Interpolation theorems, lower bounds for proof systems, and independence results for bounded arithmetic

Jan Krajíček*

Mathematical Institute of the Academy of Sciences
Žitná 25, Praha 1, 115 67, Czech Republic
krajicek@earn.cvut.cz

Abstract

A proof of the (propositional) Craig interpolation theorem for cut-free sequent calculus yields that a sequent with a cut-free proof (or with a proof with cut-formulas of restricted form; in particular, with only analytic cuts) with k inferences has an interpolant whose circuit-size is at most k. We give a new proof of the interpolation theorem based on a communication complexity approach which allows a similar estimate for a larger class of proofs. We derive from it several corollaries:

1. Feasible interpolation theorems for the following proof systems:
   (a) resolution.
   (b) a subsystem of LK corresponding to the bounded arithmetic theory $S^2_2(\alpha)$.
   (c) linear equational calculus.
   (d) cutting planes.
2. New proofs of the exponential lower bounds (for new formulas)
   (a) for resolution ([15]).
   (b) for the cutting planes proof system with coefficients written in unary ([4]).
3. An alternative proof of the independence result of [43] concerning the provability of circuit-size lower bounds in the bounded arithmetic theory $S^2_2(\alpha)$.

*1991 Mathematics Subject Classification. Primary 03F20, 03B05, 03F30; Secondary 68Q25. Partially supported by the *US - Czechoslovak Science and Technology Program* grant # 93025, and by the grant #119107 of the *AV ČR*.

In the other direction we show that a depth 2 subsystem of LK does not admit feasible monotone interpolation theorem (the so called Lyndon theorem), and that a feasible monotone interpolation theorem for the depth 1 subsystem of LK would yield new exponential lower bounds for resolution proofs of the weak pigeonhole principle.

Introduction

The interpolation theorem proved by Craig [11, 12] is a basic result in logic. It says that whenever an implication

$$A \longrightarrow B$$

is valid then there is a third formula I, an *interpolant*, which contains only those symbols of the language occurring in both A and B and such that the two implications

$$A \longrightarrow I \qquad I \longrightarrow B$$

are both valid. The theorem holds for propositional logic as well as for the first order logic but we shall confine our attention to propositional logic in this paper.

The question of finding an interpolant for the implication is quite relevant to computational complexity theory. To see this let U and V be two disjoint $\mathcal{NP}$-subsets $\{0,1\}^*$. It is well known that there are sequences of propositional formulas $A_n(p_1,\ldots,p_n,q_1,\ldots,q_{t_n})$ and $B_n(p_1,\ldots,p_n,r_1,\ldots,r_{s_n})$ such that the size of A_n and B_n is $n^{O(1)}$ and

$$U \cap \{0,1\}^n \;=\; \{(\epsilon_1,\ldots,\epsilon_n) \in \{0,1\}^n \mid \exists \alpha_1,\ldots,\alpha_{t_n}\ A_n(\overline{\epsilon},\overline{\alpha}) \text{ holds}\}$$

and

$$V \cap \{0,1\}^n \;=\; \{(\epsilon_1,\ldots,\epsilon_n) \in \{0,1\}^n \mid \exists \beta_1,\ldots,\beta_{s_n}\ B_n(\overline{\epsilon},\overline{\beta}) \text{ holds}\}\ .$$

The assumption that $U \cap V = \emptyset$ is equivalent to the statement that the implications

$$A_n \longrightarrow \neg B_n$$

are all tautologically valid. If $I_n(\overline{p})$ is an interpolant (hence only atoms $p_1,\ldots,p_n$ occur in I_n) then the set

$$W := \bigcup_n \{\overline{\epsilon} \in \{0,1\}^n \mid I_n(\overline{\epsilon}) \text{ holds }\}$$

separates U from V:

$$U \subseteq W \text{ and } W \cap V = \emptyset\ .$$

Hence an estimate of the complexity of propositional interpolation formulas in terms of the complexity of an implication yields an estimate to the computational complexity of a set separating U from V. For example, if one could always find such an interpolant whose formula-size (or circuit-size; recall that a circuit-size of a formula I is the number of different subformulas occurring in it) is polynomial in the size of the implication then $\mathcal{NP} \cap co\mathcal{NP} \subseteq \mathcal{NC}^1/poly$ (or $\mathcal{NP} \cap co\mathcal{NP} \subseteq \mathcal{P}/poly$). This is because for $U \in \mathcal{NP} \cap co\mathcal{NP}$ we may take for V the complement of U and hence it must hold that $W = U$. This example can be understood as a conditional lower bound to the size of interpolants; it was first noted by Mundici [30, 31, 32]. For predicate logic there are lower bounds in terms of recursion theory, see [28, 13], for other connections to computer science see [14].

The question we shall study in this paper is a bit different.

> **Problem:** *Given a propositional proof system, estimate the circuit-size of an interpolant of an implication in terms of the size of the shortest proof of the implication.*

Presumably one gets different estimates for different proof systems and, in particular, not all proof systems should admit polynomial upper bounds. However, this is an open problem. The proof of Craig interpolation theorem [11, 12] via cut-elimination (see, for example, [45] or [20, 4.3]) shows that an implication whose cut-free proof in the sequent calculus has k steps has an interpolant with circuit-size at most k.

The reason for studying this problem is that a good *upper bound* for a proof system P yields *lower bounds* on the size of P-proofs. In particular, a pair of $\mathcal{NP}$-sets U and V inseparable by a set of small complexity yields a sequence of implications $A_n \longrightarrow \neg B_n$ which cannot have short P-proofs (as the assumed good interpolation yields feasible upper bound to the complexity of I_n and hence of W). This idea was discussed in Krajíček [19] but no lower bounds were obtained there in this way.

Our interest in this question was renewed by a remark in Razborov [43] that the results underlying the unprovability results there are certain interpolation theorems for fragments of second order bounded arithmetic. It occurred to us that these interpolation theorems (and problems) are more rudimentary in the propositional setting, and that a sufficiently sharp estimate to the complexity of the interpolation theorem for resolution - together with the known relations of propositional proof systems to bounded arithmetic theories - might yield an alternative proof of the main result of [43]. We prove such an interpolation theorem (in fact, a polynomial bound for resolution follows already from the bound for cut-free systems via a translation of resolution refutations into cut-free derivations, see 6.1(second proof)). In fact, we give a new proof of the Craig interpolation theorem (as well as of the Lyndon version) allowing us to deduce in a new way exponential lower bounds to the size of proofs in various systems (a subsystem of LK stronger than the cut-free fragment, resolution, a version of cutting planes). We formulate a general, syntax-free, framework for which our proof of the interpolation theorem yields good bounds.

The paper is organized as follows. In the first section we define several proof systems (sequent calculus, resolution, linear equational calculus and cutting planes). In the second section we recall some communication complexity (Karchmer-Wigderson game) and we reformulate a bit the characterization of the circuit-size in terms of PLS-problems from [43]. In the third section we give a new proof of Craig interpolation theorem for cut-free sequent calculus. The proof applies in a general, syntax-free, context. This is formalized by the notion of *semantic derivations* defined in section 4. A general form of the interpolation theorem for semantic derivations is proved in section 5. In section 6 we deduce from it polynomial upper bounds for interpolation for resolution, a subsystem of sequent calculus relevant to bounded arithmetic, linear equational calculus and a variant of cutting planes. In section 7 we obtain new proofs of exponential lower bounds for some of these systems and in section 8 we give an alternative treatment of the proof of the main independence result of [43].

A question for which proof systems one can prove a non-trivial *lower bound* for interpolation is discussed in section 9. It is linked with two topics, the existence of optimal propositional proof systems relative to a given theory (studied earlier in [22]) and the question of implicit definability of inverse functions to one-way functions. We also prove there that the depth 2 subsystem of LK does not admit feasible monotone interpolation theorem and that the validity of such a theorem for the depth 1 subsystem would imply new exponential lower bounds to the resolution proofs of the weak pigeonhole principle.

The reader is assumed to have some familiarity with the subjects involved, in particular with some basic notions of complexity theory. A familiarity with bounded arithmetic is assumed only in the last two sections. References to original papers are often accompanied by a reference to a place in [20] which offers a survey of basic results in the field.

A remark on notation: we denote n-tuples of numbers or bits simply $a, b, x, y, \dots$ rather than $\overline{a}, \dots$, and the elements or the bits of a are denoted $a_1, a_2, \dots$. Logarithm log is base 2.

1 Propositional proof systems

The propositional language of the *sequent calculus* LK contains the following connectives: constants 0 (false) and 1 (true), the negation $\neg$, the conjunction $\bigwedge$ and the disjunction $\bigvee$. The negation is allowed only in front of atoms, the conjunction and the disjunction are of unbounded arity. The symbol $\neg A$ denotes the formula obtained from the formula A by interchanging 0 and 1, $\bigvee$ and $\bigwedge$ and p_i and $\neg p_i$.

The *size* $|A|$ of A is the number of occurrences of connectives and atoms in it. The *depth* $dp(A)$ of A is the maximal nesting of $\bigvee$ and $\bigwedge$ in A:

$$dp(0) = dp(1) = dp(p_i) = dp(\neg p_i) = 0$$

$$dp(\bigvee_i A_i) = dp(\bigwedge_i A_i) = 1 + max_i(dp(A_i)).$$

We shall adopt the following version of the sequent calculus LK. The particular modification is unimportant and used just for technical reasons, similarly as in [19]. We shall keep the name LK as well. Further information on LK can be found in [45] or [20, Section 4.3] (contains also information about resolution, Section 4.2, and cutting planes, Section 13.1).

A *cedent* is is a finite (possibly empty) sequence of formulas denoted $\Gamma, \Delta, \ldots$. The basic object of LK is a *sequent*, an ordered pair of cedents written $\Gamma \longrightarrow \Delta$. A sequent is satisfied if at least one formula in Δ is satisfied or at least one formula in Γ is falsified. In particular, the empty sequent cannot be satisfied.

The inference rules are the following:

1. the *initial sequents* are:

$$\longrightarrow 1 \qquad \neg 1 \longrightarrow \qquad 0 \longrightarrow \qquad \longrightarrow \neg 0$$

$$p \longrightarrow p \qquad \neg p \longrightarrow \neg p \qquad p, \neg p \longrightarrow \qquad \longrightarrow p, \neg p$$

2. the weak structural rules are:

   the exchange:
$$\frac{\Gamma \longrightarrow \Delta}{\Gamma' \longrightarrow \Delta'}$$
   where Γ', Δ' are any permutations of Γ, Δ

   the contraction:
$$\frac{\Gamma \longrightarrow \Delta}{\Gamma' \longrightarrow \Delta'}$$
   where Γ', Δ' are obtained from Γ, Δ by deleting any multiple occurrences of formulas

   the weakening:
$$\frac{\Gamma \longrightarrow \Delta}{\Gamma' \longrightarrow \Delta'}$$
   where $\Gamma' \supseteq \Gamma$ and $\Delta' \supseteq \Delta$

3. *the propositional rules* are:

   $\bigwedge$:introduction

$$\frac{A, \Gamma \longrightarrow \Delta}{\bigwedge A_i, \Gamma \longrightarrow \Delta} \qquad \frac{\Gamma \longrightarrow \Delta, A_1 \quad \ldots \quad \Gamma \longrightarrow \Delta, A_m}{\Gamma \longrightarrow \Delta, \bigwedge_{i \leq m} A_i}$$

   where A is one of A_i in the left rule.

$\bigvee$:introduction

$$\frac{A_1, \Gamma \longrightarrow \Delta \quad \ldots \quad A_m, \Gamma \longrightarrow \Delta}{\bigvee_{i \leq m} A_i, \Gamma \longrightarrow \Delta} \qquad \frac{\Gamma \longrightarrow \Delta, A}{\Gamma \longrightarrow \Delta, \bigvee_i A_i}$$

where A is one of A_i in the right rule.

