Electronic Colloquium on Computational Complexity, Report No. 149 (2013)

The Ordering Principle in a Fragment of Approximate Counting

Albert Atserias[*] Neil Thapen[†]

October 28, 2013

Abstract

The ordering principle states that every finite linear order has a least element. We show that, in the relativized setting, the surjective weak pigeonhole principle for polynomial time functions does not prove a Herbrandized version of the ordering principle over T^1_2. This answers an open question raised in [Buss, Kołodziejczyk and Thapen, 2012] and completes their program to compare the strength of Jeřábek's bounded arithmetic theory for approximate counting with weakened versions of it.

1 Introduction

We show that, in the relativized setting, the surjective weak pigeonhole principle for polynomial time functions does not prove the Herbrandized ordering principle over T^1_2. This answers an open question from [2]. In the rest of this section we will give a brief introduction to this problem. We will assume a basic knowledge of the language and theories of bounded arithmetic; standard references are [1] and [10].

The *Herbrandized ordering principle* HOP is a formula in the vocabulary $\alpha = (\prec, h)$, where $\prec$ is a binary relation symbol and h is a unary function symbol. It asserts that if $\prec$ is a strict linear ordering of an interval $[n] :=$

[*] Departament de Llenguatges i Sistemes Informàtics. Universitat Politècnica de Catalunya, Barcelona, Spain (atserias@lsi.upc.edu). Research partially supported by TIN2010-20967-C04-05 (TASSAT).

[†] Institute of Mathematics. Academy of Sciences of the Czech Republic, Prague, Czech Republic (thapen@math.cas.cz). Research partially supported by grant IAA100190902 of GA AV ČR, and by Center of Excellence CE-ITI under grant P202/12/G061 of GA ČR and RVO: 67985840.

ISSN 1433-8092

$\{0, \ldots, n-1\}$ and h maps $[n]$ into $[n]$, then there is some $x \in [n]$ such that $h(x)$ is not the immediate predecessor of x. In other words, either there exists a witness that $\prec$ is not a strict linear ordering of $[n]$, or there exists $x \in [n]$ such that $h(x) \not\prec x$, or there exist $x, y \in [n]$ such that $h(x) \prec y \prec x$. The principle is expressed by a $\Sigma^b_1(\alpha)$ formula $\exists z < n^3\, (\theta(z, n))$ where z codes a possible witness (the biggest of which would be a triple witnessing that $\prec$ is not transitive) and θ is a quantifier-free $\mathrm{PV}(\alpha)$ formula. We note that the natural Herbrandization of the ordering principle would not contain the last condition, about h giving the immediate predecessor. As in [2], including this condition is convenient and also makes the principle weaker, and hence makes our unprovability result stronger. Similar principles to HOP (without this condition) have appeared as the *generalized iteration principle* in [4] and as *Herbrandized minimization* in [5].

To define the surjective weak pigeonhole principle sWPHP, we first introduce some notation: for a function $f(\bar{z}, x)$ of several arguments, we will sometimes treat some of the arguments as parameters and write them as subscripts, writing, for example, $f_{\bar{z}}(x)$ for f considered as a family of one-argument functions parametrized by $\bar{z}$.

Given a function symbol f, the formula $\mathrm{sPHP}^a_b(f)$ expresses that, if $b > a$, then f is not a surjection from $[a]$ onto $[b]$. We take the principle $\mathrm{sWPHP}(\mathrm{PV}(\alpha))$ to be the formula

$$\forall a\, \forall e\ \left(\mathrm{sPHP}^a_{a^2}(g_e)\right)$$

where $g(e, x)$ is a "universal" polynomial time function (with oracle α) which we can think of as, for example, evaluating the Boolean circuit e on input x (where e is allowed "oracle gates" for computing queries to α). The principle expresses that no $\mathrm{PV}(\alpha)$ function with parameters is a surjection from $[a]$ onto $[a^2]$, for any $a > 1$.

It is a long-standing open problem to separate Buss' hierarchy T^i_2 of bounded arithmetic theories by sentences of fixed complexity. This is unknown even for the relativized hierarchy $\mathrm{T}^i_2(\alpha)$ (although there is such a separation known for the hierarchy $\mathrm{T}^i_1(\alpha)$, which has polynomial rather than quasipolynomial growth rate [6]). Techniques exist to separate $\mathrm{PV}(\alpha)$ from $\mathrm{T}^1_2(\alpha)$, and $\mathrm{T}^1_2(\alpha)$ from $\mathrm{T}^2_2(\alpha)$, by $\forall\Sigma^b_1(\alpha)$ sentences [4]. But these do not seem to be useful in the case of $\mathrm{T}^2_2(\alpha)$ and $\mathrm{T}^3_2(\alpha)$.

