BRICS

Basic Research in Computer Science

BRICS RS-94-36 A. A. Razborov: On provably disjoint NP-pairs

On provably disjoint NP-pairs

Alexander A. Razborov

BRICS Report Series RS-94-36

ISSN 0909-0878 November 1994

See back inner page for a list of recent publications in the BRICS Report Series. Copies may be obtained by contacting:

BRICS
Department of Computer Science
University of Aarhus
Ny Munkegade, building 540
DK - 8000 Aarhus C
Denmark

Telephone: +45 8942 3360
Telefax: +45 8942 3255
Internet: BRICS@brics.dk

BRICS publications are in general accessible through WWW and anonymous FTP:

`http://www.brics.dk/`
`ftp ftp.brics.dk (cd pub/BRICS)`

On provably disjoint **NP**-pairs

Alexander A. Razborov[*]
Steklov Mathematical Institute
Vavilova 42, 117966, GSP–1, Moscow, RUSSIA

November 11, 1994

Abstract

In this paper we study the pairs (U, V) of disjoint **NP**-sets representable in a theory T of Bounded Arithmetic in the sense that T proves $U \cap V = \emptyset$. For a large variety of theories T we exhibit a natural disjoint **NP**-pair which is complete for the class of disjoint **NP**-pairs representable in T. This allows us to clarify the approach to showing independence of central open questions in Boolean complexity from theories of Bounded Arithmetic initiated in [11]. Namely, in order to prove the independence result from a theory T, it is sufficient to separate the corresponding complete **NP**-pair by a (quasi)poly-time computable set. We remark that such a separation is obvious for the theory $\mathcal{S}(S_2) + \mathcal{S}\Sigma^b_2 - PIND$ considered in [11], and this gives an alternative proof of the main result from that paper.

1. Introduction

In this paper we study the class of pairs (U, V), where U and V are disjoint **NP**-sets. There are at least two good reasons to be interested in this issue.

Firstly, the question of existence of such a pair not separable by a set in **P** is closely connected to the existence of public-key cryptosystems [5].

[*]Part of this work was done while the author was visiting BRICS, Basic Research in Computer Science, Centre of the Danish National Research Foundation. Supported by the grant # 93-011-16015 of the Russian Foundation for Fundamental Research, and by an AMS-FSU grant.

The second motivation comes from the attempts to understand on the formal level the machinery existing in non-uniform Boolean complexity for proving lower bounds [10, 12, 11]. Of the main importance for this approach is the following observation.

Let U consist of truth-tables of all "simple" Boolean functions, and let

$$V \rightleftharpoons \{ f \oplus s \mid f \in U \},$$

where s is a supposedly complex function in the same number of variables as f. Then proving that s is indeed complex is equivalent to showing that $U \cap V = \emptyset$.

Based upon the notion of a natural proof [12], it was implicitly shown in [11] that if sufficiently strong pseudo-random generators exist then these U and V can not be separated by a quasipolynomial time computable set. It was (also implicitly) shown there that if some particular system $\mathcal{S}(S_2) + \mathcal{S}\Sigma_2^b - PIND$ of Bounded Arithmetic can prove that $U \cap V = \emptyset$ for *some* **NP**-pair (U, V) then this pair can *not* be separated by a quasipolynomial time computable set. Putting things together, we obtain the independence result modulo the hardness assumption.

The question if there exist disjoint **NP**-pairs which can not be separated by a set in **P** is open. Moreover, it was shown in [6] that there exists an oracle relative to which $\mathbf{P} \neq \mathbf{NP}$, and still such pairs do not exist. Thus, the assumption of the existence of **P**-inseparable disjoint **NP**-pairs seems to be stronger than merely $\mathbf{P} \neq \mathbf{NP}$. It should be noted, however, that this assumption is implied by both $\mathbf{P} \neq \mathbf{UP}$ (see e.g. [13, Theorem 9]) and, for obvious reasons, by $\mathbf{P} \neq \mathbf{NP} \cap co - \mathbf{NP}$.

It is known [5, Theorem 6] that every disjoint **NP**-pair is many-one reducible to another disjoint **NP**-pair in which both components are **NP**-complete. However, it is open whether there exists an **NP**-pair which is complete in the class of all disjoint **NP**-pairs under a natural reduction. The reason lies in the highly non-constructive nature of the condition $U \cap V = \emptyset$: e.g. we apparently can not enumerate pairs of nondeterministic poly-time machines producing all disjoint **NP**-pairs.

In this paper we try to build the hierarchy of disjoint **NP**-pairs based upon the strength of logical tools needed for *proving* the fact $U \cap V = \emptyset$. Namely, for a variety of systems T of Bounded Arithmetic, we consider the class of **NP**-pairs for which this fact is provable in T. We exhibit a natural **NP**-pair which is complete in this class under the many-one reduction. Roughly speaking, the first component in this pair consists of all satisfiable CNF, and the second component consists of those unsatisfiable CNF which allow a short refutation in the propositional proof system associated with T. This reduces the approach suggested in [11] to the very concrete algorithmic question: for which theories T the associated complete **NP**-pair can be separated by a quasipolynomial time computable set? Whenever such a separation exists, we have the independence of $\mathbf{NP} \not\subseteq \mathbf{P}/poly$ from the theory T modulo the hardness assumption. For the theory $\mathcal{S}(S_2) + \mathcal{S}\Sigma_2^b - PIND$ the separating set is fairly obvious, and this gives us an alternative, and, perhaps, more natural (not to be confused with the concept from [12]!) proof of the main result from [11].

The paper is organized as follows. In Section 2 we recall necessary facts from Bounded Arithmetic and propositional calculus. In Section 3 we formulate the main concept of an **NP**-pair representable in a theory T and formulate our main result. In Section 4 we demonstrate one nice feature of the split versions introduced in [11]: we show that they allow some sort of elimination of sharply bounded quantifiers. The next section 5 contains the proof of our main theorem. In Section 6 we show how to reduce the approach to proving independence results in Bounded Arithmetic to purely complexity questions. The paper is concluded by a brief discussion of their status in Section 7.

2. Background from Logic

2.1. Systems of Bounded Arithmetic

We assume the familiarity with [1], and use the now-standard notation for denoting various hierarchies and fragments of Bounded Arithmetic from that book. L_1 is the first order language which consists of the constant 0, function symbols $S, +, \cdot, \lfloor \frac{1}{2}x \rfloor, |x|$, and of the predicate symbol $\leq$. L_2 is obtained from L_1 by augmenting it with the smash symbol $\#$ which has the intended meaning $x\#y = 2^{|x|\cdot|y|}$. $L_k(\alpha, \beta)$ $(k = 1, 2)$ is the first-order language obtained from L_k by appending to the latter two new unary predicate symbols $\alpha(a), \beta(a)$, and $\mathcal{L}_k$ is the second-order language based on L_k. To simplify the notation, we will sometimes be using several predicate symbols (second-order variables in the case of $\mathcal{L}_k$) like $\alpha_1, \alpha_2, \ldots$ or $\beta_1, \beta_2, \ldots$: they can always be combined into a single α or β using an easy encoding.

The theories we are interested in will be either in the language $L_k(\alpha, \beta)$ or in $\mathcal{L}_k$ $(k = 1, 2)$. All they contain the set $BASIC_k$ of simple open axioms describing basic properties of symbols from L_k. On the top of it, second-order theories also always include the comprehension axiom scheme $\Sigma_0^{1,b} - CA$. The difference between theories is specified by the amount of induction allowed.

Behind the standard hierarchy Σ_i^b, Π_i^b of bounded formulae we also need its split version $\mathcal{S}\Sigma_i^b, \mathcal{S}\Pi_i^b$ in the language $L_2(\alpha, \beta)$ [11]. $\mathcal{S}\Sigma_0^b = \mathcal{S}\Pi_0^b$ is the set of *all* bounded formulae which contain either only occurrences of α or only occurrences of β. The inductive definition of $\mathcal{S}\Sigma_{i+1}^b, \mathcal{S}\Pi_{i+1}^b$ is the same as for $\Sigma_{i+1}^b, \Pi_{i+1}^b$.

The hierarchy E_i, U_i (see e.g. [17]) was defined as the ordinary hierarchy of bounded formulae in the language of Peano Arithmetic (where we do not have the notion of a sharply bounded quantifier at all). A bounded formula is D_i in a theory T if it is provably equivalent to an E_i- and U_i-formula in T. We extend this hierarchy to the language $L_1(\alpha, \beta)$ simply by counting sharply bounded quantifiers exactly as ordinary quantifiers. The split versions $\mathcal{S}E_i, \mathcal{S}U_i, \mathcal{S}D_i$ of this hierarchy in the language $L_1(\alpha, \beta)$ are defined analogously to $\mathcal{S}\Sigma^b_i, \mathcal{S}\Pi^b_i$.

The following table summarizes the definitions of the theories of Bounded Arithmetic considered in this paper[1]:

Theory	Underlying language	Induction scheme
$S^i_2(\alpha, \beta)$	$L_2(\alpha, \beta)$	$\Sigma^b_i(\alpha, \beta) - PIND$
$\mathcal{S}S^i_2$	$L_2(\alpha, \beta)$	$\mathcal{S}\Sigma^b_i - PIND$
$IE_i(\alpha, \beta)$	$L_1(\alpha, \beta)$	$E_i(\alpha, \beta) - IND$
$\mathcal{S}IE_i$	$L_1(\alpha, \beta)$	$\mathcal{S}E_i - IND$
$T^i_2(\alpha, \beta)$	$L_2(\alpha, \beta)$	$\Sigma^b_i(\alpha, \beta) - IND$
$\mathcal{S}T^i_2$	$L_2(\alpha, \beta)$	$\mathcal{S}\Sigma^b_i - IND$
$I\Delta_0(\alpha, \beta)$	$L_1(\alpha, \beta)$	$\Delta_0(\alpha, \beta) - IND$
$S_2(\alpha, \beta)$	$L_2(\alpha, \beta)$	$\Sigma^b(\alpha, \beta) - PIND$
U^1_1	$\mathcal{L}_1$	$\Sigma^{1,b}_1 - PIND$
U^1_2	$\mathcal{L}_2$	$\Sigma^{1,b}_1 - PIND$
V^1_1	$\mathcal{L}_1$	$\Sigma^{1,b}_1 - IND$
V^1_2	$\mathcal{L}_2$	$\Sigma^{1,b}_1 - IND$

Table 1: Summary of fragments of Bounded Arithmetic

We will need the following easy generalization of [2, Theorem 5] to our setting (see [11]):

Proposition 2.1. *For $i \geq 1$, $\mathcal{S}S^{i+1}_2$ is $\forall\mathcal{S}\Sigma^b_{i+1}$-conservative over $\mathcal{S}T^i_2$.*

[1] we have introduced the natural notation $\mathcal{S}S^i_2, \mathcal{S}T^i_2$ for the theories $\mathcal{S}(S_2) + \mathcal{S}\Sigma^b_i - PIND, \mathcal{S}(S_2) + \mathcal{S}\Sigma^b_i - IND$ from [11] and $\mathcal{S}IE_i$ for the split versions of the theories $IE_i(\alpha, \beta)$. Also, $\Sigma^b \rightleftharpoons \bigcup_{i \geq 0} \Sigma^b_i$.

2.2. Propositional proof systems

In this paper we will be exclusively working with sequential (= natural deduction) proof systems. The cut rule will be always present.