4. *the cut rule*:
$$\frac{\Gamma \longrightarrow \Delta, A \quad A, \Gamma \longrightarrow \Delta}{\Gamma \longrightarrow \Delta}$$

An LK-*proof* of a sequent S from the sequents $S_1, \ldots, S_m$ is a sequence $Z_1, \ldots, Z_k$ such that $Z_k = S$ and each Z_i is either an initial one or from $\{S_1, \ldots, S_m\}$, or derived from the previous ones by an inference rule.

A *proof-graph* of an LK-proof π is a directed acyclic graph whose nodes are the sequents of π and a directed edge goes from a hypothesis of a rule to its conclusion. Hence the initial sequents correspond to the leaves.

A proof is *tree-like* if its proof-graph is a forest, i.e., if every sequent is a hypothesis of at most one inference.

$k(\pi)$ is the number of sequents in π. The *size* of a proof is the sum of the sizes of the formulas in it (counting multiple occurrences of a formula separately).

A *resolution refutation* of sequents $S_1, \ldots, S_m$ which contain no $\bigvee, \bigwedge$ is an LK-proof of the empty sequent from $S_1, \ldots, S_m$ in which no $\bigvee, \bigwedge$ occur. This is obviously (essentially) equivalent to the more usual definition of resolution with clauses and the resolution rule as a resolution clause

$$\{\neg p_{i_1}, \ldots, \neg p_{i_a}, p_{j_1}, \ldots, p_{j_b}\}$$

can be represented by the sequent

$$p_{i_1}, \ldots, p_{i_a} \longrightarrow p_{j_1}, \ldots, p_{j_b}$$

and the resolution rule by the cut rule (and vice versa). We shall freely slide between the two definitions of resolution.

We define a *linear equational calculus* (LEC) to be a proof system working with linear equations
$$a_1 x_1 + \ldots + a_n x_n = b$$
over a field F. F is either a finite field or the field of rationals $\mathbf{Q}$. The rules allow to add two equations and to multiply an equation by an element of F. An LEC-refutation of equations $E_1, \ldots, E_m$ is an LEC-derivation of the equation $0 = 1$ from $E_1, \ldots, E_m$. The size of an equation is $\sum_i ||a_i|| + ||b||$ where $||\frac{u}{v}||$ is the sum of the absolute values of u and v if $F = \mathbf{Q}$ and $||a|| = 1$ for all $a \in F$ if F is finite.

LEC is sound and complete (by Gauss elimination), if by completeness we mean that every system of equations unsolvable in F is refutable. When completeness is considered only w.r.t. the systems with no 0-1 solution then LEC

is complete only for the two-element field $\mathbf{F}_2$. To get such completeness also for other fields one would have to expand LEC to an equational logic working with general polynomials and based on ring axioms.

However, not even all Boolean functions can be represented by a conjunction of linear equations and so LEC cannot be considered, even for $\mathbf{F}_2$, as a full propositional proof system in the sense of [9].

An important example of a formula which can be so represented is the negation of the pigeonhole principle, formalizing that there is a bijection between $\{1, \ldots, n+1\}$ and $\{1, \ldots, n\}$. This formula is represented by the following set of equations (over any F) with variables x_{ij}, $i = 1, \ldots n+1$ and $j = 1, \ldots, n$:

$$\sum_j x_{ij} = 1 \ , \text{ for all } i$$

$$\sum_i x_{ij} = 1 \ , \text{ for all } j \ .$$

It is easy to see that there is an LEC-refutation of this set of size polynomial in n.

A system stronger than the resolution system is the *cutting planes* proof system introduced in [10]. This system CP works with inequalities of the form $a_1 x_1 + \ldots a_n x_n \geq b$, where $a_i, b \in \mathbf{Z}$ and x_i represent truth values of atoms. CP has few obvious rules: adding two inequalities, multiplying an inequality by a positive constant, the division rule:

$$\frac{a_1 x_1 + \ldots a_n x_n \geq b}{\frac{a_1}{c} x_1 + \ldots \frac{a_n}{c} x_n \geq \lceil \frac{b}{c} \rceil}$$

provided $c | a_i$, all i, and few initial inequalities: $x \geq 0$, $-x \geq -1$. CP is a refutation system which derives from an unsatisfiable system of inequalities the inequality $0 \geq 1$. The term *unsatisfiable* means that the system has no 0-1 solution. It is sound and complete and polynomially simulates resolution, see [10] or [20, 13.1].

2 Protocols for Karchmer-Wigderson game

Karchmer-Wigderson game (see [16]) is played as follows. Let $U, V \subseteq \{0,1\}^n$ be two disjoint sets. The game is played by two players A and B. Player A receives $u \in U$ while B receives $v \in V$. They communicate bits of information (following a protocol previously agreed on) until both players agree on the same $i \in \{1, \ldots, n\}$ such that $u_i \neq v_i$. Their objective is to minimize (over all protocols) the number of bits they need to communicate in the worst case. This minimum is called the *communication complexity* of the game and it is denoted by $C(U, V)$.

We say that the Boolean function $B(p_1, \ldots, p_n)$ *separates U from V* if and only if $B(x) = 1$ holds (resp. $= 0$) for all $x \in U$ (resp. for all $x \in V$).

The following is a rather simple but quite important result.

Theorem 2.1 ([16]) *Let $U, V \subseteq \{0,1\}^n$ be two disjoint sets. Then $C(U,V)$ is precisely the minimal depth of a formula with binary $\vee, \wedge$ separating U from V.*

We shall need a bit finer version of the theorem. For that we need to define the notion of a protocol in a particular way.

Definition 2.2 *Let $U, V \subseteq \{0,1\}^n$ be two disjoint sets. A* protocol *for the game on the pair (U,V) is a labelled directed graph G satisfying the following four conditions:*

1. *G is acyclic and has one source (the in-degree 0 node) denoted $\emptyset$.*

   The nodes with the out-degree 0 are leaves, *all other are* inner nodes.

2. *All leaves are labelled by one of the following formulas:*

   $$u_i = 1 \wedge v_i = 0 \quad or \quad u_i = 0 \wedge v_i = 1$$

   for some $i = 1, \ldots, n$.

3. *There is a function $S(u,v,x)$ (the* strategy*) such that S assigns to a node x and a pair $u \in U$ and $v \in V$ the edge $S(u,v,x)$ leaving from the node x.*

Every pair $u \in U$ and $v \in V$ defines for every node x a directed path P^x_{uv} in G from the node x to a leaf: $P^x_{uv} = x_1, \ldots, x_h$, where $x_1 = x$, the edge $S(u,v,x_i)$ goes from x_i to x_{i+1}, and x_h is a leaf.

4. *For every $u \in U$ and $v \in V$ there is a set $F(u,v) \subseteq G$ satisfying:*

   (a) *$\emptyset \in F(u,v)$*

   (b) *$x \in F(u,v) \rightarrow P^x_{u,v} \subseteq F(u,v)$*

   (c) *the label of any leaf from $F(u,v)$ is valid for u, v.*

   Such a set F is called the consistency condition.

A protocol is called monotone *iff every leaf in it is labelled by one of the formulas $u_i = 1 \wedge v_i = 0$, $i = 1, \ldots, n$.*

The communication complexity of G is the minimal number t such that for every $x \in G$ the players (one knowing u and x, the other one v and x) decide whether $x \in F(u,v)$ and compute $S(u,v,x)$ with at most t bits exchanged in the worst case.

This definition is a variant of the formulation from [43] using PLS-problems. We would like to replace the consistency condition 4. by a simpler one: for all u, v the label of the leaf in $P^{\emptyset}_{u,v}$ is valid for u, v. However, the example of a linear size branching program finding $i \leq n$ such that $u_i \neq v_i$ for *all* different u, v shows that that is not enough.

Important examples of protocols are protocols formed from a circuit (later we shall define protocols from proofs). Assume that C is a circuit separating U from V. Reverse the edges in C, take for $F(u, v)$ those subcircuits differing in the value on u and v, and define the strategy and the labels of the leaves in an obvious way. This determines a protocol for the game on (U, V) whose communication complexity is 2. The next theorem says that there is a similar converse construction.

Theorem 2.3 ([43]) *Let $U, V \subseteq \{0,1\}^n$ be two disjoint sets. Let G be a protocol for the game on U, V which has k nodes and the communication complexity t.*

Then there is a circuit C of size $k2^{O(t)}$ separating U from V. Moreover, if G is monotone so is C.

On the other hand, any circuit (monotone circuit) C of size m separating U from V determines a protocol (a monotone protocol) G with m nodes whose communication complexity is 2.

Proof :

Let G be a protocol from the game. The number of nodes reachable from x via the edges defines a *cost* of x. For any u, v, the set $F(u, v)$ together with the cost function and with the *neighborhood function* given by the strategy is a PLS-problem. By [43, Thm. 3.1] there is a circuit separating U from V of size at most

$$\mid \bigcup_{u,v} F(u, v) \mid \ \cdot 2^{O(t)} = k \cdot 2^{O(t)} \ .$$

If the protocol is monotone so is the circuit.

The second part of the statement was noted above.

q.e.d.

3 The Craig interpolation theorem

We define an *interpolant* of a valid implication

$$A(p, q) \to B(p, r)$$

where $p = (p_1, \ldots, p_n)$ are the atoms occurring in both A and B, while $q = (q_1, \ldots, q_s)$ occur only in A and $r = (r_1, \ldots, r_t)$ only in B, to be any Boolean function $I(p)$ such that both implications

$$A(p, q) \to (I(p) = 1) \quad \text{and} \quad (I(p) = 1) \to B(p, r)$$

are tautologically valid. If $I(p)$ is defined by a formula (also denoted I) this means that both implications

$$A \to I \quad \text{and} \quad I \to B$$

are tautologies.

In the calculus LK the implication $A \to B$ is represented by the sequent $A \longrightarrow B$ and, in general, the sequent $A_1, \ldots, A_m \longrightarrow B_1, \ldots, B_\ell$ represents the implication $\bigwedge_i A_i \to \bigvee_j B_j$.

Craig [11, 12] proved that every tautologically valid implication has an interpolant. In fact, the argument via the cut-elimination (see [45] or [20, 4.3]) gives the following theorem (with the bound $\leq k(\pi)$ instead of $k(\pi)^{O(1)}$, in fact). We shall give a new proof of the theorem which will later allow some generalizations not offered by Craig's original proof. For completeness we recall the standard proof as well (the second proof below).

Theorem 3.1 ([11, 12]) *Let π be a cut-free LK-proof of the sequent:*

$$A_1(p, q), \ldots, A_m(p, q) \longrightarrow B_1(p, r), \ldots, B_\ell(p, r)$$

with $p = (p_1, \ldots, p_n)$ the atoms occurring simultaneously in some A_i and B_j, and $q = (q_1, \ldots, q_s)$ and $r = (r_1, \ldots, r_t)$ all other atoms occurring only in some A_i or in some B_j respectively.

Then there is an interpolant $I(p)$ of the implication:

$$\bigwedge_{i \leq m} A_i \longrightarrow \bigvee_{j \leq \ell} B_\ell$$

whose circuit-size is at most $k(\pi)^{O(1)}$.

Moreover, if the atoms p occur only positively in all A_i or in all B_j then there is a monotone interpolant whose monotone circuit-size is at most $k(\pi)^{O(1)}$.

First proof:

Define two sets $U, V \subseteq \{0, 1\}^n$ by:

$$U = \{u \in \{0,1\}^n \mid \exists q^u \in \{0,1\}^s, \bigwedge_{i \leq m} A_i(u, q^u)\}$$

$$V = \{v \in \{0,1\}^n \mid \exists r^v \in \{0,1\}^t, \bigwedge_{j \leq \ell} \neg B(v, r^v)\} \ .$$

Note that the fact that the sequent $A_1, \ldots, A_m \longrightarrow B_1, \ldots, B_\ell$ is tautologically valid is equivalent to the fact that the sets U, V are disjoint, and that any Boolean function separates U from V iff it is an interpolant of the sequent.

