Feasible interpolation for lifted sequents

Phuong Nguyen [*]
McGill University

May 13, 2011

Abstract

The idea of feasible interpolation for propositional proof systems is to derive lower bounds for propositional proofs using circuit lower bounds for Craig's interpolant. However, as far as we know, proof systems such as constant-depth Frege do not admit feasible interpolation. We extend the notion of feasible interpolation so that it is admitted by a number of treelike propositional proof systems. This allows us to derive new lower bounds for treelike Frege proof systems and new conditional lower bounds for treelike Frege proof systems with modular counting connectives (all of constant depth). We obtain our results by augmenting Krajíček's argument from [Kra97] with the idea from Maciel and Pitassi [MP06].

1 Introduction

Craig interpolation theorem for propositional logic states that if $A(\vec{p},\vec{q}) \supset B(\vec{p},\vec{r})$ is valid, where $A(\vec{p},\vec{q})$ and $B(\vec{p},\vec{r})$ are propositional formulas with all free variables displayed, then there is a formula $I(\vec{p})$, i.e., an interpolant, such that both $A(\vec{p},\vec{q}) \supset I(\vec{p})$ and $I(\vec{p}) \supset B(\vec{p},\vec{r})$ are valid. The problem of determining the complexity of the interpolant is interesting because of the following observation, due to Mundici, that if we can always find an interpolant I of size polynomial in the size of A and B, then $\mathbf{NP} \cap co\text{-}\mathbf{NP} \subseteq \mathbf{NC}^1/poly$. So far superpolynomial lower bounds on the size of I are known only under the restriction that it is a monotone formula (or more generally a monotone circuit) [AB87].

Krajíček's notion of feasible interpolation for a propositional proof system $\mathcal{P}$ stems from a related problem, where now a $\mathcal{P}$-proof of the implication $A(\vec{p},\vec{q}) \supset B(\vec{p},\vec{r})$ is also given [Kra94, Kra97]. A propositional proof system $\mathcal{P}$ is said to admit feasible interpolation if given any $\mathcal{P}$-proof π of an implication $A(\vec{p},\vec{q}) \supset B(\vec{p},\vec{r})$, an interpolant $I(\vec{p})$ can be computed by a boolean circuit of size polynomial in the size of π. If $\mathcal{P}$ has feasible interpolation, then the task of proving super polynomial lower bounds for $\mathcal{P}$ can be reduced to proving

[*] Supported by an NSERC Postdoctoral Fellowship. Email: pnguyen@cs.toronto.edu

super polynomial lower bound for circuits that compute the interpolants. (The idea of such a reduction is also used in [BPR97].) As a result, the exponential lower bound for monotone circuits [Raz85, AB87, Pud97] gives exponential lower bounds for several propositional proof systems, including Resolution [Kra97] and Cutting Plane [BPR97, Pud97]. See also [Raz95, Kra98, IPU94, Pud99, PS98].

On the other hand, under some plausible complexity assumption the propositional proof systems for which we are currently not able to prove super polynomial lower bounds do not admit feasible interpolation. For example, it is shown in [KP98] that unless the RSA cryptographic scheme is not secure, extended Frege systems do not have feasible interpolation. By a similar approach, it is shown in [BPR00] that unless the Diffie–Hellman secret key exchange can be broken by polynomial size circuits, $\mathbf{TC}^0$-Frege systems do not have feasible interpolation. Recently, it is shown that a certain depth-3 Frege system does not have feasible interpolation unless there are polytime algorithm solving the so-called mean-payoff game problem [AM11]. Note, however, that these results are proved by showing that the existence of polynomial size interpolants for some particular tautologies would violate some complexity hypothesis (i.e., the security of RSA or Diffie–Hellman protocols, or intractability of the mean-payoff games). It is still plausible that the above systems have feasible interpolation for *other* tautologies.

A recent paper reviving interest in feasible interpolation is [Kra10]. In this paper Krajíček introduces a notion called *chain feasible interpolation*. The intuition behind this is as follows. Suppose that there are constant-size Frege proofs of a tautology of the form

$$\neg\big(C_1 \vdash \Phi \wedge (\bigwedge_{i=1}^{m-1} C_i \cong C_{i+1}) \wedge C_m \vdash \Psi\big)$$

where C_i are first-order L-structures (for some signature L) and Φ, Ψ are two $\mathbf{\Sigma}_1^1$ sentences that cannot be satisfied simultaneously in an L-structure. Then there ought to be a first-order L-sentence γ that separates Φ and Ψ, i.e., such that for all structures A:

$$(A \vdash \Phi \Rightarrow A \vdash \gamma) \wedge (A \vdash \Psi \Rightarrow A \vdash \neg\gamma)$$

Now if we can show that no such γ exists for some signature L and sentences Φ, Ψ, then the above tautology does not admit short constant-size Frege proofs.

Here we will consider another approach by using the idea of [MP06]; see also [Ngu07, MNP11]. The intuition behind these papers is as follows. Suppose that some sequent $\mathcal{S}$ requires cut-free Frege proofs of size at least s. Then a "lifted" version of $\mathcal{S}$, i.e., the sequent that is obtained from $\mathcal{S}$ by substituting formulas that express the Parity function for all variables appearing in $\mathcal{S}$, should require constant-depth Frege proofs of size at least s as well. So far this is still an open problem, although under some special settings lower bounds for the proofs of the lifted sequents can be proved, see [MNP11]. This kind of "lifting" has already be considered in [Kra94], although there the lifting functions are Sipser's functions.

The problem can be stated more generally for other propositional proof systems, such as constant-depth Frege augmented with modulo counting connectives (the so-called **ACC**-Frege systems) or threshold connectives ($\mathbf{TC}^0$-Frege). The Parity function must be replaced by other functions that satisfy certain hardness conditions against the class of cut formulas under consideration (i.e., **ACC** circuits or $\mathbf{TC}^0$ circuits). We call these functions the *lifting* functions.

The kind of interpolation property we consider in this paper is as follows. A propositional proof system $\mathcal{P}$ is said to have *lifted feasible interpolation*, or just *lifted interpolation*, provided that given a proof π of a lifted version of the implication $A(\vec{p}, \vec{q}) \supset B(\vec{p}, \vec{r})$, there is an interpolant $C(\vec{p})$ that can be computed by a circuit of size polynomial in the size of $A(\vec{p}, \vec{q}) \supset B(\vec{p}, \vec{r})$ and the *length* of π (i.e., the number of sequents in π).

