Short refutations for an equivalence-chain principle for constant-depth formulas

Sam Buss*
Department of Mathematics
University of California, San Diego
La Jolla, CA 92093-0112, USA
sbuss@ucsd.edu

Ramyaa Ramyaa†
Department of Computer Science
New Mexico Institute of Mining and Technology
Socorro, NM 87801
ramyaa@cs.nmt.edu

July 21, 2018

Abstract

We consider tautologies expressing equivalence-chain properties in the spirit of Thapen and Krajíček, which are candidates for exponentially separating depth k and depth $k+1$ Frege proof systems. We formulate a special case where the initial member of the equivalence chain is fully specified and the equivalence-chain implications are actually equivalences. This special case is shown to lead to polynomial size resolution refutations. Thus it cannot be used for separating depth k and depth $k+1$ propositional systems. We state some Håstad switching lemma conditions that restrict the possible propositional proofs in more general situations.

*Supported in part by NSF grant CCR-1213151. Buss also thanks the Simons Institute for a supporting a visit to the Special Program on *Logical Structures in Computation* in August and September 2016 where a main portion of this work was carried out.

†Ramyaa gratefully acknowledges the support from Simons Institute in Theoretical Computer Science and its Fall 2016 Program on *Logical Structures in Computation* for hosting her during the period of this work.

1 Introduction

Frege systems are the usual "textbook-style" proof systems for propositional logic with modus ponens as the only rule of inference. It is often more elegant to use the propositional sequent calculus PK instead, with connectives $\neg$, $\wedge$ and $\vee$. We can require that negations ($\neg$) be applied only to variables, and then define the depth of a formula as the number of alternating levels of $\wedge$'s and $\vee$'s in the formula. A depth k PK proof is a PK proof in which all formulas have at most k alternating blocks of $\wedge$'s and $\vee$'s. One of the central open questions in propositional proof complexity is the question of whether, for proofs of depth 2 (say) tautologies, the constant-depth Frege (or PK) systems form a hierarchy with respect to quasipolynomial size or other sub-exponential size simulations. The best results so far give a $n^{(\log n)^{\epsilon}}$ (superpolynomial) separation between depth $d-1$ and depth d PK proofs, where the constant $\epsilon > 0$ depends on d. This was first proved in a uniform setting by Impagliazzo and Krajíček [4] working with bounded arithmetic; it was generalized to the non-uniform setting by Krajíček [6] following a suggestion of Thapen.

There are two main approaches to the problem of giving better separations between depth $d-1$ and depth d PK proofs. The first approach considers the uniform version of this problem via the Paris-Wilkie translation by studying the $\forall\Sigma^b_1(\alpha)$-consequences of the fragments $T^k_2(\alpha)$ on bounded arithmetic (see [2, 5] for overviews). There has been extensive work in this direction, with the modern line of work initiated by the game induction principles of Skelley and Thapen [8], which characterized the total NP search problems of the fragments of the bounded arithmetic theories $T^k_2(\alpha)$. The second and less well-developed approach is the "isomorphism-chain" proposal of Krajíček [7].

The present paper considers a special case of equivalence-chain tautologies EC_D. These can also be construed as tautologies expressing a restricted case of the game induction principles in which the induction step is reversible and very simple and in which the first player has a unique strategy for the initial stage. The first versions of these tautologies were defined in [7], attributed to Thapen. The paper [7] was motivated by feasible interpolation; we are grateful to Pudlák [private communication] for bringing them to our attention in the framework of seeking separations of constant-depth proofs.

To define the EC_D tautologies, consider constant-depth Boolean sentences which are constructed by alternating levels of $\wedge$ and $\vee$ gates of fanin f, with the inputs all constants 1 (True) or 0 (False). Two such Boolean sentences φ and ψ are called "isomorphic" if one can be obtained from the other

by permuting the orders of inputs to gates. As a special case of this, we say that φ and ψ are "locally equivalent" if there is a gate g in φ and integers i and j such that interchanging the i-th and j-th inputs to g transforms φ into ψ. We will work with tautologies EC_D which state that if $\varphi_0, \varphi_1, \ldots, \varphi_T$ are depth D Boolean sentences with each φ_i locally equivalent to φ_{i+1}, then it is impossible for φ_0 to evaluate to true in a "minimal" way and for φ_T to evaluate to false. (See Section 2 for the precise definition.)

It is not hard to see that the tautologies EC_D have polynomial size, depth D PK proofs which are tree-like and log depth. Our main result is that the clausal forms of the complement $\overline{\mathrm{EC}}_D$ actually have polynomial size resolution refutations, even if D is allowed to vary. Thus the principles EC_D do not serve as candidates for a depth hierarchy for Frege systems or PK proof systems.

In Section 4, we describe alternative forms IC_D of the EC_D tautologies which we conjecture do provide a depth hierarchy for Frege systems and PK systems: these replace the conditions that φ_{i+1} is locally equivalent to φ_i with the conditions that φ_{i+1} is "locally implied" by φ_i. ("EC" and "IC" stand for "equivalence chain" and "'implication chain".) An alternate formulation called EC^-_D keeps the "locally equivalent" conditions but weakens the conditions on the first formula φ_0.