Different proof systems are usually specified by the syntactic requirements placed on the sequents allowed in the proof:

- For a fixed constant $w > 0$, we denote by R_w the system of *bounded resolutions.* All sequents in the proof must have the form $\ell_1, \ldots, \ell_p \longrightarrow \ell_{p+1}, \ldots, \ell_q$, where ℓ_is are *literals* (that is, either propositional variables or their negations) and, moreover, $q \leq w$. Applying cosmetic ($\neg$:right) rule, we can always move all literals to the succedent, after which the cut rule turns into the familiar resolution rule.

- R, *resolutions* is the same system as R_w, only without any restrictions on the length of the sequents.

- F_d is the *depth-d Frege system*: all formulae appearing in the proof must either have the form
$$\bigvee_{i_1=1}^{r_1} \bigwedge_{i_2=1}^{r_2(i_1)} \cdots \mathop{\bigvee\!\!\!\!\!\!\bigwedge}_{i_d=1}^{r_d(i_1,\ldots,i_{d-1})} \ell_{i_1 \ldots i_d} \tag{1}$$
(Σ_d-*formulae*) or
$$\bigwedge_{i_1=1}^{r_1} \bigvee_{i_2=1}^{r_2(i_1)} \cdots \mathop{\bigvee\!\!\!\!\!\!\bigwedge}_{i_d=1}^{r_d(i_1,\ldots,i_{d-1})} \ell_{i_1 \ldots i_d} \tag{2}$$
(Π_d-*formulae*), where $\ell_{i_1 \ldots i_d}$ are literals. The inference rules are modified for unbounded fan-in, e.g. ($\wedge$:right) looks like
$$\frac{\Gamma \longrightarrow A_i, \Delta \ (i \in I)}{\Gamma \longrightarrow \bigwedge_{i \in I} A_i, \Delta}.$$
Note that $F_0 = R$.

- F is the ordinary *Frege system.* At this point it is no longer important that we work in the sequential calculus, but we prefer to stick to this for the sake of uniformity.

- EF is the *extended Frege proof system* [18, 4]. It additionally allows us to use *extension axioms* of the form $p \equiv A$, where p is a new propositional variable (called *extension atom*) which did not appear earlier in the proof.

For an unsatisfiable CNF $\phi = \bigwedge_{i\in I}\bigvee_{j\in J_i}\ell_{ij}$ and a proof system P we denote by $s_P(\phi)$ the minimal possible number of logical symbols in a P-derivation of the empty sequent from the sequents $\longrightarrow \{\ell_{ij} \mid j \in J_i\}$ $(i \in I)$.

2.3. Correspondence between theories of Bounded Arithmetic and propositional proof systems

For many theories of Bounded Arithmetic T there exists a propositional proof system P_T closely associated with T in the following sense:

a) T proves the soundness of P_T,

b) every proof in T of a formula A with appropriately low logical complexity can be efficiently transformed into a short P_T-proof of the propositional variant of A.

In this section we recall those details of this correspondence which will be important in the sequel.

Let $Tr(a, \alpha, \beta)$ be the predicate asserting that the truth assignment α makes the Boolean formula encoded by the string $\beta(0) \ldots \beta(a-1)$ true (*truth definition*).

Proposition 2.2 ([7]). *$Tr(a, \alpha, \beta)$ has a $\Delta_1^{1,b}$-definition in U_1^1 about which U_1^1 proves the usual Tarski's conditions.*

Let $Tr_d(a, \alpha, \beta)$ be the variant of $Tr(a, \alpha, \beta)$ in which $\beta(0) \ldots \beta(a-1)$ encodes a Boolean formula from $\Sigma_d \cup \Pi_d$. The following is straightforward:

Lemma 2.3. *For any fixed $d \geq 0$, $Tr_d(a, \alpha, \beta)$ has a $\mathcal{SD}_{d+1}$-definition in $\mathcal{S}IE_0$ about which $\mathcal{S}IE_0$ proves the usual Tarski's conditions.*

Note for the record that this truth definition can also be assumed to satisfy the natural property

$$\mathcal{S}IE_0 \vdash \forall x < a(\alpha_1(x) \equiv \beta_1(x)) \supset (Tr_d(a, \alpha, \alpha_1) \equiv Tr_d(a, \alpha, \beta_1)). \tag{3}$$

Let the $\Sigma_0^{1,b}$-formula $Ref_P(a_0, a_1, \beta_0, \beta_1)$ assert that the string $\beta_0(0) \ldots \beta_0(a_0 - 1)$ encodes an inference of length $\leq a_0$ in the propositional proof system P of the empty sequent from the clauses of the CNF encoded by $\beta_1(0) \ldots \beta_1(a_1-1)$. The following two propositions are slight modifications of [7, Theorem 2.4] and [7, Theorem 2.5] respectively (the latter also follows from earlier results of Cook [3] via the correspondence between PV and S_2^1 [1, Chapter 6] and $RSUV$-isomorphism [14, 15, 9]):

Proposition 2.4. $U^1_1 \vdash Ref_F(a_0, a_1, \beta_0, \beta_1) \supset \neg Tr_2(a_1, \alpha, \beta_1)$.

Proposition 2.5. $V^1_1 \vdash Ref_{EF}(a_0, a_1, \beta_0, \beta_1) \supset \neg Tr_2(a_1, \alpha, \beta_1)$.

Paris and Wilkie [8] showed that $I\Delta_0(\alpha, \beta) \vdash Ref_{F_d}(a_0, a_1, \beta_0, \beta_1) \supset \neg Tr_2(a_1, \alpha, \beta_1)$ for any fixed $d \geq 0$. We will need the following refinement of their result:

Lemma 2.6. *For any fixed $d \geq 0$, $\mathcal{S}IE_{d+2}(\alpha, \beta) \vdash Ref_{F_d}(a_0, a_1, \beta_0, \beta_1) \supset \neg Tr_2(a_1, \alpha, \beta_1)$.*

Proof. Assuming $Tr_2(a_1, \alpha, \beta_1)$, we prove by induction on c that in ANY one of the first c sequents of the inference encoded by β_0 there EXISTS either a formula ϕ in the antecedent such that $\neg Tr_d(a_1, \alpha, \phi)$ or a formula ϕ in the succedent such that $Tr_d(a_1, \alpha, \phi)$. By Lemma 2.3, the formula expressing this fact is in $\mathcal{S}U_{d+2}$, and $\mathcal{S}U_{d+2} - IND$ is available in $\mathcal{S}IE_{d+2}$.■

Let us now fix propositional variables $p_1, p_2, \ldots, p_n, \ldots, q_1, q_2, \ldots$

Definition 2.7 (see e.g. [7]). For every $A(\vec{a}, \alpha, \beta) \in \Sigma^{1,b}_0$, where all free variables are displayed, and a tuple of integers $\vec{n}$ we define the propositional formula $\langle A(\vec{a})\rangle_{\vec{n}}$ by induction on the complexity of A:

a) if A does not contain occurrences of α and β, and $A(\vec{n})$ is true [false] on integers then $\langle A(\vec{a})\rangle_{\vec{n}} \rightleftharpoons 1$ [0, respectively];

b) if $A(\vec{a}) = \alpha(t(\vec{a}))$ $[\beta(t(\vec{a}))]$ then $\langle A(\vec{a})\rangle_{\vec{n}} \rightleftharpoons p_{t(\vec{n})}$ $[q_{t(\vec{n})}$, respectively];

c) $\langle \neg A(\vec{a})\rangle_{\vec{n}} \rightleftharpoons \neg\langle A(\vec{a})\rangle_{\vec{n}}$;

d) $\langle A(\vec{a}) * B(\vec{a})\rangle_{\vec{n}} \rightleftharpoons \langle A(\vec{a})\rangle_{\vec{n}} * \langle B(\vec{a})\rangle_{\vec{n}}$ for $* \in \{\wedge, \vee, \supset\}$;

e) $\langle (\exists x \leq t(\vec{a}))A(\vec{a}, x)\rangle_{\vec{n}} \rightleftharpoons \bigvee_{m \leq t(\vec{n})} \langle A(\vec{a}, b)\rangle_{\vec{n},m}$;

f) $\langle (\forall x \leq t(\vec{a}))A(\vec{a}, x)\rangle_{\vec{n}} \rightleftharpoons \bigwedge_{m \leq t(\vec{n})} \langle A(\vec{a}, b)\rangle_{\vec{n},m}$.

The following two propositions slightly modify and strengthen [7, Theorems 3.2,3.1] (the latter also follows from [3]):

Proposition 2.8. *Let $U^1_2 \vdash A(\vec{a}, \alpha, \beta)$, where $A(\vec{a}, \alpha, \beta)$ is a $\Sigma^{1,b}_0$-formula with all free variables displayed. Then there exists a quasipolynomial time*[2] *algorithm which for any tuple of integers $\vec{n}$ given in the unary form $1^{\vec{n}}$ produces an F-proof of the propositional formula $\langle A(\vec{a})\rangle_{\vec{n}}$.*

[2] that is with running time $2^{(\log n)^{O(1)}}$. The corresponding class of functions/predicates computable in quasipolynomial time will be denoted by **QP**.

Proposition 2.9. *Let $V_1^1 \vdash A(\vec{a}, \alpha, \beta)$, where $A(\vec{a}, \alpha, \beta)$ is in $\Sigma_0^{1,b}$. Then there exists a polynomial time algorithm which for any $1^{\vec{n}}$ produces an EF-proof of $\langle A(\vec{a}) \rangle_{\vec{n}}$.*

The same remains true after replacing "V_1^1" by "V_2^1", and "polynomial time" by "quasipolynomial time".

A similar result about the provability in $I\Delta_0(\alpha, \beta)$ was established in [8]. It, however, requires more serious adjustment to our purposes, so we defer this until Section 5.

3. Representations of disjoint NP-pairs in systems of Bounded Arithmetic

Definition 3.1. Let U and V be two disjoint sets in **NP**, and T be either a first-order theory in the language $L_k(\alpha, \beta)$ or a second-order theory in the language $\mathcal{L}_k$ ($k = 1, 2$). The pair (U, V) is *representable* in T if there exist $\Sigma_0^{1,b}$-formulae
$A(a, \alpha), B(a, \beta), C(a, b, \alpha), D(a, b, \beta)$ with all free variables displayed such that:

a) for every $w = (w_0, w_1, \ldots, w_{N-1}) \in \{0, 1\}^N$, if $w \in U$ then

$$\mathbf{N} \models \exists \boldsymbol{\alpha} \left(A(N, \boldsymbol{\alpha}) \wedge \forall i < N (C(N, i, \boldsymbol{\alpha}) \equiv w_i = 1) \right),$$

and if $w \in V$ then

$$\mathbf{N} \models \exists \boldsymbol{\beta} \left(B(N, \boldsymbol{\beta}) \wedge \forall i < N (D(N, i, \boldsymbol{\beta}) \equiv w_i = 1) \right);$$

b)

$$T \vdash (A(a, \alpha) \wedge B(a, \beta)) \supset \exists x < a (C(a, x, \alpha) \not\equiv D(a, x, \beta)).$$

Informally, condition a) says that A and B specify some $\widetilde{U} \supseteq U$ and $\widetilde{V} \supseteq V$ as projections of **P**-sets if $k = 1$ and **QP**-sets if $k = 2$. b) means that $\widetilde{U} \cap \widetilde{V} = \emptyset$ is provable in T.

