Dual weak pigeonhole principle, Boolean complexity, and derandomization

Emil Jeřábek
Mathematical Institute, AS CR, Prague

November 25, 2003

Abstract

We study the extension (introduced as BT in [5]) of the theory S^1_2 by instances of the dual (onto) weak pigeonhole principle for p-time functions, $dWPHP(PV)^x_{x^2}$. We propose a natural framework for formalization of randomized algorithms in bounded arithmetic, and use it to provide a strengthening of Wilkie's witnessing theorem for $S^1_2+dWPHP(PV)$. We construct a propositional proof system WF (based on a reformulation of Extended Frege in terms of Boolean circuits), which captures the $\forall\Pi^b_1$-consequences of $S^1_2+dWPHP(PV)$. We also show that WF p-simulates the Unstructured Extended Nullstellensatz proof system of [2].

We prove that $dWPHP(PV)$ is (over S^1_2) equivalent to a statement asserting the existence of a family of Boolean functions with exponential circuit complexity. Building on this result, we formalize the Nisan-Wigderson construction (derandomization of probabilistic p-time algorithms) in a conservative extension of $S^1_2+dWPHP(PV)$.

Preliminaries

We assume the reader is familiar with basic concepts of bounded arithmetic, such as the hierarchy of Σ^b_i and Π^b_i formulas, and theories S^i_2. (Section 5.2 of [4] is a good introduction to these topics.)

The theory PV has function symbols for all polynomial-time algorithms (introduced inductively via Cobham's limited recursion on notation), its axiom set consists of defining equations for all these functions, and an open axiom schema equivalent to open induction. (This theory is called PV_1 in [4]. Our usage of the symbol PV is nonstandard, it usually denotes an equational theory.) The hierarchy of $\Sigma^b_i(PV)$-formulas in the language of PV is defined as usual, and we define $S^i_2(PV)$ as the extension of PV by the $\Sigma^b_i(PV)$-*PIND* schema.

PV-functions have well-behaved Δ^b_1-definitions in S^1_2. Under this interpretation, every $\Sigma^b_i(PV)$-formula is equivalent to a Σ^b_i-formula, in particular $S^1_2(PV)$ is a definable (hence conservative) extension of S^1_2. We will thus ignore the distinction between S^1_2 and $S^1_2(PV)$, and use PV-functions freely to simplify the presentation. If the reader is unfamiliar with PV, she may simply identify PV-functions with functions Δ^b_1-definable in S^1_2.

If M is a model of PV or S^1_2, $Log(M)$ denotes the cut $\{|a|^M;\, a \in M\}$. We will often use this notation outside the model-theoretical context, in which case $x \in Log$ is a shortcut for $\exists y\, x = |y|$. Similarly, $x \in LogLog$ means $\exists y\, x = ||y||$.

Let f be a function. *Dual weak pigeonhole principle* for f is the formula

$$\forall a > 1\; dPHP(f)^a_{a^2},$$

where $dPHP(f)^a_b$ stands for

$$\exists v < b\, \forall u < a\; f(u) \neq v.$$

The schema $dWPHP(PV)$ is the dual weak pigeonhole principle for all PV-functions f (with parameters). This schema is finitely axiomatizable: it is equivalent to $dWPHP(eval)$, where $eval(C,u)$ is the PV-function which evaluates a circuit C on input u. The exact bound $b = a^2$ in the definition of $dWPHP(PV)$ is inessential, since the following are equivalent over S^1_2 (this is essentially due to [11]):

(*i*) $\forall a\, \exists b\; dPHP(PV)^a_b$,

(*ii*) $dWPHP(PV)$,

(*iii*) $\forall a > 0\, \forall n \in Log\; dPHP(PV)^{an}_{a(n+1)}$.

In particular, we will often use the principle with $b = 2a$.

Sharply bounded collection is the schema

$$\forall i \leq |x|\, \exists v \leq y\, \varphi(i,v) \to \exists w\, \forall i \leq |x|\, \varphi(i,(w)_i).$$

The symbol $BB\Sigma^b_i$ denotes sharply bounded collection for all Σ^b_i-formulas φ. $BB\Sigma^b_i$ is provable in S^i_2 (see [4]), and $S^i_2 + BB\Sigma^b_{i+1}$ is $\forall\Sigma^b_{i+1}$-conservative over S^i_2, by [12].

We will occasionally need another schema, the *length-minimization principle*:

$$\varphi(x) \to \exists u \leq x\, (\varphi(u)\ \&\ \forall v \leq x\, (|v| < |u| \to \neg\varphi(v))).$$

Length minimization for Σ^b_i-formulas, Σ^b_i-*LENGTH-MIN*, is provable in S^i_2 (in fact, it is equivalent to S^i_2 over a weak base theory). Similarly, there is a *length-maximization principle*, Σ^b_i-*LENGTH-MAX*.

We recall two indispensable tools in bounded arithmetic: *Parikh's theorem*, and *Buss's witnessing theorem*.

0.1 Proposition (Parikh [10]) *Let T be a $\forall\Sigma^b_\infty$-axiomatizable extension of S^1_2. If T proves $\forall x\, \exists y\, \varphi(x,y)$, where φ is bounded, then there exists a term t such that*

$$T \vdash \forall x\, \exists y \leq t(x)\, \varphi(x,y). \qquad \square$$

0.2 Proposition (Buss [1]) *Assume that $S^1_2 \vdash \forall x\, \exists y\, \varphi(x,y)$, with $\varphi \in \Sigma^b_1$. Then there exists a PV-function f such that $PV \vdash \forall x\, \varphi(x, f(x))$.* $\square$

We will also need some notions from proof complexity. Recall that a *propositional proof system* is a p-time computable function P, whose range is the set $TAUT$ of all classical propositional tautologies in De Morgan language. A proof system P *polynomially simulates* a proof system Q, in symbols $Q \leq_p P$, iff there is a p-time function f such that $Q = P \circ f$.

The set of propositional tautologies is definable by the Π_1^b-formula

$$Taut(\varphi) \equiv \forall x < 2^{|\varphi|}\, eval(\varphi, x) = 1.$$

If P is PV-function which defines a propositional proof system, the *consistency* and *reflection* principles for P are the $\forall\Pi_1^b$-sentences

$$Con(P) \equiv \forall \pi\, P(\pi) \neq \bot,$$
$$0\text{-}RFN(P) \equiv \forall \pi\, Taut(P(\pi)).$$

An important link between bounded arithmetic and propositional proof complexity is given by translation of bounded formulas into propositional logic. For any Π_1^b-formula φ, there is a (canonically constructed) sequence of propositional formulas $\{\|\varphi\|^n;\, n \in \omega\}$, such that $\forall x\, \varphi(x)$ is true in the standard model iff all $\|\varphi\|^n$ are tautologies. The following theorem is a prominent example of a connection between an arithmetical theory, and a propositional proof system:

0.3 Proposition (Cook [3])

(*i*) *If $PV \vdash \varphi(x)$, $\varphi \in \Pi_1^b$, then tautologies $\|\varphi\|^n$ have polynomial-time constructible proofs in the Extended Frege (EF) proof system.*

(*ii*) $PV \vdash 0\text{-}RFN(EF)$ □

Other notation: we denote the set of natural numbers by ω. We also borrow from set theory the convention $n = \{0, 1, \dots, n-1\}$, in particular a "function $f\colon a \to b$" is actually $f\colon [0, a-1] \to [0, b-1]$.

Many of our results are formalizations of known statements in fragments of bounded arithmetic, like PV or $S_2^1 + dWPHP(PV)$. To make the notation more compact, we indicate the theory by the symbol "$(T \vdash:)$", which can be read as "theory T proves:".

1 Randomized computation in bounded arithmetic

The main purpose of the present section is to develop a framework for expressing (defining) a certain kind of probabilistic algorithms in bounded arithmetic. As described in the next definition, we deal with a slightly nonstandard class of randomized algorithms; this choice was motivated by two demands: (i) we want ZPP, RP, and $coRP$ languages to fit in, (ii) we want to consider functions as well as predicates. Moreover, it is not natural for randomized algorithms to compute *univalued* functions, hence we allow also *multifunctions*. Formally, an n-ary partial multifunction is just an $(n+1)$-ary relation; by an abuse of language, we write $F(\vec{x}) = y$ as a shorthand for "y is one of the possible values of $F(\vec{x})$". Also notice that we left out *BPP* algorithms; they would require a different treatment, which does not blend smoothly with our approach to *RP*-like algorithms (in particular, the concept of *BPP*-like multifunctions does not seem to make much sense).

1.1 Definition Let F be a partial multifunction, $\alpha\colon \omega \to [0,1]$, and M a randomized Turing machine. We say that *M is an α-PPTM* (*probabilistic polynomial-time Turing machine*) *for F* iff the following conditions are satisfied:

- (i) The time of any computation of M is polynomial in the length of its input.
- (ii) On any input x, either M computes a number y such that $F(x) = y$, or it halts in a special state labeled "Sorry, try again" (S.T.A.)
- (iii) If $x \in \mathrm{dom}(F)$, the probability that the computation of M on input x stops in the S.T.A. state is bounded from above by $\alpha(|x|)$.

Let *MFRP* be the class of all partial multifunctions (pmf) computable by a 1/2-*PPTM*.

1.2 Remarks

- Trivial amplification shows that the definition of *MFRP* does not change, if we replace the constant 1/2 by any function $\alpha(n)$ such that $1 - n^{-c} \geq \alpha(n) \geq 2^{-n^c}$ for some $c > 0$.
- $L \in \mathit{ZPP}$ iff the characteristic function of L is in *MFRP*.
- $L \in \mathit{RP}$ iff $L = \mathrm{dom}(F)$ for some $F \in \mathit{MFRP}$ iff the function which is constantly 0 on L and undefined on its complement is in *MFRP*.
- An α-*PPTM* for F is also an α-*PPTM* for any pmf G such that $\mathrm{dom}(F) = \mathrm{dom}(G)$ and $F \subseteq G$.

Our next step is to formalize this definition in bounded arithmetic. First we give an informal description. We take a PV-function $f(\vec{x}, w)$, which simulates the computation of M on input $\vec{x}$ and a string of random bits w. The machine may touch only a polynomial number of these random bits, we thus fix an explicit bound $w < r(\vec{x})$. The output of $f(\vec{x}, w)$ is either a number, or a special symbol "$*$", which corresponds to halting in the S.T.A. state (we may encode it as a number by putting "$*$" $= 0$, "n" $= n + 1$). Now we need to express the condition (iii). Assume that $F(\vec{x})$ is defined, and let us say that a random string w is *good*, if $f(\vec{x}, w) \neq *$, otherwise it is *bad*. We will consider an *onto* mapping $m\colon t \times r(\vec{x}) \twoheadrightarrow s \times \mathit{Bad}$, where *Bad* is the set of all bad random strings; such a mapping explicitly witnesses that the ratio of bad strings is at most t/s, hence (iii) holds with $\alpha = t/s$. A formal definition follows:

1.3 Definition Let T be a theory containing PV, and $t(\vec{x})$ and $s(\vec{x})$ any PV-functions. A *definable t/s-PPTM* consists of PV functions f and r such that T proves

$$(\bigstar)\quad \exists w < r(\vec{x})\ \ f(\vec{x}, w) \neq * \to$$
$$\to \exists\, \text{circuit } C\, \forall w < r(\vec{x})\, (f(\vec{x}, w) = * \to \forall i < s(\vec{x})\, \exists v < r(\vec{x})\, \exists j < t(\vec{x})\ \ C(v, j) = \langle w, i \rangle),$$

where the size of C is tacitly bounded by a polynomial in the length of $\vec{x}$. A *t/s-PPTM* is *uniformly witnessed* if the formula above holds with $C(v,j)$ replaced by $m(\vec{x},v,j)$, where m is a PV-function symbol.

A definable *t/s-PPTM* computes a pmf $F(\vec{x})$, defined by

$$F(\vec{x}) = y \quad \text{iff} \quad \exists w < r(\vec{x}) \;\; f(\vec{x},w) = y \neq *.$$

(Notice that this is Σ^b_1. Condition ($\bigstar$) itself is $\forall\Sigma^b_3$ for general *PPTM*'s, and $\forall\Sigma^b_1$ for uniformly witnessed *PPTM*'s.) We will call such a function *definable t/s-MFRP*, or shortly *t/s-definable*. A definable *MFRP* is *weakly total* iff ($\bigstar$) holds with the condition "$\exists w < r(\vec{x}) \;\; f(\vec{x},w) \neq * \rightarrow$" *dropped.*

1.4 Observation Assuming $dWPHP(PV)$, a weakly total definable *t/s-MFRP* is total, provided $2t \leq s$.

Proof: If $F(\vec{x})$ were undefined, the circuit C from 1.3 would represent a surjective mapping of $t(\vec{x})r(\vec{x})$ onto $s(\vec{x})r(\vec{x})$, contradicting $dWPHP(PV)$. □

1.5 Lemma ($PV \vdash$:) *Let t, s and p be PV-functions such that $p(x) \geq 1$. Any tp/sp-definable MFRP F has a t/s-definition, which is uniformly witnessed and/or weakly total, whenever F is. (Hence the symbol t/s may be interpreted as a quotient.)*

Proof: Let $f(x,w)$ and $r(x)$ be as in Definition 1.3. Put

$$r'(x) := r(x) \cdot p(x),$$
$$f'(x,w') := f(x,w'_1),$$

where we consider w' as a pair $[w'_0, w'_1]$, $w'_0 < p(x)$, $w' = w'_1 \cdot p(x) + w'_0$. Let $f'(x,w') \neq *$ for some $w' < r'(x)$. This means that $f(x,w'_1) \neq *$, hence there is a circuit C such that $C(v,j) = \langle w,i \rangle$ for some $j < t(x) \cdot p(x)$ and $v < r(x)$, whenever $i < s(x) \cdot p(x)$, $w < r(x)$ and $f(x,w) = *$. Define a new circuit C' by

$$C'(v',j') := \langle [i_0,w], i_1 \rangle, \qquad \text{where } C(v'_1, [v'_0, j']) = \langle w,i \rangle.$$

(As above, we decompose $i = [i_0,i_1]$, $v' = [v'_0, v'_1]$, etc.) Given $i' < s(x)$ and $w' < r(x) \cdot p(x)$ such that $f'(x,w') = *$, we have $[w'_0, i'] < s(x) \cdot p(x)$ and $f(x,w'_1) = *$, hence there is $j < t(x) \cdot p(x)$ and $v < r(x)$ such that $C(v,j) = \langle w'_1, [w'_0, i'] \rangle$. Therefore $C'([j_0,v], j_1) = \langle [w'_0, w'_1], i' \rangle = \langle w', i' \rangle$, $j_1 < t(x)$, and $[j_0, v] < r'(x)$ as required. □

1.6 Lemma ($PV + BB\Sigma^b_1 \vdash$:) *Let t, s and p be PV-functions such that $p(x) \geq 1$. Then any t/s-definable MFRP F has a $t^{|p|}/s^{|p|}$-definition, which is uniform and/or weakly total, if the original one was. (This lemma also holds in plain PV, if p is constant.)*

Proof: For any fixed numbers a and b, we may identify $w < a^{|b|}$ with a sequence $\langle w_k \rangle_{k<|b|}$ of numbers less than a, namely $w_k = \lfloor \frac{w}{a^k} \rfloor \bmod a$. Given f and r defining F, we put

$$r'(x) := r(x)^{|p(x)|},$$

$$f'(x,w) := \begin{cases} *, & \text{if } \forall k < |p(x)| \;\; f(x,w_k) = *, \\ f(x,w_k), & \text{if } k < |p(x)| \text{ is minimal such that } f(x,w_k) \neq *. \end{cases}$$

Clearly, $\exists w < r(x)\ f(x,w) = y$ iff $\exists w < r'(x)\ f'(x,w) = y$. Assume that C is a circuit satisfying ($\bigstar$). Define

$$C'(v,j) = \langle w,i\rangle, \qquad \text{where } C(v_k,j_k) = \langle w_k,i_k\rangle \text{ for each } k < |p(x)|.$$