Using the proof π we define a particular protocol for the game on U, V. Assume that player A received $u \in U$ and B received $v \in V$. Player A fixes some $q^u \in \{0,1\}^s$ such that $\bigwedge_{i \leq m} A_i(u, q^u)$ holds and player B fixes some $r^v \in \{0,1\}^t$ for which $\bigwedge_{j \leq \ell} \neg B_j(v, r^v)$ holds.

Exchanging some bits they will construct the path $P = S_0, \ldots, S_h$ of sequents of π satisfying the following conditions:

1. S_0 in the end-sequent.

2. S_{i+1} is an upper sequent of the inference giving S_i.

3. S_h is an initial sequent.

4. For any $a = 0, \ldots, h$: if S_a has the form:

$$E_1(p,q), \ldots, E_e(p,q) \longrightarrow F_1(p,r), \ldots, F_f(p,r)$$

   then $\bigwedge_{i \leq e} E_i(u, q^u)$ holds while $\bigvee_{j \leq f} F_j(v, r^v)$ fails.

Note that as the proof is cut-free and there are no $\neg$-rules, no formula in the antecedent (resp. the succedent) of a sequent in the proof contains an atom r_i (resp. an atom q_i).

To find S_{a+1} they proceed as follows:

(a) If S_a was deduced by an inference with only one hypothesis, they put S_{a+1} to be that hypothesis and they exchange no bits.

(b) If the inference yielding S_a was the introduction of $\bigwedge_{i \leq g} D_i$ to the succedent the player B, who thinks that $\bigwedge_{i \leq g} D_i$ is false, sends to A $\lceil \log(g) \rceil$ bits identifying one particular $D_i(v, r^v)$, $i \leq g$, which is false. They take for S_{a+1} the upper sequent of the inference containing the minor formula D_i.

(c) If the inference yielding S_a was the introduction of $\bigvee_{i \leq g} D_i$ to the antecedent then analogously with (b) player A identifies to B a true $D_i(u, q^u)$ and they take for S_{a+1} the sequent containing that minor formula.

Let S_h be the initial sequent the players arrive at in the path P. It must be one of the following:

$$p_i \to p_i \text{ or } \neg p_i \to \neg p_i$$

for some $i = 1, \ldots, n$. This is because all other the initial sequents either contain an atom r_i in the antecedent or an atom q_i in the succedent, or violate the condition 4. from the definition of P.

If S_h is the former then by 4. $u_i = 1 \wedge v_i = 0$, if it is the latter then $u_i = 0 \wedge v_i = 1$.

Formally, the protocol G is defined as follows: the nodes are the sequents of π, the strategy is given by the definition of S_{a+1} from S_a and the sequent is in the consistency condition $F(u,v)$ iff it satisfies the condition 4. above. The communication complexity of G is $\leq \lceil \log(g) \rceil + 2 \leq \lceil \log(k(\pi)) \rceil + 2$.

By Theorem 2.3 there is a circuit of size $k(\pi)^{O(1)}$ separating U from V. Note that if the atoms p_i occur only positively in the antecedent or in the succedent of the end-sequent then the players always arrive to an initial sequent of the form $p_i \longrightarrow p_i$. This yields the monotone case.
This concludes the first proof.

Second proof:

Let S be a sequent in the cut-free proof π. By induction on the number of inferences before S we define explicitly the interpolant $I^S(p)$ for S (this makes sense as no r_i occurs in an antecedent and no q_i occurs in a succedent).

If S is initial then I^S is one of $0, 1, p_i, \neg p_i$ (because the only initial sequents in π where q_i or r_i occur are $q_i, \neg q_i \longrightarrow$ and $\longrightarrow r_i, \neg r_i$ which have interpolants 0 and 1 respectively).

If S was derived from one hypothesis S_1 put $I^S := I^{(S_1)}$. If S was derived by the right $\bigwedge$:introduction (resp. by the left $\bigvee$:introduction) from $S_1, \ldots, S_g$ then put $I^S := \bigwedge_{i \leq g} I^{(S_i)}$ (resp. $\bigvee_{i \leq g} I^{(S_i)}$).

The monotone case follows as then $\neg p_i$ cannot occur as the interpolant of an initial sequent.
This concludes the second proof.

q.e.d.

Both proofs of Theorem 3.1 can be modified for the case when π is not necessarily cut-free but no cut-formula contains atoms q and r at the same time. To maintain the condition that q (resp. r) do not occur in the succedent (resp. in the antecedent) we picture a cut-inference with the cut-formula D as

$$\frac{\neg D, \Gamma \longrightarrow \Delta \qquad D, \Gamma \longrightarrow \Delta}{\Gamma \longrightarrow \Delta}$$

or

$$\frac{\Gamma \longrightarrow \Delta, D \qquad \Gamma \longrightarrow \Delta, \neg D}{\Gamma \longrightarrow \Delta}$$

according to whether atoms q do or do not occur in D (this is equivalent to considering a partition of formulas in a sequent S into the ancestors of the formulas from the antecedent and from the succedent of the end-sequent, rather than just into the antecedent and the succedent of S itself).

A modification of the first proof is then straightforward as the truth-value of any cut-formula is known to one of the players and he can direct the path by sending one bit. To modify the second proof note that if I_1 and I_2 are interpolants of the hypotheses of a cut-inference as above then $I_1 \vee I_2$, respectively $I_1 \wedge I_2$, is an interpolant of the lower sequent.

Corollary 3.2 *Let π be an LK-proof of the sequent:*

$$A_1(p,q), \ldots, A_m(p,q) \longrightarrow B_1(p,r), \ldots, B_\ell(p,r)$$

with occurrences of atoms as shown. Assume that in no cut-formula some q_i and r_j occur simultaneously.

Then there is an interpolant $I(p)$ of the implication:

$$\bigwedge_{i \le m} A_i \longrightarrow \bigvee_{j \le \ell} B_\ell$$

whose circuit-size is at most $k(\pi)^{O(1)}$.

Moreover, if the atoms p occur only positively in all A_i or in all B_j then there is a monotone interpolant whose monotone circuit-size is at most $k(\pi)^{O(1)}$.

A third way how to obtain the corollary from Theorem 3.1 is provided by the second proof of Theorem 6.1.

4 Semantic derivations

In the first proof of Theorem 3.1 we have not used any particular syntactic property of formulas in an LK-proof. Rather the proof worked with the sets of truth assignments satisfying the formulas in π. The following notion is intended to formalize a general situation in which a similar argument may work for proofs allowing some form of cut-rule.

Definition 4.1 *Let N be a fixed natural number.*

1. *The* semantic rule *allows to infer from two subsets $A, B \subseteq \{0,1\}^N$ a third one:*

   $$\frac{A \quad B}{C}$$

   iff $C \supseteq A \cap B$.

2. *A* semantic derivation *of the set $C \subseteq \{0,1\}^N$ from the sets $A_1, \ldots, A_m \subseteq \{0,1\}^N$ is a sequence of sets $B_1, \ldots, B_k \subseteq \{0,1\}^N$ such that $B_k = C$, each B_i is either one of A_j or derived from two previous B_{i_1}, B_{i_2} by the semantic rule.*

3. *Let $\mathcal{X} \subseteq exp(\{0,1\}^N)$ be a family of subsets of $\{0,1\}^N$. A semantic derivation $B_1, \ldots, B_k$ is an $\mathcal{X}$-derivation iff all $B_i \in \mathcal{X}$.*

Recall that a *filter* of subsets of $\{0,1\}^N$ is a family $\mathcal{X}$ closed upwards ($A \in \mathcal{X} \wedge B \supseteq A \rightarrow B \in \mathcal{X}$) and closed under intersection ($A, B \in \mathcal{X} \rightarrow A \cap B \in \mathcal{X}$). The following lemma is obvious.

Lemma 4.2 *Let $A_1, \ldots, A_m, C \in \{0,1\}^N$. Then the following three conditions are equivalent:*

1. *C can be semantically derived from $A_1, \ldots, A_m$.*
2. *C can be semantically derived from $A_1, \ldots, A_m$ in $m-1$ steps.*
3. *C is in the smallest filter containing $A_1, \ldots, A_m$.*

It means that to have a non-trivial meaning of the length of semantic derivations we must restrict to $\mathcal{X}$-derivations for some family $\mathcal{X}$ which itself is not a filter. For example, a family formed by subsets of $\{0,1\}^N$ definable by disjunctions of literals yields a non-trivial notion. The following technical definition abstracts a property of sets of truth assignments used in the first proof of Theorem 3.1.

Definition 4.3 *Let $N = n + s + t$ be fixed and let $A \subseteq \{0,1\}^N$. Let $u, v \in \{0,1\}^n$, $q^u \in \{0,1\}^s$ and $r^v \in \{0,1\}^t$.*

Consider three tasks:

1. *Decide whether $(u, q^u, r^v) \in A$.*
2. *Decide whether $(v, q^u, r^v) \in A$.*
3. *If $(u, q^u, r^v) \in A \not\equiv (v, q^u, r^v) \in A$ find $i \leq n$ such that $u_i \neq v_i$.*

These tasks can be solved by two players, one knowing u, q^u and the other one knowing v, r^v. The communication complexity *of A, $CC(A)$, is the minimal number of bits they need to exchange in the worst case in solving any of these three tasks.*

Consider two more tasks:

4. *If $(u, q^u, r^v) \in A$ and $(v, q^u, r^v) \notin A$ either find $i \leq n$ such that*

$$u_i = 1 \wedge v_i = 0$$

or learn that there is some u' satisfying

$$u' \geq u \wedge (u', q^u, r^v) \notin A$$

($u \leq u'$ means $\bigwedge_{i \leq n} u_i \leq u'_i$.)

5. *If* $(u, q^u, r^v) \notin A$ *and* $(v, q^u, r^v) \in A$ *either find* $i \leq n$ *such that*

$$u_i = 1 \wedge v_i = 0$$

or learn that there is some v' *satisfying*

$$v' \leq v \wedge (v', q^u, r^v) \notin A \ .$$

(Note that the players are not required to find u' *and* v' *in 4. and 5. and that the two cases in each task are not necessarily exclusive.)*
The monotone communication complexity w.r.t. U *of* A, $MCC_U(A)$, *is the minimal* $t \geq CC(A)$ *such that the task 4. can be solved communicating* $\leq t$ *bits in the worst case.*

The monotone communication complexity w.r.t. V *of* A, $MCC_V(A)$, *is the minimal* $t \geq CC(A)$ *such that the task 5. can be solved communicating* $\leq t$ *bits in the worst case.*

In the example above, any set definable by a disjunction of literals has both the communication complexity and the monotone communication complexity at most $\lceil \log(n) \rceil + 2$.

Note that proofs in any of the usual propositional calculi translate into semantic derivations: simply replace a sequent (a formula, an equation, etc.) by the set of its satisfying truth assignments. The inference rules translate into instances of the semantic rule as they are all sound.

