscr{C}_n^{\ell_d,q}$.

Assume n is sufficiently large. Hence (7) implies $s \geq 2$. Moreover, let

$$p_i := \Pr\Big[\text{there is a gate } g \text{ of depth} \leq i \text{ with } \mathrm{DTdepth}_{\mathrm{vertex}}(\mathsf{C}_g\restriction_{\mu_i}) > s \;\Big|\; \text{all gates of depth} \leq i \text{ have } \mathrm{DTdepth}_{\mathrm{vertex}}(\mathsf{C}_g\restriction_{\mu_{i-1}}) \leq s\Big],$$

where C_g is the subcircuit of C with root g.

Since every gate g at height 0 depends only on one edge variable X_e, and $\mu_0(X_e) = \star$, we conclude that $\mathrm{DTdepth}_{\mathrm{vertex}}(\mathsf{C}_g\restriction_{\mu_0}) = 2$ and $p_0 = 0$ (recall $s \geq 2$).

Now assume that for all gates g of depth $\leq i$ we have $\mathrm{DTdepth}_{\mathrm{vertex}}(C_g\restriction_{\mu_i}) \leq s$. Let g be an $\bigvee$-gate of depth $i+1 < d$. [3] It follows that $C_g\restriction_{\mu_i}$ can be expressed as a DNF-formula with terms of vertex length at most s. By (9)

$$p := \frac{\ell_{i+1}}{\ell_i} \leq n^{-5s/k}.$$

[3] In particular, g is *not* the output gate.

Then by Beame's Clique Switching Lemma (with $r \leftarrow s$ and $s \leftarrow s$) and assuming n is sufficiently large,

$$\begin{aligned}
&\Pr_{\pi_{i+1} \in \mathscr{C}_{\ell_i}^{\ell_{i+1},q}} \Big[\mathrm{DTdepth}_{\mathrm{vertex}}\big((C_g \restriction_{\mu_i}) \restriction_{\pi_{i+1}}\big) > s\Big] \\
&\qquad \le \frac{8}{3}\left((2/q)^{s-1/2} ps\right)^s \\
&\qquad \le 3\left(s\big(2n^{1/k}\big)^{s-1/2} n^{-5s/k}\right)^s \\
&\qquad \le 3n^{3s^2/k} n^{-5s^2/k} \\
&\qquad = 3n^{-2s^2/k} \\
&\qquad \le 3n^{-1.9 \cdot \log_n Sd} \qquad \big(\text{by (7) and for } n, \text{ and hence } k, \text{ sufficiently large}\big) \\
&\qquad = \frac{3}{(Sd)^{1.9}} \le \frac{o(1)}{Sd} \qquad \big(\text{by } S(n) \ge n\big).
\end{aligned}$$

To obtain the third inequality, i.e.,

$$3\left(s\big(2n^{1/k}\big)^{s-1/2} n^{-5s/k}\right)^s \le 3n^{3s^2/k} n^{-5s^2/k},$$

we argue, using $k \le \log_2 n$,

$$s\big(2n^{1/k}\big)^{s-1/2} \le s\big(2n^{1/k}\big)^s = s2^s n^{s/k} \le 2^{2s} n^{s/k} \le n^{2s/k} n^{3s/k} = n^{5s/k}.$$

Since there are at most S many $\bigvee$-gates at depth $i+1$ in the circuit C, we have $p_{i+1} \le o(1)/d$ by a union bound.

Finally, let o be the output gate of C of depth d, i.e., $\mathsf{C}_o = \mathsf{C}$. Assume $\mathrm{DTdepth}_{\mathrm{vertex}}(\mathsf{C}_g \restriction_{\mu_i}) \le s$ for all gates g of depth $< d$. Again applying Beame's Clique Switching Lemma with parameters $r \leftarrow s$ and $s \leftarrow 1$, we obtain

$$\begin{aligned}
&\Pr_{\pi_d \in \mathscr{C}_{\ell_{d-1}}^{\ell_d,q}} \Big[\mathrm{DTdepth}_{\mathrm{vertex}}\big((\mathsf{C}_o \restriction_{\mu_{d-1}}) \restriction_{\pi_d}\big) > 1\Big] \\
&\qquad \le \frac{8}{3}\left((2/q)^{s/2} ps\right) < 3\left(s\big(2n^{1/k}\big)^{s/2} n^{-5s/k}\right) \\
&\qquad \le 3\left(2^s\big(2n^{1/k}\big)^{s/2} n^{-5s/k}\right) = 3\left(2^{1.5s} n^{s/2k} n^{-5s/k}\right) \\
&\qquad \le 3n^{2s/k} n^{-5s/k} = 3n^{-3s/k} \le 3 \cdot 2^{-3s} = o(1) \qquad \big(\text{by } \lim_{n\to\infty} k(n) = \infty\big).
\end{aligned}$$