The recent paper [2] tries to approach this problem from a different direction by considering, rather than T^2_2, Jeřábek's theory $\mathrm{T}^1_2 + \mathrm{sWPHP}(\mathrm{PV}_2)$ of approximate counting [9]. Here PV_2 stands for a set of terms naming all $\mathrm{FP}^{\mathrm{NP}}$ functions. This theory is called APC_2 in [2], and sits at a similar level in the hierarchy to T^2_2. The authors do not show any separation of APC_2 from anything higher, but do give $\forall\Sigma^b_1(\alpha)$ separations of certain subtheories of $\mathrm{APC}_2(\alpha)$ from $\mathrm{T}^2_2(\alpha)$ and $\mathrm{APC}_2(\alpha)$ itself. In particular, they observe that HOP is provable in both $\mathrm{APC}_2(\alpha)$ and $\mathrm{T}^2_2(\alpha)$ and show, among other things, that $\mathrm{PV}(\alpha) + \mathrm{sWPHP}(\mathrm{PV}_2(\alpha)) \nvdash \mathrm{HOP}$.[1]

Our result is that

$$\mathrm{T}^1_2(\alpha) + \mathrm{sWPHP}(\mathrm{PV}(\alpha)) \nvdash \mathrm{HOP}.$$

This resolves an open question in [2] and shows that weakening $\mathrm{APC}_2(\alpha)$, either by reducing the amount of induction from $\Sigma^b_1(\alpha)$ to $\Sigma^b_0(\alpha)$ (that is, from $\mathrm{T}^1_2(\alpha)$ to $\mathrm{PV}(\alpha)$, see [7]), or by reducing the functions for which sWPHP holds from $\mathrm{PV}_2(\alpha)$ to $\mathrm{PV}(\alpha)$, gives a strictly weaker theory.

In Section 2 below we give the high-level proof of our result, and in Sections 3 and 4 we prove some necessary technical lemmas. The most important of these is Lemma 3, which shows how decision trees computing functions in $\mathrm{PV}(\alpha)$ are simplified under a certain random restriction. In Section 5 we discuss a propositional version of our result.

We are grateful to Emíl Jeřábek and Leszek Kołodziejczyk for helpful comments on earlier versions of this work.

2 Main theorem

We begin with some standard manipulations.

Lemma 1. *Suppose $\phi(n)$ is a $\Sigma^b_1(\alpha)$ formula and*

$$\mathrm{T}^1_2(\alpha) + \mathrm{sWPHP}(\mathrm{PV}(\alpha)) \vdash \forall n\,(\phi(n)).$$

Then there is a term $t = t(n)$ and a function symbol $f \in \mathrm{PV}(\alpha)$ such that

$$\mathrm{T}^1_2(\alpha) \vdash \forall n\,(t > 2 \wedge \forall v{<}t^2\, \exists u{<}t\,(f_n(u) = v \vee \phi(n))).$$

[1] There is a technical issue here concerning our definition of sWPHP. There are three versions considered in the paper [2]: that there is no surjection from $[a]$ onto $[a(1+1/|a|)]$; from $[a]$ onto $[2a]$; or from $[a]$ onto $[a^2]$. The notation sWPHP is used formally in [2] only for the first (and strongest) version, while we, for the sake of simplicity, use it to mean the third (and weakest). For our result, as in most cases, this makes no difference, because the three versions are equivalent over $\mathrm{S}^1_2(\alpha)$ for $\mathrm{PV}(\alpha)$ functions. However in the result from [2] referred to here it does matter which version is used, because for $\mathrm{PV}_2(\alpha)$ functions they are unlikely to be equivalent over $\mathrm{PV}(\alpha)$ (see [8]). Precisely, the result is for $[a]$ onto $[2a]$, and hence also for $[a]$ onto $[a^2]$. It is not known for $[a]$ onto $[a(1+1/|a|)]$.

Proof. This is proved by standard tricks about amplifying failures of the weak pigeonhole principle (see for example [13]). In detail, we are given a $\mathrm{PV}(\alpha)$ function symbol $g(e,x)$ such that

$$\mathrm{T}^1_2(\alpha) + \forall a\, \forall e\, \left(\mathrm{sPHP}^a_{a^2}(g_e)\right) \vdash \forall n\, (\phi(n)).$$

By Parikh's theorem, there is a term $p = p(n)$ such that

$$\mathrm{T}^1_2(\alpha) \vdash \forall n\, \exists a{<}p\, \exists e{<}p\, \left(\neg \mathrm{sPHP}^a_{a^2}(g_e) \vee \phi(n)\right).$$

We may assume without loss of generality that $\mathrm{T}^1_2(\alpha)$ proves $p > 1$. By Corollary 2.2 of [13] there is a $\mathrm{PV}(\alpha)$ function symbol $G(e,a,b,x)$ such that, in any model of $\mathrm{S}^1_2(\alpha)$ (and in particular any model of $\mathrm{T}^1_2(\alpha)$), if g_e is a surjection from a onto a^2 then $G_{e,a,b}$ is a surjection from a onto b. Let $f(n,u)$ be the function

$$f : (n, (e,a,x)) \mapsto G(e,a,p^6,x)$$

where f interprets its second argument u as a triple (e,a,x) of numbers, each in $[p]$. Then whenever there exist e and a in $[p]$ such that g_e is a surjection from a onto a^2, we also have that f_n is a surjection from $[p^3]$ onto $[p^6]$. The lemma follows by putting $t = p^3$. □

Theorem 1. $\mathrm{T}^1_2(\alpha) + \mathrm{sWPHP}(\mathrm{PV}(\alpha)) \nvdash \mathrm{HOP}$.

Proof. We assume the opposite to reach a contradiction. By Lemma 1 we have a term $t = t(n)$ and function symbol $f \in \mathrm{PV}(\alpha)$ such that

$$\mathrm{T}^1_2(\alpha) \vdash \forall n\, (t > 2 \wedge \forall v{<}t^2\, \exists u{<}t\, \exists z{<}n^3\, (f_n(u) = v \vee \theta(z,n))) \qquad (1)$$

where w and $\theta(z,n)$ are the bounding term and the quantifier-free $\mathrm{PV}(\alpha)$-formula, respectively, from the $\Sigma^b_1(\alpha)$-formula expressing HOP.