5 An interpolation theorem for semantic derivations

Let $N = n + s + t$ be fixed for the whole section. For $A \subseteq \{0,1\}^{n+s}$ define the set $\tilde{A}$ by:

$$\tilde{A} := \bigcup_{(a,b) \in A} \{(a, b, c) \mid c \in \{0,1\}^t\}$$

where a, b, c range over $\{0,1\}^n$, $\{0,1\}^s$ and $\{0,1\}^t$ respectively, and similarly for $B \subseteq \{0,1\}^{n+t}$ define $\tilde{B}$:

$$\tilde{B} := \bigcup_{(a,c) \in B} \{(a, b, c) \mid b \in \{0,1\}^s\} \ .$$

Theorem 5.1 *Let* $A_1, \ldots, A_m \subseteq \{0,1\}^{n+s}$ *and* $B_1, \ldots, B_\ell \subseteq \{0,1\}^{n+t}$. *Assume that there is a semantic derivation* $\pi = D_1, \ldots, D_k$ *of the empty set* $\emptyset = D_k$ *from the sets* $\tilde{A}_1, \ldots, \tilde{A}_m, \tilde{B}_1, \ldots, \tilde{B}_\ell$ *such that:*

$$CC(D_i) \leq t$$

for all $i \leq k$.
Then the two sets

$$U = \{u \in \{0,1\}^n \mid \exists q^u \in \{0,1\}^s ; (u, q^u) \in \bigcap_{j \leq m} A_j\}$$

and

$$V = \{v \in \{0,1\}^n \mid \exists r^v \in \{0,1\}^t ; (v, r^v) \in \bigcap_{j \leq \ell} B_j\}$$

can be separated by a circuit of size at most $(k+2n)2^{O(t)}$.

Moreover, if the sets $A_1, \ldots, A_m$ *satisfy the following monotonicity condition w.r.t.* U:

$$(u, q^u) \in \bigcap_{j \leq m} A_j \wedge u \leq u' \;\rightarrow\; (u', q^u) \in \bigcap_{j \leq m} A_j$$

and $MCC_U(D_i) \leq t$ *for all* $i \leq k$, *or if the sets* $B_1, \ldots, B_\ell$ *satisfy:*

$$(v, r^v) \in \bigcap_{j \leq \ell} B_j \wedge v \geq v' \;\rightarrow\; (v', r^v) \in \bigcap_{j \leq \ell} B_j$$

and $MCC_V(D_i) \leq t$ *for all* $i \leq k$, *then there is a monotone circuit separating* U *from* V *of size at most* $(k+n)2^{O(t)}$.

Proof :
Let $\pi = D_1, \ldots, D_k$ be a semantic derivation of $\emptyset$ from $\tilde{A}_1, \ldots, \tilde{B}_\ell$. We shall construct a protocol G for the Karchmer-Wigderson game on U, V. Before we define it formally we first explain its idea in terms of two players constructing a path through π.

The two players A and B, one knowing $(u, q^u) \in \bigcap_j A_j$ and the other one knowing $(v, r^v) \in \bigcap_j B_j$, attempt to construct a path $P = S_0, \ldots, S_h$ through π. $S_0 = \emptyset = D_k$, S_{a+1} is one of the two sets which are the hypotheses of the semantic inference yielding S_a and $S_h \in \{\tilde{A}_1, \ldots, \tilde{B}_\ell\}$. Moreover, both tuples (u, q^u, r^v) and (v, q^u, r^v) are *not* in S_a, $a = 0, \ldots, h$.

If the players know S_a which was deduced in the inference:

$$\frac{X \quad Y}{S_a}$$

then they first determine whether $(u, q^u, r^v) \in X$ and $(v, q^u, r^v) \in X$. There are three possible outcomes:

1. both (u, q^u, r^v) and (v, q^u, r^v) are in X
2. none of $(u, q^u, r^v), (v, q^u, r^v)$ is in X
3. only one of $(u, q^u, r^v), (v, q^u, r^v)$ is in X.

In the first case none of the two tuples can be in Y and the players put $S_{a+1} := Y$. In the second case they take $S_{a+1} := X$. In the third case they stop constructing the path and enter a protocol aimed at finding $i \leq n$ such that $u_i \neq v_i$. Such i must exists as necessarily $u \neq v$. As none of the initial sets $\tilde{A}_1, \ldots, \tilde{B}_\ell$ avoids both $(u, q^u, r^v), (v, q^u, r^v)$ the players must sooner or later enter the possibility 3. and find $i \leq n$ such that $u_i \neq v_i$.

Now we define the protocol G formally. G has $(k+2n)$ nodes, the k steps of the derivation π plus $2n$ additional nodes labelled by formulas $u_i = 1 \wedge v_i = 0$ and $u_i = 0 \wedge v_i = 1$, $i = 1, \ldots, n$. The consistency condition $F(u, v)$ consists of those D_j such that $(v, q^u, r^v) \notin D_j$ and of those of the additional $2n$ nodes whose label is valid for the particular pair u, v.

The strategy function (for D_j derived from X and Y) is defined as follows:

1. If $(u, q^u, r^v) \notin D_j$ then

$$S(u, v, D_j) := \begin{cases} X & \text{if } (v, q^u, r^v) \notin X \\ Y & \text{if } (v, q^u, r^v) \in X \text{ (and hence } (v, q^u, r^v) \notin Y). \end{cases}$$

2. If $(u, q^u, r^v) \in D_j$ then the players use the protocol (whose existence is guaranteed by the definition of $CC(D_j)$) for finding $i \leq n$ such that $u_i \neq v_i$. $S(u, v, D_j)$ is then the one of the two nodes labelled by $u_i = 1 \wedge v_i = 0$ and $u_i = 0 \wedge v_i = 1$ whose label is valid for the pair u, v.

Note that both the strategy function $S(u, v, x)$ and the membership relation $x \in F(u, v)$ can be determined by the players exchanging at most $O(t)$ bits. As G has $(k+2n)$ nodes, Theorem 2.3 yields the wanted circuit separating U from V and having the size at most $(k+2n) \cdot 2^{O(t)}$.

The protocol requires a modification for the monotone case. Assume that the sets $A_1, \ldots, A_m$ satisfy the monotonicity condition w.r.t. U and that $MCC_U(D_i) \leq t$ for all $i \leq k$ (the case of the monotonicity w.r.t. V is analogous). The protocol has $(k+n)$ nodes, the k steps of the derivation π plus n additional nodes labelled by formulas $u_i = 1 \wedge v_i = 0$, $i = 1, \ldots, n$. The consistency condition $F(u, v)$ is defined as before.

The strategy function is defined in a bit different way. The players use the protocol for solving the task 4. from Definition 4.3. There are two possible outcomes:

1. They decide that the condition:

$$\exists u' \geq u, (u', q^u, r^v) \notin D_j$$

is true for u, v. Then they put:

$$S(u, v, D_j) := \begin{cases} X & \text{if } (v, q^u, r^v) \notin X \\ Y & \text{if } (v, q^u, r^v) \in X. \end{cases}$$

2. They find $i \leq n$ such that $u_i = 1 \wedge v_i = 0$. $S(u, v, D_j)$ is then the additional node with the label $u_i = 1 \wedge v_i = 0$.

By the monotonicity condition imposed on $A_1, \ldots, A_m$, for every u' occurring above it holds:

$$(u', q^u, r^v) \in \bigcap_{j \leq m} A_j \ .$$

This implies that the players have to find sooner or later $i \leq n$ such that $u_i = 1 \wedge v_i = 0$.

By the assumption about the monotone communication complexity of all D_j, both the relation $x \in F(u, v)$ and the function $S(u, v, x)$ can be computed exchanging $O(t)$ bits. As G has $(k + n)$ nodes, Theorem 2.3 yields the wanted monotone circuit separating U from V and having the size at most $(k+n) \cdot 2^{O(t)}$.

q.e.d.

6 Upper bounds for some interpolation theorems

In this section we derive from Theorem 5.1 feasible bounds for interpolation theorems for resolution, a subsystem of LK relevant to bounded arithmetic, and for LEC and CP.

Theorem 6.1 *Assume that the set of clauses $\{A_1, \ldots, A_m, B_1, \ldots, B_\ell\}$ where:*

1. *$A_i \subseteq \{p_1, \neg p_1, \ldots, p_n, \neg p_n, q_1, \neg q_1, \ldots, q_s, \neg q_s\}$, all $i \leq m$*
2. *$B_j \subseteq \{p_1, \neg p_1, \ldots, p_n, \neg p_n, r_1, \neg r_1, \ldots, r_t, \neg r_t\}$, all $j \leq \ell$*

has a resolution refutation with k clauses.

Then the implication:

$$\bigwedge_{i \leq m} (\bigvee A_i) \ \longrightarrow \bigvee_{j \leq \ell} (\bigwedge \neg B_j)$$

(where $\bigvee A_i$ denotes the disjunction of the literals in A_i and $\bigwedge \neg B_j$ denotes the conjunction of the negations of the literals in B_j) has an interpolant $I(p)$ whose circuit-size is $kn^{O(1)}$.

Moreover, if all atoms p occur only positively in all A_i, or if all p occur only negatively in all B_j, then there is a monotone interpolant whose monotone circuit-size is $kn^{O(1)}$.

First proof:

Let $\pi = C_1, \ldots, C_k$ be a resolution refutation of $A_1, \ldots, B_\ell$. For a clause C denote by $\tilde{C}$ the subset of $\{0, 1\}^{n+s+t}$ of all those truth assignments satisfying C. Then $\tilde{\pi} = \tilde{C}_1, \ldots, \tilde{C}_k$ is a semantic derivation of $\emptyset$ from $\tilde{A}_1, \ldots, \tilde{B}_\ell$.

Obviously, for any clause C both the communication complexity and the monotone communication complexity of $\tilde{C}$ is at most $CC(\tilde{C}) \leq \lceil \log(n) \rceil + 2$. Hence Theorem 5.1 yields circuit of size $(k + 2n) \cdot n^{O(1)} \leq k \cdot n^{O(1)}$ Similarly for the monotone case. This concludes the first proof.

Second proof:
We give a second proof of a slightly worse bound via a translation of resolution refutation into cut-free proofs. Assume that $C_1, \ldots, C_k$ is a resolution refutation of clauses $A_1, \ldots, B_\ell$. We show that for every $a \leq k$ there are cedents Γ_a, Δ_a such that the following conditions hold:

1. Each formula in Γ_a has the form either $p_i \vee \neg p_i$ or $q_i \vee \neg q_i$.
2. Each formula in Δ_a has the form $r_i \wedge \neg r_i$.
3. The sequent

$$\Gamma_a, \bigvee A_1, \ldots, \bigvee A_m \longrightarrow \bigwedge \neg B_1, \ldots, \bigwedge \neg B_\ell, \Delta_a, \bigvee C_a$$

   has a cut-free LK-proof with $O(a \cdot N)$ sequents, where $N = n + s + t$.

This is readily established by induction on a. For $C_a \in \{A_1, \ldots, B_\ell\}$ apply $\bigvee$: left or $\bigwedge$: right ($O(N)$ sequents), otherwise replace the cut inference yielding C_a with the cut formula p_i or q_i by introduction of $p_i \vee \neg p_i$ or $q_i \vee \neg q_i$ respectively into the antecedent, and the cut inference with the cut formula r_i by introduction $r_i \wedge \neg r_i$ to the succedent. These new formulas form the cedents Γ_a and Δ_a respectively.

By Theorem 3.1 the implication:

$$\bigwedge \Gamma_k \wedge \bigwedge_{i \leq m} (\bigvee A_i) \longrightarrow \bigvee_{j \leq \ell} (\bigwedge \neg B_j) \vee \bigvee \Delta_k$$

has an interpolant $I(p)$ of circuit size $(kN)^{O(1)}$. Now note that, as $\bigwedge \Gamma_k$ is a tautology while $\bigvee \Delta_k$ is unsatisfiable, $I(p)$ is, in fact, an interpolant for the implication

$$\bigwedge_{i \leq m} (\bigvee A_i) \longrightarrow \bigvee_{j \leq \ell} (\bigwedge \neg B_j)$$

as well.
This concludes the second proof.

q.e.d.

The following statement extends the previous theorem to a larger class of LK-proofs. This class appears naturally in connection with bounded arithmetic (see [19, 2.2] or [20, 12.2.1]).

