

Polynomial time ultrapowers and the consistency of circuit lower bounds*

Jan Bydžovský

Institute of Discrete Mathematics and Geometry
Technische Universität Wien
Vienna, Austria
jan.bydz@gmail.com

Moritz Müller

Department of Computer Science
Universitat Politècnica de Catalunya
Barcelona, Spain
moritz@cs.upc.edu

Abstract

A polynomial time ultrapower is a structure given by the set of polynomial time computable functions modulo some ultrafilter. They model the universal theory $\forall\text{PV}$ of all polynomial time functions. Generalizing a theorem of Hirschfeld (1975), we show that every countable model of $\forall\text{PV}$ is isomorphic to an existentially closed substructure of a polynomial time ultrapower. Moreover, one can take a substructure of a special form, namely a *limit* polynomial time ultrapower in the classical sense of Keisler (1963). Using a polynomial time ultrapower over a nonstandard Herbrand saturated model of $\forall\text{PV}$ we show that $\forall\text{PV}$ is consistent with a formal statement of a polynomial size circuit lower bound for a polynomial time computable function. This improves upon a recent result of Krajíček and Oliveira (2017).

1 Introduction

In [5], according to Pudlák “the founding paper of the field of proof complexity” [32, p.540], Cook introduced the theory PV as a theory formalizing the intuitive concept of feasible provability. The language of PV , also denoted PV , has symbols for all polynomial time functions. While Cook defined PV as an equational theory, a variant as a universal first-order theory has been given in [25]. Roughly, it is given by equations following Cobham’s characterization of polynomial time [7] and some form of quantifier-free induction. In this paper we work with the larger theory $\forall\text{PV}$, the theory of all universal sentences true in the standard PV -model \mathbb{N} .¹

In Krajíček and Oliveira’s words, “ PV or its mild extensions seem to formalize most of contemporary complexity theory [...] It is thus of interest to understand, given an

*Based on the first author’s Master Thesis [4] written under the supervision of the second author.

¹All relevant technical notions will be defined precisely later.

established conjecture, whether it is provable in one of these theories or at least consistent with them.” [24, p.1] An example of particular interest is circuit lower bounds. Razborov [35] argued that PV can formalize existing lower bounds for restricted circuit models, so showing unprovability in PV of general lower bounds would somehow explain the difficulty of obtaining such results. Assuming the existence of strong pseudorandom generators, Razborov proved unprovability in the theory $\text{S}_2^2(\alpha)$ in [36] based on the natural proof barrier [37]; simpler proofs using feasible interpolation have been given in [34, 3, 20]. Razborov used a peculiar formalization of circuit lower bounds as Π_1^b -statements. A more usual formalization that only allows to formalize polynomial size lower bounds and has higher quantifier complexity has been proposed by Krajíček in his 1995 monograph [19, Section 15.2] and it is this “succinct” [30] formalization that we use in this paper. Some discussion can be found in [30] where existing lower bounds for restricted circuits are formalized in a mild extension of PV , namely Jeřábek’s theory APC_1 of approximate counting [13].

The first and final words of [19] motivate the task to show the consistency, as opposed to unprovability, of complexity theoretic conjectures with bounded arithmetics [19, p.xii, p.326]. Recently, it has been shown [24] that general circuit lower bounds are consistent with $\forall\text{PV}$ (see [6] for earlier conditional results under certain complexity theoretic assumptions). More precisely, for a Boolean (i.e., 0/1-valued) function $g(x) \in \text{PV}$ let $LB[g](s, n)$ be a PV -formula with variables s, n that expresses

For all circuits C of size at most s there is x of length $n > 0$ such that $g(x) \neq C(x)$.

Krajíček and Oliveira [24] proved that for every $k \in \mathbb{N}$ there is a Boolean $g(x) \in \text{PV}$ such that for every $c \in \mathbb{N}$ the sentence $\exists n LB[g](c \cdot n^k, n)$ is consistent with $\forall\text{PV}$. We give a new proof that yields the following seemingly stronger² result (cf. [24, Remark 2.2]):

Theorem 1.1. *For every $k \in \mathbb{N}$ there is a Boolean $g(x) \in \text{PV}$ such that $\forall z \exists n LB[g](|z| \cdot n^k, n)$ is consistent with $\forall\text{PV}$.*

On a high level, the idea of the proof is to infer the consistency of a non-uniform lower bound from the truth of a uniform lower bound. The true lower bound in question has been recently established by Santhanam and Williams [38, Theorem 1.1]:

Theorem 1.2 (Santhanam, Williams 2014). *For every $k \in \mathbb{N}$ there is a Boolean $g(x) \in \text{PV}$ which is not computable by PTIME -uniform size $O(n^k)$ circuit families.*

Now, if a weak theory would prove $\neg LB[g](c \cdot n^k, n)$, then it should be possible (by witnessing) to extract from the proof a polynomial time function mapping n (in unary) to circuit C_n of size $c \cdot n^k$ computing $g(x)$ on inputs of length n and thereby contradict Theorem 1.2. However, for $\forall\text{PV}$ such witnessing is known only for Σ_1^b -formulas while $\neg LB[g](c \cdot n^k, n)$ is Σ_2^b . For Σ_2^b the KPT-theorem [25] gives a witnessing function computable by a polynomial time

²For $\ell, k \in \mathbb{N}$, the theory $\forall\text{PV} \cup \{\exists n LB[g](c \cdot n^\ell, n) \mid c \in \mathbb{N}\}$ does not seem to imply $\forall z \exists n LB[g](|z| \cdot n^k, n)$. For all we know, there could be a nonstandard model of the former theory that contains only standard n witnessing the lower bounds, while every nonstandard model of $\forall z \exists n LB[g](|z| \cdot n^k, n)$ witnesses the lower bound with some nonstandard n .

student interacting with an omnipotent teacher. Krajíček and Oliveira’s proof [24] is based on the KPT-theorem. For Theorem 1.1 a new argument is required since the negation $\neg\forall z\exists n LB[g](|z| \cdot n^k, n)$ of our sentence is Σ_4^b . Our proof is model-theoretic and based on a polynomial time ultrapower over a nonstandard model of $\forall PV$ that is Herbrand saturated in the sense of [1].

For now restrict attention to the standard model \mathbb{N} interpreting PV . Note one can apply polynomial time functions (from \mathbb{N} to \mathbb{N}) to others by means of composition, so the set of polynomial time functions naturally interprets PV . Given an ultrafilter U on \mathbb{N} one forms a *polynomial time ultrapower* (over \mathbb{N}) by identifying two functions that agree on some set in U . It is easy to see that polynomial ultrapowers satisfy $\forall PV$.

Of course, one can form *restricted ultrapowers* F/U for other function families F and such constructions have been frequently used to study arithmetic and its fragments [39, 27, 18, 12, 29, 33, 21, 10] ever since Skolem’s *definable ultrapower* [41]. We give a brief historical survey in Section 2.2. For example, Hirschfeld studies *recursive ultrapowers* built from the set of computable functions and shows that they “are the basic models from which all models of” $Th_{\Pi_2}(\mathbb{N})$, the Π_2 -fragment of true arithmetic, “are composed” [12, p.112]:

Theorem 1.3 (Hirschfeld 1975). *A countable $\{+, \cdot, 0, 1, <\}$ -structure is a model of $Th_{\Pi_2}(\mathbb{N})$ if and only if it is isomorphic to an existentially closed substructure of some recursive ultrapower; moreover, this substructure can be taken to be a limit recursive ultrapower.*

The second claim refers to Keisler’s classical notion of limit ultrapowers [14].

Polynomial time ultrapowers deserve some interest and we develop their theory to some extent beyond what is needed to prove Theorem 1.1. We prove:

Theorem 1.4. *A countable PV -structure is a model of $\forall PV$ if and only if it is isomorphic to an existentially closed substructure of some polynomial time ultrapower; moreover, this substructure can be taken to be a limit polynomial time ultrapower.*

Corollary 1.5. *Every countable model of $\forall PV$ is isomorphic to a restricted ultrapower F/U for a family F of polynomial time computable functions and an ultrafilter U on \mathbb{N} .*

We shall prove a more general Characterization Theorem 2.26 that implies both Theorems 1.3 and 1.4, and might be of some independent interest. The hope is that this description of the models of $\forall PV$ helps understanding these models and thereby eventually understanding the status of complexity theoretic conjectures in $\forall PV$ – and its mild extensions: as in [24, p.2] we ask whether Theorem 1.1 holds for APC_1 .

The paper is organized as follows. Section 2 gives some general theory of restricted ultrapowers. Section 2.1 collects some of their basic properties in full generality, and Section 2.2 gives examples and a historical survey. Section 2.3 defines $\forall PV$ and polynomial time ultrapowers. The Characterization Theorem 2.26 is proved in Section 2.4.

Section 3 proves Theorem 1.1 in its final Section 3.4. To emphasize the simplicity of our proof we explain the idea on a high level in Section 3.1. Section 3.2 defines the formulas $LB[g](s, n)$ and Section 3.3 formalizes Santhanam and Williams’ proof of Theorem 1.2.

2 Restricted ultrapowers

2.1 Basics

Fix a language L and an L -structure M . We do not distinguish M from its universe notationally, and denote the interpretation of a symbol $s \in L$ in M by s^M . We view constants as 0-ary function symbols. Writing $\varphi(\bar{x})$ or $t(\bar{x})$ for a formula or a term means its free variables are among those in the tuple \bar{x} . In such a context, and if \bar{x} has r variables, we write $\varphi(M)$ for the r -ary relation defined by $\varphi(\bar{x})$ in M , and t^M for the r -ary function given by the interpretation of $t(\bar{x})$ in M .

