Separation results for the size of constant-depth propositional proofs

Arnold Beckmann* Samuel R. Buss†

December 17, 2004

Abstract

This paper proves exponential separations between depth d-LK and depth $(d+\frac{1}{2})$-LK for every $d \in \frac{1}{2}\mathbb{N}$ utilizing the order induction principle. As a consequence, we obtain an exponential separation between depth d-LK and depth $(d+1)$-LK for $d \in \mathbb{N}$. We investigate the relationship between the sequence-size, tree-size and height of depth d-LK-derivations for $d \in \frac{1}{2}\mathbb{N}$, and describe transformations between them.

We define a general method to lift principles requiring exponential tree-size $(d+\frac{1}{2})$-LK-refutations for $d \in \mathbb{N}$ to principles requiring exponential sequence-size d-LK-refutations, which will be described for the Ramsey principle and $d = 0$. From this we also deduce width lower bounds for resolution refutations of the Ramsey principle.

Constant-depth Frege systems are propositional proof systems in which formulas have a bounded alternation of unbounded fanin conjunctions and disjunctions. These proof systems have been extensively studied because of their connection with the complexity of constant-depth circuits and fragments of bounded arithmetic (c.f. [19, 1, 3, 11, 15, 16, 20, 4]). Some of the strongest lower bounds on proof complexity apply to constant-depth Frege proofs and these lower bounds are intimately related to lower bounds on circuit complexity. The Paris-Wilkie translation [19] gives a direct connection between fragments of bounded arithmetic and constant-depth

*Partially supported by a Marie Curie Individual Fellowship #HPMF-CT-2000-00803 from the European Commission, and by FWF-grants #P16264-N05 and #P16539-N04 of the Austrian Science Fund. Institute of Algebra and Computational Mathematics, Vienna University of Technology, A-1040 Vienna, Austria. Email: beckmann@logic.at

†Supported in part by NSF grant DMS-0100589. Dept. of Mathematics, Univ. of California, San Diego, La Jolla, CA 92093-0112. Email: sbuss@ucsd.edu.

Frege systems that permits proof complexity lower bounds to translate to independence results for bounded arithmetic, and vice-versa.

Krajíček [11] defined an alternative notion of constant-depth that is particularly well suited to the Paris-Wilkie translation for the theories S^1_2 and T^i_2 of bounded arithmetic: A formula is defined to have Σ-depth d iff if is depth $d+1$ and the bottommost level of connectives have fanin $\leq \log S$, where S is a size parameter. A proof is defined to have Σ-depth d provided every formula in the proof has Σ-depth d.

It is common to unify the notions of depth and Σ-depth by defining Σ-depth d to be the same as depth $d + 0.5$ (c.f. [18]). Definition 1 below does this by defining a notion of Θ-depth: the Θ-depth d formulas will be defined so as to be the same as depth-d formulas, whereas Θ-depth $d + \frac{1}{2}$ will be the same as Σ-depth d. This will define the notions of "Θ-depth d" for all half-integers d, i.e., all $d \in \frac{1}{2}\mathbb{N} = \{0, \frac{1}{2}, 1, 1\frac{1}{2}, \ldots\}$. Since "$\Theta$-depth d" and "depth d" are synonymous for $d \in \mathbb{N}$, we generally drop the modifier "Θ-" and just speak of "depth d" for $d \in \frac{1}{2}\mathbb{N}$.

$$
\begin{array}{ccccccccc}
\text{depth } 0 & \subset & \Sigma\text{-depth } 0 & \subset & \text{depth } 1 & \subset & \Sigma\text{-depth } 1 & \subset & \cdots \\
\| & & \| & & \| & & \| & & \\
\Theta\text{-depth } 0 & \subset & \Theta\text{-depth } \frac{1}{2} & \subset & \Theta\text{-depth } 1 & \subset & \Theta\text{-depth } 1\frac{1}{2} & \subset & \cdots
\end{array}
$$

We also consider propositional proofs of (Θ-)depth d for all $d \in \frac{1}{2}\mathbb{N}$. There are many ways to measure the complexity of these proofs. First, the complexity of proofs can be measured by the number of symbols or the number of lines appearing in the proof. We will define the "size" of a proof to be the number of symbols in the proof, and the "cedent-size" to be the number of lines or cedents in the proof. Furthermore, proofs can be either sequence-like or tree-like. This paper studies in particular the relationships between (quasi-)polynomial size sequence-like and tree-like depth d proofs, and depth d proofs of (poly-)logarithmic height, for $d \in \frac{1}{2}\mathbb{N}$. ("Height" refers to the length of the longest branch in the proof.) It is well-known that quasipolynomial size, constant depth proofs arise when using the Paris-Wilkie translation of bounded arithmetic to propositional logic, see [12]. Poly-logarithmic height restricted proofs also occur naturally in this way, c.f. [5, 6].

For sequence-like proofs, Krajíček [11] gives a separation of Σ-depth $(d-1)$ proofs from Σ-depth d proofs, for $d \in \mathbb{N}$, $d > 0$, using a modification of the pigeonhole principle formulated as a set of clauses of formulas of depth d. But in terms of depth d proofs this only gives us a separation of depth $(d-1)$ proofs from depth $(d+1)$ proofs. Now it is possible to improve this argument of Krajíček's to give a separation of Σ-depth $(d-1)$ proofs from depth d proofs, or in our preferred terminology, a separation of depth $d-\frac{1}{2}$ proofs from depth d proofs, for $2 \leq d \in \mathbb{N}$. Here is a sketch of the proof: It only requires strengthening the upper bound on the size of proofs of the "Sipserised" weak pigeonhole principles. The (negations of the) weak pigeonhole principles used by Krajíček have quasipolynomial sequence-size refutations of depth $\frac{1}{2}$. Substituting the Sipser functions $f^{d,n}_{i,j}$, also called $\mathrm{S}_{d,n}(p_{i,j})$ in the present paper's notation, gives a refutation of the "Sipserized" weak pigeonhole principle, which has quasipolynomial sequence-size and in which every formula is either depth $\leq d$ or has depth $d+1$ with topmost connective having fanin $\log n^{O(1)}$. Applying the distributive law to the connectives at the *top* of the latter formulas, transforms them into depth d formulas, still of quasipolynomial size. These transformed cedents can be patched together to form a valid refutation of depth d.

