Extensions of models of PV

Jan Krajíček*

Mathematical Institute and Institute of Computer Science
Academy of Sciences of the Czech Republic

Abstract

We prove that certain models of PV in which $NP \not\subseteq P/poly$ have a Π_1^b-elementary extension to a model of (PV and) $NP \not\subseteq coNP/poly$.

If S_2 proves a particular fact about bipartite graphs then, in fact, all models of PV in which $NP \not\subseteq P/poly$ have a Π_1^b-elementary extension to a model of $NP \not\subseteq coNP/poly$.

Introduction

PV is a bounded arithmetic theory with function symbols for all polynomial time algorithms, and axiomatized by a particular set of universal formulas, cf. [3]. Models of PV are a natural environment for notions of computational complexity theory around deterministic and non-deterministic polynomial time. Major open problems in this part of complexity theory have their counterparts in bounded arithmetic and propositional logic. We are interested in proving some notorious open conjectures for a model of bounded arithmetic, and not so much in showing that some of these conjectures might be unprovable in bounded arithmetic. For a general motivation (for this author, at least) for research in this area see the preface to [4].

In a model M of the theory PV the class P of the polynomial-time sets is the class of subsets of M definable by an atomic PV-formula with parameters from M (in S_2^1 this would be provably Δ_1^b-formulas with parameters), equivalently: recognizable by a *standard* DTM with an extra input (the parameter) which may be non-standard, equivalently: recognizable by a DTM possibly with a non-standard description but whose time is bounded by a standard degree polynomial.

The class $P/poly$ is defined in the same way except that the parameters may vary with the length of the inputs, and the classes $NP, NP/poly$ and

*Partially supported by the *US - Czechoslovak Science and Technology Program* grant # 93025, and by grant #$A1019602$ of the *AVČR*.

$coNP, coNP/poly$ are defined analogously using $NDTM$'s. In particular, NP-subsets of M (resp. $coNP$) are those definable by Σ_1^b-formulas (resp. by Π_1^b-formulas) with parameters, that may vary with the length in case of $NP/poly$ and $coNP/poly$.

It is not important whether we require that the length of parameters in the non-uniform classes is polynomial in the length of the input. This is because we are concerned with definability of sets of inputs of a fixed length. In general one may restrict to those models of PV in which lengths are polynomial (with a standard degree) in one fixed length.

The problem whether PV equals to S_2^1 is closely related to the circuit complexity of NP-sets. In particular, $PV \neq S_2^1$ if $NP \not\subseteq P/poly$ (by [8]) or if there is a model of PV in which $NP \not\subseteq coNP/poly$ (by [2, 9]).[1]

Constructions of extensions of models of PV (or of S_2^1) are also closely related to length-of-proofs problems about the extended Frege systems, cf. [4, 5, 6].

In this paper we study the problem to construct a model of PV in which $NP \not\subseteq coNP/poly$. We give three versions of a construction showing that certain models of PV in which $NP \not\subseteq P/poly$ have a Π_1^b-elementary extension to a model of (PV and) $NP \not\subseteq coNP/poly$. An ultimate goal is to make the construction work under weaker assumptions on models than those in Theorem 2.1.

A relevant background can be found in [4]. In particular, necessary facts from all other references can be also found there.

1 Preliminaries

Given a length $n = |y|$ of $y \in M$, $SAT_n(M)$ denotes the set of satisfiable formulas in M of length n; this set is defined by a canonical Σ_1^b-formula $Sat_n(x)$ with a parameter of the same length as y. $Log(M)$ is the set of lengths of elements of M.

For a formula a and a truth assignment w the relation $w \models a$ denotes that w *satisfies* a, and is definable by a fixed open formula. We shall assume that $w \models a$ implies (in PV) that a is a formula from $SAT_n(M)$ and w is a truth-assignment to its atoms.

Let $Circuit_M$ denote the set of multi-output circuits in M and for $C \in Circuit_M$ and $a \in M$ of appropriate length, $C(a) = b$ is a function definable by a ternary PV-symbol stating that b is the output of the computation of circuit C on input a (when numbers are identified with their binary encodings).

The following lemma follows from the fact that PV can define binary search.