We exploit the ordinary notion of $\leq_m^p$-reducibility in the context of promise problems. Namely, $(U, V) \leq_m^p (U', V')$ means that there is a polynomially time computable function $f : \{0, 1\}^* \longrightarrow \{0, 1\}^*$ such that $f(U) \subseteq U'$ and $f(V) \subseteq V'$. The variant $\leq_m^{qp}$ of this reducibility is defined in the same way with the difference that we only require f to be in **QP**.

Theorem 3.2. **a)** *Let T be one of the theories*

$$\mathcal{S}S_2^i, \mathcal{I}E_i, \mathcal{S}T_2^i\ (i \geq 1), I\Delta_0(\alpha,\beta), S_2(\alpha,\beta), U_1^1, U_2^1, V_1^1, V_2^1.$$

Then the class of **NP**-*pairs representable in T is closed under $\leq_m^p$-reducibility.*

b) *If, moreover, $T \in \{\mathcal{S}S_2^i, \mathcal{S}T_2^i, S_2(\alpha,\beta), U_2^1, V_2^1\}$ then this class is closed under $\leq_m^{qp}$-reducibility.*

Proof. a). Assume that (U,V) is representable in T via bounded formulae $A(a,\alpha), B(a,\beta), C(a,b,\alpha), D(a,b,\beta)$, and let $(U',V') \leq_m^p (U,V)$ via a polynomial time computable function f. Then for a suitable polynomial $p(a)$ we have $\Sigma_0^{1,b}$-formulae $Prot(a,\gamma_0,\gamma_1), Output(a,b,\gamma_0,\gamma_1)$ and Δ_0-definable in $I\Delta_0(\gamma_0,\gamma_1)$ function symbol $Length(a,\gamma_0,\gamma_1)$ expressing the following:

- $Prot(a,\gamma_0,\gamma_1)$ – "$\gamma_1(0)\ldots\gamma_1(p(a)-1)$ is (the encoding of) the protocol of the poly-time computation of f on the input string $\gamma_0(0)\ldots\gamma_0(a-1)$";
- $Length(a,\gamma_0,\gamma_1)$ is the length of the output of γ_1 if $Prot(a,\gamma_0,\gamma_1)$ and 0 otherwise;
- $Output(a,b,\gamma_0,\gamma_1)$ – "$Prot(a,\gamma_0,\gamma_1)$, $b < Length(a,\gamma_0,\gamma_1)$ and the bth bit of γ_1's output is equal to 1".

We now set:

$$\begin{aligned}
A'(a,\alpha_0,\alpha_1,\alpha_2) &\rightleftharpoons Prot(a,\alpha_0,\alpha_1) \wedge A(Length(a,\alpha_0,\alpha_1),\alpha_2)\\
&\qquad \wedge \forall x < Length(a,\alpha_0,\alpha_1)(C(a,x,\alpha_2) \equiv Output(a,x,\alpha_0,\alpha_1))\\
B'(a,\beta_0,\beta_1,\beta_2) &\rightleftharpoons Prot(a,\beta_0,\beta_1) \wedge B(Length(a,\beta_0,\beta_1),\beta_2)\\
&\qquad \wedge \forall x < Length(a,\beta_0,\beta_1)(D(a,x,\beta_2) \equiv Output(a,x,\beta_0,\beta_1))\\
C'(a,b,\alpha_0,\alpha_1,\alpha_2) &\rightleftharpoons \alpha_0(b)\\
D'(a,b,\beta_0,\beta_1,\beta_2) &\rightleftharpoons \beta_0(b).
\end{aligned}$$

We claim that A', B', C', D' provide a representation of (U',V') in the theory T.

Condition a) from Definition 3.1 is straightforward.

In order to see b), suppose, arguing informally in T, that $\forall x < a(\alpha_0(x) \equiv \beta_0(x))$, $A'(a,\alpha_0,\alpha_1,\alpha_2)$ and $B'(a,\beta_0,\beta_1,\beta_2)$. Applying $\mathcal{S}U_1 - IND$ on $c \leq p(a)$ (which is available in T) to the formula $\forall x < c(\alpha_1(x) \equiv \beta_1(x))$, we find $\forall x < p(a)(\alpha_1(x) \equiv \beta_1(x))$. Thus,

T	P_T	reducibility
$\mathcal{S}IE_i\ (i \geq 2)$	F_{i-2}	$\leq_m^p$
$\mathcal{S}T_2^i, \mathcal{S}S_2^{i+1}\ (i \geq 2)$	F_{i-2}	$\leq_m^{qp}$
U_2^1	F	$\leq_m^{qp}$
V_1^1	EF	$\leq_m^p$
V_2^1	EF	$\leq_m^{qp}$

Table 2: $(SAT^*, REF(P_T))$ is complete in the class corresponding to T

$Length(a, \alpha_0, \alpha_1) = Length(a, \beta_0, \beta_1)$ and $\forall x < Length(a, \alpha_0, \alpha_1)(Output(a, x, \alpha_0, \alpha_1) \equiv Output(a, x, \beta_0, \beta_1))$. From the definition of A', B' we conclude

$$\forall x < Length(a, \alpha_0, \alpha_1)(C(a, x, \alpha_2) \equiv Output(a, x, \beta_2)),$$

and this contradicts condition b) for the original pair (U, V) (after substituting $a := Length(a, \alpha_0, \alpha_1), \alpha := \alpha_2, \beta := \beta_2$).

Part b) is proved in exactly the same way.■

Let now $SAT^* \rightleftharpoons \{\langle \phi, 1^t \rangle \mid \phi \text{ is a satisfiable CNF}\}$. For a propositional proof system P, let $REF(P) \rightleftharpoons \{\langle \phi, 1^t \rangle \mid \phi \text{ is an unsatisfiable CNF and } s_P(\phi) \leq t\}$. Obviously, $SAT^*, REF(P) \in \mathbf{NP}$ and $SAT^* \cap REF(P) = \emptyset$. The following theorem is the main result of this paper.

Theorem 3.3. *Let T be one of the theories in the left column of Table 2, and P_T be the corresponding proof system in the middle column. Then $(SAT^*, REF(P_T))$ is complete in the class of disjoint* **NP***-pairs representable in T with respect to the reducibility given in the right column.*

The proof of this theorem will be given in two subsequent sections.

We conclude this section with the following corollary asserting a certain symmetry of pairs $(SAT^*, REF(P))$:

Corollary 3.4. $(REF(F_d), SAT^*) \leq_m^p (SAT^*, REF(F_d))$ $(d \geq 0)$*,* $(REF(F), SAT^*) \leq_m^{qp} (SAT^*, REF(F))$*, and* $(REF(EF), SAT^*) \leq_m^p (SAT^*, REF(EF))$*.*

Proof. Immediately follows from Theorem 3.3 since the notion of a pair representable in a theory T is symmetric with respect to the two components U, V.■

4. Elimination of sharply bounded quantifiers in split versions

Let us consider the analogue $E_i^{\#}, U_i^{\#}$ of the hierarchy E_i, U_i in the language L_2, and its split versions $\mathcal{S}E_i^{\#}, \mathcal{S}U_i^{\#}$ in the language $L_2(\alpha, \beta)$. Thus, $\mathcal{S}E_i^{\#}, \mathcal{S}U_i^{\#}$ differ from $\mathcal{S}E_i, \mathcal{S}U_i$ only in the underlying language, whereas the syntactic inductive definitions for both hierarchies are the same. The theories $IE_i^{\#}, \mathcal{S}IE_i^{\#}$ have the obvious meaning. In this section we prove the following:

Theorem 4.1. *$\mathcal{S}IE_i^{\#} = \mathcal{S}T_2^i$ for all $i \geq 0$.*

Proof. Since $\mathcal{S}E_i^{\#} \subseteq \mathcal{S}\Sigma_i^b$, it suffices to show that $\mathcal{S}IE_i^{\#} \vdash \mathcal{S}\Sigma_i^b - IND$. This will be immediately implied by the following

Claim 4.2. *Let $0 \leq j \leq i$. Then every $\mathcal{S}\Sigma_j^b$-formula is equivalent in $\mathcal{S}IE_i^{\#}$ to a $\mathcal{S}E_j^{\#}$-formula.*

Proof of Claim 4.2. W.l.o.g. we may assume that $A \in \mathcal{S}\Sigma_j^b$ contains only connectives $\{\neg, \wedge, \vee\}$ and, moreover, that negations appear on atomic subformulae only. Now we apply induction on $\langle j, |A| \rangle$.

Base $j = 0$ is obvious since $\mathcal{S}\Sigma_0^b = \mathcal{S}E_0^{\#}$.

Inductive step. Let $j > 0$ and $A \in \mathcal{S}\Sigma_j^b$. If $A \in \mathcal{S}\Pi_{j-1}^b$, we convert $(\neg A)$ into the equivalent form $\bar{A} \in \mathcal{S}\Sigma_{j-1}^b$ obeying the above restrictions, and apply to $\bar{A}$ the inductive assumption with $j := j - 1$. If $A = B * C$ or $A = (\exists x \leq t) B(x)$, the inductive step is obvious ($\mathcal{S}E_j^{\#}$ is closed under these operations).

The only nontrivial case is $A = (\forall x \leq |t|) B(x)$. By the inductive assumption, $B(a)$ is equivalent in $\mathcal{S}IE_i^{\#}$ to a $\mathcal{S}E_j^{\#}$-formula, and we can further assume that this formula is in the prenex normal form. That is to say,

$$\mathcal{S}IE_i^{\#} \vdash A \equiv \forall x \leq |t| \exists y_1 \leq s_1 \ldots \exists y_\ell \leq s_\ell \forall \vec{z}^{(2)} \leq \vec{r}^{(2)} \ldots Q \vec{z}^{(j)} \leq \vec{r}^{(j)} C(x, \vec{y}, \vec{z}^{(2)}, \ldots, \vec{z}^{(j)}),$$

where C is a Boolean combination of $\mathcal{S}E_0^{\#}$-formulae. The crucial point is that since $\mathcal{S}IE_i^{\#}$ contains S_2^1, it can also define all Σ_1^b-definable in S_2^1 function symbols. Moreover, usage of this symbols does not increase the logical complexity of formulae in terms of the hierarchy $\mathcal{S}E_i^{\#}$ (remember that $\mathcal{S}E_0^{\#}$ consists of *all* bounded formula either not containing α or not containing β).

We claim that the formula[3]

$$D(a,\vec{b}) \rightleftharpoons \forall x \le |a| \forall \vec{z}^{(2)} \le \vec{r}^{(2)} \dots Q\vec{z}^{(j)} \le \vec{r}^{(j)} C(x,(b_1)_{x+1},\dots,(b_\ell)_{x+1},\vec{z}^{(2)},\dots,\vec{z}^{(j)})$$

is equivalent to a formula in $\mathcal{S}E_j^{\#}$. This is obvious if $j \ge 2$ (in fact, D is even in $\mathcal{S}U_{j-1}^{\#}$). If $j = 1$, we can represent $C(a,\vec{b})$ in the equivalent form

$$C(a,\vec{b}) \equiv \bigwedge_{i=1}^{m} \left(C_i'(a,\vec{b},\alpha) \vee C_i''(a,\vec{b},\beta) \right),$$

and we are left to show that $\forall x \le |a| \left(C_i'(a,\vec{b},\alpha) \vee C_i''(a,\vec{b},\beta) \right)$ is equivalent to a $\mathcal{S}E_1^{\#}$-formula. The required formula is simply

$$\begin{aligned} &\exists y' \le 4a \exists y'' \le 4a \Big(\forall x \le |a| (C_i'(x,\vec{b},\alpha) \equiv Bit(x,y')) \\ &\quad \wedge \forall x \le |a| (C_i''(x,\vec{b},\beta) \equiv Bit(x,y'')) \wedge \forall x \le |a| (Bit(x,y') = 1 \vee Bit(x,y'') = 1) \Big). \end{aligned}$$

Now, when we know that $D(a,\vec{b})$ is provably equivalent to a $\mathcal{S}E_j^{\#}$-formula, we can apply $\mathcal{S}E_j^{\#} - PIND$ on a to the formula $(\exists y_1 \le SqBd(a,s_1)) \dots (\exists y_\ell \le SqBd(a,s_\ell)) D(a,\vec{y})$ to see that $\mathcal{S}IE_i^{\#} \vdash A \equiv (\exists y_1 \le SqBd(t,s_1)) \dots (\exists y_\ell \le SqBd(t,s_\ell)) D(t,\vec{y})$.

