Multifunction Algebras and the Provability of $PH \downarrow$

Chris Pollett

Department of Mathematics,
University of California, Los Angeles, 90095-1555 CA
cpollett@willow.math.ucla.edu

We introduce multifunction algebras B_i^τ where τ is a set of 0 or 1-ary terms used to bound recursion lengths. We show that if for all $\ell \in \tau$ we have $\ell \in O(|x|)$ then $B_i^\tau = FP^{\Sigma_{i-1}^p}(wit, \dot{\tau})$, those multifunctions computable in polynomial time with at most $O(p(\ell(x)))$ queries to a Σ_{i-1}^p witness oracle for $\ell \in \tau$ and p a polynomial. We use our algebras to obtain independence results in bounded arithmetic. In particular, we show if S_2^i proves $\Sigma_j^b = PH$ for some $j \geq i$ then $S_2^i \preceq_{B(\hat{\Sigma}_{i+1}^b)} S_2$. This implies if $P^{NP} \neq P^{NP}(\log)$ then S_2^1 does not prove the polynomial hierarchy collapses. We then consider a subtheory, Z, of the well-studied bounded arithmetic theory $S_2 = \cup_i S_2^i$. Using our algebras (mainly the $i = 1$ variants of our algebras) we establish the following properties of this theory: (1) Z cannot prove the polynomial hierarchy collapses. In fact, even $Z + \hat{\Pi}_0^b$-consequences of S_2 cannot prove the hierarchy collapses. (2) If $Z \subseteq S_2^i$ for any i then the polynomial hierarchy collapses. (3) If Z proves the polynomial hierarchy is infinite then for all i, $S_2^i \vdash \Sigma_i^p \neq \Pi_i^p$.

Key words: bounded arithmetic, complexity theory, multivalued functions, conservation results, independence results, polynomial hierarchy
1991 MSC: 03F30, 68Q15

1 Introduction

Over the past thirty years many techniques have been developed to try to resolve the $P = NP$ question. Recently there has been some research in how much mathematics is needed to formalize these techniques. The goal of such research would be to show there is a theory which on the one hand can formalize the currently available methods yet on the other hand can be shown to be unable to resolve $P = NP$. Since circuit lower bound results tend to involve probability arguments over families of finite spaces, it turns out the necessary counting arguments needed to perform these proofs can be done in

relatively weak fragments of arithmetic. Razborov [16] argues that monotone circuit lower bounds for clique whose proof uses the sunflower lemma can be carried out in V_1^1 and that Hastad [6] style results can be carried out in U_1^1 or $S_2(\alpha)$. Further Pudlak [15] has shown Ramsey's theorem can be proven in S_2 and Paris, Wilkie, Woods [12] have shown that for every n there is a prime between n and $2n$ is provable in S_2. For those unfamiliar with this area, V_1^1 roughly has induction for NE-predicates up to lengths of some number, U_1^1 has induction for NE-predicates up to lengths of lengths of a number, and $S_2(\alpha)$ has induction for predicates in the polynomial hierarchy with oracle predicate α.

There has also been some work trying to show independence in connection with these theories. Results of Takeuti [18] and Johannsen [7,8] have shown simple functions such as $x \dot{-} 1$ or $\lfloor \frac{x}{3} \rfloor$ are not definable in certain theories with induction on formulas where all the quantifiers are bounded by a length. Some slightly stronger theories involving weak quantifier replacement for $\hat{\Pi}_0^b$-formulas unable to define $\lfloor \frac{x}{3} \rfloor$ were given in the author's thesis [14]. For still stronger theories, conditional results are known. Razborov [17] has shown assuming the existence of pseudorandom number generators secure against attacks by quasi-polynomial sized circuit families that $S_2^2(\alpha)$ cannot prove super-polynomial lower bounds on circuit size for NP-predicates. In general, though, it seems hard to show that these larger fragments of S_2 cannot prove $P = NP$, $NP = co\text{-}NP$, or $P \neq NP$. Razborov's result is not as strong as one could hope in that the superpolynomial sized circuits must be coded using a second order predicate α and $S_2^2(\alpha)$ has limited ability to reason about such objects. Also, $S_2^2(\alpha)$ might still be able to prove no circuit family of size n^m can decide SAT for each integer m, which would still imply $P \neq NP$. In fact, assuming the existence of pseudo-random number generators Buss [4] shows the Σ_1^b-formula which says that C does not code a $|x|^m$-size circuit which computes satisfying assignments to any satisfiable instance x of SAT can be witnessed by a probabilistic p-time algorithm with error bounded away from a half. It is therefore not unreasonable to conjecture that if there are quasipolynomial sized circuit (or even p-time) secure pseudorandom generators (which is a strong form of $P \neq NP$) and $P = BPP$ (which is true with respect to a random oracle [1]) then in fact $S_2^1 \vdash P \neq NP$. So independence proofs for S_2^1 may be hard to prove. Nevertheless, proving a better independence result for this fragment or weaker ones is important. Not only would it rule out some methods of proving $P \neq NP$, but given the recent work on automatizability [2] such a result could lead to efficient theorem proving methods for these systems.

In this paper, we show one can allow a limited amount of induction on predicates at every level of the the polynomial hierarchy and still end up with a theory that cannot prove the polynomial hierarchy collapses. This theory Z is non-trivial for the following reasons: (1) As we have argued above S_2 can

formulate interesting complexity theoretic results. (2) The natural fragments S_2^i of S_2 do not contain Z unless the polynomial hierarchy collapses. (3) If Z can prove the polynomial hierarchy is infinite then for all i, $S_2^i \vdash \Sigma_i^p \neq \Pi_i^p$. Notice (3) above is non-trivial since S_2 might prove the hierarchy is infinite, yet it might be the case that $NP \neq co\text{-}NP$ is not provable in S_2^i for $i \leq 100$. The theory Z roughly consists of open axioms for the symbols in our language together with induction up to terms of form $|t|_{i+3}$ on Σ_i^p predicates for all $i > 0$. Here $|t|_{i+3}$ is $i+3$ applications of the length function to the term t.

Our proof method may be of independent interest. To show Z cannot prove the hierarchy collapses we first show: (1) if Z proves $PH \downarrow$ then $Z = S_2$. Then to get a contradiction, we show: (2) S_2 can $\hat{\Sigma}_1^b$-define $\lfloor \frac{x}{3} \rfloor$ and Z cannot. The result (1) holds for any theory contained in S_2 which for each i proves $\hat{\Sigma}_i^b$ induction up to some term which is $\Omega(|x|_m)$ for some m. (m may increase with i.) So this result may be useful for stronger independence result as well as conditional independence results. As one such application we show that if $P^{NP} \neq P^{NP}(\log)$ then S_2^1 cannot prove the hierarchy collapses. To show (2) we introduce multifunction algebras B_i^τ. We then give a collection of growth rates τ_1^Z such that $B_1^{\tau_1^Z}$ is precisely the $\hat{\Sigma}_1^b$-definable multifunctions of Z and show this class does not contain $\lfloor \frac{x}{3} \rfloor$. These algebras are also of interest in that for $i > 1$, they correspond to the multifunction classes $FP^{\Sigma_i^p}(wit, \dot{\tau})$ provided that $\ell \in O(|x|)$ for all $\ell \in \tau$. That is, the multifunctions computable in polynomial time with at most $p(\ell)$ queries to a Σ_{i-1}^p witnessing oracle where $\ell \in \tau$ and p is a polynomial. (The B in B_i^τ is for bounded query class.) These algebras may be useful to those who study machine independent characterizations of complexity classes. A less direct proof that these algebras are $FP^{\Sigma_i^p}(wit, \dot{\tau})$ was given in the author's thesis [14].

We now outline the format of the rest of this paper. In Section 2 we introduce the algebras B_i^τ and show for $i > 1$ they are the same class as $FP^{\Sigma_i^p}(wit, \dot{\tau})$. In Section 3 we characterize the $\hat{\Sigma}_i^b$-definable multifunctions of the theories $\hat{T}_2^{i,\tau}$. Then in Section 4, we use this characterization to establish the properties of Z and S_2^i mentioned at the start of this paper.

2 The algebras B_i^τ and B_i

Before we introduce our algebras let us make precise what we mean by multifunction.

Definition 1 *A* multifunction *is a set $f \subseteq \mathbb{N} \times \mathbb{N}$ such that for all $x \in \mathbb{N}$ there exists $\langle x, y \rangle \in f$. We express $\langle x, y \rangle \in f$ as $f(x) = y$. The composition of f, g is the relation $(f \circ g)(x) = z$ which holds if there is a $y \in \mathbb{N}$ such that $f(x) = y$ and $g(y) = z$. If f is a multifunction and r is a function, we write*

$f(x) > r(x)$ *if there exists* $y > r(x)$ *such that* $f(x) = y$. *We define* $f(x) < r(x)$ *and* $f(x) = r(x)$ *similarly.*

We now define some operations necessary to present our algebras.

Definition 2 *Let* e *be a multifunction.*

(i) *(*μ*-operator)* $(\mu x \leq z)[e(x, \mathbf{y}, z) = 0]$ *returns the least* $x \leq z$ *such that* $e(x, \mathbf{y}, z) = 0$ *holds and returns* $z + 1$ *if no such value exists.*

(ii) *(*W*-operator) The multifunction* $(Wx \leq z)[C(x, \mathbf{y}, z) = 0]$ *is the set of tuples*

$$\{\langle\langle \mathbf{y}, z\rangle, x\rangle | (C(x, \mathbf{y}, z) = 0 \wedge x \leq z) \vee \neg(\exists x \leq z)(C(x, \mathbf{y}, z) = 0 \wedge x = z+1)\}$$

(iii) *(*BPR^τ*)* f *is defined by* τ*-bounded primitive recursion from multifunctions* g, h, t, *and* r *if*

$$\begin{aligned} F(0, \mathbf{x}) &= g(\mathbf{x}) \\ F(n+1, \mathbf{x}) &= \min(h(n, \mathbf{x}, F(n, \mathbf{x})), r(n, \mathbf{x})) \\ f(n, \mathbf{x}) &= F(\ell(t(n, \mathbf{x})), \mathbf{x}) \end{aligned}$$

for some $r \in B_0$ *and for some* $t \in B_0$ *and* $\ell \in \tau$.

If g, h, t, and r are multifunctions then f obtained by BPR^τ results by viewing each step in the above iteration as a composition of multifunctions.

Definition 3

(i) $B_0^\tau = B_0$ *is the smallest class containing* $zero(x) = 0$, $S(x) := x + 1$, $MSP(x, i) := \lfloor x/2^i \rfloor$, $+$, $\cdot$, $\dot{-}$, $|x| := \lceil \log_2(x) + 1 \rceil$, $x\#y := 2^{|x||y|}$, *and closed under composition.*

(ii) B_1 *is the smallest class containing* B_0, *containing* $(Wx \leq |z|)[C(x, \mathbf{y}) = 0]$ *for* $C \in B_0$, *and closed under composition.*

(iii) B_1^τ *is the smallest class containing* B_0^τ, *containing* $(Wx \leq |z|)[C(x, \mathbf{y}) = 0]$ *for any* C *in* B_0, *closed under composition, and closed under* BPR^τ.

(iv) $(i > 1)$ B_i *is the smallest class containing* B_{i-1}, *containing* $(Wx \leq z)[D(x, \mathbf{y}) = 0]$ *for* $D \in B_{i-1}$ *and closed under composition.*

(v) $(i > 1)$ B_i^τ *is the smallest class containing* B_{i-1}, *containing* $(Wx \leq z)[D(x, \mathbf{y}) = 0]$ *for* $D \in B_{i-1}$, *closed under composition, and closed under* BPR^τ.

Definition 4 *Let* τ *be a set of iterms (*0 *or* 1*-ary* L_2*-terms).* $FP^{\Sigma_i^p}(wit, \tau)$ *is the class of multifunctions computable in polynomial time with fewer than* $O(\ell(h(x)))$ *witness queries to a* Σ_i^p*-oracle for some* $\ell \in \tau$ *and* $h \in B_0$. $FP^{\Sigma_i^p}(wit, s)$ *for some single function* s *is the class where the number of queries on inputs* x *of length* n *is bounded by* $O(s(n))$.

To guarantee the class $FP^{\Sigma_i^p}(wit, \tau)$ is closed under BPR^τ we next define a notion of a product closed set of iterms.