The intuition for why the principles EC^-_D and IC_D should be hard for depth D' PK proofs with D' sufficiently smaller than D is that the only possible proofs seem to require (implicitly) using truth definitions for depth D Boolean formulas to prove inductively that φ_i is true for all $i = 1, 2, \ldots, T$. But, Yao and Håstad showed that the truth of depth D Boolean formulas cannot be expressed by depth $D-1$ Boolean formulas of subexponential size [3]. The same intuition would seem to hold for the principles EC_D: but here the intuition fails, as there is an alternate way to indirectly express the truth of the formulas φ_i. For this, see the discussion at the beginning of Section 3.

The next section gives the main definitions. Section 3.1 describes the polynomial resolution refutations informally. Section 3.2 explains how the informal proof is translated into polynomial size resolution refutations. Section 3.3 outlines how the resolution refutations can be also obtained via formalization in the system $U_{2,1}$-IND of [1]. Section 4 uses the Håstad switching lemma to restrict what kinds of low depth proofs can be used to prove EC^-_D or IC_D. Ultimately, improved switching lemmas might be able to show that the EC^-_D or IC_D tautologies do yield a depth hierarchy for the Frege and PK proof systems.

We assume the reader has some basic knowledge of propositional logic

and resolution. Section 3.3 gives an alternate proof of Theorem 4 using formalization in bounded arithmetic, but this can be skipped by the reader unfamiliar with these systems.

2 Definitions of the Tautologies

This section formulates the EC tautologies as a unsatisfiable set of clauses. The EC tautologies are defined using three integer parameters: (1) $D > 1$ is the depth of the Boolean sentences φ_i (implicitly) described by the tautologies. (2) $f \geq 2$ is the fanin of each $\wedge/\vee$ gate in the Boolean sentences. (3) $T \geq 1$ gives the number of φ_i's in the chain of Boolean sentences.

We use standard conventions for concepts such as "clause", "resolution inference", etc. We use $[k]$ to denote $\{0, 1, \ldots, k-1\}$.

By convention, a depth D Boolean sentence φ is the same as a tree-like Boolean circuit with constants as inputs. The topmost (output) gate is a $\vee$ gate and is at depth 0. More generally, gates at depth $d < D$ are $\wedge$ gates if d is odd, and $\vee$ gates if d is even. The inputs are at depth D and are constants 0 or 1. Gates and inputs of φ are specified by sequences $\pi = \langle a_1, \ldots, a_k \rangle$ for $0 \leq k \leq D$ and $a_i \in [f]$. We call π a "path" and it specifies the node which is reached by starting at the root (the output gate) of φ, then walking to its child number a_1, then to that gate's child number a_2, etc., repeating this for k steps. Paths of even (resp., odd) length $< D$ specify an $\vee$ gate (resp., an $\wedge$ gate). Paths of length D specify input nodes.

Definition 1. A *path* is a sequence π of length $|\pi| \leq D$ of the form

$$\langle d_1, c_1, d_2, c_2, \ldots, c_{\ell-1}, d_\ell \rangle \qquad \text{or} \qquad \langle d_1, c_1, d_2, c_2, \ldots, d_\ell, c_\ell \rangle$$

where c_i and d_i are in $[f] = \{0, \ldots, f-1\}$. Thus $|\pi|$ equals $2\ell-1$ or 2ℓ (respectively).

The letters "d" and "c" stand for "disjunction" and "conjunction": a d_i selects an input to a disjunction; a c_i selects an input to a conjunction.

The tautologies $\mathrm{EC}_{D,f,T}$, or just EC_D for short, implicitly express properties about Boolean sentences $\varphi_0, \ldots, \varphi_T$. For $0 \leq t \leq T$, φ_t is the Boolean sentence at time t. These tautologies involve the following variables:

1. (*Leaf variables.*) For each $0 \leq t \leq T$, and each path π of length D, there is a variable x^t_π. The intuition is that Boolean value of x^t_π gives the true/false value of the input of φ_t indexed by π. There are $f^D \cdot (T+1)$ many leaf variables x^t_π.

2. (*Swap variables.*) For each path π with length $< D$, each $0 \leq t \leq T$, and each $0 \leq i < j < f$, there is a variable $y^t_{\pi,i,j}$. The intuition is that $y^t_{\pi,i,j}$ indicates that φ_{t+1} is obtained from φ_t by interchanging (swapping) the i-th and j-th inputs of the gate g_π indexed by π.

Definition 2. An *initial* $\vee$*-path* is any path π of length D of the form $\langle 0, c_1, 0, c_2, \ldots\rangle$, i.e., with each $d_i = 0$. Dually, a *terminal* $\wedge$*-path* is any path π of length D of the form $\langle d_1, 0, d_2, 0, \ldots\rangle$, i.e., with each $c_i = 0$.

The idea for initial $\vee$-paths is that they are the paths that can be obtained by always selecting the first input (indexed by $d_i = 0$) of $\vee$-gates. The same holds for terminal $\wedge$-paths, except using the first inputs of $\wedge$-gates. The initial $\vee$-paths are called "initial" since we will use them for the first Boolean sentence φ_0, that is, with $t = 0$. Likewise, the "terminal" $\wedge$-paths will be used for the last Boolean sentence φ_T.