In this paper, we obtain an exponential separation of Θ-depth d proofs from Θ-depth $(d+\frac{1}{2})$ proofs for all $d \in \frac{1}{2}\mathbb{N}$. This improves the results in [11] in several ways: First, our separation gives exponential separations between depth d and Σ-depth d proofs, and between Σ-depth d and depth $(d+1)$ proofs. In particular, this yields an exponential separation between depth d and depth $(d+1)$ proofs. Second, our exponential separation is stronger than that of [11]; instead of an separation of $\exp(\exp(\Omega(\log^{1/2} n)))$, we obtain a separation of $\exp(n^{\Omega(1)})$.

1 The Proof Systems

The propositional proof system LK is a variant of constant-depth Frege proofs. Since we work in classical logic, not intuitionistic logic, we can w.l.o.g. formalize LK as a Tait-style calculus. Formulas of LK are formed from propositional variables $p_0, p_1, p_2, \ldots$, and negation, $\neg$, and unbounded fanin conjunctions and disjunctions, $\bigwedge$ and $\bigvee$.

Formulas are defined inductively as follows: atoms p_i and negated atoms $\neg p_i$ are formulas called *literals*. If Φ is a non-empty finite set of formulas, then $\bigwedge\Phi$ and $\bigvee\Phi$ are formulas. For φ a formula, $\neg\varphi$ is an abbreviation of the formula formed from φ by interchanging $\bigwedge$ and $\bigvee$, and interchanging atoms and their negations. The *depth*, $\mathrm{dp}(\varphi)$, of a formula φ is the maximal nesting of $\bigwedge$ and $\bigvee$ in φ. Thus, literals have depth 0, and $\mathrm{dp}(\bigwedge\Phi) = \mathrm{dp}(\bigvee\Phi) = 1 + \max_{\varphi\in\Phi} \mathrm{dp}(\varphi)$.

Each line in an LK-proof is a finite set of formulas, called a *cedent*.

We use capital Greek letters Γ, $\Delta, \ldots$ as names for cedents. The intended meaning of a cedent Γ is $\bigvee\Gamma$. Cedents are sometimes also called *clauses* (in the case of refutations). We often abuse notation by writing Γ, φ or $\Gamma \vee \varphi$ instead of $\Gamma \cup \{\varphi\}$, or by writing $\varphi_1, \ldots, \varphi_k$ instead of $\{\varphi_1, \ldots, \varphi_k\}$.

Let $\mathcal{A}$ be an additional set of axioms, that is, $\mathcal{A}$ is a set of cedents. The intended meaning of $\mathcal{A}$ is $\bigwedge\{\bigvee\Gamma\,;\,\Gamma \in \mathcal{A}\}$. An LK proof from the hypotheses $\mathcal{A}$ has the following axioms and inference rules:

$$\text{Logical Axiom: } \frac{}{\varphi, \neg\varphi}$$

$$\text{Non-Logical Axiom: } \frac{}{\Gamma} \text{ and } \frac{}{\bigvee\Gamma}\text{, for } \Gamma \in \mathcal{A}$$

$$\bigvee\ \frac{\Gamma, \varphi}{\Gamma, \bigvee\Phi}\text{, where } \varphi \in \Phi \qquad\qquad \bigwedge\ \frac{\Gamma, \varphi \qquad \text{for all } \varphi \in \Phi}{\Gamma, \bigwedge\Phi}$$

$$\text{Weakening: } \frac{\Gamma}{\Gamma'}\text{, provided } \Gamma' \supseteq \Gamma \qquad\qquad \text{Cut: } \frac{\Gamma, \neg\varphi \qquad \Gamma, \varphi}{\Gamma}$$

The formulas $\bigvee\Phi$ and $\bigwedge\Phi$ in the lower cedents of the $\bigvee$ and $\bigwedge$ rules are called the *principal* formulas of those inferences. Note that we allow both Γ and $\bigvee\Gamma$ as non-logical axioms; this is reasonable enough since they have the same meaning. In addition, this convention simplifies the statements of our theorems and avoids some technical problems.

An LK-*derivation from* $\mathcal{A}$ is a tree; each node in the tree is labeled with a cedent. Cedents at leaf nodes must be logical or non-logical axioms, and cedents at internal nodes are inferred by one of the rules of inference from the cedents on the child nodes. We picture the tree with the root at the bottom; if the root is labeled with the cedent Γ, the derivation is called a derivation of Γ from $\mathcal{A}$. If the last cedent is the empty cedent, then the derivation is called an LK-*refutation of* $\mathcal{A}$.

The complexity of derivations can be measured in five different ways: the main distinction between the measures is by counting number of cedents vs. number of symbols. We have three ways of counting cedents of a derivation: by their *sequence-cedent-size* (number of distinct cedents), their *tree-cedent-size* (number of occurrences of cedents, i.e., no cedent is used as a hypothesis more than once), and by the *height* of the derivation tree. Height is defined to equal the maximum number of cedents along any path in the refutation. We define the *sequence-size* of a derivation to be the number of symbols of the distinct cedents, and the *tree-size* to be the number of symbols in all occurrences of cedents.

The constant depth LK proof systems will be defined by restricting the depth of formulas appearing in refutations. Fix a size parameter S. It is useful to treat $\bigwedge_{i<\log S}\varphi_i$ and $\bigvee_{i<\log S}\varphi_i$ for literals φ_i as being depth 1/2. This motivates the following definition:

Definition 1. *Let S, d be in $\mathbb{N}$. The classes Θ_d^S and $\Theta_{d+0.5}^S$ are inductively defined by the following:*

1. *$\varphi \in \Theta_0^S$ iff φ is a literal.*
2. *$\varphi \in \Theta_{0.5}^S$ iff φ is a literal or a $\bigwedge$ or $\bigvee$ of at most $\log S$ many literals.*
3. *$\varphi \in \Theta_{d+1}^S$ iff φ is in Θ_d^S, or it has the form of a $\bigvee$ or $\bigwedge$ of at most S many formulas from Θ_d^S.*

For $d \in \mathbb{N}$, it is easy to show, by induction on d, that a Θ_d^S-formula has size $O(S^d)$, and a $\Theta_{d+\frac{1}{2}}^S$-formula has size $O(S^d \log S)$. In particular, for all $d \in \frac{1}{2}\mathbb{N}$ we have that a Θ_d^S-formula has size $O(S^d)$.