[1] We inessentially abuse the notation here; instead of PV, which is an equational theory as defined in [3], we work with its first-order conservative extension PV_1 defined in [8, 4], and in place of S_2^1 we should use its conservative extension $S_2^1(PV)$ in the language of PV, cf. [4].

Lemma 1.1 *For any length n and any circuit $C \in Circuit_M$ there exists another circuit $C' \in Circuit_M$ such that if*

$$M \models (\forall x, |x| = n), Sat_n(x) \equiv (C(x) = 1)$$

then

$$M \models (\forall x, |x| = n), Sat_n(x) \to (C'(x) \models x) \ .$$

In particular, the property that $SAT_n(M)$ is recognized in a model M of PV by a circuit is preserved to Π_1^b-elementary extensions of M.

This means that only M in which $NP \not\subseteq P/poly$ can possibly have a cofinal Π_1^b-elementary extension in which $NP \not\subseteq coNP/poly$.

Definition 1.2 *Let M be a model of PV and assume that for some length $n \in Log(M)$ the set $SAT_n(M)$ is not recognized in M by a circuit.*

A counter-example *function (for $SAT_n(M)$ in M, tacitly) is a function ξ that assigns to any circuit $C \in Circuit_M$ with n inputs a pair*

$$\xi(C) = (a, w)$$

such that

1. $w \models a$
2. $C(a) \not\models a$.

We say that ξ is in $P/poly$ of M if for every length $m \in Log(M)$ there is a circuit $D_m \in Circuit_M$ with m input bits and $2n$ output bits computing $\xi(C)$ for any C of size at most m.

Note that the statement that $SAT_n(M)$ is not recognized by a circuit of size at most m is $\Pi_1^b(\xi)$, whenever ξ is a counter-example function. Hence we have the following lemma.

Lemma 1.3 *Let M be a model of PV in which the set $SAT_n(M)$ is not recognized by a circuit. Let ξ be a corresponding counter-example function. Then $SAT_n(M')$ is not recognized by a circuit in any $\Pi_1^b(\xi)$-elementary, cofinal extension (M', ξ') of (M, ξ).*

In particular, if M admits a counter-example function in $P/poly$ then the set $SAT_n(M')$ is not recognized by a circuit in any Π_1^b-elementary, cofinal extension M' of M, and M' admits a counter-example function in $P/poly$.

2 An ultrapower

Theorem 2.1 *Let M be a countable model of PV and assume that for some length $n \in Log(M)$ the set $SAT_n(M)$ is not recognized in M by a circuit. Assume that M admits a counter-example function ξ in $P/poly$.*

Then there is a Π_1^b-elementary, cofinal extension M' of M, a model of PV, such that the set $SAT_n(M')$ is not recognized in M' by a co-non-deterministic circuit.

Proof

For $C \in Circuit_M$ let f_C be the function from M to M computed by the circuit C.

Take:

$$\mathcal{F}_M := \{f_C : SAT_n(M) \to M \mid C \in Circuit_M\}.$$

We shall construct an ultrapower of the form $\mathcal{F}_M/\mathcal{U}$, with $\mathcal{U} \subseteq exp(SAT_n(M))$ a particular ultrafilter. The following claim is obvious.

Claim 1 *For any ultrafilter $\mathcal{U}$, Loš's theorem holds for all open PV- formulas, and $\mathcal{F}_M/\mathcal{U}$ is a Π_1^b-elementary, cofinal extension of M. In particular, $\mathcal{F}_M/\mathcal{U}$ is a model of PV.*

Define a particular element of $\mathcal{F}_M$:

$$a_{\mathcal{U}} = id_{SAT_n(M)}/\mathcal{U}.$$

Claim 2 *Let $\psi(x)$ be a Π_1^b-formula with parameters from M such that:*

$$M \models \psi(a)$$

for all $a \in SAT_n(M)$. Then:

$$\mathcal{F}_M/\mathcal{U} \models \psi(a_{\mathcal{U}}).$$

Claim 2 follows by Loš's theorem for all open PV- formulas.