Let $i < s(x)^{|p(x)|}$ and $f'(x,w) = *$, $w < r'(x)$. This means that for any k, $f(x,w_k) = *$, hence there are $v' < r(x)$ and $j' < t(x)$ such that $C(v',j') = \langle w_k,i_k\rangle$. By $BB\Sigma^b_1$ there are sequences v and j such that $C(v_k,j_k) = \langle w_k,i_k\rangle$ for any $k < |p(x)|$. Then $C'(v,j) = \langle w,i\rangle$. □

1.7 Corollary ($PV + BB\Sigma^b_1 \vdash$:) *Assuming $s(x) \geq 1$ and $t(x)(|p(x)|+1) \leq s(x)|p(x)|$ for some p, any t/s-definable MFRP has a $1/q$-definition for any $q(x)$. (I.e., as in the real world, we can boost the probability of error from $1 - 1/poly(n)$ to $1/2^{poly(n)}$.)*

Proof: Straightforward induction shows that $(a+1)^b \geq a^b + ba^{b-1}$ for any $b \geq 1$, $b \in Log$, in particular $(|p(x)|+1)^{|p(x)|} \geq 2|p(x)|^{|p(x)|}$. This implies $s^{|p|}|p|^{|p|} \geq t^{|p|}(|p|+1)^{|p|} \geq 2t^{|p|}|p|^{|p|}$, hence $s^{|p|} \geq 2t^{|p|}$. Thus using Lemmas 1.6 and 1.5, any t/s-definable *MFRP* has a $1/2$-definition, and also a $1/q$-definition by Lemma 1.6 again, as $2^{|q|} > q$. □

1.8 Lemma ($PV + BB\Sigma^b_1 \vdash$:) *Any 1/2-definable MFRP has a uniformly witnessed 1/2-definition.*

Proof: Let f and r be the 1/2-definition of F, and let $C \leq c(x)$ be the circuit size bound implicit in ($\bigstar$). Put $p(x) = 2^{|c(x)|}$ and define f' and r' as in the proof of Lemma 1.6. Finally, define

$$m(x,v,j) := \langle w,i\rangle, \qquad \text{where } eval(j,\langle v_k,0\rangle) = \langle w_k,i_k\rangle \text{ for each } k < |p(x)|.$$

(Here $eval(C,x)$ is the value computed by a Boolean circuit C on input x.) Assuming $C \leq c(x)$ is a circuit satisfying ($\bigstar$), the proof of Lemma 1.6 shows that m witnesses that f' and r' form a $2^{|c|}/2^{|p|}$-definition of F (the third argument of m will be C for all w and i). Lemma 1.5 implies that F has a uniform 1/2-definition, because $2^{|p|} = 2 \cdot 2^{|c|}$. □

1.9 Definition Let $F(\vec{x})$ and $G(y)$ be partial multifunctions. We say that G is *composable* with F, if for all $\vec{x}$, y and y' such that $F(\vec{x}) = y$ and $F(\vec{x}) = y'$, $y \in \mathrm{dom}(G)$ iff $y' \in \mathrm{dom}(G)$. Similarly for $G(y_1,\dots,y_n)$ and $F_1(\vec{x}),\dots,F_n(\vec{x})$.

1.10 Remark There is a total multifunction F and a partial function G, both in *MFRP* (using no randomness at all, in fact), such that their composition $G \circ F$ is a constant partial function with an *NP*-complete domain (hence $G \circ F \notin MFRP$, unless $NP = RP$). Indeed, choose an *NP*-complete predicate $Q(x) \leftrightarrow \exists y\,(|y| \leq |x|^n \,\&\, R(x,y))$ with $R \in P$, and put

$$\begin{aligned}
&F(x) = y \quad \text{iff} \quad y = 0 \text{ or } R(x,y-1),\ |y-1| \leq |x|^n,\\
&G(0) \text{ is undefined},\\
&G(x+1) = 0.
\end{aligned}$$

Clearly, G is a partial p-time function. Also $F \in MFRP$, because F contains the constant 0 function. However,

$$(G \circ F)(x) = \begin{cases} 0, & \text{if } Q(x), \\ \text{undefined} & \text{otherwise.} \end{cases}$$

This shows that dealing with a condition like 1.9 is unavoidable, if we want *MFRP* to be closed under composition (or even to formalize this in bounded arithmetic).

1.11 Lemma ($PV + BB\Sigma^b_1 \vdash$:) *Let $F_1(\vec{x}), \dots, F_n(\vec{x})$, $G(y_1, \dots, y_n)$ be 1/2-definable p.m.f., such that G is composable with $F_1, \dots, F_n$. Then their composition $G(F_1(\vec{x}), \dots, F_n(\vec{x}))$ is also 1/2-definable.* (PV *suffices, if G and F_i's are uniformly witnessed.*)

Proof: For simplicity we will assume $n = 1$. By 1.7 and 1.8 there is a 1/3-definition of F given by functions $f(x, w)$ and $r(x)$, uniformly witnessed by $m(x, v)$. Similarly let $f'(y, w)$ and $r'(y)$ be a 1/3-definition of G, uniformly witnessed by $m'(y, v)$. Using the idea of the proof of Lemma 1.5, we may assume that $r'(y) \mid r'(z)$ whenever $y \leq z$. Let $b(x)$ be a PV-function such that $f(x', w) \leq b(x)$ for all $x' \leq x$ and $w < r(x')$. Define

$$r''(x) := r'(b(x)) \cdot r(x),$$

$$f''(x, w) := \begin{cases} *, & \text{if } f(x, w_0) = *, \\ f'(f(x, w_0), w_1 \bmod r'(f(x, w_0))) & \text{otherwise,} \end{cases}$$

$$m''(x, v, 0) := \langle [w, v_1], i \rangle, \qquad \text{where } m(x, v_0) = \langle w, i \rangle,$$

$$m''(x, v, 1) := \begin{cases} \langle 0, 0 \rangle, & \text{if } f(x, v_0) = *, \\ \langle [v_0, w + \lfloor \frac{v_1}{q} \rfloor \cdot q], i \rangle, & \text{if } q = r'(f(x, v_0)),\ m'(f(x, v_0), v_1 \bmod q) = \langle w, i \rangle. \end{cases}$$

We claim that f'' and r'' is a 2/3-definition of $G \circ F$, witnessed by m''. Clearly, non-$*$ values of $f''(x, w)$ are just the values of $G \circ F$. Assume that $f''(x, u) \neq *$ for some $u < r''(x)$, and let $w < r''(x)$, $i < 3$, and $f''(x, w) = *$. This means that either $f(x, w_0) = *$, or $f'(y, w_1 \bmod r'(y)) = *$, where $y = f(x, w_0)$. In the former case, we put $j = 0$, and we find $v < r(x)$ such that $m(x, v) = \langle w_0, i \rangle$, then we have $m''(x, [v, w_1], j) = \langle w, i \rangle$. In the latter case, put $q = r'(y)$ and $j = 1$. Since $f(x, w_0) \in \mathrm{dom}(G)$ and G is composable with F, we have $y \in \mathrm{dom}(G)$, hence there is $v < q$ such that $m'(y, v) = \langle w_1 \bmod q, i \rangle$. Then $m''(x, [w_0, v + q\lfloor \frac{w_1}{q} \rfloor], j) = \langle [w_0, (w_1 \bmod q) + q\lfloor \frac{w_1}{q} \rfloor], i \rangle = \langle w, i \rangle$.

By 1.7, we may turn a 2/3-definition of $G \circ F$ into a 1/2-definition. □

1.12 Lemma ($PV + BB\Sigma^b_1 + dWPHP(PV) \vdash$:) *Let $F(x)$ be a 1/2-definable pmf. For every $n \in Log$ there exists a (polynomial size) circuit $C \colon 2^n \to 2^m \cup \{*\}$ such that*

$$F(x) \text{ is defined } \leftrightarrow C(x) \neq *,$$
$$y = C(x) \neq * \rightarrow F(x) = y,$$

for any x of length n.

Proof: Fix n, and a uniformly witnessed $1/2^{n+1}$-definition of F. We may assume that $r(x) = r$ is independent on x (for x of length n). The witnessing function

$$m(x,v)\colon 2^n \times r \to 2^{n+1} \times r$$

cannot be onto (by $dWPHP$), we may thus fix $\langle i,w\rangle \in (2^{n+1} \times r) \smallsetminus \mathrm{rng}(m)$, and define $C(x) = f(x,w)$. Clearly $F(x) = C(x)$ if $C(x) \neq *$. Moreover, if $C(x) = *$ then $F(x)$ is undefined, because otherwise there would be a $v < r$ such that $m(x,v) = \langle i,w\rangle$, a contradiction. $\square$

1.13 Example (Rabin's algorithm.) There is a *coRP*-predicate $P(x)$, 1/2-definable in S^1_2, such that S^1_2 proves

$$P(x) \quad \text{iff} \quad x > 1 \,\&\, \forall y < x \ \ (y \neq 0 \to y^{x-1} \equiv 1 \pmod{x}).$$

Any number satisfying this condition is provably prime, but the converse is equivalent to the Little Fermat's Theorem (hence unlikely to be provable in S^1_2, by [7]).

Proof: Define

$$r(x) := \begin{cases} 1, & \text{if } x \leq 1 \vee 2 \mid x, \\ x-2 & \text{otherwise,} \end{cases}$$

$$f(x,w) := \begin{cases} *, & \text{if } x = 2, \\ 0, & \text{if } x \leq 1 \vee (2 \mid x \,\&\, x > 2), \\ *, & \text{if } x > 2 \text{ odd, and } \forall k\, (2^k \mid x-1 \to (w+1)^{(x-1)/2^k} \equiv 1 \pmod{x}), \\ *, & \text{if } x > 2 \text{ odd, and } j \neq 0 \,\&\, (w+1)^{(x-1)/2^j} \equiv -1 \pmod{x}, \\ 0 & \text{otherwise,} \end{cases}$$

where $j < |x|$ is the least number such that $2^j \mid x-1$ and

$$(w+1)^{(x-1)/2^j} \not\equiv 1 \pmod{x},$$

$$P(x) :\leftrightarrow \forall w < r(x) \ \ f(x,w) = *,$$
$$Q(x) :\leftrightarrow x > 1 \,\&\, \forall y < x \ \ (y \neq 0 \to y^{x-1} \equiv 1 \pmod{x}).$$

(It would be more natural to consider random choices $w \in [1,x)$.) From now on, we will assume that $x > 2$ and x is odd, other cases are trivial.

Let $x - 1 = y2^k$, where y is odd and $k > 0$. Clearly $(-1)^y \equiv -1 \not\equiv 1 \pmod{x}$, let $i \leq k$ be the least number such that $\exists a \in \mathbb{Z}^*_x \ a^{(x-1)/2^i} \not\equiv 1 \pmod{x}$, which exists by the Σ^b_1-*LENGTH-MIN* principle. (Here $a \in \mathbb{Z}^*_x$ means $a < x \,\&\, (a,x) = 1$, which is in PV equivalent to $a < x \,\&\, \exists b < x \ \ ab \equiv 1 \pmod{x}$.)

Case 1: $i = 0$. Clearly, neither $P(x)$ nor $Q(x)$ holds. Let $b \in \mathbb{Z}^*_x$ be such that $b^{x-1} \not\equiv 1 \pmod{x}$. (Forgetting about S^1_2 for a moment, $\{w;\ f(x,w+1) = *\}$ is contained in a proper subgroup $\{w;\ w^{x-1} \equiv 1\} \lneq \mathbb{Z}^*_x$, thus there are at most $|\mathbb{Z}^*_x|/2$ of them, and we may witness this using multiplication by a fixed element b of a nontrivial coset of this subgroup.) Define

$$C(x,v) := \begin{cases} \langle (b(v+1) \bmod x) - 1, 0\rangle, & \text{if } (b(v+1))^{x-1} \equiv 1 \pmod{x}, \\ \langle v, 1\rangle & \text{otherwise.} \end{cases}$$

Assume $w < r(x)$ is such that $f(x,w) = *$. Then we have $(w+1)^{x-1} \equiv 1 \pmod{x}$ and $(b(w+1))^{x-1} \not\equiv 1 \pmod{x}$, hence $C(x,w) = \langle w, 1\rangle$. Since $b, w+1 \in \mathbb{Z}_x^*$, there is $v < r(x)$ such that $b(v+1) \equiv w+1 \pmod{x}$. We have $(b(v+1))^{x-1} \equiv 1 \pmod{x}$, thus $C(x,v) = \langle w, 0\rangle$.

Case 2: $i > 0 \,\&\, \exists b \in \mathbb{Z}_x^* \; b^{(x-1)/2^i} \not\equiv \pm 1 \pmod{x}$. Fix any such b. Obviously $\neg P(x)$, moreover if we put $c \equiv b^{(x-1)/2^i} \pmod{x}$, we have $x \mid (c^2 - 1) = (c-1)(c+1)$ by minimality of i, but neither $x \mid c-1$ nor $x \mid c+1$. This means that x is not prime, and *a fortiori* $\neg Q(x)$ (as $Q(x)$ would imply $\mathbb{Z}_x^* = [1, x)$). Similarly to the Case 1, we define

$$C(x,v) := \begin{cases} \langle (b(v+1) \bmod x) - 1, 0\rangle, & \text{if } (b(v+1))^{(x-1)/2^i} \equiv \pm 1 \pmod{x}, \\ \langle v, 1\rangle & \text{otherwise.} \end{cases}$$

If $w < r(x)$ and $f(x,w) = *$, we have $(w+1)^{(x+1)/2^i} \equiv \pm 1 \pmod{x}$, hence $C(x,w) = \langle w, 1\rangle$ and $C(x,v) = \langle w, 0\rangle$ for some $v < r(x)$, by essentially the same argument as above.

Case 3: $i > 0 \,\&\, \forall b \in \mathbb{Z}_x^* \; b^{(x-1)/2^i} \equiv \pm 1 \pmod{x}$. We need some elementary number theory.

Claim 1 *PV proves the Chinese Remainder Theorem: if $a = \langle a_j \rangle_{j<\ell}$ is a sequence of pairwise coprime numbers and $b = \langle b_j \rangle_{j<\ell}$, then there is c such that $c \equiv b_j \pmod{a_j}$ for all $j < \ell$.*

Proof: For any $j < \ell$, $(a_j, \prod_{j' \neq j} a_{j'}) = 1$ (by Δ_1^b-*LIND*), hence there is $d_j < a_j$ such that $c_j := d_j \prod_{j' \neq j} a_{j'} \equiv 1 \pmod{a_j}$. Put $c = \sum_{j<\ell} b_j c_j$. We have $c_j \equiv 1 \pmod{a_j}$ and $c_j \equiv 0 \pmod{a_{j'}}$ for all $j' \neq j$, hence $c \equiv b_j \pmod{a_j}$. □ (Claim 1)

Claim 2 *S_2^1 proves that*

(i) $x > 0$ is a prime power iff there are no coprime proper divisors u and v of x such that $uv = x$.

(ii) Any $x > 0$ is uniquely representable as $x = \prod_{j<\ell} p_j^{e_j}$, where $\langle p_j \rangle_{j<\ell}$ is an increasing sequence of primes and each e_j is nonzero.

Proof: Every number $x > 1$ is divisible by a prime. To see this, choose $p > 1$, $p \mid x$ with minimal length (using Δ_1^b-*LENGTH-MIN*). If $p = uv$, $u > 1$, then $v \mid x$ and $|v| < |p|$, hence $v = 1$, i.e., p is a prime.