This completes the proof of Claim 4.2.■

As we noted above, Theorem 4.1 follows.■

Remark 4.3. It is worth noting that the similar question $T_2^i \overset{?}{=} IE_i^{\#}$ is open.

5. Proof of Theorem 3.3

We start by showing that $(SAT^*, REF(P_T))$ is representable in T (this part is easier). It is sufficient to consider the cases $(T, P_T) = (\mathcal{S}IE_i, F_{i-2}), (U_2^1, F)$ or (V_1^1, EF) (in fact, for the second case we will be able to show that $(SAT^*, REF(F))$ is representable already in U_1^1). This is actually almost explicitly contained in Propositions 2.4, 2.5 and Lemma 2.6.

Formally, we construct the representation $A(a,\alpha_0,\alpha), B(a,\beta_0,\beta), C(a,b,\alpha), D(a,b,\beta)$ of $(SAT^*, REF(P_T))$ in T as follows:

[3] to avoid collision with another usage of β, we denote the xth member of a sequence b by $(b)_x$ rather than by $\beta(x,b)$

- $A(a, \alpha_0, \alpha)$ asserts that the string $\alpha(0) \dots \alpha(a-1)$ encodes a pair of the form $\langle \phi, 1^t \rangle$, where ϕ is a CNF such that $Tr_2(|\phi|, \alpha_0, \phi)$;
- $B(a, \beta_0, \beta)$ asserts that the string $\beta(0) \dots \beta(a-1)$ encodes $\langle \phi, 1^t \rangle$, where ϕ is a CNF such that $Ref_{P_T}(t, |\phi|, \beta_0, \phi)$;
- $C(a, b, \alpha) \rightleftharpoons \alpha(b)$;
- $D(a, b, \beta) \rightleftharpoons \beta(b)$.

Then condition a) of Definition 3.1 is straightforward. Condition b) is also easy to see: arguing informally in T, if we have $\forall x < a(\alpha(x) \equiv \beta(x))$, where $\alpha(0) \dots \alpha(a-1)$ encodes a pair $\langle \phi_\alpha, 1^{t_\alpha} \rangle$, and $\beta(0) \dots \beta(a-1)$ encodes a pair $\langle \phi_\beta, 1^{t_\beta} \rangle$, then $|\phi_\alpha| = |\phi_\beta|$ and $\forall x < |\phi_\alpha|(\phi_\alpha(x) \equiv \phi_\beta(x))$. This, along with $Tr_2(|\phi_\alpha|, \alpha_0, \phi_\alpha)$, implies by (3) $Tr_2(|\phi_\beta|, \alpha_0, \phi_\beta)$, and now we only have to apply Lemma 2.6, Proposition 2.4 or Proposition 2.5 (depending on T) with $a_0 := t, a_1 := |\phi_\beta|, \beta_1 := \phi_\beta$.

Now we prove the second part of Theorem 3.3. Namely, assume that (U, V) is representable in T, where T is one of the theories in the left column of Table 2. We want to show that (U, V) is reducible to $(SAT^*, REF(P_T))$.

For this we need to modify Definition 2.7. Firstly we enlarge our alphabet of propositional variables. Now it will consist of all variables of the form $p_{A(\vec{a},\alpha),\vec{n}}$, $q_{B(\vec{a},\beta),\vec{n}}$, all free variables in $A, B \in \Sigma_0^{1,b}$ being displayed, and we identify original p_n, q_n with $p_{\alpha(a),n}, q_{\alpha(a),n}$. Note that this time we have two different alphabets corresponding to the languages L_1, L_2; it will be always clear from the context which one is used. Also we assume for simplicity that A and B contain the connectives from $\{\neg, \wedge, \vee\}$ only.

We define the modification $\{A(\vec{a})\}_{\vec{n}}$ of $\langle A(\vec{a}) \rangle_{\vec{n}}$ by extending item b) in Definition 2.7 to

b)* *if* $A(\vec{a}, \alpha)$ $[B(\vec{a}, \beta)]$ *contains occurrences of* α $[\beta]$ *but does not contain occurrences of* β $[\alpha]$ *then* $\{A(\vec{a}, \alpha)\}_{\vec{n}} \rightleftharpoons p_{A(\vec{a},\alpha),\vec{n}}$ $[\{B(\vec{a}, \beta)\}_{\vec{n}} \rightleftharpoons q_{B(\vec{a},\beta),\vec{n}}$, *respectively*].

In accordance with this, items c)-f) are restricted to the case when the formula on the left-hand side contains occurrences of *both* α and β.

Denote be Def_α the following set of propositional sequents, where A, B run over all $\Sigma^b(\alpha)$-formulae, and t runs over all first-order terms[4]:

$$p_{A(\vec{t}(\vec{a})),\vec{n}} \longleftrightarrow p_{A(\vec{a}),\vec{t}(\vec{n})};$$
$$p_{\neg A(\vec{a}),\vec{n}} \longleftrightarrow \bar{p}_{A(\vec{a}),\vec{n}};$$

[4] we will use the notation $\Gamma \longleftrightarrow \Delta$ for denoting the pair of sequents $\Gamma \longrightarrow \Delta$ and $\Delta \longrightarrow \Gamma$

$$
\begin{aligned}
&p_{A(\vec{a})\wedge B(\vec{a}),\vec{n}} \longrightarrow p_{A(\vec{a}),\vec{n}};\\
&p_{A(\vec{a})\wedge B(\vec{a}),\vec{n}} \longrightarrow p_{B(\vec{a}),\vec{n}};\\
&p_{A(\vec{a}),\vec{n}}, p_{B(\vec{a}),\vec{n}} \longrightarrow p_{A(\vec{a})\wedge B(\vec{a}),\vec{n}};\\
&p_{A(\vec{a})\vee B(\vec{a}),\vec{n}} \longrightarrow p_{A(\vec{a}),\vec{n}}, p_{B(\vec{a}),\vec{n}};\\
&p_{A(\vec{a}),\vec{n}} \longrightarrow p_{A(\vec{a})\vee B(\vec{a}),\vec{n}};\\
&p_{B(\vec{a}),\vec{n}} \longrightarrow p_{A(\vec{a})\vee B(\vec{a}),\vec{n}};\\
&p_{(\exists x\leq a)A(x,\vec{b}),n,\vec{m}} \longrightarrow p_{A(a,\vec{b}),0,\vec{m}}, \ldots, p_{A(a,\vec{b}),n,\vec{m}}; \qquad (4)\\
&p_{A(a,\vec{b}),n,\vec{m}} \longrightarrow p_{(\exists x\leq a)A(x,\vec{b}),n',\vec{m}} \ (n \leq n');\\
&p_{(\forall x\leq a)A(x,\vec{b}),n',\vec{m}} \longrightarrow p_{A(a,\vec{b}),n,\vec{m}} \ (n \leq n');\\
&p_{A(a,\vec{b}),0,\vec{m}}, \ldots, p_{A(a,\vec{b}),n,\vec{m}} \longrightarrow p_{(\forall x\leq a)A(x,\vec{b}),n,\vec{m}}. \qquad (5)
\end{aligned}
$$

Def_β is defined in the same way.

We also consider the variant Σ'_d, Π'_d of the hierarchy Σ_d, Π_d of Boolean formulae (see Section 2.2) by allowing $\ell_{i_1\ldots i_d}$ in (1), (2) to have the form $p * q$, where $* \in \{\wedge, \vee\}$, and p, q are propositional variables from the corresponding alphabets. Let F'_d be the variant of the proof system F_d in which we allow the formulae from $\Sigma'_d \cup \Pi'_d$ in the proofs.

Lemma 5.1. *Let T be one of the theories in the left column of Table 2. Assume that*

$$T \vdash \exists \vec{x} \leq \vec{t}(\vec{a})(A(\vec{a},\vec{x},\alpha) \wedge B(\vec{a},\vec{x},\beta)),$$

where $A, B \in \Sigma_0^{1,b}$ with all free variables displayed, and $\vec{t}(\vec{a})$ are arbitrary terms of the underlying language. Then there exists a polynomial or quasipolynomial, depending on the entry in the right column, algorithm which for every tuple of integers $\vec{n}$ written in unary produces a proof of the empty sequent in the system F'_{i-2}, F or EF determined by the middle column from the set of axioms

$$\mathrm{Def}_\alpha, \mathrm{Def}_\beta, \left\{ \longrightarrow \bar{p}_{A(\vec{a},\vec{b},\alpha),\vec{n},\vec{m}}, \bar{q}_{B(\vec{a},\vec{b},\beta),\vec{n},\vec{m}} \,\middle|\, \vec{m} \leq \vec{t}(\vec{n}) \right\}. \qquad (6)$$

Proof. We start with the case of second-order theories (lines 3-5) as it rather easily follows from known results. Namely, we can construct in polynomial or quasipolynomial (depending on the underlying language) time F-proofs

$$\mathrm{Def}_\alpha \vdash \langle A(\vec{a},\vec{b},\alpha)\rangle_{\vec{n},\vec{m}} \equiv p_{A(\vec{a},\vec{b},\alpha),\vec{n},\vec{m}}$$

and

$$\mathrm{Def}_\beta \vdash \langle B(\vec{a},\vec{b},\beta)\rangle_{\vec{n},\vec{m}} \equiv q_{B(\vec{a},\vec{b},\beta),\vec{n},\vec{m}}.$$

Using these, we construct F-proofs of the formulae $\langle\neg(A(\vec{a},\vec{b},\alpha)\wedge B(\vec{a},\vec{b},\beta))\rangle_{\vec{n},\vec{m}}$ $(\vec{m}\le\vec{t}(\vec{n}))$ from the axioms (6). Then we construct, using Propositions 2.8, 2.9, an F-proof or EF-proof, depending on the theory T, of the formula $\langle\exists\vec{x}\le\vec{t}(\vec{a})(A(\vec{a},\vec{x},\alpha)\wedge B(\vec{a},\vec{x},\beta))\rangle_{\vec{n}}$, and apply a sequence of cuts to derive the empty sequent.

Assume now that T is a first-order theory from the first two lines of Table 2. If T comes from the second line, then we can, using Proposition 2.1 and Theorem 4.1, replace it by $\mathcal{S}IE_i^{\#}$. Now, the theories $\mathcal{S}IE_i$ and $\mathcal{S}IE_i^{\#}$ differ only in the underlying language, and the rest of the proof is absolutely identically for them. So, we consider only the case of $\mathcal{S}IE_i$.