Now, let $d \in \frac{1}{2}\mathbb{N}$. An LK-derivation is called a Θ_d^S-LK-derivation if all formulas in the derivation are in Θ_d^S. Often we are interested in formulas whose sizes grow polynomially or quasi-polynomially. We define Θ_d^{poly} (resp., Θ_d^{qp}) to be the set of all sequences of formulas $(\varphi_n)_n$ such that there is some $c \in \mathbb{N}$ with $\varphi_n \in \Theta_d^{n^c}$ (resp., $\varphi_n \in \Theta_d^{2^{\log^c n}}$) for all but finitely many n.

Given a bound on formula size, we can bound the size of a refutation in terms of its cedent-size: A Θ_d^S-LK-refutation of (sequence- or tree-) cedent-size S has (sequence- or tree-, resp.) size $S^{d+O(1)}$. This is because at most S distinct formulas can appear in the proof, so each cedent in such a derivation contains at most S Θ_d^S-formulas, each of size $O(S^d)$.

Krajíček [11] defined the Σ-depth of a derivation P of size S. In our terms, this can be defined as the minimal $d \in \mathbb{N}$ such that every formula in P is in $\Theta_{d+\frac{1}{2}}^S$. Let us define the Θ-depth of a derivation P of size S to be the minimal $d \in \frac{1}{2}\mathbb{N}$ such that every formula in P is in Θ_d^S. Thus for $d \in \mathbb{N}$, Σ-depth d equals Θ-depth $(d+\frac{1}{2})$. An LK-derivation is called a Θ-d-LK-derivation if its Θ-depth is bounded by d. When there is no chance of confusion, we drop the "Θ-" and use "depth" or "d-LK" instead of "Θ-depth" or "Θ-d-LK".

Usually we are interested in asymptotic growth rates of sizes of refutations of sequences of sets of clauses. For this, let $\{\mathcal{A}_n\}_n$ be a family of sets of cedents and let $S\colon \mathbb{N} \to \mathbb{N}$. $\{\mathcal{A}_n\}_n$ is said to be in $\Theta_d^{S(n)}$ provided each cedent in the set $\mathcal{A}_n$ contains only $\Theta_d^{S(n)}$-formulas. Further, $\{\mathcal{A}_n\}_n$ is in Θ_d^{poly} (resp., Θ_d^{qp}) provided $\{\mathcal{A}_n\}_n \in \Theta_d^{S(n)}$ for some $S(n)$ of polynomial (resp., quasipolynomial) growth rate, i.e., $S(n) = n^{O(1)}$ (resp., $S(n) = 2^{(\log n)^{O(1)}}$). Let $\sigma\colon \mathbb{N} \to \mathbb{N}$. We are generally interested in asymptotic bounds on proof size; namely, the question of whether each set of cedents $\mathcal{A}_n$ has a $\Theta_d^{S(n)}$-LK refutation of size $\sigma(n)$. Often we are interested in $S(n)$ of (quasi-)polynomial growth rate; in this case we write: $\mathcal{A}_n$ has a Θ_d^{poly}-LK-refutation (resp., Θ_d^{qp}-LK-refutation) of size $\sigma(n)$. As usual, proof size may be measured in terms of sequence-size, tree-size, or proof height, and cedent-size may be used instead of size.

These definitions allow us to identify refutation systems, which correspond in some sense to the set of consequences of a logical theory. By a *refutation system* we mean a collection of sequences of sets of cedents which all have a common asymptotic provability complexity. We define $\{\mathcal{A}_n\}_n \in \mathcal{Q}_d^{\text{seq}}$ iff each $\mathcal{A}_n$ has a d-LK-refutation of quasi-polynomial sequence-size. Most of the time we sloppily write $\mathcal{A}_n \in \mathcal{Q}_d^{\text{seq}}$ instead of $\{\mathcal{A}_n\}_n \in \mathcal{Q}_d^{\text{seq}}$. Similarly, we define $\mathcal{A}_n \in \mathcal{Q}_d^{\text{tree}}$ (resp., $\mathcal{A}_n \in \mathcal{Q}_d^{\text{height}}$) iff each $\mathcal{A}_n$ has a d-LK-refutation of quasi-polynomial tree-size (resp., simultaneously of quasi-polynomial tree-size and of poly-logarithmic height). Note that $\mathcal{A}_n \in \mathcal{Q}_d^{\text{seq}}$ iff each $\mathcal{A}_n$ has a d-LK-refutation of quasi-polynomial sequence-size, and iff each $\mathcal{A}_n$ has a Θ_d^{qp}-LK-refutation of quasi-polynomial sequence-*cedent*-size. Similarly, $\mathcal{A}_n \in \mathcal{Q}_d^{\text{tree}}$ iff each $\mathcal{A}_n$ has a d-LK-refutation of quasi-polynomial tree-size, and iff each $\mathcal{A}_n$ has a Θ_d^{qp}-LK-refutation of quasi-polynomial tree-*cedent*-size.

Similar notions can be defined with respect to Θ_d^{poly}, with the exception of height restricted refutation systems. I.e., $\mathcal{A}_n \in \mathcal{P}_d^{\text{seq}}$ (resp., $\mathcal{A}_n \in \mathcal{P}_d^{\text{tree}}$) iff each $\mathcal{A}_n$ has a d-LK-refutation of polynomial sequence-size (resp., of polynomial tree-size). We define $\mathcal{A}_n \in \mathcal{P}_d^{\text{height}}$ iff each $\mathcal{A}_n$ has a d-LK-refutation which simultaneously has polynomial tree-size, logarithmic height and has $O(1)$ many formulas in each cedent. The remark on size versus cedent-size remains true also in the polynomial case.