Thus the probability of the event

either for a gate g of depth $< d$ we have $\mathrm{DTdepth}_v(C_g \restriction_\mu) > s$ or $\mathrm{DTdepth}_v(C \restriction_\mu) > 1$

is $o(1)$, which implies the desired result. □

4.2. A strong $\mathbf{AC}^0$ version of the planted clique conjecture. In the standard planted clique problem, we are given a graph G whose edges are generated by starting with a random graph with universe $[n]$ and edge probability $1/2$, then "planting" (adding edges to make) a random clique on k vertices; the problem asks for efficient algorithms finding such a clique of size k. The problem was addressed

in [18, 19, 2], among many others. It is conjectured that no such algorithm exists for $k = o(\sqrt{n})$. Here, as a consequence of Lemma 4.5, we prove a statement considerably stronger than the AC^0 version of this conjecture.

Let us be more precise. The Erdős-Rényi probability space $\mathrm{ER}(n,p)$, where $n \in \mathbb{N}$ and $p \in \mathbb{R}$ with $0 \leq p \leq 1$, is obtained as follows. We start with the set $[n]$ of vertices. Then we choose every $e \in \binom{[n]}{2}$ as an edge of G with probability p, independently of the choices of other edges.

For $G \in \mathrm{ER}(n,1/2)$ the expected size of a maximum clique is approximately $2\log\, n$. Therefore G almost surely has no clique of size, say, $4\log\, n$. For any graph G with vertex set $[n]$ and any $A \subseteq [n]$ we denote by $G + C(A)$ the graph obtained from G by adding edges such that the subgraph induced on A is a clique. For $n, c \in \mathbb{N}$ with $c \in [n]$ and $p \in \mathbb{R}$ with $0 \leq p \leq 1$ we consider a second distribution $\mathrm{ER}(n,p,c)$: Pick a random graph $G \in \mathrm{ER}(n,p)$ and a uniformly random subset A of $[n]$ of size c and plant in G a clique on A, thus getting the graph $G + C(A)$. The notation $(G,A) \in \mathrm{ER}(n,p,c)$ should give the information that the random graph was G and that the random subset of $[n]$ of size c was A.

Theorem 4.6. *Let $k : \mathbb{N} \to \mathbb{R}^+$ with $\lim_{n\to\infty} k(n) = \infty$, and $c : \mathbb{N} \to \mathbb{N}$ with $c(n) \leq n^\xi$ for some $0 \leq \xi < 1$. Then for all AC^0 circuits $\big(\mathsf{C}_n\big)_{n\in\mathbb{N}}$,*

$$\lim_{n\to\infty} \Pr_{(G,A)\in \mathrm{ER}(n,n^{-1/k(n)},\, c(n))} \big[\mathsf{C}_n(G) = \mathsf{C}_n(G + C(A))\big] = 1.$$

We first deal with the case where $k(n) \leq \log_2 n$ for all sufficiently large n. The general case will be reduced to it by standard techniques from probability theory.

Lemma 4.7. *Let $k : \mathbb{N} \to \mathbb{R}^+$ with $k(n) \leq \log_2 n$ for all sufficiently large n and $\lim_{n\to\infty} k(n) = \infty$, and $c : \mathbb{N} \to \mathbb{N}$ with $c(n) \leq n^\xi$ for some $0 \leq \xi < 1$. Then for all AC^0 circuits $\big(\mathsf{C}_n\big)_{n\in\mathbb{N}}$,*

$$\lim_{n\to\infty} \Pr_{(G,A)\in \mathrm{ER}(n,n^{-1/k(n)},\, c(n))} \big[\mathsf{C}_n(G) = \mathsf{C}_n(G + C(A))\big] = 1.$$

Moreover, the convergence rate is uniform for all AC^0 circuits of a fixed depth and size.