For a set X , silently assumed to be disjoint from L , we write $L(X)$ for $L \cup X$ and view elements from X as constants. We interpret $L(M)$ -formulas and $L(M)$ -terms in M understanding that each *parameter* $a \in M$ is interpreted by itself. We call a relation definable in M if it is definable in M *with parameters*, i.e, by an $L(M)$ -formula. An (r -ary) function is definable in M if so is its graph (viewed as a $(r + 1)$ -ary relation).

Let Ω be a nonempty set and $F \subseteq M^\Omega$ a set of functions from Ω to M . We let α, β, \dots range over F and ω over Ω . We say F is *closed under* W for an r -ary function $W : M^r \rightarrow M$ if for every $\bar{\alpha} = (\alpha_0, \dots, \alpha_{r-1}) \in F^r$ also the function $W \circ \bar{\alpha}$ that maps ω to $W(\bar{\alpha}(\omega))$ is in F . Here, $\bar{\alpha}(\omega)$ denotes $(\alpha_0(\omega), \dots, \alpha_{r-1}(\omega)) \in M^r$.

Definition 2.1. A set of functions $F \subseteq M^\Omega$ is L -closed if it is non-empty and closed under f^M for every function symbol $f \in L$.

The following canonical examples are going to play a central role.

Example 2.2. For $\Omega := M$, the smallest L -closed F that contains id_Ω , the identity function on Ω , is

$$T_L^M := \{t^M \mid t(x) \text{ is an } L\text{-term}\}.$$

Similarly,

$$T_{L(M)}^M := \{t^M \mid t(x) \text{ is an } L(M)\text{-term}\}.$$

is the smallest L -closed F that contains id_Ω and for every $a \in M$ the constant function

$$\alpha_a : \Omega \rightarrow M : \omega \mapsto \alpha_a(\omega) = a.$$

Restricted ultrapowers are associated with L -closed function families and ultrafilters on Ω . We recall some terminology: a (*proper*) *filter* U on Ω is a nonempty collection of nonempty subsets of Ω which is closed under taking intersections and supersets. An *ultrafilter* is a maximal filter. A collection \mathcal{X} has the *finite intersection property* if intersections of finitely many members of \mathcal{X} are nonempty; it then *generates* the filter consisting of the supersets of these intersections. Recall further that every filter is contained in an ultrafilter.

For an ultrafilter U on Ω we let F/U denote the set of equivalence classes α^U for $\alpha \in F$ with respect to \sim^U where $\alpha \sim^U \beta$ if and only if $\{\omega \mid \alpha(\omega) = \beta(\omega)\} \in U$. For $\bar{\alpha} = (\alpha_0, \dots, \alpha_{r-1}) \in F^r$ we let $\bar{\alpha}^U$ denote $(\alpha_0^U, \dots, \alpha_{r-1}^U)$.

Definition 2.3. Let $F \subseteq M^\Omega$ be L -closed and U an ultrafilter on Ω . The *restricted ultrapower* F/U is the following L -structure with universe the set of equivalence classes F/U . It interprets r -ary relation and function symbols R and f from L by, respectively,

- the set of $\bar{\alpha}^U \in (F/U)^r$ such that $\{\omega \mid \bar{\alpha}(\omega) \in R^M\} \in U$, and
- the function that maps $\bar{\alpha}^U \in (F/U)^r$ to $(f^M \circ \bar{\alpha})^U \in F/U$.

It is easily checked that F/U is well-defined more generally for every filter U ; however, we shall consider only ultrafilters.

Keisler’s definition [15] of limit ultrapowers works verbatim also for restricted ultrapowers; the terminology stems from [14, Theorem 3]. Recall, a substructure A of a structure B is *existentially closed (in B)* if every universal $L(A)$ -sentence true in A is true in B ; equivalently, if every quantifier-free $L(A)$ -formula, which is satisfiable in B , is satisfiable in A .

Definition 2.4. Let $F \subseteq M^\Omega$ be L -closed, U an ultrafilter on Ω and E a filter on Ω^2 . The *limit restricted ultrapower* $F/U/E$ is the substructure of F/U whose universe is the set of those \sim^U -equivalence classes that contain some $\alpha \in F$ with $eq(\alpha) \in E$ where

$$eq(\alpha) := \{(\omega, \omega') \in \Omega^2 \mid \alpha(\omega) = \alpha(\omega')\}.$$

We call $F/U/E$ *existentially closed*³ if it is existentially closed as a substructure of F/U .

It is easily checked that this is well-defined in the sense that the defined set is indeed the universe of some substructure of F/U ([15] considers only relational languages). Intuitively, using Keisler’s words [15, p.383], $F/U/E$ is the substructure of F/U given by equivalence classes of functions $\alpha \in F$ that are ‘almost constant’ in the sense that $\alpha(\omega) = \alpha(\omega')$ holds throughout some member of E .

Remark 2.5. Let F, U, E be as above. Then $G := \{\alpha \in F \mid eq(\alpha) \in E\}$ is L -closed and $F/U/E \cong G/U$ via the canonical isomorphism $\alpha^U \mapsto \alpha^U \cap G$. In particular, for $F := M^\Omega$, this shows that Keisler’s limit ultrapowers are restricted ultrapowers.

To exemplify the notation, note that for every L -term $t(\bar{x})$ and tuple $\bar{\alpha}$ from F

$$t^{F/U}(\bar{\alpha}^U) = (t^M \circ \bar{\alpha})^U. \tag{1}$$

We need some more notation. For an L -formula $\varphi(\bar{x})$ and a tuple $\bar{\alpha}$ from F we get an $L(F)$ -sentence $\varphi(\bar{\alpha})$ and, if $\omega \in \Omega$, then $\varphi(\bar{\alpha}(\omega))$ is an $L(M)$ -sentence. We define

$$\langle\langle \varphi(\bar{\alpha}) \rangle\rangle := \{\omega \mid M \models \varphi(\bar{\alpha}(\omega))\}.$$

³This is a slight abuse of standard terminology according to which a structure is *existentially closed* if it is existentially closed as a substructure of any of its extensions. We shall not use this terminology.

Definition 2.6. Let $F \subseteq M^\Omega$ be L -closed and U an ultrafilter on Ω . An L -formula $\varphi(\bar{x})$ is *Los for* (F, U) if for all tuples $\bar{\alpha}$ from F :

$$F/U \models \varphi(\bar{\alpha}^U) \iff \langle\langle \varphi(\bar{\alpha}) \rangle\rangle \in U.$$

Being *Los for* F means being Los for (F, U) for every ultrafilter U , and being *Los* means being Los for F for every Ω and every L -closed $F \subseteq M^\Omega$.

Proposition 2.7. *Quantifier-free formulas are Los.*

Proof. It follows easily from (1) that atomic formulas are Los, and Los formulas are closed under Boolean combinations. \square

Proposition 2.8. *Let $F \subseteq M^\Omega$ be L -closed and U be an ultrafilter on Ω . Then every universal sentence true in M is also true in F/U . More generally, if $\varphi(\bar{x})$ is Los for (F, U) and $M \models \forall \bar{x} \varphi(\bar{x})$, then $F/U \models \forall \bar{x} \varphi(\bar{x})$.*

Proof. If $M \models \forall \bar{x} \varphi(\bar{x})$, then $\langle\langle \varphi(\bar{\alpha}) \rangle\rangle = \Omega$ for all $\bar{\alpha}$. If $\varphi(\bar{x})$ is Los for (F, U) , this implies $F/U \models \varphi(\bar{\alpha}^U)$ for all $\bar{\alpha} \in F$. \square

An L -structure A is *generated by* $A_0 \subseteq A$ if every $a \in A$ is the value of a closed $L(A_0)$ -term in A ; it is *generated by one point* if it is generated by a (i.e. by $\{a\}$) for some $a \in A$.

Lemma 2.9. *Let $\Omega = M$ and U be an ultrafilter on Ω . Then T_L^M/U is generated by id_M^U , and $T_{L(M)}^M/U$ is generated by $\{id_M^U\} \cup \{\alpha_a \mid a \in M\}$.*

Proof. For every L -term $t(x)$ we have $\langle\langle t(id_M) = t^M \rangle\rangle = \Omega \in U$. As $t(x) = y$ is Los by Proposition 2.7, this implies $T_L^M/U \models t(id_M^U) = (t^M)^U$.