A refutation system $\mathcal{R}$ is Θ_d^{qp}*-equivalent to* a refutation system $\mathcal{S}$ iff $\mathcal{S} \cap \Theta_d^{\text{qp}} = \mathcal{R} \cap \Theta_d^{\text{qp}}$. We denote this by $\mathcal{R} =_{\Theta_d^{\text{qp}}} \mathcal{S}$. Similarly, $\mathcal{R} \subsetneq_{\Theta_d^{\text{qp}}} \mathcal{S}$ means that $\mathcal{R} \cap \Theta_d^{\text{qp}} \subsetneq \mathcal{S} \cap \Theta_d^{\text{qp}}$. The notions Θ_d^{poly}*-equivalence* and $\subsetneq_{\Theta_d^{\text{poly}}}$ can be defined similarly. In these terms our main results of this paper are, for $d \in \frac{1}{2}\mathbb{N}$,

$$\mathcal{Q}_{d-1}^{\text{seq}} \quad =_{\Theta_{d-1}^{\text{qp}}} \quad \mathcal{Q}_d^{\text{tree}} \quad =_{\Theta_d^{\text{qp}}} \quad \mathcal{Q}_{d+1}^{\text{height}},$$

where the first equality only holds if $d - 1 \geq 0$, and

$$\mathcal{Q}_{d+\frac{1}{2}}^{\text{tree}} \quad \subsetneq_{\Theta_d^{\text{qp}}} \quad \mathcal{Q}_{d+1}^{\text{tree}} \;;$$

and, for polynomial size refutation systems,

$$\mathcal{P}^{\mathrm{seq}}_{d-1} \quad =_{\Theta^{\mathrm{poly}}_{d-1}} \quad \mathcal{P}^{\mathrm{tree}}_{d} \quad =_{\Theta^{\mathrm{poly}}_{d}} \quad \mathcal{P}^{\mathrm{height}}_{d+1},$$

where the first equality only holds if $d-1 \geq 0$, and

$$\mathcal{P}^{\mathrm{tree}}_{d+\frac{1}{2}} \quad \subsetneq_{\Theta^{\mathrm{poly}}_{d}} \quad \mathcal{P}^{\mathrm{tree}}_{d+1} \ .$$

2 Sequence-size, tree-size and height of proofs

The next theorem summarizes the main results of this section: variants of this theorem are already known for other versions of constant depth proof systems. Theorem 2 is stated in terms of quasipolynomial size proofs. At the end of this section, after Theorem 2 has been proved, we shall formulate Theorem 10, which will be a refined version of Theorem 2 that applies to polynomial size proofs.

Theorem 2. *Let $d \in \frac{1}{2}\mathbb{N}$, and $\{\mathcal{A}_n\}_n \in \Theta^{\mathrm{qp}}_d$ be a family of sets of cedents. Then the following conditions* (1) *and* (2) *are equivalent:*

(1) *$\mathcal{A}_n$ has a d-*LK *refutation of sequence-size quasi-polynomial in n, for all n.*

(2) *$\mathcal{A}_n$ has a $(d+1)$-*LK *refutation of tree-size quasi-polynomial in n, for all n.*

Furthermore, the following conditions (3) *and* (4) *are equivalent:*

(3) *$\mathcal{A}_n$ has d-*LK *refutation of tree-size quasi-polynomial in n, for all n.*

(4) *$\mathcal{A}_n$ has a $(d+1)$-*LK *refutation which simultaneously has tree-size quasi-polynomial in n and height poly-logarithmic in n, for all n.*

Before beginning the proof of Theorem 2, we present without proof a simple lemma which states that bounding the height of a derivation also bounds the tree-size and the size of cedents in the proof. This lemma is proved by induction on the height.

Lemma 3. *Assume $\mathcal{A}$ has an* LK *refutation of height bounded by η, and the connectives of every formula in the refutation have fanin $\leq S$, i.e., every formula is in Θ^S_d for some d. Then, the tree-cedent-size of this refutation is bounded by S^η, and each cedent in this proof consists of at most η many formulas.*

Theorem 2 is proved in stages as a series of lemmas, beginning with (1)$\Rightarrow$(2). Fix $d \in \frac{1}{2}\mathbb{N}$. The next lemma is due essentially to Krajíček.

Lemma 4. *Assume $\mathcal{A}$ has a Θ_d^σ-LK refutation R of sequence-size $\leq \sigma$. Then $\mathcal{A}$ has a Θ_{d+2}^σ-LK refutation R' which is simultaneously of height $\log\sigma + O(1)$ and tree-size $O(\sigma^4)$. Furthermore, each cedent in R' has $O(1)$ many formulas.*

A similar statement holds in terms of cedent-size. That is, a Θ_d^σ-LK refutation of sequence-cedent-size $\leq \sigma$ can be transformed into a Θ_{d+2}^σ-LK refutation which simultaneously has height $\log\sigma + O(1)$ and tree-cedent-size $O(\sigma^3)$ and has $O(1)$ many formulas in each cedent.

Proof. We only give a sketch of the proof. A refutation of $\mathcal{A}$ of sequence-size $\leq \sigma$ can be written as a sequence of cedents $\Gamma_1, \ldots, \Gamma_\sigma = \emptyset$ where each Γ_i is a logical axiom, a non-logical axiom Γ or $\bigvee\Gamma$ with $\Gamma \in \mathcal{A}$, or is formed from previous cedents $\Gamma_1, \ldots, \Gamma_{i-1}$ by applying one of the rules of LK. Let γ_i be $\bigwedge_{j<i}\bigvee\Gamma_j$. We claim that the cedent $\neg\gamma_i, \gamma_{i+1}$ has a cut-free LK derivation from $\mathcal{A}$ of constant height and of size $O(\sigma^3)$. For example, in the case that Γ_i is the cedent $\Delta, \bigwedge_{j<J}\varphi_j$ and is derived by an $\bigwedge$-inference from the cedents $\Gamma_{i_j} = \Delta, \varphi_j$, there is a derivation

$$
(\textstyle\bigwedge)\,\dfrac{
(\textstyle\bigvee)\,\dfrac{
(\textstyle\bigwedge)\,\dfrac{
(\textstyle\bigvee)\,\dfrac{
(\textstyle\bigwedge)\,\dfrac{\neg\varphi_j, \varphi_j \qquad (\textstyle\bigvee)\,\dfrac{\neg\gamma, \gamma}{\neg\gamma, \bigvee\Gamma_i} \quad \gamma \in \Delta}{\bigwedge\neg\Gamma_{i_j}, \bigvee\Gamma_i, \varphi_j}
}{\neg\gamma_i, \bigvee\Gamma_i, \varphi_j} \;,\; j < J
}{\neg\gamma_i, \bigvee\Gamma_i, \bigwedge_{j<J}\varphi_j}
}{\neg\gamma_i, \bigvee\Gamma_i}
\qquad
(\textstyle\bigvee)\,\dfrac{\neg\bigvee\Gamma_j, \bigvee\Gamma_j}{\neg\gamma_i, \bigvee\Gamma_j} \;,\; j < i
}{\neg\gamma_i, \gamma_{i+1}}
$$

Clearly this derivation has height $O(1)$, the number of cedents is $O(\sigma^2)$, and the total number of symbols $O(\sigma^3)$.