If $x = p^e$ for a prime p, then any proper divisor of x is divisible by p (by Δ_1^b-*LIND* on e), hence $p \leq (u,v)$ for any $u, v > 1$ which divide x. On the other hand, assume that the right hand side of (i) holds, and w.l.o.g. $x > 1$. Let p be a prime divisor of x, and let $e < |x|$ be maximal such that $p^e \mid x$. If $p^e < x$, we have $(p^e, x/p^e) > 1$, thus $p \mid (x/p^e)$ and $p^{e+1} \mid x$, a contradiction. Hence $x = p^e$ is a prime power.

By Σ_1^b-*LENGTH-MAX*, there is the maximal $k < |x|$ such that there exists a sequence $a = \langle p_j \rangle_{j<k}$ of numbers greater than 1, such that $\prod_{j<k} p_j = x$. Every p_j in any maximal sequence is obviously prime. The sequence of p_j's may be arranged in non-decreasing order, and we may group together occurences of the same prime, yielding $x = \prod_{j<\ell} p_j^{e_j}$ as in the statement of the Claim. If $x = \prod_{j<m} q_j^{f_j}$ is another such representation, Δ_1^b-*LIND* on $j < \min(\ell, m)$ shows that $p_j = q_j$ and $e_j = f_j$, hence also $\ell = m$. □ (Claim 2)

Let us return to the analysis of the Case 3.

First assume that x is not a prime power. Choose coprime $a_1, a_2 > 1$ such that $a_1 a_2 = x$, and $b \in \mathbb{Z}_x^*$ such that $b^{(x-1)/2^i} \equiv -1 \pmod{x}$ (by the definition of i). The Chinese Remainder Theorem gives us c such that $c \equiv 1 \pmod{a_1}$ and $c \equiv b \pmod{a_2}$. We claim that $c \in \mathbb{Z}_x^*$: we have $(b, a_2) = 1$, hence we may find d such that $d \equiv 1 \pmod{a_1}$ and $bd \equiv 1 \pmod{a_2}$, then $cd \equiv 1 \pmod{x}$. By our assumption, $c^{(x-1)/2^i} \equiv 1 \pmod{x}$ or $c^{(x-1)/2^i} \equiv -1 \pmod{x}$. However, the former contradicts $c^{(x-1)/2^i} \equiv -1 \pmod{a_2}$, while the latter contradicts $c^{(x-1)/2^i} \equiv 1 \pmod{a_1}$, because a_1 and a_2 are odd.

We may thus write $x = p^e$, where p is an odd prime. Assume that $e > 1$. Notice that for any u and v, $(up^{e-1}+1)(vp^{e-1}+1) \equiv (u+v)p^{e-1}+1 \pmod{x}$. This gives $(p^{e-1}+1)^u \equiv up^{e-1}+1 \pmod{x}$ by Δ_1^b-*PIND*, in particular $(p^{e-1}+1)^{x-1} \equiv 1-p^{e-1} \pmod{x}$. However $p^{e-1}+1 \in \mathbb{Z}_x^*$, hence $(p^{e-1}+1)^{x-1} \equiv 1 \pmod{x}$. This means $x \mid p^{e-1}$, a contradiction. Therefore $e = 1$ and $x = p$ is a prime.

We have $Q(x)$, because $\mathbb{Z}_p^* = [1, p)$ and $b^{x-1} \equiv 1 \pmod{x}$ for any $b \in \mathbb{Z}_x^*$ by our assumption. Also $P(x)$ holds: if $f(x, w) \neq *$, we would have $b^2 \equiv 1 \pmod{x}$ and $b \not\equiv \pm 1 \pmod{x}$ (where $b \equiv (w+1)^{(x-1)/2^j} \pmod{x}$ for some $j > 0$), which we know is impossible for any prime x. $\square$

1.14 Proposition (Thapen [13]) *Assume that $S_2^1 + dWPHP(PV) \vdash \forall x\, \exists y\, \varphi(x, y)$, where φ is Σ_1^b. Then for any ℓ there are $k \geq \ell$ and PV-function symbols G, g, and h such that*

$$PV \vdash \forall x\, \forall w < 2^{2|x|^k}\, (g(x, w) < 2^{|x|^k}\, \&\, (G(g(x, w)) = w \vee \varphi(x, h(x, w)))).$$

More generally, there are PV-functions g, h, and a constant k such that

$$PV \vdash \forall x\, \forall b \geq 2^{|x|^k}\, \forall w < b^2\, (g(x, w, b) < b\, \&\, (G(g(x, w, b)) = w \vee \varphi(x, h(x, w, b)))).$$

Proof (*sketch*)*:* If $f(x, y)$ is a PV-function, there is a parameter-free PV-function $G(z)$ such that G maps b^4 onto b^8, whenever there are $a, c < b$ such that $f(c, \cdot)$ maps a onto a^2. (This is Lemma 3.8 of [13].) Take a "universal" function (e.g., a circuit evaluator) for f. Our assumption on φ gives

$$S_2^1 \vdash \forall x\, (\exists y\, \varphi(x, y) \vee \exists a, c\, \neg dWPHP(f(c))^a_{a^2}).$$

By Parikh's theorem, all existential quantifiers may be bounded by a term $t(x)$, and the properties of G imply

$$S_2^1 \vdash \forall x\, (\exists y \leq t(x)\, \varphi(x, y) \vee \forall b \geq t(x)^4\, \neg dWPHP(G)^b_{b^2}).$$

We may write this as

$$S_2^1 \vdash \forall x\, \forall w\, \forall b\, (\exists y \leq t(x)\, \varphi(x, y) \vee (b \geq t(x)^4\, \&\, w < b^2 \to \exists v < b\, G(v) = w)),$$

and an application of Buss's witnessing theorem gives us g and h as required. $\square$

1.15 Corollary *The $\forall\Sigma^b_1$-consequences of $S^1_2 + dWPHP(PV)$ can be axiomatized over PV by $dWPHP'(PV)$, where $dWPHP'(f,g)$ denotes the formula*

$$a > 1 \to \exists x < a^2\,(g(x,a) \geq a \vee f(g(x,a),a) \neq x). \qquad \square$$

By A. Wilkie's witnessing theorem (see [4] for a proof), the $\forall\Sigma^b_1$-consequences of the theory $S^1_2 + dWPHP(PV)$ are witnessed by randomized p-time functions (total *MFRP* in our notation). Our next proposition ensures that these witnessing functions can be chosen so that they are definable and provably total in $S^1_2 + dWPHP(PV)$. (Conversely, the statement that certain PV-functions define a uniformly witnessed total *MFRP* is $\forall\Sigma^b_1$.)

1.16 Proposition *Let $\varphi(\vec{x},y)$ be a Σ^b_1-formula such that $\forall\vec{x}\,\exists y\,\varphi(\vec{x},y)$ is provable in $S^1_2 + dWPHP(PV)$. Then for every PV-function s there is $F \in MFRP$ such that*

(i) F has a uniformly witnessed $1/s$-definition in PV,

(ii) F is weakly total in PV (in particular, F is total in $PV + dWPHP(PV)$),

(iii) $PV \vdash F(\vec{x}) = y \to \varphi(\vec{x},y)$.

In particular, every formula which is Δ^b_1 in $S^1_2 + dWPHP(PV)$ is in $PV + dWPHP(PV)$ equivalent to a definable ZPP-predicate.

Proof: Fix ℓ such that $PV \vdash s(x) \leq 2^{|x|^\ell}$, and find $k \geq \ell$, and PV-functions G, g, and h according to the Proposition 1.14. Define

$$r(x) := 2^{|x|^k} \cdot 2^{|x|^k},$$

$$f(x,w) := \begin{cases} h(x,w), & \text{if } G(g(x,w)) \neq w, \\ * & \text{otherwise,} \end{cases}$$

$$m(x,v) := \langle G(v_0), v_1\rangle,$$

where $v = [v_0, v_1] = v_1 \cdot 2^{|x|^k} + v_0$ as in 1.5. We claim that f and r define in PV a weakly total $1/2^{|x|^k}$-*MFRP*, witnessed by m. To see this, let $w < 2^{2|x|^k}$ be such that $f(x,w) = *$, and $i < 2^{|x|^k}$. Put $v = [g(x,w), i]$. Then we have $m(x,v) = \langle w, i\rangle$, because $G(g(x,w)) = w$, and $v < r(x)$, because $g(x,w) < 2^{|x|^k}$.

If we put $F(x) = y$ iff $\exists w < r(x)\ f(x,w) = y \neq *$, then any value y of $F(x)$ satisfies $\varphi(x,y)$, because $y = h(x,w)$ for some $w < r(x)$ such that $G(g(x,w)) \neq w$. $\square$

2 A propositional proof system corresponding to *dWPHP*

In this section, we will present a propositional proof system *WF* which corresponds to the theory $S^1_2 + dWPHP(PV)$, i.e., *WF* is the strongest proof system whose consistency is provable in $S^1_2 + dWPHP(PV)$, and tautologies resulting from translation of $\forall\Pi^b_1$-consequences of $S^1_2 + dWPHP(PV)$ have polynomial-size proofs in *WF*. Obviously, such a system has to contain Extended Frege; we could indeed formulate WF as an extension of EF, but it will be more convenient to use a variant of EF which manipulates Boolean circuits instead of formulas, to get rid of EF's extension axioms. We will describe this variant first[1].

2.1 Definition Any Boolean circuit C can be "unfolded" into a unique (possibly huge) formula φ_C. Circuits C and D are *similar*, written as $C \simeq D$, if φ_C and φ_D are the same formulas.

2.2 Lemma *Similarity of circuits is polynomial-time decidable.*

Proof: As $NLOG \subseteq P$, it suffices to show $\simeq \in coNLOG$, which is clearly accomplished by the following algorithm:

$c \leftarrow$ output node of C, $\quad d \leftarrow$ output node of D
loop
  $\ell_c \leftarrow$ label of c, $\quad \ell_d \leftarrow$ label of d $\quad \{$*connective or variable*$\}$
  if $\ell_c \neq \ell_d$ **then** REJECT
  if ℓ_c is a variable or a constant **then** ACCEPT
  non-deterministically choose i smaller than the arity of ℓ_c
  $c \leftarrow i^{\text{th}}$ input of c, $\quad d \leftarrow i^{\text{th}}$ input of d
end loop

□

2.3 Definition A CF (*circuit Frege*) proof system is defined as follows: choose a finite basis $\mathcal{B}$ of Boolean connectives, and a finite, sound, and implicationally complete set $\mathcal{R}$ of Frege rules over $\mathcal{B}$. A CF-proof of a circuit A is a sequence of $\mathcal{B}$-circuits $A_0, \dots, A_k = A$, such that for every $i \leq k$, either there are $j_1, \dots, j_\ell < i$ such that

$$\frac{A_{j_1} \cdots A_{j_\ell}}{A_i}$$

is an instance of a rule $R \in \mathcal{R}$, or there is $j < i$ such that $A_j \simeq A_i$. (Lemma 2.2 ensures that CF indeed fulfills the definition of a propositional proof system. Also, when we work with CF in bounded arithmetic, we cannot use Definition 2.1 directly as it involves exponentially large objects, we thus use the algorithm from Lemma 2.2 instead.)

2.4 Lemma *Any CF system p-simulates any EF system.*

Proof: All EF systems simulate each other, hence we may assume w.l.o.g. that both proof systems use the same set of connectives and Frege rules. Let $\pi\colon \varphi_0, \dots, \varphi_k$ be an EF-proof, and let

$$q_1 \equiv \psi_1$$
$$q_2 \equiv \psi_2(q_1)$$
$$\dots$$
$$q_\ell \equiv \psi_\ell(q_1, \dots, q_{\ell-1})$$

[1] Although it is folklore that EF is essentially "a Frege system operating with circuits", we were unable to find a reference making this explicit.

be all extension axioms used in π. We define circuits $Q_{i,j}(q_1, \dots, q_j)$, $0 \le j < i \le \ell$, as follows:

$$
\begin{aligned}
Q_{i,i-1}(q_1, \dots, q_{i-1}) &:= \psi_i(q_1, \dots, q_{i-1}), \\
Q_{i,j-1}(q_1, \dots, q_{j-1}) &:= Q'_{i,j}(q_1, \dots, q_{j-1}, \psi_j(q_1, \dots, q_{j-1})),
\end{aligned}
$$

where $Q'_{i,j}$ differs from $Q_{i,j}$ by joining all occurrences of q_j together. We put $Q_i := Q_{i,0}$. It is easy to see that $Q_i \simeq \psi_i(Q_1, \dots, Q_{i-1})$.

We modify the proof π by putting a (constant size) Frege proof of $q_i \equiv q_i$ before every extension axiom $q_i \equiv \psi_i$, and then we substitute circuits $Q_1, \dots, Q_\ell$ for variables $q_1, \dots, q_\ell$ in the whole proof. This makes up a correct CF proof π': substitution does not break Frege rules, and extension axioms translate to circuits $Q_i \equiv \psi_i(Q_1, \dots, Q_{i-1})$, each preceded by a similar circuit $Q_i \equiv Q_i$.

The size of $Q_{i,j}$ is bounded by $|\psi_{j+1}| + \dots + |\psi_i|$, in particular the size of Q_i is bounded by $|\pi|$, hence the size of π' is $O(|\pi|^2)$. □

2.5 Lemma *Any EF system p-simulates proofs of formulas in any CF system.*

Proof: Let $\pi\colon A_0, \dots, A_k = \varphi$ be a CF proof, where φ is a formula. We assign an extension variable $q_i =: q[C]$ to each subcircuit C of each A_j in such a way that similar circuits get the same variable, and every circuit gets a variable with higher index than all its subcircuits. The EF proof π' will start with extension axioms for q_i's, which describe the relation of the corresponding circuits to their subcircuits. For example, if $C = p_1 \lor \neg(p_2 \to p_1)$, we could have

$$
\begin{aligned}
&q_1 \equiv p_1 \\
&q_2 \equiv p_2 \\
&q_3 \equiv q_2 \to q_1 \\
&q_4 \equiv \neg q_3 \\
&q_5 \equiv q_1 \lor q_4
\end{aligned}
$$

Then we extend the proof to contain the sequence $q[A_0], \dots, q[A_k]$. If $A_i \simeq A_j$, $j < i$, we have nothing to do, because $q[A_i] = q[A_j]$. Assume that $A_i = \chi(B_1, \dots, B_m)$ was inferred by a Frege rule R from $A_{j_1} = \psi_1(B_1, \dots, B_m)$, $\dots$, $A_{j_\ell} = \psi_\ell(B_1, \dots, B_m)$, where $j_1, \dots, j_\ell < i$. There is a constant size Frege proof of

$$
\begin{gathered}
q[A_{j_1}] \equiv \psi_1(q[B_1], \dots, q[B_m]) \\
\dots \\
q[A_{j_\ell}] \equiv \psi_\ell(q[B_1], \dots, q[B_m]) \\
q[A_i] \equiv \chi(q[B_1], \dots, q[B_m])
\end{gathered}
$$

from the extension axioms. By the induction hypothesis our proof already contains the formulas $q[A_{j_1}],\dots,q[A_{j_\ell}]$, hence we get a proof of

$$\begin{gathered}\psi_1(q[B_1],\dots,q[B_m])\\ \dots\\ \psi_\ell(q[B_1],\dots,q[B_m])\\ \chi(q[B_1],\dots,q[B_m])\\ q[A_i]\end{gathered}$$

by a constant-size simulation of R and Modus Ponens (or rather its variant for $\equiv$).

We thus have an $O(|\pi|)$ proof of $q[\varphi]$, and we finish it by an $O(|\varphi|^2)$ proof of $q[\varphi]\equiv\varphi$ and Modus Ponens. $\square$

2.6 Definition The *WF* (*WPHP Frege*) proof system is defined as follows: a *WF*-proof of a circuit A is a sequence of circuits $A_0,\dots,A_k$ such that $A_k=A$, and every A_i is inferred from some $A_{j_1},\dots,A_{j_\ell}$, $j_1,\dots,j_\ell<i$ by a Frege rule, or it is similar to some A_j, $j<i$, or it is a special axiom

$$\bigvee_{\ell=1}^{m}(r_\ell\not\equiv C_{i,\ell}(D_{i,1},\dots,D_{i,n})),$$

where $n<m$, and r_ℓ are pairwise distinct variables which do not occur in A, $C_{i,\ell'}$, or A_j for $j<i$, but may occur in $D_{i,1},\dots,D_{i,n}$.