By [3], this is witnessed by a PLS problem[2], as follows. The problem is given by a term $s = s(v,n)$, a cost function C and a neighborhood function N on $[s]$, where C and N take n and v as parameters, run in time polynomial in $|n|$, and have oracle access to $\prec$ and h. A solution to the problem is a number $x \in [s]$ such that $C(N(x)) \geq C(x)$. There is a polynomial time reduction function g (which does not access the oracles) such that for all

[2] Our version of the PLS witnessing theorem is slightly different from the one that appears in [3]. However their PLS problem (F_L, c_L, N_L) is easily reducible to ours, by putting $C(x) = c_L(x)$ and $N(x) = N_L(x)$ for $x \in F_L$, and putting $C(x) = q + 1$ and $N(x) = 0$ for $x \notin F_L$, where q is an upper bound on the cost in their instance.

choices of oracles $\prec$ and h, for all n and v and all $x \in [s]$, if x is a solution to the problem then $g(x)$ is a witness $\langle u, z\rangle$ for the existential quantifiers on the right-hand side of (1).

Before we continue we need some definitions. Let n, p and q be positive integers such that q divides p and $p/q < n - p$. Let $m = p/q$. A *random restriction* ρ with these parameters is a partition of $[n]$ into linearly ordered sets chosen randomly as follows:

1. choose a random set $B_0 \subseteq [n]$ of cardinality $n - p$,
2. randomly partition $[n] \setminus B_0$ into *blocks* $B_1, \dots, B_q$ of cardinality m,
3. choose a random linear ordering $\prec_i$ of each B_i for $i \in \{0, \dots, q\}$.

The conditions on n, p and q imply that $m < n - p$. Consequently we will call B_0 the *big block* and the other blocks *small blocks*. For every $x \in [n]$, we write B^x for the unique block that contains x. Let $\mathcal{R}(n, p, q)$ be the set of all restrictions ρ with parameters n, p and q as above. If $\rho = (B_0, \dots, B_q, \prec_0, \dots, \prec_q)$ is a restriction in $\mathcal{R}(n, p, q)$ with blocks $B_0, \dots, B_q$ and linear orderings $\prec_0, \dots, \prec_q$, we say that a total linear ordering $\prec$ of $[n]$ is *compatible* with ρ if it satisfies three conditions:

1. $\prec$ extends $\prec_i$ for every $i \in \{0, \dots q\}$,
2. B_i is $\prec$-convex[3] for every $i \in \{0, \dots, q\}$, and
3. $x \prec y$ for every $x \in [n] \setminus B_0$ and every $y \in B_0$.

Notice that there are always exactly $q!$ total linear orderings compatible with ρ, corresponding to the $q!$ possible ways of arranging the small blocks $B_1, \dots, B_q$ below the big block B_0.

We continue with the proof. Let n_0 be a large integer, and let $n \geq n_0$ be an exact eighth power, so that $p := n^{1/2}$ and $q := n^{1/8}$ are both integers. Note that $m := p/q = n^{3/8}$ is also an integer. Given a total linear ordering $\prec$ of $[n]$ let $h_\prec$ be the predecessor function arising from $\prec$, except that $h(z) = z$ for the $\prec$-minimum element $z \in [n]$, and also $h(z) = z$ for every $z \notin [n]$. We will call oracles $(\prec, h_\prec)$ of this form *standard*. We say that such an oracle is compatible with a restriction ρ if $\prec$ is compatible with ρ.

Lemma 2. *There is a restriction $\rho \in \mathcal{R}(n, p, q)$ with n, p and q as specified, and a number $v \in [t^2]$, such that for every $u \in [t]$ we have $f_n(u) \neq v$ under every standard oracle $(\prec, h_\prec)$ compatible with ρ.*

[3] If $(S, <)$ is a linearly ordered set, we say that a subset $C \subseteq S$ is $<$-*convex* if whenever x and y belong to C and z in S is such that $x < z$ and $z < y$, then also z belongs to C.

The proof of this lemma takes up Sections 3 and 4 of this paper. We show now how we use it to obtain a contradiction.

Let ρ and v be given by the lemma. Let $|n|^k$ be a bound on the number of oracle queries and replies that can occur in a computation of the cost or neighbourhood functions C or N. For large enough n, we may assume that $3|n|^k + 3 < q$. Let A be the set of partially-defined oracles $\alpha = (\prec, h)$ arising in the following way. Choose $\ell \leq |n|^k$ of the small blocks of ρ and arrange them in any order as $B_{i_1}, \ldots, B_{i_\ell}$. Let the domain of $\prec$ be the union of all these blocks together with B_0, and let $\prec$ be a total linear ordering of this domain, with the ordering inside each block given by ρ and the ordering between blocks given by $B_{i_1} \prec \cdots \prec B_{i_\ell} \prec B_0$. Let h be defined everywhere on this domain except for its $\prec$-minimum element, as the predecessor function arising from $\prec$.