Our proof is an adaptation of Krajíček's proof of the feasible interpolation property for Resolution and cut-free **LK** [Kra97, Theorem 3.1]. The idea is to show that the proof of the implication can be used to construct a protocol for the associated KW game. In order to make this construction possible in the presence of the cut rule, we need to use the fact that the lifting function is hard to compute by the cut formulas. In essence, the two players in the KW game can always make progress just as in the case of cut-free proofs, because there are always enough truth assignments for the variables in the lifted sequents that satisfy the players' choice as well as the cut formulas. Our argument, however, only works for treelike proofs, because it seems indispensable that the players be able to figure our their past moves.

Our result is presented and proven in Section 2. First, we give some basic definitions.

1.1 Proof systems

We use sequent calculi. **PK** is Gentzen's propositional calculus over the set of connectives $\{\wedge, \vee, \neg\}$; see for example [CN10, Chapter II]. The axioms of **PK** consist of

$$p \longrightarrow p, \qquad \bot \longrightarrow \ , \qquad \longrightarrow \top$$

In Figure 1 we list only the logical rules of **PK**. Apart from these it has also the usual structural rules: exchange, weakening, and contraction.

We will also discuss extensions of **PK** using connectives $\oplus_r^b$. The meaning of $\oplus_r^b(A_1, A_2, \ldots, A_m)$ is that, modulo r, exactly b formulas among $A_1, A_2, \ldots, A_m$ are true. For $r \geq 2$, $\mathbf{PK}[r]$ denotes the extension of **PK** by having the connectives $\oplus_r^b$ for $0 \leq b < r$. The axioms of $\mathbf{PK}[r]$ consist of those of **PK**, together with

$$A \longrightarrow \oplus_r^1(A), \qquad \oplus_r^1(A) \longrightarrow A$$

The rules for these connectives are given in Figure 2. Here F stands for a list of formulas.

The depth of a *literal* is 0. The depth of other formulas is defined inductively as usual. For a system $\mathcal{P}$ among **PK** and $\mathbf{PK}[r]$, depth-d $\mathcal{P}$ (or just d-$\mathcal{P}$) is the subsystem where all cut formulas have depth at most d; note that we can

$$\frac{\Gamma \longrightarrow \Delta, A}{\neg A, \Gamma \longrightarrow \Delta}\ \neg\text{-left} \qquad \frac{A, \Gamma \longrightarrow \Delta}{\Gamma \longrightarrow \Delta, \neg A}\ \neg\text{-right}$$

$$\frac{A, B, \Gamma \longrightarrow \Delta}{(A \wedge B), \Gamma \longrightarrow \Delta}\ \wedge\text{-left} \qquad \frac{\Gamma \longrightarrow \Delta, A \qquad \Gamma \longrightarrow \Delta, B}{\Gamma \longrightarrow \Delta, (A \wedge B)}\ \wedge\text{-right}$$

$$\frac{A, \Gamma \longrightarrow \Delta \qquad B, \Gamma \longrightarrow \Delta}{(A \vee B), \Gamma \longrightarrow \Delta}\ \vee\text{-left} \qquad \frac{\Gamma \longrightarrow \Delta, A, B}{\Gamma \longrightarrow \Delta, (A \vee B)}\ \vee\text{-right}$$

$$\frac{\Gamma \longrightarrow \Delta, A \qquad A, \Gamma \longrightarrow \Delta}{\Gamma \longrightarrow \Delta}\ \text{cut}$$

Figure 1: Logical rules of **PK**

$$\frac{A, \oplus_r^{b-1}(F), \Gamma \longrightarrow \Delta \qquad \oplus_r^{b}\,(F), \Gamma \longrightarrow A, \Delta}{\oplus_r^{b}(A, F), \Gamma \longrightarrow \Delta}\ \text{MOD-left}$$

$$\frac{A, \Gamma \longrightarrow \oplus_r^{b-1}(F), \Delta \qquad \Gamma \longrightarrow A, \oplus_r^{b}(F), \Delta}{\Gamma \longrightarrow \oplus_r^{b}(A, F), \Delta}\ \text{MOD-right}$$

$$\frac{\Gamma \longrightarrow \oplus_r^{a}(F), \Delta \qquad \Gamma \longrightarrow \oplus_r^{b}(G), \Delta}{\Gamma \longrightarrow \oplus_r^{a+b}(F, G), \Delta}\ \text{MOD-add}$$

$$\frac{\Gamma \longrightarrow \oplus_r^{a}(F, G), \Delta \qquad \Gamma \longrightarrow \oplus_r^{b}(G), \Delta}{\Gamma \longrightarrow \oplus_r^{a-b}(F), \Delta}\ \text{MOD-substract}$$

Figure 2: Rules for modular connectives

still prove sequent of arbitrary depth in this system. The size of a formula A, $size(A)$, is the total number of symbols in it; the size of a sequent $\mathcal{S}$ (or a proof π), denoted by $size(\mathcal{S})$ (resp. $size(\pi)$), is the sum of the size of all formulas in $\mathcal{S}$ (or π). Also, the number of sequents in π, i.e. the *length* of π, is denoted by $st(\pi)$. For a proof system $\mathcal{P}$, $\mathcal{P}^{\star}$ denotes its subsystem where we require that the proofs are treelike.

Definition 1.1. *Suppose that $\mathcal{S}(\vec{p}, \vec{q}, \vec{r})$ is a valid sequent, with all free variables shown, of the form*

$$\Gamma(\vec{p}, \vec{q}) \longrightarrow \Delta(\vec{p}, \vec{r})$$

Then an interpolant for $\mathcal{S}$ is a formula $I(\vec{p})$ such that both

$$\Gamma(\vec{p}, \vec{q}) \longrightarrow I(\vec{p}), \qquad \text{and} \qquad I(\vec{p}) \longrightarrow \Delta(\vec{p}, \vec{r})$$

are valid.

We will view I as a boolean function, and will be interested in its circuit complexity.

1.2 Protocol for Karchmer–Wigderson game

Given two disjoint sets $U, V \subset \{0,1\}^n$, the Karchmer–Wigderson (KW) game [KW88] is played by two players, called Alice and Bob, as follows. Alice gets an element $\vec{u} \in U$ and Bob gets an element $\vec{v} \in V$. They exchange bits of information in order to agree on an index i where u_i and v_i differ, and we are interested in the minimal number of bits that they need to communicate.

The notion of a protocol from [Kra97] is very well suited for proving feasible interpolation theorem for propositional proof systems. A protocol for the KW game consists of:

- a directed acyclic graph G,
- a *strategy function*, *Next*, that takes inputs of the form $(\vec{u}, \vec{v}, x)$ where $\vec{u} \in U, \vec{v} \in V$ and x is a vertex of G, and outputs a node y such that (x, y) is an edge of G,
- a collection $\{F(\vec{u}, \vec{v}) \ : \ \vec{u} \in U, \vec{v} \in V\}$ where each $F(\vec{u}, \vec{v})$ is a subset of vertices of G and is called the *consistency condition* for $(\vec{u}, \vec{v})$.