For a circuit $D \in Circuit_M$ with n bits of input define the set:

$$D^* := \{a \in SAT_n(M) \mid D(a) \models a\} \ .$$

Claim 3 *Assume that an ultrafilter $\mathcal{U} \subseteq exp(SAT_n(M))$ satisfies the condition:*

$$\forall D \in Circuit_M; D^* \notin \mathcal{U}$$

Then :

$$\mathcal{F}_M/\mathcal{U} \models \neg Sat_n(a_{\mathcal{U}}) \ .$$

The claim follows from Loš's theorem again: an element $f_D/\mathcal{U}$ satisfies the formula $a_{\mathcal{U}}$ in $\mathcal{F}_M/\mathcal{U}$ iff $D^* \in \mathcal{U}$.

Claim 4 *$SAT_n(\mathcal{F}_M/\mathcal{U})$ is not recognized in $\mathcal{F}_M/\mathcal{U}$ by a circuit and $\mathcal{F}_M/\mathcal{U}$ admits a counter-example function in $P/poly$.*

Assume on the contrary that $SAT_n(\mathcal{F}_M/\mathcal{U})$ is recognized in $\mathcal{F}_M/\mathcal{U}$ by a circuit, hence by Lemma 1.1 it holds in $\mathcal{F}_M/\mathcal{U}$:

$$f_W/\mathcal{U} \models f_A/\mathcal{U} \Rightarrow f_C/\mathcal{U}(f_A/\mathcal{U}) \models f_A/\mathcal{U}$$

for some $f_C/\mathcal{U} \in Circuit_{\mathcal{F}_M/\mathcal{U}}$ and all $f_A/\mathcal{U}, f_W/\mathcal{U} \in \mathcal{F}_M$.

For an arbitrary f_C define particular f_A, f_W by:

$$(f_A(a), f_W(a)) \ := \ \xi(C(a))$$

For those $a \in SAT_n(M)$ for which $C(a)$ is a circuit with n inputs, $f_W(a) \models f_A(a)$ but $C(a)(f_A(a)) \not\models f_A(a)$ by the definition of ξ. Hence $f_C/\mathcal{U}$ cannot have the property stated earlier.

Note that by Claim 1 the circuits D_m computing ξ in M compute a counter-example function in $\mathcal{F}_M/\mathcal{U}$ as well.

Let $\mathcal{U}_0 \subseteq exp(SAT_n(M))$ consist of all sets X containing some set of the form:

$$SAT_n(M) \setminus D^*$$

for some $D \in Circuit_M$. By the hypothesis that $SAT_n(M)$ is not recognized in M by a circuit, the class $\mathcal{U}_0$ is closed under intersections and $\emptyset \notin \mathcal{U}_0$, i.e., it is a non-trivial filter. Let $\mathcal{U} \supseteq \mathcal{U}_0$ be arbitrary ultrafilter.

Define M^1 to be the countable model $\mathcal{F}_M/\mathcal{U}$. By Claims 2 and 3 no Π^b_1-formula with parameters from M defines the set $SAT_n(M^1)$ in M^1.

By Claim 4 the set $SAT_n(M^1)$ is not recognized in M^1 by a circuit. We may therefore repeat this construction countably many times to obtain a chain:

$$M \subseteq M^1 \subseteq M^2 \subseteq \dots$$

of Π^b_1-elementary, cofinal extensions (killing all potential Π^b_1-definitions of $Sat_n(x)$ with all possible parameters from all M^t) such that its union:

$$M' := \bigcup_t M^t$$

is a Π^b_1-elementary, cofinal extension of M in which $SAT_n(M')$ is not defined by any Π^b_1-formula with parameters from M', i.e., it is not recognized by a co-non-deterministic circuit.

Q.E.D.

Note that the version of the theorem with $P, NP, coNP$ in place of the non-uniform classes is a simple corollary of Herbrand's theorem.

3 A compactness argument

In this section we give another proof of Theorem 2.1.

Let $\pi(x)$ be a Π^b_1 - formula with parameters from M. We want to find a Π^b_1-elementary, cofinal extension of M in which $\exists x; \neg(\pi(x) \equiv Sat_n(x))$ holds. Note that we may assume w.l.o.g. that in $PV + Th_{\forall\Pi^b_1}(M)$ it holds that

$$\pi(c) \rightarrow |c| = n$$

(otherwise just replace $\pi(c)$ by $\pi(c) \wedge |c| = n$).