An $L(M)$ -term $t(x)$ can be written $s(x, a, b, \dots)$ for some L -term $s(x, y, z, \dots)$ and finitely many parameters $a, b, \dots \in M$. Then $\langle\langle s(id_M, \alpha_a, \alpha_b, \dots) = t^M \rangle\rangle = \Omega \in U$ and hence, as before, $T_{L(M)}^M/U \models s(id_M^U, \alpha_a^U, \alpha_b^U, \dots) = (t^M)^U$. \square

Proposition 2.10. *Let $F \subseteq M^\Omega$ be $L(M)$ -closed and U be an ultrafilter on Ω . Then $a \mapsto \alpha_a^U$ defines an isomorphism of M onto an existentially closed substructure of F/U .*

Proof. The map $a \mapsto \alpha_a^U$ is an embedding of M into F/U because it preserves quantifier free formulas by Proposition 2.7. That its image is existentially closed in F/U follows from Proposition 2.8 with $L(M)$ in place of L : view M and F/U as $L(M)$ -structures and note $a^{F/U} = \alpha_a^U$ for $a \in M$. \square

The following proposition implies that the truth of a $\forall\exists$ -sentence is preserved if F is closed under a suitable Skolem function for the sentence. Here, $W : M^r \rightarrow M$ is a *Skolem function* for $\exists y \varphi(x_0, \dots, x_{r-1}, y)$ if for all $\bar{a} \in M^r$ we have

$$M \models (\exists y \varphi(\bar{a}, y) \rightarrow \varphi(\bar{a}, W(\bar{a}))).$$

Proposition 2.11. *Let $F \subseteq M^\Omega$ be L -closed, U an ultrafilter on Ω . If F is closed under some Skolem function for $\exists y\varphi(\bar{x}, y)$ and $\varphi(\bar{x}, y)$ is Los for (F, U) , then so is $\exists y\varphi(\bar{x}, y)$; if additionally $M \models \forall \bar{x}\exists y\varphi(\bar{x}, y)$, then $F/U \models \forall \bar{x}\exists y\varphi(\bar{x}, y)$.*

Proof. Given $\bar{\alpha}$ we have to show

$$\langle\langle \exists y\varphi(\bar{\alpha}, y) \rangle\rangle \in U \iff F/U \models \varphi(\bar{\alpha}^U, \beta^U) \text{ for some } \beta \in F.$$

If $F/U \models \varphi(\bar{\alpha}^U, \beta^U)$ for some $\beta \in F$, then $\langle\langle \varphi(\bar{\alpha}, \beta) \rangle\rangle \in U$ since $\varphi(\bar{x}, y)$ is Los for (F, U) , so $\langle\langle \exists y\varphi(\bar{\alpha}, y) \rangle\rangle \in U$ since $\langle\langle \varphi(\bar{\alpha}, \beta) \rangle\rangle \subseteq \langle\langle \exists y\varphi(\bar{\alpha}, y) \rangle\rangle$. Conversely, assume $\langle\langle \exists y\varphi(\bar{\alpha}, y) \rangle\rangle \in U$ and note $\langle\langle \exists y\varphi(\bar{\alpha}, y) \rangle\rangle = \langle\langle \varphi(\bar{\alpha}, \beta) \rangle\rangle$ for $\beta := W \circ \bar{\alpha} \in F$ where W is a suitable Skolem function. Then $F/U \models \varphi(\bar{\alpha}^U, \beta^U)$ since $\varphi(\bar{x}, y)$ is Los for (F, U) .

If $M \models \forall \bar{x}\exists y\varphi(\bar{x}, y)$, then $F/U \models \forall \bar{x}\exists y\varphi(\bar{x}, y)$ by Proposition 2.8 as $\exists y\varphi(\bar{x}, y)$ is Los for (F, U) . \square

2.2 Examples and historical survey

This subsection illustrates the basic theory developed in the previous one by examples and embeds their treatment in a brief historical survey. On the way, we introduce some notions for later use: Example 2.15 defines (limit) recursive ultrapowers from Hirschfeld's Theorem 1.3, and Definition 2.16 defines *unbounded* ultrafilters.

Ultrapowers and -products have been extensively investigated in model theory in the 60's. For arbitrary (first-order) structures they have first been defined by Los [26] in 1955:

Example 2.12 (Los's Theorem). The usual ultrapower of M modulo U is F/U for $F := M^\Omega$. Since M^Ω is closed under Skolem functions for all formulas, Proposition 2.11 implies that all formulas are Los for M^Ω . This is Los's theorem [26].

According to Keisler, the "initial interest in ultraproducts in the late 1950's was sparked by the discovery of a proof of the Compactness Theorem for first order logic via ultraproducts (see [9]). This proof was attractive because it gave a direct algebraic construction of the required model." [16, Section 4]. Ultraproducts seemed to offer a syntax-free approach to concepts and results of mathematical logic (cf. e.g. [17]). For example, a driving conjecture was that elementarily equivalent structures have isomorphic ultrapowers, known at the time only under the generalized continuum hypothesis. Kochen [17] proved it for direct limits of ultrapowers, and Keisler [15] for *limit ultrapowers* – certain special substructures of ultrapowers (see Definition 2.4). A decade later Shelah [40] finally settled the conjecture.

We refer to [16] for a survey and turn to restricted ultrapowers. Similar to the compactness theorem, Herbrand's theorem has a proof via restricted ultrapowers. We include the simple argument as it illustrates a typical use of Propositions 2.7 and 2.8.

Example 2.13 (Herbrand's Theorem). Let T be a universal theory in the language L and assume it proves $\exists y\varphi(x, y)$ for φ quantifier-free. Then T proves a disjunction of the form $\bigvee_{i < k} \varphi(x, t_i(x))$ where $k \in \mathbb{N}$ and the t_i are L -terms.

Proof. Otherwise, by compactness, there is a model M of T falsifying the universal closures of all these disjunctions. Set $\Omega := M$ and $F := T_L^M$. The family $\langle\langle \neg\varphi(id_\Omega, t^M) \rangle\rangle$ with $t(x)$ ranging over L -terms has the finite intersection property, so is contained in some ultrafilter U . Since T is universal, $F/U \models T$ by Proposition 2.8 and thus $F/U \models \forall x\exists y\varphi(x, y)$. Then, by choice of F , $F/U \models \varphi(id_\Omega^U, (t^M)^U)$ for some L -term $t(x)$, so $\langle\langle \varphi(id_\Omega, t^M) \rangle\rangle \in U$ by Proposition 2.7. But, by choice, U contains the complement of this set, a contradiction. \square

Historically, already Skolem’s [41] nonstandard model of arithmetic from 1934 was a restricted ultrapower:

Example 2.14 (Definable ultrapowers). Let M be a model of Peano arithmetic. Let $\Omega := M$ and F be the set of all functions definable in M (with parameters). Originally, Skolem [41] used for M the standard model \mathbb{N} (in a possibly richer language). Then again F is closed under Skolem functions for all formulas, so all formulas are Los for F . As in Proposition 2.10 one sees that M is isomorphic, via $a \mapsto \alpha_a^U$, to an elementary submodel of F/U .

As in this example, a typical choice for F is the set of functions of some bounded logical or computational complexity. Possibly side-stepping Los’s theorem, restricted ultrapowers offer a syntax-free approach, in Kripke and Kochen’s words, “to prove independence not by the self-referencing technique of Gödel but rather by the older model building method used in geometry” [18, p.211]. E.g. Scott [39, p.244] suggests and discusses this possibility.

Example 2.15 (Recursive ultrapowers). Let M be the standard model \mathbb{N} of arithmetic in the language $\{+, \cdot, 0, 1, <\}$. Let $\Omega := M$ and F be the set of computable functions. A *recursive ultrapower* is a model of the form F/U for U an ultrafilter on \mathbb{N} , and a *limit recursive ultrapower* is a model of the form $F/U/E$ for U an ultrafilter on \mathbb{N} and E a filter on \mathbb{N}^2 . Note F contains Skolem functions for $\exists y\varphi(\bar{x}, y)$ whenever $\mathbb{N} \models \forall \bar{x}\exists y\varphi(\bar{x}, y)$ and φ is existential. By Proposition 2.11, recursive ultrapowers are models of $Th_{\forall\exists}(\mathbb{N})$, the set of $\forall\exists$ -sentences true in \mathbb{N} .

Scott [39, p.244] mentions that non-standard recursive ultrapowers cannot model PA. Indeed, Hirschfeld [12, Theorem 2.6] shows that in such models the standard cut is Σ_2 -definable. Two decades after [39], Kripke and Kochen [18] proved the Paris-Harrington theorem [31] via some restricted ultrapower.

As mentioned in the Introduction, Hirschfeld studied recursive ultrapowers in order to characterize models of $Th_{\Pi_2}(\mathbb{N})$. McLaughlin [29] extends Hirschfeld’s work to functions definable higher up in the arithmetical hierarchy.⁴ These constructions can meaningfully start with a nonstandard model M of a sufficiently large fragment of arithmetic (cf. [11, IV.1]). A famous example is Mac Dowell and Specker’s [27] construction of end extensions as definable ultrapowers over a nonstandard model of Peano arithmetic. Since we shall use some of the notions we include some details.

⁴In contrast to the setting of [18] and of this paper, [39, 33, 12, 29] consider ultrafilters over restricted Boolean algebras, namely those consisting of sets with characteristic function in F .

Assume M interprets a binary relation symbol $< \in L$ by a linear order $<^M$. For $a \in M$ let

$$[a] := \{b \in M \mid b <^M a\}.$$

A subset $X \subseteq M$ is *bounded* if $X \subseteq [a]$ for some $a \in M$, and otherwise *unbounded*.

Definition 2.16. Let $\Omega \subseteq M$ be unbounded. An ultrafilter U on Ω is *unbounded* if it contains only unbounded sets, equivalently, if $\{\Omega \setminus [a] \mid a \in M\} \subseteq U$.

Lemma 2.17. Let $\Omega \subseteq M$ be unbounded. Every collection of unbounded subsets \mathcal{X} of Ω which is closed under intersections is contained in an unbounded ultrafilter on Ω .