These must satisfy the following conditions:

1. G has a single vertex of in-degree 0 which we call *root*. The vertices of G of out-degree 0 are called leaves and are labelled by a formula of the form $u_i = 0 \wedge v_i = 1$ or $u_i = 1 \wedge v_i = 0$.
2. For every pair $\vec{u} \in U, \vec{v} \in V$ the consistency condition $F(\vec{u}, \vec{v})$ has the property that it contains *root*, the labels at the leaves in $F(\vec{u}, \vec{v})$ are valid for $(\vec{u}, \vec{v})$, and that for each $x \in F(\vec{u}, \vec{v})$ the path $P^x_{\vec{u},\vec{v}}$ in G, which starts at x and is determined by the strategy function *Next* with inputs $(\vec{u}, \vec{v})$, is contained in $F(\vec{u}, \vec{v})$.

A protocol is *monotone* if the labels of the leaves are of the form $u_i = 1 \wedge v_i = 0$ only.

The communication complexity of the protocol is defined to be the minimum number of communication bits that is required for the two players, one knowing u and the other knowing v, to (i) determine whether an arbitrary node x belongs to $F(\vec{u}, \vec{v})$, and (ii) to computes the next node *Next*$(\vec{u}, \vec{v}, x)$. We will also call a vertex of G a vertex of the protocol. (Because of (i), the consistency condition $F(\vec{u}, \vec{v})$ above cannot be taken to be simply the path $P^{root}_{\vec{u},\vec{v}}$.)

The following theorem relates the complexity of a protocol to the size of a circuit that separates U and V:

Theorem 1.2 ([Raz95, Kra97]). *Let U, V be two disjoint subsets of $\{0,1\}^n$. Suppose that there is a protocol (resp. a monotone protocol) for the KW game on (U, V) with k vertices and communication complexity t. Then there is a circuit (resp. a monotone circuit) of size $k2^{\mathcal{O}(t)}$ that separates U and V.*

Conversely, suppose that there is a circuit (resp. monotone circuit) of size s that separates U and V. Then there is a protocol (resp. monotone protocol) for the KW game on (U, V) that has s vertices and communication complexity 2.

This theorem can be used to prove the feasible interpolation theorem for resolution and some other proof systems [Kra97, Theorem 3.1]. The important property of these systems that is needed is that the sequents in a proof can be evaluated using only a small amount of communication by the two players who are given appropriate truth assignments to the variables. This property no longer holds for proof systems where there are cuts, even when the cut formulas are of depth one.

Here we observe that if the proof that we have is a *treelike* proof of the lifted version of a tautology $\Gamma(\vec{p}, \vec{q}) \longrightarrow \Delta(\vec{p}, \vec{r})$ (see Definition 2.1 below), then the two players can play in much the same way as before by virtually ignoring the cut formulas. This is because they can always find truth assignments to the variables that make their move consistent. The treelikeness is crucial for our argument. Essentially, it allows the two players, who are given a sequent in the proof, to completely determine the unique path from the endsequent to the given sequent. This means, in particular, that the players know earlier moves in the protocol, so that they can determine (without communication) any *future* moves that are on a sequent which is derived by the same inference as in a past move.

Once feasible interpolation property is established, the following corollary of [AB87] suffices to gives concrete lower bound:

Theorem 1.3. *For n sufficient large, any monotone circuit that decides whether a graph on n vertex contains a clique of size $\sqrt{n}$ has size at least $2^{n^{1/4}}$.*

Feasible interpolation and the above circuit lower bound are combined to give lower bounds for proofs of the following sequent:

$$\mathit{Clique}_{n,\sqrt{n}} \longrightarrow \neg \mathit{Color}_{n,\sqrt{n}-1} \tag{1}$$

where $\mathit{Clique}_{n,k}$ and $\mathit{Color}_{n,\ell}$ are defined for $n, k, \ell \geq 1$ as follows.

First, $\mathit{Clique}_{n,k}$ denotes the set of the following formulas, which together express that the graph on $\{1, 2, \ldots, n\}$ that is encoded by $\vec{p}$ has a clique of size at least k (there is an edge between two vertices j_1, j_2 iff p_{j_1,j_2} is true):

1. $\bigvee_{j=1}^{n} q_{i,j}$ for $1 \leq i \leq k$ ($q_{i,j}$ says that the i-th vertex in the clique is vertex j),

2. $\neg q_{i_1,j} \vee \neg q_{i_2,j}$ for $1 \leq j \leq n$, $1 \leq i_1 < i_2 \leq k$ (no vertex is listed twice in the clique),

3. $\neg q_{i_1,j_1} \vee \neg q_{i_2,j_2} \vee p_{j_1,j_2}$ for $1 \leq i_1 < i_2 \leq k$ and $1 \leq j_1 \neq j_2 \leq n$ (there is an edge between any two vertices in the clique).

Next, $\mathit{Color}_{n,\ell}$ denotes the set of the following formulas, which together express that the graph encoded by $\vec{p}$ as above can be colored properly with ℓ colors:

1. $\bigvee_{i=1}^{\ell} r_{i,j}$ for $1 \leq j \leq n$ (each vertex j is colored by some color),

2. $\neg r_{i_1,j} \vee \neg r_{i_2,j}$ for $1 \leq i_1 < i_2 \leq \ell$, $1 \leq j \leq n$ (each vertex is colored by only one color),
3. $\neg r_{i,j_1} \vee \neg r_{i,j_2} \vee \neg p_{j_1,j_2}$ for $1 \leq i \leq \ell$, $1 \leq j_1 < j_2 \leq n$ (no edge has two endpoints of the same color).

Also, let $\neg Color_{n,\ell}$ denote the set of the negation of the formulas in $Color_{n,\ell}$, written in negation normal form.

2 Interpolation theorem for lifted sequents

Definition 2.1. *Let $\mathcal{S}$ be a sequent with variables $\vec{p} = p_1, p_2, \ldots, p_n$ and f a formula on m variables. Then $\mathcal{S}(f)$ denotes the sequent obtained from $\mathcal{S}$ by replacing each variable p_i by $f(\mathbf{x}_i)$, where $\mathbf{x}_i = x_{i,1}, x_{i,2}, \ldots, x_{i,m}$ is a new set of distinct variables.*

Below we will denote the new set of variables for p_i, q_i, r_i by $\mathbf{x}_i, \mathbf{y}_i, \mathbf{z}_i$, respectively.