If already

$$M \models \exists x; \neg(\pi(x) \equiv Sat_n(x))$$

then this will be preserved in every Π_1^b -elementary extension. If

$$PV + Th_{\forall\Pi_1^b}(M) \vdash \forall x(\pi(x) \equiv Sat_n(x))$$

then by Herbrand's theorem there is a PV-symbol $f(x, y)$ and $b \in M$ such that:

$$PV + Th_{\forall\Pi_1^b}(M) \vdash \forall x(Sat_n(x) \equiv (f(x, b) \models x)) \ ,$$

so the set $SAT_n(M)$ is recognized in M by a circuit, contradicting the hypothesis of the theorem.

So the only case creating difficulties is when

$$M \models \forall x \ (\pi(x) \equiv Sat_n(x))$$

but

$$PV + Th_{\forall\Pi_1^b}(M) \not\vdash \forall x(\pi(x) \equiv Sat_n(x))$$

which implies:

$$PV + Th_{\forall\Pi_1^b}(M) \not\vdash \forall x \ (\pi(x) \to Sat_n(x))$$

(as the opposite implication is in $Th_{\forall\Pi_1^b}(M)$).

Take a new constant c and a formula

$$\pi(c) \wedge \neg Sat_n(c) \ .$$

Claim *The theory*

$$PV + Th_{\forall\Pi_1^b}(M) + \pi(c) \wedge \neg Sat_n(c)$$

does not prove that $Sat_n(x)$ is recognized by a polynomial size circuit.

Assume on the contrary that

$$PV + Th_{\forall\Pi_1^b}(M) + \pi(c) + \neg Sat_n(c) \vdash \exists D(\forall x, |x| = n); \ Sat_n(x) \to D(x) \models x$$

By the hypothesis $PV + Th_{\forall\Pi_1^b}(M) + \pi(c) + \neg Sat_n(c)$ is consistent and hence has a model N (that contains M as a submodel). Take N^* to be the unique substructure of N generated from elements of $M \cup \{c\}$ by PV-function symbols. Thus $N^* \models PV + Th_{\forall\Pi_1^b}(M) + \pi(c) + \neg Sat_n(c)$ and hence

$$N^* \models \exists D(\forall x, |x| = n); \ Sat_n(x) \to D(x) \models x$$

Moreover, N^* is a Π_1^b-elementary and cofinal (as $|c| = n$) extension of M.

However, that is a contradiction with Lemma 1.3, as by the hypothesis of the theorem M admits a counter-example function in $P/poly$.

By the claim we may take M^1, a Π^b_1-elementary, cofinal extension of M that is a model of $\pi(c) \wedge \neg Sat_n(c)$, and such that there is no circuit in M^1 recognizing $SAT_n(M^1)$. Then we construct a countable chain $M \subseteq M^1 \subseteq M^2 \subseteq \ldots$ killing all potential Π^b_1-definitions (with all possible parameters from all M^i) of $Sat_n(x)$. Thus $M' := \bigcup_i M^i$ is the required extension.

Q.E.D.

4 A Boolean-valued extension

Boolean-valued extensions of S^1_2 were defined in [5], see also [4, Chpt. 9.4]. For PV in place of S^1_2 the construction has a particular formulation.

Let M be a model of PV and let $(p_1, \ldots, p_n) \in M$ be a sequence of propositional atoms. Let $Circuit_M(\overline{p})$ be all circuits with one output formed from atoms p_i, and let $\mathbf{B}(\overline{p})$ be the Boolean algebra obtained by factoring $Circuit_M(\overline{p})$ by the equivalence relation $C_1 \sim C_2$ that holds for C_1, C_2 iff there is an EF-proof in M of $C_1 \equiv C_2$ (see [5] for a formalization of this notion).

Given an ultrafilter $\mathcal{G}$ on $\mathbf{B}(\overline{p})$, let $\nu_{\mathcal{G}}(C)$ be equal 1 if $(C/\sim) \in \mathcal{G}$ and equal to 0 otherwise.