We can now define the clauses that will be used formalize the propositional principles EC, EC$^-$, and IC. These clauses will be chosen from the following:

a. For each initial $\vee$-path π, the unit clause x^0_π.

b. For each terminal $\wedge$-path π, the unit clause $\neg x^T_\pi$.

c. For each π, $|\pi| = D$, not an initial $\vee$-path, the unit clause $\neg x^0_\pi$.

d. For each swap variable $y^t_{\pi,i,j}$ and each path σ with $|\pi| + 1 + |\sigma| = D$ (so $0 \leq |\sigma| < D$), the clauses[1]

   i. $y^t_{\pi,i,j} \wedge x^t_{\pi i\sigma} \to x^{t+1}_{\pi j\sigma}$
   ii. $y^t_{\pi,i,j} \wedge x^{t+1}_{\pi j\sigma} \to x^t_{\pi i\sigma}$
   iii. $y^t_{\pi,i,j} \wedge x^t_{\pi j\sigma} \to x^{t+1}_{\pi i\sigma}$
   iv. $y^t_{\pi,i,j} \wedge x^{t+1}_{\pi i\sigma} \to x^t_{\pi j\sigma}$

   and for each path τ with $|\tau| = D$ not of the above form $\pi i\sigma$ or $\pi j\sigma$, the clauses

   v. $y^t_{\pi,i,j} \wedge x^t_\tau \to x^{t+1}_\tau$
   vi. $y^t_{\pi,i,j} \wedge x^{t+1}_\tau \to x^t_\tau$

[1] These are written as implications but of course are equivalent to clauses. E.g., d.i. is the clause $\neg y^t_{\pi,i,j} \vee \neg x^t_{\pi i\sigma} \vee x^{t+1}_{\pi j\sigma}$.

e. For each fixed $t < T$, the clause $\bigvee y^t_{\pi,i,j}$. The disjunction is taken over all paths π of length $< D$ and all $i < j$; that is, over all swap variables at time t.

f. For each $t < T$, the clauses $\neg y^t_{\pi,i,j} \vee \neg y^t_{\pi',i',j'}$, for all pairs of distinct swap variables $y^t_{\pi,i,j}$ and $y^t_{\pi',i',j'}$.

The clauses of type a. suffice to prove that the Boolean sentence φ_0 evaluates to true. Indeed, the $\vee$-gates are all satisfied by their first inputs. Alternately, we can define the truth of φ using a two player D-round game played by an *existential* player and a *universal* player. The players select a path from the root gate-by-gate with the existential and universal players selecting the inputs of $\vee$- and $\wedge$-gates, respectively. Then φ_0 is true means the existential player has a winning strategy. The clauses of type a. ensures that the existential player can win by always choosing the first input to the current $\vee$-gate. Dually, the clauses of type b. ensure that the final formula φ_T is false.

The clauses of type c. state that the rest of the inputs to φ_0 are false; in other words, that φ_0 is true in a minimal way. These clauses are not necessary for the truth of the equivalence chain tautologies, but are crucial for our resolution refutations. One could dually define unit clauses x^T_π for π not a terminal $\wedge$-path, but it turns out we do not need them.

The clauses of type e. and f. state that exactly one swap variable is true at time t. The true swap variable $y^t_{\pi,i,j}$ controls how φ_{t+1} is formed as an isomorphic copy of φ_t. The clauses of type d. force the correct isomorphic copy. The clauses d.i-iv. state that the inputs that are "moved" by the swap correctly maintain their values, and the clauses of type d.v-vi. state that the rest of the inputs maintain their values.

Definition 3. Fix values for the depth $D > 1$, the fanin f, and the number of swaps T. The $\mathrm{EC}_{D,f,T}$ clauses are the clauses of types a.-f. The $\mathrm{EC}^-_{D,f,T}$ are the same clauses except omitting the clauses of type c. The $\mathrm{IC}_{D,f,T}$ clauses omit d.ii., d.iv., and d.vi.; that is, they are the clauses a., b., c., d.i., d.iii., d.v., e. and f.

Each of these sets of clauses is unsatisfiable. If we fix the value of D, we get families of sets of clauses, we denote EC_D, EC^-_D and IC_D. The clauses of type c. are superfluous for IC_D, and may as well be omitted. It is not clear that the clauses of type f. are ever useful; indeed, our resolution refutations do not use them. Nonetheless, it seems more elegant to include them.

3 Resolution refutations of EC_D clauses

We now state our main result. We measure the size of a resolution refutation in terms of the number of clauses appearing in the refutation.

Theorem 4. *Letting all three of D, f, and T vary, there are polynomial size resolution refutations of the set of clauses* $\mathrm{EC}_{D,f,T}$.

To be more specific about the size bounds, note that there are $O(f \cdot f^D \cdot T)$ many swap variables $y^t_{\pi,i,j}$; there are $O(f^2 f^{2D} \cdot T)$ many clauses of type f.; there are $O(f \cdot f^D \cdot T)$ many clauses of type d., and fewer clauses of the other types. Thus, Theorem 4 states that the sizes of the resolution refutations can be polynomially bounded in terms of f^D and T; this is the same as being polynomially bounded in the size of $\mathrm{EC}_{D,f,T}$.