The sequent $\mathcal{S}(f)$ is called a *lifted* sequent of $\mathcal{S}$. We will use $Parity_m$ (or just $Parity$) for f, where $Parity_m(x_1, x_2, \ldots, x_m)$, is the parity of the number of 1 among $x_1, x_2, \ldots, x_m$. We need the following crucial property of this function:

Theorem 2.2 (Håstad [Hås86]). *If C is a depth-d $\mathbf{AC}^0$ circuit of size $2^{n^{1/(d+1)}}$, then, for sufficiently large n, C cannot compute Parity correctly on more than a $1/2 + 1/2^{n^{1/(d+1)}}$ fraction of the inputs.*

Using this we can prove the following lemma; for a proof see [Ngu07, Cor. 4.8], [MNP11, Lemma 5.4].

Lemma 2.3. *Fix a $d \geq 1$. Let m be sufficiently large, and $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ be disjoint sets of variables, where each $\mathbf{x}_i$ has m variables $x_{i,1}, x_{i,2}, \ldots, x_{i,m}$. Let C be any formula of depth $d+1$ and size at most $2^{m^{1/(d+2)}}$ (C may contain variables other than the $\mathbf{x}$). Suppose that there are at least a fraction of $4^k/2^{m^{1/(d+2)}}$ truth assignments to the variables in C that satisfy C. Then, given any values $u_1, u_2, \ldots, u_k \in \{0,1\}^k$, there exists a truth assignment that satisfies C and that gives $Parity(\mathbf{x}_i)$ the value u_i, for $1 \leq i \leq k$.*

This lemma is used for showing that, under proper setting of the parameters, the function *Next* is well defined for an internal node of the protocol.

Now we state our main theorem. Recall that $st(\pi)$ denotes the number of sequents in a proof π.

Theorem 2.4. *Let $\mathcal{S}_0 = \Gamma_0(\vec{p}, \vec{q}) \longrightarrow \Delta_0(\vec{p}, \vec{r})$ be a valid sequent, and π a d-$\mathbf{PK}^\star$ proof of the lifted sequent $\mathcal{S}_0(Parity_m)$, for some constant $d \geq 1$. Suppose that the number of steps in π, $st(\pi)$, satisfies $4st(\pi) \leq 2^{m^{1/(d+2)}/size(\mathcal{S}_0)}$, and that no sequent in π has size greater than $2^{m^{1/(d+2)}}$. Then there is a circuit $I(\vec{p})$ of size at most $\mathcal{O}(st(\pi))$ that computes an interpolant for $\mathcal{S}_0$. Furthermore, if $\vec{p}$ only occur positively (or negatively) in Γ_0 or Δ_0, then the circuit is monotone.*

Note that the theorem can be stated using $\mathit{fst}(\pi)$ instead of $\mathit{st}(\pi)$, where $\mathit{fst}(\pi)$ is the number of *fresh* inferences in π, see Definition 2.12 below. This is always at most the total number of logical inferences in π (i.e., excluding the structural inferences).

Before proving the theorem, let us mention a standard application for a lifted version of the sequent (1).

Corollary 2.5. *Fix a constant* $1 \leq d \in \mathbb{N}$*. Let* n *be sufficiently large, and* $m \geq n^{4(d+2)}$*. Then any* d-$\mathbf{PK}^{\star}$ *proof of the sequent*

$$\mathit{Clique}_{n,\sqrt{n}}(\mathit{Parity}_m) \longrightarrow \neg \mathit{Color}_{n,\sqrt{n}-1}(\mathit{Parity}_m) \tag{2}$$

has at least $\Omega\big(2^{n^{1/4}}\big)$ *many sequents.*

Proof. Apply Theorem 2.4. Here let $\mathcal{S}_0$ be the sequent (1) which has size at most $3n^3$. Thus if $4\mathit{st}(\pi) \geq 2^{m^{1/(d+2)}/\mathit{size}(\mathcal{S}_0)}$, then $\mathit{st}(\pi) > 2^{n/3}/4$ and we are done. Also, if any sequent in π has size greater than $2^{m^{1/(d+2)}}$ then we are also done. So suppose that $\mathit{st}(\pi) < 2^{m^{1/(d+2)}/\mathit{size}(\mathcal{S}_0)}$, and that no sequent in π has size greater than $2^{m^{1/(d+2)}}$. The theorem gives an upper bound of $\mathcal{O}(\mathit{st}(\pi))$ for the size of a monotone circuit that computes the clique problem. So the conclusion follows from Theorem 1.3. □

The next theorem is proved in the same way as Theorem 2.4 and Corollary 2.5. Here we want to get a conditional lower bound for d-$\mathbf{PK}^{\star}[r]$, and we need a hypothetical hard function for $\mathbf{AC}^0[r]$ in the same way that *Parity* is hard for depth-d $\mathbf{ACC}$ circuits.

Conjecture 2.6. *Let* $r \geq 2$ *and* $d \geq 1$*. Suppose that there is a balanced function* f *that satisfies the following condition: if* C *is any depth-*$(d+1)$ $\mathbf{AC}^0[r]$ *circuit of size less than* $2^{n^{1/(d+2)}}$*, then for sufficiently large* n*,* C *does not compute* f *correctly for more than a fraction of* $1/2 + 1/2^{n^{1/(d+2)}}$ *inputs.*

For a function f we let f_m denote f when it takes exactly m inputs.

Theorem 2.7. *Suppose that Conjecture 2.6 is true, and let* f *be a function that satisfies the conjecture. Let* $\mathcal{S}_0 = \Gamma_0(\vec{p}, \vec{q}) \longrightarrow \Delta_0(\vec{p}, \vec{r})$ *be a valid sequent, and* π *a* d-$\mathbf{PK}^{\star}[r]$ *proof of the lifted sequent* $\mathcal{S}_0(f_m)$*. Suppose that the number of steps in* π*,* $\mathit{st}(\pi)$*, satisfies* $4\mathit{st}(\pi) \leq 2^{m^{1/(d+2)}/\mathit{size}(\mathcal{S}_0)}$*, and that no sequent in* π *has size greater than* $2^{m^{1/(d+2)}}$*. Then there is a circuit* $I(\vec{p})$ *of size at most* $\mathit{st}(\pi)^c$*, for some constant* $c > 0$*, that computes an interpolant for* $\mathcal{S}_0$*. Furthermore, if* $\vec{p}$ *only occur positively (or negatively) in* Γ_0 *or* Δ_0*, then the circuit is monotone.*

The proof of this theorem is almost identical to the proof of Theorem 2.4 given below. Note that although the new connectives have unbounded fanin, they have binary introduction rules.

The next corollary is derived in the same way as the previous corollary:

Corollary 2.8. *Suppose that Conjecture 2.6 is true, and let f be a function that exists by the conjecture. Then any d-$\mathbf{PK}^\star[r]$ proof of*

$$\mathit{Clique}_{n,\sqrt{n}}(f_m) \longrightarrow \neg \mathit{Color}_{n,\sqrt{n}-1}(f_m)$$

has size at least $2^{cn^{1/4}}$ for some constant $c > 0$.