Now, application of $\sigma - 1$ many cuts to the cedents $\neg\gamma_i, \gamma_{i+1}$ gives a Θ_{d+2}^σ-LK refutation of $\mathcal{A}$; performing the cuts in a balanced pattern makes the refutation have height $\log\sigma + O(1)$ and size $O(\sigma^4)$. This proves the lemma. $\square$

Lemma 4 combined with the implication (4)$\Rightarrow$(3), which we prove next, suffices to prove the implication (1)$\Rightarrow$(2).

The next lemma is based on Razborov [22] and Krajicek [13, §12.2].

Lemma 5. *Assume* $\bigcup\mathcal{A} \subset \Theta_d^\sigma$ *and* $\mathcal{A}$ *has a* Θ_{d+1}^σ-LK *refutation of tree-cedent-size* $\le \sigma$, *where each cedent in the refutation consists of at most* λ *many formulas. Then* $\mathcal{A}$ *has a* Θ_d^σ-LK *refutation of tree-cedent-size* $\le \sigma^{\lambda+1}$.

Proof. We shall give only a sketch of the idea of the proof. Each cedent Γ in the LK refutation R is of the form $\Gamma = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$ where $\Gamma_1 = \Gamma \cap \Theta_d^\sigma$, $\Gamma_2 = \{\bigvee_{j<n_i} A_{i,j} : i < N\}$ and $\Gamma_3 = \{\bigwedge_{j<m_i} B_{i,j} : i < M\}$. Each cedent Γ is then replaced by the collection of cedents

$$\Gamma_1 \cup \{A_{i,j} : i < N, j < n_i\} \cup \{B_{i,f(i)} : i < M\}, \tag{1}$$

for each function f such that $f(i) < m_i$ for all $i < M$. These cedents are evidently all in Θ_d^σ; furthermore, it is straightforward to modify the refutation R so as to be a refutation on the cedents of the form (1). In addition, the modified refutation is still tree-like. This gives the desired Θ_d^σ-LK refutation. To bound the tree-size of the new refutation, we claim that a single cedent Γ becomes at most σ^λ many cedents. This is because $M \le \lambda$ and since each $m_i \le \sigma$. □

Lemmas 3 and 5 together show (4)⇒(3). The implication (2)⇒(1) is a consequence of the following lemma:

Lemma 6. *Assume* $\bigcup\mathcal{A} \subset \Theta_d^\sigma$, *and* $\mathcal{A}$ *has a* Θ_{d+1}^σ-LK *refutation of tree-size* σ. *Then* $\mathcal{A}$ *has a* Θ_d^σ-LK *refutation of sequence-size* $3\sigma^2$.

The proof also establishes a similar assertion about cedent-size: A Θ_{d+1}^σ-LK refutation of tree-cedent-size σ can be transformed into a Θ_d^σ-LK refutation of sequence-cedent-size 3σ.

Proof. Krajíček [11] proves this lemma with a bound of σ^4. We can argue the stronger bound of $3\sigma^2$ as follows. As in the previous proof, each cedent Γ in the Θ_{d+1}^σ-LK refutation can be written in the form $\Gamma = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$ with $\Gamma_1 = \Gamma \cap \Theta_d^\sigma$, $\Gamma_2 = \{\bigvee_{j<n_i} A_{i,j} : i < N\}$ and $\Gamma_3 = \{\bigwedge_{j<m_i} B_{i,j} : i < M\}$. Let Δ be the cedent $\Delta = \Gamma_1 \cup \{A_{i,j} : i < N, j < n_i\}$. Also, for $i < M$, let Π_i be the cedent $\Pi_i = \{\neg B_{i,j} : j < m_i\}$; note Π_i expresses the negation of $\bigwedge_{j<m_i} B_{i,j}$.

We claim that if Γ has a Θ_{d+1}^σ-LK proof from the hypotheses $\mathcal{A}$ with tree-cedent-size σ, then Δ has a Θ_d^σ-LK proof from the hypotheses $\mathcal{A} \cup \{\Pi_i : i < M\}$ with sequence-cedent-size $\le 3\sigma$. This claim is straightforward to prove by induction on σ, with the argument splitting into cases depending on the type of inference that derives the endcedent of the refutation. (We leave the proof of the claim to the reader.)

We observe that each cedent in the constructed derivation has size $\leq \sigma$, because each formula in the cedent occurred in the original derivation. Thus, the sequence-size is $\leq 3\sigma^2$.

The lemma follows from this claim by letting Γ be the empty set. $\square$

To finish the proof of Theorem 2, we still need to establish (3)$\Rightarrow$(4). For $d \geq 1$, we can already deduce this from (2)$\Rightarrow$(1) and Lemma 4, but we also want it for $d = 0$ and $d = \frac{1}{2}$. This is accomplished by the next two lemmas. We temporarily use a modified version LK$'$ of the Gentzen-Tait calculus. LK$'$ is the system LK augmented with the following generalized cut rule:

$$\text{G-Cut: } \frac{\Gamma, \bigwedge_{i=0}^{n} \neg\varphi_i \qquad \Gamma, \varphi_0, \ldots, \varphi_n}{\Gamma} \tag{2}$$

Lemma 7. *Suppose $\Gamma, \bigcup \mathcal{A} \subset \Theta_d^\sigma$ and that P is a Θ_d^σ-LK$'$ derivation of the cedent Γ from the hypotheses $\mathcal{A}$. Also suppose the tree-size of P is $\leq \sigma$, and let τ equal the tree-cedent-size of P. Then there is a Θ_{d+1}^σ-LK$'$ derivation of Γ from $\mathcal{A}$ which has height $O(\log \tau)$ and tree-size $O(\sigma^3)$.*

Proof. The proof is a Spira-Brent style divide-and-conquer construction, but with some complications caused by the fact that the lines in derivations are cedents instead of formulas. We shall prove that P can be transformed into a derivation of height $\leq 4\log\tau$, by induction on τ. (The logarithm is base two and by convention $\log 1 = 1$.) Since $\tau < 4\log\tau$ for $\tau \leq 15$, we may assume $\tau \geq 16$.