Let M be the set of pairs (x, α) of $x \in [s]$ and $\alpha \in A$ for which the cost $C^\alpha(x)$ of x under α is defined. We claim that M is non-empty, and that for any $(x, \alpha) \in M$, there is $(y, \beta) \in M$ such that $C^\beta(y) < C^\alpha(x)$. Together these imply a contradiction, since costs must be positive.

To see that M is non-empty, we simulate a computation of $C(0)$, constructing a partial oracle α as we go. At the beginning of the simulation, we set α to be the ordering $\prec$ given by ρ on B_0 and undefined elsewhere, with the corresponding predecessor function h defined on B_0 without its $\prec$-minimum element. Each time a query is made about any element z currently outside the domain of $\prec$, we add the block B^z containing z to the bottom of our ordering $\prec$ and extend h appropriately, in particular setting $h(w)$ to be w', where w is the minimum element of our current ordering and w' is the maximum element of the new block. If $h(z)$ is queried where z is currently the $\prec$-minimum element, we take any unused block and similarly add it to the bottom of the ordering. We add at most $|n|^k < q$ blocks over the course of the simulation, so never run out of unused blocks to add in this second case.

Given (x, α) in M, to find a suitable (y, β) in M we simulate a computation of $N(x)$, save this value as y, and then simulate a computation of $C(y)$, all using a partial oracle γ which we construct as we go. At the beginning we set γ to be α. We extend γ as needed during the simulation, as in the previous paragraph. As $3|n|^k < q$, we never run out of unused blocks. To construct β, first remove from γ every block that does not appear in oracle queries or replies in the computation of $C^\gamma(y)$. Then adjust h to skip over any holes, so that for each block B except for the bottom-most, $h(w) = w'$ where w is the minimum element of B and w' is the maximum element of the block below B. This cannot change the computation of $C(y)$, since if

$h(w)$ had been queried in that computation then the reply would have come from the block B' immediately below B in γ, and hence we would not have removed B'. Since the computation of $C^\gamma(y)$ makes at most $|n|^k$ queries, the resulting β belongs to A.

It remains to show that $C^\beta(y) < C^\alpha(x)$. It is enough to show $C^\gamma(y) < C^\gamma(x)$. Suppose not. Then under any standard oracle γ' which extends γ and is compatible with ρ we have $C^{\gamma'}(N^{\gamma'}(x)) \geq C^{\gamma'}(x)$, implying that x is a solution of our PLS problem in γ' and thus, by the properties of ρ, that $g(x)$ is a pair $\langle u, z\rangle$ where z is a witness to HOP. Considered as such a witness, z mentions at most three elements of $[n]$. We construct a particular such γ' from γ by first adding to the bottom of our ordering the blocks containing those of the three elements which are not yet in the domain of γ, and then all the remaining blocks in any order. Note that, as $3|n|^k + 3 < q$, we added at least one block below the three elements mentioned in z. Finally we let $h(x) = x$ for the minimal element x of the total ordering we have constructed and for every $x \notin [n]$. Thus γ' is a standard oracle, compatible with ρ, in which z does not witness HOP because $h(w)$ is the immediate predecessor of w for each element of $[n]$ mentioned in z. This completes the proof. □

3 Frames and frame decision trees

The computation of an oracle Turing machine can be modeled by a decision tree, in which each internal node is labeled with an oracle query and has children corresponding to the possible replies, and each leaf is labeled with an output value. In this section we define a particular kind of tree computing the function $f_n(u)$ from Lemma 2 under standard oracles. In Section 4 we will show that, for a random restriction ρ, the expected number of paths through the tree consistent with ρ, and hence the expected number of output values of $f_n(u)$ that can occur over all standard oracles consistent with ρ, is small (in fact it is less than 2, if n is large enough). Lemma 2 will follow easily.

The proof of this bound is complicated by the fact that in our random restrictions oracle replies are not independent of each other, which will require us to work with conditional probabilities. To consider, briefly, a simpler example, suppose that our machine only queried a unary oracle for a subset A of $[n]$. Fix a small probability p and let σ be the usual random restriction which independently sets each bit of A to 0 or 1, each with probability $(1-p)/2$, or leaves it unset with probability p. Let T be a tree modeling computations of the machine, where we may assume that no bit of A

is queried more than once along any path. Let v be any node in T. The expected number of replies consistent with σ to the query $x \in A$? labeling v is then $2p + (1-p) = 1+p$. It is straightforward to show (as in the first part of the proof of Lemma 3 below) that the expected number of paths through T consistent with σ is thus at most $(1+p)^d$, where d is the height of T.

Returning to our proof, let n, f_n and t be as in the proof of Theorem 1. We may assume without loss of generality, by adding dummy oracle queries as necessary, that the machine computing $f_n(u)$ on a standard oracle works in the following way. It maintains a set $S \subseteq [n]$ of points and some total ordering $\prec$ of S. Furthermore for some pairs $x, y \in S$ it knows that $h(x) = y$. It can ask two kinds of oracle query. The first is an *ordering query*, for $x \in [n] \setminus S$. This involves writing x? on the oracle tape, and getting as a reply the position of x in the oracle's ordering $\prec$ with respect to all elements of S. Assuming that $\prec$ is a linear ordering, there are at most $|S|+1$ possible replies. The second is a *predecessor query*, for $x \in S$ where $h(x)$ is not known. This involves writing $h(x)$? on the oracle tape, and getting as a reply some $y \in [n]$. There are at most n possible replies. Finally the computation outputs some $v \in [t^2]$.