Note that in order to get super-polynomial lower bounds, it suffices to consider such an f that can be computed by sub-exponential size circuits (e.g., circuits of size $2^{o(n)}$).

2.1 Proof of Theorem 2.4

Let n, s, t denote respectively the length of $\vec{p}, \vec{q}, \ \vec{r}$. Let

$$U = \{\vec{u} \in \{0,1\}^n \ : \ \exists \vec{q} \in \{0,1\}^s \bigwedge_{A \in \Gamma_0} A(\vec{u}, \vec{q})\}$$
$$V = \{\vec{v} \in \{0,1\}^n \ : \ \exists \vec{r} \in \{0,1\}^t \bigwedge_{B \in \Delta_0} \neg B(\vec{v}, \vec{r})\}$$

Because $\mathcal{S}_0$ is a valid sequent, $U \cap V = \varnothing$. By Theorem 1.2 it suffices to define a protocol for the game on (U, V) with at most $st(\pi)$ many nodes and constant communication complexity.

The nodes of the protocol are sequents of π. To describe the consistency condition $F(\vec{u}, \vec{v})$ we need a few notations. First, given $\vec{u}$, Alice fixes $\vec{q}$ that satisfy

$$\bigwedge_{A \in \Gamma_0} A(\vec{u}, \vec{q})$$

Similarly, given $\vec{v}$, Bob picks $\vec{r}$ that satisfy

$$\bigwedge_{B \in \Delta_0} \neg B(\vec{v}, \vec{r})$$

Because the proof is treelike, each formula in any sequent of π has at most one descendant in the endsequent $\mathcal{S}_0(\mathit{Parity})$. We say that a formula is *owned by Alice* if it has a descendant in the antecedent of $\mathcal{S}_0(\mathit{Parity})$. Any other formula is *owned by Bob.* For convenience, we will write $\mathbf{x}^A$ (resp. $\mathbf{x}^B$) for the variables $\mathbf{x}$ that appear in a formula owned by Alice (resp. Bob). A *pseudo truth assignment* for a sequent $\mathcal{S}$ in π assigns a single value for each variable in $\mathbf{y}$ and $\mathbf{z}$, but for each variable $x_{i,j}$ in $\mathbf{x}$ it may give two different values, one for $x^A_{i,j}$ and one for $x^B_{i,j}$.

The reason why our result apply for proofs with cut (instead of cut-free proofs) is essentially because we can ignore small depth formulas by looking only at the set of pseudo (partial) truth assignments that satisfy all depth-d formulas in the antecedents and falsify all depth-d formulas in the succedents. We call these *critical* pseudo (partial) truth assignments.

Definition 2.9. *A pseudo (partial) truth assignment for a sequent $\mathcal{S} = \Gamma \longrightarrow \Delta$ is said to be* critical *if it satisfies*

$$\bigwedge_{A \in \Gamma,\ depth(A) \le d} A \wedge \bigwedge_{B \in \Delta,\ depth(B) \le d} \neg B \tag{3}$$

Note that if there is no depth-d formula in $\mathcal{S}$, then all pseudo truth assignments are critical.

Ideally we want to focus only on "big" formulas in $\mathcal{S}$, i.e., formulas of the form $A(\mathit{Parity})$ which are obtained from substituting *Parity* formulas for the variables of a formula A. By restricting to critical pseudo truth assignments of $\mathcal{S}$ we almost achieve this; however, $\mathcal{S}$ may contain subformulas of *Parity* that have depth greater than d, and thus a critical pseudo truth assignment may become useless because, for example, it falsifies some such a formula in the antecedent of $\mathcal{S}$. To handle these formulas we make use of the treelike structure of the proof. The idea is that these formulas must come from some *Parity* formula that at some point on the path from the root of the proof stands by itself in the sequent (as opposed to being part of another formula). This means that the value of such a *Parity* formula must be known to both players. For example, if $\mathit{Parity}(\mathbf{x}_1^A)$ is a formula in the antecedent, then both players know that it is true. Consequently, the players can agree on a common truth assignment to the variables $\mathbf{x}_1^A$ without any communication.

Definition 2.10. *We say that a set of variables $\mathbf{x}_j^A$ (or $\mathbf{x}_j^B$, $\mathbf{y}_j$, $\mathbf{z}_j$) is* determined *for a sequent $\mathcal{S}$ if some occurrence of $\mathit{Parity}(\mathbf{x}_j^A)$ appears as a formula (as opposed to a proper subformula) in some sequent on the path from the endsequent $\mathcal{S}_0(\mathit{Parity})$ to $\mathcal{S}$.*

For each sequent $\mathcal{S}$ in π, the two players fix a common partial pseudo truth assignment to its determined variables without communication as follows. Suppose without loss of generality that $\mathbf{x}_1^A, \mathbf{y}_2, \mathbf{x}_3^B, \mathbf{z}_4, \ldots$ are determined variables for $\mathcal{S}$ and that they become determined in that order as we follow the path starting from the endsequent to $\mathcal{S}$. The partial pseudo truth assignment is defined inductively. For example, suppose that some truth values for $\mathbf{x}_1^A$ have been agreed upon between the two players. Then at the moment when $\mathbf{y}_2$ become determined, the two players choose a truth assignment to $\mathbf{y}_2$ so that it makes $\mathit{Parity}(\mathbf{y}_2)$ true or false depending on whether the formula appears in the antecedent or succedent, and such that the fraction of critical pseudo truth assignments to the remaining variables is as large as possible. If there are more than one such partial truth assignment, the players break tie by taking the first in some lexicographical order. In the set $F(\vec{u}, \vec{v})$ for the protocol defined below, the sequents are also chosen so that this fraction is large, i.e., at least $\frac{1}{st(\pi)^{\sigma(\mathcal{S})}}$ for a function $\sigma(\mathcal{S})$ that we will now define.

The two players can be viewed as constructing a path in the tree π that starts from its root $\mathcal{S}_0(\mathit{Parity})$ and goes toward some leaf. Every time they meet a binary inference they have to choose between the two top sequents. In general the total fraction of critical pseudo truth assignments may decrease, but the players will always try to construct paths with a heuristic to maximize this fraction. We use the following notion.