Every $\mathcal{S}E_j$-formula $(j\ge1)$ is equivalent to a formula in the prenex normal form and it is easily seen to be further equivalent in $\mathcal{S}IE_0$ to a formula of the form

$$\left.\begin{array}{c}\exists\vec{x}^{(1)}\le\vec{t}^{(1)}(\vec{a})\forall\vec{x}^{(2)}\le\vec{t}^{(2)}(\vec{a})\dots Q\vec{x}^{(j)}(\vec{a})\le\vec{t}^{(j)}(\vec{a})\\ \Big(C(\vec{a},\vec{x}^{(1)},\dots,\vec{x}^{(j)},\alpha)*D(\vec{a},\vec{x}^{(1)},\dots,\vec{x}^{(j)},\beta)\Big),\end{array}\right\}\quad(7)$$

where $*\in\{\wedge,\vee\}$. Denote by $\mathcal{S}E'_j$ the class of formulae having the form (7), and let $\mathcal{S}U'_j$ be the dual class. For $C\in\mathcal{S}E'_j$ $[C\in\mathcal{S}U'_j]$ we denote by $\bar{C}$ the dual formula in $C\in\mathcal{S}U'_j$ $[C\in\mathcal{S}E'_j$, respectively$]$ logically equivalent to $(\neg C)$. Note that for $C(\vec{a})\in\mathcal{S}U'_{i-2}$ and every tuple $\vec{n}$, the propositional formula $\{C(\vec{a})\}_{\vec{n}}$ is in Π'_{i-2}.

For $C(\vec{a})\in\mathcal{S}E'_{i-1}\setminus\mathcal{S}U'_{i-2}$; $C(\vec{a})=(\exists\vec{x}\le\vec{t}(\vec{a}))D(\vec{a},\vec{x})$, where $D(\vec{a},\vec{b})$ is in $\mathcal{S}U'_{i-2}$, denote by $\Gamma_{C(\vec{a}),\vec{n}}$ the cedent consisting of the formulae $\Big\{D(\vec{a},\vec{b})\Big\}_{\vec{n},\vec{m}}$ $(\vec{m}\le\vec{t}(\vec{n}))$. If $C(\vec{a})\in\mathcal{S}U'_{i-2}$, we let $\Gamma_{C(\vec{a}),\vec{n}}$ consist of the single formula $\{C(\vec{a})\}_{\vec{n}}$.

For $C(\vec{a})\in\mathcal{S}U'_i\setminus\mathcal{S}E'_{i-1}$; $C(\vec{a})=(\forall\vec{x}\le\vec{t}(\vec{a}))D(\vec{a},\vec{x})$, where $D(\vec{a},\vec{b})$ is in $\mathcal{S}E'_{i-1}$, denote by $\mathcal{G}_{C(\vec{a}),\vec{n}}$ the collection of sequents $\Big\{\longrightarrow\Gamma_{D(\vec{a},\vec{b}),\vec{n},\vec{m}}\Bigm|\vec{m}\le\vec{t}(\vec{n})\Big\}$. In the case $C(\vec{a})\in\mathcal{S}E'_{i-1}$, we let $\mathcal{G}_{C(\vec{a}),\vec{n}}$ consist of the single sequent $\longrightarrow\Gamma_{C(\vec{a}),\vec{n}}$.

The following two statements are proven by an easy induction on the logical complexity of C:

Statement 5.2. *For every $C(\vec{a},\vec{b})\in\mathcal{S}U'_{i-2}$ and terms $\vec{t}(\vec{a})$ there is a polynomial time algorithm which for any tuple of integers $\vec{n}$ (written in unary) produces an F'_{i-2}-proof of $\Big\{C(\vec{a},\vec{b})\Big\}_{\vec{n},\vec{t}(\vec{n})}\longleftrightarrow\Big\{C(\vec{a},\vec{t}(\vec{a}))\Big\}_{\vec{n}}$ from* $\mathrm{Def}_\alpha,\mathrm{Def}_\beta$.

Statement 5.3. *Let $C(\vec{a})\in\mathcal{S}E'_{i-1}$.*

a) *There exists a polynomial time algorithm which for any $1^{\vec{n}}$ and any formula $L\in\Gamma_{C(\vec{a}),\vec{n}}$ produces an F'_{i-2}-proof*

$$\mathrm{Def}_\alpha,\mathrm{Def}_\beta,\mathcal{G}_{\bar{C}(\vec{a}),\vec{n}}\vdash L\longrightarrow .$$

b) *There exists a polynomial time algorithm which for any* $1^{\vec{n}}$ *and any sequent* $(\longrightarrow \Gamma) \in \mathcal{G}_{\bar{C}(\vec{a}),\vec{n}}$ *produces an* F'_{i-2}*-proof*

$$\mathrm{Def}_\alpha, \mathrm{Def}_\beta \vdash\!\longrightarrow \Gamma_{C(\vec{a}),\vec{n}}, \Gamma.$$

We are going to prove the following generalization of Lemma 5.1:

Statement 5.4. *Suppose that*

$$\left.\begin{array}{rl} \mathcal{S}IE_i \ \vdash & A_1(\vec{a}), \ldots, A_k(\vec{a}), B_1(\vec{a}), \ldots, B_\ell(\vec{a}) \\ & \qquad \longrightarrow A_{k+1}(\vec{a}), \ldots, A_r(\vec{a}), B_{\ell+1}(\vec{a}), \ldots, B_s(\vec{a}), \end{array}\right\} \quad (8)$$

where $A_1, \ldots, A_k, B_{\ell+1}, \ldots, B_s \in \mathcal{S}U'_i$; $B_1, \ldots, B_\ell, A_{k+1}, \ldots, A_r \in \mathcal{S}E'_i$, *and all free variables are explicitly displayed. Then there exists a polynomial time algorithm which for any tuple of integers* $\vec{n}$ *written in unary and any cedents* $\Gamma_1, \ldots, \Gamma_s$, *where* $(\longrightarrow \Gamma_\nu) \in \begin{cases} \mathcal{G}_{\bar{B}_\nu(\vec{a}),\vec{n}} \text{ if } 1 \le \nu \le \ell \\ \mathcal{G}_{B_\nu(\vec{a}),\vec{n}} \text{ if } \ell+1 \le \nu \le s, \end{cases}$ *produces an* F'_{i-2}*-proof*

$$\mathrm{Def}_\alpha, \mathrm{Def}_\beta, \mathcal{G}_{A_1(\vec{a}),\vec{n}}, \ldots, \mathcal{G}_{A_k(\vec{a}),\vec{n}}, \mathcal{G}_{\bar{A}_{k+1}(\vec{a}),\vec{n}}, \ldots, \mathcal{G}_{\bar{A}_r(\vec{a}),\vec{n}} \vdash\!\longrightarrow \Gamma_1, \ldots, \Gamma_s. \quad (9)$$

Proof of Statement 5.4. As we noticed above, every $\mathcal{S}E_i$-formula is equivalent in $\mathcal{S}IE_0$ to an $\mathcal{S}E'_i$-formula. Thus we can assume that $\mathcal{S}E_i - IND$ in the proof (8) is applied only to $\mathcal{S}E'_i$-formulae. By the Cut Elimination Theorem (see e.g. [1, Theorem 4.3]) we can also assume that all formulae appearing in this proof belong to $\mathcal{S}E'_i \cup \mathcal{S}U'_i$. Let P be this reduced proof.

Now we apply induction on the number of inferences in P. As usual, the argument splits into many cases depending on the final inference (the case when P consists of a single axiom is completely trivial). Most of these cases are straightforward, so we consider explicitly only a few of them. We can assume w.l.o.g. that the final sequent of P has the form $A_1(\vec{a}), \ldots, A_r(\vec{a}) \longrightarrow B_1(\vec{a}), \ldots, B_s(\vec{a})$, where $A_1, \ldots, A_r, B_1, \ldots, B_s \in \mathcal{S}U'_i$. Suppose that we are given integers $\vec{n}$ and $(\longrightarrow \Gamma_\nu) \in \mathcal{G}_{B_\nu(\vec{a}),\vec{n}}$ $(1 \le \nu \le s)$, and we have to construct efficiently an F'_{i-2}-proof (9).

($\vee$:left). Assume that the final inference of P has the form

$$\frac{A'(\vec{a}), A_2(\vec{a}), \ldots, A_r(\vec{a}) \longrightarrow B_1(\vec{a}), \ldots, B_s(\vec{a}) \quad A''(\vec{a}), \ldots, A_r(\vec{a}) \longrightarrow B_1(\vec{a}), \ldots, B_s(\vec{a})}{A'(\vec{a}) \vee A''(\vec{a}), A_2(\vec{a}), \ldots, A_r(\vec{a}) \longrightarrow B_1(\vec{a}), \ldots, B_s(\vec{a})}.$$

Due to the syntactic structure of $\mathcal{S}U'_i$-formulae, $(A'(\vec{a}) \vee A''(\vec{a})) \in \mathcal{S}U'_0$. Hence, by induction hypothesis we have F'_{i-2}-proofs of the sequent $\longrightarrow \Gamma_1, \ldots, \Gamma_s$ from both

$$\mathrm{Def}_\alpha, \mathrm{Def}_\beta, \{A'(\vec{a})\}_{\vec{n}}, \mathcal{G}_2, \ldots, \mathcal{G}_r$$

and

$$\mathrm{Def}_\alpha, \mathrm{Def}_\beta, \{A''(\vec{a})\}_{\vec{n}}, \mathcal{G}_2, \ldots, \mathcal{G}_r.$$

We modify the first proof by adding $\{A'(\vec{a})\}_{\vec{n}}$ to antecedents of all its sequents. This will result in an F'_{i-2}-proof of $\{A'(\vec{a})\}_{\vec{n}} \longrightarrow \Gamma_1, \ldots, \Gamma_s$ from axioms $\mathrm{Def}_\alpha, \mathrm{Def}_\beta, \mathcal{G}_2, \ldots, \mathcal{G}_r$. A similar procedure applied to the second proof gives us a proof of $\{A''(\vec{a})\}_{\vec{n}} \longrightarrow \Gamma_1, \ldots, \Gamma_s$ from the same axioms. The sequent $\longrightarrow \{A'(\vec{a})\}_{\vec{n}}, \{A''(\vec{a})\}_{\vec{n}}$, however, has an obvious proof from $\mathrm{Def}_\alpha, \mathrm{Def}_\beta, \{A'(\vec{a}) \vee A''(\vec{a})\}_{\vec{n}}$. Applying twice the cut rule, we will find the desired proof $\mathrm{Def}_\alpha, \mathrm{Def}_\beta, \{A'(\vec{a}) \vee A''(\vec{a})\}_{\vec{n}}, \mathcal{G}_2, \ldots, \mathcal{G}_r \vdash\!\!\longrightarrow \Gamma_1, \ldots, \Gamma_s$. It is easy to see that the whole construction is polynomial time computable.

($\forall \leq$:left). Assume that the final inference of P has the form

$$\frac{A(\vec{a}, t(\vec{a})), A_2(\vec{a}), \ldots, A_r(\vec{a}) \longrightarrow B_1(\vec{a}), \ldots, B_s(\vec{a})}{t(\vec{a}) \leq s(\vec{a}), (\forall x \leq s(\vec{a}))A(\vec{a}, x), A_2(\vec{a}), \ldots, A_r(\vec{a}) \longrightarrow B_1(\vec{a}), \ldots, B_s(\vec{a})}.$$

If $t(\vec{n}) \leq s(\vec{n})$ is false, everything is obvious. Otherwise, it is easy to see that every sequent in $\mathcal{G}_{A(\vec{a},b),\vec{n},t(\vec{n})}$ has a short proof from $\mathrm{Def}_\alpha, \mathrm{Def}_\beta, \mathcal{G}_{(\forall x \leq s(\vec{a}))A(\vec{a},x),\vec{n}}$, and, by Statement 5.2, the same is true for every sequent in $\mathcal{G}_{A(\vec{a},t(\vec{a})),\vec{n}}$. Hence we can apply the inductive assumption.

