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 $\forall\mathsf{PV}$ of all polynomial time functions. Generalizing a theorem of Hirschfeld (1975), we show that every countable model of $\forall\mathsf{PV}$ 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 $\forall\mathsf{PV}$ we show that $\forall\mathsf{PV}$ 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 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. 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 $\forall\mathsf{PV}$, 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 established conjecture, whether it is provable in one of these theories or at least consistent

*Based on the first author's Master Thesis [4] written under the supervision of the second author.
[1] All relevant technical notions will be defined precisely later.

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 $\mathsf{S}^2_2(\alpha)$ 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^b_1-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" [31] formalization that we use in this paper. Some discussion can be found in [31] 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 $\forall$PV (see [6] for earlier conditional results under certain complexity theoretic assumptions). More precisely, for a Boolean (i.e., 0/1-valued) function $g(x) \in \mathsf{PV}$ let $LB[g](s,n)$ be a PV-formula with variables s, n that expresses

For all circuits C with $\leqslant n$ inputs and size $\leqslant s$ there is x of length n such that $g(x) \neq C(x)$.

Krajíček and Oliveira [24] proved that for every $k \in \mathbb{N}$ there is a Boolean $g(x) \in \mathsf{PV}$ such that for every $c \in \mathbb{N}$ the sentence $\exists n\ LB[g](c \cdot n^k, n)$ is consistent with $\forall$PV. We give a new proof that yields the following stronger result (cf. [24, Remark 1]):

Theorem 1.1. *For every $k \in \mathbb{N}$ there is a Boolean $g(x) \in \mathsf{PV}$ such that $\forall z \exists n\ LB[g](|z| \cdot n^k, n)$ is consistent with $\forall$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 $g(x) \in \mathsf{PV}$ 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 (witnessing) to extract from the proof a polynomial time function mapping n to a size $c \cdot n^k$ circuit C_n computing $g(x)$ on inputs of length n and thereby contradict Theorem 1.2. However, for $\forall$PV such witnessing is known only for Σ^b_1-formulas while $\neg LB[g](c \cdot n^k, n)$ is Σ^b_2. For Σ^b_2 the KPT-theorem [25] gives a witnessing function computable by a polynomial time 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^b_4. Our proof is model-theoretic and based on a *polynomial time ultrapower*.

Note one can apply polynomial time functions 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* $\mathsf{PTIME}(\mathbb{N})/U$ by identifying two functions that agree on some set in U. It is easy to see that polynomial ultrapowers satisfy $\forall\mathsf{PV}$. One can also start with a nonstandard model M of $\forall\mathsf{PV}$ instead of $\mathbb{N}$. As we shall see, to get something new, one must grant the polynomial time functions access to nonstandard parameters from M (see Theorem 2.23). We shall use a polynomial time ultrapower over a nonstandard model of $\forall\mathsf{PV}$ that is Herbrand saturated in the sense of [1].

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 it fragments [39, 26, 18, 12, 27, 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" the Π_2-fragment of true arithmetic "are composed" [12, p.112]. See Corollary 2.30 for a precise statement.

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 the following analogon of Hirschfeld's result:

Theorem 1.3. *A countable* PV*-structure is a model of* $\forall\mathsf{PV}$ *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.*

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

Corollary 1.4. *Every countable model of* $\forall\mathsf{PV}$ *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 the above as well as Hirschfeld's characterizations [12] and might be of some independent interest. The hope is that this description of the models of $\forall\mathsf{PV}$ helps understanding these models and thereby eventually understanding the status of complexity theoretic conjectures in $\forall\mathsf{PV}$ – 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 $\forall\mathsf{PV}$ 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 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 $\alpha, \beta, \dots$ range over F and ω over Ω. We say F is *closed under* W for an r-ary function $W : M^r \to M$ 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

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

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

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

We call $F/U/E$ *existentially closed*[2] 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).

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

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

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

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

Definition 2.6. Let $F \subseteq M^\Omega$ be L-closed and U an ultrafilter on Ω. An L-formula $\varphi(\bar{x})$ is *Łos 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 *Łos for F* means being Łos for (F, U) for every ultrafilter U, and being *Łos* means being Łos for F for every Ω and every L-closed $F \subseteq M^\Omega$.