Proof. If $\mathcal{X} \cup \{\Omega \setminus [a] \mid a \in M\}$ is not contained in some ultrafilter, then it does not have the finite intersection property, that is, there are finite $\mathcal{Y} \subseteq \mathcal{X}$ and $A \subseteq M$ such that $\bigcap \mathcal{Y} \cap \bigcap_{a \in A} (\Omega \setminus [a]) = \emptyset$, so $\bigcap \mathcal{Y} \subseteq [a^*]$ where a^* is the $<^M$ -maximum of A . Hence \mathcal{X} contains a bounded set or is not closed under intersections. \square

Recall that N is an *end extension* of M if M is a substructure of N and for all $a, b \in N$ we have that $a <^N b \in M$ implies $a \in M$.

Example 2.18 (Definable ultrapowers, continued). Let M, F be as in Example 2.14 and assume M is countable. Then there exists an ultrafilter U such that, up to isomorphism, F/U is an elementary end extension of M . This is Mac Dowell and Specker's theorem [27].

Proof. Let $(\alpha_0, a_0), (\alpha_1, a_1), \dots$ enumerate $F \times M$ and define a sequence $X_0 \supseteq X_1 \supseteq \dots$ of unbounded definable subsets of $\Omega = M$, as follows. Set $X_0 := M$ and assume X_i is defined. The function $\omega \mapsto \min\{\alpha_i(\omega), a_i\}$ is constant on some unbounded subset of X_i , say, equal to $b_i \leq a_i$ (see [11, II.1]); set $X_{i+1} := \{\omega \in X_i \mid \min\{\alpha_i(\omega), a_i\} = b_i\}$.

Choose an ultrafilter U containing every X_i . By Example 2.14, we are left to verify that F/U is an end extension of the image of the map $a \mapsto \alpha_a^U$, that is: for all $\alpha \in F$ and $a \in M$, if $\alpha^U <^{F/U} \alpha_a^U$ then there is $b <^M a$ such that $\alpha^U = \alpha_b^U$. So assume $\alpha^U <^{F/U} \alpha_a^U$, i.e., $X := \{\omega \mid \alpha(\omega) <^M a\} \in U$. Then $a \neq 0$. Choose $i \in \mathbb{N}$ such that $\alpha_i = \alpha$ and $a_i +^M 1 = a$. Then $b_i <^M a$ and $X \cap X_{i+1} \subseteq \{\omega \mid \alpha(\omega) = b_i\} = \{\omega \mid \alpha(\omega) = \alpha_{b_i}(\omega)\} \in U$. \square

Restricted ultrapowers have also been used in bounded arithmetic. There, a natural choice for F is the set of polynomial time computable functions yielding *polynomial time ultrapowers*: see the next section. E.g. Pudlák [33] uses such a structure. One can also start with a nonstandard model M of $\forall\text{PV}$. Such structures are constructed in [21] and [10] who present their powers for functions restricted to some M -finite $\Omega \subseteq M$ (cf. Remark 2.21); in [10] F is additionally restricted to straight-line programs of a certain (nonstandard) length.

Krajíček's book [23] is dedicated to related constructions, and we are partly following its notation in order to stress the similarity. In [23], the typically M -finite index set Ω is interpreted as a sample space and functions in F as random variables, typically of low computational complexity. Instead of dividing by an ultrafilter one constructs a Boolean valued model with values in a carefully chosen Boolean algebra.

Our proof of Theorem 1.1 relies on a polynomial time ultrapower over a nonstandard model M of $\forall\text{PV}$, namely one that is Herbrand saturated in the sense of [1]; our index set Ω is not M -finite but an unbounded definable subset of M .

2.3 \forall PV and polynomial time ultrapowers

We define the language PV to contain a binary relation symbol $<$ and each polynomial time computable function (on \mathbb{N}) as a symbol; of course, if $f : \mathbb{N}^r \rightarrow \mathbb{N}$ is such a function, then the arity of the symbol is r . As usual, $\text{TIME}(n^k) \subseteq \text{PV}$ is the set of functions computable in time $O(n^k)$. The *standard PV-model* \mathbb{N} has universe \mathbb{N} and interprets $<$ by the natural order and each function symbol by itself. The set of universal sentences true in \mathbb{N} is

$$\forall \text{PV}.$$

To fix our notation we list some functions in PV. It contains Buss' language of arithmetic $x + y, x \cdot y, \lfloor x/2 \rfloor, |x|, x \# y$ and constants $0, 1, 2, \dots$. The *length* $|n| := \lceil \log_2(n + 1) \rceil$ is the length of the binary encoding of n ; that is, $n = \sum_{i < |n|} 2^i \cdot \text{bit}(n, i)$ where $\text{bit}(n, i) := 0$ for $i \geq |n|$ and $\text{bit}(x, y) \in \text{PV}$. The *smash* function is $n \# m := 2^{|n| \cdot |m|}$. For every fixed $k \in \mathbb{N}$ there is $\langle x_0, \dots, x_{k-1} \rangle \in \text{PV}$ such that every $(n_0, \dots, n_{k-1}) \in \mathbb{N}^k$ is *coded* by $n := \langle n_0, \dots, n_{k-1} \rangle$; namely, we have $(n)_i = n_i$ where $(x)_0, (x)_1, \dots$ are unary functions in PV.

It is easy to see that the functions in PV are \forall PV-provably closed under composition and definitions by quantifier-free case distinctions:

Lemma 2.19. *For each PV-term $t(\bar{x})$ there is $f(\bar{x}) \in \text{PV}$ such that $\forall \text{PV}$ proves $t(\bar{x}) = f(\bar{x})$. For every quantifier-free PV-formula $\varphi(\bar{x})$ and PV-terms $t(\bar{x}), s(\bar{x})$ there is $f(\bar{x}) \in \text{PV}$ such that $\forall \text{PV}$ proves $(f(\bar{x}) = t(\bar{x}) \wedge \varphi(\bar{x})) \vee (f(\bar{x}) = s(\bar{x}) \wedge \neg \varphi(\bar{x}))$.*

Recall Example 2.2. For a model M of \forall PV we write

$$\text{PTIME}(M) := T_{\text{PV}(M)}^M.$$

Using sequence coding and the above lemma one sees that every $t^M(x) \in \text{PTIME}(M)$ equals $f^M(x, a)$ for some $f(x, y) \in \text{PV}$ (independent of M) and some parameter $a \in M$. We shall write $f_a(x)$ for $f(x, a)$. The phrase “for all (there is) $f_a^M(x) \in \text{PTIME}(M) \dots$ ” stands for “for all (there is) $f(x, y) \in \text{PV}$ and for all (there is) $a \in M \dots$ ”.

Definition 2.20. Let M be a model of \forall PV. A *polynomial time ultrapower over M* is a PV-structure of the form $\text{PTIME}(M)/U$ for some ultrafilter U on M .

A *limit polynomial time ultrapower over M* is a PV-structure of the form $\text{PTIME}(M)/U/E$ for some ultrafilter U on M and some filter E on M^2 .

For $M = \mathbb{N}$, the standard PV-model, we omit the phrase “over \mathbb{N} ”.

Remark 2.21. As mentioned in the end of Section 2.2 the functions in $\text{PTIME}(M)$ can be restricted to some non-empty $\Omega \subseteq M$, i.e., take for F the set

$$\text{PTIME}(M) \upharpoonright \Omega := \{\alpha \upharpoonright \Omega \mid \alpha \in \text{PTIME}(M)\}.$$

This might be convenient but, for general reasons (cf. [9, Corollary 1.3]), does not lead to anything new. Indeed, if U is an ultrafilter on Ω and V is any ultrafilter on M containing U , then $(\alpha \upharpoonright \Omega)^U \mapsto \alpha^V$ is an isomorphism from $(\text{PTIME}(M) \upharpoonright \Omega)/U$ onto $\text{PTIME}(M)/V$.

By Propositions 2.8 and 2.10:

Proposition 2.22. *Let M be a model of $\forall\text{PV}$. Every polynomial time ultrapower over M is a model of $\forall\text{PV}$ and, in fact, has an existentially closed substructure isomorphic to M .*

If we disallow parameters from M in the definition of polynomial time ultrapowers over M , then we get nothing new from starting with M instead \mathbb{N} :

Theorem 2.23. *Let M be a model of $\forall\text{PV}$. For every ultrafilter U over M there exists an ultrafilter V over \mathbb{N} such that $T_{\text{PV}}^M/U \cong \text{PTIME}(\mathbb{N})/V$.*

We give a proof in the next subsection.

2.4 Characterization theorem

It turns out that Hirschfeld’s [12] results can be proved in a much more general setting, and not only for fragments of arithmetic. The theories should however be able to do some sequence coding. The following ad hoc notion isolates what is needed. *Sequential* theories (see [11, Definition III.1.12]) satisfy it up to a conservative addition of some function symbols.

Definition 2.24. A theory is *weakly sequential* if for every countable model of M there is a family of L -terms $(t_a(x))_{a \in M}$ such that for every finite subset $A \subseteq M$ there is $b \in M$ such that $M \models t_a(b) = a$ for all $a \in A$.

Lemma 2.25. *$\forall\text{PV}$ is weakly sequential.*

Proof. Enumerate a countable $M \models \forall\text{PV}$ by m_0, m_1, \dots and set $t_{m_i}(x) := (x)_i$. Given a finite $A \subseteq M$ choose $k \in \mathbb{N}$ larger than all $i \in \mathbb{N}$ with $m_i \in A$. Then $b := \langle m_0, \dots, m_{k-1} \rangle^M$ is as required. \square

For the rest of this section fix a countable first-order language L and a countable L -structure M . We consider only ultrapowers with $\Omega := M$. Let $\text{Th}_{\forall}(M)$ be the set of universal sentences which are true in M .