Proof : Let $(\mathsf{C}_n)_{n\in\mathbb{N}}$ be a family of circuits such that for some $\bar{d}, t \in \mathbb{N}$ every C_n has depth at most $\bar{d}$ and size bounded by n^t. In order to apply Lemma 4.5, we set for $n \in \mathbb{N}$,

$$S(n) = n^t \quad \text{and} \quad d(n) = \bar{d}. \tag{10}$$

By Lemma 4.5, it follows that (recall that $q(n) = n^{-1/k(n)}$)

$$\Pr_{\mu\in\mathscr{C}_n^{\ell_{\bar{d}}(n),q(n)}} \Big[\mathsf{C}_n\restriction_\mu \text{ is constant}\Big] = 1 - o(1), \tag{11}$$

where the $o(1)$ term only depend on S, d, and k, i.e., t, $\bar{d}$ and k. Furthermore,

$$\ell_{\bar{d}}(n) = n^{1-\Theta\left(5d(n)\sqrt{(\log_n S(n)d(n))/k(n)}\right)} = n^{1-o(1)}; \tag{12}$$

the first equality holds by Lemma 4.5 and the second by (10). The key step consists of the following random process, which generates $(G,A) \in \mathrm{ER}(n, n^{-1/k(n)},\, c(n))$ from $\mu \in \mathscr{C}_n^{\ell_{\bar{d}}(n),q(n)}$.

(a) Let $V(G) := [n]$.

(b) Add edges $e \in \binom{[n]}{2}$ with $\mu(e) = 1$ to $E(G)$.

(c) Recall that $\mu^{-1}(\star) = \binom{U}{2}$, where $U \in \binom{[n]}{\ell_{\bar{d}}(n)}$ was chosen uniformly at random. For every $e \in \binom{U}{2}$, add e to $E(G)$ with probability $q(n)$.

(d) Choose $A \in \binom{U}{c(n)}$ uniformly at random. Note that this is possible as $|U| = \ell_{\bar{d}}(n) = n^{1-o(1)} > n^{\xi} \geq c(n)$ for sufficiently large n.

By (b)–(d), G and $G + C(A)$ contain the same edges from $\binom{[n]}{2} \setminus \mu^{-1}(\star)$. Thus, by (11), $\mathsf{C}_n(G) = \mathsf{C}_n(G + C(A))$ with high probability. By (c) and (d), A can be viewed as being chosen in $\binom{[n]}{c(n)}$ uniformly at random. □

Reduction to small edge probability. We fix the size $c(n) \leq n^{\xi}$ for the planted clique in Theorem 4.6. Assume $k, k' : \mathbb{N} \to \mathbb{R}^+$ and $k'(n) \geq k(n)$ for all $n \in \mathbb{N}$. We set

$$p(n) := \frac{n^{-1/k'(n)} - n^{-1/k(n)}}{1 - n^{-1/k(n)}}.$$

Then $0 \leq p(n) < 1$. It is easy to see that for $H \in \mathrm{ER}(n, p(n))$ and $G \in \mathrm{ER}(n, n^{-1/k(n)})$ the graph $H \cup G$ has the distribution $\mathrm{ER}(n, n^{-1/k'(n)})$. Here $H \cup G = \big([n], E(H) \cup E(G)\big)$.

Now let $(\mathsf{C}_n)_{n\in\mathbb{N}}$ be any sequence of circuits of depth d and circuit size n^t. For every $H \in \mathrm{ER}(n, p(n))$ one can define a circuit C_n^H of depth $d + 1$ and size $n^t + n^2$ such that for all graphs G with vertex set $[n]$,

$$\mathsf{C}_n^H(G) := \mathsf{C}_n(H \cup G).$$

Therefore, we have

$$\begin{aligned}
&\Pr_{(H',A)\in\mathrm{ER}(n,n^{-1/k'(n)},\, c(n))}\big[\mathsf{C}_n(H') = \mathsf{C}_n(H' + C(A))\big] \\
&\quad= \Pr_{\substack{H\in\mathrm{ER}(n,p(n)),\\ (G,A)\in\mathrm{ER}(n,n^{-1/k(n)},\, c(n))}} \big[\mathsf{C}_n(H \cup G) = \mathsf{C}_n((H \cup G) + C(A))\big] \\
&\quad= \sum_{H_0\in\mathbf{G}(n)} \Pr_{(G,A)\in\mathrm{ER}(n,n^{-1/k(n)},\, c(n))} \big[\mathsf{C}_n^{H_0}(G) = \mathsf{C}_n^{H_0}(G + C(A))\big] \cdot \Pr_{H\in\mathrm{ER}(n,p(n))}[H = H_0],
\end{aligned}$$

where $\mathbf{G}(n)$ denotes the set of graphs with vertex set $[n]$. So from the equality between the first and last term, we see the following.