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

Proof. It follows easily from (1) that atomic formulas are Łos, and Łos formulas are closed under Boolean combinations. □

[2]This 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.

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 Łos 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 Łos for (F,U), this implies $F/U \models \varphi(\bar{\alpha}^U)$ for all $\bar{\alpha} \in F$. □

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

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. View M and F/U as $L(M)$-structures with $a^{F/U} = \alpha_a^U$ for $a \in M$. Then apply Proposition 2.8. □

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, $W : M^r \to M$ is a *Skolem function* for $\exists y\varphi(x_0,\ldots,x_{r-1},y)$ if for all $\bar{a} \in M^r$ we have

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

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 Łos 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

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

The direction from right to left follows from $\varphi(\bar{x},y)$ being Łos for (F,U) and $\langle\!\langle\varphi(\bar{\alpha},\beta)\rangle\!\rangle \subseteq \langle\!\langle\exists y\varphi(\bar{\alpha},y\rangle\!\rangle$ for every β. For the converse note $\langle\!\langle\exists y\varphi(\bar{\alpha},y)\rangle\!\rangle = \langle\!\langle\varphi(\bar{\alpha},\beta)\rangle\!\rangle$ for $\beta := W \circ \bar{\alpha}$ where W is a Skolem function for $\exists y\varphi(\bar{x},y)$ such that F is closed under W.

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 2.8 as $\exists y\varphi(\bar{x},y)$ is Łos for (F,U). □

2.2 Examples and historical remarks

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 Łos [29] in 1955:

Example 2.12 (Łos'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 Łos for M^Ω. This is Łos's theorem [29].

According 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 Łos 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 Łos'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. Models of the form F/U for U an ultrafilter on $\mathbb{N}$ are *recursive ultrapowers* [12]. 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 [28] via some restricted ultrapower.

As mentioned in the Introduction, Hirschfeld studied recursive ultrapowers in order to characterize models of the true Π_2-theory of $\mathbb{N}$. McLaughlin [27] extends Hirschfeld's work to functions definable higher up in the arithmetical hierarchy.[3] 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 [26] 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.

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. □

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. 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 [26].

[3] In contrast to the setting of [18] and of this paper, [39, 33, 12, 27] consider ultrafilters over restricted Boolean algebras, namely those consisting of sets with characteristic function in F.

Proof. Let $(\alpha_0, a_0), (\alpha_1, a_1), \ldots$ enumerate $F \times M$ and define a sequence $M = X_0 \supseteq X_1 \supseteq \cdots$ of unbounded definable subsets of $\Omega = M$. For every (α_i, a_i) the function $\omega \mapsto \min\{\alpha_i(\omega), a_i\}$ is constant on an unbounded subset of X_i, say, equal to $b_i \leqslant a_i$ (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. We are left to verify that for all $\alpha \in F, a \in M$, 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 $X \cap X_{i+1} \subseteq \{\omega \mid \alpha(\omega) = b_i\} = \{\omega \mid \alpha(\omega) = \alpha_{b_i}(\omega)\} \in U$. □

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 $\forall\mathsf{PV}$. 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 $\forall\mathsf{PV}$, 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 $\forall\mathsf{PV}$ 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 $f : \mathbb{N}^r \to \mathbb{N}$ is such a function, then the arity of the symbol is r. As usual, $\mathsf{TIME}(n^k) \subseteq \mathsf{PV}$ 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

$$\forall\mathsf{PV}.$$

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 $0, 1, 2, \ldots$. The *length* $|n| := \lceil \log_2(n+1) \rceil$ is the length of the binary encoding of n; that is, $n = \sum_{i<|n|} 2^i \cdot bit(n,i)$ where $bit(n,i) := 0$ for $i > |n|$ and $bit(x,y) \in \mathsf{PV}$. The *smash* function is $n\#m := 2^{|n|\cdot|m|}$. every fixed $k \in \mathbb{N}$ there is $\langle x_0, \ldots, x_{k-1} \rangle \in \mathsf{PV}$ such that every $(n_0, \ldots, n_{k-1}) \in \mathbb{N}^k$ is *coded* by $n := \langle n_0, \ldots, n_{k-1} \rangle$; namely, we have $(n)_i = n_i$ where $(x)_0, (x)_1, \ldots$ are unary functions in PV.