Theorem 2.26 (Characterization). *Assume $\text{Th}_{\forall}(M)$ is weakly sequential. Then a countable L -structure is a model of $\text{Th}_{\forall}(M)$ if and only if it is isomorphic to an existentially closed limit restricted ultrapower $T_L^M/U/E$ where U is an ultrafilter on M and E is a filter on M^2 .*

Remark 2.27. The backward direction is clear by Proposition 2.8, and holds even when “existentially closed” is deleted.

Our main results concerning polynomial time ultrapowers and models of $\forall\text{PV}$ are direct consequences of Theorem 2.26: Theorem 1.4 follows setting $L := \text{PV}$ and $M := \mathbb{N}$, and Corollary 1.5 follows by Remark 2.5.

The following two lemmas comprise the two main steps in the proof of Theorem 2.26. They do not need the assumption of weak sequentiality.

Lemma 2.28. *Assume N is a model of $Th_{\forall}(M)$ generated by one point. Then there exists an ultrafilter U on M such that N is isomorphic to T_L^M/U .*

Proof. Let N be generated by $a \in N$. Since $N \models Th_{\forall}(M)$, the collection

$$\mathcal{X} := \{\varphi(M) \mid \varphi(x) \text{ is a quantifier-free } L\text{-formula and } N \models \varphi(a)\}$$

has the finite intersection property. Let U be an ultrafilter containing \mathcal{X} . We claim that N is isomorphic to T_L^M/U via an isomorphism that maps a to id_M^U . Since T_L^M/U is generated by id_M^U (Lemma 2.9) it suffices to show that a and id_M^U satisfy the same quantifier-free formulas in their respective structures. But for quantifier-free $\varphi(x)$ we have

$$\begin{aligned} N \models \varphi(a) &\iff \varphi(M) \in \mathcal{X} \\ &\iff \langle\langle \varphi(id_M) \rangle\rangle \in U \\ &\iff T_L^M/U \models \varphi(id_M^U). \end{aligned}$$

For the second equivalence note $\varphi(M) = \langle\langle \varphi(id_M) \rangle\rangle$, so the forward direction is clear; conversely, if $\varphi(M) \notin \mathcal{X}$, then $\neg\varphi(M) \in \mathcal{X}$, so $\langle\langle \neg\varphi(id_M) \rangle\rangle = M \setminus \langle\langle \varphi(id_M) \rangle\rangle \in U$, so $\langle\langle \varphi(id_M) \rangle\rangle \notin U$. The third equivalence follows from $\varphi(x)$ being Los by Proposition 2.7. \square

Lemma 2.29. *Let U be an ultrafilter on M and N be an existentially closed substructure of T_L^M/U . Then there exists a filter E on M^2 such that N equals $T_L^M/U/E$.*

Proof. Let E be the filter on M^2 generated by the sets $eq(t^M)$ where $t(x)$ ranges over L -terms such that $(t^M)^U \in N$. Obviously, $N \subseteq T_L^M/U/E$, and we show the converse.

Let $(t^M)^U \in T_L^M/U/E$ for some L -term $t(x)$. Then there are L -terms $t_0(x), \dots, t_{k-1}(x)$ with $(t_0^M)^U, \dots, (t_{k-1}^M)^U \in N$ such that

$$eq(t^M) \supseteq \bigcap_{i < k} eq(t_i^M). \tag{2}$$

Using Proposition 2.7 we get

$$T_L^M/U \models t(id_M^U) = (t^M)^U \wedge \bigwedge_{i < k} t_i(id_M^U) = (t_i^M)^U.$$

Since N is existentially closed in T_L^M/U there are $\alpha^U, \beta^U \in N$ such that

$$N \models t(\alpha^U) = \beta^U \wedge \bigwedge_{i < k} t_i(\alpha^U) = (t_i^M)^U.$$

Then T_L^M/U models this sentence too, so by Proposition 2.7

$$X := \langle\langle t(\alpha) = \beta \wedge \bigwedge_{i < k} t_i(\alpha) = t_i^M \rangle\rangle \in U.$$

It suffices to show that $t^M(\omega) = \beta(\omega)$ for all $\omega \in X$. But $\omega \in X$ means

$$M \models t(\alpha(\omega)) = \beta(\omega) \wedge \bigwedge_{i < k} t_i(\alpha(\omega)) = t_i(\omega).$$

The second conjunct and (2) imply $(\omega, \alpha(\omega)) \in eq(t^M)$, that is, $t^M(\omega) = t^M(\alpha(\omega))$. Thus, the first conjunct gives $t^M(\omega) = \beta(\omega)$, as claimed. \square

We are ready to prove the Characterization Theorem 2.26:

Proof of Theorem 2.26. For the backward direction see Remark 2.27. To see the forward direction, let N be a countable model of $Th_{\forall}(M)$. Recall that α_a denotes the function which is constantly a . For N let $(t_a(x))_{a \in N}$ witness that $Th_{\forall}(M)$ is weakly sequential. This means that the collection $\{\langle\langle t_a(id_N) = \alpha_a \rangle\rangle \mid a \in N\}$ has the finite intersection property. Let V be an ultrafilter extending it.

By Lemma 2.9, $T_{L(N)}^N/V$ is generated by id_N^V together with $\alpha_a^V, a \in N$. But, in fact, it is generated by id_N^V alone: $T_{L(N)}^N/V \models t_a(id_N^V) = \alpha_a^V$ because $\langle\langle t_a(id_N) = \alpha_a \rangle\rangle \in V$ and $t_a(x) = y$ is Los by Proposition 2.7.

As $T_{L(N)}^N/V$ models $Th_{\forall}(M)$ by Proposition 2.8 and is generated by one point we can apply Lemma 2.28 and get an ultrafilter U on M such that

$$T_{L(N)}^N/V \cong T_L^M/U.$$

By Proposition 2.10, N is isomorphic to an existentially closed substructure of $T_{L(N)}^N/V$ and thus of T_L^M/U . By Lemma 2.29, $N \cong T_L^M/U/E$ for some filter E on M^2 . \square

Theorem 2.23 from the previous subsection follows from Lemma 2.28:

Proof of Theorem 2.23. Consider a structure of the form T_{PV}^N/U where N is a countable model of $\forall\text{PV}$ and U is an ultrafilter on N . This structure models $\forall\text{PV}$ by Proposition 2.8 and is generated by id_N^U by Lemma 2.9. Applying Lemma 2.28 shows T_{PV}^N/U is isomorphic to $\text{PTIME}(\mathbb{N})/V$ for some ultrafilter V on \mathbb{N} . \square

For completeness we show how to derive Hirschfeld's Theorem 1.3. Recall (limit) recursive ultrapowers have been defined in Example 2.15.

Proof of Theorem 1.3. Let \mathbb{N}^* be the standard model of arithmetic in the language consisting of $<$ and symbols for all computable functions. Clearly, every model of $Th_{\Pi_2}(\mathbb{N})$ has an expansion to a model of $Th_{\forall}(\mathbb{N}^*)$. Now, for the forward direction, apply Theorem 2.26 and take the reduct to $\{\cdot, +, <, 0, 1\}$. Conversely, recursive ultrapowers model $Th_{\forall\exists}(\mathbb{N})$ by Example 2.15, and hence, so do their existentially closed substructures. It follows from the MRDP theorem [8] that $Th_{\forall\exists}(\mathbb{N})$ is equivalent to $Th_{\Pi_2}(\mathbb{N})$ (see [12, Corollary 1.7.1.b]). \square

Remark 2.30. Note Theorem 2.26 applies directly to $Th_{\forall}(\mathbb{N}^*)$ which is a conservative extension of $Th_{\forall\exists}(\mathbb{N})$. We noted in Remark 2.27 that Theorem 2.26 holds true with “existentially closed” deleted. A similar remark holds true for Theorem 1.3 because actually all limit recursive ultrapowers are existentially closed [12, Theorem 3.5].

3 Consistency of circuit lower bounds

3.1 Herbrand saturation and proof outline

Building on earlier work of Zambella [42], Avigad [1] proposed a model-theoretic approach to witnessing theorems based on the notion of Herbrand saturation. An L -structure M is *Herbrand saturated* if every universal $L(M)$ -formula $\varphi(\bar{x})$ which is consistent with the universal diagram of M is satisfiable in M . The universal diagram of M is the set of all universal $L(M)$ -sentences true in M . Avigad [1, Theorem 3.2] showed

Theorem 3.1. *Every universal theory has a Herbrand saturated model.*

The crucial property of Herbrand saturated models is [1, Theorem 3.3]. We state it for $\forall\text{PV}$, simplified using Lemma 2.19.

Theorem 3.2. *Let M be a Herbrand saturated model of $\forall\text{PV}$. Assume $M \models \forall\bar{x}\exists y\varphi(\bar{x}, y)$ where $\varphi(\bar{x}, y)$ is a quantifier-free $\text{PV}(M)$ -formula. Then there exists $f_a^M(\bar{x}) \in \text{PTIME}(M)$ such that $M \models \forall\bar{x}\varphi(\bar{x}, f_a(\bar{x}))$*

We describe the idea of the proof of Theorem 1.1. As explained in the Introduction the crucial step is to infer the consistency of a non-uniform lower bound from the truth of a uniform lower bound, namely from Theorem 1.2. This is done by switching back and forth between two perspectives on $\alpha \in \text{PTIME}(M)$ given an ultrapower $\text{PTIME}(M)/U$: as a point α^U in the structure $\text{PTIME}(M)/U$ or as a function $\alpha : M \rightarrow M$ in M . See [23, Section 24.4] or [28] for a use of these views in the context of forcing with random variables.