We say that an inference in π is *tall* if its principal formula contains an instance of *Parity*. Note that these formulas cannot be cut, so they must remain in the final sequent $\mathcal{S}_0(\mathit{Parity})$ (though some identical copies can be eliminated using contraction). The players will make consistent choice when they meet a binary tall inference, in the sense that the paths that they construct satisfy the following condition. Suppose that a binary tall inference with the same principal formula appears twice, then the path follows the same auxiliary formula in both case. For example, suppose that the path contains both $\mathcal{S}_3$ and $\mathcal{S}_6$ which are derived using the following binary tall inferences:

$$\frac{\mathcal{S}_1 \qquad \mathcal{S}_2}{\mathcal{S}_3} = \frac{\Gamma_1 \longrightarrow \Delta_1, A(\mathit{Parity}) \qquad \Gamma_1 \longrightarrow \Delta_1, B(\mathit{Parity})}{\Gamma_1 \longrightarrow \Delta_1, A(\mathit{Parity}) \wedge B(\mathit{Parity})}$$

$$\frac{\mathcal{S}_4 \qquad \mathcal{S}_5}{\mathcal{S}_6} = \frac{\Gamma_2 \longrightarrow \Delta_2, A(\mathit{Parity}) \qquad \Gamma_2 \longrightarrow \Delta_2, B(\mathit{Parity})}{\Gamma_2 \longrightarrow \Delta_2, A(\mathit{Parity}) \wedge B(\mathit{Parity})}$$

Then either both $\mathcal{S}_1$ and $\mathcal{S}_4$ are on the path, or both $\mathcal{S}_2$ and $\mathcal{S}_5$ are on the path.

Definition 2.11. *An $\mathcal{S}_0(\mathit{Parity})$–$\mathcal{S}$ path in π is said to be* consistent *if it satisfies the above condition.*

In other words, using the notion of a *fresh* inference defined below, the choices made at tall binary inferences on a consistent path can be completely determined by the choices made at these inferences.

Definition 2.12. *A tall binary inference with bottom sequent $\mathcal{S}$ in π is called* fresh *if the principal formula contains some undetermined variables and has not appeared on the $\mathcal{S}_0(\mathit{Parity})$–$\mathcal{S}$ path as the principal formula of the same rule.*

Essentially, the fresh inferences are the only places where the players need to communicate in order to compute the *Next* function.

Definition 2.13. *Let $\mathcal{S}$ be a sequent in π such that the $\mathcal{S}_0(\mathit{Parity})$–$\mathcal{S}$ path, named P, is consistent. Then $\sigma(\mathcal{S})$ is the total number of contiguous blocks of sequents on P that are not the bottom sequent of a fresh inference.*

Note that $\sigma(\mathcal{S}_0(\mathit{Parity})) = 0$, and $\sigma(\mathcal{S})$ increases as we go further from $\mathcal{S}_0(\mathit{Parity})$, but that $\sigma(\mathcal{S})$ is always at most the total number of distinct subformulas of $\mathcal{S}_0$.

Now we define the consistency condition $F(\vec{u}, \vec{v})$. The leaves in $F(\vec{u}, \vec{v})$ are those sequents $\mathcal{S} = \Gamma \longrightarrow \Delta$ such that for some determined variables $\mathbf{x}_i^A$ and $\mathbf{x}_i^B$ we have

$$u_i = \mathit{Parity}(\mathbf{x}_i^A) \neq v_i = \mathit{Parity}(\mathbf{x}_i^B)$$

To define the internal nodes of $F(\vec{u}, \vec{v})$ we use the following notation. For each sequent $\mathcal{S} = \Gamma \longrightarrow \Delta$ we let $\mathcal{S}' = \Gamma' \longrightarrow \Delta'$ be the sequent obtained from $\mathcal{S}$ by simplifying it using the pseudo partial truth assignments to the determined variables of $\mathcal{S}$ as specified above.

Now the internal nodes of $F(\vec{u}, \vec{v})$ are those sequents $\mathcal{S} = \Gamma \longrightarrow \Delta$ that are not the leaves and that satisfy the following conditions:

(i) The $\mathcal{S}_0(\mathit{Parity})$–$\mathcal{S}$ path in π is consistent.

(ii) $\mathcal{S}$ is involved (either as the bottom or as a top sequent) in a fresh (binary tall) inference.

(iii) The fraction of critical pseudo truth assignments to the variables in $\mathcal{S}'$ is at least $\frac{1}{st(\pi)^{\sigma(\mathcal{S})}}$.

(iv) Any critical pseudo truth assignment (to the variables in $\mathcal{S}'$) that gives the *Parity* formulas owned by Alice the values from $\vec{u}, \vec{q}$ and the *Parity* formulas owned by Bob the values from $\vec{v}, \vec{r}$ falsifies $\mathcal{S}'$. In other words, the following are True:

$$\bigwedge_{A(\vec{P},\vec{Q})\in\Gamma',\text{ owned by Alice}} A(\vec{u},\vec{q}) \;\wedge \bigwedge_{A(\vec{P},\vec{Q})\in\Delta',\text{ owned by Alice}} \neg A(\vec{u},\vec{q}) \quad (4)$$

$$\bigwedge_{B(\vec{P},\vec{R})\in\Gamma',\text{ owned by Bob}} B(\vec{v},\vec{r}) \;\wedge \bigwedge_{B(\vec{P},\vec{R})\in\Delta',\text{ owned by Bob}} \neg B(\vec{v},\vec{r}) \quad (5)$$

Lemma 2.14. *Any sequent $\mathcal{S}$ that satisfies (iii) and (iv) above must contain both $\mathit{Parity}(\mathbf{x}_i^A)$ and $\mathit{Parity}(\mathbf{x}_i^B)$ for some i.*

In particular, an internal node of $F(\vec{u}, \vec{v})$ is necessarily not an axiom. Later we will actually show that the *Next* function is well defined for such a sequent.

Proof of Lemma 2.14. Let k denote the total number of occurrences of all $\vec{p}, \vec{q}, \vec{r}$ variables in $\mathcal{S}_0$. Then $k \leq \mathit{size}(\mathcal{S}_0)$. Since $4st(\pi) \leq 2^{m^{1/(d+1)}/\mathit{size}(\mathcal{S}_0)}$ and $\sigma(\mathcal{S}) \leq \mathit{size}(\mathcal{S}_0)$, it can be verified that $1/st(\pi)^{\sigma(\mathcal{S})} \geq 4^k/2^{m^{1/(d+1)}}$. Lemma 2.3 implies that we can set the values of the *Parity* formulas to be $\vec{u}, \vec{v}, \vec{q}, \vec{r}$ appropriately. Here C is the conjunction of all depth-d formulas in the antecedent, and the negation of all depth-d formulas in the succedent of $\mathcal{S}$. So if no pair $\mathit{Parity}(\mathbf{x}_i^A)$ and $\mathit{Parity}(\mathbf{x}_i^B)$ appear simultaneously, then (iv) implies that the sequent $\mathcal{S}$ is not a tautology, contradicts the fact that it is in π. □

Now it is easy to verify that the endsequent $\mathcal{S}_0(\mathit{Parity})$ is an internal node of $F(\vec{u}, \vec{v})$. Also, deciding whether a sequent is a leaf of $F(\vec{u}, \vec{v})$ can be done by the two players without communication, because they both agree on some common values for the determined variables. In addition, conditions (i–iii) can be checked by each player separately, while (4) and (5) can be evaluated separately by the appropriate players, so to verify (iv) each player only needs one bit to transmit their result. As a result, membership in $F(\vec{u}, \vec{v})$ can be computed using two bits of communication.