Corollary 6.2 *Let π be an LK-proof of the sequent:*

$$A_1(p,q),\ldots,A_m(p,q) \longrightarrow B_1(p,r),\ldots,B_\ell(p,r)$$

with atoms p,q,r occurring as displayed and such that the formulas A_i (resp. B_j) are literals or disjunctions (resp. conjunctions) of literals.

Assume that π satisfies:

1. *π is tree-like.*
2. *Every formula in π has the depth at most two.*
3. *Every sequent in π contains at most c depth 2 formulas, where c is an independent constant.*

Then there is an interpolant $I(p)$ of the implication:

$$\bigwedge_{i\le m} A_i \longrightarrow \bigvee_{j\le\ell} B_j$$

whose circuit-size is at most $k(\pi)^{O(c)}n^{O(1)}$.

Moreover, if the atoms p occur only positively in all A_i or in all B_j then there is a monotone interpolant whose monotone circuit-size is $k(\pi)^{O(c)}n^{O(1)}$.

Proof :

Assume that π satisfies the hypothesis of the corollary. It was shown in [19, 2.2] (or see [20, 12.2.1]) that π can be transformed into a tree-like proof π' with $k(\pi') = k(\pi)^{O(c)}$ in which every formula is of depth ≤ 1.

Furthermore, by [19, 1.2] (or see [20, 12.2.1]) such π' can be transformed into a resolution refutation π'' (which is not necessarily tree-like) of clauses representing the sequents $\longrightarrow A_1,\ldots,\longrightarrow A_m$ and $\neg B_1 \longrightarrow,\ldots,\neg B_\ell \longrightarrow$, and such that $k(\pi'') = k(\pi')^{O(1)} = k(\pi)^{O(c)}$. The corollary then follows from Theorem 6.1.

q.e.d.

Next we deduce interpolation theorems for LEC and CP.

Theorem 6.3 *Let $E_1(x,y),\ldots,E_m(x,y)$ and $F_1(x,z),\ldots,F_\ell(x,z)$ be a system of linear equations over a finite field F in which occur only the displayed variables $x=(x_1,\ldots,x_n)$, $y=(y_1,\ldots,y_s)$ and $z=(z_1,\ldots,z_t)$.*

Assume that there is an LEC-refutation π of the system with $k(\pi)$ inferences.

Then there is an interpolant $I(x)$ of the implication:

$$\bigwedge_{i\le m} E_i(x,y) \longrightarrow \bigvee_{j\le\ell} \neg F_j(x,z)$$

whose circuit-size is at most $k(\pi)n^{O(1)}$.

Proof :

Put $N := n + s + t$. For an equation $C(x, y, z)$ denote by $\tilde{C}$ the subset of $\{0,1\}^N$ of those tuples satisfying C. If $\pi = C_1, \ldots, C_k$ is an LEC-refutation then $\tilde{C}_1, \ldots, \tilde{C}_k$ is a semantic refutation of $\tilde{E}_1, \ldots, \tilde{F}_\ell$. For any linear equation C the communication complexity of $\tilde{C}$ is at most $O(\log(n))$. The theorem then follows from Theorem 5.1.

q.e.d.

If $F = \mathbf{Q}$ we do not get such an estimate on $CC(\tilde{C})$ valid for all C. Rather we need also to incorporate the sizes of the coefficients (see the definition of $||a||$ in section 1). Instead of this we prove an interpolation theorem for CP; the case of LEC with $F = \mathbf{Q}$ is similar.

Theorem 6.4 *Let $E_1(x, y), \ldots, E_m(x, y), F_1(x, z), \ldots, F_\ell(x, z)$ be a system of CP-inequalities in which occur only the displayed variables $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_s)$ and $z = (z_1, \ldots, z_t)$. Let $N := n + s + t$. Assume that there is a CP-refutation π of the system such that:*

1. *π contains $k(\pi)$ inferences.*
2. *Every coefficient occurring in π has the absolute value at most M.*

Then there is an interpolant $I(x)$ of the implication:

$$\bigwedge_{i \leq m} E_i(x, y) \longrightarrow \bigvee_{j \leq \ell} \neg F_j(x, z)$$

whose circuit-size is at most $k(\pi)(MN)^{O(1)}(Mn)^{O(\log n)}$.

Moreover, if all x_i occur in all $E_1, \ldots, E_m$ with non-negative coefficients only, or if all x_i occur in all $F_1, \ldots, F_\ell$ only with non-positive coefficients only, then there is a monotone interpolant whose monotone circuit-size is at most $k(\pi)(MN)^{O(1)}(Mn)^{O(\log n)}$.

Proof :

Assume that π is a CP-refutation satisfying the hypothesis of the corollary. Every inequality D in π has the form:

$$a \cdot x + b \cdot y + c \cdot z \geq d$$

where $a \cdot x$ abbreviates the scalar product $a_1 x_1 + \ldots + a_n x_n$ (and similarly $b \cdot y$ and $c \cdot z$). Let $\tilde{D}$ be the set of assignments satisfying D.

Assume that player A received $u \in \{0,1\}^n$ such that all $E_i(u, y^u)$ are satisfied for some y^u, while B received $v \in \{0,1\}^n$ such that all $F_j(v, z^v)$ are satisfied for some z^v. As in the proof of Theorem 6.1 it is sufficient to estimate the (monotone) communication complexity of $\tilde{D}$.

For the tasks to decide whether $(u, y^u, z^v) \in \tilde{D}$ and $(v, y^u, z^v) \in \tilde{D}$ it is sufficient if A sends to B the numbers $a \cdot u$ and $b \cdot y^u$, and B sends to A $a \cdot v$ and $c \cdot z^v$. This needs at most $2 \cdot \lceil \log(MN) \rceil$ bits each.

If $(u, y^u, z^v) \in \tilde{D} \not\equiv (v, y^u, z^v) \in \tilde{D}$ then necessarily $a \cdot u \neq a \cdot v$ and the players find $i \leq n$ such that $u_i \neq v_i$ by binary search. Here $\lceil \log n \rceil \cdot \lceil \log(Mn) \rceil$ bits suffice.

This shows that for any $\tilde{D}$ the communication complexity is at most:

$$CC(\tilde{D}) \leq O(\log(MN) \ + \ \log n \cdot \log(Mn)) \ .$$

Theorem 5.1 implies that the implication has an interpolant of circuit-size at most:

$$k(\pi) \cdot (MN)^{O(1)} (Mn)^{O(\log n)} \ .$$

For the monotone case assume that the variables x_i occur only with non-negative coefficients in all $E_1, \ldots, E_m$ (the other case is analogous). The monotonicity condition of Theorem 5.1

$$u' \geq u \wedge \bigwedge_{i \leq m} E_i(u, y^u) \longrightarrow \bigwedge_{i \leq m} E_i(u', y^u)$$

is then satisfied. It is thus sufficient to estimate $MCC_U(\tilde{D})$.

Assume $(u, y^u, z^v) \in \tilde{D}$ while $(v, y^u, z^v) \notin \tilde{D}$. Then $a \cdot u > a \cdot v$. Write the vector a as a difference of two vectors with non-negative coefficients, $a = a_1 - a_2$. There are two possibilities

1. $a_1 \cdot u > a_1 \cdot v$.

2. $a_1 \cdot u \leq a_1 \cdot v \wedge a_2 \cdot u < a_2 \cdot v$.

In the first case the players use binary search to find $i \leq n$ such that $u_i = 1 \wedge v_i = 0$. This needs at most $\lceil \log n \rceil \cdot \lceil \log(Mn) \rceil$ bits. In the second case they know that for some $u' \geq u$ it holds $a_2 \cdot u' \geq a_2 \cdot v$ and hence also $a \cdot u' \leq a \cdot v$ and $(u', y^u, z^v) \notin \tilde{D}$. To decide which case applies needs at most $2\lceil \log(Mn) \rceil$ bits. Hence

$$MCC_U(\tilde{D}) \leq O(\log(MN) \ + \ \log n \cdot \log(Mn)) \ .$$

Theorem 5.1 yields the existence of a monotone interpolant with the monotone circuit-size at most

$$k(\pi) \cdot (MN)^{O(1)} (Mn)^{O(\log n)} \ .$$

q.e.d.

7 Lower bounds for proof systems

Assume that for a propositional proof system P we have a good interpolation theorem allowing, in particular, good estimates of the complexity of the monotone interpolants. Then an implication which cannot have a small monotone interpolant must have long P-proofs. A similar idea of lower bounds for proof systems was discussed in the context of counting principles in [19, Sect.5].

Non-trivial lower bounds to the circuit size are known for monotone circuits separating graphs with large cliques from those colorable by a small number of colors, see [39, 2, 1]. It is thus natural to use the implications determined by these two $\mathcal{NP}$-sets as explained in the introduction. Similar implications were discussed in [43].

Definition 7.1 *Let $n, \omega, \xi \geq 1$ be natural numbers, and let $\binom{n}{2}$ denote the set of two-element subsets of $\{1, \dots, n\}$. The set $Clique_{n,\omega}(p, q)$ is a set of the following formulas in the atoms p_{ij}, $\{i, j\} \in \binom{n}{2}$, and q_{ui}, $u = 1, \dots, \omega$ and $i = 1, \dots, n$:*

(1a) $\bigvee_{i \leq n} q_{ui}$, *all* $u \leq \omega$,

(1b) $\neg q_{ui} \vee \neg q_{vi}$, *all* $u < v \leq \omega$ *and* $i = 1, \dots, n$,

(1c) $\neg q_{ui} \vee \neg q_{vj} \vee p_{ij}$, *all* $u < v \leq \omega$ *and* $\{i, j\} \in \binom{n}{2}$.

The set $Color_{n,\xi}(p, r)$ is the set of the following formulas in the atoms p_{ij}, $\{i, j\} \in \binom{n}{2}$, and r_{ia}, $i = 1, \dots, n$ and $a = 1, \dots, \xi$:

(2a) $\bigvee_{a \leq \xi} r_{ia}$, *all* $i \leq n$,

(2b) $\neg r_{ia} \vee \neg r_{ib}$, *all* $a < b \leq \xi$ *and* $i \leq n$,

(2c) $\neg r_{ia} \vee \neg r_{ja} \vee \neg p_{ij}$, *all* $a \leq \xi$ *and* $\{i, j\} \in \binom{n}{2}$.

The expression

$$Clique_{n,\omega} \longrightarrow \neg Color_{n,\xi}$$

is an abbreviation of the sequent whose antecedent consists of all formulas in (1 a-c) and whose succedent consists of the negations of the formulas in (2 a-c).

Truth assignments to p_{ij} can be identified with graphs on n vertices. Truth assignments to q_{ui} such that $Clique_{n,\omega}(p, q)$ is satisfied can be identified with 1-to-1 maps from the set $\{1, \dots, \omega\}$ onto a clique in graph p, and truth assignments to r_{ia} such that $Color_{n,\xi}(p, r)$ is satisfied can be identified with colorings of graph p by colors $\{1, \dots, \xi\}$. Thus the set

$$\{p \mid \exists q \; Clique_{n,\omega}(p, q)\}$$

is the set of graphs with a clique of size $\geq \omega$, while the set

$$\{p \mid \exists r \; Color_{n,\xi}(p, r)\}$$

is the set of graphs colorable by $\leq \xi$ colors. Hence the sequent

$$Clique_{n,\omega} \longrightarrow \neg Color_{n,\xi}$$

is tautologically valid iff $\xi < \omega$. The following theorem just restates the bound from [1], replacing the class of graphs without a clique of size ξ used in [1] by the smaller class of ξ-colorable graphs (the bound to monotone circuits separating these two classes is what is actually proved in [1]).