The first intuition for how to prove Theorem 4 is to use truth definitions for the formulas φ_i, or more relevantly, use truth definitions for all subformulas of all the φ_i's. Normally, clauses cannot express the truth and falsity of formulas of depth D; however, the inclusion of clauses of type c., along with the presence of clauses d.ii., d.iv. and d.vi., makes this possible in our setting. Namely, for $i = 0$, any subformula ψ of φ_0 evaluates to true iff at least one input to ψ is true. More formally, for π a path of length $\leq D$, the subformula ψ of φ_0 rooted at the gate indexed by π is true iff the disjunction $\bigvee_\sigma x^0_{\pi\sigma}$ is true. The corresponding property can then be proved for each φ_i by proceeding inductively with $i = 1, 2, \ldots, T$.

This informal argument cannot be so simply carried out in resolution, nonetheless we use closely related techniques. We give an informal outline of the proof in Section 3.1, and then in Section 3.2 explain how this translates into polynomial size resolution refutations of the sets $\mathrm{EC}_{D,f,T}$. Section 3.3 discusses an alternate construction of the polynomial size resolution refutations, namely by formalizing the proof in the bounded arithmetic theory $U_{2,1}$-IND.

3.1 Informal version of the proof

The informal version of the proof is via Lemma 8 and Corollary 9 as discussed next. Section 3.2 will show how to formalize these as resolution derivations.

Definition 5. Fix $t \in [T]$. Fix a vector $\vec{c}$ of values $c_1, \ldots, c_\ell$, where $\ell = \lfloor D/2 \rfloor$. The clause $C^t_{\vec{c}}$ containing the literals x^t_π with the c_i components fixed to these values is called an $\vee$*-group at time* t:

$$C^t_{\vec{c}} \;:=\; \{x^t_{\langle d_1, c_1, d_2, c_2, \ldots\rangle} : \text{for all } i,\ d_i \in [f]\}. \tag{1}$$

Since the d_i's may equal zero, the clauses of type a. defined above specify that any $\vee$-group $C^0_{\vec{c}}$ at time $t = 0$ contains a variable which is set to true. In other words, any $\vee$-group $C^0_{\vec{c}}$, when viewed as a clause, must be true.

The next definitions express the local correctness of the informal definition of truth discussed above.

Definition 6. Fix $t \leq T$, and fix a path π of odd length $|\pi| = k < D$, so that π ends at an $\wedge$ gate. Let $i < f$ indicate an input to that gate. The (π, i)-*equivalence at time* t, denoted $E^t_{\pi,i}$, is the implication

$$\bigvee_{|\sigma|=|D|-k} x^t_{\pi\sigma} \;\rightarrow\; \bigvee_{|\tau|=|D|-k-1} x^t_{\pi i\tau}. \tag{2}$$

This is called an "equivalence" instead of an "implication" since the reverse implication is automatic.

Let also $j \in [f]$. The (π, i, j)-*equivalence at time* t, denoted $E^t_{\pi,i,j}$, is the equivalence

$$\bigvee_{|\tau|=D-k-1} x^t_{\pi i\tau} \;\equiv\; \bigvee_{|\tau|=D-k-1} x^t_{\pi j\tau}. \tag{3}$$

Definition 7. Fix t and π as in the previous definition. Let $k = 2k' + 1$, so $\pi = \langle d_1, c_1, \ldots, d_{k'}, c_{k'}, d_{k'+1}\rangle$. Also fix $\vec{c} = \langle c_{k'+1}, c_{k'+2}, \ldots, c_{k''}\rangle$ where $k' < k''$ and $2k'' \leq D$. Set $\ell = D - 2k''$. Letting $\vec{d}$ range over vectors of the form $\langle d_{k'+2}, d_{k'+3}, \ldots, d_{k''}\rangle$, and τ range over vectors with $|\tau| = \ell$, the $(\pi, \vec{c})$-*equivalence at time* t, denoted $E^t_{\pi,\vec{c}}$, is

$$\bigvee_{|\sigma|=D-k} x^t_{\pi\sigma} \;\rightarrow\; \bigvee_{\vec{d},\,\tau} x^t_{\pi c_{k'+1} d_{k'+2} \cdots c_{k''-1} d_{k''} c_{k''} \tau}. \tag{4}$$

Note that $E^t_{\pi,i}$ is a special case of $E^t_{\pi,\vec{c}}$ by taking $\vec{c}$ to be just $\langle i\rangle$ with $k'' = k' + 1$.

Lemma 8. *Fix π_0, i_0, j_0, and $\vec{c}$. The equivalences $E^t_{\pi_0,i_0,j_0}$ and $E^t_{\pi_0,\vec{c}}$ are consequences of the equivalences of the form $E^t_{\pi,i}$.*

Proof. By $E^t_{\pi_0,i_0}$ and $E^t_{\pi_0,j_0}$, the right-hand side and the left-hand side of the equivalence $E^t_{\pi_0,i_0,j_0}$ are both equivalent to $\bigvee_\sigma x^t_{\pi_0\sigma}$. Hence the equivalence $E^t_{\pi_0,i_0,j_0}$ holds.