($\forall \leq$:right). Assume that the final inference of P is

$$\frac{b \leq t(\vec{a}), A_1(\vec{a}), \ldots, A_r(\vec{a}) \longrightarrow B_1(\vec{a}), \ldots, B_{s-1}(\vec{a}), B(\vec{a}, b)}{A_1(\vec{a}), \ldots, A_r(\vec{a}) \longrightarrow B_1(\vec{a}), \ldots, B_{s-1}(\vec{a}), (\forall x \leq t(\vec{a}))B(\vec{a}, x)}.$$

If $(\forall x \leq t(\vec{a}))B(\vec{a}, x) \in \mathcal{S}E'_{i-1}$ then it is actually in $\mathcal{S}U'_{i-2}$. By inductive assumption, we have efficient F'_{i-2}-proofs $\mathrm{Def}_\alpha, \mathrm{Def}_\beta, \mathcal{G}_1, \ldots, \mathcal{G}_r \vdash\!\!\longrightarrow \Gamma_1, \ldots, \Gamma_{s-1}, \{B(\vec{a}, b)\}_{\vec{n},m}$ for all $m \leq t(\vec{n})$. Applying (5) followed by a sequence of cuts in the case $B(\vec{a}, b) \in \mathcal{S}U'_0$, and ($\wedge$:right) otherwise, we find an efficient proof of $\longrightarrow \Gamma_1, \ldots, \Gamma_{s-1}, \{(\forall x \leq t(\vec{a}))B(\vec{a}, x)\}_{\vec{n}}$ from the same axioms.

If $(\forall x \leq t(\vec{a}))B(\vec{a}, x) \notin \mathcal{S}E'_{i-1}$ then $(\longrightarrow \Gamma_s) \in \mathcal{G}_{B(\vec{a},b),\vec{n},m}$ for some $m \leq t(\vec{n})$, and we simply use the proof of $\longrightarrow \Gamma_1, \ldots, \Gamma_{s-1}, \Gamma_s$ available by inductive assumption.

($\exists \leq$:left). The final inference has the form

$$\frac{b \leq t(\vec{a}), A(\vec{a}, b), A_2(\vec{a}), \ldots, A_r(\vec{a}) \longrightarrow B_1(\vec{a}), \ldots, B_s(\vec{a})}{(\exists x \leq t(\vec{a}))A(\vec{a}, x), A_2(\vec{a}), \ldots, A_r(\vec{a}) \longrightarrow B_1(\vec{a}), \ldots, B_s(\vec{a})}.$$

$(\exists x \leq t(\vec{a}))A(\vec{a},x)$ should necessarily belong to $\mathcal{SE}'_{i-1}$, hence $\mathcal{G}_{(\exists x\leq t(\vec{a}))A(\vec{a},x),\vec{n}}$ and $\mathcal{G}_{A(\vec{a},b),\vec{n},m}$ consist of single sequents with empty antecedents. Denote by Δ and Δ_m, respectively, their succedents.

By inductive assumption, for any $m \leq t(\vec{n})$ we have an F'_{i-2}-proof $\mathrm{Def}_\alpha, \mathrm{Def}_\beta, (\longrightarrow \Delta_m), \mathcal{G}_2, \ldots, \mathcal{G}_r \vdash\longrightarrow \Gamma_1, \ldots, \Gamma_s$. These proofs give raise to proofs

$$\mathrm{Def}_\alpha, \mathrm{Def}_\beta, L, \mathcal{G}_2, \ldots, \mathcal{G}_r \vdash \Gamma_1, \ldots, \Gamma_s$$

for every $L \in \bigcup_{m\leq t(\vec{n})} \Delta_m$. Also, $\longrightarrow \Delta_0, \Delta_1, \ldots, \Delta_{t(\vec{n})}$ has an efficient proof from $\mathrm{Def}_\alpha, \mathrm{Def}_\beta, (\longrightarrow \Delta)$. Now we argue as in the case ($\vee$:left).

$(\mathcal{SE}'_i - IND)$. The last inference has the form

$$\frac{A_1(\vec{a}), \ldots, A_r(\vec{a}), A(\vec{a},b) \longrightarrow A(\vec{a},b+1), B_1(\vec{a}), \ldots, B_s(\vec{a})}{A_1(\vec{a}), \ldots, A_r(\vec{a}), A(\vec{a},0) \longrightarrow A(\vec{a},t(\vec{a})), B_1(\vec{a}), \ldots, B_s(\vec{a})},$$

where $A(\vec{a},b)$ is in $\mathcal{SE}'_i$. Replacing $A(\vec{a},b)$ by $\bar{A}(\vec{a}, t(\vec{a}) \dot{-} b)$ if necessary, we may assume that A is instead in $\mathcal{SU}'_i$ and, moreover, one of the following is true:

a) $A(\vec{a},0)$ is on the list $A_1, \ldots, A_k, B_{\ell+1}, \ldots, B_s$, and $A(\vec{a},t(\vec{a}))$ is on the list $B_1, \ldots, B_\ell, A_{k+1}, \ldots, A_r$ in (8);

b) $A(\vec{a},0), A(\vec{a},t(\vec{a}))$ are on the same list, and $A \in \mathcal{SE}'_{i-1}$.

Let us first analyze case a).

Denote by $\mathcal{D}_m$ the set of sequents $\mathcal{G}_{A(\vec{a},b),\vec{n},m}$. Then we know from the inductive assumption that for every $m < t(\vec{n})$ and every $(\longrightarrow \Delta_{m+1}) \in \mathcal{D}_{m+1}$, the sequent $\longrightarrow \Delta_{m+1}, \Gamma_1, \ldots, \Gamma_s$ has an efficient F'_{i-2}-proof from the axioms $\mathrm{Def}_\alpha, \mathrm{Def}_\beta, \mathcal{G}_1, \ldots, \mathcal{G}_r, \mathcal{D}_m$. Appending to the succedents of all sequents in this proof $\Gamma_1, \ldots, \Gamma_s$, we will construct F'_{i-2}-proofs

$$\mathrm{Def}_\alpha, \mathrm{Def}_\beta, \mathcal{G}_1, \ldots, \mathcal{G}_r, \{\longrightarrow \Delta_m, \Gamma_1, \ldots, \Gamma_s \mid (\longrightarrow \Delta_m) \in \mathcal{D}_m\} \vdash\longrightarrow \Delta_{m+1}, \Gamma_1, \ldots, \Gamma_s.$$

Now we combine these proofs together and get a polynomially time constructible proof

$$\mathrm{Def}_\alpha, \mathrm{Def}_\beta, \mathcal{G}_1, \ldots, \mathcal{G}_r, \{\longrightarrow \Delta_0, \Gamma_1, \ldots, \Gamma_s \mid (\longrightarrow \Delta_0) \in \mathcal{D}_0)\} \vdash\longrightarrow \Delta_{t(\vec{n})}, \Gamma_1, \ldots, \Gamma_s$$

for every $(\longrightarrow \Delta_{t(\vec{n})}) \in \mathcal{D}_{t(\vec{n})}$. This completes the analysis of the induction rule in the case when $A(\vec{a},0)$ is on the list $A_1, \ldots, A_k, B_{\ell+1}, \ldots, B_s$ in (8), and $A(\vec{a},t(\vec{a}))$ is on the list $B_1, \ldots, B_\ell, A_{k+1}, \ldots, A_r$.

In the remaining case b), A is in $\mathcal{S}E'_{i-1}$. This implies that $\mathcal{D}_m$ consists of a single sequent $(\longrightarrow \Delta_m)$, and we have already constructed above a proof

$$\mathrm{Def}_\alpha, \mathrm{Def}_\beta, \mathcal{G}_1, \ldots, \mathcal{G}_r, (\longrightarrow \Delta_0) \vdash\!\!\longrightarrow \Delta_{t(\vec{n})}, \Gamma_1, \ldots, \Gamma_s. \tag{10}$$

Let $\bar{\mathcal{D}}_m \rightleftharpoons \mathcal{G}_{\bar{A}(\vec{a},b),\vec{n},m}$. Then, depending on which one of the two lists in (8) contains the formulae $A(\vec{a},0), A(\vec{a},t(\vec{a}))$, we have to construct efficiently either a proof

$$\mathrm{Def}_\alpha, \mathrm{Def}_\beta, \mathcal{G}_1, \ldots, \mathcal{G}_r, (\longrightarrow \Delta_0), \bar{\mathcal{D}}_{t(\vec{n})} \vdash\!\!\longrightarrow \Gamma_1, \ldots, \Gamma_s$$

or proofs

$$\mathrm{Def}_\alpha, \mathrm{Def}_\beta, \mathcal{G}_1, \ldots, \mathcal{G}_r \vdash\!\!\longrightarrow \bar{\Delta}_0, \Delta_{t(\vec{n})}, \Gamma_1, \ldots, \Gamma_s$$

for all $(\longrightarrow \bar{\Delta}_0) \in \bar{\mathcal{D}}_0$. These modifications of (10) are easily obtained using Statement 5.3.

This completes the proof of Statement 5.4.■

In order to get Lemma 5.1 for the remaining case $T = I\mathcal{S}E_i$, we only have to apply Statement 5.4 with $k := r := 1,\ s := 0,\ A_1(\vec{a}) \rightleftharpoons \forall \vec{x} \leq \vec{t}(\vec{a})(A(\vec{a},\vec{x},\alpha) \vee B(\vec{a},\vec{x},\beta))$ (for $i > 2$ notice that axioms (6) imply $\{A_1(\vec{a})\}_{\vec{n}}$ via one application of ($\wedge$:right)). Thus, the proof of Lemma 5.1 is also completed.■

Now we are ready to finish the proof of Theorem 3.3. Recall that we have an **NP**-pair (U, V) representable in T, and let $A(a,\alpha), B(a,\beta), C(a,b,\alpha), D(a,b,\beta)$ be the corresponding formulae from Definition 3.1. Then

$$\begin{aligned} T \ \vdash\ \ & \exists x \leq a+1 \exists y \leq 1 \Big((x < a \wedge y = 0 \wedge C(a,x,\alpha) \wedge \neg D(a,x,\beta)) \\ & \vee (x < a \wedge y = 1 \wedge \neg C(a,x,\alpha) \wedge D(a,x,\beta)) \\ & \vee (x = a \wedge \neg A(a,\alpha)) \vee (x = a+1 \wedge \neg B(a,\beta)) \Big). \end{aligned}$$

We apply to this proof Lemma 5.1 and find, within (quasi)polynomial in N time a propositional proof P_N

$$\begin{aligned} & \mathrm{Def}_\alpha, \mathrm{Def}_\beta, p_{A(a),N}, q_{B(a),N}, (\longrightarrow \bar{p}_{C(a,b),N,i}, q_{D(a,b),N,i})\ (i < N), \\ & \quad (\longrightarrow p_{C(a,b),N,i}, \bar{q}_{D(a,b),N,i})\ (i < N) \vdash\!\!\longrightarrow \end{aligned}$$

in the corresponding system F'_{i-2}, F or EF. Let $t(N)$ be the size of P_N, and let $\mathrm{Def}'_{\alpha,N}$ be the CNF which is obtained by taking sequents in Def_α actually used as axioms in P_N, and moving their antecedents to the right-hand side with the ($\neg$:right) rule.