Theorem 7.2 ([1]) *Assume that* $3 \leq \xi < \omega$ *and* $\sqrt{\xi}\omega \leq \frac{n}{8\log n}$. *Then the sequent*

$$Clique_{n,\omega} \longrightarrow \neg Color_{n,\xi}$$

has no interpolant of the monotone circuit-size smaller than:

$$2^{\Omega(\sqrt{\xi})} \ .$$

For the next statement note that all formulas in the set $Clique_{n,\omega} \cup Color_{n,\xi}$ are disjunctions of literals and thus can be identified with resolution clauses. A resolution clause $\{x_{i_1}, \ldots, x_{i_a}, \neg x_{j_1}, \ldots, \neg x_{j_b}\}$ can be represented by a CP-inequality

$$x_{i_1} + \ldots + x_{i_a} - x_{j_1} \ldots - x_{j_b} \geq 1 - b \ .$$

Hence the set $Clique_{n,\omega} \cup Color_{n,\xi}$ can be considered also as a set of CP-inequalities in p, q, r.

Corollary 7.3 *Let* n *be sufficiently large and let* $\xi = \lceil \sqrt{n}\ \rceil$, $\omega = \xi + 1$. *Then:*

1. *Every resolution refutation of the clauses* $Clique_{n,\omega} \cup Color_{n,\xi}$ *must have at least* $2^{\Omega(n^{\frac{1}{4}})}$ *clauses.*

2. *Every* CP*-refutation of the clauses* $Clique_{n,\omega} \cup Color_{n,\xi}$ *with all coefficients in the absolute value* $\leq M$ *must have at least*

$$\frac{2^{\Omega(n^{\frac{1}{4}})}}{M^{O(\log n)}}$$

inequalities.

In particular, if $M \leq 2^{n^{\epsilon}}$ *then for* ϵ *a sufficiently small constant the number of inequalities is at least* $2^{n^{\Omega(1)}}$.

Proof :

By Theorem 6.1 a resolution refutation with k clauses would imply the existence of an interpolant with monotone circuit size $kn^{O(1)}$. The hypothesis of Theorem 7.2 is fulfilled and so it must hold:

$$kn^{O(1)} \geq 2^{\Omega(n^{\frac{1}{4}})}$$

and hence

$$k \geq 2^{\Omega(n^{\frac{1}{4}})}$$

as well.

The second part is proved analogously using Theorem 6.4 in place of 6.1. By 6.4 and 7.2:

$$k(\pi)(MN)^{O(1)}(Mn)^{O(\log n)} \geq 2^{\Omega(n^{\frac{1}{4}})}$$

where $N = \binom{n}{2} + n(s+t) = O(n^2)$. This implies:

$$k(\pi)M^{O(\log n)}n^{O(\log n)} \geq 2^{\Omega(n^{\frac{1}{4}})}$$

and so

$$k \geq \frac{2^{\Omega(n^{\frac{1}{4}})}}{M^{O(\log n)}} \ .$$

For $M \leq 2^{n^{\epsilon}}$, ϵ suitably small, the right-hand side is $2^{n^{\Omega(1)}}$.

q.e.d.

Note that by a suitable choice of ξ we can get a lower bound of the form $2^{\Omega(n^{\frac{1}{3}-\epsilon})}$, for arbitrary small $\epsilon > 0$.

8 An independence result for the bounded arithmetic theory $S^2_2(\alpha)$

The first bounded arithmetic theory was introduced in [34]. Current research is centered around the theories defined in [5]. In this section we give a new presentation of the proof of the independence result for the theory $S^2_2(\alpha)$ obtained in [43]. For the definition of the theory as well as for the details of bounded arithmetic the reader should consult [5] or [20, Chpt. 5] (in particular, the language $L(\alpha)$ of $S^2_2(\alpha)$ contains, in fact, countably many unary predicates α_i). In the latter can also be found details of various relations between the arithmetic systems and the propositional proof systems (in [20, Chpt. 9] in particular). We shall recall briefly a translation of bounded $L(\alpha)$ formulas; [35] used it first in a connection with bounded arithmetic.

A bounded formula $A(a, \alpha_1, \ldots, \alpha_k)$ with the predicate parameters α_i and the number parameter a can be for every value $a := N$ translated into a constant-depth, size $2^{(\log N)^{O(1)}}$ formula: the atomic sentence $j \in \alpha_i$ translates into the atom p^i_j, a true (resp. false) first-order atomic sentence translates into 1 (resp. into 0) and a bounded universal (resp. existential) quantifier $\forall x < tB(x)$ resp. $\exists x < tB(x)$ translates into a conjunction (resp. a disjunction) of the translations of $B(x)$, $x = 0, \ldots, t-1$. We shall denote the translation of formula A for $a = N$ by $\langle A\rangle_N(p^1, \ldots, p^k)$.

There are rather sophisticated relations between bounded arithmetic theories and propositional proof systems, see [8, 35, 22, 23, 26, 27, 19, 21] or [20, Chpts.9 and 11-15].

The class of first-order bounded formulas in the language of bounded arithmetic (no predicates α_i) is denoted Σ^b_∞. We call a bounded $L(\alpha)$-formula $E_1(\alpha, \Sigma^b_\infty)$ if it has the form $\exists^{\leq} A$, where A is a disjunction of conjunctions of atomic formulas and Σ^b_∞-formulas. $U_1(\alpha, \Sigma^b_\infty)$-formulas are defined dually, replacing $\exists^{\leq}$ by $\forall^{\leq}$ and a disjunction of conjunctions by a conjunction of disjunctions. The following theorem is known (see, for example, the simulation as proved in [21] or [20, Chpt. 9]).

Theorem 8.1 *Assume that $\forall x \leq s(a) A(a, x, \alpha_1, \ldots, \alpha_k)$ is a bounded $U_1(\alpha_1, \ldots, \alpha_k, \Sigma^b_\infty)$-formula and that $\exists y \leq t(a) B(a, y, \alpha_1, \ldots, \alpha_k)$ is a bounded $E_1(\alpha_1, \ldots, \alpha_k, \Sigma^b_\infty)$-formula.*

Assume that the theory $S^2_2(\alpha)$ proves the sequent:

$$\forall x \leq s(a)\ A(a, x, \alpha_1, \ldots, \alpha_k) \quad \longrightarrow \exists y \leq t(a)\ B(a, y, \alpha_1, \ldots, \alpha_k)\ .$$

Then for every N the propositional sequent:

$$\langle A\rangle_{N,0}, \ldots, \langle A\rangle_{N,s(N)} \ \longrightarrow\ \langle B\rangle_{N,0}, \ldots, \langle B\rangle_{N,t(N)}$$

(where the formulas are built from atoms $p^1, \ldots, p^k$) has an LK-proof π_N satisfying the following conditions:

1. *π is tree-like.*
2. *$k(\pi_N) = 2^{(\log N)^{O(1)}}$.*
3. *Every formula in π_N has depth at most 2.*
4. *Every sequent in π_N contains at most c depth 2 formulas (c an independent constant).*

We turn now our attention to the provability of circuit-size lower bounds in bounded arithmetic.

Razborov [42, 43] studies a formalization of Boolean complexity methods in the bounded arithmetic theory V^1_1 and in its fragments. In that formalization Boolean functions and circuits are coded by sets while Boolean inputs are coded by numbers. This allows to speak directly about exponential size circuits. In [42] he demonstrated that all major lower bounds to the circuit-size of restricted circuit models known at present can be also proved in V^1_1. On the other hand, in [43] he showed - under a cryptographic assumption about the existence of strong pseudo-random number generators - that the subtheory $S^2_2(\alpha)$ of V^1_2 does not prove a superpolynomial lower bound to the size of general unrestricted circuits computing the satisfiability predicate.

His proof relies on an interpolation theorem for second order bounded arithmetic derived with the help of technical *split versions*[1] of bounded arithmetic theories. We shall give below a direct proof via propositional interpolation.

We first very briefly recall the formalization of Boolean functions and circuits adopted in [42, 43]. Let N be a number of the form $N = 2^n$. Any subset $f \subseteq \{1, \ldots, N\}$ is thought of as a truth-table of a Boolean function $f : \{0,1\}^n \to \{0,1\}$. A first order bounded formula (no second order variables) $E(x, a)$ determines for every N the function $E_N = \{1 \leq i \leq N \mid E(i, N)\}$. We shall call such functions *explicit*.

The formula:

$$Comp(\alpha, N, t(N), f)$$

formalizes that α codes a circuit (with $\vee, \wedge$ of fan-in two) of size $t(N)$ in inputs $x_1, \ldots, x_n$, together with the computations of the circuits on all $a \leq N$ of length n, and that the circuit computes the function $f \subseteq \{1, \ldots, N\}$. It is a $U_1(\alpha, \Sigma^b_\infty)$-formula.

The formalization of the lower bound $t(N)$ to the circuit-size of f is the formula:

$$LB(N, t(N), f) := \forall \alpha, \neg \mathrm{Comp}(\alpha, N, t(N), f).$$

It should be noted that the formalization of the notions of complexity theory adopted in [42, 43] differs from the one usually accepted in bounded arithmetic [36, 5, 8, 20, 22] in which all combinatorial objects (inputs, circuits,...) are coded at the same level (by sets in the case of V^1_1) while (Boolean) functions are identified with definable classes. In the latter formalization one can speak only about functions from the polynomial-time hierarchy (as only those are definable) and only about circuits of polynomial size (as the existence of objects of superpolynomial size is not provable in bounded arithmetic). It is this latter version in which bounded arithmetic theories enjoy close relationship with propositional proof systems (mutual simulation) and with computational complexity (definability, witnessing and natural forms of major problems).

Results in the two frameworks can be compared using the known relation between first and second order bounded arithmetic (the $RSUV$-isomorphism, see [17, 40, 46]). Positive results like formalizations of particular exponential lower bounds in the former formalization are apparently weaker than the corresponding proofs of the non-existence of polynomial size upper bounds in the

[1]Not only are the split theories reader-unfriendly but their notation is unpleasant as well. I suggest to replace the original notation of [43] by the following more customary one: $\mathcal{S}\Sigma^b_0 = \Sigma^b_\infty(\alpha) + \Sigma^b_\infty(\beta)$, $\mathcal{S}\Sigma^b_i = \Sigma^b_i(\Sigma^b_\infty(\alpha), \Sigma^b_\infty(\beta))$, $\mathcal{S}(S_2) = S_2(\alpha) + S_2(\beta)$, $\mathcal{S}S^i_2 = S^i_2(\Sigma^b_\infty(\alpha), \Sigma^b_\infty(\beta))$, etc.

latter formalization (in the same theory). On the other hand, negative results like the unprovability of superpolynomial bounds are apparently stronger than the corresponding statements in the latter formalization (the unprovability of polynomial size upper bounds). For the latter formalization various (unconditional) independence results for the theory PV are known, see [20, Sections 7.6 and 15.3] and [24]. The theories occurring in [43] correspond to very weak subtheories of PV.

The cryptographic assumption mentioned above is expressed by the phrase: *strong pseudo-random number generators do exist*, which says that there is $\varepsilon > 0$ such that for all n there is a function (a pseudo-random generator):

$$G_n : \{0,1\}^n \to \{0,1\}^{2n}$$

which is computable by a circuit of size $n^{O(1)}$ and has *the hardness* $H(G_n) \geq 2^{n^\varepsilon}$.

Here, $H(G_n)$ is the minimal S such that there is a circuit C of size S computing a Boolean function $C : \{0,1\}^{2n} \to \{0,1\}$ such that:

$$\mid Prob_{x\in\{0,1\}^n}[C(G_n(x)) = 1] - Prob_{y\in\{0,1\}^{2n}}[C(y) = 1] \mid \geq \frac{1}{S}.$$

The reader should consult [43] or [29] for further details.

We are prepared to reprove the main conditional independence result from [43]. The idea of the reduction of the independence to interpolation is the same as in [43], we only use more friendly propositional version proved in earlier sections. For the benefit of the reader we give all details.