Assume $\text{PTIME}(M)/U \models \forall\ell\neg\text{LB}[g](\text{small}, \ell)$. Plug $|id_M^U|$ for ℓ and choose $\zeta \in \text{PTIME}(M)$ such that ζ^U is in $\text{PTIME}(M)/U$ a small circuit computing g on inputs of length $|id_M^U|$. We restrict attention to unary strings 1^n as arguments for the functions in $\text{PTIME}(M)$. Now, ζ is a function on M and its value is U -often a small circuit in the sense of M . In this sense, $(\zeta(1^n))_n$ is “almost” a uniform family of small circuits in M . If g satisfies Theorem 1.2 in M , this should give x of some length n such that the circuit $\zeta(1^n)$ evaluated on x disagrees with $g(x)$ in M . If g satisfies an “almost everywhere” version of Theorem 1.2 in M we get such a counterexample x at all sufficiently large lengths n . If M is Herbrand saturated, then there is $\gamma \in \text{PTIME}(M)$ computing such counterexamples x from 1^n in M . We intend to take this counterexample function γ in M as a single counterexample γ^U in $\text{PTIME}(M)/U$ at length $|id_M^U|$.

Lemma 3.6 makes the above sketch precise. It infers a non-uniform lower bound in some polynomial time ultrapower over M from the truth of a uniform lower bound in the Herbrand saturated M . The latter is proved as Lemma 3.5 by carrying out Santhanam and Williams’ proof of Theorem 1.2 in M .

3.2 Formalization of circuit lower bounds

Let M be a model of $\forall\text{PV}$. We let $\text{Log}(M)$ denote the set of $n \in M$ such that $n = |N|$ for some $N \in M$. For such n we write 1^n for $1\#N - 1$ (indicating its binary expansion). Here

and below we often omit superscripts writing e.g. $\#$ instead $\#^M$. We shall use the phrase “for all large enough $n \in \text{Log}(M) : \dots n \dots$ ” for

“there is $n_0 \in \text{Log}(M)$ such that for every $n \in \text{Log}(M)$ with $n_0 < n : \dots n \dots$ ”

We shall need some details of how circuits are coded by numbers. In this paper all circuits have gates of fan-in at most 2 and, unless specified otherwise, exactly one output gate. We code a circuit C of size (number of gates) s by a number coding the set of tuples $\langle u, v, w \rangle$ where $u, v < s$ are (numbers of) gates such that u is wired to gate v and $w < 3$ specifies the label \wedge, \vee or \neg of v . There are $O(s)$ such tuples, each of length $O(|s|)$, so

$$|C| \leq O(s \cdot |s|), \quad (3)$$

where we blur the distinction between C and its code. It is convenient to allow 0 as a code of a circuit of size 0. Given a pair (a, C) it is decidable in polynomial time whether a is in the set coded by C . Hence, there is a quantifier free PV-formula $x \in y$ that defines the set of these pairs (a, C) in the standard PV-structure \mathbb{N} . Similarly, there is a quantifier free PV-formula $\text{Circuit}(x, y)$ defining in \mathbb{N} the set of pairs (C, s) such that C is a circuit of size at most s . By convention, $\text{Circuit}(0, 0) \in \forall\text{PV}$.

There is $\text{eval}(x, y) \in \text{TIME}(n^2)$ such that $\text{eval}(C, a)$ is the output bit of the circuit C when its inputs are assigned the bits $\text{bit}(a, 0), \text{bit}(a, 1), \dots$; if C is not a circuit, then $\text{eval}(C, a) = 0$. Note eval is Boolean where we call $f(\bar{x}) \in \text{PV Boolean}$ if $\forall \bar{x} f(\bar{x}) < 2 \in \forall\text{PV}$.

The following formula expresses that a Boolean $f(x) \in \text{PV}$ is not computable by size $\leq s$ circuits on inputs of length $n > 0$:

$$LB[f](s, n) := n > 0 \wedge \forall C (\text{Circuit}(C, s) \rightarrow \exists x (|x| = n \wedge f(x) \neq \text{eval}(C, x))).$$

For readability we use C, s, n as variables. Given $a \in M$ we can plug $f(x, a)$ for $f(x)$ and get a formula $LB[f_a](s, n)$ with parameter a (recall the notation f_a from Section 2.3).

3.3 Santhanam and Williams’ proof

The following is [24, Lemma 3.1], a formalization of a “folklore result about a time hierarchy for deterministic time, where the lower bound holds against sublinear advice.” [38, Proposition 1]

Lemma 3.3. *For every $d \in \mathbb{N}$ there is a Boolean $g_d(x) \in \text{TIME}(n^{d+1})$ such that for every $h(x, y) \in \text{TIME}(n^d)$ there is $c_h \in \mathbb{N}$ such that $\forall\text{PV}$ proves*

$$n > c_h \wedge |a| = n^{2/3} \rightarrow \exists x (|x| = n \wedge h(x, a) \neq g_d(x)).$$

In the following, by n^δ for some rational δ we mean $\lfloor n^\delta \rfloor$.

Definition 3.4. Let $\delta \geq 0$ be rational, and M be a model of $\forall\text{PV}$. We call $f_a^M(x) \in \text{PTIME}(M)$ a *uniform size n^δ circuit family in M* if for all $n \in \text{Log}(M)$

$$M \models \text{Circuit}(f_a(1^n), n^\delta).$$

Lemma 3.5. *Let $k \geq 3$ be natural, $0 < \epsilon < 1/3$ rational, and M a model of $\forall\text{PV}$. There is a Boolean $g(x) \in \text{TIME}(n^{3k})$ such that for every uniform size $n^{k+\epsilon}$ circuit family $f_a^M \in \text{PTIME}(M)$ there is a Boolean $\tilde{f}_a^M \in \text{PTIME}(M)$ such that at least one of the following holds:*

(a) $M \models \forall z \exists n \text{LB}[\tilde{f}_a](|z| \cdot n^k, n)$, or,

(b) $M \models \exists x(|x| = n \wedge g(x) \neq \text{eval}(f_a(1^n), x))$ for all large enough $n \in \text{Log}(M)$.

Proof. Let $g(x)$ be the function $g_d(x)$ from the previous Lemma 3.3 with $d := 3k - 1$, and let f_a^M be as stated. It suffices to show that, if (a) fails for suitable \tilde{f}_a^M chosen below, then there is $h(x, y) \in \text{TIME}(n^{3k-1})$ such that for all large enough $n \in \text{Log}(M)$ there is $D \in M$ with $|D| = n^{2/3}$ such that

$$M \models \forall x(|x| = n \rightarrow h(x, D) = \text{eval}(f_a(1^n), x)). \quad (4)$$

Argue in M . Recall $f_a(1^n)$ codes a set of triples $\langle u, v, w \rangle$ with $u, v < n^{k+\epsilon}$ and $w < 3$. Every such triple has length at most $c \cdot |n|$ for some $c \in \mathbb{N}$. Hence, provided $n \in \text{Log}(M)$ is larger than some standard constant, every tuple $\langle n, u, v, w \rangle$ with $u, v < n^{k+\epsilon}$ and $w < 3$ can be *padded* to length exactly $n^{1/(2k)}$. More precisely, there is a $\text{TIME}(n)$ function that computes the padded version of $\langle n, u, v, w \rangle$ given $\langle n, u, v, w \rangle$ and 1^n ; further, there is a $\text{TIME}(n)$ function that computes the tuple $\langle n, u, v, w \rangle$ when given its padded version and 1^n . It is easy to see that the characteristic function of the set of padded versions of tuples $\langle n, u, v, w \rangle$ with $\langle u, v, w \rangle \in f_a(1^n)$ has the form $\tilde{f}_a^M(x) \in \text{PTIME}(M)$ (with the same parameter a).

Now assume (a) fails, so there are $b \in M$ and for every $n \in \text{Log}(M) \setminus \{0\}$ and $m := n^{1/(2k)}$ a circuit D_m of size $\leq |b| \cdot n^{1/2}$ that computes $\tilde{f}_a(x)$ on inputs x of length m . Then $|D_m| < n^{2/3}$ for large enough $n \in \text{Log}(M)$. Indeed, for some $e \in \mathbb{N}$ we have in M

$$|D_m| \leq e \cdot n^{1/(2k)} \cdot |b|n^{1/2} \cdot ||b|n^{1/2}| \leq e \cdot n^{1/(2k)+1/2+1/100} \cdot |n| < n^{2/3};$$

the first inequality holds by (3), the second holds for $n \geq |b|^{100}$, and the third for n larger than some standard constant. Similarly as above we *pad* codes of circuits of length $< n^{2/3}$ to length exactly $n^{2/3}$.

Let $h(x, y)$ be computed by the following algorithm: first check that $|y| = n^{2/3}$ for $n := |x|$ and y is the padded version of a circuit D with $|D| < |y|$; if the check fails, output 0; else for all $u, v < n^{k+\epsilon}$ and $w < 3$ compute $\text{eval}(D, p)$ where p is the padded (to length $m = n^{1/(2k)}$) version of $\langle n, u, v, w \rangle$; define C as the (number coding the) set containing for every such p with $\text{eval}(D, p) = 1$ the tuple $\langle u, v, w \rangle$; finally, output $\text{eval}(C, x)$.