Now we are ready to describe the reduction from (U, V) to $(SAT^*, REF(P_T))$. Namely, this reduction takes a binary string $w = (w_0 w_1 \dots w_{N-1})$ of length N to $\langle \phi(w), 1^{t(N)} \rangle$, where $\phi(w)$ is the CNF obtained from $\mathrm{Def}'_{\alpha,N}$ by applying to it the restriction ρ_w assigning $p_{A(a),N}$ to 1 and assigning all $p_{C(a,b),N,i}$ to w_i $(i < N)$.

Assume that $w \in U$. Then, by Definition 3.1 a), there exists $\boldsymbol{\alpha} \subseteq \mathbf{N}$ such that $\mathbf{N} \models A(N, \boldsymbol{\alpha})$ and for every $i < N$, $\mathbf{N} \models C(N, i, \boldsymbol{\alpha}) \equiv w_i = 1$. The total assignment of ps which sends every $p_{E(\vec{a},\alpha),\vec{n}}$ to 1 if $\mathbf{N} \models E(\vec{n}, \boldsymbol{\alpha})$ and to 0 otherwise, satisfies $\mathrm{Def}'_{\alpha,N}$ and extends ρ_w. Thus, $\phi(w) \in SAT$.

Assume that $w \in V$, and take $\boldsymbol{\beta} \subseteq \mathbf{N}$ so that $\mathbf{N} \models B(N, \boldsymbol{\beta})$ and for every $i < N$, $\mathbf{N} \models D(N, i, \boldsymbol{\beta}) \equiv w_i = 1$. Hit the proof P_N with the restriction which extends ρ_w by additionally sending every $q_{E(\vec{a},\beta),\vec{n}}$ to 1 if $\mathbf{N} \models E(\vec{n}, \boldsymbol{\beta})$ and to 0 otherwise. This restriction assigns the same values to $p_{C(a,b),N,i}$ and $q_{C(a,b),N,i}$, hence it forces to 1 all axioms of P except for, possibly, those in $\mathrm{Def}'_{\alpha,N}$. Thus we get a proof of the empty sequent from the clauses of $\phi(w)$, and its size is at most $t(N)$. For the first-order case we additionally note that every F'_{i-2}-proof becomes an F_{i-2}-proof if we assign truth values to *all* q-variables. Hence $\langle \phi(w), 1^{t(N)} \rangle \in REF(P_T)$.

This completes the proof of Theorem 3.3.

6. Application to independence results

The purpose of this section is to recast one approach to proving independence results in Bounded Arithmetic in purely complexity terms.

Let us fix an integer-valued superpolynomially-growing function $t(n)$ computable in time $2^{O(n)}$. Denote by $SIMPLE_t$ the language consisting of truth-tables of those Boolean functions f_n which have circuit size at most $t(n)$, where n is the number of variables of f_n. Obviously, $SIMPLE_t \in \mathbf{NP}$. It turns out that the *computational* hardness of this language to a certain extent captures the hardness of *proving* lower bounds on the circuit size of explicit functions.

For example, in [12] Razborov and Rudich introduced the notion of a natural proof justified by a careful analysis of existing proofs for restricted models. This notion can be reformulated in terms of purely structural properties of $SIMPLE_t$: a natural proof (against the class $\mathbf{P}/poly$) consists of a set $L \in \mathbf{P}$ such that $L \cap SIMPLE_t = \emptyset$ for some superpolynomial function $t(n)$, and L is "dense" in the sense that $\mathbf{P}[\boldsymbol{f}_n \in L] \geq 2^{-O(n)}$, where $\boldsymbol{f}_n$ is the random function in n variables. The main result from [12] says that if there exists a pseudo-random number generator with hardness $2^{n^{\Omega(1)}}$ then there exists no L with these properties even in $\mathbf{P}/poly$ (and it was observed in [11] that this further extends to sets L computable by *quasipolynomial* size circuits).

Let $s = \{ s_n \mid n \in \omega \}$ be any sequence of Boolean functions from the class $\mathbf{E}$ ($= DTIME(2^{O(n)})$). We define $SIMPLE_t^{\oplus s}$ as the language

$$\{ f_n \oplus s_n \mid n \in \omega,\ f_n \in SIMPLE_t \}.$$

Note that $SIMPLE_t^{\oplus s}$ is in $\mathbf{NP}$.

If $SIMPLE_t \cap SIMPLE_t^{\oplus s} = \emptyset$ then, in particular, $s_n \not\in SIMPLE_t$ for all n. On the other hand, if $SIMPLE_t \cap SIMPLE_t^{\oplus s} \neq \emptyset$, and f_n belongs to the intersection, then we can combine the two size-$t(n)$ circuits for f_n and $f_n \oplus s_n$ with a single $PARITY$ gate at the top to get a size-$O(t(n))$ circuit for s_n. This means that, roughly speaking, the function s is hard if and only if $SIMPLE_t \cap SIMPLE_t^{\oplus s} = \emptyset$.

Let now T be one of the theories of Bounded Arithmetic considered in this paper. We additionally assume that the function $t^\sharp$ given by $t^\sharp(N) \rightleftharpoons t(|N|)$ and the predicate $S^\sharp(N, a) \rightleftharpoons s_{|N|}(a) = 1$ can be defined by bounded formulae of the underlying language. Let $LB_{t,s}(N, \gamma)$ be a $\Sigma_0^{1,b}$-formula asserting that γ does *not* encode a circuit of size $t(|N|) = t^\sharp(N)$ computing $s_{|N|}$ (our $LB_{t,s}(N, \gamma)$ corresponds to $LB(t^\sharp, s^\sharp, \gamma)$ in the notation of [11]). Thus, $\forall \phi LB_{t,s}(2^n \dot- 1, \phi)$ exactly expresses the fact $s_n \not\in SIMPLE_t$. Let $SLB_{t,s}(N, \alpha, \beta)$ assert that α and β do not encode circuits of size $t(|N|)$ each such that the $PARITY$ of their outputs is $s_{|N|}$. Thus, $\forall \phi \forall \psi SLB_{t,s}(2^n \dot- 1, \phi, \psi)$ means that $SIMPLE_t \cap SIMPLE_t^{\oplus s} = \emptyset$. Since the argument from the above paragraph is easy to formalize, we can study the provability of $SLB_{t,s}(N, \alpha, \beta)$ instead of $LB_{t,s}(N, \gamma)$ (and the split versions were designed in [11] exactly for this purpose). Given Theorem 3.3, we can now reduce the question about provability of $SLB_{t,s}(N, \alpha, \beta)$ in T to the purely complexity question

$$(SIMPLE_t, SIMPLE_t^{\oplus s}) \overset{?}{\leq}_m (SAT^*, REF(P_T)), \tag{11}$$

where $\leq_m$ is the appropriate reducibility.

The following easy result (implicit in [11, Proof of Theorem 6.1]) shows that this complexity question is at least not meaningless:

Proposition 6.1. *If there exists a pseudo-random number generator with hardness $2^{n^{\Omega(1)}}$ then for any t, s with the above properties the pair $(SIMPLE_t, SIMPLE_t^{\oplus s})$ can not be separated by quasipolynomial size circuits.*

Proof. Assume that $E = \left\{ E_n \subseteq \{0,1\}^{(2^n)} \mid n \in \omega \right\}$ is such a separator: $SIMPLE_t \subseteq E$, $E \cap SIMPLE_t^{\oplus s} = \emptyset$. Then for any n either $|E_n| \geq \frac{1}{2} \cdot 2^{2^n}$ or $|E_n| \leq \frac{1}{2} \cdot 2^{2^n}$. In the first case we let $L_n \rightleftharpoons (E_n \oplus s_n)$, and in the second case we let $L_n \rightleftharpoons \{0,1\}^{(2^n)} \setminus E_n$. Then $L \rightleftharpoons \bigcup_{n\in\omega} L_n$ is computable by quasipolynomial size circuits since one extra bit of information telling us which of the two cases takes place can be hardwared into the circuit. Also, $L \cap SIMPLE_t = \emptyset$ and $\mathbf{P}[\boldsymbol{f}_n \in L] \geq 1/2$. As we noticed above, this contradicts the main result from [12].■

For completeness we also include an unconditional form of this proposition based upon [12, Theorem 4.4]. Recall [12] that a non-decreasing integer-valued function $t(n)$ is *half-exponential* if

$$t^{-1}(n^C) \leq o(\log t(n))$$

for every $C > 0$, where

$$t^{-1}(n) \rightleftharpoons \max\{\, x \mid t(x) \leq n \,\}.$$

It is easy to see that any half-exponential function has superpolynomial rate of growth.

Let us call $t(n)$ *strongly half-exponential* if it satisfies

$$t^{-1}(n^C) \leq (\log t(n))^{o(1)}$$

for every $C > 0$.

Theorem 6.2. *Let $t(n)$ be any half-exponential function, and $s = \{\, s_n \mid n \in \omega \,\}$ be such that for some sequence of primes $\{\, p_n \mid n \in \omega \,\}$ and some primitive roots g_n mod p_n, s_n is poly-time nonuniformly Turing reducible to computing discrete logarithm mod p_n base g_n. Then there is no $E \in \mathbf{P}$ such that $SIMPLE_t \subseteq E$ and $SIMPLE_t^{\oplus s} \cap E = \emptyset$. Moreover, if $t(n)$ is strongly half-exponential, then no such E exists even in $\mathbf{QP}$.*

Proof. Assuming the contrary, we, like in the previous proof, would have a natural proof $L \in \mathbf{P}/poly$ with the additional property $s_n \in L$ for all $n \in \omega$. It can not exist (without any unproven assumptions!) by [12, Theorem 4.4]. It is also easy to see that if $t(n)$ is strongly half-exponential then [12, Theorem 4.4] extends to L computable by quasipolynomial size circuits.■

Proposition 6.1 and Theorem 6.2 show that in order to prove the independence of $SLB_{t,s}(N, \alpha, \beta)$ from a theory T, it is sufficient to separate the pair $(SAT^*, REF(P_T))$ by a (quasi)polynomial time computable set. We conclude this section by showing another proof of the main result from [11] which goes exactly along these lines.