Definition 5 *A set τ of terms is* product closed *if for all $\ell(x), \ell'(x) \in \tau$ and $s, t \in B_0$ there is an $(\ell \cdot \ell') \in \tau$ and an $r \in B_0$ such that $(\ell \cdot \ell')(r(x)) \geq \ell(s(x)) \cdot \ell'(t(x))$.*

An example of a product closed set of iterms is $\{id\}$ since $id(s(x) \cdot t(x)) = id(s(x)) \cdot id(t(x))$.

Given a set τ of iterms it is not hard to define inductively a minimal set of iterms containing $\tau \cup cl$ which is product closed. Here cl is the set of all closed L_2-terms. We write $\dot{\tau}$ for the product closure of τ and $(|\dot{\tau}|)$ for the product closure of $|\tau|$.

We will frequently use the following B_0 functions:

$$
\begin{aligned}
2^{|y|} = 2^{|y|^1} &:= 1\#y & \max(x,y) &:= cond(K_{\leq}(x,y), y, x)\\
2^{|y|^n} = 2^{1\cdot|y|^n} &:= 2^{|y|^{n-1}}\#y & \min(x,y) &:= cond(K_{\leq}(x,y), x, y)\\
2^{k\cdot|y|^n} &:= 2^{|y|^n} \cdot 2^{(k-1)\cdot|y|^n} & 2^{\min(|y|,x)} &:= MSP(2^{|y|}, |y| \dot{-} x)\\
\mathrm{mod2}(a) &:= a \dot{-} 2 \cdot \lfloor \tfrac{1}{2} a \rfloor & LSP(x,i) &:= x \dot{-} MSP(x,i) \cdot 2^{\min(|x|,i)}\\
K_\neg(x) &:= 1 \dot{-} x. & \hat{\beta}(x, |t|, w) &:= MSP(LSP(w, Sx|t|), x|t|)\\
K_{\leq}(x,y) &:= K_\neg(y \dot{-} x) & Bit(i,x) &:= \hat{\beta}(i, 1, x)\\
K_\wedge(x,y) &:= x \cdot y & \dot{\beta}(x, |t|, s, w) &:= min(\hat{\beta}(x, |t|, w), s).
\end{aligned}
$$

$$
\begin{aligned}
K_=(x,y) &:= K_\wedge(K_{\leq}(x,y), K_{\leq}(y,x))\\
cond(x,y,z) &:= K_\neg(x) \cdot y + K_\neg(K_\neg(x)) \cdot z
\end{aligned}
$$

The k and n in $2^{k\cdot|y|^n}$ are fixed integers. Taking products of terms $2^{k\cdot|s|^n}$ we can construct terms representing $2^{p(|s|)}$ where p is any polynomial. $\hat{\beta}$ and $\dot{\beta}$ allow block sequence coding. Roughly, $\hat{\beta}(x, |t|, w)$ projects out the xth block (starting with a 0th block) of $|t|$ bits from w. $\dot{\beta}(x, |t|, s, w)$ returns the minimum of $\hat{\beta}(x, |t|, w)$ and s. For clarity, we write $2^{\ell}(x)$ for $2^{\min(|t(x)|, \ell(x))}$, if $\ell(x)$ is a term which is obviously less than $|t(x)|$ for some $t \in L_2$.

We define a pairing operation which will sometimes be more convenient than block coding.

Let $B = 2^{|\max(x,y)|+1}$. So B will be longer than either x or y. Hence, we can code pairs as $\langle x, y\rangle := (2^{|\max(x,y)|}+y)\cdot B + (2^{|\max(x,y)|}+x)$. To project out the coordi-

nates from an ordered pairs we use $\beta(1,w) := \hat{\beta}(0, \lfloor \frac{1}{2}|w| \rfloor \dot{-} 1, \hat{\beta}(0, \lfloor \frac{1}{2}|w| \rfloor, w))$ and $\beta(2,w) := \hat{\beta}(0, \lfloor \frac{1}{2}|w| \rfloor \dot{-} 1, \hat{\beta}(1, \lfloor \frac{1}{2}|w| \rfloor, w))$ which returns the left and right coordinate of the pair w. (The real Gödel beta function projects out $\beta(i,w)$, the ith element of a sequence w. However, as we never use this function we allow the suggestive notation.) To check if w is a pair we use $ispair(w) :=$

$$Bit(w, \lfloor \frac{1}{2}|w| \rfloor \dot{-} 1) = 1 \wedge 2 \cdot |\max(\beta(1,w).\beta(2,w))| + 2 = |w|.$$

Notice the above functions are all in B_0.

Definition 6 *Given $t \in B_0$ we define a monotonic term t^+ called the* dominator for t *by induction on the complexity of t. $t = t^+$ if t is constant or a variable. If t is $S(f)$ then t^+ is $S(f^+)$. If t is $f \circ g$ for $\circ$ a binary operation other than $\dot{-}$ or MSP then t^+ is $f^+ \circ g^+$. Lastly, if t is $f \dot{-} g$ or $MSP(f,g)$ then t^+ is f^+.*

Lemma 7 *($i \geq 1$) B_i^τ is closed under the following type of recursion:*

$$\begin{aligned} F(0,\mathbf{x}) &= g(\mathbf{x}) \\ F(n+1,\mathbf{x}) &= \min(h(n,\mathbf{x},F(n,\mathbf{x})), r(n,\mathbf{x})) \\ f(n,\mathbf{x}) &= F(\min(n, \ell(t(n,\mathbf{x}))), \mathbf{x}) \end{aligned}$$

where g and h are in B_i^τ, $r, t \in L_2$ and $\ell \in \tau$.

PROOF. Let r^+ denote $r^+(\ell(t), \mathbf{x})$. To define f we first define f' as

$$\begin{aligned} F'(0,\mathbf{x}) &= g(\mathbf{x}) \\ F'(n+1,\mathbf{x}) &= \min(F'(n,\mathbf{x}) + \\ &\qquad \min(h(n,\mathbf{x},\hat{\beta}(n,|r^+|,F'(n,\mathbf{x}))), r)(2^{i|r^+|}), 2^{(\ell(t)+1)|r^+|}) \\ f'(n,\mathbf{x}) &= F'(\ell(t(n,\mathbf{x})),\mathbf{x}) \end{aligned}$$

From f' we can define f as $\hat{\beta}(\min(n,\ell(t)), |r^+|, f'(n,\mathbf{x}))$. $\square$

To show $B_i^\tau = FP^{\Sigma_i^p}(wit, \dot{\tau})$ we use an arithmetization of the polynomial hierarchy which is essentially due to Kent-Hodgson [10]. Let L_2 be the language which consists of the initial functions of B_0. (The 2 in L_2 is due to the presence of $\# := \#_2$ in the language. In general, $x\#_k y := 2^{|x|\#_{k-1}|y|}$ and L_k where $k > 2$ is the language containing L_{k-1} together with $\#_k$.) We call a quantifier of the form $(\forall x \leq t)$ or $(\exists x \leq t)$ where t is an L_2-term not containing x a *bounded quantifier*. A formula is *bounded* if all it quantifiers are. A quantifier of the form $(\forall x \leq |t|)$ or of the form $(\exists x \leq |t|)$ is called *sharply bounded* and similarly a formula is *sharply bounded* if all its quantifiers are.

The bounded arithmetic hierarchy is defined as follows: $\Sigma^b_0 = \Pi^b_0$ is the class of sharply bounded formulas. Σ^b_i is the least class containing Π^b_{i-1}, closed under conjunction, disjunction, sharply bounded universal quantifiers, and bounded existential quantifiers. Similarly, Π^b_i is the least class containing Σ^b_{i-1}, closed under conjunction, disjunction, sharply bounded existential quantifiers, and bounded universal quantifiers. This hierarchy corresponds in a natural way to the polynomial time hierarchy: In the standard model Σ^b_i-formulas describe exactly predicates in Σ^p_i. Similarly, Π^b_i-formulas correspond to Π^p_i-predicates. This correspondence is proven in Buss [3].

The prenex bounded arithmetic hierarchy is defined similarly: $\hat{\Sigma}^b_0$ are those formulas of the form $(\exists x \leq |s|)\phi$ and $\hat{\Pi}^b_0$ are those formulas of the form $(\forall x \leq |s|)\phi$ where ϕ is an open formula. $\hat{\Sigma}^b_i$ are those formulas of the form $(\exists x \leq t)\phi$ where $\phi \in \hat{\Pi}^b_{i-1}$-formula. $\hat{\Pi}^b_i$ are those formulas of the form $(\forall x \leq t)\phi$ where $\phi \in \hat{\Sigma}^b_{i-1}$. For $i \geq 1$, the sets described by $\hat{\Sigma}^b_i$-formulas and Σ^b_i-formulas are equivalent [14,13]. (Given that we can do block coding in B_0 it is not hard to see one can do the necessary quantifier replacements and pairing. See Remark 18) Similarly, sets described by $\hat{\Pi}^b_i$-formulas and Π^b_i-formulas are equivalent. We call any formula in $\bigcup_i \hat{\Sigma}^b_i \cup \hat{\Pi}^b_i$ a *prenex formula.*

Definition 8 *For $\mathcal{C}$ a class of multifunctions, we write $\mathcal{C} = y$ to denote the relations of the form $f = y$ where $f \in \mathcal{C}$. We define $\mathcal{C} > y$ similarly.*

The next lemma connects the algebras B_i to the polynomial hierarchy.

Lemma 9 *(1) $B_0 = 0$ expresses precisely the open formulas of L_2. (2) For $i \geq 1$ $B_i = 0$ expresses any predicate which is a Boolean combination of $\hat{\Sigma}^b_{i-1}$-formulas. (3) For $i \geq 0$, $B_i = y$ can be expressed by a $\hat{\Sigma}^b_i$-formula.*

PROOF.

(1) The functions in B_0 are precisely the L_2-terms as $B^\tau_0 = B_0$ is just the closure of the initial functions of L_2 under composition. In particular, $B^\tau_0 = B_0$ can define $K_=$, $K_\leq$, $K_\wedge$, and $K_\neg$. From these terms one can express any open formula. Now suppose $t \in B_0$, then since $t \in L_2$, $t = 0$ is an open formula.

(2) The proof is by induction on i. When $i = 1$ by (1) it suffices to show B_1 can express sharply bounded quantifiers. Consider $A := (\exists x \leq |t|)B$ where B is an open formula equivalent to $f_B = 0$ in $B_0 = 0$. Then A can be expressed as

$$[K_\neg(K_=((Wx \leq |t|)[f_B = 0], |t| + 1))] = 0.$$

For all $j < i$ assume B_j expresses Boolean combinations of $\hat{\Sigma}^b_{j-1}$-predicates. Consider the $\hat{\Sigma}^b_{i-1}$-formula $A := (\exists x \leq t)B$ where B is a $\hat{\Pi}^b_{i-2}$-predicate which by assumption can be expressed in $B_{i-2} = 0$ as $f_B = 0$. Then A can be

expressed as $[K_{\neg}(K_{=}((Wx \leq t)[f_B = 0], t+1))] = 0$.

(3) We show the graph of any $f(\mathbf{x}) \in B_i$ can be expressed in the form $A_f(\mathbf{x}, y)$ where A_f is a $\hat{\Sigma}^b_i$-formula and y is bounded in A_f by a term t. In the $i = 1$ case, by using dummy quantifiers we can express the base functions of B_0 with $\hat{\Sigma}^b_1$-formulas. So it remains to show $\hat{\Sigma}^b_1$-formulas can express the graphs of functions defined by sharply bounded W-operator on open formulas (by (1)).