Theorem 8.2 ([43]) *Assume that strong pseudo-random number generators do exist. Let $E(x,a)$ be a first order bounded formula and $t(N)$ a function such that $t(N) = (\log N)^{\omega(1)}$ holds in the natural numbers.*

Then the theory $S^2_2(\alpha)$ does not prove the circuit-size lower bound $t(N)$ for the function E_N, i.e. the theory $S^2_2(\alpha)$ does not prove the formula:

$$\forall a; LB(a, t(a), E_a) \ .$$

Proof :

Let $s(N)$ be any function such that $s(N) = (\log N)^{\omega(1)}$ and the parity of two circuits of size $\leq s(N)$ is computable in size $\leq t(N)$.

Consider the implication:

$$Comp(\alpha, N, s(N), f) \to \neg Comp(\beta, N, s(N), f \oplus E_N)$$

where $f \oplus E_n$ is the parity of functions f and E_N definable from f and E_N by the bounded formula:

$$\forall x, x \in f \oplus E_N \equiv (x \in f \not\equiv x \in E_N).$$

Assume that the implication fails. Then α codes, in particular, a circuit computing f while β codes a circuit computing $f \oplus E_n$, and hence the two circuits joined by $\oplus$ compute E_N. By the choice of $s(N)$ this implies $\neg LB(N, t(N), E_N)$. Hence to show that the lower bound to E_N is not provable in $S^2_2(\alpha)$ it is sufficient to demonstrate that $S^2_2(\alpha)$ does not prove the implication above.

Assume, for the sake of contradiction, that it does. Denote by

$$A_i(p_1, \dots, p_N, q_1, \dots, q_s)$$

the propositional formula formalizing that the computation on i coded in α yields the value $f(i)$ (with the atoms p_i translating $i \in f$ and q_j translating $j \in \alpha$), and similarly denote by:

$$B_i(p_1, \dots, p_N, r_1, \dots r_s)$$

the formula formalizing that the computation on i coded in β does not yield the value $f(j) \oplus E_N(j)$ (here r_j translate $j \in \beta$). We have $s = s(N)$.

By Theorem 8.1 the sequent:

$$A_1, \dots, A_N \longrightarrow B_1, \dots, B_N$$

has an LK-proof π of size $2^{(\log N)^{O(1)}}$ satisfying the hypothesis of Corollary 6.2.[2] By that corollary the implication admits an interpolant

$$I(p_1, \dots, p_N)$$

whose circuit-size is $\leq 2^{(\log N)^{0(1)}}$.

We shall further freely slight between the bits $p_1, \dots, p_N$ and the function $f \in \{0,1\}^N$ they define. Define the set U by:

$$U = \{f \in \{0,1\}^N \mid \exists q \bigwedge_i A_i(p, q)\}$$

and the set V by:

$$V = \{f \in \{0,1\}^N \mid \exists r \bigwedge_j \neg B_j(p, r)\} \ .$$

Note that the sets U and V are disjoint and that:

$$f \in U \ \equiv \ (f \oplus E_N) \in V$$

holds for any $f \in \{0,1\}^N$. This implies that the interpolant I, which separates U from V, has the following two properties:

[2] The reader familiar with the split theories of [43] might note at this point that our method yields a propositional counter-part of the interpolation for full split version of S^2_2 as well. The propositional proof obtained analogously through Theorem 8.1 and Corollary 6.2 will not be a resolution proof but rather a proof satisfying the hypothesis of Corollary 3.2; that is sufficient.

1. $\neg I(f) \rightarrow f \notin U$, for any f,

2. $I(f \oplus E_N) \rightarrow f \notin U$, for any f.

Define the property $P(f)$ of functions $f \in \{0,1\}^N$ by:

$$P(f) := \begin{cases} \neg I(f) & \text{if at least a half of function satisfy } \neg I \\ I(f \oplus E_N) & \text{otherwise.} \end{cases}$$

This property satisfies clearly all three conditions of the definition of a *natural property against $P/poly$*, see [44], with the effectivity condition weakened to the requirement that P is computable in quasi-polynomial size. However, as noted in [43], the proof of [44, Thm 4.1] works for this modification as well (take $k = n^\epsilon$ in place of $k = \epsilon n$ in its proof). Hence we conclude that the provability of the lower bound in $S_2^2(\alpha)$ implies the failure of the cryptographic assumption.

q.e.d.

9 A possibility of lower bounds for interpolation

It is important and interesting to find out for which proof systems one can prove a good interpolation theorem. As mentioned in the introduction it was noted in [30, 31, 32] that unless $\mathcal{NP} \cap co\mathcal{NP} \subseteq \mathcal{NC}^1/poly$ (resp. $\subseteq \mathcal{P}/poly$) one cannot bound the size (formula- or circuit-) of an interpolant in terms of the size of the implication. However, it appears to be more difficult to find a natural computational complexity conjecture which would rule out, for example, that an interpolant of size polynomial in the number of sequents of an LK-proof of the implication always exists. The size of the interpolant means circuit-size as by the example in [19, Sect. 5] one cannot expect good bounds to formula-size unless $P/poly \subseteq NC^1/poly$.

In this section we discuss the possibility of such (conditional) lower bounds for interpolation. We shall freely slide between corresponding pairs of a propositional proof system P and a bounded arithmetic theory T. For these correspondences see [8, 22, 26, 27, 35, 21] or [20, Chpts 9 and 14]. Informally, this correspondence essentially says that a formula $A(a)$ is provable in T iff the propositional translations $\langle A \rangle_N$ of its instances have short P-proofs.

We shall first examine this question via the method of [22]. Let $TAUT$ be the set of propositional tautologies in the language of LK. Let U, V be two disjoint $\mathcal{NP}$-sets. Take a polynomial-time reduction f of U to the complement of $TAUT$ which is an $\mathcal{NP}$-complete set. Then:

$$f''V \subseteq TAUT$$

and if $V = \{x \mid \exists y B(x,y)\}$, where $B(x,y)$ is a polynomial-time predicate implicitly bounding y, then the function:

$$Q_f(w) := \begin{cases} f(x) & \text{if } w = \langle x, y \rangle \ \wedge B(x,y) \\ 1 & \text{otherwise.} \end{cases}$$

is a propositional proof system in the sense of [9] (not necessarily complete). A statement that $U \cap V = \emptyset$ merely means that Q_f is sound and, in fact, it is equivalent to it. Moreover, an interpolant I for the implication

$$Q_f(w) = a \longrightarrow a \in TAUT$$

gives immediately an interpolant $J(b) := I(f(b))$ for the original implication:

$$b \in V \longrightarrow b \notin U \ .$$

The strongest such reflection principle provable in T is for P itself (see [22]). That is, the instances of the reflection principle for P:

$$\langle Prf_P(w,a) \rangle_{M,N} \longrightarrow \langle b \models a \rangle_{N,N}$$

($Prf_P(w,a)$ formalizes that w is a P-proof of a and $b \models a$ formalizes that b is a satisfying assignment for a) are the strongest reflection principles admitting short P-proofs. (In fact, for most P these are as well the strongest tautologies - over some base system P_0 - with short P-proofs, see [22] or [20, Chpt. 14]). Hence P admits a feasible interpolation theorem iff these particular implications admit feasible interpolants. This yields the completeness result for the disjoint $\mathcal{NP}$-pair $(SAT^*, REF(P))$ proved directly in [41]. To make use of this observation one has to establish the correspondence (P,T) for usual P and T. For most P the corresponding T is known. For example, LK with the size measured by the number of steps (that is, the Extended Frege system, see [9]) corresponds to the theory V_1^1 (the particular correspondence of depth i subsystems of LK to $S_2^{i+2}(\alpha)$ relevant to [41] follows from [19, 2.2 and 1.2] as in the proof of Corollary 6.2).

Now we shall consider another approach to lower bounds for interpolation theorems. A simple corollary of Craig interpolation theorem is Beth definability theorem, see [3]. It says, in particular, that if the formula $A(p,q)$ *implicitly defines* the bit q_1, i.e. the implication

$$A(p,q), A(p,r) \longrightarrow q_1 \equiv r_1$$

is a tautology, then there is a function $E(p)$, an *explicit definition*, such that

$$A(p,q) \longrightarrow E(p) \equiv q_1$$

is a tautology as well. This follows from the interpolation theorem immediately considering the implication

$$(A(p, q) \wedge q_1) \longrightarrow (r_1 \vee \neg A(p, r))$$

whose any interpolant is an explicit definition of q_1. This simple reduction implies that the same bounds which hold for Craig interpolation theorem hold for Beth definability theorem (and vice versa, in fact). Hence to prove a lower bound for the interpolation theorem for a proof system P it is enough to find an implicit definition $A(p, q)$ of a function $g(p)$ admitting polynomial size P-proofs of the above implication, but which itself cannot be computed in $P/poly$.

The following particular definition of one-way functions is a slightly biased definition from [33]. It defines an apparently weaker notion then the more customary probabilistic versions (see e.g. [29]) but it suffices for our purposes.

Definition 9.1 *A function $f : \{0,1\}^* \to \{0,1\}^*$ is* one-way *iff*

1. *f is polynomial-time computable.*
2. *f is* 1-*to*-1.
3. *There are constants ϵ, k such that*

$$|x|^\epsilon \leq |f(x)| \leq |x|^k$$

holds for all x.

4. *The inverse function f^{-1}:*

$$f^{-1}(y) := \begin{cases} x & \text{if } y \in Rng(f) \text{ and } f(x) = y \\ 0 & \text{if } y \notin Rng(f) \end{cases}$$

is not in $P/poly$.

It is a simple observation that a one-way function exists iff $UP \not\subseteq P/poly$, where UP is the class of $\mathcal{NP}$-sets acceptable by a polynomial-time non-deterministic Turing machine with at most one accepting computation on every input. The following theorem follows from the remarks on connections between Craig interpolation and Beth definability theorems.

Theorem 9.2 *Let a propositional proof system P and an arithmetic theory T be a pair of corresponding proof systems (in the sense of [22] or [20, Chpts. 9 and 14]). Assume that $f(x)$ is a one-way function with an $\mathcal{NP}$-graph such that the theory T proves that f is* 1-*to*-1*:*

$$T \vdash f(x) = y, f(x') = y \longrightarrow x = x' \ .$$

Then the proof system P does not admit polynomial upper bound to the circuit-size of interpolants.

No one-way function is known at present but there appear to be two chief candidates (see [33, 29]), namely the (inverse function to the) *factoring*:

$$(p, q) \longrightarrow p \cdot q$$

mapping two primes p, q to the product $p \cdot q$, and the (inverse function to the) *discrete logarithm*:

$$(p, g, x) \longrightarrow (p, g, g^x \bmod p)$$

mapping a prime p, a primitive root g mod p and $x \in \mathbf{Z}_p^*$ to p, g and g^x mod p. These function have $\mathcal{NP}$-graphs as primes are in $\mathcal{NP} \cap co\mathcal{NP}$, see [38].

The hypothesis of the theorem can be fulfilled for Extended Frege system and a conjectured one-way function f based on the RSA encryption scheme, see [25].

The theorem makes sense, however, also for proof systems P working only with constant depth formulas (or clauses) as the graph of f can be reduced to the satisfiability of a set of clauses. It would be interesting to know whether the hypothesis of the theorem can be satisfied for the depth 1 subsystem of LK.

For the monotone interpolation we have an unconditional lower bound, essentially contained in [43, Section 8].

Theorem 9.3 *The depth 2 subsystem of LK does not admit a polynomial bound for the monotone interpolation theorem (Lyndon theorem).*
In particular, the set of clauses:

$$Clique_{n,\omega} \cup Color_{n,\xi}$$

for $n := \xi^4$ and $\omega := \xi^2$ has a depth 2 LK-refutation of size $2^{(\log n)^{O(1)}}$ but the implication:

$$\bigwedge Clique_{n,\omega} \longrightarrow \bigvee \neg Color_{n,\xi}$$

has no monotone interpolant of monotone circuit-size smaller than $2^{n^{\Omega(1)}}$.