Define the extension $M[\mathcal{G}]$ of M as follows. Let $Names_M(\overline{p})$ be the set of sequences $\langle C_1, \ldots, C_\ell \rangle \in M$ of elements of $Circuit_M(\overline{p})$. The elements of $M[\mathcal{G}]$ are tuples

$$\langle \nu_{\mathcal{G}}(C_1), \ldots, \nu_{\mathcal{G}}(C_\ell) \rangle$$

one for each $\langle C_1, \ldots, C_\ell \rangle \in Names_M(\overline{p})$.

For $f(x_1, \ldots, x_k)$ a PV-function and $\ell \in Log(M)$ a length, let $D^t_{f,\ell}(y_{ij})$ ($i \leq k$ and $j \leq \ell$) be a circuit in M computing (provably in PV) the t^{th} bit of $f(x_1, \ldots, x_k)$ for inputs x_i of length at most ℓ with bits $y_{i1}, \ldots, y_{i\ell}$. Define $f(w_1, \ldots, w_k)$ for elements w_i of $M[\mathcal{G}]$

$$w_i = \langle \nu_{\mathcal{G}}(C_{i1}), \ldots, \nu_{\mathcal{G}}(C_{i\ell}) \rangle$$

to be

$$\langle \nu_{\mathcal{G}}(D^1_{f,\ell}(y_{ij}/C_{ij}), \nu_{\mathcal{G}}(D^2_{f,\ell}(y_{ij}/C_{ij}), \ldots \rangle$$

The following is a special case of [5, Thm. 5.1]. See also [7] or [5, Sec. 9.4] for another treatment of the construction.

Theorem 4.1 *Let M be a model of PV, $(p_1, \ldots, p_n) \in M$ propositional atoms, and let $\mathcal{G}$ be an ultrafilter on $\mathbf{B}(\overline{p})$. Assume that $\mathcal{G}$ is closed under EF-provability in M, i.e., whenever there is an EF-proof in M of D from $C_1, \ldots, C_k$ and $\nu_{\mathcal{G}}(C_i) = 1$ then $\nu_{\mathcal{G}}(D) = 1$ too.*

Then $M[\mathcal{G}]$ is a cofinal extension of M and it is a model of PV.

Moreover, if $\nu_{\mathcal{G}}(C) = 1$ whenever $C \in Circuit_M(\overline{p})$ computes the function constantly 1 in M, then $M[\mathcal{G}]$ is a Π^b_1-elementary, cofinal extension of M.

We give now another proof of Theorem 2.1.

Let M be a countable model of PV in which $SAT_n(M)$ is not recognized by a circuit, and that admits a counter-example function ξ in $P/poly$.

We shall denote by $y \models x$ also the circuit in M that computes on two n-bit inputs x, y whether they satisfy the relation $y \models x$. Let $\phi(x)$ be a Π_1^b-formula with parameters from M of the form $\forall z, |z| \leq |x|^k \rightarrow \phi_0(x, z)$, where ϕ_0 is open.

Let $\overline{p} = (p_1, \ldots, p_n)$ be mutually different propositional atoms in M. Consider the set T of propositional formulas of the form

$$\neg(\langle W_1, \ldots, W_n\rangle \models \overline{p})$$

and of the form

$$\phi_0(\overline{p}, \langle Z_1, \ldots, Z_m\rangle)$$

where $\overline{W} = \langle W_1, \ldots, W_n\rangle$, $\overline{Z} = \langle Z_1, \ldots, Z_m\rangle$ are all elements of $Names_M(\overline{p})$ of the length n and $m = n^k$ respectively.

Claim 1 *There is no EF-refutation of T in M.*

Assume otherwise, i.e., there is an EF-proof of

$$\bigvee_{\overline{W}} \overline{W}(\overline{p}) \models \overline{p} \quad \vee \quad \bigvee_{\overline{Z}} \neg\phi_0(\overline{p}, \overline{Z})$$

for some $\overline{W}$'s and $\overline{Z}$'s. As EF is sound in any model of PV, the $\overline{W}$'s and $\overline{Z}$'s may be combined into a circuit in M recognizing the set $SAT_n(M)$. That is a contradiction.

Claim 2 *There is an ultrafilter $\mathcal{G}$ on $\mathbf{B}(\overline{p})$ that is closed under EF-provability in M and such that*

1. $\nu_{\mathcal{G}}(C) = 1$*, for all* $C \in T$.
2. $\nu_{\mathcal{G}}(C) = 1$*, for all* $C \in Circuit_M(\overline{p})$ *computing in M constantly* 1.