2.7 Remark In principle, different choices of connectives and Frege rules give different variants of *WF*. We ignore this ambiguity, as all such systems are polynomially equivalent.

We will see in 2.13 that we could restrict *WF*-proofs to contain only *one* special axiom, and still get an equivalent system. On the other hand, we could allow special axioms with the same C's to share the same sequence of special variables: the proof of 2.8 can be easily modified to show the consistency of such a system in $S^1_2+dWPHP(PV)$, hence it is polynomially equivalent to the original *WF* by 2.13.

2.8 Proposition *$S^1_2+dWPHP(PV)$ proves 0-RFN(WF).*

Proof: Let $\pi=\langle A_0,\dots,A_k\rangle$ be a *WF*-proof of a circuit $A=A_k$, and let e be a truth assignment to the variables occurring in A. W.l.o.g. we may assume that every variable in π either occurs in A, or it is a special variable of a *WF*-axiom from π. We will show by induction on $i\le k<|\pi|$ that there is an assignment $e'\supseteq e$, which makes A_j true for every $j\le i$ (this is Σ^b_1-*LIND*).

If A_i is inferred by a Frege rule from $A_{j_1},\dots,A_{j_\ell}$, $j_1,\dots,j_\ell<i$, the induction step from $i-1$ to i is easy because the rule is sound: its verification consists of checking only finitely many cases involving the inductive definition of satisfaction for some top-level subcircuits of the A_j's, hence it goes through in S^1_2.

If $A_i\simeq A_j$, $j<i$, we have $e'(A_i)=e'(A_j)$ by induction on the depth of the circuit.

Assume that A_i is the special axiom $\bigvee_{j=1}^m(r_j\not\equiv C_j(D_1,\dots,D_n))$. Notice that the truth value of all variables occurring in $C_j(s_1,\dots,s_n)$ is fixed by e', except for the placeholders

$s_1, \ldots, s_n$ (the definition of WF implies that special variables from $A_{i'}$, $i' \geq i$, cannot occur in C_j). Hence the sequence of circuits $C = \langle C_1, \ldots, C_m \rangle$ computes a function $g\colon 2^n \to 2^m$. More precisely, there is a PV-function symbol $f(u, v, x)$ with the following property: if u is a sequence of circuits, and v a partial truth assignment, then the j-th bit of $f(u, v, x)$ is $u_j(a)$, where a extends v and the j'-th variable not assigned by v is given the value $bit(x, j')$ by a. Then we put $g(x) = f(\langle C_1, \ldots, C_m \rangle, e', x)$. By definition $n < m$, i.e. $2 \cdot 2^n \leq 2^m$, hence $dWPHP(PV)$ implies that there is $y < 2^m$ such that $y \neq g(x)$ for any $x < 2^n$. We extend e' by putting $e'(r_j) = bit(y, j-1)$, and we claim that $e'(A_i) = 1$: if $x < 2^n$ is such that $bit(x, j') = e'(D_{j'+1})$, then the value of $C_j(\vec{D}_{j'})$ under e' is $bit(g(x), j-1)$, which is distinct from $e'(r_j)$ for some $j \leq m$. $\square$

Recall that G is a propositional proof system operating with quantified Boolean formulas, defined (in [6]) as an extension of the usual Gentzen sequent calculus by rules for introducing existential and universal quantifiers. G_2 is a fragment of G, which allows only sequents consisting of Σ^q_2-formulas (these are, roughly, formulas of the form $\exists x_1 \cdots \exists x_k \forall y_1 \cdots \forall y_\ell\, \varphi$, with φ quantifier-free).

2.9 Corollary *G_2 polynomially simulates WF.*

Proof: By [11] (see also [8], and Chapter 11.2 of [4]), T^2_2 proves $dWPHP(PV)$, hence also $T^2_2 \vdash 0\text{-}RFN(WF)$. By [6], this implies $S^1_2 \vdash WF \leq_p G_2$. See also [4], Chapters 9.2, 9.3. $\square$

Recall the definition of the $\|\varphi\|$ translation of $\Pi^b_1(PV)$-formulas into propositional logic: first, we assign to every PV-function $f(x_1, \ldots, x_k)$ and numbers $n_1, \ldots, n_k$ a (p-size) circuit $\{\!\{f\}\!\}^{\vec{n}}(\vec{p})$, which computes the restriction $f\colon 2^{n_1} \times \cdots \times 2^{n_k} \to 2^{b(n_1,\ldots,n_k)}$, where b is a bounding polynomial to f. The formula

$$\|f\|^{\vec{n}}(\vec{p}; \vec{r}; \vec{q})$$

expresses that the circuit $\{\!\{f\}\!\}^{\vec{n}}$ computes $\vec{r}$ on input $\vec{p}$, with $\vec{q}$ being the intermediate steps of the computation (there is an atom q_i for every node of the circuit).

Then we define Boolean formulas $\|\varphi(\vec{x})\|^{\vec{n}}(\vec{p}; \vec{q})$ by induction on complexity of a $\Pi^b_1(PV)$-formula φ. (Atoms $\vec{p}$ correspond to the variables $\vec{x}$. Atoms $\vec{q}$ are auxiliary, you may think of them as being universaly quantified; they arise from universal quantifiers of φ, and from the output and intermediate atoms $\vec{q}$, $\vec{r}$ of $\|f\|^{\vec{n}}$, for functions f appearing in φ). The induction steps are straightforward, and for atomic formulas and their negations we have

$$\|f(\vec{x}) = g(\vec{x})\|^{\vec{n}} := \|f\|^{\vec{n}}(\vec{p}; \vec{r}; \vec{q}) \mathbin{\&} \|g\|^{\vec{n}}(\vec{p}; \vec{r'}; \vec{q'}) \to \bigwedge_i (r_i \equiv r'_i),$$

$$\|\neg f(\vec{x}) = g(\vec{x})\|^{\vec{n}} := \|f\|^{\vec{n}}(\vec{p}; \vec{r}; \vec{q}) \mathbin{\&} \|g\|^{\vec{n}}(\vec{p}; \vec{r'}; \vec{q'}) \to \neg \bigwedge_i (r_i \equiv r'_i),$$

$$\|f(\vec{x}) \leq g(\vec{x})\|^{\vec{n}} := \|f\|^{\vec{n}}(\vec{p}; \vec{r}; \vec{q}) \mathbin{\&} \|g\|^{\vec{n}}(\vec{p}; \vec{r'}; \vec{q'}) \to \bigwedge_i \bigl(r_i \mathbin{\&} \bigwedge_{j>i} (r_j \equiv r'_j) \to r'_i\bigr),$$

$$\|\neg f(\vec{x}) \leq g(\vec{x})\|^{\vec{n}} := \|f\|^{\vec{n}}(\vec{p}; \vec{r}; \vec{q}) \mathbin{\&} \|g\|^{\vec{n}}(\vec{p}; \vec{r'}; \vec{q'}) \to \neg \bigwedge_i \bigl(r_i \mathbin{\&} \bigwedge_{j>i} (r_j \equiv r'_j) \to r'_i\bigr).$$

In this translation, it is necessary to encode the computation of the circuit $\{\!\{f\}\!\}^{\vec{n}}$ by a formula introducing extra auxiliary variables, as it is unlikely that $PV \vdash P \subseteq NC^1$. This seems to obfuscate things a bit, and we will use a proof system handling Boolean circuits directly, we thus avoid this inconvenience by introducing a more natural modified translation, which produces circuits instead of formulas. It is defined as follows:

2.10 Definition Let $\varphi(\vec{x})$ be a $\Pi_1^b(PV)$-formula, and $b(\vec{x})$ its bounding polynomial. We define a Boolean circuit $\{\!\{\varphi\}\!\}^{\vec{n}}(\vec{p};\vec{q})$ by induction on complexity of φ:

$$
\begin{aligned}
\{\!\{f(\vec{x}) = g(\vec{x})\}\!\}^{\vec{n}}(\vec{p}) &:= \bigwedge_{i<b(\vec{n})} (\{\!\{f\}\!\}_i^{\vec{n}}(\vec{p}) \equiv \{\!\{g\}\!\}_i^{\vec{n}}(\vec{p})),\\
\{\!\{f(\vec{x}) \le g(\vec{x})\}\!\}^{\vec{n}}(\vec{p}) &:= \bigwedge_{i<b(\vec{n})} \big(\{\!\{f\}\!\}_i^{\vec{n}} \,\&\, \bigwedge_{j>i}(\{\!\{f\}\!\}_j^{\vec{n}} \equiv \{\!\{g\}\!\}_j^{\vec{n}}) \to \{\!\{g\}\!\}_i^{\vec{n}}\big),\\
\{\!\{\neg\varphi\}\!\}^{\vec{n}}(\vec{p}) &:= \neg\{\!\{\varphi\}\!\}^{\vec{n}}(\vec{p}), \quad \varphi \in \Sigma_0^b(PV),\\
\{\!\{\varphi \circ \psi\}\!\}^{\vec{n}}(\vec{p};\vec{q}) &:= \{\!\{\varphi\}\!\}^{\vec{n}}(\vec{p};\vec{q}) \circ \{\!\{\psi\}\!\}^{\vec{n}}(\vec{p};\vec{q}), \quad \circ \in \{\&, \vee\},\\
\{\!\{\forall y \le |t(\vec{x})|\, \varphi(\vec{x},y)\}\!\}^{\vec{n}}(\vec{p};\vec{q^0},\dots,\vec{q^m}) &:= \bigwedge_{j\le m} \{\!\{y \le |t(\vec{x})| \to \varphi(\vec{x},y)\}\!\}^{\vec{n},|m|}(\vec{p},\vec{\varepsilon};\vec{q^j}),\\
\{\!\{\exists y \le |t(\vec{x})|\, \varphi(\vec{x},y)\}\!\}^{\vec{n}}(\vec{p};\vec{q^0},\dots,\vec{q^m}) &:= \bigvee_{j\le m} \{\!\{y \le |t(\vec{x})| \,\&\, \varphi(\vec{x},y)\}\!\}^{\vec{n},|m|}(\vec{p},\vec{\varepsilon};\vec{q^j}),\\
\{\!\{\forall y \le t(\vec{x})\, \varphi(\vec{x},y)\}\!\}^{\vec{n}}(\vec{p};\vec{q},\vec{p'}) &:= \{\!\{y \le t(\vec{x}) \to \varphi(\vec{x},y)\}\!\}^{\vec{n},m}(\vec{p},\vec{p'};\vec{q}),
\end{aligned}
$$

where $m = m(\vec{n})$ is a bounding polynomial to $t(\vec{x})$, and $\vec{\varepsilon}$ is the representation of j as a sequence of $|m|$ binary digits (= truth constants). Notice that auxiliary variables $\vec{q}$ are introduced only for (non-sharply) bounded universal quantifiers.

2.11 Lemma *Let $\varphi(\vec{x}) \in \Pi_1^b(PV)$. There are circuits $\vec{C}_\varphi^{\vec{n}}$, and a p-time constructible sequence of CF-proofs of*

$$
\begin{aligned}
\{\!\{\varphi\}\!\}^{\vec{n}}(\vec{p};\vec{q}) &\to \|\varphi\|^{\vec{n}}(\vec{p};\vec{q},\vec{q'}),\\
\|\varphi\|^{\vec{n}}(\vec{p};\vec{q},\vec{C}_\varphi^{\vec{n}}(\vec{p},\vec{q})) &\to \{\!\{\varphi\}\!\}^{\vec{n}}(\vec{p};\vec{q}).
\end{aligned}
$$

Proof: This follows by straightforward induction on complexity of φ. We need the following property for the base case: for any PV-function f, there are circuits $\vec{C}_f^{\vec{n}}$, and p-time constructible CF-proofs of

$$
\begin{aligned}
&\|f\|^{\vec{n}}(\vec{p};\{\!\{f\}\!\}^{\vec{n}}(\vec{p});\vec{C}_f^{\vec{n}}(\vec{p})),\\
&\|f\|^{\vec{n}}(\vec{p};\vec{r};\vec{q}) \to \bigwedge_i (r_i \equiv \{\!\{f\}\!\}_i^{\vec{n}}(\vec{p})).
\end{aligned}
$$

We may take subcircuits of $\{\!\{f\}\!\}$ for $\vec{C}_f$. The second part essentially states that the computation of $\{\!\{f\}\!\}$ is unique, and its proof in CF may be constructed by induction on the size of $\{\!\{f\}\!\}$. $\square$

2.12 Proposition *If $S^1_2 + dWPHP(PV) \vdash \forall x\, \varphi(x)$, where $\varphi \in \Pi^b_1$, then tautologies $\|\varphi\|^n$ have polynomial size WF-proofs. Actually, these proofs are constructible by a p-time function, and PV proves this fact.*

Proof: Assume that $S^1_2 + dWPHP(PV) \vdash \forall x\, \varphi(x)$, where $\varphi \in \Sigma^b_0(PV)$ for simplicity. By Proposition 1.14, there is a constant k, and PV-functions G and g such that

$$PV \vdash 2^{|x|^k} \le b \,\&\, w < b^2 \to (G(g(x,w,b)) = w \vee \varphi(x)),$$
$$PV \vdash g(x,w,b) < b.$$

Given n (bounding x), and $m = 2n^k$ (bounding w and b), there are $poly(n)$-size CF-proofs (constructible in PV) of the circuits

$$\{\!\{2^{|x|^k} \le b\}\!\}^{n,m}(\vec{p},\vec{q}) \,\&\, \{\!\{w < b^2\}\!\}^{m,m}(\vec{r},\vec{q}) \to$$
$$\to \{\!\{G(g(x,w,b)) = w\}\!\}^{n,m,m}(\vec{p},\vec{r},\vec{q}) \vee \{\!\{\varphi(x)\}\!\}^n(\vec{p}),$$

$$\{\!\{g(x,w,b) < b\}\!\}^{n,m,m}(\vec{p},\vec{r},\vec{q}),$$

using the simulation of PV by EF [3], and Lemmas 2.4, 2.11. We substitute the binary representation of $b := 2^{n^k}$ for the variables $\vec{q}$, i.e., $q_{n^k} = 1$, $q_j = 0$ for $j \ne n^k$. Then there are poly-size CF-proofs of $\{\!\{2^{|x|^k} \le b\}\!\}^{n,m}$ and $\{\!\{w < b^2\}\!\}^{m,m}$, hence by modus ponens

$$\{\!\{G(g(x,w,b)) = w\}\!\}^{n,m,m} \vee \{\!\{\varphi(x)\}\!\}^n,$$

which is the circuit

$$\bigwedge_{i<m} (r_i \equiv \{\!\{G\}\!\}^{q(n)}_i(\{\!\{g\}\!\}_0, \dots, \{\!\{g\}\!\}_{q(n)-1})) \vee \{\!\{\varphi(x)\}\!\}^n,$$

where $q(n)$ is the bounding polynomial for g. However, $\{\!\{g(x,w,b) < b\}\!\}^{n,m,m}$ implies

$$\bigwedge_{i=n^k}^{q(n)-1} \neg\{\!\{g\}\!\}^{n,m,m}_i,$$

thus we get a proof π of

$$\bigwedge_{i<m} (r_i \equiv \{\!\{G\}\!\}^{q(n)}_i(\{\!\{g\}\!\}_0, \dots, \{\!\{g\}\!\}_{n^k-1}, 0, \dots, 0)) \vee \{\!\{\varphi(x)\}\!\}^n.$$

If we define $C_j(s_0, \dots, s_{n^k-1}) = \{\!\{G\}\!\}^{q(n)}_j(\vec{s}, 0, \dots, 0)$, and $D_i(\vec{p},\vec{r}) = \{\!\{g\}\!\}^{n,m,m}_i(\vec{p},\vec{r},\vec{q})$, we may rewrite this as

$$\bigwedge_{i<m} (r_i \equiv C_i(D_0, \dots, D_{n^k-1})) \vee \{\!\{\varphi(x)\}\!\}^n.$$

Since $m = 2n^k > n^k$ for every $n > 0$, and C_j does not contain any of the $r_{j'}$, we may put a special axiom

$$\bigvee_{i<m} (r_i \not\equiv C_i(D_0, \dots, D_{n^k-1}))$$

before the first line of π, and we finish the proof by De Morgan rules and modus ponens to get a *WF*-proof of

$$\{\!\{\varphi(x)\}\!\}^n.$$

Lemma 2.11, and another modus ponens give

$$\|\varphi(x)\|^n. \qquad \square$$

2.13 Corollary

(*i*) *For any Π_1^b-formula $\varphi(x)$, $S_2^1 + dWPHP(PV) \vdash (WF \vdash \|\varphi\|^{|x|}) \to \varphi(x)$.*

(*ii*) *$PV + Con(WF)$ axiomatizes strict $\forall\Pi_1^b$-consequences of $S_2^1 + dWPHP(PV)$.*

(*iii*) *If $S_2^1 + dWPHP(PV) \vdash$ 0-$RFN(P)$, where P is a propositional proof system, then $PV \vdash (P \leq_p WF)$.*

(*iv*) *WF is polynomially simulated by a modified WF proof system, in which we allow only the first formula of the proof to be a special axiom.*

Proof: (*i*) follows from 2.8 together with $S_2^1 \vdash Taut(\|\varphi\|^{|x|}) \to \varphi(x)$.