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

Lemma 2.19. *For each* PV*-term* $t(\bar{x})$ *there is* $f(\bar{x}) \in \mathsf{PV}$ *such that* $\forall\mathsf{PV}$ *proves* $t(\bar{x}) = f(\bar{x})$. *For every quantifier-free* PV*-formula* φ *and* PV*-terms* $t(\bar{x}), s(\bar{x})$ *there is* $f(\bar{x}) \in \mathsf{PV}$ *such that* $\forall\mathsf{PV}$ *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 $\forall\mathsf{PV}$ we write

$$\mathsf{PTIME}(M) \ := \ T^M_{\mathsf{PV}(M)}.$$

Using sequence coding and the above lemma one sees that every $t^M(x) \in \mathsf{PTIME}(M)$ equals $f^M(x,a)$ for some $f(x,y) \in \mathsf{PV}$ (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) $f^M_a(x) \in \mathsf{PTIME}(M) \ldots$" stands for

"for all (there is) $f(x,y) \in \mathsf{PV}$ and for all (there is) $a \in M \ldots$"

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

A *limit polynomial time ultrapower over* M is a PV-structure of the form $\mathsf{PTIME}(M)/U/E$ 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 $\mathsf{PTIME}(M)$ can be restircted to some non-empty $\Omega \subseteq M$, i.e., take for F the set

$$\mathsf{PTIME}(M)\restriction\Omega := \{\alpha\restriction\Omega \mid \alpha \in \mathsf{PTIME}(M)\}.$$

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 $(\alpha\restriction\Omega)^U \mapsto \alpha^V$ is an isomorphism from $(\mathsf{PTIME}(M)\restriction\Omega)/U$ onto $\mathsf{PTIME}(M)/V$.

By Propositions 2.8 and 2.10:

Proposition 2.22. *Let* M *be a model of* $\forall\mathsf{PV}$. *Every polynomial time ultrapower over* M *is a model of* $\forall\mathsf{PV}$ *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* $\forall\mathsf{PV}$. *For every ultrafilter* U *over* M *there is an ultrafilter* V *over* $\mathbb{N}$ *such that* $T^M_{\mathsf{PV}}/U \cong \mathsf{PTIME}(\mathbb{N})/V$.

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. $\forall\mathsf{PV}$ *is weakly sequential.*

Proof. Enumerate a countable $M \models \forall\mathsf{PV}$ by $m_0, m_1, \ldots$ 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 $b := \langle m_0, \ldots, m_{k-1}\rangle^M$ is as required. □

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

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

Theorem 1.3 follows setting $L := \mathsf{PV}$ and $M := \mathbb{N}$. Corollary 1.4 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 $\mathit{Th}_\forall(M)$ generated by one point. Then there exists an ultrafilter U on M such that N is isomorphic to T^M_L/U.*

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

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

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

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

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 Łos by Proposition 2.7. □

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 $t_0(x), \ldots, t_{k-1}(x)$ with $(t_0^M)^U, \ldots, (t_{k-1}^M)^U \in N$ such that

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

Using Proposition 2.7 we get

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

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

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

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

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

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

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

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. □

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 Łos 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. □

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

For completeness we show how to derive the main result of [12]. Temporarily let $\mathbb{N}$ denote the standard model of arithmetic in the language $\{\cdot, +, <, 0, 1\}$ (instead PV). Hirschfeld considers the Π_2-fragment of true arithmetic. It follows from the MRDP theorem [8] that this theory is equivalent to $Th_{\forall\exists}(\mathbb{N})$, the set of $\forall\exists$-sentences true in $\mathbb{N}$ (see [12, Corollary 1.7.1.b]). Recall Example 2.15 for a definition of recursive ultrapowers; *limit* recursive ultrapowers are explained in the obvious way.