Suppose $f(x) = y$ is a function in B_0. Hence, f is an L_2-term. So we can define the graph of $((Wx \leq |t|)[f(x) = 0]) = y$ with the following formula which when prenexified is equivalent a $\hat{\Sigma}^b_1$-formula

$$[(\exists x \leq |t|)(f(x) = 0 \wedge y = x) \vee (\forall x \leq |t|)(f(x) > 0 \wedge y = |t| + 1)].$$

Now suppose $f = h(g_1(\mathbf{x}_1), \ldots, g_n(\mathbf{x}_n))$ and we can $\hat{\Sigma}^b_1$-define the functions $h(z_1, \ldots z_n)$ and $g_j(\mathbf{x}_j)$ with graphs H, G_j. Then we can define f with the following formula which when prenexified is a $\hat{\Sigma}^b_1$-formula:

$$\begin{aligned} y \leq t \wedge (\exists y_1 \leq t_1) \cdots (\exists y_n \leq t_n)[&G_1(\mathbf{x}_1, y_1) \\ &\wedge \cdots \wedge G_n(\mathbf{x}_1, y_1) \wedge H(y_1, \ldots, y_n, y)]. \end{aligned}$$

For $i \geq 1$ the same argument shows the graphs of $\hat{\Sigma}^b_i$-formulas are closed composition. What is left to show is that one can express with $\hat{\Sigma}^b_i$-formulas the graphs of multifunctions defined by W-operator. Suppose $f_C(x) \in B_{i-1,2}$. Our induction hypothesis is that the graph of $f_C(x) = y$ can be defined with some $\hat{\Sigma}^b_{i-1}$-formula, $C(x, y)$. We can define $(Wy \leq t)[f_C(x) = 0] = z$ with the following formula which when prenexified is a $\hat{\Sigma}^b_i$-formula

$$[(C(x, 0) \wedge x = z) \vee (\forall x \leq t)(\neg(C(x, 0) \wedge z = t + 1)].$$

□

The following lemma follows from the fact that we can compose multifunctions defined using BPR^{τ} in B^{τ}_i.

Lemma 10 *($i \geq 0$) $B^{\tau}_i = B^{\dot{\tau}}_i$.*

Theorem 11 *($i > 1$) $B^{\tau}_i = FP^{\Sigma^p_{i-1}}(wit, \dot{\tau})$ provided $\ell \in O(\{\dot{|x|}\})$ for all $\ell \in \tau$.*

PROOF. The condition on τ insures that B^{τ}_i can only perform polynomially many witness queries on a given input. First we argue $B^{\tau}_i \subseteq FP^{\Sigma^p_{i-1}}(wit, \dot{\tau})$.

By Lemma 9 (3), the graph of any $f \in B_{i-1}$ is contained in Σ^b_{i-1}. Hence, with one witness query to a Σ^p_{i-1}-oracle we can compute the value of any $f \in B_{i-1}$. Similarly with one witness query to a Σ^p_{i-1}-oracle we can compute $(Wy \leq t)(f(x,y) = 0)$ where $f \in B_{i-1}$. Suppose $M_f, M_g \in FP^{\Sigma^p_i}(wit, \dot{\tau})$, the machine that runs first M_g on input x followed by M_f on the result is still in this class since the number of queries will just be the sum of M_g's and M_f's queries which is boundable by a term in $\dot{\tau}$. Similarly, for closure under BPR^τ to compute M_f from M_g and M_h with bound $\ell(t)$ where $\ell \in \tau$, we first run M_g on x then run M_h on the output, then M_h on that output, and so on $\ell(t)$ times. The total number of queries will be $\ell(t)$ times the maximum number of queries M_h makes in an step. Since by definition $\dot{\tau}$ is product closed this total can be bounded by some term in $\dot{\tau}$.

Now we show $B^\tau_i \supseteq FP^{\Sigma^p_{i-1}}(wit, \dot{\tau})$. By Lemma 9 (2), any $B(\Sigma^p_{i-1})$-predicate can be expressed in $B_i = 0$. Let $M \in FP^{\Sigma^p_{i-1}}(wit, \dot{\tau})$. Let $(\exists y \leq t)C(q,y)$ be M's oracle and let $p(|x|)$ bound M's runtime and $\ell(t(x))$ where $\ell \in \dot{\tau}$ and $t \in L_2$ bound the number of queries M makes. Consider the following Π^p_{i-2}-predicate $Comp(x, w, v, j)$ "w is a valid computation of M on the input x with the first j queries answered by the first j bits of v and if the query k answer is a 1 then the witness w_k returned satisfies $C(q_k, w_k)$?" We assume the coding of a w is done using block coding and the maximum block length is $|k(x)|$ where $k \in L_2$. We assume block i contains a tuple that can be decoded using the pairing operations and this tuple gives the configuration of the machine at time i. Since we have a bound t on the size of witness returned by a query, we can bound the size of any computation w of M on x by some function $2^{p(|x|)}$ where p is a polynomial. Let $g(x) \in B_i$ be $\langle 1, ((Wy < 2^{p(|x|)})Comp(x, y, 1, 1))\rangle$. The first coordinate in this case is being used to say that $v = 1$. The y given in the remaining coordinate returned by g will be a computation on x where the oracle always responded 'no' except on the first query or y will be $2^{p(|x|)} + 1$ if there is no such computation. Now let $h(j, x, w')$ be

$$\langle 2 \cdot \beta(1, w') + 1, ((Wy < 2^{p(|x|)})Comp(x, y, j, 2 \cdot \beta(1, w') + 1)\rangle$$

if $\beta(2, w') \leq 2^{p(|x|)}$ and let $h(j, x, w')$ be

$$\begin{aligned}&\langle 4 \cdot MSP(\beta(1, w'), 1) + 1,\\ &((Wy < 2^{p(|x|)})Comp(x, y, j, 4 \cdot MSP(\beta(1, w'), 1) + 1)\rangle\end{aligned}$$

otherwise. Clearly h can be defined in B_i using *cond*. The coordinate of w' stores the value of the current v. The two cases of h correspond to the case where there was a computation of M on x with the first j queries answered according to v, and where there wasn't. In the first case, we shift v one bit to the left and put a 1 as the low order bit and then query whether there is a computation of M on x with the first j queries answered according to this v'.

The second case, is similar except to make v' we set the low order bit of v to 0, shift left and add 1. We can now define a multifunction $f \in B_i^\tau$ which returns a computation of M on input x. This function is defined from g, h and $r(x)$ a bound on the size of pairs that can occur in the above using the recursion of Lemma 7 up to $\ell(t)$. Now using the $\hat{\beta}$ function we can project out the last block of M's computation on x and so get the output of M on x. $\square$

Corollary 12 $(i > 1)$ $(B_i^\tau = 0) = P^{\Sigma_{i-1}^p}(\dot{\tau})$ *provided* $\ell \in O(\{|\dot{x}|\})$ *for all* $\ell \in \tau$.

PROOF. Suppose $M \in P^{\Sigma_{i-1}^p}(\dot{\tau})$ then by Theorem 11, M can be computed by some $f \in B_i^\tau$, since $P^{\Sigma_{i-1}^p}(\dot{\tau}) \subseteq FP^{\Sigma_{i-1}^p}(wit, \dot{\tau})$. Now $f = 1$ is equivalent to $1 \dot{-} f = 0$. So $B_i^\tau = 0$ contains $P^{\Sigma_{i-1}^p}(\dot{\tau})$. For the other direction consider some predicate $f = 0$ in $B_i^\tau = 0$. By Theorem 11, f can be computed by some M_f in $FP^{\Sigma_{i-1}^p}(wit, \dot{\tau})$. Let M be the $P^{\Sigma_i^p}(\dot{\tau})$ machine which uses the oracle $\exists wComp(x, w, v, j)$ (this is a non-witnessing oracle it just answers 1 or 0) and performs the same kind of search for a v as in Theorem 11. After having determined v for a correct computation it then asks the query $\exists w(Comp(x, w, v, j) \wedge Out(w) = 1)$ where $Out(w)$ is the output of M_f on input x in this computation. If the answer if 1, M outputs 0 otherwise M outputs 1. $\square$

3 Bounded Arithmetic

We now introduce some bounded arithmetic theories including Z. Then we characterize their $\hat{\Sigma}_i^b$-multifunctions.

We begin with $BASIC$ which consists of all substitution instances finite set of quantifier free axioms for the non-logical symbols of L_2. These axioms are listed in Buss [3] with the exception of the axioms for MSP and $\dot{-}$ which are listed in Takeuti [19].

Definition 13 *$EBASIC$ is obtained from $BASIC$ by adding the following three axioms:*

(1) $b < 2^{\min(k\cdot|d|,|d|^2)} \supset MSP(a \cdot 2^{\min(k\cdot|d|,|d|^2)} + b, \min(k \cdot |d|, |d|^2)) = a.$

(2) $(b < 2^{|d|} \wedge a < 2^{|d|}) \supset (\hat{\beta}(0, |d|, a \cdot 2^{|d|} + b) = b \wedge \hat{\beta}(1, |d|, a \cdot 2^{|d|} + b) = a).$

(3) $Si \cdot |a| \leq k \supset \hat{\beta}(i, |a|, w) = \hat{\beta}(i, |a|, LSP(w, k))$

The three new axioms allow $EBASIC$ to do simple reasoning about block codings of sequences. (see [14,13]). For example, they allow $EBASIC$ to prove

the following lemma the proof of which appears in [14,13].

Lemma 14 *Let* $m = \max(s(a), t(a,s))$ *and let* $t^* := t(a, \dot\beta(0, |m|, s(a), w))$ *where* $s(a), t(a,b) \in L_2$. *Then LIOpen and EBASIC prove:*

(a) $(\exists w \le 2^{2\cdot|m|})A(\dot\beta(0,|m|,s,w), \dot\beta(1,|m|,t^*,w))$
$\Leftrightarrow (\exists x \le s)(\exists y \le t)A(x,y)$
(b) $(\forall w \le 2^{2\cdot|m|})A(\dot\beta(0,|m|,s,w), \dot\beta(1,|m|,t^*,w))$
$\Leftrightarrow (\forall x \le s)(\forall y \le t)A(x,y)$.

We now define more powerful theories by adding various types of induction axioms to $BASIC$ and $EBASIC$.

Definition 15 *The* Ψ-IND^τ *axioms are the axioms* IND^ℓ_α:

$$\alpha(0) \wedge (\forall x)(\alpha(x) \supset \alpha(Sx)) \supset (\forall x)\alpha(\ell(x))$$

where $\alpha \in \Psi$ *and* $\ell \in \tau$.

As an example, let $id(a) = a$. Then Ψ-$IND^{\{id\}}$ is the usual induction for Ψ-formulas. Other common sets of terms are $\{|id|\}$, $\{||id||\}$ or $\{|id|_m\}$ where $|id|_0 = id$ and $|id|_m = ||id|_{m-1}|$. We often write IND, $LIND$ and $LLIND$ instead of $IND^{\{id\}}$, $IND^{\{|id|\}}$, and $IND^{\{||id||\}}$. The set $\{|id|_m\}$ for fixed m is just a singleton set; however, we will consider sets of terms such as $\{2^{p(|id|_i)}\}$ or $\{2^{2^{p(|id|_i)}}\}$ where p is a polynomial.

Definition 16 $(i \ge 0)$ T^i_2, S^i_2 *and* R^i_2 *are axiomatized as* $BASIC{+}\Sigma^b_i$-IND, $BASIC{+}\Sigma^b_i$-$LIND$, *and* $BASIC{+}\Sigma^b_i$-$LLIND$ *respectively.*

$\hat T^i_2$, $\hat S^i_2$, *and* $\hat R^i_2$ *are defined similarly except with* $\hat\Sigma^b_i$ *induction axioms.*

Let τ *be a set of iterms. We define* $\hat T^{i,\tau}_2$ *to be*

$$EBASIC{+}\hat\Sigma^b_i\text{-}IND^\tau$$

We define $S_2 := \cup_i S^i_2$ *and define* $Z := \cup_i Z_i$ *where* $Z_i := \hat T^{i,\{|id|_{i+3}\}}_2$.

It is shown in Pollett [13] that $T^i_2 = \hat T^i_2$ and $S^i_2 = \hat S^i_2$. It is not known if $R^i_2 = \hat R^i_2$. However, $\hat R^i_2 \subseteq R^i_2$ since one can show R^0_2 proves the axioms of $EBASIC$ [13]. Finally, it follows by the recursive doubling trick used to show $S^{i+1}_2 \supseteq T^i_2$ in Buss [3] that $Z_{i+1} \supseteq Z_i$ for each i.

Proofs in our theories will be carried out in the sequent calculus system LKB of Buss [3], together with the theories' axioms as initial sequents. It is often convenient, however, to reformulate inductions axioms as induction rules of inference:

Definition 17 *A Ψ-IND^τ inference is an inference:*

$$\frac{A(b), \Gamma \to A(Sb), \Delta}{A(0), \Gamma \to A(\ell(t(x))), \Delta}$$

where b is an eigenvariable and must not appear in the lower sequent, $t \in L_2$, $\ell \in \tau$, and $A \in \Psi$.