(*ii*): if $\varphi \in strict\Pi_1^b$, the formula $Taut(\|\varphi\|^{|x|}) \to \varphi(x)$ just mentioned is provable already in PV, and $Con(WF)$ implies 0-$RFN(WF)$ as *WF* is provably closed under substitution and modus ponens. This, together with 2.12, shows the harder inclusion of (*ii*), the other one follows from 2.8.

(*iii*): we have $PV \vdash (WF \vdash \{\!\{P(p) = f \to Taut(f)\}\!\})$ by 2.12, and it is easy to see that $PV \vdash (P(\pi) = \varphi \to CF \vdash \{\!\{P(p) = f\}\!\}(\pi, \varphi))$ and $PV \vdash (WF \vdash \{\!\{Taut\}\!\}(\varphi) \to WF \vdash \varphi)$, hence $PV \vdash (P(\pi) = \varphi \to WF \vdash \varphi)$.

(*iv*): the proof of (*iii*) works for the modified *WF*-system from (*iv*) as well, because the proof constructed in 2.12 used only one special axiom; then (*iv*) follows from 2.8. $\square$

2.14 Definition Let p be a prime. *Unstructured Extended Nullstellensatz* of [2] is a proof system for multivariate polynomials over $\mathbb{Z}_p$: a $UENS_p$-refutation of a set of polynomials $f_0, \dots, f_{n-1} \in \mathbb{Z}_p[x_0, \dots, x_{m-1}]$ shows that the f_i's do not have a 0–1 solution (i.e., a common zero at a point from $\{0,1\}^m$). A $UENS_p$-refutation is given by two sequences of polynomials $g_0, \dots, g_{\ell-1}$ and $g'_0, \dots, g'_{\ell+n+m-1}$, such that

$$\sum_{i<\ell} g_i g'_i + \sum_{i<n} f_i g'_{i+\ell} + \sum_{i<m} (x_i^2 - x_i) g'_{i+\ell+n} = 1,$$

and each g_i has the form

$$\prod_{j<k} (h_{i,j} - r_{i,j}),$$

where $r_{i,j}$ are pairwise distinct variables not occurring among $x_0, \dots, x_{m-1}$, $h_{i,j}$ does not contain any of $r_{i,0}, \dots, r_{i,k-1}$, and $\ell < e^{k/p}$ (where e is the Euler number).

The *UENS* proof system simulates Extended Frege, but the converse is an open problem. In fact, it was not clear whether *any* "traditional" proof system simulates *UENS*. We show that it is possible to simulate *UENS* in *WF* (hence also in G_2).

2.15 Proposition *For any prime p, the WF proof system polynomially simulates $UENS_p$.*

Proof: By 2.13 it suffices to prove the soundness of $UENS_p$ in $S^1_2 + dWPHP(PV)$. It is not clear how to express base e exponentiation in bounded arithmetic, however we may simply relax the last condition of 2.14 to $\ell < \beta^{k/p}$, where β is any fixed rational such that $e < \beta < (1-1/p)^{-p}$.

Consider any $UENS_p$-refutation as in 2.14, and assume for contradiction that $f_i(\vec{a}) = 0$ for all $i < n$, with $a_j \in \{0,1\}$. Put $t = k\ell$. W.l.o.g. we assume that every variable in g_i and g'_i is one of x_j or $r_{i',j}$. We will find an assignment $b_{0,0}, \dots, b_{\ell-1,k-1} \in \mathbb{Z}_p$ to $\{r_{i,j}\}_{i<\ell,j<k}$ such that $g_i(\vec{a}, \vec{b}) = 0$ for all i, then

$$\sum_{i<\ell} g_i(\vec{a},\vec{b})g'_i(\vec{a},\vec{b}) + \sum_{i<n} f_i(\vec{a})g'_{i+\ell}(\vec{a},\vec{b}) + \sum_{i<m}(a_i^2 - a_i)g'_{i+\ell+n}(\vec{a},\vec{b}) = 0 \neq 1,$$

contradicting the definition of a $UENS_p$-proof.

We define a function

$$F \colon \ell \times (p-1)^k \times p^{t-k} \to p^t$$

by $F(i, u_0, \dots, u_{k-1}, v_0, \dots, v_{t-k-1}) = \langle b_{0,0}, \dots, b_{\ell-1,k-1}\rangle$, where $b_{i',j}$ are assigned according to $\vec{v}$ if $i' \neq i$, and

$$b_{i,j} = \begin{cases} u_j, & \text{if } u_j < h_{i,j}(\vec{a},\vec{b}), \\ u_j + 1, & \text{otherwise.} \end{cases}$$

Notice that the value of $h_{i,j}(\vec{a},\vec{b})$ depends only on $\vec{v}$, as $r_{i,0}, \dots, r_{i,k-1}$ do not occur in $h_{i,j}$.

It is clear from the definition that the values of $F(i, \bullet)$ are exactly the assignments $\vec{b}$ such that $g_i(\vec{a},\vec{b}) \neq 0$, hence it suffices to show that $\operatorname{rng}(F) \neq p^t$. Choose a rational constant $\alpha > 1$ such that $\beta\alpha^p < (p/(p-1))^p$. Then $\alpha\ell(p-1)^k p^{t-k} < \beta^{k/p}\alpha^k(p-1)^k p^{t-k} < p^t$, hence F is not onto by $dWPHP(PV)^x_{\alpha x}$. □

2.16 Remark By an easy modification of the proof of 2.15, we could simulate a slightly stronger system than *UENS*: the extension variables $r_{i,j}$ could be reused in $g_{i'}$, $i' \neq i$, and we could allow $r_{i,0}, \dots, r_{i,j-1}$ to occur in $h_{i,j}$. (However, it is quite possible that this modification is polynomially equivalent to the original *UENS*.)

3 Hard Boolean functions

3.1 Definition Let $\varepsilon > 0$. A number x (viewed as an n-bit binary string, $n = |x|$) is *ε-hard,* if there is no Boolean circuit C on $|n|$ variables such that $|C| \leq n^\varepsilon$, and $C(u) = bit(x,u)$ for all $u < n$. We write $Hard_\varepsilon(x)$ in such a case.

A Boolean function f on $k \in LogLog$ variables is identified with its truth table, i.e., a 2^k-bit number.

A function f is *ε-hard on average* (abbreviated $Hard^{\varnothing}_{\varepsilon}(f)$), if there does not exist a circuit C of size $|C| \le 2^{\varepsilon k}$ which approximates f, i.e., $|\{u < 2^k;\, C(u) = f(u)\}| \ge (1/2 + 2^{-\varepsilon k})2^k$.

Notice that $Hard_{\varepsilon}(x)$ and $Hard^{\varnothing}_{\varepsilon}(f)$ are Π^b_1.

3.2 Lemma ($PV + dWPHP(PV) \vdash$:) *For every $n \in Log$, there is an x of length n such that x cannot be computed by a circuit of size $n/(2|n|)$.*

Proof: Let $e\colon 2^{n-1} \to 2^n$ be a PV-function, which interprets its input as a circuit on $|n|$ variables, and outputs the truth table of the circuit. By $dWPHP(e)$ there is an $x \in 2^n \smallsetminus \mathrm{rng}(e)$. Since any circuit of size $m = n/(2|n|)$ may be described by a number of length at most $2m(|m|+1) \le n-1$, x is not computable by a circuit of size $\le m$. □

3.3 Corollary ($PV + dWPHP(PV) \vdash$:) *For every $k \in LogLog$, there is a Boolean function $f\colon 2^k \to 2$ such that $Hard_{1-o(1)}(f)$.* □

3.4 Lemma ($PV + dWPHP(PV) \vdash$:) *For any $k \in LogLog$, there are $(1/3 - o(1))$-hard on average functions $f\colon 2^k \to 2$.*

Proof: Put $n = 2^k$, and $m = (n/k)^{1/3}$. Consider the function

$$g\colon 2^{2m|m|} \times \sum_{i=0}^{n(1/2-1/m)} \binom{n}{i} \to 2^n,$$

whose first argument is a circuit $C\colon 2^k \to 2$ of size m, its second argument is a string $x \in 2^n$ containing at most $n(1/2 - 1/m)$ 1's, and its output is the truth-table of C XOR'ed by x. Clearly, a function $f\colon 2^k \to 2$ is $(1/3 - |k|/k)$-hard on average if $f \notin \mathrm{rng}(g)$. By Chernoff's inequality, provable in PV by Proposition A.5, the domain of g is a number bounded by

$$d2^{\frac{2}{3}n^{1/3}k^{2/3}}2^n2^{-2n^{4/3}k^{2/3}n^{-1}} = d2^{n-\frac{4}{3}n^{1/3}k^{2/3}}$$

for some constant d. Since $d2^{n-\frac{4}{3}n^{1/3}k^{2/3}} < 2^{n-1}$ for $n \gg 0$, the function g cannot be onto, by $dWPHP(PV)$. □

3.5 Proposition ($S^1_2 \vdash$:) *Assume that $dWPHP(PV)$ fails. Then there is $s \in Log$ such that every string x is computable by a circuit of size at most s.*

Proof: Let $h\colon 2^m \twoheadrightarrow 2^{2m}$ be a surjection, computable by a circuit C. For any $i \in LogLog$, h may be amplified in i steps into a surjection $2^m \twoheadrightarrow (2^m)^{2^i}$, and this will allow us to express any $x \in 2^{2^i m}$ by a circuit of size $O(|C|i)$.

Let $D\colon 2 \times 2^m \to 2^m$ be the circuit defined by $D_j(b,y) = (\neg b \,\&\, C_j(y)) \vee (b \,\&\, C_{j+m}(y))$ for all $j < m$, where v_j is a shorthand for $bit(v,j)$. Fix $i \in LogLog$, and define a sequence of circuits $E^k\colon 2^k \times 2^m \to 2^m$, $k \le i$ by

$$\begin{aligned} E^0(0,y) &:= y,\\ E^{k+1}(u,y) &:= E^k(u \restriction k, D(u_k, y)), \quad \text{where } u \restriction k = u \bmod 2^k, \end{aligned}$$

and put $E := E^i$. Notice that the size of E is bounded by $i|D|$. We claim that E represents an onto map $2^m \twoheadrightarrow 2^{2^i m}$ in the following sense: for any $x < 2^{2^i m}$, there is $y < 2^m$ such that $E_j(u, y) = x_{um+j}$ holds for every $u < 2^i$ and $j < m$. Indeed, we show by induction on $k \leq i$ that there is a sequence w of numbers less than 2^m such that

$$lh(w) = 2^{i-k} \;\&\; \forall v < 2^{i-k}\, \forall u < 2^k\, \forall j < m \;\; E^k_j(u, (w)_v) = bit(x, (v2^k + u)m + j).$$

(This is Σ^b_1, because $i \in LogLog$, i.e., all universal quantifiers are sharply bounded.) The base step is trivial, we simply view x as a sequence of 2^i numbers less than 2^m. Assume that we have found a suitable w for $k < i$. Since C is onto, there is a sequence w' such that $C((w')_v) = [(w)_{2v}, (w)_{2v+1}]$ for any $v < 2^{i-k-1}$ (using $BB\Sigma^b_1$). We claim that w' works for $k+1$: given numbers $v < 2^{i-k-1}$, $u < 2^{k+1}$, and $j < m$, we have

$$\begin{aligned} E^{k+1}_j(u, (w')_v) &= E^k_j(u \restriction k, D(u_k, (w')_v)) = E^k_j(u \restriction k, (w)_{2v+u_k}) = \\ &= bit(x, ((2v + u_k)2^k + u \restriction k)m + j) = bit(x, (v2^{k+1} + u)m + j). \end{aligned}$$

Let $x < 2^{2^i m}$, and let $y < 2^m$ be its "inverse image" as described above. We may construct a small Boolean circuit $B\colon 2^{|n|} \to 2$ computing x as follows: $B(u) = E_{u \bmod m}(\lfloor \frac{u}{m} \rfloor, y)$. For simplicity, we may assume that m is a power of two, which means that the size of B is bounded by $2m|m| + i|D|$.

In other words, any x of length n is computable by a circuit of size $\leq 2m|m| + |D| \cdot |\lceil n/m \rceil| \leq c|n|$ for a suitable $c \in Log$. Take any $d \in Log \smallsetminus LogLog$ (this is possible, because $S^1_2 + Exp \vdash dWPHP(PV)$). Then $d > |n|$, hence x is computable by a circuit of size at most $s := c \cdot d \in Log$. $\square$

3.6 Corollary *Let $0 < \varepsilon < 1$. There exists a standard constant c such that the following are equivalent over S^1_2:*

(*i*) $dWPHP(PV)$,

(*ii*) $\forall k \in LogLog\;\; (k \geq c \to \exists f\colon 2^k \to 2\;\; Hard_\varepsilon(f))$,

(*iii*) $\forall k_0 \in LogLog\;\; \exists k \in LogLog\;\; (k \geq k_0 \;\&\; \exists f\colon 2^k \to 2\;\; Hard_\varepsilon(f))$.

The same holds for hard on average functions, if $\varepsilon < 1/3$.

Proof: (*i*) $\to$ (*ii*) follows from 3.3 and 3.4, (*ii*) $\to$ (*iii*) is trivial. The implication (*iii*) $\to$ (*i*) follows from 3.5, because numbers $2^{\varepsilon k}$, $k \in LogLog$, are cofinal in Log for any fixed ε. $\square$

3.7 Corollary *There is a PV-function $C(a, x)$ such that $PV + \neg dWPHP'(PV)$ proves*

$$\exists a\, \forall x\;\; C(a, x) \text{ is a circuit of size } \leq |a| \text{ computing } x.$$

Actually, a can be itself computed by a PV-function from a counterexample to $dWPHP'(PV)$.

Proof: Let g and h be counterexamples to $dWPHP'(PV)$, i.e., $h\colon b \to b^2$, $g\colon b^2 \to b$, $g \circ h = id$. Given x, we proceed as in the proof of 3.5 to construct a small circuit for x, but instead of nondeterministically guessing preimages under h, we use g to find them explicitly (this way we also get rid of $BB\Sigma^b_1$, and Σ^b_1-$LIND$).