Corollary 2.30 (Hirschfeld 1975)**.** *A countable $\{\cdot, +, <, 0, 1\}$-structure is a model of $Th_{\forall\exists}(\mathbb{N})$ if and only if it is isomorphic to an existentially closed limit recursive ultrapower.*

Proof. Let $\mathbb{N}^*$ be the standard model of arithmetic in the language consisting of $<$ and symbols for all recursive functions. By the MRDP theorem every such function is definable in $\mathbb{N}$ by an existential formula. Obviously, $Th_{\forall\exists}(\mathbb{N})$ proves the functionality of this formula. Hence every model of $Th_{\forall\exists}(\mathbb{N})$ 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, we already noted in Example 2.15 that recursive ultrapowers model $Th_{\forall\exists}(\mathbb{N})$. But existentially closed substructures inherit the $\forall\exists$-theory of their superstructures. □

Remark 2.31. 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. The same holds true for Corollary 2.30 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 $\forall\mathsf{PV}$, simplified using Lemma 2.19.

Theorem 3.2. *Let M be a Herbrand saturated model of $\forall\mathsf{PV}$. Assume $M \models \forall x\exists y\varphi(\bar{x}, y)$ where $\varphi(\bar{x}, y)$ is a quantifier free $\mathsf{PV}(M)$-formula. Then there exists $f^M_a(\bar{x}) \in \mathsf{PTIME}(M)$ 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 $\alpha \in \mathsf{PTIME}(M)$ given an ultrapower $\mathsf{PTIME}(M)/U$: as a point α^U in the structure $\mathsf{PTIME}(M)/U$ or as a function $\alpha : M \to M$ in M. See [23, Section 24.4] or [30] for a use of these views in the context of forcing with random variables.

Assume $\mathsf{PTIME}(M)/U \models \forall \ell \neg LB[g](small, \ell)$. Plug $|id_M^U|$ for ℓ and choose $\zeta \in \mathsf{PTIME}(M)$ such that ζ^U is in $\mathsf{PTIME}(M)/U$ 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 $\mathsf{PTIME}(M)$. 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 $\gamma \in \mathsf{PTIME}(M)$ 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 $\mathsf{PTIME}(M)/U$ 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 $\forall\mathsf{PV}$. An element $a \in M$ *codes* the set $\{b \in M \mid bit(a, b) = 1\}$. Here and below we often omit superscripts writing e.g. *bit* instead bit^M. Coded sets are subsets of $Log(M)$, i.e., the set of $n \in M$ such that $n = |N|$ for some $N \in M$; for such n we write 2^n for $1\#N$ and 1^n for $2^n - 1$ (the number coding the string coding n in unary). We shall use the phrase "for all large enough $n \in Log(M) : \ldots n \ldots$" for

"there is $n_0 \in Log(M)$ such that for every $n \in Log(M)$ with $n_0 < n : \ldots n \ldots$"

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) $\leqslant s$ with $\leqslant n$ inputs by a number coding the set of tuples

$$\langle 1^n, a, b, c \rangle \tag{3}$$

where $a, b < s$ are (numbers of) gates and either a is wired to gate b and c specifies the label $\wedge, \vee$ or $\neg$ of b, or $a = b$ is the c-th input gate. Note the coding depends on the choice of n which will always be clear from context. There are $O(s)$ such tuples, each of length $O(n + |s|)$. It is convenient to allow 0 as a code of a circuit with 0 inputs and size 0. We blur the distinction of C and its code, and note

$$n \leqslant |C| \leqslant O(n \cdot s \cdot |s|). \tag{4}$$

The set of triples (C, s, n) such that C is a circuit of size $\leqslant s$ with $\leqslant n$ inputs is in the standard PV-model $\mathbb{N}$ defined by a quantifier free formula $Circuit(x, y, z)$. By convention,

$$Circuit(0,0,0) \tag{5}$$

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

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

$$LB[f](s, n) \;\; := \;\; \forall C\big(Circuit(C, s, n) \to \exists x(|x| = n \wedge f(x) \neq eval(C, x))\big).$$

For readability we use C, n as variables. Given a model M we can plug a function $f_a(x)$ for $f(x)$ (see Section 2.3) and get a formula $LB[f_a](s, n)$ with parameter a.