In the rest of this section we give a formal definition of *frames* and *frame decision trees*, which we will use to model computations of $f_n(u)$ on standard oracles. A *frame* consists of:

1. a set $S \subseteq [n]$,
2. a linear ordering $\prec$ of S,
3. a partition $C_0, \dots, C_{r-1}$ of S into $\prec$-convex sets.

Each C_i is called an *h-chain.* If C is an h-chain, we write $\min(C)$ and $\max(C)$ for its minimum and maximum elements in the linear ordering $\prec$, respectively. We always assume that the partition of S into h-chains $C_0, \dots, C_{r-1}$ is given in the order induced by $\prec$ on their least elements, which by $\prec$-convexity means that

$$\min(C_{i-1}) \preceq \max(C_{i-1}) \prec \min(C_i) \preceq \max(C_i)$$

for every $i \in \{1, \dots, r-1\}$. For every x in S, we write C^x for the unique h-chain that contains x. The unique frame with empty S is called the *empty frame.* A total linear ordering $\prec'$ of $[n]$ is *compatible* with S if $\prec'$ extends $\prec$, and every h-chain in S remains $\prec'$-convex.

A *frame decision tree* (FDT) is a tree whose nodes are labeled by frames satisfying certain conditions, which we describe below. Each leaf node is assigned an *output*, which is an element of $[t^2]$. Each non-leaf node labeled by a frame $(S, \prec, C_0, \dots, C_{r-1})$ is assigned one of two types of *queries*: an ordering query x? for some x in $[n] \setminus S$, or a predecessor query $h(x)$? for some x in S with $x = \min(C^x)$.

A node labeled by an ordering query has $r+1$ children, one for each i in $\{0, \dots, r\}$. The child corresponding to $i \in \{0, \dots, r\}$ is an FDT whose root is labeled by a frame derived from the frame of its parent by extending S to $S \cup \{x\}$, extending the linear ordering $\prec$ to $y \prec x$ for every $y \in C_0 \cup \dots \cup C_{i-1}$ and $x \prec y$ for every $y \in C_i \cup \dots \cup C_{r-1}$, and with the following partition of $S \cup \{x\}$ into h-chains:

$$C_0, \dots, C_{i-1}, \{x\}, C_i, \dots, C_{r-1}.$$

For a node labeled by a predecessor query, there are two cases. The first case occurs if $C^x = C_0$, representing the situation in which x is the smallest element of S. In this case the tree has one child for each possible reply z in $[n] \setminus S$, and one extra child, representing the possibility that $h(x) = x$. The child corresponding to $z \in [n] \setminus S$ is an FDT whose root is labeled by the frame that has S extended to $S \cup \{z\}$, the linear ordering $\prec$ extended to $z \prec y$ for every y in S, and the following partition of $S \cup \{z\}$ into h-chains:

$$\{z\} \cup C_0, C_1, \dots, C_{r-1}.$$

The child corresponding to the reply $h(x) = x$ is an FDT whose root is labeled by the same frame as its parent.

The second case occurs if $C^x = C_i$ for some $i > 0$, representing the situation in which there are already elements of S smaller than x. In this case the tree has one child for each z in $[n] \setminus S$ and one child for $z = \max(C_{i-1})$. The child corresponding to $z \in [n] \setminus S$ is an FDT whose root is labeled by the frame that has S extended to $S \cup \{z\}$, the linear ordering $\prec$ extended to $y \prec z$ for every $y \in C_0 \cup \dots \cup C_{i-1}$ and $z \prec y$ for every $y \in C_i \cup \dots \cup C_{r-1}$, and the following partition of $S \cup \{z\}$ into h-chains:

$$C_0, \dots, C_{i-1}, \{z\} \cup C_i, C_{i+1}, \dots, C_{r-1}.$$

The child corresponding to $z = \max(C_{i-1})$ is an FDT whose root is labeled by the frame that has the same set S, the same linear ordering $\prec$, and the following partition of S into h-chains:

$$C_0, \dots, C_{i-2}, C_{i-1} \cup C_i, C_{i+1}, \dots, C_{r-1}.$$

If $\pi = (S, \prec, C_0, \dots, C_{r-1})$ is a frame and $\rho = (B_0, \dots, B_q, \prec_0, \dots, \prec_q)$ is a restriction from $\mathcal{R}(n, p, q)$, we say that ρ and π are *compatible*, denoted by $\rho \parallel \pi$, if there exists a total linear ordering $\prec'$ of $[n]$ which is compatible with both ρ and π. If T is an FDT and ρ is a restriction, the *restricted tree* $T|_\rho$ is defined by deleting each subtree whose root is labeled by a frame that is not compatible with ρ.

4 Decision trees under random restrictions

Let n_0, n, p, q, m, f_n and t be as in the proof of Theorem 1. In particular $p = n^{1/2}$, $q = n^{1/8}$ and $m = p/q = n^{3/8}$ are integers. We prove the following technical lemma:

Lemma 3. *If n_0 is large enough, T is an FDT modeling the computation of $f_n(u)$ for some $u \in [t]$, and ρ is a random restriction from $\mathcal{R}(n,p,q)$ with n, p and q as specified, then the expected number of leaves in $T|_\rho$ is at most $(1+n^{-1/10})^d$, where d is the height of T.*

Using this lemma we can now prove Lemma 2. Recall that our goal is to find a restriction $\rho \in \mathcal{R}(n,p,q)$ and a number $v \in [t^2]$ such that $f_n(u) \neq v$ for every $u \in [t]$ under every standard oracle compatible with ρ.