Alternatively, we may use Buss's witnessing theorem. Proposition 3.5 tells us

$$S^1_2 \vdash \exists b\, \forall v < 2b\, (h(v) < b \mathbin{\&} g(h(v)) = v) \to \exists S\, \forall x\, \exists C \leq S\, (C \text{ computes } x),$$

and it is easy to see from its proof that S is actually bounded by a term $t(b)$, thus

$$S^1_2 \vdash \forall b \forall x\, (\exists v < 2b\, (h(v) \geq b \vee g(h(v)) \neq v) \vee \exists C \leq t(b)\, (C \text{ computes } x)).$$

The formula in parenthesis is Σ^b_1, hence there is a PV-function f such that

$$\begin{aligned} PV \vdash (f(b,x) < 2b \mathbin{\&} (h(f(b,x)) \geq b \vee g(h(f(b,x))) \neq f(b,x)) \vee \\ \vee\, (f(b,x) \leq t(b) \mathbin{\&} f(b,x) \text{ computes } x), \end{aligned}$$

which means

$$PV + \neg dWPHP'(PV) \vdash \exists b\, \forall x\, (f(b,x) \leq t(b) \mathbin{\&} f(b,x) \text{ computes } x).$$

It suffices to define $C(a,x) = \min\{f((a)_0, x), a\}$, as we can take $a = \langle b, t(b) \rangle$.

Notice that the converse to this corollary holds too, in a similar fashion to Lemma 3.2. $\square$

4 The Nisan-Wigderson generator

This section presents a derandomization result for definable probabilistic algorithms within bounded arithmetic. We will follow closely the Nisan-Wigderson construction [9]; however, we will present the derandomization in a relativized form: rather than postulating the existence of an explicit language in E with exponential average-case hardness, we will use an *oracle* for a family of hard Boolean functions, and our derandomized algorithms will have access to this oracle. We thus work in a theory with an extra unary function symbol α:

4.1 Definition Let $0 < \varepsilon < 1$ and c be standard constants. The theory $HARD^{\varnothing}_{\varepsilon,c}$ is an extension of $S^1_2(\alpha)$ by the following axioms:

$$\alpha(x)\colon 2^{||x||} \to 2,$$
$$x > c \to Hard^{\varnothing}_{\varepsilon}(\alpha(x)).$$

The theory $HARD_{\varepsilon,c}$ is defined similarly. We will usually ignore c in the sequel. (To avoid confusion: here $||x||$ means double iteration of the length function, it has nothing to do with the translation of Π^b_1-formulas into propositional logic from Section 2. We will not use this translation any more.)

4.2 Observation $HARD^{\varnothing}_{\varepsilon}$ implies $HARD_{\varepsilon}$, and $HARD_{\varepsilon}$ proves $dWPHP(PV)$. $\square$

First, notice that we get a certain derandomization for free, namely for definable *MFRP* which are *provably total* in $S^1_2 + dWPHP(PV)$:

4.3 Lemma *Let F be a definable MFRP, provably total in $S^1_2+dWPHP(PV)$, and let $\varepsilon > 0$. Then there is a $PV(\alpha)$ function f such that*

$$HARD_\varepsilon \vdash f(\vec{x}) = y \to F(\vec{x}) = y.$$

Proof: By our assumptions $HARD_\varepsilon$ proves $\forall\vec{x}\,\exists y\, F(\vec{x}) = y$, which is $\forall\Sigma^b_1$. Moreover, $HARD_\varepsilon$ is a $\forall\Pi^b_1(\alpha)$ extension of $S^1_2(\alpha)$, hence the result follows from the relativized Buss's witnessing theorem. □

However, we want to derandomize also functions which are not provably total (e.g., *RP*-predicates). Moreover, the Nisan-Wigderson construction will give a stronger result (see 4.9): f needs only *one* oracle query.

4.4 Definition ([9]) Let $k, \ell, t, m \in Log$, $k \le \ell \le t$. A $\langle k, \ell, t, m\rangle$-*design* is a sequence $\langle S_i\rangle_{i<m}$ of subsets $S_i \subseteq t$, such that $|S_i| = \ell$ and $|S_i \cap S_j| \le k$ for all $i < j < m$.

4.5 Lemma *Let $0 < \gamma < 1$. There are constants $\delta > 0$, $c > 1$, and a PV-function d such that*

$$PV \vdash d(x) \textit{ is a } \langle\gamma\ell, \ell, c\ell, 2^{\delta\ell}\rangle\textit{-design, where } \ell = ||x||.$$

Proof: Put $c = 2/\gamma$, $\delta = c^{-2}$, and let $k = \gamma\ell$, $t = c\ell$, and $m = 2^{\delta\ell}$. The function d will iterate through all subsets $S \subseteq t$, putting S into the design if $|S| = \ell$ and its intersection with all elements of the design so-far constructed is at most k. We have to show that this algorithm will not stop with a design shorter than m. Clearly, it suffices to prove that for any design $\langle S_j\rangle_{j<i}$, $i < m$, there is an $S_i \subseteq t$ such that $\langle S_j\rangle_{j\le i}$ is also a design. We will do this by a counting argument (which works directly without any *PHP*, as $m \in Log$ and $k, \ell, t \in LogLog$). However, it turns out that instead of counting subsets $S \subseteq t$, it is easier to count functions $f\colon t \to t$ which represent $S = S(f) := f^{-1}{}''\ell$ (thus, choosing uniformly a random f means to choose S in such a way that $\Pr(a \in S) = \ell/t$ for all $a < t$).

The number of $f\colon t \to t$ such that $|S(f)| \ge \ell$ is

$$\sum_{i\ge\ell} \binom{t}{i} \ell^i (t-\ell)^{t-i} \ge \varepsilon t^t$$

for some constant $\varepsilon > 0$, by A.4. If S is a subset of t of size ℓ, the number of f such that $|S(f) \cap S| \ge k$ is

$$\sum_{i=k}^{\ell} \binom{\ell}{j} \ell^j (t-\ell)^{\ell-j} t^{t-\ell} = t^t \ell^{-\ell} \sum_{i=k}^{\ell} \binom{\ell}{j} (k/2)^j (\ell - k/2)^{\ell-j} \le t^t 4^{-(k/2)^2/\ell} = t^t 2^{-\gamma^2\ell/2}$$

by Chernoff's inequality (A.5). The number of f such that $|S(f) \cap S_j| \ge k$ for some $j < i$ is thus at most

$$t^t m 2^{-\gamma^2\ell/2} = t^t 2^{(\delta-\gamma^2/2)\ell} \le t^t 2^{-\gamma^2\ell/4} < \varepsilon t^t,$$

hence there is f such that we may put $S(f)$ into the design. (If $|S(f)| > \ell$, we discard some of its elements.) □

4.6 Definition ([9]) Let $x < 2^t$, and $S \subseteq t$, $|S| = \ell$. Let $\{s_i\}_{i<\ell}$ be the increasing enumeration of the set S. Then we put $x \restriction S := y$, where $y < 2^\ell$ and $bit(y,i) = bit(x,s_i)$ for all $i < \ell$.

If $f\colon 2^\ell \to 2$, and $S = \langle S_i\rangle_{i<m}$ is a $\langle k,\ell,t,m\rangle$-design, the *Nisan-Wigderson generator* is a function $NW_{f,S}\colon 2^t \to 2^m$ defined by

$$bit(NW_{f,S}(x), i) = f(x \restriction S_i).$$

Let NW be a PV-function such that $NW(f,S,x) = NW_{f,S}(x)$.

4.7 Proposition *There is a PV-function $\pi(f,S,D,a,z)$, such that S^1_2 proves the following property:*

Let $f\colon 2^\ell \to 2$ be a Boolean function such that $|\{x < 2^\ell;\, C(x) = f(x)\}| \le 2^{\ell-1} + a$ for any circuit C of size $|C| \le s$. Let S be a $\langle k,\ell,t,m\rangle$-design, and let $D\colon 2^m \to 2$ be a circuit of size $|D| < s - m2^k$. Put $e = am2^{m+t-\ell}$. Then

$$\pi(f,S,D,a,\cdot)\colon e \,\dot\cup\, (2^m \times \{x < 2^t;\, D(NW_{f,S}(x)) = 1\}) \twoheadrightarrow 2^t \times \{r < 2^m;\, D(r) = 1\}.$$

4.8 Remark The function π witnesses that $\Pr_x(D(NW_{f,S}(x)) = 1) \ge \Pr_r(D(r) = 1) - m\varepsilon$, where $\varepsilon = a2^{-\ell}$.

Proof: We will find (uniformly in $i < m$) surjections

$$G_i\colon a2^{m+t-\ell} \,\dot\cup\, M_{i+1} \twoheadrightarrow M_i,$$

where $M_i = \{\langle\vec r, x\rangle;\, D(f(x \restriction S_0), \dots, f(x \restriction S_{i-1}), r_i, \dots, r_{m-1}) = 1\}$. Notice that $M_0 = \{\vec r;\, D(\vec r) = 1\} \times 2^t$, and $M_m = 2^m \times \{x;\, D(NW_{f,S}(x)) = 1\}$.

Fix $i < m$, $y < 2^{t-\ell}$, and $r_{i+1}, \dots, r_{m-1} < 2$. For any $u < 2^\ell$ and $j < m$ define $f^y_j(u) = f(x \restriction S_j)$, where $x \restriction S_i = u$ and $x \restriction (t \smallsetminus S_i) = y$. Finally put

$$\begin{aligned} A_0(u) &= D(f^y_0(u), \dots, f^y_{i-1}(u), 0, r_{i+1}, \dots, r_{m-1}),\\ A_1(u) &= \neg D(f^y_0(u), \dots, f^y_{i-1}(u), 1, r_{i+1}, \dots, r_{m-1}). \end{aligned}$$

Each $f^y_j(u)$, $j < i$, depends only on $|S_j \cap S_i| \le k$ variables, hence it is computable by a circuit of size 2^k. This allows A_0 and A_1 to be represented as circuits of size at most $1 + |D| + i2^k \le 1 + |D| + m2^k \le s$, hence

$$|\{u;\, A_r(u) = f(u)\}| \le 2^{\ell-1} + a, \qquad r = 0, 1.$$

By summing these two inequalities we get

$$
\begin{aligned}
2a &\geq |\{u;\ f(u)=A_0(u)\}| + |\{u;\ f(u)=A_1(u)\}| - 2^\ell = \\
&= |\{u;\ (A_0(u) \mathbin{\&} \neg(\neg f(u) \mathbin{\&} A_0(u))) \vee (\neg f(u) \mathbin{\&} \neg(\neg f(u) \mathbin{\&} A_0(u)))\}| + \\
&\qquad + |\{u;\ (\neg A_1(u) \mathbin{\&} \neg(f(u) \mathbin{\&} \neg A_1(u))) \vee (f(u) \mathbin{\&} \neg(f(u) \mathbin{\&} \neg A_1(u)))\}| - 2^\ell = \\
&= |\{u;\ A_0(u)\}| - |\{u;\ \neg f(u) \mathbin{\&} A_0(u)\}| + |\{u;\ \neg f(u)\}| - |\{u;\ \neg f(u) \mathbin{\&} A_0(u)\}| + \\
&\qquad + |\{u;\ \neg A_1(u)\}| - |\{u;\ f(u) \mathbin{\&} \neg A_1(u)\}| + |\{u;\ f(u)\}| - |\{u;\ f(u) \mathbin{\&} \neg A_1(u)\}| - 2^\ell = \\
&= |\{u;\ A_0(u)\}| + |\{u;\ \neg A_1(u)\}| - 2|\{u;\ (\neg f(u) \mathbin{\&} A_0(u)) \vee (f(u) \mathbin{\&} \neg A_1(u))\}| = \\
&= |\{u;\ D(f_0^y(u), \dots, f_{i-1}^y(u), 0, r_{i+1}, \dots, r_{m-1})\}| + |\{u;\ D(f_0^y(u), \dots, 1, r_{i+1}, \dots)\}| - \\
&\qquad - 2|\{u;\ D(f_0^y(u), \dots, f(u), r_{i+1}, \dots)\}| = \\
&= |\{\langle r, u\rangle;\ D(f_0^y(u), \dots, r, r_{i+1}, \dots)\}| - |\{\langle r, u\rangle;\ D(f_0^y(u), \dots, f(u), r_{i+1}, \dots)\}|.
\end{aligned}
$$

Employing counting functions for the two sets in the last line, we get a surjection

$$
\begin{aligned}
g_{i,y,r_{i+1},\dots,r_{m-1}} \colon 2a \mathbin{\dot\cup} \{\langle r, u\rangle;\ D(f_0^y(u), \dots, f(u), r_{i+1}, \dots, r_{m-1})\} \twoheadrightarrow \\
\{\langle r, u\rangle;\ D(f_0^y(u), \dots, r, r_{i+1}, \dots, r_{m-1})\}.
\end{aligned}
$$

Define $G_i \colon M_{i+1} \mathbin{\dot\cup} a2^{m+t-\ell} \to M_i$ by

$$
\begin{aligned}
G_i(\vec{r}, x) &= \langle r_0, \dots, r_{i-1}, r_i', r_{i+1}, \dots, r_{m-1}, x'\rangle, \\
&\qquad \text{if } g_{i,y,r_{i+1},\dots,r_{m-1}}(r_i, x \restriction S_i) = \langle r_i', x' \restriction S_i\rangle, \text{ and } x' \restriction (t \smallsetminus S_i) = y, \\
G_i(2av + w) &= \langle r_0, \dots, r_{i-1}, r_i', r_{i+1}, \dots, r_{m-1}, x'\rangle, \\
&\qquad \text{if } v = \langle y, r_0, \dots, r_{i-1}, r_{i+1}, \dots, r_{m-1}\rangle, \\
&\qquad\qquad g_{i,y,r_{i+1},\dots,r_{m-1}}(w) = \langle r_i', x' \restriction S_i\rangle, \text{ and } x' \restriction (t \smallsetminus S_i) = y,
\end{aligned}
$$

It is straightforward to check that the functions G_i are well defined and onto, using $f(x \restriction S_j) = f_j^{x \restriction (t \smallsetminus S_i)}(x \restriction S_i)$.

Now we define π as a composition of $G_0, \dots, G_{m-1}$. More precisely, we put

$$\pi(f, S, D, a, z) = G^m(z),$$

where $G^i \colon M_i \mathbin{\dot\cup} ai2^{m+t-\ell} \to M_0$ is defined inductively by

$$
\begin{aligned}
G^0(z) &= z, \\
G^{i+1}(z) &= \begin{cases} w - a2^{m+t-\ell}, & a2^{m+t-\ell} \leq z < a(i+1)2^{m+t-\ell}, \\ G^i(G_i(z)), & \text{otherwise.} \end{cases}
\end{aligned}
$$

Given $z \in M_0$, we prove by Σ_1^b-*LIND* on $i \leq m$ that there is a $w \in M_i \mathbin{\dot\cup} ai2^{m+t-\ell}$ such that $G^i(w) = z$, in particular $\pi \colon M_m \mathbin{\dot\cup} am2^{m+t-\ell} \to M_0$ is onto, as required. □

4.9 Proposition *Let F be a MFRP definable in $S_2^1 + dWPHP(PV)$, and let $\varepsilon > 0$. Then there are PV-functions h and g such that $HARD_\varepsilon^\varnothing$ proves*

$$
\begin{aligned}
&\exists y\ \ y = F(x) \ \leftrightarrow\ h(x, \alpha(g(x))) \neq *, \\
&\exists y\ \ y = F(x) \to h(x, \alpha(g(x))) = F(x).
\end{aligned}
$$

Proof: Fix a 1/2-definition of $F(x)$ given by $f(x,w)$, $w < r(x)$. We may assume w.l.o.g. that $r(x) \geq x$. Choose a constant $b \geq 1$ such that for all $n \gg 0$, there is a circuit $C\colon 2^n \times 2^m \to 2$ of size at most m^b such that

$$C(x,w) = 1 \quad \text{iff} \quad f(x,w) \neq *,$$

where $m = |r(x)|$. Choose $\gamma < \varepsilon$, and let c, δ, and d be as in Lemma 4.5. We may assume $\gamma + \delta < \varepsilon$ and $\delta < \varepsilon/b$, because we may shorten the design produced by d if necessary.