Then $h(x, y) \in \text{TIME}(n^{3k-1})$. Indeed, its time is dominated by the evaluations $\text{eval}(D, p)$. There are $(n^{k+\epsilon})^2 \cdot 3$ many of them and each needs time $O((|p| + |D|)^2) \leq O(n^{4/3})$ (plus $O(n)$ for the computation of p). But $n^{2k+2\epsilon+4/3} \leq n^{3k-1}$ by the assumptions on k and ϵ .

If in M we plug for y the padded (to length $n^{2/3}$) version of D_m , then the computed C equals the circuit $f_a(1^n)$. This implies (4). \square

Note, setting $M = \mathbb{N}$ in this lemma implies Theorem 1.2.

3.4 Proof of Theorem 1.1

Lemma 3.6. *Let $k \in \mathbb{N}$, $0 < \epsilon < 1$ be rational, and M a Herbrand saturated model of $\forall\text{PV}$. Suppose $g(x) \in \text{PV}$ is Boolean and such that for every uniform size $n^{k+\epsilon}$ circuit family $f_a^M \in \text{PTIME}(M)$ we have for all large enough $n \in \text{Log}(M)$:*

$$M \models \exists x (|x| = n \wedge g(x) \neq \text{eval}(f_a(1^n), x)). \quad (5)$$

Then there exists a polynomial time ultrapower over M that satisfies $\forall z \exists \ell \text{LB}[g](|z| \cdot \ell^k, \ell)$.

Proof. Let

$$\Omega := \{1^n \mid n \in \text{Log}(M)\},$$

and U be an unbounded ultrafilter on Ω (Lemma 2.17). By Remark 2.21 it suffices to show:

$$(\text{PTIME}(M)\upharpoonright\Omega)/U \models \forall z \exists \ell \text{LB}[g](|z| \cdot \ell^k, \ell).$$

Let $\alpha \in \text{PTIME}(M)$, a value for z , be given, say $\alpha = z_{a_0}^M$ for $z(x, y) \in \text{PV}$ and $a_0 \in M$. We choose a witness for ℓ as follows.

It is not hard to see that there is $s \in \mathbb{N}$ such that $M \models |z_{a_0}(1^n)| \leq n^s$ for all large enough $n \in \text{Log}(M)$ (with threshold depending on $|a_0|$). Then set $t := \lceil s/\epsilon \rceil \in \mathbb{N}$ and choose $\ell(x) \in \text{PV}$ such that $\ell(1^n) = 1^{n^t}$ for all $n \in \text{Log}(M)$. We then have

$$M \models |z_{a_0}(1^n)| \leq |\ell(1^n)|^\epsilon \quad (6)$$

for all large enough $n \in \text{Log}(M)$. We set $\lambda := \ell^M(x)$ and aim to verify $\text{LB}[g](|\alpha^U| \cdot |\lambda^U|^k, |\lambda^U|)$ in $(\text{PTIME}(M)\upharpoonright\Omega)/U$. So let $\zeta \in \text{PTIME}(M)$ be such that

$$(\text{PTIME}(M)\upharpoonright\Omega)/U \models \text{Circuit}(\zeta^U, |\alpha^U| \cdot |\lambda^U|^k). \quad (7)$$

We are looking for $\gamma \in \text{PTIME}(M)$ such that

$$(\text{PTIME}(M)\upharpoonright\Omega)/U \models |\gamma^U| = |\lambda^U| \wedge g(\gamma^U) \neq \text{eval}(\zeta^U, \gamma^U). \quad (8)$$

Let $\zeta = C_{a_1}^M(x)$ for $C(x, y) \in \text{PV}$ and $a_1 \in M$. By Proposition 2.7, (7) implies that

$$X := \{1^n \in \Omega \mid M \models \text{Circuit}(C_{a_1}(1^n), |z_{a_0}(1^n)| \cdot |\ell(1^n)|^k)\} \in U.$$

By Lemma 2.19 we find $\tilde{C}(x, y) \in \text{PV}$ such that for all $n \in \text{Log}(M)$

$$M \models \tilde{C}_{a_1}(1^n) = \begin{cases} C_{a_1}(1^{n^{1/t}}) & \text{if } \text{Circuit}(C_{a_1}(1^{n^{1/t}}), n^{k+\epsilon}) \\ 0 & \text{else.} \end{cases} \quad (9)$$

Since $\text{Circuit}(0, 0) \in \forall\text{PV}$ by convention, $\tilde{C}_{a_1}^M(x)$ is a uniform sequence of size $n^{k+\epsilon}$ circuits in M . By assumption (5) we have for all large enough $n \in \text{Log}(M)$:

$$M \models \exists x (|x| = n \wedge g(x) \neq \text{eval}(\tilde{C}_{a_1}(1^n), x)) \quad (10)$$

It follows from Theorem 3.2 that there exists a function $w_{a_2}^M(x) \in \text{PTIME}(M)$ such that for all large enough $n \in \text{Log}(M)$:

$$M \models |w_{a_2}(1^n)| = n \wedge g(w_{a_2}(1^n)) \neq \text{eval}(\tilde{C}_{a_1}(1^n), w_{a_2}(1^n)) \quad (11)$$

Indeed, by (10) there is $b \in \text{Log}(M)$ such that

$$M \models \forall u \exists x (b < |u| \rightarrow |x| = |u| \wedge g(x) \neq \text{eval}(\tilde{C}_{a_1}(1^{|u|}), x))$$

and Herbrand saturation gives $w_{a_2}^M(u) \in \text{PTIME}(M)$ witnessing $\exists x$ as in Theorem 3.2.

Choose $\tilde{w}(x, y) \in \text{PV}$ such that $\forall \text{PV}$ proves $\tilde{w}(x, y) = w(\ell(x), y)$, so $M \models \tilde{w}_{a_2}(1^n) = w_{a_2}(1^{n^t})$ for all $n \in \text{Log}(M)$. Define

$$\gamma := \tilde{w}_{a_2}^M.$$

We are left to verify (8). Plugging n^t for n in (11) gives

$$M \models |\tilde{w}_{a_2}(1^n)| = |\ell(1^n)| \wedge g(\tilde{w}_{a_2}(1^n)) \neq \text{eval}(\tilde{C}_{a_1}(\ell(1^n)), \tilde{w}_{a_2}(1^n))$$

for all large enough $n \in \text{Log}(M)$. Since U is unbounded, this implies for $\tilde{\zeta} := \tilde{C}_{a_1}^M(\ell^M(x))$

$$(\text{PTIME}(M) \upharpoonright \Omega) / U \models |\gamma^U| = |\lambda^U| \wedge g(\gamma^U) \neq \text{eval}(\tilde{\zeta}^U, \gamma^U).$$

It now suffices to show that $\tilde{\zeta}^U = \zeta^U$: in M we have

$$|z_{a_0}(1^m)| \cdot |\ell(1^m)|^k \leq |\ell(1^m)|^{k+\epsilon} = (m^t)^{k+\epsilon}$$

for all $m \in \text{Log}(M)$ large enough such that (6) holds; since U is unbounded, it contains the set of $1^m \in X$ such that m is large enough such that (6) holds; for 1^m in this set we have $M \models \text{Circuit}(C_{a_1}(1^m), (m^t)^{k+\epsilon})$ and thus $\tilde{C}_{a_1}^M(1^{m^t}) = C_{a_1}^M(1^m)$ by (9). \square

We are ready to prove our main result:

Proof of Theorem 1.1. By Theorem 3.1 there exists a Herbrand saturated model M of $\forall \text{PV}$. We can assume without loss of generality that $k > 3$. Let $0 < \epsilon < 1/3$ and choose $g(x) \in \text{PV}$ according to Lemma 3.5 and distinguish two cases.

If (5) holds for every uniform size $n^{k+\epsilon}$ circuit family $f_a^M \in \text{PTIME}(M)$, then Lemma 3.6 states that $\forall z \exists n \text{LB}[g](|z| \cdot n^k, n)$ holds in some polynomial time ultrapower over M , so is consistent with $\forall \text{PV}$ (Proposition 2.22).

Otherwise there is a uniform size $n^{k+\epsilon}$ circuit family $f_a^M \in \text{PTIME}(M)$ such that (5) fails. Then Lemma 3.5 implies

$$M \models \forall z \exists n \text{LB}[\tilde{f}_a](|z| \cdot n^k, n) \quad (12)$$

some $\tilde{f}_a^M(x) \in \text{PTIME}(M)$, i.e., some $\tilde{f}(x, y) \in \text{PV}$ and some $a \in M$. To prove the theorem we have to get rid of the parameter a . Choose $h(z) \in \text{PV}$ such that $\forall \text{PV}$ proves $h(\langle x, y \rangle) = \tilde{f}(x, y)$. We are left to show $M \models \forall z \exists n \text{LB}[h](|z| \cdot n^k, n)$.