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 $g_d(x) \in \mathsf{TIME}(n^{d+1})$ such that for every $h(x, y) \in \mathsf{TIME}(n^d)$ there is $c_h \in \mathbb{N}$ such that $\forall\mathsf{PV}$ proves*

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

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

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

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

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

$$M \models \forall z \exists n LB[\tilde{f}_a](|z| \cdot n^k, n) \tag{6}$$

or for all large enough $n \in Log(M)$

$$M \models \exists x\big(|x| = n \wedge g(x) \neq eval(f_a(1^n), x)\big).$$

Proof. Choose $g(x) := g_{3k-1}(x)$ as in the previous lemma and let f_a^M be as stated. It suffices to show that, if (6) fails for suitable $\tilde{f}_a^M$ chosen below, then there is $h(x,y) \in \mathsf{TIME}(n^{3k-1})$ such that for all large enough $n \in \mathit{Log}(M)$ there is $D \in M$ with $|D| = n^{2/3}$ such that

$$M \models \forall x\big(|x| = n \to h(x, D) = \mathit{eval}(f_a(1^n), x)\big). \tag{7}$$

Argue in M. The circuits $f_a(1^n)$ are coded by the (numbers coding the) tuples (3) for $a, b < n^{k+\epsilon}, c < n$. They could be coded also by the *succinct tuples* $\langle n, a, b, c\rangle$ of length $O(|n|)$. For $n \in \mathit{Log}(M)$ larger than some standard constant, each succinct tuple can be *padded* to a number of length exactly $n^{1/(2k)}$. The details are unimportant, we only need $\mathsf{TIME}(n)$ functions that compute such padded tuples from a, b, c and 1^n, and recover the original from the padded version. Then it is easy to see that the characteristic function of these padded tuples has the form $\tilde{f}_a^M(x) \in \mathsf{PTIME}(M)$ (with the same parameter a).

Now assume (6) fails, so there are $b \in M$ and for every $n \in \mathit{Log}(M)$ and $m := n^{1/(2k)}$ a circuit D_m with $\leqslant m$ inputs and size $\leqslant |b| \cdot n^{1/2}$ that compute $\tilde{f}_a(x)$ on inputs x of length m. Then $|D_m| \leqslant n^{2/3}$ for large enough $n \in \mathit{Log}(M)$. Indeed, for some $e \in \mathbb{N}$ we have in M

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

the first inequality holds by (4), the second holds for $n \geqslant |b|^{100}$, 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 $a, b < n^{k+\epsilon}$ and $c < n$ compute $\mathit{eval}(D, u)$ where u is the padded (to length $m = n^{1/(2k)}$) version of the succinct tuple $\langle n, a, b, c\rangle$; define C as the (number coding the) set containing for every such u with $\mathit{eval}(D, u) = 1$ the tuple $\langle 1^n, a, b, c\rangle$; finally, output $\mathit{eval}(C, x)$.

Then $h(x,y) \in \mathsf{TIME}(n^{3k-1})$. Indeed, its time is dominated by the evaluations $\mathit{eval}(D, u)$. There are $(n^{k+\epsilon})^2 \cdot n$ many of them and each needs time $O((|u|+|D|)^2) \leqslant O(n^{4/3})$ (plus $O(n)$ for the computation of u). But $n^{2k+2\epsilon+1+4/3} \leqslant n^{3k-1}$ 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 (7). □

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 $\forall\mathsf{PV}$. Suppose $g(x) \in \mathsf{PV}$ is Boolean and such that for every uniform size $n^{k+\epsilon}$ circuit family $f_a^M \in \mathsf{PTIME}(M)$ we have for all large enough $n \in \log(M)$:*

$$M \models \exists x\big(|x| = n \wedge g(x) \neq \mathit{eval}(f_a(1^n), x)\big). \tag{8}$$