Proposition 4.8. *Let $c : \mathbb{N} \to \mathbb{N}$ with $c(n) \leq n^{\xi}$ for some $0 \leq \xi < 1$ and let $k, k' : \mathbb{N} \to \mathbb{R}^+$ with $k'(n) \geq k(n)$ for all $n \in \mathbb{N}$. If for every AC^0 circuits $(\mathsf{C}_n)_{n\in\mathbb{N}}$*

$$\lim_{n\to\infty} \Pr_{(G,A)\in\mathrm{ER}(n,n^{-1/k(n)},\, c(n))} \big[\mathsf{C}_n(G) = \mathsf{C}_n(G + C(A))\big] = 1$$

and the convergence rate is uniform for all AC^0 circuits of a fixed depth and size, then

$$\lim_{n\to\infty} \Pr_{(G,A)\in\mathrm{ER}(n,n^{-1/k'(n)},\, c(n))} \big[\mathsf{C}_n(G) = \mathsf{C}_n(G + C(A))\big] = 1.$$

Now Theorem 4.6 follows immediately from Lemma 4.7 and Proposition 4.8.

Remark 4.9. For a random graph $G \in \mathrm{ER}(n, n^{-1/k(n)})$, the expected size of a maximum clique is $O(k(n))$. Thus if $k(n) = o(c(n))$, to distinguish G and $G+C(A)$ for $(G, A) \in \mathrm{ER}(n, n^{-1/k(n)}, c(n))$, some constant-depth circuits of size

$$n^{O(k(n))}$$

suffice. By a careful inspection of the proof of Theorem 4.6, in particular, the equation (12) in Lemma 4.7, it is easy to see that any constant-depth circuits of size

$$n^{o(k(n))}$$

cannot distinguish G and $G + C(A)$.

Furthermore, if the depth of polynomial-size circuits is

$$o\left(\sqrt{k(n)}\right),$$

then (12) still holds. Hence, polynomial-size circuits of depth $o(\sqrt{\log n})$ cannot distinguish G and $G+C(A)$ for $(G, A) \in \mathrm{ER}\big(n, 1/2, O(\sqrt{n})\big)$.[4] These arguments are based on (12), an equality holding under the hypothesis $k(n) \leq \log_2 n$. Again the general case is reduced by the standard techniques from probability theory used to prove Proposition 4.8.

4.3. Proof of Theorem 4.3. Let $\big(\mathsf{C}_{n,k}\big)_{n,k\in\mathbb{N}}$ be a family of circuits such that for some function $f : \mathbb{N} \to \mathbb{N}$ and $d, c \in \mathbb{N}$ every $\mathsf{C}_{n,k}$ has depth at most d and size bounded by $f(k) \cdot n^c$. Then we show that there are some $n, k \in \mathbb{N}$ such that $\mathsf{C}_{n,k}(G)$ does not decide p-GAP_ρ-CLIQUE on instances (G, k) with $n := |V(G)|$. Hence, our proof even works for a nonuniform version of para-AC^0: We neither assume that the family $(\mathsf{C}_{n,k})$ is computable from n and k nor that f is computable.

We may assume that f is nondecreasing and unbounded. We choose a nondecreasing and unbounded function $k : \mathbb{N} \to \mathbb{N}$ such that for sufficiently large $n \in \mathbb{N}$ we have

$$2k(n) + 1 \leq \min\left\{f^{-1}(n), \frac{\sqrt{n}}{\rho(\sqrt{n})}\right\}, \tag{13}$$

where $f^{-1}(n) := \max\big(\{\ell \mid f(\ell) \leq n\} \cup \{0\}\big)$, and such that $k(n) \leq \log_2 n$ for $n \geq 1$. It follows that the circuit

$$\mathsf{C} := \mathsf{C}_{n,2k(n)+1}$$

has size bounded by $S(n) := O(n^{c+1})$, i.e., $\big\{\mathsf{C}_{n,2k(n)+1}\big\}_{n\in\mathbb{N}}$ are AC^0-circuits.

We consider the distribution $(G, A) \in \mathrm{ER}(n, n^{-1/k(n)}, \lceil\sqrt{n}\,\rceil)$. The next claim is easy to verify.

Claim 1. $G + C(A)$ contains a clique of size $\lceil\sqrt{n}\rceil$, i.e., $\omega(G + C(A)) \geq \sqrt{n}$. On the other hand, $\Pr\big[\omega(G) < 2k(n) + 1\big] = 1 - o(1)$.

Assume n is sufficiently large, and recall $m \mapsto m/\rho(m)$ is increasing, so (13) implies

$$2k(n) + 1 \leq \frac{\omega(G + C(A))}{\rho(\omega(G + C(A)))}.$$

[4] For the distribution $(G, A) \in \mathrm{ER}\big(n, 1/2, \Theta(\sqrt{n})\big)$ there are polynomial time algorithms [2] (thus also polynomial-size circuits) which can detect the planted clique $C(A)$ in $G + C(A)$, hence distinguish G and $G + C(A)$.