Assume that the height of P is greater than $4\log\tau$. To start the divide-and-conquer reduction of P, choose an inference I with lower cedent Δ with the following two properties:

(a) The total number of cedents in the subderivation ending with Δ is $> \lceil \tau/2 \rceil$.

(b) Further, each hypothesis Π of I is derived with a subderivation of tree-cedent-size $\leq \lceil \tau/2 \rceil$.

We have chosen Δ so that if the subderivation ending with I is discarded, leaving Δ as an initial sequent, then the resulting derivation has tree-cedent-size $\leq \tau/2$.

By the induction hypothesis, each hypothesis Π to I has a proof of height $\leq 4\log\lceil\tau/2\rceil < 4\log\tau - 3$, since $\tau \geq 16$. By putting these proofs together with the inference I, and applying an additional weakening with Γ, the cedent Γ, Δ has a derivation Q from $\mathcal{A}$ of height less than $4\log\tau - 1$.

Now consider the rest of the derivation P. Discarding the subderivation above Δ, we have a derivation of Γ from the hypotheses $\mathcal{B} = \mathcal{A} \cup \{\Delta\}$. By the induction hypothesis, there is a derivation R of Γ from $\mathcal{B}$ of height $\leq 4\log(\tau/2) = 4\log\tau - 4$. Let Δ be the cedent $\delta_1, \ldots, \delta_k$. We can convert R into a derivation R' of $\bigwedge_i \neg\delta_i, \Gamma$ from (only) the hypotheses $\mathcal{A}$ by the following construction: First, replace the cedent Δ, which is a non-logical axiom in R, with the derivation

$$\frac{\dfrac{\delta_i, \neg\delta_i}{\Delta, \neg\delta_i} \qquad , i = 1, \ldots, k}{\Delta, \bigwedge_i \neg\delta_i} \tag{3}$$

Second, insert weakening inferences just below every other initial cedent Π in R to derive $\bigwedge_i \neg\delta_i, \Pi$. Further add the formula $\bigwedge_i \neg\delta_i$ to every other cedent in R. The resulting derivation R' has height at most two greater than the height of R, that is, has height at most $4\log\tau - 2$.

Finally, combine the derivations Q and R' with a *G-Cut* inference to get the desired derivation, P', of Γ from the hypotheses $\mathcal{A}$. Clearly, the resulting derivation has height $< 4\log\sigma$ as desired.

Next we bound the tree-cedent-size of the proof P'. For this, let λ be the maximum number of formulas in any cedent of P. Second, define $csz(P')$ to be the number of cedents in P', but excluding from the count any cedents $\delta_i, \neg\delta_i$ and $\Delta, \neg\delta_i$ introduced in the subderivations (3). (In effect, by using csz, we are temporarily allowing initial sequents of the form $\Delta, \bigwedge_i \neg\delta_i$.) We claim that $csz(P') < \tau^2$ and prove this by induction on τ. Let S be the tree-cedent-size of the derivation obtained by discarding the subderivation above Δ, and let S_i be the tree-cedent-size of the i-th premise to I. W.l.o.g. we may assume that we have exactly three premises to I, because we could group premises appropriately. Then $\tau = S + S_1 + S_2 + S_3$ and $S \leq \frac{\tau}{2}$ and $S_i \leq \frac{\tau}{2}$ for $i = 1, 2, 3$ by assumption. By the induction hypothesis, $csz(Q) \leq \sum_{i=1}^{3} S_i^2 + 2$. Also, $csz(R') < 2csz(R) \leq 2(S^2 - 1)$. Thus, $csz(P') = csz(Q) + csz(R') + 1 \leq 2S^2 + \sum_{i=1}^{3} S_i^2$. In case $S \leq \frac{\tau}{4}$ we obtain from this

$$csz(P') \leq 2\left(\frac{\tau}{4}\right)^2 + 3\left(\frac{\tau}{2}\right)^2 < \tau^2.$$

In the other case $S > \frac{\tau}{4}$, hence one of S_1, S_2, S_3 must be $< \frac{\tau}{4}$, and another $< \frac{3\tau}{8}$, thus

$$csz(P') \leq 2\left(\frac{\tau}{2}\right)^2 + \left(\frac{\tau}{4}\right)^2 + \left(\frac{3\tau}{8}\right)^2 + \left(\frac{\tau}{2}\right)^2 < \tau^2.$$

To bound the actual tree-cedent-size of P', note that at most $\tau \cdot (2\lambda)$ cedents in P' do not count to towards $csz(P')$. Thus the tree-cedent-size of P is bounded by $\tau^2 + 2\tau\lambda = O(\sigma^2)$.

Finally, it is clear from the construction, that every sequent in P' has size $\leq \sigma$. Thus, P' has tree-size $O(\sigma^3)$. □

The next lemma lets us eliminate the use of *G-Cut* inferences.

Lemma 8. *Suppose P is a Θ_d^σ-LK$'$ derivation of Δ from the hypotheses $\mathcal{A}$, where $\bigcup \mathcal{A} \subset \Theta_{d-1}^\sigma$. Further suppose P has height η and tree-size σ. Then there is a Θ_d^σ-LK derivation of Δ from the hypotheses $\mathcal{A}$, which has height at most $2\eta + 3$ and tree-size $O(\sigma^2)$.*

Formulating this lemma in terms of cedent-size would give us a derivation of height at most $2\eta + 3$ and tree-cedent-size $O(\sigma^2)$.

Proof. We sketch a proof of a strengthened form of the lemma to let the induction run smoothly. Suppose that P is a proof from the hypotheses $\mathcal{A}$ with endcedent $\Delta = \Gamma, \Phi_1, \ldots, \Phi_k$, where each Φ_i is a cedent. Also suppose $\Phi_i' \supseteq \Phi_i$, for $1 \leq i \leq k$, and let Δ' be the cedent $\Gamma, \bigvee \Phi_1', \ldots, \bigvee \Phi_k'$. Let η be the height of P and let σ bound both the tree-size of P and the size of the sequent Δ'. We claim there is a Θ_d^σ-LK proof of Δ' from the hypotheses $\mathcal{A}$ which has height at most $2\eta + 3$ and tree-size $O(\sigma^2)$. The lemma follows from this claim by letting $k = 0$.