Define $g(x) = 2^{m^{1/\delta}}$, so that $\varphi = \alpha(g(x))$ is a Boolean function on $\ell = |m|/\delta$ variables. Put $t = c\ell$, $k = \gamma\ell$, and let $S = d(g(x))$ (hence S is a $\langle k, \ell, t, m\rangle$-design). Finally, define

$$h(x,\varphi) = \begin{cases} f(x, NW_{\varphi,S}(u)), & \text{if } u \text{ is the smallest } u < 2^t \text{ such that } f(x, NW_{\varphi,S}(u)) \neq *, \\ *, & \text{if no such } u \text{ exists.} \end{cases}$$

Notice that $2^t = m^{c/\delta} = n^{O(1)}$, so the loop over all $u < 2^t$ may be done by a PV-function (i.e., it is p-time computable).

Clearly $h(x, \alpha(g(x))) = *$ if $F(x)$ does not have a value, and y is a value of $F(x)$ if $y = h(x, \alpha(g(x))) \neq *$. It remains to show that $h(x, \alpha(g(x))) \neq *$ if $F(x)$ is defined.

Put $s = 2^{\varepsilon\ell}$ and $a = 2^{(1-\varepsilon)\ell}$, so that φ satisfies the assumptions of Proposition 4.7. The size bound on $D(w) := C(x,w)$ is also satisfied: $|D| + m2^k \leq m^b + m^{1+\gamma/\delta} < m^{\varepsilon/\delta} = s$, because $1 + \gamma/\delta < \varepsilon/\delta$ and $b < \varepsilon/\delta$.

Assume that we do not find a suitable $u < 2^t$. This means that $D(NW_{\varphi,S}(u)) = 0$ for all $u < 2^t$, hence by Proposition 4.7 the function $\pi(\varphi, S, D, a, \cdot)$ is a surjection from $e = am2^{m+t-\ell}$ to $2^t \times \{w;\, D(w) = 1\}$. On the other hand, f is a 1/2-definition of F and we assume that $F(x)$ is defined, hence we also have a surjection of 2^m onto $2 \times \{w;\, D(w) = 0\}$. We may modify this function to map 2^{m+t-1} onto $2^t \times \{w;\, D(w) = 0\}$, and combine it with π to get a surjection from $2^{m+t-1} + e$ onto 2^{m+t}.

However, $e = 2^{m+t+(\delta-\varepsilon)\ell} < 2^{m+t-2}$ because $\delta < \varepsilon$, hence we obtain a mapping of $3 \cdot 2^{m+t-2}$ onto $4 \cdot 2^{m+t-2}$. This contradicts $dWPHP(PV)$, which is available in $HARD^{\varnothing}_{\varepsilon}$. □

A subtle point arises here: we have shown derandomization in $HARD^{\varnothing}_{\varepsilon}$, but we do not know (yet) the strength of this theory as compared to unrelativized bounded arithmetical theories, in particular $S^1_2 + dWPHP(PV)$. In fact, the Nisan-Wigderson theorem is true, hence its formalized version 4.9 *trivially* holds in some theory similar to $HARD^{\varnothing}_{\varepsilon}$. To see that no such cheating is involved here, we will show that $HARD^{\varnothing}_{\varepsilon}$ is a conservative extension of $S^1_2 + dWPHP(PV)$.

4.10 Lemma *Let $\varepsilon < 1$. There is a constant c such that $PV + dWPHP(PV)$ proves*

(*i*) $\forall k \in LogLog\, \exists w\, \forall i < k\, (i \geq c \to (w)_i \colon 2^i \to 2 \,\&\, Hard_{\varepsilon}((w)_i))$,

(*ii*) $\forall k \in LogLog\, \exists w\, \forall i < k\, (i \geq c \to (w)_i \colon 2^i \to 2 \,\&\, Hard^{\varnothing}_{\varepsilon/3}((w)_i))$.

Proof: This is a refinement of 3.2 and 3.4. (Notice that we cannot use these lemmas directly, as $BB\Pi^b_1$ is not available. The conclusion is $\forall\Sigma^b_2$, but it is not *a priori* clear that the

Σ^b_2-conservativity of $BB\Sigma^b_2$ over S^1_2 extends to $S^1_2 + dWPHP(PV)$, although see 4.12.) As in Lemma 3.4, choose d such that

$$\sum_{j=0}^{2^i(1/2-2^{-\delta i})} \binom{2^i}{j} \leq 2^{d+2^i-2\cdot 2^{i(1-2\delta)}}$$

for all $i > 0$, where $\delta = \varepsilon/3$, and choose $c \geq 2$ such that $2 \cdot 2^{i(1-2\delta)} \geq \frac{2}{3} i 2^{\delta i} + d + i$ for all $i \geq c$. Put $k = ||b||$, and define a PV-function

$$g\colon \sum_{i=c}^{k-1} 2^{2^k-2^c-i} \to 2^{2^k-2^c}$$

as follows: given $i < k$ and $x < 2^{2^k-2^c-i}$, interpret the first $2^k - 2^c - 2^i$ bits of x as a sequence $\langle f_j;\, c \leq j < k, j \neq i\rangle$ of functions $f_j\colon 2^j \to 2$. The next $2\delta i 2^{\delta i}$ bits of x describe a circuit $C\colon 2^i \to 2$ of size $2^{\delta i}$, and the rest of x defines a binary string y of length 2^i with at most $2^{i-1} - 2^{i(1-\delta)}$ ones. (We need $d + 2^i - 2 \cdot 2^{i(1-2\delta)}$ bits for y, and we have $2^i - 2\delta i 2^{\delta i} - i \geq 2^i - 2 \cdot 2^{i(1-2\delta)} + d$ bits left.) We create a function $f_i\colon 2^i \to 2$ by taking the truth-table of C XOR'ed by y, and we let g output the sequence $\langle f_j;\, c \leq j < k\rangle$.

If $f = \langle f_j;\, c \leq j < k\rangle$ is a sequence of functions outside of the range of g, then all f_j are δ-hard on average. The domain of g is at most $2^{2^k-2^c-c+1} \leq 2^{2^k-2^c-1}$, hence g is not onto by $dWPHP(PV)$. A similar argument works for ε-hard functions. $\square$

4.11 Proposition *Let T denote the theory $HARD_\varepsilon$ or $HARD^{\varnothing}_{\varepsilon/3}$, with $0 < \varepsilon < 1$. Then T is fully conservative over $S^1_2 + dWPHP(PV)$. More generally, for any $i \geq 1$, $T + S^i_2(\alpha)$ and $T + T^i_2(\alpha)$ are conservative extensions of $S^i_2 + dWPHP(PV)$ and $T^i_2 + dWPHP(PV)$, respectively. Every countable model of $S^1_2 + dWPHP(PV)$ has an expansion into a model of T.*

Proof: Let $\mathcal{A}$ be a countable model of $S^1_2 + dWPHP(PV)$. Choose an increasing chain $p^0 \subseteq p^1 \subseteq p^2 \subseteq \dots$ of sequences $p^n \in A$ such that

$$\forall i < lh(p^n)\, (i \geq c \to (p^n)_i\colon 2^i \to 2 \;\&\; Hard_\varepsilon((p^n)_i)),$$

where c is the constant from Lemma 4.10, and such that $\{lh(p^n);\, n \in \omega\}$ is cofinal in $LogLog(A)$. Define $\alpha^A = \bigcup_{n\in\omega} p^n$, i.e.

$$\alpha^A(a) := (p^n)_{||a||}, \qquad \text{for any } n \text{ s.t. } lh(p^n) > ||a||.$$

Clearly, $\langle \mathcal{A}, \alpha^A\rangle$ satisfies the hardness conditions from T.

Claim 1 *Let $\varphi(\vec{x})$ be a $\Sigma^b_\infty(\alpha)$-formula. Denote by $\tilde{\varphi}(p, \vec{x})$ the Σ^b_∞-formula which results from φ by substitution of $(p)_{||t||}$ for every subterm $\alpha(t)$. There is a constant c_φ such that*

$$\langle \mathcal{A}, \alpha^A\rangle \vDash \varphi(\vec{a}) \quad \textit{iff} \quad \mathcal{A} \vDash \tilde{\varphi}(p^n, \vec{a})$$

for any n such that $lh(p^n) > c_\varphi ||\vec{a}||$.

Proof: By straightforward induction on complexity of φ. If φ is atomic, it suffices to choose c_φ so that all $(p^n)_{||t||}$ are defined. If e.g. $\varphi(\vec{x}) = \exists y \leq s(\vec{x})\, \psi(y, \vec{x})$, take $c_\varphi = (d+1)c_\psi$, where d is such that $||s(\vec{x})|| < d||\vec{x}||$ for all $\vec{x}$. The assertion then follows from the induction hypothesis, because $y \leq s(\vec{x})$ and $lh(p^n) > c_\varphi ||\vec{x}||$ imply $lh(p^n) > c_\psi(||y|| + ||\vec{x}||)$. □ (Claim 1)

As a corollary of the Claim we get that $\langle \mathcal{A}, \alpha^A \rangle \models \forall \vec{x}\, \varphi(\vec{x})$, whenever φ is a bounded $L(\alpha)$-formula, and $\mathcal{A} \models \forall \vec{x}\, \forall p\, \tilde{\varphi}(p, \vec{x})$. In particular, $\langle \mathcal{A}, \alpha^A \rangle \models S^1_2(\alpha)$, and additionally it is a model of $S^i_2(\alpha)$ or $T^i_2(\alpha)$, if S^i_2 or T^i_2 holds in $\mathcal{A}$. □

4.12 Corollary *$S^1_2 + dWPHP(PV) + BB\Sigma^b_2$ is $\forall\Sigma^b_2$-conservative over $S^1_2 + dWPHP(PV)$.*

Proof: This follows from 4.11, and $\Sigma^b_2(\alpha)$-conservation of $BB\Sigma^b_2(\alpha)$ over $S^1_2(\alpha)$ [12], because $HARD_{1/2}$ is a $\forall\Pi^b_1(\alpha)$-axiomatized extension of $S^1_2(\alpha)$. □

5 Acknowledgements

The author wishes to thank Jan Krajíček for his guidance, clarifying discussions on the topic, and many suggestions and improvements. I would like also to thank Pavel Pudlák and Neil Thapen, for reading and commenting a preliminary version of this paper.

References

[1] S. Buss: *Bounded Arithmetic,* Bibliopolis, Naples 1986.

[2] S. Buss, R. Impagliazzo, J. Krajíček, P. Pudlák, A. Razborov, J. Sgall: Proof complexity in algebraic systems and bounded depth Frege systems with modular counting, *Computational Complexity* 6 (1996/97), pp. 256–298.

[3] S. Cook: Feasibly constructive proofs and the propositional calculus, in: *Proceedings of the 3rd Annual ACM Symposium on Theory of Computing,* ACM Press, 1975, pp. 83–97.

[4] J. Krajíček: *Bounded Arithmetic, Propositional Logic, and Complexity Theory,* Cambridge University Press, 1995.

[5] J. Krajíček: On the weak pigeonhole principle, *Fundamenta Mathematicae* 170 (2001), pp. 123–140.

[6] J. Krajíček, P. Pudlák: Quantified propositional calculi and fragments of bounded arithmetic, *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* 36 (1990), pp. 29–46.

[7] J. Krajíček, P. Pudlák: Some Consequences of Cryptographical Conjectures for S^1_2 and EF, *Information and Computation* 140 (1998), pp. 82–94.

[8] A. Maciel, T. Pitassi, A. Woods: A New Proof of the Weak Pigeonhole Principle, *Journal of Computer and System Sciences* 64 (2002), pp. 843–872.

[9] N. Nisan, A. Wigderson: Hardness vs. Randomness, *Journal of Computer and System Sciences* 49 (1994), pp. 149–167.

[10] R. Parikh: Existence and feasibility in arithmetic, *Journal of Symbolic Logic* 36 (1971), pp. 494–508.

[11] J. Paris, A. Wilkie, A. Woods: Provability of the pigeonhole principle and the existence of infinitely many primes, *Journal of Symbolic Logic* 53 (1988), pp. 1235–1244.

[12] J.-P. Ressayre: A conservation result for system of bounded arithmetic, unpublished manuscript, 1986.

[13] N. Thapen: *The Weak Pigeonhole Principle in Models of Bounded Arithmetic,* PhD thesis, Oxford University, 2002.

A Some bounds on binomial coefficients

Here we show that several well-known inequalities, useful for counting, are actually provable in PV, when their parameters are restricted to logarithmically small numbers. We could not get around proving these technical results, even though they are not unexpected.

A.1 Definition Let $n \in Log$, $k, i \leq n$. Define

$$\begin{bmatrix} n \\ i \end{bmatrix}_k := \binom{n}{i} k^i (n-k)^{n-i},$$

$\binom{n}{<i} := \sum_{j<i} \binom{n}{j}$, $\begin{bmatrix} n \\ <i \end{bmatrix}_k := \sum_{j<i} \begin{bmatrix} n \\ j \end{bmatrix}_k$.

A.2 Proposition (Stirling's bound) *There is a $c > 1$ such that PV proves*

$$0 < k < n \in Log \;\rightarrow\; \frac{1}{c}\binom{n}{k} \leq \frac{n^n}{k^k (n-k)^{n-k}} \sqrt{\frac{n}{k(n-k)}} \leq c\binom{n}{k}.$$

(We will abbreviate this as "$\binom{n}{k} = \Theta(\cdots)$.")

Proof: Define $f(i) := \begin{bmatrix} n \\ i \end{bmatrix}_k$, and $\gamma(i) := f(i+1)/f(i) = k(n-i)/((n-k)(i+1))$. We have $j < i \rightarrow \gamma(j) > \gamma(i)$, because $(j+1)(n-i) < (i+1)(n-j)$. Also

$$\gamma(k-1) = (n-k+1)/(n-k) > 1 > k/(k+1) = \gamma(k),$$

hence $f(i+1) > f(i)$ for $i < k$, and $f(i+1) < f(i)$ for $i \geq k$.