Next, given an internal node $\mathcal{S}$ of $F(\vec{u}, \vec{v})$, we show how to compute the next sequent $\mathit{Next}(\mathcal{S})$, which must be a sequent in $F(\vec{u}, \vec{v})$. There are two cases.

Case I: $\mathcal{S}$ is the bottom sequent of a fresh inference.

Without loss of generality, suppose that the principal formula in $\mathcal{S}$ is $A(\mathit{Parity}) \wedge B(\mathit{Parity})$, so $\mathcal{S}$ is derived by a $\wedge$-right inference:

$$\frac{\mathcal{T}_1 \qquad \mathcal{T}_2}{\mathcal{S}} \quad = \quad \frac{\Gamma \longrightarrow \Delta, A(\mathit{Parity}) \qquad \Gamma \longrightarrow \Delta, B(\mathit{Parity})}{\Gamma \longrightarrow \Delta, A(\mathit{Parity}) \wedge B(\mathit{Parity})}$$

Moreover, on the $\mathcal{S}_0(\mathit{Parity})$–$\mathcal{S}$ path $A(\mathit{Parity}) \wedge B(\mathit{Parity})$ has not been used as the principal formula in any $\wedge$-right inference.

Note that the formula $A(\mathit{Parity}) \wedge B(\mathit{Parity})$ cannot be cut and will appear in the final sequent. Therefore it is owned by either Alice or Bob. The player who owns this formula knows the values of $A(\mathit{Parity})$ and $B(\mathit{Parity})$, and will choose the one that has the same value as $A(\mathit{Parity}) \wedge B(\mathit{Parity})$ and communicates one bit to indicate whether $\mathit{Next}(\mathcal{S})$ is $\mathcal{T}_1$ or $\mathcal{T}_2$ accordingly.

Without loss of generality, suppose that $A(\mathit{Parity}) \wedge B(\mathit{Parity})$ and $A(\mathit{Parity})$ are both false, and hence $\mathit{Next}(\mathcal{S}) = \mathcal{T}_1$. First consider the case where A is a variable and the variables in $A(\mathit{Parity})$ are not determined in $\mathcal{S}$. So the variables in $A(\mathit{Parity})$ become determined in $\mathcal{T}_1$. If $\mathcal{T}_1$ is not a leaf of $F(\vec{u}, \vec{v})$, then the choice of the players' common truth assignment to these variables guarantees that the fraction of critical pseudo truth assignments for $\mathcal{T}_1'$ is at least that of $\mathcal{S}'$. In this case it is easy to verify that both (4) and (5) remain true for $\mathit{Next}(\mathcal{S})$, so (iv) holds for $\mathit{Next}(\mathcal{S})$. In addition, conditions (i–iii) hold for $\mathit{Next}(\mathcal{S})$ because they hold for $\mathcal{S}$, and because of the fact that $\mathcal{S}$ and $\mathit{Next}(\mathcal{S})$ have the same set of critical pseudo truth assignments and that $\sigma(\mathcal{S}) = \sigma(\mathit{Next}(\mathcal{S}))$. Thus $\mathit{Next}(\mathcal{S})$ is an internal node of $F(\vec{u}, \vec{v})$.

The case where A is not a variable, or if A is a variable but the variables in $A(\mathit{Parity})$ are already determined, is straightforward, because there are no new determined variables.

Case II: $\mathcal{S}$ is not the bottom sequent of any fresh inference.

In this case $\mathcal{S}$ must be the top sequent of a fresh inference. Basically we want to start from $\mathcal{S}$ and follow a path in π until we reach (the bottom sequent of) a fresh inference. So consider applying the following operations repeatedly on the subtree of π rooted at $\mathcal{S}$:

1. Prune all branches that are not consistent.

2. Prune all subtrees that are rooted at a leaf of $F(\vec{u}, \vec{v})$.

3. If a sequent $\mathcal{T}$ is obtained by a binary rule whose principal formula is a subformula of a Parity formula and has depth greater than d, then keep only the subtree rooted at the top sequent whose auxiliary formula has the same value as the principal formula (if both of them have the same value, then keep only the left one).

4. Prune the subtrees rooted at those sequents $\mathcal{T}$ that are derived by a fresh (binary tall) inference (so that such a $\mathcal{T}$ becomes a leaf of T).

The players can perform these operations individually without communication. This is because, for example, for 3. the principal formula only involved determined variables, and the players agree on a common truth assignment to the determined variables.

Let T denote the resulting tree. The sequent $\mathit{Next}(\mathcal{S})$ is defined as follows. If T contains a leaf of $F(\vec{u}, \vec{v})$, then let $\mathit{Next}(\mathcal{S})$ be the lexicographically first such a sequent. Otherwise, $\mathit{Next}(\mathcal{S})$ is the leaf of T that has the most fraction of critical pseudo truth assignments. Note that $\mathit{Next}(\mathcal{S})$ can be computed by the players without communication. We now argue that it also belongs to $F(\vec{u}, \vec{v})$. We consider the case where T does not contain any leaf of $F(\vec{u}, \vec{v})$. We will show that in this case $\mathit{Next}(\mathcal{S})$ is an internal node of $F(\vec{u}, \vec{v})$.

Consider an internal node $\mathcal{T}$ of the tree T. Note that $\mathcal{T}$ has two parents just in case it is derived by a binary inference whose principal formula has depth at most d. In this case, it can be seen that the set of critical pseudo truth assignments for $\mathcal{T}$ is the union of that of its parents. On the other hand, in the case where $\mathcal{T}$ has a single parent, the fraction of critical pseudo truth assignment of its parent is always the same as that of $\mathcal{T}$. From this it follows that the fraction of critical pseudo truth assignments for $\mathit{Next}(\mathcal{S})$ is at least $\frac{1}{st(\pi)}$ that of $\mathcal{S}$, because the size of T is at most $st(\pi)$. Observe that $\sigma(\mathcal{T}) = \sigma(\mathcal{S}) + 1$. This shows that condition (iii) holds for $\mathit{Next}(\mathcal{S})$. The fact that conditions (i), (ii) and (iv) hold for $\mathit{Next}(\mathcal{S})$ can be easily verified.