The claim is proved by induction on the height of P, with the proof splitting into cases depending on the last inference of P. All cases but axioms and *G-Cut* follow immediately from the induction hypothesis and eventually one application of $(\bigvee)$ if the derived formula is in one of the Φ_i's. These cases ($\bigvee$, $\bigwedge$ and weakening inferences) add at most one cedent to the height of the proof. The case of *G-Cut* follows immediately from induction hypothesis and a cut, with no change to the height of the proof, because we formulated the strengthened form of our lemma for just this purpose.

In the case of an axiom Φ such that $\Phi = \Gamma, \Phi_1, \ldots, \Phi_k$ for $\Phi \in \mathcal{A}$, we build the following derivations: For $\varphi \in \Gamma$, a single logical axiom followed by a weakening inference gives

$$\Gamma, \bigvee \Phi_1', \ldots, \bigvee \Phi_k', \neg\varphi.$$

For $\varphi \in \Phi_i$, we derive, in height 3,

$$\bigvee + \text{Weakening} \frac{\varphi, \neg\varphi}{\Gamma, \bigvee \Phi_1', \ldots, \bigvee \Phi_k', \neg\varphi}$$

Putting the derivations together with a ($\bigwedge$) gives us the following derivation from $\mathcal{A}$:

$$
\text{Cut}\ \dfrac{\text{Weakening}\ \dfrac{\bigvee\Phi}{\Gamma, \bigvee\Phi'_1, \ldots, \bigvee\Phi'_k, \bigvee\Phi} \qquad \bigwedge\ \dfrac{\ddots\ \vdots\ \iddots}{\Gamma, \bigvee\Phi'_1, \ldots, \bigvee\Phi'_k, \bigwedge\neg\Phi}}{\Gamma, \bigvee\Phi_1, \ldots, \bigvee\Phi_k}
$$

This is a Θ_d^σ-LK derivation of height 5.

In the case of an axiom $\neg\varphi, \varphi$, or an axiom $\bigvee\Phi$ for some $\Phi \in \mathcal{A}$, the assertion follows by the same axiom plus at most two applications of $\bigvee$-inferences for formulas in Φ_i and possibly a weakening inference. This derivation has height at most 4.

It is clear that the constructed proof has height at most $2\eta + 3$. Likewise, the tree-size bound $O(\sigma^2)$ is evident from the construction. $\Box$

We combine the last two lemmas.

Corollary 9. *Suppose $\bigcup\mathcal{A} \subset \Theta_d^\sigma$ and that $\mathcal{A}$ has a Θ_d^σ-LK refutation of tree-size σ. Then $\mathcal{A}$ has a Θ_{d+1}^σ-LK refutation of height $O(\log\sigma)$.* $\Box$

We end this section with a refined version of Theorem 2 that deals with polynomial size instead of quasi-polynomial size.

Theorem 10. *Let $d \in \frac{1}{2}\mathbb{N}$, and $\{\mathcal{A}_n\}_n \in \Theta_d^{\mathrm{poly}}$ be a family of sets of cedents. Then the following conditions (1) and (2) are equivalent:*

(1) *$\mathcal{A}_n$ has a d-LK refutation of sequence-size polynomial in n, for all n.*

(2) *$\mathcal{A}_n$ has a $(d+1)$-LK refutation of tree-size polynomial in n, for all n.*

Furthermore, the following conditions (3) and (4) are equivalent:

(3) *$\mathcal{A}_n$ has a d-LK refutation of tree-size polynomial in n, for all n.*

(4) *$\mathcal{A}_n$ has a $(d+1)$-LK refutation which simultaneously has tree-size polynomial in n, has height logarithmic in n, and has $O(1)$ many formulas in each cedent, for all n.*

This theorem follows immediately from Lemmas 3 through Corollary 9.

3 The Ordering Principle

This section introduces the ordering principle tautologies, and their generalizations obtained by substituting Sipser functions for variables. The ordering principle is a well-known propositional tautology; it was introduced by Krishnamurthy [17], and has been studied by Stålmarck [25], Bonet-Galesi [8] et. al. It is also the same as the Graph Ordering Principles used by Segerlind et. al [24] when restricted to complete graphs. The idea of using Sipser functions to transform hard sets of tautologies into hard sets of depth d formulas comes from Krajíček [11].

Let $\prec$ be a binary relation. The *Ordering Principle* $\mathrm{OP}(\prec)$ states that if $\prec$ is a transitive and irreflexive relation on the set $[n] = \{0, \ldots, n-1\}$, i.e., a partial ordering, then $\prec$ has a minimal element on $[n]$:

$$\bigwedge_{x \in [n]} \neg x \prec x \;\wedge \bigwedge_{x,y,z \in [n]} (x \prec y \wedge y \prec z \rightarrow x \prec z) \;\rightarrow\; \bigvee_{x \in [n]} \bigwedge_{y \in [n]} \neg y \prec x.$$

To write $\mathrm{OP}(\prec)$ as a propositional formula, we use variables $p_{x,y}$ to stand for the condition $x \prec y$. The negated formula $\neg\,\mathrm{OP}(\prec)$ can be written as the following set of clauses:

M_x	$\{\, p_{0,x}, \ldots, p_{n-1,x} \,\}$	<u>M</u>inimal element clause
I_x	$\{\, \overline{p}_{x,x} \,\}$	<u>I</u>rreflexive relation clause
$T_{x,y,z}$	$\{\, \overline{p}_{x,y}, \overline{p}_{y,z}, p_{x,z} \,\}$	<u>T</u>ransitive relation clause.

Since x, y, z range over $[n]$, we call this set of clauses $\neg\,\mathrm{OP}(n)$.

When $\mathrm{OP}(\prec)$ is rewritten in disjunctive normal form as a formula involving the variables $p_{x,y}$, it has depth 2. Note that this depth comes essentially from the last part of the formula, which corresponds to the M_x clauses; furthermore, these clauses have only positive occurrences of variables. For this reason, if we replace each variable $p_{x,y}$ (expressing the condition $x \prec y$) by a Sipser function of the form $\bigvee_{i \in [n]} q^i_{x,y}$, then, after distributing out, we obtain a formula of the same depth and still of polynomial size. More exactly, replacing $p_{x,y}$ by $\bigvee_{i \in [n]} q^i_{x,y}$ in the formula $\neg\,\mathrm{OP}(\prec)$ and distributing out, we obtain

$$\bigwedge_{x \in [n]} \bigwedge_{i \in [n]} \overline{q}^i_{x,x} \;\wedge \bigwedge_{x,y,z \in [n]} \bigwedge_{i_1 \in [n]} \bigwedge_{i_2 \in [n]} (\overline{q}^{i_1}_{x,y} \vee \overline{q}^{i_2}_{y,z} \vee \bigvee_{i_3 \in [n]} q^{i_3}_{x,z})$$
$$\wedge \bigwedge_{x \in [n]} \bigvee_{y \in [n]} \bigvee_{i \in [n]} q^i_{y,x} \;.$$

This formula is designated $\neg\,\mathrm{OP}^0(n)$.