Buss [3] shows that one gets the same theory if one formulates S_2^i or T_2^i with inductions axioms or induction rules. The same proof works in the $\hat{T}_2^{i,\tau}$ case.

We will sometimes casually argue that a given formula is equivalent to a $\hat{\Sigma}_i^b$ formula so we can do induction on it.

Remark 18 *The following was shown in Pollett [13]. The proof is a straightforward induction argument. Let τ be a set of iterms all of which are $O(|x|)$ then $\hat{T}_2^{i,\tau}$ proves the $\hat{\Pi}_i^b$-$REPL^\tau$ axioms*

$$(\forall x \leq \ell(s))(\exists y \leq t(x,a))\alpha(x,y,a) \Leftrightarrow$$
$$(\exists w \leq 2 \cdot (t^+(\ell(s),a)\#(2^{\ell(s)})))(\forall x \leq \ell(s))\alpha(x, \dot{\beta}(x, |t^+(\ell(s),a)|, t, w))$$

where $\alpha \in \hat{\Pi}_i^b$, $\ell \in \tau$, and $s, t \in L_2$. Using the above kind of replacement where $\tau = \{id\}$ and Lemma 14 one can show the result we stated earlier that every Σ_i^b-predicate is equivalent to a $\hat{\Sigma}_i^b$-predicate. In fact, this is provable in $\hat{S}_2^i$. So this also gives $\hat{S}_2^i = S_2^i$, $\hat{T}_2^i = T_2^i$.

Let Ψ be a set of formulas. A theory T can Ψ-*define* a multifunction $f(x)$, if there is a Ψ-formula $A_f(x,y)$ such that $T \vdash \forall x \exists y A_f(x,y)$ and $\mathbb{N} \models A_f(x,y) \Leftrightarrow f(x) = y$. If T proves y is unique then we say T Ψ-defines the function f. We will be interested in Σ_i^b and $\hat{\Sigma}_i^b$-definability. A predicate is Δ_i^b with respect to a T if it is provably equivalent in T to both a Σ_i^b-formula and a Π_i^b-formula. A predicate is $\hat{\Delta}_i^b$ with respect to a T if it is provably equivalent to both a $\hat{\Sigma}_i^b$-formula and a $\hat{\Pi}_i^b$-formula.

Theorem 19 *($i \geq 0$) Suppose $\ell \in O(\{|\dot{x}|\})$ for all $\ell \in \tau$. Then $\hat{T}_2^{i,\tau}$ can $\hat{\Sigma}_i^b$-define the multifunctions in B_i^τ.*

PROOF. Since functions in $B_0^\tau = B_0$ are all L_2-terms, $EBASIC \subseteq \hat{T}_2^{i,\tau}$ can $\hat{\Sigma}_0^b$-define them. For $i \geq 1$, it suffices to show that $\hat{T}_2^{i,\tau}$ proves the class B_i^τ contains the appropriate W-operators, and is closed under composition and BPR_2^τ.

(W-operator) We first show $EBASIC$ can $\hat{\Sigma}_1^b$-define $(Wx \leq |t|)[f(x,\mathbf{z}) = 0]$ for $f(x,\mathbf{z})$ a function in $B_0 = B_0^\tau$. i.e., f is just an L_2-terms. To see this notice

$EBASIC$ proves

$$\exists y \leq |t|+1[(\exists x \leq |t|)(f(x,\mathbf{z}) = 0 \wedge y = x) \vee (\forall x \leq |t|)(f(x,\mathbf{z}) > 0 \wedge y = |t|+1)]$$

and the formula inside the $(\exists y \leq |t|+1)$ is equivalent to a $\hat{\Sigma}^b_1$-formula. Next we show $EBASIC$ can $\hat{\Sigma}^b_i$-define $(Wx \leq t)[f_C(x,\mathbf{z}) = 0]$ for $f_C(x,\mathbf{z}) \in B_{i-1}$. By Lemma 9, $f_C = 0$ is expressible by a $\hat{\Sigma}^b_{i-1}$-formula $C(x,\mathbf{z})$. So $EBASIC$ can $\hat{\Sigma}^b_i$-define $(Wx \leq t)[f_C(x,\mathbf{z}) = 0]$ since it proves

$$\begin{aligned}(\exists y \leq t+1)[(\exists x \leq t)(C(x,\mathbf{z}) \wedge y = x) \vee \\ (\forall x \leq t)(\neg C(x,\mathbf{z}) \wedge y = t+1)].\end{aligned}$$

and the formula inside the $(\exists y \leq t+1)$ is equivalent to a $\hat{\Sigma}^b_i$-formula. Since $\hat{T}^{i,\tau}_2 \supset EBASIC$, this shows $\hat{T}^{i,\tau}_2$ is also closed under the appropriate W-operators.

(Composition) Suppose $f = h(g_1(\mathbf{x}_1), \ldots g_n(\mathbf{x}_n))$ and that $\hat{T}^{i,\tau}_2$ can $\hat{\Sigma}^b_i$-define $h(z_1, \ldots z_n)$ and $g_j(\mathbf{x}_j)$ where $1 \leq j \leq n$ and where $h, g_j \in B_i^{|\tau|}$. Then there are $\hat{\Sigma}^b_i$-formulas H, $G_1, \ldots, G_n$ such that $\hat{T}^{i,\tau}_2 \vdash (\forall \mathbf{z})(\exists y \leq t)H(\mathbf{z}, y)$ and $\hat{T}^{i,\tau}_2 \vdash (\forall \mathbf{x}_j)(\exists y \leq t_j)G_j(\mathbf{x}_j, y)$, for $1 \leq j \leq n$. So

$$\begin{aligned}\hat{T}^{i,\tau}_2 \vdash (\forall \mathbf{x}_1) \cdots (\forall \mathbf{x}_n)(\exists y \leq t)(\exists y_1 \leq t_1) \cdots (\exists y_n \leq t_n)(G_1(\mathbf{x}_1, y_1) \\ \wedge \cdots \wedge G_n(\mathbf{x}_1, y_1) \wedge H(y_1, \ldots y_n, y)).\end{aligned}$$

The formula inside the $(\exists y \leq t)$ is equivalent to a $\hat{\Sigma}^b_i$-formula in $EBASIC$ and it defines f.

(BPR^τ) Suppose f is obtained by BPR^τ from g and h which are $\hat{\Sigma}^b_i$-definable, $r, t \in L_2$, and $\ell \in \tau$. Let G and H be the $\hat{\Sigma}^b_i$-graphs of g and h such that $\hat{T}^{i,\tau}_2 \vdash (\forall \mathbf{x})(\exists y \leq t_1)G(\mathbf{x}, y)$ and $\hat{T}^{i,\tau}_2 \vdash (\forall n, \mathbf{x}, u)(\exists v \leq t_2)H(n, \mathbf{x}, u, v)$. We can assume $r(0, \mathbf{x}) \leq t_1(\mathbf{x})$. So let $A(n, \mathbf{x}, w, y)$ be

$$\begin{aligned}&G(\mathbf{x}, \dot{\beta}(0, |r^+(\ell(t), \mathbf{x})|, r(0, \mathbf{x}), w))) \wedge \\ &\dot{\beta}(n, |r^+(\ell(t), \mathbf{x})|, r(\ell(t)), \mathbf{x}), w) = y \wedge \\ &(\forall j < \ell(t))((H(j, \mathbf{x}, \hat{\beta}(j, |r^+(\ell(t), \mathbf{x})|, w), \hat{\beta}(Sj, |r^+(\ell(t), \mathbf{x})|, w)) \\ &\wedge \hat{\beta}(Sj, |r^+(\ell(t), \mathbf{x})|, w) < r(n, \mathbf{x})) \vee \hat{\beta}(Sj, |r^+(\ell(t), \mathbf{x})|, w) = r(n, \mathbf{x}))\end{aligned}$$

and let $B(n, \mathbf{x})$ be $(\exists y \leq r)(\exists w \leq 2^{\ell(t)\cdot(|r^+|+1)})A(n, \mathbf{x}, z, w, y)$. Let $F(n, \mathbf{x}, y)$ denote the formula within the $(\exists y \leq r)$. Since $\ell \in O(|x|)$, this formula is equivalent to a $\hat{\Sigma}^b_i$-formula in $\hat{T}^{i,\tau}_2$ and we can define f if we can show

$$(\forall \mathbf{x}, n)(\exists y \leq r)F(\ell(t(n, \mathbf{x})), \mathbf{x}, y).$$

So it suffices to show $(\forall \mathbf{x}, n)B(\ell(t), \mathbf{x})$. Now B is also equivalent to a $\hat{\Sigma}^b_i$-formula, so $\hat{T}^{i,\tau}_2$ can use IND^τ_B axioms. Since $\hat{T}^{i,\tau}_2$ proves $(\forall \mathbf{x})(\exists y \leq t_1)G$, it proves $B(0, \mathbf{x})$. Suppose $\hat{T}^{i,\tau}_2 \vdash B(m, \mathbf{x})$ where $m \leq \ell(t)$. So there are v, w, y satisfying $A(m, \mathbf{x}, w, y)$. If we set $y' = h(m, \mathbf{x}, y)$, and

$$w' = y' \cdot 2^{\min((m+1)\cdot|r^+|, \ell(c)\cdot|r^+|)} + LSP(w, (m+1)\cdot|r^+|)$$

then by axioms (1) and (3) of $EBASIC$, $\hat{T}^{i,\tau}_2 \vdash A(m+1, \mathbf{x}, z, w', y')$. Hence, $\hat{T}^{i,\tau}_2 \vdash B(m+1, \mathbf{x})$. By the IND^τ_B axioms, $\hat{T}^{i,\tau}_2 \vdash (\forall \mathbf{x}, n)B(\ell(t), \mathbf{x})$. $\square$

Let T be $EBASIC$ or $\hat{T}^{i,\tau}_2$. By Parikh's Theorem [11], T can $\hat{\Sigma}^b_m$-define a function f if and only if there is a $\hat{\Sigma}^b_m$-formula $A_f(x, y)$ such that T proves $(\forall x)(\exists! y \leq t)A_f(x, y)$. For a multifunction one does not have to show uniqueness. An $E\hat{\Sigma}^b_m$-formula is a formula $(\exists y \leq t)A$ where $A \in \hat{\Sigma}^b_m$. We write $L\Psi$ (lexicographically Ψ) for the set of formulas that can be made into Ψ-formulas by introducing dummy quantifiers. We define a witness predicate as follows.

If $A(\mathbf{a}) \in L\hat{\Pi}^b_{m-1}$ then $Wit^m_A(w, \mathbf{a}) := w = 0 \wedge A(\mathbf{a})$

If $A(\mathbf{a})$ is $(\exists x \leq t(\mathbf{a}))B$ and $A \in \hat{\Sigma}^b_m$ then $Wit^m_A(w, \mathbf{a}) := w \leq t(\mathbf{a}) \wedge B(w, \mathbf{a})$

If $A(\mathbf{a})$ is $(\exists x_1 \leq t_1)(\exists x_2 \leq t_2)B$ and $A \in E\hat{\Sigma}^b_i$ then

$$\begin{aligned} Wit^m_A(w, \mathbf{a}) := & \mathit{ispair}(w) \wedge \beta(1, w) \leq t_1 \wedge \beta(2, w) \leq t_2 \wedge \\ & B(\beta(1, 2), \beta(2, w), \mathbf{a}). \end{aligned}$$

Thus, if $A \in LE\hat{\Sigma}^b_m$ then Wit^m_A is equivalent in $EBASIC$ to a $\hat{\Pi}^b_{m-1}$-formula. The witness predicate above is simplified from Buss [3]. The simplification arises because we are in the prenex setting. From the definition of witness the next useful properties follow:

Lemma 20 *($m \geq 1$) If $A(\mathbf{a}) \in LE\hat{\Sigma}^b_m$, then: (a) $EBASIC \vdash Wit^m_A(w, \mathbf{a}) \supset A(\mathbf{a})$. (b) There is a t_A so that $EBASIC \vdash A(\mathbf{a}) \Leftrightarrow (\exists w \leq t_A(\mathbf{a}))Wit^m_A(w, \mathbf{a})$. (c)For t_A, $EBASIC \vdash Wit^m_A(w, \mathbf{a}) \supset w \leq t_A$.*

PROOF. (a) This statement is immediate from the definition of Wit^m_A.