We prove $E^t_{\pi_0,\vec{c}}$ by induction on the length $k'' - k'$ of $\vec{c}$. The base case where $\vec{c}$ has length 1 is the same as $E^t_{\pi,c_{k''}}$. For the induction step, the induction hypothesis states that

$$\bigvee_{|\sigma|=D-|\pi_0|} x^t_{\pi_0\sigma} \;\rightarrow\; \bigvee_{\vec{d}^-,\,\tau'} x^t_{\pi_0 c_{k'+1} d_{k'+2} \cdots d_{k''-1} c_{k''-1} \tau'}$$

where $\vec{d}^-$ means $d_{k'+2}, \ldots, d_{k''-1}$, and τ' ranges over vectors of length $D-2k''+2 > 0$. Rewriting τ' as $d_{k''}\tau''$. the induction hypothesis becomes

$$\bigvee_{|\sigma|=D-|\pi_0|} x^t_{\pi_0\sigma} \;\rightarrow\; \bigvee_{\vec{d},\,\tau''} x^t_{\pi_0 c_{k'+1} d_{k'+2} \cdots d_{k''-1} c_{k''-1} d_{k''} \tau''} \tag{5}$$

Now apply the $E^t_{\pi,i}$ equivalence (2) with π equal to the odd length sequence $\pi_0 c_{k'+1} d_{k'+2} \cdots d_{k''-1} c_{k''-1} d_{k''}$ and $i = c_{k''}$ to obtain

$$\bigvee_{\tau''} x^t_{\pi_0 c_{k'+1} d_{k'+2} \cdots d_{k''-1} c_{k''-1} d_{k''} \tau''} \;\rightarrow\; \bigvee_{\tau} x^t_{\pi_0 c_{k'+1} d_{k'+2} \cdots d_{k''-1} c_{k''-1} d_{k''} c_{k''} \tau}. \tag{6}$$

The implications (5) and (6) immediately imply $E^t_{\pi_0,\vec{c}}$. □

When we consider $|\pi_0|$ equal to one, Lemma 8 gives:

Corollary 9. *For fixed $\vec{c}$ of length $\lfloor D/2 \rfloor$, view $C^t_{\vec{c}}$ as a disjunction of variables. The following implication follows from the equivalences of the form $E^t_{\pi,i}$:*

$$\bigvee_{|\tau|=D} x^t_\tau \;\rightarrow\; C^t_{\vec{c}}. \tag{7}$$

Proof. Suppose D is odd (the argument when D is even is very similar). Let $k' = 0$ and $k'' = \lfloor D/2 \rfloor$. For $d_1 < f$, let $\pi = \langle d_1 \rangle$. Let σ range over sequences of length $D-1$, and let $\vec{d}$ range over sequences $\langle d_2, \ldots, d_{k''} \rangle$. Then $E^t_{\pi,\vec{c}}$ as defined in (4) yields

$$\bigvee_{\sigma} x^t_{d_1\sigma} \;\rightarrow\; \bigvee_{\vec{d},d_{k''+1}} x^t_{d_1 c_1 d_2 c_2 \ldots d_{k''} c_{k''} d_{k''+1}}.$$

Combining these f many implications for $d_1 < f$, gives immediately the implication (7) as desired. The argument for D even is identical, except that $d_{k''+1}$ is omitted and (4) is used with τ empty. □

We now give an informal semantic version of the arguments that will be used in the resolution refutations that satisfy Theorem 4. Section 3.2 shows discusses how these arguments can be realized by polynomial size resolution derivations.

Assume the $\mathrm{EC}_{D,f,T}$ clauses hold. We shall prove, for each successive $t = 0, 1, \ldots, T$, first $\bigvee_\pi x^t_\pi$ and second every (π, i)-equivalence $E^t_{\pi,i}$.

Time 0. We first derive $\bigvee_\pi x^0_\pi$. This is immediate from the initial clauses of type a.

We next derive every $E^0_{\pi,i}$ at time 0. Fix a (π,i)-equivalence $E^0_{\pi,i}$ as in (2). First, suppose $\pi = \langle d_1, c_1, d_2, c_2, \ldots, d_k\rangle$ with $d_j \neq 0$ for some $1 \leq j \leq k$. Then by initial clauses c., every $x^0_{\pi\sigma}$ is false; thus $E^0_{\pi,i}$ holds. Suppose conversely that each d_j in π is 0. Choosing τ to be of the form $\langle 0, c_{k+2}, 0, c_{k+3}, \ldots\rangle$ makes $\pi i\tau$ an initial $\vee$-path. Therefore $x^0_{\pi i\tau}$ is an initial clause of type a.; Thus $E^0_{\pi,i}$ again holds.

Time t+1. In this step, we derive $\bigvee_\pi x^{t+1}_\pi$ and the $E^{t+1}_{\pi,i}$'s. The arguments split into cases depending on which $y^t_{\pi_0,i_0,j_0}$ is true.[2] When $y^t_{\pi_0,i_0,j_0}$ is true, we define a bijection $h : \{x^t_\pi : |\pi| = D\} \to \{x^{t+1}_\pi : |\pi| = D\}$ by letting

$$h(x^t_{\pi_0 i_0 \sigma}) \;=\; x^{t+1}_{\pi_0 j_0 \sigma} \qquad \text{and} \qquad h(x^t_{\pi_0 j_0 \sigma}) \;=\; x^{t+1}_{\pi_0 i_0 \sigma}$$

for $|\sigma| = D - |\pi_0| - 1$, and letting $h(x^t_{\pi'}) = x^{t+1}_{\pi'}$ for all other $x^t_{\pi'}$. Overloading h, we also write $x^{t+1}_{h(\pi)}$ for $h(x^t_\pi)$. The mapping h extends in the natural way to clauses containing the literals x^t_π and more generally to propositional formulas. Furthermore, $h(\pi)$ can be uniquely defined for $|\pi| < D$ by letting $h(\pi) = \pi'$ provided $h(\pi\sigma) = \pi'\sigma'$ holds for some σ, σ'.