Lemma 6.3. *If a pair (U, V) of disjoint $\mathbf{NP}$-sets is representable in $\mathcal{S}IE_1$ $[\mathcal{S}S_2^2]$ then there exists a constant $w > 0$ such that $(U, V) \leq_m^p (SAT^*, REF(R_w))$
$[(U, V) \leq_m^{qp} (SAT^*, REF(R_w))$, respectively$]$.*

Proof. By modifying the proof of Lemma 5.1 for the case $\mathcal{S}IE_2$. Namely, we replace the axioms (4),(5) by

$$\begin{gathered}
p_{(\exists x \leq a)A(x,\vec{b}),0,\vec{m}} \longrightarrow p_{A(a,\vec{b}),0,\vec{m}} \\
p_{(\exists x \leq a)A(x,\vec{b}),n+1,\vec{m}} \longrightarrow p_{(\exists x \leq a)A(x,\vec{b}),n,\vec{m}}, p_{A(a,\vec{b}),n+1,\vec{m}} \\
p_{A(a,\vec{b}),0,\vec{m}} \longrightarrow p_{(\forall x \leq a)A(x,\vec{b}),0,\vec{m}} \\
p_{(\forall x \leq a)A(x,\vec{b}),n,\vec{m}}, p_{A(a,\vec{b}),n+1,\vec{m}} \longrightarrow p_{(\forall x \leq a)A(x,\vec{b}),n+1,\vec{m}}
\end{gathered}$$

so that all sequents in $\mathrm{Def}_\alpha, \mathrm{Def}_\beta$ have bounded length. The important point is that if we can deduce $(n+1)$ sequents $\Gamma \longrightarrow p_{A(a,\vec{b}),0,\vec{m}}, \Delta; \ldots; \Gamma \longrightarrow p_{A(a,\vec{b}),n,\vec{m}}, \Delta$ in R_w then we can deduce $\Gamma \longrightarrow p_{(\forall x \leq a)A(x,\vec{b}),n,\vec{m}}, \Delta$ in $R_{w'}$ for some w' depending only on w, and similarly for $\Gamma, p_{(\exists x \leq a)A(x,\vec{b}),n,\vec{m}} \longrightarrow \Delta$. For $C(\vec{a}) \in \mathcal{S}E_0'$, the cedent $\Gamma_{C(\vec{a}),\vec{n}}$ in our case always consists of the single formula $\{C(\vec{a})\}_{\vec{n}}$.

With these observations in mind, it is easy to see that the procedure described in the proof of Statement 5.4 for $i = 2$, actually gives in the case $i = 1$ a resolution proof in which the length of all clauses is bounded by some absolute constant (depending on the original proof P in $\mathcal{S}IE_1$). The only additional remark which should be made is that the "bad" rules ($\exists \leq$:left), ($\exists \leq$:right) now simply do not occur in the proof.■

Lemma 6.4. *For every fixed constant $w > 0$, SAT^* and $REF(R_w)$ can be separated by a poly-time computable set.*

Proof. The separator is

$$\left\{ \langle \phi, 1^t \rangle \mid \text{there is no derivation of the empty sequent from } \phi \text{ in the system } R_w \right\}.$$

It is poly-time computable simply by producing the list of all sequents of length at most w which can be derived from ϕ.■

Theorem 6.5. *A disjoint* **NP***-pair is representable in $\mathcal{S}IE_1$ $[\mathcal{S}_2^2]$ if and only if it can be separated by a polynomial [quasipolynomial, respectively] time computable set.*

Proof. Immediate from Theorem 3.2, Lemma 6.3 and Lemma 6.4.■

The first part of the following theorem is exactly [11, Theorem 6.4]:

Theorem 6.6. *If there exists a pseudo-random number generator with hardness $2^{n^{\Omega(1)}}$, then for any t, s with the properties stated at the beginning of this section,*

$$\mathcal{S}S^2_2 \not\vdash SLB_{t,s}(N, \alpha, \beta).$$

If, in addition, t is half-exponential [strongly half-exponential], and s is reduced to the discrete logarithm problem as described in the statement of Theorem 6.2*, then $\mathcal{S}IE_1 \not\vdash SLB_{t,s}(N, \alpha, \beta)$ [$\mathcal{S}S^2_2 \not\vdash SLB_{t,s}(N, \alpha, \beta)$, respectively] without any unproven assumptions.*

Proof. Immediate from Theorem 6.5, Proposition 6.1 and Theorem 6.2.■

7. Discussion

This paper brings to attention the question for which propositional proof systems P the pair $(SAT^*, REF(P))$ can be separated by a (quasi)polynomial time computable set. In this section we try to locate this question with respect to more familiar hypothesis.

Let us first point out that the affirmative answer implies the following alternative:

Theorem 7.1. *Assume that for some proof system P, SAT^* and $REF(P)$ can be separated by a poly-time computable set. Then one of the following is true:*

a) $\mathbf{P} = \mathbf{NP}$,

b) *the proof system P is not optimal in the sense that the function*

$$s_P(n) \rightleftharpoons \max\{\, s_P(\phi) \mid \phi \text{ is an unsatisfiable CNF of length } \le n\}$$

is not bounded by any polynomial.

Proof. Let $SAT^* \subseteq L$; $L \cap REF(P) = \emptyset$; $L \in \mathbf{P}$, and assume that b) does not take place. Then $s_P(n) \le p(n)$ for some polynomial p, and $\phi \in SAT \equiv \langle \phi, 1^{p(|\phi|)} \rangle \in L$. Thus, $SAT \in \mathbf{P}$.■

This theorem might be taken as an evidence that any attempts to prove the existence of the separator by known methods are doomed to fail. We should be, however, somewhat careful with this conclusion. For example, the proof of Lemma 6.4, whatever simple, still *does not tell us which of the two alternatives* a) *and* b) *is true for the system R_w.* Of course, we know that b) is true, and, moreover, R_w is not even complete – but this has to be proved separately. Thus, simply knowing that either a) or b) is true might be surprising approximately to the same extent as knowing that one of the two alternatives **LOGSPACE** $\neq$ **P** or **P** $\neq$ **PSPACE** is true.

But, of course, we can not hope to show by the existing methods that $(SAT^*, REF(P))$ (as well as any other disjoint **NP**-pair) is *not* separable. So, if we are interested in evidence toward the negative solution, the best we can hope for is to reduce to $(SAT^*, REF(P))$ another pair which is believed to be hard.

I do not know of any example of a reduction from a presumably hard **NP**-pair to $(SAT^*, REF(EF))$, which is the same, due to our main result, as an example of such pair representable in V_1^1.

There is, however, a number of "plain" reductions from (U, V) to $(SAT^*, REF(EF))$, where (U, V) is separable but this fact is highly non-trivial. The best example of this kind (in the sense that it is applicable to the weakest system P) is provided by [11, Example 1]. Namely, let $CHR \rightleftharpoons \{< G, s > \mid G \text{ is an } s-\text{colourable graph}\}$, and $CL^2 \rightleftharpoons \{< G, s > \mid G \text{ contains a clique of size } s^2\}$. Then (CHR, CL^2) is representable in $\mathcal{S}T_2^3$ and, thus, $(CHR, CL^2) \leq_m^{qp} (SAT^*, REF(F_1))$. On the other hand, the known poly-time computable separator for (CHR, CL^2) is based upon very deep combinatorial ideas [16].

I do not know of any evidence of this sort that $(SAT^*, REF(R))$ is hard. This could be the next accessible question.

8. Acknowledgement

I am indebted to Jan Krajíček for his initial suggestion to look for a propositional counterpart of the machinery from [11]. My thanks are also due to Søren Riis and Alan Selman for several useful remarks.

References

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

[2] S. R. Buss. Axiomatizations and conservations results for fragments of Bounded Arithmetic. In *Logic and Computation, Contemporary Mathematics* 106, pages 57–84. American Math. Society, 1990.

[3] S. A. Cook. Feasibly constructive proofs and the propositional calculus. In *Proceedings of the 7th Annual ACM Symposium on the Theory of Computing*, pages 83–97, 1975.

[4] S. A. Cook and A. R. Reckhow. The relative efficiency of propositional proof systems. *Journal of Symbolic Logic*, 44(1):36–50, 1979.

[5] J. Grollmann and A. L. Selman. Complexity measures for public-key cryptosystems. *SIAM Journal on Computing*, 17(2):309–335, April 1988.

[6] S. Homer and A. L. Selman. Oracles for structural properties: The isomorphism problem and public-key cryptography. *Journal on Computer and System Sciences*, 44(2):287–301, April 1992.

[7] J. Krajíček. On Frege and extended Frege proof systems. Manuscript, 1993.

[8] J. Paris and A. Wilkie. Counting problems in bounded arithmetic. In *Methods in Mathematical Logic, Lecture Notes in Mathematics 1130*, pages 317–340. Springer-Verlag, 1985.

[9] A. Razborov. An equivalence between second order bounded domain bounded arithmetic and first order bounded arithmetic. In P. Clote and J. Krajíček, editors, *Arithmetic, Proof Theory and Computational Complexity*, pages 247–277. Oxford University Press, 1992.

[10] A. Razborov. Bounded Arithmetic and lower bounds in Boolean complexity. To appear in the volume *Feasible Mathematics II*, 1993.

[11] A. Razborov. Unprovability of lower bounds on circuit size in certain fragments of Bounded Arithmetic. To appear in *Izvestiya of the RAN*, 1994.

[12] A. Razborov and S. Rudich. Natural proofs. Preliminary version appeared in *Proceedings of the 26th ACM Symposium on Theory of Computing, pp.* 204-213, 1994.

[13] A. L. Selman. Complexity issues in cryptography. *Proceedings of Symposia in Applied Mathematics*, 38:92–107, 1989.

[14] G. Takeuti. S_3^i and $\overset{\circ}{V}{}_2^i(BD)$. *Archive for Math. Logic*, 29:149–169, 1990.

[15] G. Takeuti. *RSUV* isomorphisms. In P. Clote and J. Krajíček, editors, *Arithmetic, Proof Theory and Computational Complexity*, pages 364–386. Oxford University Press, 1992.

[16] É. Tardos. The gap between monotone and nonmonotone circuit complexity is exponential. *Combinatorica*, 8:141–142, 1988.

[17] G. Wilmers. Bounded existential induction. *The Journal of Symbolic Logic*, 50(1):72–90, March 1985.

[18] Г. С. Цейтин. О сложности вывода в исчислении высказываний. In А. О. Слисенко, editor, *Исследования по конструктивной математике и математической логике,* II*; Записки научных семинаров ЛОМИ, т. 8*, pages 234–259. Наука, Ленинград, 1968. Engl. translation: G. C. Tseitin, On the complexity of derivations in propositional calculus, in: *Studies in mathematics and mathematical logic, Part II*, ed. A. O. Slissenko, pp. 115-125.

Recent Publications in the BRICS Report Series

RS-94-36 Alexander A. Razborov. *On provably disjoint* NP*-pairs*. November 1994. 27 pp.

RS-94-35 Gerth Stølting Brodal. *Partially Persistent Data Structures of Bounded Degree with Constant Update Time*. November 1994. 24 pp.

RS-94-34 Henrik Reif Andersen, Colin Stirling, and Glynn Winskel. *A Compositional Proof System for the Modal μ-Calculus*. October 1994. 18 pp. Appears in: Proceedings of LICS '94, IEEE Computer Society Press.

RS-94-33 Vladimiro Sassone. *Strong Concatenable Processes: An Approach to the Category of Petri Net Computations*. October 1994. 40 pp.

RS-94-32 Alexander Aiken, Dexter Kozen, and Ed Wimmers. *Decidability of Systems of Set Constraints with Negative Constraints*. October 1994. 33 pp.

RS-94-31 Noam Nisan and Amnon Ta-Shma. *Symmetric Logspace is Closed Under Complement*. September 1994. 8 pp.

RS-94-30 Thore Husfeldt. *Fully Dynamic Transitive Closure in Plane Dags with one Source and one Sink*. September 1994. 26 pp.

RS-94-29 Ronald Cramer and Ivan Damgård. *Secure Signature Schemes Based on Interactive Protocols*. September 1994. 24 pp.

RS-94-28 Oded Goldreich. *Probabilistic Proof Systems*. September 1994. 19 pp.

RS-94-27 Torben Braüner. *A Model of Intuitionistic Affine Logic from Stable Domain Theory (Revised and Expanded Version)*. September 1994. 19 pp. Full version of paper appearing in: ICALP '94, LNCS 820, 1994.

RS-94-26 Søren Riis. *Count(q) versus the Pigeon-Hole Principle*. August 1994. 3 pp.