(b) If $A \in \hat{\Sigma}^b_m$ then t_A is just the bounds on the outermost existential quantifier. Otherwise, if the outermost two existential quantifiers are bounded by t_1 and t_2, their pair is bounded by $2^{2\cdot(\max(t_1,t_2)+1)}$.

(c) Follows from (b) and the definition of Wit^m_A. In particular, the definition of *ispair* forces any pair for a witness to be unique. $\square$

For a cedent $\Gamma = \{A_1, \ldots, A_n\}$ we use $\vee\Gamma$ (resp. $\wedge\Gamma$) to denote the disjunction (resp. conjunction) of its formulas. We write $w = \langle\langle w_1, \cdots, w_n\rangle\rangle$ to denote pairings of the form $\langle w_1, \langle w_2, \cdots, \langle w_{n-1}, w_n\rangle \cdots\rangle\rangle$. We will use this convention in defining witnesses for $Wit^m_{\wedge\Gamma}$ and $Wit^m_{\vee\Gamma}$.

We define $Wit^m_{\wedge\Gamma}(w, \mathbf{a})$ by induction. If $\Gamma = \emptyset$, define $Wit^m_{\wedge\Gamma}(w, \mathbf{a})$ to be $0 = 0$. If $\Gamma = \{A\}$ then $Wit^m_{\wedge\Gamma}(w, \mathbf{a})$ is $Wit^m_A(w, \mathbf{a})$. If $\Gamma = \{A_1, \ldots, A_n\}$, let Γ' be $\{A_2, \ldots A_n\}$ and set $Wit^m_{\wedge\Gamma}(w, \mathbf{a})$ to be $Wit^m_{A_1}(\beta(1, w), \mathbf{a}) \wedge Wit^m_{\wedge\Gamma'}(\beta(2, w), \mathbf{a})$.

Now we define $Wit^m_{\vee\Gamma}(w, \mathbf{a})$. If $\Gamma = \emptyset$, define $Wit^m_{\vee\Gamma}(w, \mathbf{a})$ to be $\neg(0 = 0)$. If $\Gamma = \{A\}$ then $Wit^m_{\vee\Gamma}(w, \mathbf{a})$ is $Wit^m_A(w, \mathbf{a})$. Otherwise, if $\Gamma = \{A_1, \ldots, A_n\}$, let Γ' be $\{A_2, \ldots A_n\}$ and define $Wit^m_{\vee\Gamma}(w, \mathbf{a})$ to be $(Wit^m_{A_1}(\beta(1, w), \mathbf{a}) \wedge w_1 \leq t_{A_1}) \vee Wit^m_{\vee\Gamma'}(\beta(2, w), \mathbf{a})$ where t_{A_j} are from Lemma 20.

Both $Wit^m_{\wedge\Gamma}$ and $Wit^m_{\vee\Gamma}$ are equivalent to $\hat{\Pi}^b_{m-1}$-formulas in $EBASIC$.

Lemma 21 ($m \geq 1$) *Let Γ be a cedent of $LE\hat{\Sigma}^b_m$-formula with free variables $\mathbf{a}$. There is a term t_Γ such that $EBASIC \vdash Wit^m_{\wedge\Gamma}(w, \mathbf{a}) \supset w \leq t_\Gamma$ and $EBASIC \vdash Wit^m_{\vee\Gamma}(w, \mathbf{a}) \supset w \leq t_\Gamma$.*

We also have

$$EBASIC \vdash (\exists w \leq t_\Gamma) Wit^m_{\wedge\Gamma}(w, \mathbf{a}) \rightarrow (\exists w \leq t_\Delta) Wit^m_{\vee\Delta}(w, \mathbf{a})$$

if and only if $EBASIC \vdash \Gamma \rightarrow \Delta$.

PROOF. This follows from the definition of witness for a cedent, the fact that witnesses for a cedent are made up of pairs, and by the bounds for witnesses for formulas given by Lemma 20. □

Theorem 22 ($i \geq 1$) *Suppose $\ell \in O(\{|\dot{x}|\})$ for all $\ell \in \tau$ and $\hat{T}^{i,\tau}_2 \vdash \Gamma \rightarrow \Delta$ where Γ and Δ are cedents of $LE\hat{\Sigma}^b_i$-formulas. Let $\mathbf{a}$ be the free variables in this sequent. Then there is an $f \in B^\tau_i$ such that:*

$$\hat{T}^{i,\tau}_2 \vdash Wit^i_{\wedge\Gamma}(w, \mathbf{a}) \supset Wit^i_{\vee\Delta}(f(w, \mathbf{a}), \mathbf{a}).$$

PROOF. This is proved by induction on the number of sequents in an $\hat{T}^{i,\tau}_2$-proof of $\Gamma \rightarrow \Delta$. By cut elimination, we can assume all the sequents in the proof are in $LE\hat{\Sigma}^b_i$. In the base case, the proof consists of sequent $\rightarrow A$ where A is a logical axiom, an equality axiom, or an $EBASIC$ axiom. In each of these cases the witness predicate is $A \wedge w = 0$. So we can choose f to be the zero function. The weak inferences, structural inferences, and cut can be handled in essentially the same way as in the S^i_2 case of the witnessing argument in Buss [3]. The remaining cases are the bounded quantifier rules and induction.

We show the ($\exists \leq$:left) and ($\exists \leq$:right) – the ($\forall \leq$:left) and($\forall \leq$:right) are similar – and, of course, we show the $\hat{\Sigma}^b_i$-IND^τ case.

($\exists$:left case)

$$\frac{b \leq t, A(b), \Gamma \rightarrow \Delta}{\exists x \leq t A(x), \Gamma \rightarrow \Delta}$$

By hypothesis there is a $g \in B^\tau_i$ such that

$$\hat{T}^{i,\tau}_2 \vdash Wit^i_{b \leq t \wedge A \wedge \Gamma}(w, \mathbf{a}, b) \supset Wit^i_{\vee \Delta}(g(w, \mathbf{a}, b), \mathbf{a}, b).$$

There are three subcases. In each case, we need to determine a value for the free variable b and then run g using that value. First, suppose $(\exists x \leq t)A(x) \in E\hat{\Sigma}^b_i$. If w is a witness for $(\exists x \leq t)A(x) \wedge \Gamma$, then $\beta(1, (\beta(1, w))$ is a value for b such that $A(b)$ holds and $\beta(2, \beta(1, w))$ is a witness for $A(b)$. Let our new witness function be

$$f(w, \mathbf{a}) = g(\langle\langle 0, \beta(2, \beta(1, w)), \beta(2, w)\rangle\rangle, \mathbf{a}, \beta(1, \beta(1, w))).$$

It is easy to see that

$$\hat{T}^{i,\tau}_2 \vdash Wit^i_{(\exists x \leq t)A \wedge \Gamma}(w, \mathbf{a}) \supset Wit^i_{\vee \Delta}(f(w, \mathbf{a}), \mathbf{a}).$$

In the second case suppose $(\exists x \leq t)A(x) \in \hat{\Sigma}^b_i$. If w is a witness for $(\exists x \leq t)A(x) \wedge \Gamma$, then $\beta(1, w)$ is a value for b such that $A(b)$ holds. Let our new witness function be

$$f(w, \mathbf{a}) = g(\langle\langle 0, 0, \beta(2, w)\rangle\rangle, \mathbf{a}, \beta(1, w)).$$

It follows that

$$\hat{T}^{i,\tau}_2 \vdash Wit^i_{(\exists x \leq t)A \wedge \Gamma}(w, \mathbf{a}) \supset Wit^i_{\vee \Delta}(f(w, \mathbf{a}), \mathbf{a}).$$

The last case is when $(\exists x \leq t)A(x) \in L\hat{\Sigma}^b_{i-2}$. (Notice by the definitions of $\hat{\Sigma}^b_i$ and $\hat{\Pi}^b_i$ if $(\exists x \leq t)A(x) \in L\hat{\Pi}^b_{i-1}$ then $(\exists x \leq t)A(x) \in L\hat{\Sigma}^b_{i-2}$. So $(\exists x \leq t)A(x) \in L\hat{\Sigma}^b_{i-2}$ is the only remaining case.) In this case, let f_A be the multifunction in B_{i-1} which by Lemma 9 has the property that $f_A(x) = 0$ iff $A(x)$. We define f to be the same as in the above case except rather than use $\beta(1, \beta(1, w))$ to give a value b we instead use the $B_i \subset B^\tau_i$ multifunction $(Wx \leq t)[f_A(x) = 0]$ to give a value for b. Note if $(\exists x \leq t)A(x) \in \hat{\Sigma}^b_0$ then t is sharply bounded and A is open so this function is definable in B_1.

($\exists$:right case)

$$\frac{\Gamma \rightarrow A(t), \Delta}{t \leq s, \Gamma \rightarrow (\exists x \leq s)A(x), \Delta}$$

By hypothesis there is a $g \in B_i^\tau$ such that

$$\hat{T}_2^{i,\tau} \vdash Wit^i_{\wedge\Gamma}(w, \mathbf{a}) \supset Wit^i_{A(t)\vee\Delta}(g(w, \mathbf{a}), \mathbf{a}).$$

The definition of Wit^i implies

$$\hat{T}_2^{i,\tau} \vdash Wit^i_{t\leq s\wedge\Gamma}(w, \mathbf{a}) \supset t \leq s \wedge Wit^i_{\wedge\Gamma}(\beta(2, w), \mathbf{a}).$$

So if $A \in \hat{\Sigma}_i^b$ define $f := \langle\langle t(\mathbf{a}), \beta(1, g(\beta(2, w), \mathbf{a})\rangle, \beta(2, g(\beta(2, w), \mathbf{a}))\rangle$.

If $A \in \hat{\Pi}_{i-1}^b$ define $f := \langle t(\mathbf{a}), \beta(2, g(\beta(2, w), \mathbf{a}))\rangle$.

For all other A define $f := g(\beta(2, w), \mathbf{a}))$.

These functions are all B_i^τ and note that

$$\hat{T}_2^{i,\tau} \vdash Wit^i_{t\leq s\wedge\Gamma}(w, \mathbf{a}) \supset Wit^i_{(\exists x\leq s)A(x)\vee\Delta}(f(w, \mathbf{a}), \mathbf{a}).$$

($\hat{\Sigma}_i^b$-IND^τ case)

$$\frac{A(b), \Gamma \to A(Sb), \Delta}{A(0), \Gamma \to A(\ell(t)), \Delta}$$

where $\ell \in \tau$. By hypothesis there is a $g \in B_i^\tau$ such that

$$\hat{T}_2^{i,\tau} \vdash Wit^i_{A(b)\wedge\Gamma}(w, \mathbf{a}) \supset Wit^i_{A(Sb)\vee\Delta}(g(w, \mathbf{a}), \mathbf{a}).$$

Let $h(m, w, \mathbf{a}, b)$ be

$$cond(Wit^i_{A(Sb)\vee\Delta}(m, \mathbf{a}, b), m, g(\langle m, \beta(2, w)\rangle, \mathbf{a}, b))$$

Define f by BPR_2^τ in the following way

$$\begin{aligned} F(0, w, \mathbf{a}) &= \langle\beta(1, w), 0\rangle \\ F(b + 1, w, \mathbf{a}) &= \min(h(F(b, w, \mathbf{a}), w, \mathbf{a}, b), t_{\vee A(Sb)\vee\Delta}) \end{aligned}$$

Define $f(u, w, \mathbf{a}) := h(\min(u, \ell(t)), w, \mathbf{a})$. Recall $t_{\vee A(Sb)\vee\Delta}$ is the term guaranteed to bound a witness for $A(Sb) \vee \Delta$ by Lemma 21. It is easy to see

$$\hat{T}_2^{i,\tau} \vdash Wit^i_{A(0)\wedge\Gamma}(w, \mathbf{a}) \supset Wit^i_{A(0)\vee\Delta}(f(0, w, \mathbf{a}), \mathbf{a})$$

From this one can then show that

$$\begin{aligned} \hat{T}_2^{i,\tau} \vdash Wit^i_{A(0)\wedge\Gamma}(w, \mathbf{a}) \wedge Wit^i_{A(b)\vee\Delta}(f(b, w, \mathbf{a}), b, \mathbf{a}) \\ \supset Wit^i_{A(Sb)\vee\Delta}(f(Sb, w, \mathbf{a}), Sb, \mathbf{a}). \end{aligned}$$