Take $S \subseteq Circuit_M(\overline{p})$ the set of all circuits C' majorizing (as Boolean functions) in M some $C \in T$. By Claim 1 the subset of $\mathbf{B}(\overline{p})$ of $\sim$-classes of all $C' \in S$ is a non-trivial filter. Any ultrafilter extending this set satisfies the requirements of the claim.

Take $M^1 := M[\mathcal{G}]$ for any $\mathcal{G}$ given by Claim 2. Then, by Theorem 4.1, M^1 is a model of PV in which the element

$$a_{\mathcal{G}} := \langle \nu_{\mathcal{G}}(p_1), \ldots, \nu_{\mathcal{G}}(p_n)\rangle$$

is not in $SAT_n(M^1)$ but

$$M^1 \models \phi(a_{\mathcal{G}})$$

Hence $\phi(x)$ will not define $Sat_n(x)$ in any Π_1^b-elementary extension of M^1.

By the Π_1^b-elementarity and cofinality of M^1 over M and by Lemma 1.3, no circuit in M^1 recognizes $SAT_n(M^1)$ and M^1 admits a counter-example function in $P/poly$. We may thus repeat the construction to produce a chain $M \subseteq M^1 \subseteq M^2 \subseteq \ldots$ such that $M' := \bigcup_i M^i$ is the required model, identically as in sections 2 and 3.

Q.E.D.

5 A construction of a counter-example function

Let $E \subseteq X \times Y$ be a bipartite graph, $\lceil \log_2 |X| \rceil = n$ and $\lceil \log_2 |Y| \rceil = m$. If

$$\forall y_0, \ldots, y_n \in Y \exists x \in X; \bigwedge_j \neg(xEy_j)$$

then

$$\exists x_0, \ldots, x_m \in X \forall y \in Y; \bigvee_i \neg(x_i E y)$$

This is easily proved by a pigeon-hole argument. For the purpose of bounded arithmetic we shall relax the statement a bit, removing explicit bounds on the number of x_i's and y_j's.

Definition 5.1 *Let $\alpha(x,y)$ be a binary predicate. $CE(u,\alpha)$ is an $\exists\Pi_1^b(\alpha)$- formula formalizing that either there is a sequence $(x_0, \ldots, x_k)$ of elements smaller than u such that*

$$\forall y \leq u; \bigvee_i \neg\alpha(x_i, y)$$

or there is a sequence $(y_0, \ldots, y_\ell)$ of elements smaller than u such that

$$\forall x \leq u; \bigvee_j \alpha(x, y_j)$$

Lemma 5.2 *Assume that M is a model of PV in which $SAT_n(M)$ is not recognized by a circuit. Assume also that M satisfies for all open PV-formulas $\alpha(x,y)$ the statement $\forall u; CE(u,\alpha)$ with bounds $k, \ell \leq |t(u)|$, t a term.*

Then M admits a counter-example function in $P/poly$.

Proof

Let $\alpha(x,y)$ formalizes that y is a circuit C of size at most m with n inputs, x is a pair (a,w) of $a \in SAT_n(M)$ and $w \models a$, and $C(a) \models a$.

Take the principle $CE(u,\alpha)$ for $u := max(2^{2n}, 2^m)$. The principle provides us either with circuits $C_0, \ldots, C_\ell$ of size at most m such that for every $a \in SAT_n(M)$

$$\bigvee_j C_j(a) \models a$$

or with pairs $(a_0, w_0), \ldots, (a_k, w_k)$ of $a_i \in SAT_n(M)$ and $w_i \models a_i$ such that for every circuit C of size at most m

$$\bigvee_i C(a_i) \not\models a_i$$

The former option is, however, impossible as otherwise we could combine C_j's into one circuit recognizing $SAT_n(M)$. Hence we have the pairs (a_i, w_i) and we define the circuit D_m as follows. Given as an input a circuit C, D_m tries C on all a_i and outputs the first pair (a_i, w_i) such that $C(a_i) \not\models a_i$. Clearly D_m computes a counter-example function for circuits of size at most m.