Proof of Lemma 2. For each $u \in [t]$, there is a frame decision tree T_u which computes the value of $f_n(u)$ under all standard oracles. The height d_u of T_u is bounded by a fixed polynomial function of $|n|$, say $|n|^k$. Applying Lemma 3, the expected number of leaves in $T_u|_\rho$ is at most $(1+n^{-1/10})^{d_u} \leq (1+n^{-1/10})^{|n|^k} < 2$, for large enough n_0.

Now let N_ρ be the sum, over all $u \in [t]$, of the number of leaves in $T_u|_\rho$. By linearity of expectation, the expected value of N_ρ over $\rho \in \mathcal{R}(n,p,q)$ is less than $2t$. Hence there is at least one ρ with $N_\rho < 2t$. Fix such a ρ, and observe that since $t > 2$ we have $N_\rho < t^2$ and hence there must exist some $v \in [t^2]$ which does not appear as the label of any leaf of any $T_u|_\rho$. Therefore, for every standard oracle compatible with ρ and every $u \in [t]$ we have $f_n(u) \neq v$, as required. □

Finally we prove Lemma 3.

Proof of Lemma 3. For every node v in T, let $N(v,\rho)$ be the number of leaves in the subtree of $T|_\rho$ rooted at v, if the subtree rooted at v survives in $T|_\rho$, and 0 otherwise. Let $\pi(v)$ be the frame that labels v in T. We prove that

$$\mathop{\mathbb{E}}_{\rho \parallel \pi(v)}[N(v,\rho)] \leq (1+n^{-1/10})^{\delta}, \tag{2}$$

where δ is the height of v in T and the expectation is over the probability space of random restrictions conditioned on the event that $\rho \parallel \pi(v)$. The lemma will follow since the frame at the root is the empty frame, which is compatible with every restriction.

We prove (2) by induction on the height δ of v. If v is a leaf, then $N(v, \rho) \leq 1$ with probability one and there is nothing to prove. If v is a non-leaf node and $v_0, \ldots, v_{\ell-1}$ are the children of v in T, then

$$\begin{aligned}
\mathop{\mathbb{E}}_{\rho \parallel \pi(v)}[N(v,\rho)] &= \sum_{i \in [\ell]} \mathop{\mathbb{E}}_{\rho \parallel \pi(v)}[N(v_i,\rho)] \\
&= \sum_{i \in [\ell]} \Pr_{\rho \parallel \pi(v)}[\rho \parallel \pi(v_i)] \cdot \mathop{\mathbb{E}}_{\rho \parallel \pi(v_i)}[N(v_i,\rho)] && (3) \\
&\leq (1 + n^{-1/10})^{\delta - 1} \cdot \sum_{i \in [\ell]} \Pr_{\rho \parallel \pi(v)}[\rho \parallel \pi(v_i)] && (4) \\
&= (1 + n^{-1/10})^{\delta - 1} \cdot \mathop{\mathbb{E}}_{\rho \parallel \pi(v)}[C(v,\rho)], && (5)
\end{aligned}$$

where $C(v, \rho)$ in the last equation is the random variable that counts the number of children v_i of v such that $\rho \parallel \pi(v_i)$. The identity in (3) follows from the fact that if ρ is not compatible with $\pi(v_i)$ then $N(v, \rho) = 0$. The inequality in (4) follows from the induction hypothesis, and the identity in (5) follows from the definition of $C(v, \rho)$. Thus, it suffices to show that

$$\mathop{\mathbb{E}}_{\rho \parallel \pi(v)}[C(v,\rho)] \leq 1 + n^{-1/10}. \qquad (6)$$

We proceed by cases on the type of query at v. For what follows, let $\pi(v) = (S, \prec, C_0, \ldots, C_{r-1})$ be the frame that labels v. Note that the structure of the frames that label the nodes of the tree guarantees that $r \leq |S| \leq d$. We make use of these inequalities below.

Ordering query. The query at v is of the type x? for $x \in [n] \setminus S$. In order to bound $C(v, \rho)$ for each fixed ρ we distinguish two cases: (a) $x \in B_0$ and (b) $x \notin B_0$. In case (a) we have $C(v, \rho) = 1$. This is because ρ gives a total ordering $\prec_0$ of B_0, and any ordering of $[n]$ compatible with ρ puts all elements of B_0 above all elements of $[n] \setminus B_0$. Hence exactly one child v_i of v will satisfy $\rho \parallel \pi(v_i)$, namely the child whose frame places x above all elements of $S \setminus B_0$ and in the ordering relation to the elements of $S \cap B_0$ that is given by $\prec_0$. In case (b) we have the bound $C(v, \rho) \leq r + 1 \leq d + 1$ inherited from the structure of T. To complete the argument it suffices to bound the probability that $x \notin B_0$. In order to do this, let A be the set of ρ

that are compatible with $\pi(v)$ and let B be the set of ρ that satisfy $x \notin B_0$. Then:

Claim 1.

$$\Pr_{\rho \| \pi(v)}[x \notin B_0] = \frac{|A \cap B|}{|A|} \leq \frac{(q+1) \cdot m}{n-p-d}.$$

Proof. We show this by constructing an injective map

$$F : (A \cap B) \times [n-p-d] \to A \times [q+1] \times [m]$$

as follows. We are given $\rho \in A \cap B$ and $j \in [n-p-d]$ and we want to produce $\rho' \in A$ and $(b,k) \in [q+1] \times [m]$.