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

Proof. Let

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

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

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

Let $\alpha \in \mathsf{PTIME}(M)$, a value for z, be given, say $\alpha = z^M_{a_0}$ for $z(x,y) \in \mathsf{PV}$ 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 $M \models |z_{a_0}(1^n)| \leqslant n^s$ for all large enough $n \in \mathit{Log}(M)$ (with threshold depending on $|a_0|$). Then set $t := \lceil s/\epsilon \rceil \in \mathbb{N}$ and choose $\ell(x) \in \mathsf{PV}$ such that $\ell(1^n) = 1^{n^t}$ for all $n \in \mathit{Log}(M)$. We then have

$$M \models |z_{a_0}(1^n)| \leqslant |\ell(1^n)|^\epsilon \tag{9}$$

for all large enough $n \in \mathit{Log}(M)$. We set $\lambda := \ell^M(x)$ and aim to show that

$$(\mathsf{PTIME}(M)\upharpoonright\Omega)/U \models LB[g](|\alpha^U| \cdot |\lambda^U|^k, |\lambda^U|).$$

Let $\zeta \in \mathsf{PTIME}(M)$ be such that

$$(\mathsf{PTIME}(M)\upharpoonright\Omega)/U \ \models\ \mathit{Circuit}(\zeta^U, |\alpha^U| \cdot |\lambda^U|^k, |\lambda^U|). \tag{10}$$

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

$$(\mathsf{PTIME}(M)\upharpoonright\Omega)/U \ \models\ |\gamma^U| = |\lambda^U| \wedge g(\gamma^U) \neq \mathit{eval}(\zeta^U, \gamma^U). \tag{11}$$

Let $\zeta = C^M_{a_1}(x)$ for $C(x,y) \in \mathsf{PV}$ and $a_1 \in M$. By Proposition 2.7, (10) implies that

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

By Lemma 2.19 we find $\tilde{C}(x,y) \in \mathsf{PV}$ such that for $a_2 := \langle a_0, a_1 \rangle$ (computed in M) we have for all $n \in \mathit{Log}(M)$

$$M \models \tilde{C}_{a_2}(1^n) = \begin{cases} C_{a_1}(1^{n^{1/t}}) & \text{if } \mathit{Circuit}\big(C_{a_1}(1^{n^{1/t}}), n^{k+\epsilon}, n\big) \\ 0 & \text{else.} \end{cases} \tag{12}$$

By our convention (5), $C^M_{a_1}(x)$ is a uniform sequence of size $n^{k+\epsilon}$ circuits in M. By assumption (8) we have for all large enough $n \in \mathit{Log}(M)$:

$$M \models \exists x\big(|x| = n \wedge g(x) \neq \mathit{eval}(\tilde{C}_{a_2}(1^n), x)\big) \tag{13}$$

It follows from Theorem 3.2 that there exists a function $w^M_{a_3}(x) \in \mathsf{PTIME}(M)$ such that for all large enough $n \in \mathit{Log}(M)$:

$$M \models |w_{a_3}(1^n)| = n \wedge g(w_{a_3}(1^n)) \neq \mathit{eval}(\tilde{C}_{a_2}(1^n), w_{a_3}(1^n)) \tag{14}$$

Indeed, by (13) there is $b \in Log(M)$ such that

$$M \models \forall u \exists x \big(b < |u| \to |x| = |u| \wedge g(x) \neq eval(\tilde{C}_{a_2}(1^{|u|}), x)\big)$$

and Herbrand saturation gives $w^M_{a_3}(u) \in \mathsf{PTIME}(M)$ witnessing $\exists x$ as in Theorem 3.2.