Q.E.D.

It is open whether the combinatorial principle is provable in PV or even in S_2. A corollary of the principle, namely the tournament principle (see [4, Sec. 12.1]), is also not known to be provable in bounded arithmetic.

Theorem 5.3 *Assume that S_2 proves the formula*

$$\forall u; CE(u, \alpha)$$

for the Δ_1^b-formula $\alpha(x,y)$ defined at the beginning of the proof of Lemma 5.2. Assume also that PV has a countable model in which $NP \not\subseteq P/poly$.

Then $PV \neq S_2^1$.

Proof

Take M a countable model of PV in which $SAT_n(M)$ is not recognized by a circuit. If $M \not\models S_2^1$ then we are done. So assume that $M \models S_2^1$.

Consider the theory T formed by

$$PV \ + \ Th_{\Pi_1^b}(M)$$

together with all formulas

$$\forall y \exists x; Sat_n(x) \not\equiv \phi(x,y)$$

one for each Π_1^b-formula ϕ without parameters.

If T were consistent then any of its models is a Π_1^b-elementary extension of M in which $NP \not\subseteq coNP/poly$ and thus by [2, 9] $PV \neq S_2^1$.

On the other hand, if T is inconsistent then $PV \ + \ Th_{\Pi_1^b}(M)$ proves a disjunction of formulas of the form

$$\exists y \forall x; Sat_n(x) \equiv \phi(x,y)$$

ϕ Π_1^b-formulas without parameters. This means that in M every bounded formula is equivalent to a Σ_1^b-formula and, in particular, the $PIND$ scheme for all bounded formulas holds in M as $M \models S_2^1$. Hence $M \models S_2$ and consequently $M \models \forall u; CE(u, \alpha)$.

By Lemma 5.2 and Theorem 2.1 M has an extension M' in which $NP \not\subseteq coNP/poly$. So, by [2, 9] again, $PV \neq S_2^1$.

Q.E.D.

Acknowledgements: I discussed the idea of the first proof of the theorem with S. Buss in Toronto in Spring 1993 and I thank him for pointing out then a problem with the original version. R. Impagliazzo reminded me of a construction concerning a circuit-simulation of probabilistic computations similar to the one from Lemma 5.2. I thank A. A. Razborov for critical comments on the preliminary version of this paper.

References

[1] Buss, S. R. (1986) Bounded Arithmetic. Naples, Bibliopolis.

[2] —— (1995) Relating the bounded arithmetic and polynomial time hierarchies, *Annals of Pure and Applied Logic*, **75**: 67-77.

[3] Cook, S A. (1975) Feasibly constructive proofs and the propositional calculus, in: *Proc. 7^{th} Annual ACM Symp. on Theory of Computing*, pp. 83-97. ACM Press.

[4] Krajíček, J. (1995) *Bounded arithmetic, propositional logic and complexity theory*, Encyclopedia of Mathematics and Its Applications, Vol. **60**, *Cambridge University Press*, Cambridge - New York - Melbourne, 343 p.

[5] —— (1995) On Frege and Extended Frege proof systems, in: *Feasible Mathematics II*, eds. P. Clote and J. Remmel, Birkhauser, pp. 284-319.

[6] —— (1995) On methods for proving lower bounds in propositional logic, in: Proceedings of the Tenth International Congress *Logic, Methodology and Philosophy of Science*, International Union of History and Philosophy of Science, Florence (August 19-25, 1995), Eds. M. L. Dalla Chiara, K. Doets, D. Mundici, J. van Benthem, to appear.

[7] Krajíček, J and P. Pudlák (1990) Propositional Provability and Models of Weak Arithmetic, in: *Computer Science Logic* (Kaiserlautern, Oct. '89), E. Borger, H. Kleine Buning, M.M. Richter (eds.), LNCS 440, Springer-Verlag, pp. 193-210.

[8] Krajíček, J, Pudlák, P, and Takeuti, G. (1991) Bounded arithmetic and the polynomial hierarchy, *Annals of Pure and Applied Logic*, **52**: 143–153.

[9] Zambella, D. (1996) Notes on polynomially bounded arithmetic, *J. of Symbolic Logic*, **61(3)**, pp.942-966.