Since t is in τ, it then follows by $\hat{\Sigma}^b_i$-IND^τ that

$$\hat{T}^{i,\tau}_2 \vdash Wit^i_{A(0)\wedge\Gamma}(w,\mathbf{a}) \supset Wit^i_{A(\ell(t))\vee\Delta}(f(\ell(t),w,\mathbf{a}),\mathbf{a}).$$

This completes all possible cases and the proof. □

Corollary 23 ($i \geq 1$) *Let $\ell \in O(\{|\dot{x}|\})$ for all $\ell \in \tau$. (1) The $\hat{\Sigma}^b_i$-definable functions of $\hat{T}^{i,\tau}_2$ are contained in B^τ_i. (2) For $i > 1$ the $\hat{\Sigma}^b_i$-definable multifunctions of $\hat{T}^{i,\tau}_2$ are precisely $B^\tau_i = FP^{\Sigma^p_{i-1}}(wit, \dot{\tau})$. (3) The $\hat{\Sigma}^b_1$-definable multifunctions of $\hat{T}^{1,\tau}_2$ are A^τ_1, the smallest class containing operators $(Wy \leq t)(y = y)$ where $t \in L_2$ and containing B_1 and closed under composition and BPR^τ.*

PROOF. We show (2) first. For the 'if' direction we use Theorem 19. For the other direction consider Theorem 22 when we take Γ empty and Δ to be a $E\hat{\Sigma}^b_i$ formula $(\exists y \leq t(x))A(x,y)$ provable in $\hat{T}^{i,\tau}_2$. Then we get that there is a B^τ_i function f such that $\hat{T}^{i,\tau}_2 \vdash\!\!\rightarrow Wit^i_A(x, f(x))$. Given the definition of witness we thus have $\hat{T}^{i,\tau}_2 \vdash\!\!\rightarrow A(x, \beta(1, f(x))$. So $k := \beta(1, f(x))$ give at least one value such that $A(x,y)$ holds. In the case where $A(x,y)$ defines a function in $\hat{T}^{i,\tau}_2$ this is the only value y such that $A(x,y)$ holds. This shows (1). Suppose A is multivalued. From k we next define a B^τ_i function h such that $h(x) = y$ iff $A(x,y)$. Suppose $A(x,y)$ is of the form $(\exists z \leq s)B(x,y,z)$ where $B \in \hat{\Pi}^b_{i-1}$. We do the following: (a) Compute $k(x) = y_0$. (b) Ask the queries $(Wy \leq t)(y = y)$ and $(Wz \leq s)(z = z)$. Let y_1 and z_1 be the oracle responses. (c) Ask the $\hat{\Sigma}^b_{i-1}$-query $\neg B(x, y_1, z_1)$. If the answer is '1' output y_0. Otherwise, output y_1. For $i > 1$, f can be easily constructed using *cond* and Lemma 9 as a composition of B^τ_i multifunctions so will be B^τ_i. The purpose of step (b) is to nondeterministically get values for y_1 and z_1. If these values happen to witness $(\exists y \leq t)A$ then y_1 is output, otherwise y_0 is output. Notice this argument show the $\hat{\Sigma}^b_1$-definable multifunctions of $\hat{T}^{1,\tau}_2$ are in A^τ_1. For the other direction the proof is the same as Theorem 19 once one observes that $(Wy \leq t(x))(y = y)$ can be $\hat{\Sigma}^b_1$-defined in $\hat{T}^{1,\tau}_2$ using $(\exists z \leq t(x))(z = y)$. □

The next two theorems are from Pollett [13]. We will have need of them in the next section.

Theorem 24 ($i \geq 1$) *Suppose for all $\ell \in \tau$ that $\ell \in O(\{|\dot{x}|\})$. Let $2^{\dot{\tau}}$ be the set of terms 2^ℓ where $\ell \in \dot{\tau}$. Then $\hat{T}^{i,2^{\dot{\tau}}}_2 \preceq_{B(\hat{\Sigma}^b_{i+1})} \hat{T}^{i+1,\tau}_2$. The $\hat{\Delta}^b_{i+1}$-predicates of both these classes are $P^{\Sigma^p_{i-1}}(\dot{\tau})$.*

Theorem 25 ($i \geq 0, k \geq 2$) *[Pollett [13], Corollary 57] The $\hat{\Delta}^b_{i+k}$-predicates of $\hat{T}^{i,\tau}_2$ are $P^{\Sigma^p_{i+k-1}}(1)$.*

We went to some effort establishing Corollary 23 since Pollett [13] does not classify the $\hat{\Sigma}_1^b$-definable multifunctions of $\hat{T}_2^{1,\tau}$ which we will have need of in the next section. In particular we will need the upper bound on the $\hat{\Sigma}_1^b$-definable functions of $\hat{T}_2^{1,\tau}$ below.

4 Properties of Z and S_2^i

In this section, we prove the properties of Z and S_2^i mentioned in the abstract of this paper.

Theorem 26 *If $Z \subseteq T_2^i$ for any i then the polynomial hierarchy collapses to $B(\Sigma_{i+2}^p)$. Since $S_2^i \subseteq T_2^i$, this also implies if Z is contained in S_2^i the polynomial hierarchy collapses.*

PROOF. $Z \subseteq T_2^i$ implies $Z_{i+2} \subseteq T_2^i$. The $\hat{\Delta}_{i+2}^b$-predicates of T_2^i are $P^{\Sigma_{i+1}^p}(1)$ by Theorem 25. By Corollary 24, Z_{i+2}'s $\hat{\Delta}_{i+2}^b$-predicates are $P^{\Sigma_{i+1}^p}((\{|\dot{id}|_{i+5}\}))$. It is not hard to exhibit complete problems for the latter class. Hence, if $Z_{i+2} \subseteq T_2^i$ then

$$P^{\Sigma_{i+1}^p}(1) = P^{\Sigma_{i+1}^p}((\{|\dot{id}|_{i+5}\}))$$

and so for some k, $P^{\Sigma_{i+1}^p}[k] = P^{\Sigma_{i+1}^p}[k+1]$, the result then follows from Chang and Kadin [9,5]. □

Definition 27 *Define $2 \uparrow 0(x) := x$, $2 \uparrow (i+1)(x) := 2^{2\uparrow i(x)}$. Let τ_i^Z be the set of iterms of the form $2 \uparrow j(p(|x|_j))$ for $j \geq i+3$ and p any polynomial.*

Let B_i^Z be short-hand for $B_i^{\tau_i^Z}$.

As a consequence of Theorem 24 and the fact that a statement provable in Z must in fact be provable in Z_i (recall Z_{i+1} contains Z_i) for some large enough i, we have:

Lemma 28

$(i > 0)$ $\hat{T}_2^{i,\tau_i^Z} \preceq_{B(\hat{\Sigma}_{i+1}^b)} Z$.

$(i > 0)$ *The $\hat{\Sigma}_i^b$-definable multifunctions of Z are precisely B_i^Z.*

To prove Z cannot prove the collapse of the hierarchy we first show if Z proves $PH \downarrow$ then $Z = S_2$. This requires the next lemma.

Lemma 29 *There is a $\hat{\Sigma}^b_i$-formula $U_i(e,x,z)$ such that for any $\phi(x) \in \hat{\Sigma}^b_i$ there is a fixed number e_ϕ and*

$$EBASIC \vdash U_i(e_\phi, x, 2^{|x|^{e_\phi}}) \equiv \phi(x).$$

PROOF. Note since we have pairing we are not losing any generality by only considering 1-ary ϕ's. Also by Lemma 9 we can express any open formula $A(x,\mathbf{y})$ as an equation $f(x,\mathbf{y}) = 0$ where $f \in L_2$. By induction, on the complexity of A this is provable in $EBASIC$. So any $\hat{\Sigma}^b_i$-formula $\phi(x)$ is provably equivalent in $EBASIC$ to one of the form

$$(\exists y_1 \le t_1)\cdots(Qy_i \le t_i)(Q'y_{i+1} \le |t_{i+1}|)(t_{i+2}(x,\mathbf{y}) = 0)$$

where the quantifiers Q and Q' will depend on whether i is even or odd. We fix some coding scheme for the 11 symbols of L_2 as well as for the $i+2$ variables $x, y_1, \ldots, y_{i+1}$. We use $\ulcorner\urcorner$ to denote the code for some symbol. i.e., $\ulcorner = \urcorner$ is the code for $=$. We choose our coding so that all codes require less than $|i+12|$ bits and we use 0 as $\ulcorner NOP \urcorner$ meaning no operation. The code for a term t is a sequence of blocks of length $|i+12|$ that write out t in postfix order. So $x+y_1$ would be coded as the three blocks $\ulcorner x\urcorner\ulcorner y_1\urcorner\ulcorner + \urcorner$. The code for a $\hat{\Sigma}^b_i$-formula will be $\langle\langle \ulcorner t_1\urcorner, \ldots, \ulcorner t_{i+3}\urcorner\rangle\rangle$. We now describe $U_i(e,x,z)$. It will be obtained from the formula

$$(\exists w \le z)(\exists y_1 \le z)(\forall j \le |e|)(\forall y_2 \le z)\cdots(Qy_i \le z)(Q'y_{i+1} \le |z|)\phi_i(e,j,x,\mathbf{y})$$

after pairing is applied. Here ϕ_i consists of a statement saying w is a tuple of the form $\langle\langle w_1, \ldots, w_{i+2}\rangle\rangle$ together with statements saying each w_i codes a postfix computation of t_i in $e = \langle\langle \ulcorner t_1\urcorner, \ldots, \ulcorner t_{i+3}\urcorner\rangle\rangle$. This amounts to checking conditions for each m

$$\begin{array}{l}
[\hat{\beta}(j, |i+12|, \ulcorner t_m\urcorner) = \ulcorner x\urcorner \supset \hat{\beta}(j,|z|,w_m) = x]\wedge \\
[\hat{\beta}(j, |i+12|, \ulcorner t_m\urcorner) = \ulcorner + \urcorner \supset \\
\hat{\beta}(j,|z|,w_m) = \hat{\beta}(j \dot{-} 2, |z|, w_m) + \hat{\beta}(j \dot{-} 1, |z|, w_m)] \wedge \cdots \\
\\
\cdots \\
\\
[\hat{\beta}(j, |i+12|, \ulcorner t_m\urcorner) = \ulcorner NOP\urcorner \supset \hat{\beta}(j,|z|,w_m) = \hat{\beta}(j \dot{-} 1, |z|, w_m)].
\end{array}$$

ϕ_i also has conditions $y_m \le \hat{\beta}(|e|,|z|,w_m)\wedge$ if y_m was existentially quantified and conditions $y_m \le \hat{\beta}(|e|,|z|,w_m)\supset$ if y_m was universally quantified. Finally, ϕ_i has a condition saying $\hat{\beta}(|e|,|z|,w_{i+2}) = 0$. Since $EBASIC$ can prove simple facts about projections from pairs, it can prove by induction on the complexity of the terms in any $\hat{\Sigma}^b_i$-formula $\phi(x)$ that $U_i(e_\phi, x, 2^{|x|^{e_\phi}}) \equiv \phi(x)$. $\square$

One easy corollary of the above lemma is the following:

Corollary 30 ($i \geq 1$) *The theory $\hat{T}_2^{i,\tau}$ is finitely axiomatized provided τ is finite.*

PROOF. We can axiomatize $\hat{T}_2^{i,\tau}$ as $EBASIC{+}IND_{U_i}^{\{\ell\}}$ for $\ell \in \tau$. □

Theorem 31 *Suppose Z proves $PH\downarrow$. Then $Z = S_2$.*

PROOF. Since $Z := \cup_i \hat{T}_2^{i,\{|id|_{i+3}\}}$, if Z proves $PH\downarrow$ then $\hat{T}_2^{i,\{|id|_{i+3}\}}$ must prove U_k of Lemma 29 equivalent to a $\hat{\Pi}_k^b$-formula for some i and k. Hence, $\hat{T}_2^{i,\{|id|_{i+3}\}}$ proves $\hat{\Sigma}_k^b = \hat{\Pi}_k^b$. If $k \leq i$ then $\hat{T}_2^{i,\{|id|_{i+3}\}}$ proves $\hat{\Sigma}_m^b$-$IND^{\{|id|_{i+3}\}}$ for all m. So if we choose $m := 2i + 9$ we get $\hat{T}_2^{m,\{|id|_{i+3}\}} \subseteq \hat{T}_2^{i,\{|id|_{i+3}\}}$. Then $i+3$ applications of Theorem 24 show $S_2^i \subseteq \hat{T}_2^{i,\{|id|_{i+3}\}}$. Since $\hat{T}_2^{i,\{|id|_{i+3}\}}$ proves $\hat{\Sigma}_k^b = \hat{\Pi}_k^b$ and $k < i$, $\hat{T}_2^{i,\{|id|_{i+3}\}}$ thus contains S_2^m for every m. If $k > i$, then since $\hat{T}_2^{i,\{|id|_{i+3}\}} \subseteq \hat{T}_2^{k,\{|id|_{k+3}\}}$ (you can use Theorem 24 to see this), $\hat{T}_2^{k,\{|id|_{k+3}\}}$ proves $\hat{\Sigma}_m^b$-$IND^{\{|id|_{k+3}\}}$ for all m. We can then perform the same argument as in the first case. □