This means that $\big(G + C(A), 2k(n)+1\big)$ is a yes instance of p-GAP$_\rho$-CLIQUE, while *almost surely* $\big(G, 2k(n)+1\big)$ is a no instance. Hence, by our assumption on $\big(\mathsf{C}_{n,k}\big)_{n,k\in\mathbb{N}}$ and thus on C,

$$\Pr_{(G,A)\in \mathrm{ER}(n,n^{-1/k(n)},\lceil\sqrt{n}\rceil)}\big[\mathsf{C}(G+C(A))=1\big]=1,\quad \Pr_{(G,A)\in \mathrm{ER}(n,n^{-1/k(n)},\lceil\sqrt{n}\rceil)}\big[\mathsf{C}(G)=1\big]=o(1).$$

But this contradicts Theorem 4.6. □

We prove a consequence of Theorem 4.3. For $t \geq 0$, $d \geq 1$ we denote by $\Gamma^-_{t,d}$ the subset of subformulas of $\Gamma_{t,d}$ with only negative literals. Clearly, if $\gamma \in \Gamma^-_{t,d}$ has a satisfying assignment of Hamming weight k, then it has one of weight k' for every $k' < k$. Denote by $\omega(\gamma)$ the maximum Hamming weight of assignments satisfying γ. Then p-GAP$_\rho$-WSAT$(\Gamma^-_{t,d})$ can be defined similarly as p-GAP$_\rho$-CLIQUE.

Proposition 4.10. *Let $t, d \geq 1$ with $t + d \geq 3$. Then* p-GAP$_\rho$-WSAT$(\Gamma^-_{t,d}) \notin$ para-AC^0.

Proof : Consider the reduction from p-CLIQUE to p-GAP$_\rho$-WSAT$(\Gamma_{t,d})$ in the proof of Corollary 3.10. Clearly $\delta_G \in \Gamma^-_{t,d}$ and δ_G is independent of k. Thus, the equivalence (3) preserves the approximation ratio. The result then follows immediately. □

5. The complexity of p-HALT

We already mentioned in the abstract of this article that the complexity of the parameterized halting problem p-HALT is linked to open problems in computational complexity, descriptive complexity, and proof theory [9]. For example, the membership of p-HALT in the parameterized complexity class *uniform XP* is equivalent to the existence of an almost optimal algorithm for the set of tautologies of propositional logic, or to the fact that a certain logic, presented in [17], is a logic for PTIME. Both statements are conjectured to be false. The origin of our interest in para-AC^0 was our hope to get a lower bound on the complexity of p-HALT in terms of para-AC^0, that is, to show p-HALT $\notin$ para-AC^0. But also this problem remains open. We know that AC^0 corresponds to FO$(<,+,\times)$, first-order logic with an ordering relation and built-in addition and multiplication. In this section we prove that p-HALT $\notin$ para-FO$(<,+)$, even p-HALT $\notin$ XFO$(<,+)$, hold unconditionally, to our knowledge the best known lower bound for the complexity of p-HALT.

Recall that in the paragraph following Lemma 3.11 we defined the natural ordering $<^{[n]}$ on $[n]$ and the ternary relations $+^{[n]}$ and $\times^{[n]}$ of addition and multiplication, respectively, on $[n]$. Now we address the definition of XFO$(<,+,\times)$. For this purpose we view inputs to parameterized problems as structures.

Any string $x \in \Sigma^*$ with $|x| = n$ can be identified with the $\{<,+,\times,\mathit{One}\}$-structure $\langle x\rangle^{<,+,\times} := ([n], <^{[n]}, +^{[n]}, \times^{[n]}, \mathit{One}^{[n]})$. Here $i \in [n]$ is in $\mathit{One}^{[n]}$, the interpretation of the unary relation symbol *One*, if and only if the ith bit of x is a '1'. The structures $\langle x\rangle^{<,+}$ and $\langle x\rangle^{<}$ are reducts of $\langle x\rangle^{<,+,\times}$ over the vocabularies $\{<,+,\mathit{One}\}$ and $\{<,\mathit{One}\}$, respectively.

Definition 5.1. Let (Q,κ) be a parameterized problem. Then $(Q,\kappa) \in$ XFO$(<,+,\times)$ if there is a computable function that assigns to every $k \in \mathbb{N}$ a first-order sentence φ_k such that for every instance x of (Q,κ),

$$x \in Q \iff \langle x\rangle^{<,+,\times} \models \varphi_{\kappa(x)}.$$

Analogously, the class XFO$(<,+)$ is defined.

Theorem 5.2. p-HALT $\notin$ XFO$(<,+)$.

Proof : For a contradiction we assume that p-HALT $\in$ XFO$(<,+)$ and show that then the halting problem for Turing machines would be decidable.