By initial clauses d.i-vi., we obtain $x^t_\pi \leftrightarrow x^{t+1}_{h(\pi)}$ for $|\pi| = D$. Therefore, the clause $\bigvee_\pi x^{t+1}_\pi$ follows immediately from $\bigvee_\pi x^t_\pi$.

The harder task is to establish $E^{t+1}_{\pi,i}$. We claim that this clause follows from the $E^t_{\pi,i}$'s and the initial axioms of types d.i.-d.iv. Indeed, if π_0 is an initial subsequence of π, then $E^{t+1}_{\pi,i}$ is identical to $h(E^t_{\pi',i'})$ for $\pi' i' = h(\pi i)$. Otherwise, $h(\pi) = \pi$ and thus $h(E^t_{\pi,i})$ is equivalent to $E^{t+1}_{\pi,i}$: the only difference is that the order of the literals in the disjunctions in (2) may be reordered. In other words, considering the lefthand side of (2), we have

$$\bigvee_{|\sigma| = D - |\pi|} x^t_{\pi\sigma} \;\leftrightarrow\; \bigvee_{|\sigma| = D - |\pi|} x^{t+1}_{h(\pi\sigma)}$$

by virtue of h being a bijection between the variables of the disjunctions. A similar consideration holds for the disjunction on the righthand side of (2).

Time T+1. The previous step established $\bigvee_\pi x^T_\pi$ and all clauses $E^T_{\pi,i}$. By Corollary 9, these imply $C^T_{\vec{0}}$, where $\vec{0}$ denotes the length $\lfloor D/2 \rfloor$ vector of zeros. The clause $C^T_{\vec{0}}$ is the disjunction of exactly the variables x^T_π with π a terminal $\wedge$ path. By initial clauses b., each such variable is false. This yields a contradiction, and completes the construction of the refutation of clauses a.-e.

[2]The argument here uses clauses of type e. to know that at least one of the swap variables $y^t_{\pi_0,i_0,j_0}$ is true. It does not need uniqueness, and the clauses of type f. are not used.

3.2 The resolution refutations

We now discuss how the above argument for Theorem 4 can be formulated as resolution refutations of the clauses $\mathrm{EC}_{D,f,T}$. The difficulty is that the above argument for Theorem 4, including the proofs of Lemma 8 and Corollary 9, does not work with clauses; rather it works instead with implications between clauses (e.g., the "equivalence" $E^t_{\pi,i}$) and with equivalences between clauses (e.g., $E^t_{\pi,i,j}$). To address this, we describe how to translate such implications and equivalences into sets of clauses.

We write $\overline{x}$ for $\neg x$; as usual, if y is $\overline{x}$, then $\overline{y}$ is x. We identify a clause with the disjunction of its members. By convention, clauses must be non-tautologous; i.e., a clause may not contain both of the literals z and $\overline{z}$.

Definition 10. Suppose F and G are clauses. $(F \to G)^{\mathrm{res}}$ is the set of clauses of the form $\overline{z} \vee G$ for $z \in F \setminus G$. And, $(F \equiv G)^{\mathrm{res}}$ is the set of clauses $(F \to G)^{\mathrm{res}} \cup (G \to F)^{\mathrm{res}}$.

Clearly, the set of clauses $(F \to G)^{\mathrm{res}}$ is equivalent to the formula $F \to G$, and similarly for $F \equiv G$. This allows a resolution refutation to work with, for instance, the sets of clauses $(E^t_{\pi,i})^{\mathrm{res}}$ and $(E^t_{\pi,i,j})^{\mathrm{res}}$ instead of with the formulas $E^t_{\pi,i}$ and $E^t_{\pi,i,j}$. The number of clauses in $(F \to G)^{\mathrm{res}}$ and $(F \equiv G)^{\mathrm{res}}$ is linearly bounded by the number of literals in F and G. The total number of (occurrences of) literals in $(F \to G)^{\mathrm{res}}$ and $(F \equiv G)^{\mathrm{res}}$ is quadratically bounded by the number of literals in F and G.

Lemma 11. *Let F and G be clauses. There is a resolution derivation of the clause G from the clause F and the clauses of $(F \to G)^{\mathrm{res}}$; the size of this derivation is polynomially bounded by the number of literals in F and G.*

Proof. This is essentially trivial: The derivation resolves F against the clauses $\overline{z} \vee G$ for $z \in F \setminus G$. □

Lemma 12. *Let F, G and H be clauses, and suppose the clauses of $(F \to G)^{\mathrm{res}}$ and $(G \to H)^{\mathrm{res}}$ are given as initial clauses. There is a polynomial size resolution derivation which derives all the clauses of $(F \to H)^{\mathrm{res}}$ from these initial clauses.*

By "polynomial size" is meant polynomial in the number of literals in F, G, H.