Choose $\tilde{w}(x,y) \in \mathsf{PV}$ such that $\forall\mathsf{PV}$ proves $\tilde{w}(x,y) = w(\ell(x),y)$, so $M \models \tilde{w}_{a_3}(1^n) = w_{a_3}(1^{n^t})$ for all $n \in Log(M)$. Define

$$\gamma := \tilde{w}^M_{a_3}(x).$$

We are left to verify (11). Plugging n^t for n in (14) gives

$$M \models |\tilde{w}_{a_3}(1^n)| = |\ell(1^n)| \wedge g(\tilde{w}_{a_3}(1^n)) \neq eval(\tilde{C}_{a_2}(\ell(1^n)), \tilde{w}_{a_3}(1^n))$$

for all large enough $n \in Log(M)$. Since U is unbounded, this implies for $\tilde{\zeta} := \tilde{C}^M_{a_2}(\ell^M(x))$

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

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

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

for all $m \in Log(M)$ large enough such that (9) holds; since U is unbounded, it contains the set of $1^m \in X$ such that m is large enough such that (9) holds; for 1^m in this set we have $M \models Circuit(C_{a_1}(1^m), (m^t)^{k+\epsilon}, m^t)$ and thus $\tilde{C}^M_{a_2}(1^{m^t}) = C^M_{a_1}(1^m)$ by (12). □

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

If (8) holds for every uniform size $n^{k+\epsilon}$ circuit family $f^M_a \in \mathsf{PTIME}(M)$, then Lemma 3.6 states that $\forall z \exists n LB[g](|z| \cdot n^k, n)$ holds in some polynomial time ultrapower over M, so is consistent with $\forall\mathsf{PV}$ (Proposition 2.22).

Otherwise there is a uniform size $n^{k+\epsilon}$ circuit family $f^M_a \in \mathsf{PTIME}(M)$ such that (8) fails. Then Lemma 3.5 implies

$$M \models \forall z \exists n LB[\tilde{f}_a](|z| \cdot n^k, n) \tag{15}$$

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

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

$$a_0 a_0 \cdots a_{|n|-1} a_{|n|-1} \; 01 \; b_0 b_0 \cdots b_{|m|-1} b_{|m|-1}.$$

Argue in $\mathbb{N}$: given y of length m and a size $\leqslant s$ circuit C with $2n+2+2m$ inputs one can easily construct a size $\leqslant s$ circuit computing $x \mapsto C(\langle x,y\rangle)$ on inputs x of length n.

It follows that there is a function $c(x, y) \in \mathsf{PV}$ such that $\forall\mathsf{PV}$ proves

$$|x| = n \wedge |y| = m \wedge \mathit{Circuit}(C, s, 2n + 2 + 2m) \rightarrow \\ \mathit{Circuit}(c(C, y), s, n) \wedge \mathit{eval}(C, \langle x, y\rangle) = \mathit{eval}(c(C, y), x).$$

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 $n, m \in \mathit{Log}(M)$ there is $C \in M$ such that

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

For suitable $e \in \mathbb{N}$ we have $(2n + 2 + 2m)^k \leqslant em^k n^k$ 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 $|c| \geqslant e|b|m^k$. Then

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

Since $M \models \forall x\ h(\langle x, a\rangle) = \tilde{f}_a(x)$ and $0 < n \in \mathit{Log}(M)$ is arbitrary, this contradicts (15). □

Acknowledgements

The authors thank Jan 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 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 of Symbolic Logic 72 (4): 1353-1371, 2007.

[7] A. Cobham. The instrinsic computational difficulty of functions. Proceedings of the 1964 International Congress for Logic, Methodology, and the Philosophy of Science, Y. Bar Hillel (ed.), North-Holland Publising 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. Bounded 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] 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.

[27] T. G. McLaughlin. Sub-arithmetical ultrapowers: a survey. Annals of Pure and Applied Logic 49 (2): 143-191, 1990.

[28] J. Paris and L. Harrington. A mathematical incompleteness in Peano arithmetic. Handbook of Mathematical Logic, North Holland, pp. 1133-1142, 1977.

[29] J. Łos. Quelques remarques, théorèmes et problèmes sur les classes définissables d'algèbres. Mathematical Interpretations of Formal Systems, North Holland, pp. 98-113, 1955.

[30] J. Maly and M. Müller. A remark on pseudo proof systems and hard instances of SAT. To appear in Mathematical Logic Quarterly.

[31] M. Müller and J. Pich. Feasibly constructive proofs of succinct weak circuit lower bounds. Preprint at Electronic Colloqium of Computational Complexity, Technical Report TR17-144, 2017.

[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.