Assume that there is a computable function that assigns to every $k \in \mathbb{N}$ a first-order sentence φ_k such that for every instance $(1^n, \mathbb{M})$,

$$(1^n, \mathbb{M}) \in p\text{-HALT} \iff \langle (1^n, \mathbb{M}) \rangle^{<,+} \models \varphi_{|\mathbb{M}|}.$$

Fix $\mathbb{M}$. There is a first-order interpretation I that for every $n \in \mathbb{N}$ defines an isomorphic copy of $\langle (1^n, \mathbb{M}) \rangle^{<,+}$ in $([n], <^{[n]}, +^{[n]})$: Let $c(n) := |(1^n, \mathbb{M})|$ be the length of the string $(1^n, \mathbb{M})$. We define the interpretation I stepwise. We choose s such that $c(n) \leq n^s$. As $\mathbb{M}$ is fixed, in $([n], <^{[n]}, +^{[n]})$ we can define the initial segment of $[n]^s$ of $c(n)$ elements, define on it the lexicographical ordering and the relation *One* such that we get a copy of $\langle (1^n, \mathbb{M}) \rangle^{<}$ on this initial segment. Finally, using $+^{[n]}$, we can define the corresponding built-in addition.

Then, from $\mathbb{M}$ we can compute $\varphi_{|\mathbb{M}|}$ and $\varphi^I_{|\mathbb{M}|}$ such that

$$(1^n, \mathbb{M})^{<,+} \models \varphi_{|\mathbb{M}|} \iff ([n], <^{[n]}, +^{[n]}) \models \varphi^I_{|\mathbb{M}|},$$

and thus,

$$(1^n, \mathbb{M}) \in p\text{-HALT} \iff \big([n], <^{[n]}, +^{[n]}\big) \models \varphi^I_{|\mathbb{M}|}. \tag{14}$$

By the Ginsburg-Spanier [15] improvement of Presburger's Theorem we know that for $\varphi^I_{|\mathbb{M}|}$ we may compute $n_0, p_0 \in \mathbb{N}$ such that for all $n \geq n_0$ we have

$$\big([n], <^{[n]}, +^{[n]}\big) \models \varphi^I_{\mathbb{M}} \iff \big([n + p_0], <^{[n+p_0]}, +^{[n+p_0]}\big) \models \varphi^I_{\mathbb{M}}.$$

By this equivalence and (14) we see that

$$\mathbb{M} \text{ does not hold on empty input tape} \iff \big([n_0], <^{[n_0]}, +^{[n_0]}\big) \models \neg\varphi^I_{\mathbb{M}}.$$

We can decide the halting problem by checking whether $\big([n_0], <^{[n_0]}, +^{[n_0]}\big) \models \neg\varphi^I_{\mathbb{M}}$. □

For the proof it was essential that the function assigning to every $k \in \mathbb{N}$ the FO$(<,+)$-sentence φ_k is *computable*. The class obtained if we drop the requirement of computability in the definition of XFO$(<,+)$ is called nonuniform-XFO$(<,+)$. We will see that p-HALT $\in$ nonuniform-XFO$(<,+)$ by the even stronger statement of Proposition 5.3 (1).

We note in passing that by standard modeltheoretic techniques one can show that the parameterized vertex cover problem, a standard example of a fixed-parameter tractable problem, is not in the subclass nonuniform-XFO$(<)$ of nonuniform-XFO$(<,+)$. Thus we get a lower bound for the parameterized complexity of this problem.

We come back to our claim p-HALT $\in$ nonuniform-XFO$(<,+)$. We even show p-HALT $\in$ nonuniform-para-FO$(<,+)$. By definition, a parameterized problem (Q, κ) belongs to the class nonuniform-para-FO$(<,+)$ (to para-FO$(<,+)$) if there are a sentence $\varphi \in$ FO$(<,+)$ and a (computable) function *pre* : $\mathbb{N} \to \Sigma^*$ such that for all x,

$$x \in Q \iff \langle (x, \mathit{pre}(\kappa(x))) \rangle^{<,+} \models \varphi.$$

So, in the nonuniform version we allow noncomputable precomputations. Note that para-FO$(<,+) \subseteq$ XFO$(<,+)$ as the role of the precomputation $\big($in the definition of para-FO$(<,+)\big)$ can be taken over by the sentences φ_k $\big($in the definition of XFO$(<,+)\big)$.

Proposition 5.3. *(1)* p-HALT $\in$ nonuniform-para-FO$(<,+)$.

(2) p-HALT $\notin$ nonuniform-para-FO$(<)$.