Proof. Consider any literal $z \in F \setminus H$: we need to derive $\overline{z} \vee H$. If $z \in G$, then the clause $\overline{z} \vee H$ is in $(G \to H)^{\mathrm{res}}$. Otherwise, this clause can be derived by resolving $\overline{z} \vee G$ against $\overline{u} \vee H$ for every $u \in G \setminus H$. The clauses $\overline{z} \vee G$ and $\overline{u} \vee H$ are in $(F \to G)^{\mathrm{res}}$ and $(G \to H)^{\mathrm{res}}$, respectively. □

For the next lemma, it is convenient to allow the weakening rule, namely from F infer any clause $F' \supset F$. A clause is identified with the set of literals in the clause.

Lemma 13. *Suppose F_i, G_i and H_i are clauses for $i < n$. The unions $F = \cup_i F_i$, $G = \cup_i G_i$ and $H = \cup_i H_i$ are clauses. Let the clauses in $(F_i \to G_i)^{\text{res}}$ and $(G_i \to H_i)^{\text{res}}$ for all $i < n$ be given as initial clauses. There there is a polynomial size resolution with weakening derivation of all the clauses in $(F \to H)^{\text{res}}$ from these initial clauses.*

Lemma 13 handles proofs "by cases". Informally: to prove F implies H, we argue that if F holds, then some F_i holds; from this G_i holds and therefore H_i and hence H holds.

Proof. This follows from the previous lemma. Let z be in $F \setminus H$. Choose i such that $z \in F_i \setminus H_i$. Lemma 12 gives a derivation of $\overline{z} \vee H_i$. From this, weakening gives $\overline{z} \vee H$. □

With Lemmas 11-13, it is easy to check (and we leave most of the details to the reader) that the arguments for the proof of Theorem 4 can be carried out in resolution augmented with weakening. First, using Lemmas 11-13 and adapting the proof of Lemma 8 shows that there are polynomial size derivations of the clauses of $(E^t_{\pi_0,i_0,j_0})^{\text{res}}$ and of $(E^t_{\pi_0,\vec{c}})^{\text{res}}$ from the clauses in the sets $(E^t_{\pi,i})^{\text{res}}$. Second, the proof of Corollary 9 shows there are polynomial size derivations of the clauses of $(7)^{\text{res}}$ from the clauses in the sets $(E^t_{\pi,i})^{\text{res}}$. Third, the informal arguments for Theorem 4 of Section 3.1 can be translated into a resolution with weakening derivation. For time $t = 0$, the clauses $(\bigvee_\pi x^0_\pi)$ and the clauses in $(E^0_{\pi,})^{\text{res}}$ follow from the initial clauses of type a. by weakening. For time t+1, the resolution derivation splits into cases, one for each $y^t_{\pi_0,i_0,j_0}$, as in the proof of Theorem 4. For time T+1, the resolution derivation uses the derivation obtained from Corollary 9 to derive the clause $C^T_{\vec{0}}$. Resolving with the initial clauses of type b. gives the empty clause.

Since weakenings can be removed from resolution refutations without increasing the size of the refutation, this proves Theorem 4.

3.3 Formalization in $U_{2,1}$-IND

An alternate way to describe the polynomial size resolution refutations is by formulating the arguments of Section 3.1 within the bounded arithmetic theory $U_{2,1}$-IND of Beckmann-Pudlák-Thapen [1]. For space reasons, we

omit the definitions of $U_{2,1}$-formulas and the theory $U_{2,1}$-IND. However, one can think of $U_{2,1}$-formulas as being $\forall^{\leq}\exists^{\leq}$-formulas; namely, as formulas with bounded universal quantifiers, followed by bounded existential quantifiers, followed by a quantifier-free part. For formalization in $U_{2,1}$-IND, the values D, f and T are free first-order variables. The value D is sharply bounded since the length of $\mathrm{EC}_{D,f,T}$ is polynomially bounded by f^D and T. We write $X(\pi, t)$ for x^t_π where $|\pi| = D$ is a sequence from $[f]$. A high-level sketch of the $U_{2,1}$-IND proof is that it proves the following three things by induction:

(a) The translation of $\bigvee_\pi x^t_\pi$ into first-order logic becomes the existential formula $A(t) := (\exists\pi)X(\pi, t)$, where it is required that π ranges over sequences of length D, and thus $(\exists\pi)$ is a bounded quantifier. The assertion $A(t)$ is proved by induction on t, for all $t = 0, \ldots, T$.

(b) The translations of the clauses $E^t_{\pi,i}$ are also proved by induction. These are expressed with

$$B(t) := (\forall\pi)(\forall i)(\forall\sigma)(\exists\tau)[\neg X(\pi\sigma, t) \vee X(\pi i\tau, t)],$$

with the domains of the variables π, i, σ, τ, σ constrained appropriately. The theory $U_{2,1}$-IND can express this as a $U_{2,1}$-formula. $B(t)$ is also proved by induction on t, for all $t = 0, \ldots, T$.