Corollary 32 ($i \geq 1$) *If S_2^i proves $PH\downarrow$ then $S_2^i \preceq_{B(\hat{\Sigma}_{i+1}^b)} S_2$. In particular, if $P^{NP}(\log) \subsetneq P^{NP}$ then S_2^1 does not prove $PH\downarrow$. Also, if $S_2^i \neq T_2^i$ then S_2^i does not prove $PH\downarrow$.*

PROOF. The second statement follows from the first since the $\hat{\Delta}_2^b$-predicates of S_2 contains P^{NP}; whereas, those of S_2^1 are exactly $P^{NP}(\log)$. So if $P^{NP} \supsetneq P^{NP}(\log)$ then S_2^1 cannot be $B(\hat{\Sigma}_2^b)$-conservative over S_2. The third statement similarly follows from the first since the $\hat{\Sigma}_i^b$-$IND^{\{id\}}$ axioms of T_2^i can be written as $\hat{\Sigma}_{i+1}^b$-formulas. For the first statement, we will argue that $S_2^i \preceq_{B(\hat{\Sigma}_{i+1}^b)} S_2^i + Z$. Given this if S_2^i proves $\hat{\Sigma}_k^b = \hat{\Pi}_k^b$ for some k then $S_2^i + Z$ also proves this. So by the same argument as Theorem 31, $S_2^i + Z = S_2$. Hence, $S_2^i \preceq_{B(\hat{\Sigma}_{i+1}^b)} S_2$. So suppose $S_2^i + Z$ proves a sequent of $\hat{\Sigma}_{i+1}^b$-formulas $\Gamma \rightarrow \Delta$. (We can reduce $B(\hat{\Sigma}_{i+1}^b)$-conservativety to this case using the same proof as Theorem 59 in Pollett [13].) Since $Z_m \subseteq Z_{m+1}$ for all m, we can assume in fact that $S_2^i + Z_m$ proves A for some fixed $m > 0$. Hence, also $S_2^i + \hat{T}_2^{m,\tau_m^Z}$ proves $\Gamma \rightarrow \Delta$. Since if $m \leq i$, $S_2^i \supseteq \hat{T}_2^{m,\tau_m^Z}$, the only hard case is when $m > i$. To see this case notice $S_2^i + \hat{T}_2^{m-1,\tau_{m-1}^Z}$ can simulate an $S_2^i + \hat{T}_2^{m,\tau_m^Z}$ proof of $\hat{\Sigma}_m^b$-sequents in the same way that $\hat{T}_2^{m-1,\tau_{m-1}^Z}$ can simulate a $\hat{T}_2^{m,\tau_m^Z}$-proof of $\hat{\Sigma}_m^b$ sequents in the proof of Theorem 24, the only new rule of inference to worry about is the $\hat{\Sigma}_i^b$-$IND^{\{|id|\}}$ inference. Let A be the induction formula in such an inference. Using a witness oracle for A we query if $A(|t|)$ holds. If it does we output the witness returned by the witness oracle. Otherwise we query $A(0)$. If $A(0)$ does not hold then we output 0 as the antecedent will be false. Finally if, neither of these cases

occurs, using $O(\log)$ queries to a Σ_i^p-oracle witness to binary search for a value such that $A(a)$ but $A(Sa)$ does not where A is the $\hat{\Sigma}_i^b$ induction formula and run the multifunction witnessing the upper sequent in this proof on this value to get a witness for the succedent in the lower sequent. This multifunction is $\hat{\Sigma}_{i+1}^b$-definable in S_2^i and using $\hat{\Sigma}_i^b$-$IND^{\{|id|\}}$ on how this multifunction is defined in [13], S_2^i can prove it has the desired properties. So $S_2^i + \hat{T}_2^{m-1,\tau_{m-1}^Z}$ can simulate the $S_2^i + \hat{T}_2^{m,\tau_m^Z}$ proof of $\Gamma \rightarrow \Delta$. If $m-1 > i$, then we continue proving a chain of such conservation results until we get to the $i = m-1$ case. For $i = m-1$ we have $S_2^i \supseteq \hat{T}_2^{m-1,\tau_{m-1}^Z}$, hence, S_2^i proves $\Gamma \rightarrow \Delta$ and so $S_2^i \preceq_{B(\hat{\Sigma}_{i+1}^b)} S_2^i + Z$. $\Box$

The third statement in the above corollary was pointed out to me by Sam Buss via Jan Krajicek. We now prove Z cannot prove the collapse of the hierarchy. Our method is based on the proof in Johannsen [7] that S_{2+}^0 does not Σ_1^b-define $\lfloor \frac{x}{3} \rfloor$.

Definition 33 *The function $\#_B(x)$ returns the number of alternations between 1 and 0 in reading the binary number x from left to right. We start the counting of this number at 1 so $\#_B(1) = 1$.*

As an example, let x be the binary number 1110011 then $\#_B(x) = 3$. Since the number of alternations in x's binary notation is always going to be less than the length of x we have the following easy lemma.

Lemma 34 *If $y \leq x$ then $\#_B(y) \leq |x|$.*

PROOF. This follows since $\#_B(y) \leq |y| \leq |x|$. $\Box$

To prove our results we study the way $\#_B(f(x_1, \ldots, x_n))$ depends on $\#_B(x_i)$ where f is in B_1^τ.

Lemma 35 *The following inequalities hold:*

(a) $\#_B(|x|) \leq ||x||$

(b) $\#_B(\lfloor \frac{1}{2}x \rfloor) \leq \#_B(x)$

(c) $\#_B(MSP(x,i)) \leq \#_B(x)$

(d) $\#_B(Sx) \leq \#_B(x) + 1 \leq 2 \cdot \#_B(x)$

(e) $\#_B(x\#y) = 2$

(f) $\#_B(x+y) \leq \frac{5}{2} \cdot (\#_B(x) + \#_B(y))$

(g) $\#_B(x \dot{-} y) \leq \frac{5}{2} \cdot (\#_B(x) + \#_B(y)) + \frac{7}{2} \leq \frac{9}{2} \cdot (\#_B(x) + \#_B(y))$

(h) $\#_B(x \cdot y) \leq (2 \cdot (\#_B(x) + \#_B(y))^{\log 5}(5 \cdot (\#_B(x) + \#_B(y)) + 8.5) \leq 56 \cdot (\#_B(x) + \#_B(y))^{1+\log 5}$

(i) $\#_B((Wi \leq |t(x)|)(f(x,z) = 0)) \leq ||t|| \leq k||x||$ *for some fixed* k.

PROOF.

(a) This follows from Lemma 34

(b) Since $\lfloor \frac{1}{2}x \rfloor$ chops off the low order bit of x the number of alternations can at most stay the same.

(c) This follows by similar reasoning to (b).

(d) If the low order bits of x is 0 then adding 1 can increase the number of alternations by at most one since only this bit will be flipped. Otherwise, adding 1 will toggle the low order block of 1's in x and carry the 1 to the 0 to its left. Again, at most increasing the number of alternations by 1.

(e) The number $x\#y$ is a 1 followed by $|x||y|$ zeros.

(f) First, notice that adding 2^i to or subtracting 2^i from x can only increase the number of blocks in x by at most 2. Since the blocks of 1's in y can be represented as expressions of the from $2^{j+i} - 2^i$, when we perform the addition we get at most 4 blocks in the new number for every block of 1's in y. So the new number has fewer than $\#_B(x) + 4\#_B(y)$ blocks. By symmetry it also has less than $\#_B(y) + 4\#_B(x)$ blocks. The minimum of these two values is thus bounded by their average.

(g) This follows from (f) since if $x \geq y$ then

$$x \dot{-} y = (2^{|x|+1} - 1 - ((2^{|x|+1} - 1 - x) + y))$$

and $2^{|x|+1} - 1 - x$ has at most one more block then x and

$$(2^{|x|+1} - 1 - ((2^{|x|+1} - 1 - x) + y))$$

has at most one more block then $(2^{|x|+1} - 1 - x) + y$.

(h) Consider multiplying x by a block of 1's $2^{i+j} - 2^i$. This gives $x \cdot 2^{i+j} - x2^i$ which is the subtraction of two number each with at most one more alternation than x. So we get less than $5 \cdot \#_B(x) + 8.5$ blocks by (g). There are $\lceil \frac{\#_B(y)}{2} \rceil$ blocks of 1's in y. So to compute $\#_B(x \cdot y)$ we need to add together fewer than $\#_B(y)$ numbers with fewer than $5 \cdot \#_B(x) + 8.5$ blocks. If we do this in a balanced fashion then by (f) we get fewer than $5^{|\#_B(y)|}(5 \cdot \#_B(x) + 8.5)$ blocks. This is less than $(2\#_B(y))^{\log 5}(5 \cdot \#_B(x) + 8.5)$ and in turn less than $(2(\#_B(x) + \#_B(y))^{\log 5}(5 \cdot (\#_B(x) + \#_B(y)) + 8.5)$.

(i) Follows from (a).

□

As an aside, it would be interesting to get a better bound on the number of blocks produced by multiplication. One can construct examples where $\#_B(x \cdot y)$ is $\Omega((\#_B(x) + \#_B(y))^2)$. For example, if one multiplies $(11)_2$ and $(111)_2$ one gets $(10101)_2$. Take two sequences $\{a_i\}$,$\{b_i\}$, such that a_i has i blocks of 11's alternating with blocks of 0's (these blocks of 0's increase in size with i) and b_i has i blocks of 111's alternating with blocks of 0's and such that $a_i \cdot b_i$ is of the form $(1010101\cdots)_2$.

We now bound the number of blocks that can be produced by recursion. For this define: $L_B(x) := \#_B(x) + ||x||$.

Lemma 36 *Let τ be a set of nondecreasing iterms all of which are $O(||x||)$. Assume τ has at least one unbounded term and let f be defined by BPR^τ using g, h,r, ℓ, t satisfying*

$$\begin{aligned}
\#_B(g(\mathbf{x})) &\leq c \cdot (\sum_{i=0}^{m} L_B(x_i))^{4^{\ell_1(s_1(\mathbf{x}))}} \\
\#_B(r(n,\mathbf{x})) &\leq c \cdot (L_B(n) + \sum_{i=0}^{m} (L_B(x_i)))^k \\
\#_B(h(n,\mathbf{x},y)) &\leq c \cdot (L_B(n) + (\sum_{i=0}^{m} L_B(x_i)) + L_B(y))^{4^{\ell_2(s(n,y,\mathbf{x}))}}
\end{aligned}$$

where $\ell_1, \ell_2 \in \dot{\tau}$, $s, s_1 \in L_2$ and c, k are constants. Then there is an $\ell_3 \in \dot{\tau}$, a term $t' \in L_2$ such that

$$\#_B(f(n,\mathbf{x})) \leq c \cdot (L_B(n) + \sum_{i=0}^{m} L_B(x_i))^{4^{\ell_3(t'(\mathbf{n},\mathbf{x}))}}$$

PROOF. Note we do not lose any generality in assuming the constant c is the same in the bound of each g, h, and r since if they differed we could always take the maximum of the three values. We know by the bound on h that

$$\#_B(F(n+1,\mathbf{x})) \leq c \cdot (L_B(n) + L_B(F(n,\mathbf{x})) + \sum_{i=0}^{m} L_B(x_i))^{4^{\ell'(s(n,F(n,\mathbf{x}),\mathbf{x}))}}.$$