Proof : (1) We look for a first-order sentence φ and a function *pre* : $\mathbb{N} \to \Sigma^*$ such that for all instances $(1^n, \mathbb{M})$ of p-HALT,

$$(1^n, \mathbb{M}) \in p\text{-HALT} \iff \langle((1^n, \mathbb{M}), \mathit{pre}(|\mathbb{M}|)\rangle^{<,+} \models \varphi. \tag{15}$$

For a nondeterministic Turing machine $\mathbb{M}$ define $n_{\mathbb{M}}$ as the numbers of steps of a shortest run of $\mathbb{M}$ on empty input; set $n_{\mathbb{M}} := \infty$ if every run is infinite. We turn to the definition of the (noncomputable) precomputation *pre* : $\mathbb{N} \to \Sigma^*$. For $k \in \mathbb{N}$ we enumerate all nondeterministic Turing machines of length k as

$$\mathbb{M}_1, \mathbb{M}_2, \ldots, \mathbb{M}_m$$

and set

$$\mathit{pre}(k) := \$(1^{n_{\mathbb{M}_1}}, \mathbb{M}_1)\$ \ldots \$(1^{n_{\mathbb{M}_m}}, \mathbb{M}_m)\$$$

(in a standard way we view $\mathit{pre}(k)$ as a string in Σ^*). The first-order sentence φ satisfying (15) expresses that $\mathbb{M}$ is one of the machines $\mathbb{M}_i$ and that $n_i \leq n$. Note that using addition we can express in first order logic that the substrings between the points u_1 and u_2 and between the points v_1 and v_2 have the same length (by $u_2 - u_1 = v_2 - v_1$) or distinct length, and we also can express that they coincide.

(2) For a contradiction assume that there is an FO$(<)$-sentence φ and a function *pre* : $\mathbb{N} \to \Sigma^*$ such that for all instances $(1^n, \mathbb{M})$ of p-HALT,

$$(1^n, \mathbb{M}) \in p\text{-HALT} \iff \big\langle(1^n, \mathbb{M}), \mathit{pre}(|\mathbb{M}|)\big\rangle^{<} \models \varphi.$$

We fix a Turing machine $\mathbb{M}$. Then one easily sees that there is a first-order interpretation $I = I_{\mathbb{M}}$ that for all $n \in \mathbb{N}$ defines a structure

$$\big\langle(1^n, \mathbb{M}), \mathit{pre}(|\mathbb{M}|)\big\rangle^{<}$$

from the structure $([n], <^{[n]})$. Hence, for all $n \in \mathbb{N}$,

$$\big\langle(1^n, \mathbb{M}), \mathit{pre}(|\mathbb{M}|)\big\rangle^{<} \models \varphi \iff ([n], <^{[n]}) \models \varphi^I.$$

Let q be the quantifier rank of φ^I. Note that φ^I is computable from φ and $\mathbb{M}$. Then we know that for $n, n' > 2^q$,

$$([n], <^{[n]}) \models \varphi^I \iff ([n'], <^{[n']}) \models \varphi^I.$$

So again we can decide the halting problem.

□

Let τ be a vocabulary which does not contain the reation symbols $<, +, \times$ and set $\tau_{<,+,\times} := \tau \cup \{<, +, \times\}$. Recall that a $\tau_{<,+,\times}$-structure $\mathcal{A}$ *has built-in addition and built-in multiplication* if $(A, <^{\mathcal{A}}, +^{\mathcal{A}}, \times^{\mathcal{A}})$ is isomorphic to $([|A|], <^{[|A|]}, +^{[|A|]}, \times^{[|A|]})$.

A first-order sentence φ of vocabulary $\tau_{<,+,\times}$, shortly $\varphi \in$ FO$(<,+,\times)$, is *invariant* (more precisely, $<$*-invariant*) if for every τ-structure $\mathcal{A}$ and any expansions $(\mathcal{A}, <_1, +_1, \times_1)$ and $(\mathcal{A}, <_2, +_2, \times_2)$ of $\mathcal{A}$ to structures with built-in addition and multiplication, we have:

$$(\mathcal{A}, <_1, +_1, \times_1) \models \varphi \iff (\mathcal{A}, <_2, +_2, \times_2) \models \varphi.$$

It should be clear what we mean if we say that a $\varphi \in$ FO$(<,+)$ or a $\varphi \in$ FO$(<)$ is *invariant*.