(c) Similarly, as an implication between two clauses, each $E^t_{\pi,\vec{c}}$ implication (4) can be expressed as a $U_{2,1}$-formula. These are needed only for $t = T$, and are proved by induction on the length $k'' - k'$. (Only a logarithmic length induction is needed here since D is sharply bounded and $k'' < D$.) The end result contradicts the translation of the axioms of type b.

Theorem A.2 of [1] on translations of $U_{2,1}$-IND proofs into propositional proofs then gives the existence of the polynomial size refutations needed for Theorem 4.

4 Prospects for proofs of IC_D

Theorem 4 showed that the clauses expressing (the negation of) the EC tautologies have polynomial size resolution refutations. It remains an important open question whether the IC_D or EC^-_D tautologies could have polynomial-size constant depth PK or Frege proofs. It is particularly frustrating that this question remains open since Håstad proved that the truth of depth D

Boolean sentences cannot be defined by polynomial size, depth $D-1$ formulas.

Specifically, consider the following framework for a hypothetical proof of an $\mathrm{IC}_{D,f,T}$ tautology. (The same discussion also applies to $\mathrm{EC}^{-}_{D,f,T}$.) The proof starts with the axioms of types a. and c. which assert φ_0 is true in a very strong sense, namely any true $\vee$ gate has exactly its first input true. From this it proves an assertion $A(\vec{x}^0)$ about the variables x^0_π. The formula A has polynomial size and has constant depth $D' << D$. Using $A(\vec{x}^t)$ and axioms of types d.i., d.iii., d.v., e. and f., it proves the formula $A(\vec{x}^{t+1})$ for successive values of $t = 0, 1, 2, \ldots, T-1$. Finally, it obtains a contradiction from $A(\vec{x}^T)$ and the axioms of type b. This is in fact the framework used by our proof of Theorem 4: in that proof, the formula $A(\vec{x}^t)$ consisted of the clauses $\bigvee_\pi x^t_\pi$ and $E^t_{\pi,i}$.

We claim that $\mathrm{IC}_{D,f,T}$ proofs which use this framework must have $A(\vec{x}^t)$ actually *equivalent* to the truth of the Boolean sentence φ_t. To prove this, consider the following facts:

(1) Any true formula can be modified by swapping the order of inputs to its gates so as to satisfy the initial clauses of type a. To do this, for each $\vee$-gate which evaluates to true, swap one of its true inputs to be the first input to the gate.

(2) Dually, any false formula can be modified by swapping the order of inputs to its gates so as to satisfy the initial clauses of type b. To do this, for each $\wedge$-gate which evaluates to false, swap one of its false inputs to be the first input to the gate.

(3) $A(\vec{x})$ is true if the variables $\vec{x}$ satisfy the axioms of type a.

(4) $A(\vec{x})$ is false if the variables $\vec{x}$ satisfy the axioms of type b.

(5) The property $A(\cdot)$ is preserved under swapping the order of inputs to gates.

Properties (3), (4) and (5) are needed for the proof framework.

It follows from (1), (3) and (5) that $A(\vec{x})$ must be true if the variables $\vec{x}$ define the input values for a depth D, fanin f Boolean sentence of alternating $\wedge/\vee$ gates which evaluates to true. Likewise, (2), (4) and (5) imply that $A(\vec{x})$ must be false if the values of $\vec{x}$ encode such a Boolean sentence which evaluates to false.

From Håstad [3], any such polynomial size $A(\vec{x})$ must have depth at least D.

It is tempting to try to argue that any depth $D' << D$ polynomial size proof of the IC_D or EC^-_D formulas must involve formulas $A(\vec{x})$ defining the truth of the depth D Boolean sentences, and hence must involve near-exponential size formulas (contradicting the proof being polynomial size). It is possible that such an argument can be carried out using a switching lemma, but attempts so far have been unsuccessful.

Acknowledgements. We thank the two anonymous referees for useful comments, including the suggestion to formalize the proof in $U_{2,1}$-IND. We also thank Jan Krajíček for helpful comments.

References

[1] A. Beckmann, P. Pudlák, and N. Thapen, *Parity games and propositional proofs*, ACM Transactions on Computational Logic, 15 (2014), pp. 17:1–30.

[2] S. R. Buss, *Bounded Arithmetic*, Bibliopolis, Naples, Italy, 1986. Revision of 1985 Princeton University Ph.D. thesis.

[3] J. Håstad, *Almost Optimal Lower Bounds for Small Depth Circuits*, vol. 5 of Advances in Computing Research, JAI Press, 1989, pp. 143–170.

[4] R. Impagliazzo and J. Krajíček, *A note on conservativity relations among bounded arithmetic theories*, Mathematical Logic Quarterly, 48 (2002), pp. 375–377.

[5] J. Krajíček, *Bounded Arithmetic, Propositional Calculus and Complexity Theory*, Cambridge University Press, Heidelberg, 1995.

[6] ——, *Proof Complexity*, Cambridge University Press, 20?? to appear.

[7] ——, *A form of feasible interpolation for constant depth Frege systems*, Journal of Symbolic Logic, 72 (2010), pp. 774–784.

[8] A. Skelley and N. Thapen, *The provably total search problems of bounded arithmetic*, Proceedings of the London Mathematical Society, 103 (2011), pp. 106–138.