Let $i < k$. We have $k(n-i+1)f(j-1) \leq (n-k)i\, f(j)$ for any $0 < j \leq i$, hence

$$(k(n-i+1))^\ell f(i-\ell) \leq ((n-k)i)^\ell f(i)$$

for any $0 \leq \ell \leq i$ (by Δ^b_1-induction on ℓ). Using induction once again, we find that

$$(k(n-i+1))^\ell (k(n-i+1) - (n-k)i) \sum_{j=i-\ell}^{i-1} f(j) \leq$$
$$\leq (n-k)i((k(n-i+1))^\ell - ((n-k)i)^\ell) f(i),$$

in particular,

$$(k(n-i+1))^i(kn+k-in)\sum_{j<i} f(j) \le (n-k)i((k(n-i+1))^i - ((n-k)i)^i)f(i) \le$$
$$\le (n-k)i(k(n-i+1))^i f(i),$$

hence

$$\sum_{j<i} f(j) \le \frac{(n-k)i}{kn+k-in} f(i) \le \frac{(n-k)i}{n(k-i)} f(i).$$

Similarly,

$$\sum_{j>i} f(j) \le \frac{k(n-i)}{n(i-k)} f(i) \qquad \text{for any } i>k.$$

Put $s := \left\lfloor \sqrt{\frac{k(n-k)}{n}} \right\rfloor$. Then

$$\frac{1}{f(k-s-1)} \sum_{j<k-s-1} f(j) \le \frac{n-k}{n}\left(\frac{k}{s+1}-1\right) \le \frac{n-k}{n}\left(\sqrt{\frac{kn}{n-k}}-1\right) =$$
$$= \sqrt{\frac{k(n-k)}{n}} - \frac{n-k}{n} \le s + \frac{k}{n},$$

hence

$$\sum_{j\le k} f(j) \le \sum_{j<k-s-1} f(j) + \sum_{j=k-s-1}^{k} f(j) \le \left(s+\frac{k}{n}\right) f(k-s-1) + (s+2)f(k) \le$$
$$\le \left(2s+2+\frac{k}{n}\right) f(k).$$

Similarly we may show

$$\sum_{j>k} f(j) \le \left(2s+2-\frac{k}{n}\right) f(k),$$

hence

$$n^n = \sum_{j=0}^{n} f(j) \le 4(s+1)f(k) \le 8sf(k),$$

in other words

$$\binom{n}{k} \ge \frac{n^n}{8sk^k(n-k)^{n-k}} \ge \frac{n^n}{8k^k(n-k)^{n-k}}\sqrt{\frac{n}{k(n-k)}}.$$

Claim 1 *PV proves*

$$b \in Log,\ \ b>0 \quad \rightarrow \quad (b+1)^{b+1} \le 4b^{b+1}.$$

Proof: By induction on b. The claim holds if $b = 1$. Assume $b > 1$ and $b^b \leq 4(b-1)^b$. Straightforward induction on d shows that

$$c^d \leq (c+1)^d - d(c+1)^{d-1} + \binom{d}{2}(c+1)^{d-2}, \qquad d \geq 2,$$

hence

$$\begin{aligned}
(b-1)^{b+1}(b+1)^{b+1} &= (b^2-1)^{b+1} \leq \\
&\leq b^{2b+2} - (b+1)b^{2b} + \frac{b^2+b}{2}b^{2b-2} = \\
&= b^{2b+2} - b^{2b-1} - \frac{1}{2}b^{2b} + \frac{1}{2}b^{2b-1} \leq \\
&\leq b^{2b+2} - b^{2b+1} = (b-1)b^{b+1}b^b \leq \\
&\leq 4(b-1)^{b+1}b^{b+1},
\end{aligned}$$

thus $(b+1)^{b+1} \leq 4b^{b+1}$. □ (Claim 1)

Claim 2 *PV proves*

$$a, b \in Log,\ \ b > 0 \quad \rightarrow \quad (b+a)^b \leq 4^a b^b.$$

Proof: The case $a = 0$ is trivial. If $a = 1$, we have $(b+1)^{b+1} \leq 4b^{b+1} \leq 4(b+1)b^b$ by previous Claim, hence $(b+1)^b \leq 4b^b$. We proceed by induction on a. Using the induction hypothesis for a and 1, we have

$$(b+a+1)^{b+a} \leq 4(b+a)^{b+a} \leq 4^{a+1}b^b(b+a)^a \leq 4^{a+1}b^b(b+a+1)^a,$$

hence $(b+a+1)^b \leq 4^{a+1}b^b$. □ (Claim 2)

Let $i \leq s$. Then $in(i-1) \leq k(n-k)$, hence

$$\gamma(k-i) = \frac{k(n-k+i)}{(n-k)(k-i+1)} = 1 + \frac{k+n(i-1)}{(n-k)(k-i+1)} \leq 1 + \frac{in}{k(n-k)}.$$

Since (assuming i even) $f(k-i/2) \leq f(k-i)\gamma^{i/2}(k-i)$, this implies

$$(k(n-k))^{i/2}f(k-i/2) \leq (k(n-k)+in)^{i/2}f(k-i),$$

and, using Claim 2,

$$\begin{aligned}
&(k(n-k))^{\frac{i}{2}k(n-k)}(f(k-i/2))^{k(n-k)} \leq (k(n-k)+in)^{\frac{i}{2}k(n-k)}(f(k-i))^{k(n-k)} \leq \\
&\qquad \leq (4^{in}(k(n-k))^{k(n-k)})^{i/2}(f(k-i))^{k(n-k)} = (k(n-k))^{\frac{i}{2}k(n-k)}2^{i^2n}(f(k-i))^{k(n-k)},
\end{aligned}$$

hence

$$(f(k-i/2))^{k(n-k)} \leq 2^{i^2n}(f(k-i))^{k(n-k)}.$$

Choose ℓ such that $2^\ell \leq s < 2^{\ell+1}$. Then $4^\ell \leq k(n-k)/n$, and an induction shows that

$$\begin{aligned}
&(f(k-1))^{k(n-k)} \leq (f(k-2^\ell))^{k(n-k)}2^{\frac{4}{3}(4^\ell-1)n} \leq \\
&\qquad\qquad \leq (f(k-2^\ell))^{k(n-k)}2^{\frac{4}{3}k(n-k)} \leq (3f(k-2^\ell))^{k(n-k)},
\end{aligned}$$

hence $f(k-2^\ell) \geq f(k-1)/3 \geq f(k)/6$. This implies

$$n^n \geq \sum_{j=1}^{2^\ell} f(k-j) \geq 2^\ell f(k-2^\ell) \geq \frac{2^\ell}{6} f(k) \geq \frac{s+1}{12} f(k),$$

which means

$$\binom{n}{k} \leq \frac{12n^n}{(s+1)k^k(n-k)^{n-k}} \leq \frac{12n^n}{k^k(n-k)^{n-k}}\sqrt{\frac{n}{k(n-k)}}.$$

□

A.3 Corollary *PV proves: for any $0 < k < n \in Log$,*

$$|k-i| \leq \sqrt{\frac{k(n-k)}{n}} \quad \to \quad \begin{bmatrix} n \\ i \end{bmatrix}_k = \Theta\left(\begin{bmatrix} n \\ k \end{bmatrix}_k\right).$$

(Here $|\cdot|$ denotes absolute value, not the length function.) □

A.4 Proposition *The following is provable in PV. Let $k, n \in Log$ be such that $n > k > 0$, and denote $s = \sqrt{\frac{k(n-k)}{n}}$.*

(i) Assume $i \leq s$. Then

$$\begin{bmatrix} n \\ < k-i \end{bmatrix}_k = \Theta\left(s\begin{bmatrix} n \\ k-i \end{bmatrix}_k\right) = \Theta(n^n),$$

$$\begin{bmatrix} n \\ > k+i \end{bmatrix}_k = \Theta\left(s\begin{bmatrix} n \\ k+i \end{bmatrix}_k\right) = \Theta(n^n).$$

(ii) Assume $i \geq s$.

$$\begin{bmatrix} n \\ < k-i \end{bmatrix}_k = \Theta\left(\left(1-\frac{k}{n}\right)\left(\frac{k}{i}-1\right)\begin{bmatrix} n \\ k-i \end{bmatrix}_k\right),$$

$$\begin{bmatrix} n \\ > k+i \end{bmatrix}_k = \Theta\left(\frac{k}{n}\left(\frac{n-k}{i}-1\right)\begin{bmatrix} n \\ k+i \end{bmatrix}_k\right).$$

Proof: It suffices to show the $\begin{bmatrix} n \\ <\cdots \end{bmatrix}$-part, as $\begin{bmatrix} n \\ j \end{bmatrix}_k = \begin{bmatrix} n \\ n-j \end{bmatrix}_{n-k}$.

First assume $i \leq s$. We already know from the proof of A.2 that

$$\begin{bmatrix} n \\ < k-i \end{bmatrix}_k \leq \begin{bmatrix} n \\ < k \end{bmatrix}_k = O\left(s\begin{bmatrix} n \\ k \end{bmatrix}_k\right) = O\left(s\begin{bmatrix} n \\ k-i \end{bmatrix}_k\right) = O(n^n).$$

If $i \leq s/2$, we also have

$$\begin{bmatrix} n \\ < k-i \end{bmatrix}_k \geq \begin{bmatrix} n \\ < k-s/2 \end{bmatrix}_k \geq \frac{s}{2}\begin{bmatrix} n \\ k-s \end{bmatrix}_k = \Omega\left(s\begin{bmatrix} n \\ k-i \end{bmatrix}_k\right) = \Omega(n^n).$$

The case of $s/2 < i \leq s$ is treated similarly: the proof of $\begin{bmatrix} n \\ k-s \end{bmatrix}_k = \Omega\left(\begin{bmatrix} n \\ k \end{bmatrix}_k\right)$ can be easily adapted to $\begin{bmatrix} n \\ k-2s \end{bmatrix}_k$.

Now assume $k \geq i > s$. We have already proved that

$$\begin{bmatrix} n \\ < k-i \end{bmatrix}_k \leq \frac{(n-k)(k-i)}{ni} \begin{bmatrix} n \\ k-i \end{bmatrix}_k.$$

If $i = k$, clearly $\left[{n \atop <k-i}\right]_k = 0 = \frac{k}{i} - 1$. If $k > i \geq k/4$, we have

$$\frac{\left[{n \atop <k-i}\right]_k}{\left[{n \atop k-i}\right]_k} \geq \frac{1}{\gamma(k-i-1)} = \frac{(n-k)(k-i)}{k(n-k+i+1)} \geq \frac{(n-k)(k-i)}{kn} \geq \frac{(n-k)(k-i)}{4ni}.$$

Let $k/4 > i > s$, and define f and γ as in the proof of A.2. By the monotonicity of γ and simple induction, we have

$$f(k-i-j) \geq f(k-i)(\gamma(k-2i))^{-j},$$

hence (putting $\gamma = \gamma(k-2i)$)

$$\begin{bmatrix} n \\ < k-i \end{bmatrix}_k \geq \sum_{j=1}^{i} f(k-i-j) \geq f(k-i) \sum_{j=1}^{i} \gamma^{-j} = f(k-i)\frac{1}{\gamma-1}(1-\gamma^{-i}) =$$

$$= f(k-i)\frac{(n-k)(k-2i+1)}{n(2i-1)+k}\left(1-\left(\frac{(n-k)(k-2i+1)}{k(n-k+2i)}\right)^i\right).$$

Notice that

$$\frac{(n-k)(k-2i+1)}{n(2i-1)+k} \geq \frac{(n-k)(k-2i+1)}{2ni} \geq \frac{(n-k)k}{4ni} \geq \frac{(n-k)(k-i)}{4ni}.$$

Claim 1 *$b^{a+b}2^a \leq (a+b)^{a+b}$ for any $a, b \in Log$.*

Proof: Case $a = 0$ is trivial. If $a = 1$, we have $(b+1)^{b+1} \geq b^{b+1} + (b+1)b^b \geq 2b^{b+1}$. Proceed by induction on a. Assuming the hypothesis for a, we have

$$b^{a+b+1}2^{a+1} \leq 2b(a+b)^{a+b} \leq 2(a+b)^{a+b+1} \leq (a+b+1)^{a+b+1}. \qquad \square \text{ (Claim 1)}$$

Put $a = n(2i-1)+k$, $b = (n-k)(k-2i+1)$ (hence $a+b = k(n-k+2i)$). We have

$$2^{ai}b^{i(a+b)} \leq (a+b)^{i(a+b)}.$$

On the other hand, $i^2 n \geq k(n-k)$ implies

$$ia - (a+b) = 2i^2n - in - kn + k^2 - ik \geq i^2n - i(n+k) \geq in(i-2) \geq 0,$$

hence

$$2^{a+b}b^{i(a+b)} \leq (a+b)^{i(a+b)}.$$

This means that

$$1 - \left(\frac{b}{a+b}\right)^i \geq 1 - \frac{1}{2} = \frac{1}{2},$$

thus

$$\begin{bmatrix} n \\ < k-i \end{bmatrix}_k \geq \frac{(n-k)(k-i)}{8ni} \begin{bmatrix} n \\ k-i \end{bmatrix}_k. \qquad \square$$

A.5 Proposition (Chernoff's bound) *PV proves: for any $n, k, i \in Log$ such that $k \leq n$ and $n > 0$,*

$$\frac{1}{n^n}\left(\begin{bmatrix} n \\ \leq k-i \end{bmatrix}_k + \begin{bmatrix} n \\ \geq k+i \end{bmatrix}_k\right) = O(4^{-i^2/n}).$$

Proof: The interesting case is to bound $\left[\begin{smallmatrix} n \\ <k-i \end{smallmatrix}\right]_k$ when $0 < i < k < n$. If $i \leq s = \sqrt{\frac{k(n-k)}{n}}$, there is also nothing to prove, because $i^2/n \leq (1-k/n)k/n \leq 1/4$, and the left-hand side is bounded by 1. Assume $i > s$. We know from A.2 and A.4 that

$$\frac{1}{n^n}\begin{bmatrix} n \\ < k-i \end{bmatrix}_k \leq c\left(1-\frac{k}{n}\right)\left(\frac{k}{i}-1\right)\sqrt{\frac{n}{(k-i)(n-k+i)}}\frac{k^{k-i}(n-k)^{n-k+i}}{(k-i)^{k-i}(n-k+i)^{n-k+i}}$$

for some c. Since $i > s$, we have

$$\left(1-\frac{k}{n}\right)\left(\frac{k}{i}-1\right)\sqrt{\frac{n}{(k-i)(n-k+i)}} = \frac{n-k}{i}\sqrt{\frac{k-i}{n(n-k+i)}} \leq$$
$$\leq \sqrt{\frac{(n-k)(k-i)}{k(n-k+i)}} = \sqrt{1-\frac{ni}{k(n-k+i)}} \leq 1.$$

As with the proof of A.4, it is not hard to show that $(1+1/a)^a \leq (1+1/b)^b$ and $(1+1/b)^{b+1} \leq (1+1/a)^{a+1}$ whenever $0 < a \leq b \in Log$, hence also $(1+1/a)^a \leq (1+1/b)^{b+1}$ for any a, b, in other words $(1+1/b)^{b+1}(1-1/(a+1))^a \geq 1$.

Let $a, b, j \in Log$, $0 < j < b$. Then

$$\left(1+\frac{1}{b-j}\right)^b\left(1-\frac{1}{a+j}\right)^a =$$
$$= \left(1+\frac{1}{b-j}\right)^{b-j+1}\left(1-\frac{1}{a+j}\right)^{a+j-1}\left(1+\frac{1}{b-j}\right)^{j-1}\left(1-\frac{1}{a+j}\right)^{-(j-1)} \geq$$
$$\geq \left(1+\frac{1}{b-j}\right)^{j-1}\left(1+\frac{1}{a+j-1}\right)^{j-1},$$

thus

$$\left[\left(1+\frac{1}{b-j}\right)^b\left(1-\frac{1}{a+j}\right)^a\right]^{(b-j)(a+j-1)(a+b)} \geq$$
$$\geq \left[\left(1+\frac{1}{b-j}\right)^{(b-j)(a+j-1)}\left(1+\frac{1}{a+j-1}\right)^{(b-j)(a+j-1)}\right]^{(j-1)(a+b)} \geq$$
$$\geq 2^{(j-1)(a+b-1)^2} \geq 2^{4(j-1)(b-j)(a+j-1)},$$

because $(x+y)^2 \geq 4xy$. Therefore

$$4^{2(j-1)}(a+j)^{a(a+b)}(b-j)^{b(a+b)} \leq (a+j-1)^{a(a+b)}(b-j+1)^{a(a+b)},$$

and by induction on i we have

$$4^{i^2-i}(a+i)^{(a+b)}(b-i)^{b(a+b)} \leq a^{a(a+b)}b^{a(a+b)}$$

for any $0 \leq i < b$. Put $a = k - i$ and $b = n - k + i$. Then

$$4^{i^2-n}\left(k^{k-i}(n-k)^{n-k+i}\right)^n \leq 4^{i^2-i}\left(k^{k-i}(n-k)^{n-k+i}\right)^n \leq$$
$$\leq \left((k-i)^{k-i}(n-k+i)^{n-k+i}\right)^n,$$

hence

$$\frac{k^{k-i}(n-k)^{n-k+i}}{(k-i)^{k-i}(n-k+i)^{n-k+i}} \leq 4^{(-i^2+n)/n} = 4 \cdot 4^{-i^2/n},$$

and finally

$$\frac{1}{n^n}\begin{bmatrix} n \\ < k-i \end{bmatrix}_k \leq 4c \cdot 4^{-i^2/n}. \qquad \square$$