Along the lines of [8, Theorem 10] one can show:

Proposition 5.4. *Assume that* p-HALT $\in$ XFO$(<,+,\times)$. *Let* τ *be any vocabulary not containing the symbols* $<$, $+$, *and* $\times$. *Then there is a computable function* F *defined on the class of* FO$(<,+,\times)$*-sentences of vocabulary* $\tau \cup \{<,+,\times\}$ *such that*

- *for every* $\varphi \in$ FO$(<,+,\times)$ *the sentence* $F(\varphi)$ *is invariant;*
- *if* φ *is an invariant* FO$(<,+,\times)$*-sentence, then* φ *and* $F(\varphi)$ *are equivalent.*

Thus, $\{F(\varphi) \mid \varphi$ *an invariant* FO$(<,+,\times)\}$ *is the class of sentences of vocabulary* τ *of a logic for the invariant fragment of* FO$(<,+,\times)$.

In view of Theorem 5.2, we tried, without success, to show that for FO$(<,+)$ there is no computable function F with the properties mentioned in the preceding result for FO$(<,+,\times)$, or even to show that there is no effective enumeration of the invariant sentences of FO$(<,+,\times)$.

References

[1] M. Ajtai. Σ^1_1 formulae on finite structures. *Annals of Pure and Applied Logic*, 24(3):1–48, 1983.

[2] N. Alon, M. Krivelevich, and B. Sudakov. Finding a large hidden clique in a random graph. *Random Struct. Algorithms*, 13(3-4):457–466, 1998.

[3] M. Bannach, C. Stockhusen, and T. Tantau. Fast parallel fixed-parameter algorithms via color coding. In *10th International Symposium on Parameterized and Exact Computation, IPEC 2015, September 16-18, 2015, Patras, Greece*, pages 224–235, 2015.

[4] D. A. Mix Barrington, N. Immerman, and H. Straubing. On uniformity within NC^1. *Journal of Computer and System Sciences*, 41(3):274–306, 1990.

[5] P. Beame. *A Switching Lemma Primer*. Technical Report, University of Washington, 1984.

[6] P. Beame, R. Impagliazzo, and T. Pitassi. Improved depth lower bounds for small distance connectivity. *Computational Complexity*, 7(4):325–345, 1998.

[7] J. Chen, X. Huang, I. A. Kanj, and G. Xia. Linear FPT reductions and computational lower bounds. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing, Chicago, STOC 2004, IL, USA, June 13-16, 2004*, pages 212–221, 2004.

[8] Y. Chen and J. Flum. A logic for PTIME and a parameterized halting problem. In *Proceedings of the 24th Annual IEEE Symposium on Logic in Computer Science, LICS 2009, 11-14 August 2009, Los Angeles, CA, USA*, pages 397–406, 2009.

[9] Y. Chen and J. Flum. From almost optimal algorithms to logics for complexity classes via listings and a halting problem. *Journal of the ACM*, 59(4):17, 2012.

[10] Y. Chen, M. Grohe, and M. Grüber. On parameterized approximability. *Electronic Colloquium on Computational Complexity (ECCC)*, 14(106), 2007.

[11] M. Elberfeld, C. Stockhusen, and T. Tantau. On the space and circuit complexity of parameterized problems: Classes and completeness. *Algorithmica*, 71(3):661–701, 2015.

[12] J. Flum and M. Grohe. Describing parameterized complexity classes. *Information and Computation*, 187(2):291–319, 2003.

[13] J. Flum and M. Grohe. *Parameterized Complexity Theory*. Springer, 2006.

[14] M. L. Furst, J. B. Saxe, and M. Sipser. Parity, circuits, and the polynomial-time hierarchy. *Mathematical Systems Theory*, 17(1):13–27, 1984.

[15] S. Ginsburg and E.H. Spanier. Semigroups, Presburger fomulas, and languages. *Pacific Journal of Mathematics*, 16:285–296, 1966.

[16] M. Grohe, S. Kreutzer, and S. Siebertz. Deciding first-order properties of nowhere dense graphs. In *Proceedings of the 46th Annual ACM Symposium on the Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 89–98, 2014.

[17] Y. Gurevich. Logic and the challenge of computer science. In *Current trends in Theoetical computer Science*, Computer Science Press, pages 1–57, 1988.

[18] M. Jerrum. Large cliques elude the metropolis process. *Random Structures and Algorithms*, 3(4):347–360, 1992.

[19] L. Kučera. Expected complexity of graph partitioning problems. *Discrete Applied Mathematics*, 57(2-3):193–212, 1995.

[20] A. Nash, J. B. Remmel, and V. Vianu. PTIME queries revisited. In *Database Theory - ICDT 2005, 10th International Conference, Edinburgh, UK, January 5-7, 2005, Proceedings*, volume 3363 of *Lecture Notes in Computer Science*, pages 274–288. Springer, 2005.

[21] B. Rossman. On the constant-depth complexity of k-clique. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing, STOC 2008, Victoria, British Columbia, Canada*, pages 721–730, 2008.