This completes the proof of the first part of the theorem. For the statement about monotone circuit, suppose for example that $\vec{p}$ only positively in Γ_0. Then, whenever a set $\mathbf{x}_i^A$ becomes determined, it must appear in the antecedent. This means that the leaves of $F(\vec{u}, \vec{v})$ are always labelled with labels of the type $u_i = 1 \wedge v_i = 0$. So the resulting circuit is monotone.

3 Conclusion

In using the hardness of *Parity* (against depth-d $\mathbf{AC}^0$ circuits), one issue is the presence of formulas of depth greater than (the fixed constant) d. These are large subformulas of *Parity*. In [MNP11] this is handled by requiring that the family of sequents that we consider satisfy some special condition (call the Statman property there). This condition guarantees that, as we travel the proof from the root toward a leaf, any time a compound formula is broken down we can set them to a constant; so, effectively, large subformulas of *Parity* disappear as soon as they arise. In Section 2 we utilize the treelike structure of the proof to resolve this issue, i.e., we make the two players agree on a common partial truth assignment that kill all these formulas.

Another place in our argument in Section 2 where the treelikeness of the proofs plays a crucial role is Case II. If the proof is daglike we have to make an analysis similar to this case every time we encounter a binary inference introducing a formula of the form $A(\mathit{Parity}) \wedge B(\mathit{Parity})$ or $A(\mathit{Parity}) \vee B(\mathit{Parity})$. So we can only get an exponentially small (in the length of π) lower bound on the fraction of critical pseudo truth assignments, and this is not sufficient for applying Lemma 2.3.

Let us mention again the conjecture that underlies the project of [MP06]. The idea is that the minimum number of steps in a depth-d **PK** proof of a lifted sequent $\mathcal{S}(\mathit{Parity})$ ought to be the same as the minimum number of steps in a cut-free proof of $\mathcal{S}$ (here we allow axioms of the form $A \longrightarrow A$ for any formula A.) The hope is to prove this for depth-d **PK** and extend it to proof systems with modular counting gates; this would then provide superpolynomial lower bounds for these systems, although the hard sequent for, say d-**PK**[2], is of depth much greater than d. The conjecture is open even for treelike proofs. Here, and in [MNP11], it is shown that $\mathcal{S}(\mathit{Parity})$ requires same lower bound in d-$\mathbf{PK}^\star$ as that of $\mathcal{S}$ in cut-free $\mathbf{PK}^\star$, for several sequents $\mathcal{S}$. These results are not obtained by proving the conjecture for these sequents, but instead by adapting the known proofs of lower bounds for cut-free proofs to prove lower bounds for depth-d proofs. Proving the conjecture directly even for these sequents would be interesting.

Acknowledgements

I would like to thank Jan Krajíček and an anonymous referee for comments, and for pointing out a wrong claim in an earlier version of this paper.

References

[AB87] Noga Alon and Ravi Boppana. The monotone circuit complexity of boolean functions. *Combinatorica*, 7(1):1–22, 1987.

[AM11] Albert Atserias and Elitza Maneva. Mean-payoff games and propositional proofs. accepted for publication in Information and Computation, 2011.

[BPR97] Maria Luisa Bonet, Toniann Pitassi, and Ran Raz. Lower Bounds for Cutting Planes Proofs with Small Coefficients. *Journal of Symbolic Logic*, 62(3):708–728, 1997.

[BPR00] Maria Luisa Bonet, Toniann Pitassi, and Ran Raz. On Interpolation and Automatization for Frege Systems. *SIAM Journal on Computing*, 29(6):1939–1967, 2000.

[CN10] Stephen Cook and Phuong Nguyen. *Logical Foundations of Proof Complexity.* ASL Perspectives in Logic Series. Cambridge University Press, 2010.

[Hås86] Johan Håstad. *Computational limitations for small depth circuits.* PhD thesis, Massachusetts Institute of Technology, 1986.

[IPU94] Russell Impagliazzo, Toniann Pitassi, and Alasdair Urquhart. Upper and lower bounds for tree-like cutting planes proofs. In *In Proceedings IEEE Symposium on Logic in Computer Science*, pages 220–228, 1994.

[KP98] Jan Krajíček and Pave Pudlák. Some consequences of cryptographical conjectures for S_2^1 and EF. *Information and Computation*, 140(1):82–94, January 1998.

[Kra94] Jan Krajíček. Lower bounds to the size of constant-depth propositional proofs. *Journal of Symbolic Logic*, 59:73–86, 1994.

[Kra97] Jan Krajíček. Interpolation theorems, lower bounds for proof systems, and independence results for bounded arithmetic. *Journal of Symbolic Logic*, 62(2):457–486, 1997.

[Kra98] Jan Krajíček. Discretely ordered modules as a first-order extension of the cutting planes proof system. *Journal of Symbolic Logic*, 63(4):1582–1596, 1998.

[Kra10] Jan Krajíček. A form of feasible interpolation for constant depth frege systems. *Journal of Symbolic Logic*, 75(2):774–784, 2010.

[KW88] Mauricio Karchmer and Avi Wigderson. Monotone circuits for connectivity require super-logarithmic depth. In *Proceedings of the 20th Annual ACM Symposium on Theory of Computing*, pages 539–550, 1988.

[MNP11] Alexis Maciel, Phuong Nguyen, and Toniann Pitassi. Lifting lower bounds for tree-like proofs. submitted, 2011.

[MP06] Alexis Maciel and Toniann Pitassi. A conditional lower bound for a system of constant-depth proofs with modular connectives. In *Proc. 21st IEEE Symposium on Logic in Computer Science*, 2006.

[Ngu07] Phuong Nguyen. Separating DAG-Like and Tree-Like Proof Systems. In *Proc. 22nd IEEE Symposium on Logic in Computer Science*, pages 235–244, 2007.

[PS98] Pavel Pudlák and Jirí Sgall. Algebraic models of computations and interpolation for algebraic proof systems. In Paul Beame and Sam Buss, editors, *in Proof Complexity and Feasible Arithmetic, DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, volume 39, pages 279–295. American Mathematical Society, 1998.

[Pud97] Pavel Pudlák. Lower bounds for resolution and cutting planes proofs and monotone computations. *Journal of Symbolic Logic*, 62(3):981–998, 1997.

[Pud99] Pavel Pudlák. On the complexity of propositional calculus. In *in Sets and Proofs, Invited papers from Logic Colloquium 97*, pages 197–218. Cambridge University Press, 1999.

[Raz85] Alexander A. Razborov. Lower bounds for the monotone complexity of some Boolean functions. *Mathematics of the USSR*, 31:354–357, 1985.

[Raz95] Alexander A. Razborov. Unprovability of lower bounds on the circuit size in certain fragments of bounded arithmetic. *Izvestiya of the R. A. N.*, 59(1):201–224, 1995.