Let B_i be the block B^x containing x. We know $i \neq 0$, so $|B_i| = m$. Enumerate B_i as $b_0, \dots, b_{m-1}$ using the ordering $\prec_i$ of ρ, and let $k \in [m]$ be the index of x in this enumeration. Enumerate the first $n-p-d$ elements of $B_0 \setminus S$ as $a_0, \dots, a_{n-p-d-1}$ using the ordering $\prec_0$ of ρ (recall that $|B_0| = n-p > m > d \geq |S|$). Let ρ' be ρ with x swapped with a_j. Order the blocks in ρ' by their least element in the standard ordering of $[n]$ and let $b \in [q+1]$ be the index, in this ordering of blocks, of the block in ρ to which x belonged before the swapping. Let F map (ρ, j) to (ρ', b, k).

Clearly ρ' belongs to A since from a total linear ordering $\prec'$ of $[n]$ that witnesses $\rho \parallel \pi(v)$ we can get another total linear ordering of $[n]$ that witnesses $\rho' \parallel \pi(v)$ by swapping the positions of x and a_j in $\prec'$. To show that the map is injective, suppose we are given (ρ', b, k) in its range. Then we can recover j by looking at the index of x in the block B_0 of ρ' under the ordering $\prec_0$ in ρ'. Now order the blocks of ρ' by their least element in the standard ordering of $[n]$ and let B_j be the block with index b in this ordering of blocks. Finally recover ρ by swapping x with the element with index k in B_j under the ordering $\prec_j$ of ρ'. □

Putting this together, in the case of an ordering query we get

$$\begin{aligned}\mathop{\mathbb{E}}_{\rho \| \pi(v)}[C(v,\rho)] &\leq 1 \cdot \Pr_{\rho \| \pi(v)}[x \in B_0] + (d+1) \cdot \Pr_{\rho \| \pi(v)}[x \notin B_0] \\ &\leq 1 + \frac{(d+1) \cdot (q+1) \cdot m}{n-p-d} \leq 1 + n^{-1/10},\end{aligned}$$

where the last inequality holds for large enough n_0 because d is bounded by a fixed polynomial function of $|n|$. This gives (6) as required.

Predecessor query. The query at v is of the type $h(x)$? for $x \in S$ with $x = \min(C^x)$. In order to bound $C(v,\rho)$ for each fixed ρ we again distinguish two cases: (a) $x \neq \min(B^x)$ and (b) $x = \min(B^x)$, where in both cases the minimum in B^x is taken with respect to the linear ordering on B^x defined in ρ. In case (a) we have $C(v, \rho) \leq 1$ as only one child v_i survives in the sense of satisfying $\rho \parallel \pi(v_i)$, namely the child that corresponds to the unique predecessor of x in B^x. In case (b) observe that in a standard oracle compatible with ρ, the predecessor of $x = \min(B^x)$ (if one exists) must be $\max(B_j)$ for some small block B_j distinct from B^x, where the maximum in B_j is taken with respect to $\prec_j$; there is also the possibility that $h(x) = x$. Hence there are at most $q+1$ possibilities and we have $C(v, \rho) \leq q+1$. To complete the argument it suffices to bound the probability that $x = \min(B^x)$. In order to do this, let A be the set of ρ that are compatible with $\pi(v)$ and let C be the set of ρ that have $x = \min(B^x)$. Then:

Claim 2.

$$\Pr_{\rho \parallel \pi(v)}[x = \min(B^x)] = \frac{|A \cap C|}{|A|} \leq \frac{q+1}{m-d}.$$

Proof. We show this by constructing an injective map

$$G : (A \cap C) \times [m-d] \to A \times [q+1]$$

as follows. We are given $\rho \in A \cap C$ and $j \in [m-d]$ and we want to produce $\rho' \in A$ and $b \in [q+1]$.

Order the small blocks of ρ by their least element in the standard ordering of $[n]$ and let B_i be the first small block, in this ordering of blocks, which contains no element of S. Since $|S| \leq d < q$ such a block exists. Enumerate B_i as $b_0, \dots, b_{m-1}$ using the ordering $\prec_i$ of ρ. Let ρ' be ρ with the elements of the h-chain C^x swapped with $b_j, \dots, b_{j+|C^x|-1}$. Since $j < m-d$ and $|C^x| \leq |S| \leq d$, this is well defined. Order the blocks in ρ' by their least element in the standard ordering of $[n]$ and let $b \in [q+1]$ be the index, in this ordering of blocks, of the block to which x belonged before the swapping. Let G map (ρ, j) to (ρ', b).

Using the facts that $x = \min(B^x) = \min(C^x)$ and $B_i \cap S = \emptyset$, we argue that ρ' is compatible with $\pi(v)$ as follows. From a total linear ordering $\prec'$ of $[n]$ that witnesses $\rho \parallel \pi(v)$ we construct a total linear ordering of $[n]$ that witnesses $\rho' \parallel \pi(v)$ in two stages. We first swap the positions of C^x and $b_j, \dots, b_{j+|C^x|-1}$ in $\prec'$. This gives us compatibility with ρ'. We then remove the block B_i from the place it occupies in $\prec'$ and re-insert it just before the block B^x. This preserves compatibility with ρ' and gives us compatibility with $\pi(v)$, since it puts the elements of C^x back into their correct position in the ordering with respect to the other elements of S. Thus $\rho' \in A$.