Notice by the definition of BPR^τ, $F(n,\mathbf{x}) \leq r(n,\mathbf{x})$. So also $||F|| \leq ||r||$. Since $r \in L_2$, there is a constant k' such that

$$||r(n,\mathbf{x})|| \leq k' \cdot (||n|| + \sum_{i=0}^{m} ||x_i||) \leq k' \cdot (L_B(n) + \sum_{i=0}^{m} L_B(x_i))$$

Let $k := k' + 1$. Thus, $L_B(F(n+1, \mathbf{x}))$ is less than

$$c \cdot (k \cdot (L_B(n) + \sum_{i=0}^{m} L_B(x_i)) + \#_B(F(n, \mathbf{x})))^{4^{\ell'(s(n,F(n,\mathbf{x}),\mathbf{x}))}}$$

Using our bound on $\#_B(h)$, we can then expand $\#_B(F(n, \mathbf{x}))$ and so on. We can thus bound $\#_B(f(n, \mathbf{x})) = F(\ell(t(n, \mathbf{x})), \mathbf{x})$ by $Y :=$

$$c \cdot (\sum_{i=0}^{m} L_B(x_i) + \sum_{j=0}^{\ell(t)} k \cdot (L_B(j) + \sum_{i=0}^{m} L_B(x_i)))^{4^{\psi}}$$

where ψ is

$$\ell_1(s_1(\mathbf{x})) + \sum_{j=0}^{\ell(t)} \ell_2(s(j, F(j, \mathbf{x}), \mathbf{x})$$

Since ℓ and ℓ_2 are nondecreasing and $F \leq r$ we can bound ψ by $\psi' :=$

$$\ell_1(s_1(\mathbf{x})) + \ell(t^+(n, \mathbf{x})) \cdot \ell_2(s^+(\ell(t^+(n, \mathbf{x})), r^+(\ell(t^+(n, \mathbf{x})), \mathbf{x}), \mathbf{x}).$$

Let ℓ' be an unbounded term in τ and let v be a fixed number such that $\ell'(v) > 1$ then ψ' is bounded by

$$\ell'(v) \cdot \ell_1(s_1(\mathbf{x})) \cdot \ell(t^+(n, \mathbf{x})) \cdot \ell_2(s^+(\ell(t^+(n, \mathbf{x})), r^+(\ell(t^+(n, \mathbf{x})), \mathbf{x}), \mathbf{x})$$

Since this term is just a product of terms in τ, it is bounded by some term $\ell_4(t'')$ where $\ell_4 \in \dot{\tau}$. Now consider the term W under the exponent in Y. Since ℓ is nondecreasing and $O(||x||)$, W can be bounded by

$$\sum_{i=0}^{m} L_B(x_i) + \ell(t^+(n, \mathbf{x})) \cdot k \cdot (2 \cdot ||n|| + \sum_{i=0}^{m} L_B(x_i))$$

which is less than

$$(2k+1)(L_B(n) + \sum_{i=0}^{m} L_B(x_i))^2 \leq (L_B(n) + \sum_{i=0}^{m} L_B(x_i))^{4k+2}.$$

So $\#_B(f(n, \mathbf{x}))$ is bounded by

$$c \cdot (L_B(n) + \sum_{i=0}^{m} L_B(x_i))^{(4k+2) \cdot 4^{\ell_4(t'')}}.$$

Using the unbounded term ℓ' again we can get an $\ell_3 \in \dot{\tau}$ and a t' such that $(4k+2) \cdot 4^{\ell_4(t'')} \leq 4^{\ell_3(t')}$ and thus prove the theorem. $\square$

Lemma 37 *Let τ be a set of nondecreasing iterms all of which are $O(||x||)$. Assume τ has an unbounded term. If $f(\mathbf{x}) \in B_1^\tau$ and $\#_B(x_i)) \leq ||x_i||$ then $\#_B(f(\mathbf{x})) \leq c \cdot (||x_1|| + \cdots + ||x_n||)^{4^{\ell(\mathbf{t}(\mathbf{x}))}}$, $t \in L_2$ and $\ell \in \dot{\tau}$.*

PROOF. This follows from Lemma 35 and Lemma 36 and by noticing $1 + \log 5 < 4$. $\square$

Theorem 38 *The function $\lfloor \frac{x}{3} \rfloor$ is not $\hat{\Sigma}^b_1$-definable in Z. Hence, Z cannot prove the polynomial hierarchy collapses.*

PROOF. By Lemma 28 and Corollary 23, the $\hat{\Sigma}^b_1$-functions of Z are contained in B^Z_1. Notice all the terms in $\dot{\tau}^Z_1$ are $o(|x|_3)$. Consider $y := 2^{|x|+1} - 1$ for any x. $\#_B(y) = 1$, yet $\lfloor \frac{2^{|x|+1}-1}{3} \rfloor$ is a number of length $|x| - 1$ of the form $1010\cdots$. Hence, $\#_B(\lfloor y \rfloor) = |x| - 1 > \frac{1}{4} \cdot 2^{2^{|y|_3}}$ which is greater than

$$c \cdot ||y||^{4^{p(\ell(y))}} \leq c \cdot 2^{2^{2 \cdot |y|_4 \cdot p(\ell(y))}}$$

for fixed p, c, and for $\ell \in \tau$ and large enough x. This is since $2 \cdot |y|_4 \cdot p(\ell(y))$ can be majorized by a term in $\dot{\tau}^Z_1$ and as we have already observed all these terms are $o(|x|_3)$. So by Lemma 37, $\lfloor \frac{x}{3} \rfloor$ is not in $B^{\{|id|_4\}}_1$ and hence not in B^Z_1. On the other hand, $\lfloor \frac{x}{3} \rfloor \in FP$ and by Theorem 31 if Z proves $PH \downarrow$ then $Z = S_2$. In which case, $FP \subseteq B^Z_1$ since the $\hat{\Sigma}^b_1$-definable functions of S^1_2 are FP (from Buss [3] and using the fact that S^1_2 can prove every Σ^b_1-formula equivalent to a $\hat{\Sigma}^b_1$-formula) and $S^1_2 \subseteq S_2$. This is a contradiction since B^Z_1 does not contain $\lfloor \frac{x}{3} \rfloor$. So Z does not prove $PH \downarrow$. $\square$

Remark 39 *At this point, in the spirit of Razborov's work on what fragments can formalize which lower bounds techniques, we should examine the complexity of the lower bounds proof just presented. The function $\#_B(x)$ is polynomial time computable, and hence, Σ^b_1-definable in S^1_2. The theory S^1_2 can also prove appropriate roundings on all of the inequalities in Lemma 35. For the bounds on $\#_B(x + y)$ and $\#_B(x \cdot y)$ one would fix y and perform induction on the number given by the first i blocks of x. Then reverse the roles of x and y and reason as we did above to get the bound. Lemma 36 can also be proven by $IND^{\{|id|\}}$. Hence, by induction on the complexity on any term in B^Z_1, S^1_2 can prove B^Z_1 cannot does not contain $\lfloor \frac{x}{3} \rfloor$.*

Corollary 40 *The theory $T := Z + \{\hat{\Pi}^b_0$-consequences of $S_2\}$ cannot prove the polynomial hierarchy collapses.*

PROOF. Let π denote the $\hat{\Pi}^b_0$-consequences of S_2. We claim the $\hat{\Sigma}^b_1$-definable functions of $Z + \pi$ are still in B^Z_1, and so the argument above also implies $Z + \pi$ cannot prove $PH \downarrow$. Since $Wit^{i+1}_A := w = 0 \wedge A$ for any A in π, we can choose the zero function to witness A and use the same proof as the proof of Theorem 24 to show $\hat{T}^{i,2^{\dot{\tau}}}_2 + \pi \preceq_{B(\hat{\Sigma}^b_{i+1})} \hat{T}^{i+1,\tau}_2 + \pi$. By the same reasoning as Lemma 28 we get $\hat{T}^{i,\tau^Z_i}_2 + \pi \preceq_{B(\hat{\Sigma}^b_{i+1})} Z + \pi$. So $\hat{T}^{1,\tau^Z_1}_2 + \pi \preceq_{B(\hat{\Sigma}^b_2)} Z + \pi$.

Since $Wit^1_A := w = 0 \wedge A$, essentially the same proofs of Theorem 22 and Corollary 23 show the $\hat{\Sigma}^b_1$-definable multifunctions of $\hat{T}^{1,\tau^Z_1}_2 + \pi$ and hence $Z + \pi$ are B^Z_1. $\square$

Theorem 41 *If Z proves $PH \uparrow$ then for all i, $S^i_2 \vdash \Sigma^p_i \neq \Pi^p_i$.*

PROOF. For Z to show $PH \uparrow$, Z must show $\Sigma^p_i \neq \Pi^p_i$ for $i > 0$. We take this to mean there is a $\hat{\Pi}^b_i$-formula $A(x)$ such that for each integer k, the theory Z proves the statement $\forall e \exists x \neg(A(x) \Leftrightarrow U_i(e, x, 2^{|x|^k})$. These statements are equivalent in Z to $\hat{\Sigma}^b_{i+1}$-formulas. Hence, by Lemma 28 they are provable in $\hat{T}^{i,\tau^Z_i}_2$. But $\hat{T}^{i,\tau^Z_i}_2 \subseteq S^i_2$ since terms in τ^Z_i are surpassed by $|id|_3$. So $\hat{T}^{i,\{|id|_3\}}_2 \subseteq S^i_2$ proves these statements and, thus, that $\Sigma^p_i \neq \Pi^p_i$. $\square$

5 Acknowledgements

We would like to thank Sam Buss and Jan Johannsen for constructive comments on an earlier version of this paper.

References

[1] C.H. Bennett and J. Gill. Relative to random oracle A, $P^A \neq NP^A \neq co\text{-}NP^A$ with probability 1. *SIAM Journal of Computing*, 10:96–113, 1981.

[2] M. Bonet, T. Pitassi, and R. Raz. No feasible interpolation for TC^0-frege proofs. In *Foundations of Computer Science*, volume 38, pages 254–263, 1997.

[3] S.R. Buss. *Bounded Arithmetic.* Bibliopolis, Napoli, 1986.

[4] S.R. Buss. Bounded arithmetic, complexity and cryptography. *To appear Theoria*, 1998.

[5] R. Chang and J. Kadin. The boolean hierarchy and the polynomial hierarchy: a closer connection. In *Proceedings Fifth Annual Structures in Complexity Conference*, pages 169–178, 1990.

[6] J. Hastad. Almost optimal lower bounds for small depth circuits. In *Proceedings of the Eighteenth Annual ACM Symposium on theory of Computing*, pages 6–20, 1987.

[7] J. Johannsen. On the weakness of sharply bounded polynomial induction. In *Proceedings of Gödel 1993*, pages 223–230. Springer-Verlag, 1993.

[8] J. Johannsen. A model-theoretic property of sharply bounded formula with some applications. *Mathematical Logic Quarterly*, 44(2):205–215, 1998.

[9] J. Kadin. The polynomial time hierarchy collapses if the boolean hierarchy collapses. *SIAM Journal on Computing*, 17(6):1263–1282, August 1988.

[10] C. F. Kent and B.R. Hodgson. An arithmetical characterization of np. *Theoretical Computer Science*, 21:255–267, 1982.

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

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

[13] C. Pollett. Structure and definability in general bounded arithmetic theories. To appear Annals of Pure and Applied Logic.

[14] C. Pollett. *Arithmetic Theories with Prenex Normal Form Induction.* PhD thesis, University of California, San Diego, 1997.

[15] P. Pudlak. Ramsey's theorem in bounded arithmetic. In *Computer Science Logic '90, LNCS533*, pages 308–317. Springer-Verlag, 1990.

[16] A.A. Razborov. Bounded arithmetic and lower bounds in Boolean complexity. In P. Clote and J. Remmel, editors, *Feasible Mathematics II*, pages 344–386. Birkhauser, 1995.

[17] A.A. Razborov. Lower bounds for propositional proofs and independence results in bounded arithmetic. In *Proceedings of 20th International Symposium on the Mathematical Foundations of Computer Science*, page 105. Springer-Verlag, 1995.

[18] G. Takeuti. Sharply bounded arithmetic and the function $a \dot{-} 1$. In *volume 106 of Contemporary Mathematics*, Logic and Computation, pages 281–288. American Mathematical Society, 1990.

[19] G. Takeuti. *RSUV* isomorphisms. In P. Clote and J. Krajíček, editors, *Arithmetic, Proof Theory and Computational Complexity*, pages 364–386. Oxford Science Publications, 1993.