Proof :
Work in bounded arithmetic. Let $G = (V, E)$ be a graph with n nodes, and let $F_1 : \{1, \ldots, \omega\} \longrightarrow V$ and $F_2 : V \longrightarrow \{1, \ldots, \xi\}$ be two maps.

As noted in [43] the implication that if F_1 is a 1-to-1 map onto a clique in G then F_2 cannot be a coloring of G follows from the weak pigeonhole. This principle is needed for a function f which is $\Delta_1^b(G, F_1, F_2)$-definable:

$$f(u) = a \equiv (\exists i \leq n, F_1(u) = i \wedge F_2(i) = a)$$

$$\equiv (\forall i \leq n, F_1(u) \neq i \vee F_2(i) = a) \ .$$

It says that f cannot be an injective map from $\omega = \xi^2$ into ξ and is (by [37]) provable in $T_2^3(G, F_1, F_2)$. The bounded arithmetic proof is translated into a tree-like, depth 4 LK-proof in which every sequent contains at most constantly many depth 4 formulas.

The existence of such a proof implies, analogously with the proof of Corollary 6.2, the existence of the wanted depth 2 LK-refutation.

q.e.d.

The only subsystem of LK for which the possibility of a feasible monotone interpolation is open is the depth 1 subsystem. We conjecture that this system does not admit a polynomial upper bound for the monotone interpolation, but we observe that the existence of such a bound would have interesting consequences.

Theorem 9.4 *Assume that the depth 1 subsystem of LK admits a polynomial upper bound to the monotone interpolation theorem.*

Then for any fixed ℓ, any resolution refutation of the clauses of the weak pigeonhole principle $\neg WPHP_m^{m^\ell}$:

1. $\bigvee_{a \leq m} f_{u,a}$, *all* $u \leq m^\ell$,
2. $\neg f_{u_1,a} \vee f_{u_2,a}$, *all* $u_1 < u_2 \leq m^\ell$ *and* $a \leq m$,

must have at least $2^{m^{\Omega(1)}}$ clauses.

Proof :

Take $\omega := m^\ell$, $\xi := m$ and $n := m^{\ell+1}$. Assume that π is a resolution refutation of the clauses of $\neg WPHP_m^{m^\ell}(f_{u,a})$. For any clause C in literals $f_{u,a}, \neg f_{u,a}$ denote by $\tilde{C}$ a clause in literals $q_{u,i}, \neg q_{u,i}, r_{i,a}, \neg r_{i,a}$ obtained from C as follows:

- replace every occurrence of $f_{u,a}$ by the sequence $q_{u,1} \wedge r_{1,a}, \ldots, q_{u,n} \wedge r_{n,a}$
- replace every occurrence of $\neg f_{u,a}$ by the sequence $q_{u,1} \wedge \neg r_{1,a}, \ldots, q_{u,n} \wedge \neg r_{n,a}$.

Observe two simple facts:

1. For every $C \in \neg WPHP_m^{m^\ell}(f_{u,a})$ the clause $\tilde{C}$ has a depth 1 LK-proof from the clauses $Clique_{n,\omega} \cup Color_{n,\xi}$ with $m^{O(1)}$ sequents.
2. If C can be derived by the resolution rule from C_1 and C_2 then $\tilde{C}$ has depth 1 LK-derivation from $\tilde{C}_1$ and $\tilde{C}_2$ with $m^{O(1)}$ sequents.

Hence we get a depth 1 LK-refutation of $Clique_{n,\omega} \cup Color_{n,\xi}$ with $m^{O(1)} \cdot k(\pi)$ sequents. Assuming that a polynomial upper bound holds for the monotone interpolation theorem for the system we get by Theorem 7.2:

$$(m \cdot k(\pi))^{O(1)} \geq 2^{n^{\Omega(1)}} = 2^{m^{\Omega(1)}}$$

and so

$$k(\pi) \geq 2^{m^{\Omega(1)}}$$

as well.

q.e.d.

At present such lower bounds are known only for $\ell = 2 - \Omega(1)$, see [6]. Another interesting corollary to the feasible interpolation for the above system would be an independence proof of the $\forall\Sigma_1^b(f)$-formula $WPHP(f)$ from $T_2^2(f)$ (similarly with the proof of Theorem 9.3). Such independence is known only for $S_2^2(\alpha)$ and $S_2^2(f)$ (for several formulations of $WPHP$); for $T_2^2(\alpha)$ only $\Sigma_2^b(\alpha)$-independent formulas are known (see [18, 7] or [20, Chpt. 11]).

Acknowledgement

A part of this work was done while I was visiting the Department of Mathematics of the University of California at San Diego in April 1994. I replaced the original statements about the linear equational calculus over $\mathbf{Q}$ by present Corollaries 6.4 and 7.3(2.) after learning about the proof system CP^* and the result of [4] from a lecture by M. L. Bonet at the meeting *Logic and Computational Complexity* (Indianapolis, October 1994). I thank A.A.Razborov for explaining to me the remarks on one-way functions made in [43, Sect. 8] but not described there.

I thank P. Pudlák and J. Sgall for helpful comments on the preliminary version of this paper.

References

[1] Alon, N., and Boppana, R. (1987) The monotone circuit complexity of Boolean functions, *Combinatorica*, **7(1)** : 1-22.

[2] Andreev, A. E. (1985) On a method for obtaining lower bounds for the complexity of individual monotone functions (in Russian), *Doklady AN SSSR*, **282(5)** : 1033-1037.

[3] Beth E. W. (1959) The foundations of mathematics. North - Holland, Amsterdam.

[4] Bonet, M. L., Pitassi, T., and Raz, R. (1994) Lower bounds for cutting planes proofs with small coefficients, preprint.

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

[6] Buss, S. R., and Turán, G. (1988) Resolution proofs of generalized pigeonhole principles, *Theoretical Computer Science*, **62**: 311-317.

[7] Chiari, M., and Krajíček, J. (1994) Witnessing functions in bounded arithmetic and search problems, submitted.

[8] 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.

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

[10] Cook, W., Coullard, C. R., and Turán, G. (1987) On the complexity of cutting plane proofs, *Discrete Applied Mathematics*, **18**: 25-38.

[11] Craig, W. (1957) Linear reasoning: A new form of the Herbrand-Gentzen theorem, *Journal of Symbolic Logic*, **22(3)**: 250-268.

[12] —— (1957) Three uses of the Herbrand-Gentzen theorem in relating model theory and proof theory, *Journal of Symbolic Logic*, **22(3)**: 269-285.

[13] Friedman, H. (1976) The complexity of explicit definitions, *Advances in mathematics*, **20**: 18-29.

[14] Gurevich, Y. (1984) Towards logic tailored for computational complexity, Proc. *Logic Colloquium 1983*, Springer LNM 1104: 175-216.

[15] Haken,A. (1985) The intractability of resolution, *Theoretical Computer Science*, **39**: 297-308.

[16] Karchmer, M., and Wigderson, A. (1988) Monotone circuits for connectivity require super - logarithmic depth, in: *Proc.* 20^{th} *Annual ACM Symp. on Theory of Computing*, pp.539-550. ACM Press.

[17] Krajíček, J. (1989) Exponentiation and second-order bounded arithmetic, *Annals of Pure and Applied Logic*, **48**: 261-276.

[18] —— , (1992) No counter-example interpretation and interactive computation, in: *Logic From Computer Science*, Proceedings of a Workshop held November 13-17, 1989 in Berkeley, ed. Y.N.Moschovakis, *Mathematical Sciences Research Institute Publication*, **21**: 287-293. Springer-Verlag.

[19] —— (1994) Lower bounds to the size of constant-depth propositional proofs, *Journal of Symbolic Logic*, **59(1)**: 73-86.

[20] —— (1994) *Bounded arithmetic, propositional logic and complexity theory*, Cambridge University Press, in print.

[21] —— (1995) On Frege and Extended Frege proof systems, in: *Feasible Mathematics II*, eds. P. Clote and J. Remmel, Birkhauser, pp.284-319.

[22] Krajíček, J., and Pudlák, P. (1989) Propositional proof systems, the consistency of first order theories and the complexity of computations, *J. Symbolic Logic*, **54(3)**:1063-1079

[23] —— (1990) Quantified propositional calculi and fragments of bounded arithmetic, *Zeitschrift f. Mathematikal Logik u. Grundlagen d. Mathematik*, **36**:29-46.

[24] —— (1990) Propositional provability in models of weak arithmetic, in: *Computer Science Logic (Kaiserlautern, Oct. '89)*, eds.E.Boerger, H.Kleine-Bunning and M.M. Richter, LNCS **440**: 193-210. Springer-Verlag.

[25] —— (1995) Some consequences of cryptographical conjectures for S^1_2 and EF, in: Proc. of the meeting *Logic and Computational Complexity*, (Indianapolis, October 1994), Ed. D. Leivant, to appear.

[26] Krajíček, J., and Takeuti, G. (1990) On bounded $\sum^1_1$-polynomial induction, in: *Feasible Mathematics*, eds. S.R. Buss and P.J. Scott, pp.259-280. Birkhauser.

[27] —— (1992) On induction-free provability, *Annals of Mathematics and Artificial Intelligence*, **6**: 107-126.

[28] Kreisel, G. (1961) Technical report nb.3, Applied Mathematics and Statistics Labs, Stanford University, unpublished.

[29] Luby M. (1993) Pseudo-randomness and applications, International Computer Science Institute , Berkeley, lecture notes.

[30] Mundici, D. (1983) A lower bound for the complexity of Craig's interpolants in sentential logic, *Archiv fur Math. Logik*, **23** : 27-36.

[31] —— (1984) Tautologies with a unique Craig interpolant, uniform vs. nonuniform complexity, *Annals of Pure and Applied Logic*, **27**, pp.265-273.

[32] —— (1984) NP and Craig's interpolation theorem, Proc. *Logic Colloquium 1982*, North-Holland, pp. 345-358.

[33] Papadimitriou, A. (1994) *Computational complexity*, Addison-Wesley.

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

[35] Paris, J., and Wilkie, A. J. (1985) Counting problems in bounded arithmetic, in: *Methods in Mathematical Logic*, Springer LNM **1130** : 317-340.

[36] —— (1987) On the scheme of induction for bounded arithmetic formulas, *Annals of Pure and Applied Logic*, **35** : 261-302.

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

[38] Pratt, V.R. (1975) Every prime has a succinct certificate, *SIAM J. Computing*, **4**, pp.214-220.

[39] Razborov, A. A. (1985) Lower bounds on the monotone complexity of some Boolean functions, *Soviet Mathem. Doklady*, **31** : 354-357.

[40] —— (1993) An Equivalence between second order bounded domain bounded arithmetic and first order bounded arithmetic, in: *Arithmetic,Proof Theory and Computational Complexity*, eds. P. Clote and J. Krajíček, pp.247-277. Oxford University Press.

[41] —— (1994) On provably disjoint NP-pairs, preprint.

[42] —— (1995) Bounded arithmetic and lower bounds in Boolean complexity, in : *Feasible Mathematics II*, eds. P. Clote and J. Remmel, Birkhauser, pp.344-386.

[43] —— (1995) Unprovability of lower bounds on the circuit size in certain fragments of bounded arithmetic, *Izvestiya of the R.A.N.*, **59(1)**, pp.201-224.

[44] Razborov, A. A., and Rudich, S. (1994) Natural proofs, in: *Proc. 26^{th} Annual ACM Symp. on Theory of Computing*, pp. 204-213. ACM Press.

[45] Takeuti, G. (1975) *Proof theory*, North-Holland.

[46] —— (1993) RSUV isomorphism, in: *Arithmetic, Proof Theory and Computational Complexity*, eds. P. Clote and J. Krajíček, pp.364-386. Oxford University Press.