To show that the map is injective, suppose we are given (ρ', b) in its range. Order the blocks of ρ' by their least element in the standard ordering of $[n]$. Let R_1 be the block in position b in this ordering of blocks and let R_2 be the block containing C^x. Then we can recover j by looking at the place where C^x appears in R_2, and we can recover ρ by swapping C^x from R_2 with the first $|C^x|$ elements of R_1. □

Putting this together, in the case of a predecessor query we get

$$\begin{aligned}\mathop{\mathbb{E}}_{\rho \| \pi(v)}[C(v,\rho)] &\le 1 \cdot \mathop{\Pr}_{\rho\|\pi(v)}[x \neq \min(B^x)] + (q+1) \cdot \mathop{\Pr}_{\rho\|\pi(v)}[x = \min(B^x)] \\ &\le 1 + \frac{(q+1)^2}{m-d} \le 1 + n^{-1/10},\end{aligned}$$

where the last inequality holds for large enough n_0 because d is bounded by a fixed polynomial function of $|n|$. This gives (6) as required. This completes the proof of Lemma 3. □

5 A propositional version of our result

Let $\overline{\mathrm{HOP}}_n$ be the negation of HOP on the interval $[n]$, written as a propositional formula with propositional variables for the ordering $\prec$ and for the bits of the predecessor function h. Let f be any polynomial time function with oracles $\prec$ and h and let $t = t(n)$ be any quasipolynomial term. Let $A_{n,v}$ be a propositional formula asserting that there is no $u \in [t]$ such that $f_n(u) = v$. We may take both $\overline{\mathrm{HOP}}_n$ and $A_{n,v}$ to be *narrow CNFs*, that is, conjunctions of quasipolynomially many clauses, with each clause of polylogarithmic width in n.

By a standard translation of first-order into propositional proofs (via treelike Res(log) [11, 12], or see [2] for a translation via PLS problems), if it were true that $\mathrm{T}^1_2(\alpha) + \mathrm{sWPHP}(\mathrm{PV}(\alpha)) \vdash \mathrm{HOP}$ then it would follow that

(i) $\overline{\mathrm{HOP}}_n \wedge A_{n,v}$ has polylogarithmic width resolution refutations, for all $v \in [t^2]$.

Observe that for each assignment to the propositional variables, the formula $A_{n,v}$ is true for almost all $v \in [t^2]$. Hence, given any probability distribution $\mathcal{R}$ of assignments, by an averaging argument there is some $v \in [t^2]$ for which $A_{n,v}$ is true with high probability. Hence in particular from (i) it would follow that

(ii) for every distribution $\mathcal{R}$, there is a narrow CNF A, which is true with high probability, such that $\overline{\mathrm{HOP}}_n \wedge A$ has polylogarithmic width resolution refutations.

This is called a *random narrow resolution refutation* of $\overline{\mathrm{HOP}}_n$ over $\mathcal{R}$ in [2]. The proof of our main result, with minor modifications, also shows that (i) is false. We are not able to say anything about (ii).

References

[1] S. Buss. Bounded arithmetic. Bibliopolis, 1986.

[2] S. Buss, L. Kołodziejczyk, and N. Thapen, *Fragments of approximate counting.* Manuscript, available at `www.math.cas.cz/~thapen/`, 2012.

[3] S. Buss and J. Krajíček, *An application of Boolean complexity to separation problems in bounded arithmetic.* Proceedings of the London Mathematical Society, 69:1-21, 1994.

[4] M. Chiari and J. Krajíček, *Witnessing functions in bounded arithmetic and search problems.* Journal of Symbolic Logic, 63(3):1095-1115, 1998.

[5] J. Hanika. Search problems and bounded arithmetic. Doctoral thesis, Charles University, available at `eccc.hpi-web.de/static/books/theses/`, 2004.

[6] R. Impagliazzo and J. Krajíček, *A note on conservativity relations among bounded arithmetic theories.* Mathematical Logic Quarterly, 48(3):375-377, 2002.

[7] E. Jeřábek, *The strength of sharply bounded induction.* Mathematical Logic Quarterly, 52(6):613-624, 2006.

[8] E. Jeřábek, *On independence of variants of the weak pigeonhole principle.* Journal of Logic and Computation, 17(3):587-604, 2007.

[9] E. Jeřábek, *Approximate counting by hashing in bounded arithmetic.* Journal of Symbolic Logic, 74(3):829-860, 2009.

[10] J. Krajíček. Bounded Arithmetic, Propositional Logic and Computational Complexity. Cambridge University Press, 1995.

[11] J. Krajíček, *On the weak pigeonhole principle.* Fundamenta Mathematicae, 170:123-140, 2001.

[12] M. Lauria, *Short Res*(polylog) Refutations if and only if Narrow Res Refutations*. Manuscript, available at `arXiv:1310.5714`, 2011.

[13] N. Thapen, *A model-theoretic characterization of the weak pigeonhole principle*. Annals of Pure and Applied Logic, 118:175-195, 2002.