For concreteness, let us code pairs $\langle n, m \rangle$ of numbers with binary expansions $a_0 \cdots a_{|n|-1}$ and $b_0 \cdots b_{|m|-1}$, respectively, by the number with binary expansion

$$a_0 a_0 \cdots a_{|n|-1} a_{|n|-1} 0 1 b_0 b_0 \cdots b_{|m|-1} b_{|m|-1}.$$

Argue in \mathbb{N} : given y and a size $\leq s$ circuit C one can easily construct a size $\leq s$ circuit computing $x \mapsto C(\langle x, y \rangle)$. Hence there is $c(x, y) \in \text{PV}$ such that $\forall \text{PV}$ proves

$$\text{Circuit}(C, s) \rightarrow \text{Circuit}(c(C, y), s) \wedge \text{eval}(C, \langle x, y \rangle) = \text{eval}(c(C, y), x).$$

Assume for contradiction that $M \not\models \forall z \exists n LB[h](|z| \cdot n^k, n)$. Choose $b \in M$ such that for all $n, m \in \text{Log}(M)$ there is $C \in M$ such that

$$\begin{aligned} M \models & \text{Circuit}(C, |b| \cdot (2n + 2 + 2m)^k) \\ & \wedge \forall x \forall y (|x| = n \wedge |y| = m \rightarrow \text{eval}(C, \langle x, y \rangle) = h(\langle x, y \rangle)). \end{aligned}$$

For suitable $e \in \mathbb{N}$ we have $(2n + 2 + 2m)^k \leq em^k n^k$ for all $n, m > 0$. Set $D := c^M(C, a)$ and, in M , choose $m := |a|$ (we can assume $|a| > 0$) and c of length $|c| \geq e|b|m^k$. Then

$$M \models \text{Circuit}(D, |c| \cdot n^k) \wedge \forall x (|x| = n \rightarrow \text{eval}(D, x) = h(\langle x, a \rangle)).$$

Since $M \models \forall x h(\langle x, a \rangle) = \tilde{f}_a(x)$ and $0 < n \in \text{Log}(M)$ is arbitrary, this contradicts (12). \square

Acknowledgements

We want to thank the anonymous referee for detailed comments and suggestions. We further thank Ján Pich for many helpful conversations about the topic of the current paper. The first author is currently partially supported by the Austrian Science Fund (FWF) under Project P31063. The second author is currently supported by the European Research Council (ERC) under the European Unions Horizon 2020 research programme (grant agreement ERC-2014-CoG 648276 AUTAR); the main part of the current work has been done while supported by the Austrian Science Fund (FWF) under Project P28699.

References

- [1] J. Avigad. Saturated models of universal theories. *Annals of Pure and Applied Logic* 118 (3): 219-234, 2002.
- [2] S. R. Buss. Bounded Arithmetic. Bibliopolis, Napoli. Revision of 1985 Princeton University Ph.D. thesis, 1986.
- [3] S. R. Buss. Bounded arithmetic and propositional proof complexity. In: H. Schwichtenberg (ed.), *Logic of Computation*, Springer, pp. 67-122, 1997

- [4] J. Bydžovský. Powers of Models in Weak Arithmetics. MSc. Thesis, University of Vienna, 2018. Available at <http://dmg.tuwien.ac.at/bydzovsky/mthesis.pdf> .
- [5] S. A. Cook. Feasibly constructive proofs and the propositional calculus. Proceedings of the seventh annual ACM sSymposium on Theory of Computing (STOC), ACM, pp. 83-97, 1975.
- [6] S. A. Cook and J. Krajíček. Consequences of the provability of $NP \subseteq P/poly$. Journal of Symbolic Logic 72 (4): 1353-1371, 2007.
- [7] A. Cobham. The intrinsic computational difficulty of functions. Proceedings of the 1964 International Congress for Logic, Methodology, and the Philosophy of Science, Y. Bar Hillel (ed.), North-Holland Publishing Co. Amsterdam, pp. 24-30, 1965.
- [8] M. Davis. Hilbert's tenth problem is unsolvable. The American Mathematical Monthly 80 (3): 233-269, 1973.
- [9] T. E. Frayne, A. C. Morel and D. S. Scott. Reduced direct products. Fundamenta Mathematicae 51/3: 195-228, 1962.
- [10] M. Garlík. Construction of models of bounded arithmetic by restricted reduced powers. Archive for Mathematical Logic 55: 625-648, 2016.
- [11] P. Hájek and P. Pudlák. Metamathematics of first order arithmetic. Springer/ASL Perspectives in Logic, 1993.
- [12] J. Hirschfeld. Models of arithmetic and recursive functions. Israel Journal of Mathematics 20 (2): 111-126, 1975.
- [13] E. Jeřábek. Approximate counting in bounded arithmetic. Journal of Symbolic Logic 72 (3): 959-993, 2007.
- [14] H. J. Keisler. On the class of limit ultrapowers of a relational system. Notices of the American Mathematical Society 7:878-879, 1960.
- [15] H. J. Keisler. Limit ultrapowers. Transactions of the AMS 107: 382-408, 1963.
- [16] H. J. Keisler. The ultraproduct construction. In V. Bergelson, A. Blass, M. Di Nasso, R. Jin (eds.), Ultrafilters Across Mathematics, Contemporary Mathematics 530, AMS, pp. 163-179, 2010.
- [17] S. B. Kochen. Ultraproducts in the theory of models. Annals of Mathematics 74 (2): 221-261, 1961.
- [18] S. B. Kochen and S. A. Kripke. Non-standard models of Peano arithmetic. L'Enseignement Mathématique 28: 211-231, 1982.

- [19] J. Krajíček. Bounded Arithmetic, Propositional Logic, and Complexity Theory. Encyclopedia of Mathematics and Its Applications 60, Cambridge University Press, 1995.
- [20] J. Krajíček. Interpolation theorems, lower bounds for proof systems, and independence results for bounded arithmetic. *Journal of Symbolic Logic* 62 (2): 457-486, 1997.
- [21] J. Krajíček. Extensions of models of PV. In: *Logic Colloquium'95*, J. A. Makowsky and E. V. Ravve (eds.), ASL/Springer Series Lecture Notes in Logic 11, pp.104-114, 1998.
- [22] J. Krajíček. On the weak pigeonhole principle. *Fundamenta Mathematicae* 170 (1-3): 123-140, 2001.
- [23] J. Krajíček. Forcing with random variables and proof complexity. *London Mathematical Society Lecture Note Series* 382, Cambridge University Press, 2011.
- [24] J. Krajíček and I. C. Oliveira. Unprovability of circuit upper bounds in Cook's theory PV. *Logical Methods in Computer Science* 13 (1:4), 2017.
- [25] J. Krajíček, P. Pudlák and G. Takeuti. Bounded arithmetic and polynomial hierarchy. *Annals of Pure and Applied Logic* 52: 143-154, 1991.
- [26] J. Łos. Quelques remarques, théorèmes et problèmes sur les classes définissables d'algèbres. *Mathematical Interpretations of Formal Systems*, North Holland, pp. 98-113, 1955.
- [27] R. Mac Dowell and E. Specker. Modelle der Arithmetik. In: *Infinitistic Methods. Proceedings of the Symposium on Foundations of Mathematics 1959*, Warsaw, Pergamon Press, pp. 257-263, 1961.
- [28] J. Maly and M. Müller. A remark on pseudo proof systems and hard instances of the satisfiability problem. *Mathematical Logic Quarterly* 64 (6): 418-428, 2018.
- [29] T. G. McLaughlin. Sub-arithmetical ultrapowers: a survey. *Annals of Pure and Applied Logic* 49 (2): 143-191, 1990.
- [30] M. Müller and J. Pich. Feasibly constructive proofs of succinct weak circuit lower bounds. Preprint at *Electronic Colloquium of Computational Complexity*, Technical Report TR17-144, 2017.
- [31] J. Paris and L. Harrington. A mathematical incompleteness in Peano arithmetic. *Handbook of Mathematical Logic*, North Holland, pp. 1133-1142, 1977.
- [32] P. Pudlák. *Logical Foundations of Mathematics and Computational Complexity*, a gentle introduction. Springer, 2013.
- [33] P. Pudlák. Randomness, pseudorandomness and models of arithmetic. *New Studies in Weak Arithmetics*, P. Cégielski and Ch. Cornaros (eds.), CSLI Publications, Stanford, pp.199-216, 2013.

- [34] A. A. Razborov. On provably disjoint NP-pairs. Basic Research in Computer Science BRICS RS-94-36, 1994.
- [35] A. A. Razborov. Bounded arithmetic and lower bounds in Boolean complexity. Feasible Mathematics II, pp. 344-386, 1995.
- [36] A. A. Razborov. Unprovability of lower bounds on the circuit size in certain fragments of bounded arithmetic. Izvestiya of the Russian Academy of Science, 59: 201-224, 1995.
- [37] A. A. Razborov, S. Rudich. Natural proofs. Journal of Computer and System Sciences 55 (1): 24-35, 1997.
- [38] R. Santhanam and R. Williams. On uniformity and circuit lower bounds. Computational Complexity 23 (2): 177-205, 2014.
- [39] D. Scott. On constructing models of arithmetic. In: Infinitistic Methods. Proceedings of the Symposium on Foundations of Mathematics 1959, Warsaw, Pergamon Press, p. 235-255, 1961.
- [40] S. Shelah. Every two elementarily equivalent models have isomorphic ultrapowers. Israel Journal of Mathematics 10: 224-233, 1971.
- [41] T. Skolem. Über die Nicht-charakterisierbarkeit der Zahlenreihe mittels endlich oder abzählbar unendlich vieler Aussagen mit ausschliesslich Zahlenvariablen. Fundamenta Mathematicae 23 (1): 150-161, 1934.
- [42] D. Zambella. Notes on polynomially bounded arithmetic. Journal of Symbolic Logic 61 (3): 942-966, 1996.