$\neg\,\mathrm{OP}^0(n)$ can also be written as a set of clauses: For $x, y, z, i, i_1, i_2 \in [n]$ we obtain the clauses

$$\begin{array}{ll} M_x & \bigcup_{y\in[n]}\{\, q^0_{y,x}, \ldots, q^{n-1}_{y,x} \,\} \\ I_{i,x} & \{\, \overline{q}^i_{x,x} \,\} \\ T_{i_1,i_2,x,y,z} & \{\, \overline{q}^{i_1}_{x,y}, \overline{q}^{i_2}_{y,z}, q^0_{x,z}, \ldots, q^{n-1}_{x,z} \,\} \ . \end{array}$$

The M (resp., I, T) clauses have size n^2 (resp., 1, $n+2$), and there are n (resp., n^2, n^5) of them. Thus, $\neg\,\mathrm{OP}^0(n)$ has size $n^{O(1)}$.

We now define higher depth propositional principles from $\neg OP^0(n)$ by replacing variables by Sipser functions. Let $d \in \mathbb{N}$. The Sipser functions of type $\bigwedge$ are defined by

$$\begin{aligned} \mathrm{S}^{\wedge}_{d,n}(\sigma) &= \bigwedge_{y_1<n} \ \bigvee_{y_2<n} \ \cdots \ \mathop{Q^{d-1}}_{y_{d-1}<n} \ \mathop{Q^{d}}_{y_d<n} \ p_{\sigma,y_1,\ldots,y_d} \\ \mathrm{S}^{\wedge}_{d+0.5,n}(\sigma) &= \bigwedge_{y_1<n} \ \bigvee_{y_2<n} \ \cdots \ \mathop{Q^{d-1}}_{y_{d-1}<n} \ \mathop{Q^{d}}_{y_d<n} \ \mathop{Q^{d+1}}_{y_{d+1}<(\log n)^2} \ p_{\sigma,y_1,\ldots,y_d,y_{d+1}} \end{aligned}$$

where each Q^j is $\bigvee$ or $\bigwedge$, depending on whether d is even or odd, respectively. The Sipser functions of type $\bigvee$ are obtained by exchanging $\bigwedge$ and $\bigvee$.

For $d \in \frac{1}{2}\mathbb{N}$, we define $\neg\,\mathrm{OP}^d(n)$ by replacing the variables $q^i_{x,y}$ in $\neg\,\mathrm{OP}^0(n)$ by $\mathrm{S}^{\wedge}_{d,n}(x,y,i)$. Note that $\neg\,\mathrm{OP}^d(n)$ is equivalent to $\neg\,\mathrm{OP}(n)$ where each $p_{x,y}$ is replaced by $\mathrm{S}^{\vee}_{d+1,n}(x,y)$. Clearly, $\neg\,\mathrm{OP}^d(n)$ is a polynomial size set of clauses of (Θ-)depth d formulas.

Lemma 11. *The sets $\neg\,\mathrm{OP}^0(n)$ have polynomial sequence-size* 0-LK *refutations.*

Note that for 0-refutations, there is no distinction between polynomial size and polynomial cedent-size, since each literal may occur only once in a given cedent.

Proof. It is already known that the $\neg\,\mathrm{OP}(\prec)$ clauses have polynomial sequence-size refutations (c.f. [24]); and it is possible to generalize these to form refutations of the "Sipser-ized" versions $\neg\,\mathrm{OP}^0(n)$.

However, we will sketch a different proof of the lemma, based on Theorem 2. We shall first prove there is a polynomial tree-size log-height refutation of $\neg\,\mathrm{OP}^0(n)$. For $u \in [n]$, define the depth two formula $\varphi(u)$ to be

$$\bigvee_{x\in[u]}\bigwedge_{y\in[u]}\bigwedge_{i\in[n]}\overline{q}^{i}_{y,x}.$$

$\varphi(u)$ states that there is a $\prec$-minimal element x in the set $[u]$. We claim that each cedent $\neg\varphi(u), \varphi(u+1)$ has a constant-height derivation P_u from the assumption $\neg\,\mathrm{OP}^0(n)$. The derivation P_u can be informally described as follows: Suppose $\varphi(u)$ is true by virtue of $x = x_0 \in [u]$ being a minimal element in $[u]$. Then, if $u \not\prec x_0$, x_0 is also a minimal element in $[u+1]$. Otherwise, if $u \prec x_0$, then u is a minimal element in $[u+1]$ because of the transitivity and irreflexivity of $\prec$. It is straightforward to check that P_u has constant height and polynomial tree-size. Therefore, each cedent in P_u has only constantly many formulas.

We now combine the derivations P_u of $\neg\varphi(u), \varphi(u+1)$ in a balanced tree-like fashion with cuts. This gives a proof Q of the cedent containing the sole formula $\varphi(n)$, since $\neg\varphi(0)$ is the empty (false) disjunction. By construction, Q has height $\log n + O(1)$ and every cedent in Q has a bounded number of formulas.

Finally, from the initial cedents $\bigvee_{y\in[n],\, i\in[n]} q^{i}_{x,y}$ (which are initial cedents because of the clauses M_x) we use a single $\bigwedge$-inference to derive $\neg\varphi(n)$, and then perform a cut against the endcedent of Q to derive the empty cedent. This gives a refutation R of $\neg\,\mathrm{OP}^0(n)$. Every cedent in R has a bounded number of formulas, R has height $\log n + O(1)$ and every formula in R is in Θ^n_2.

Applying Lemmas 5 and 6 to R, we get that $\neg\,\mathrm{OP}^0(n)$ has a polynomial tree-size depth one refutation, and a polynomial sequence-size, depth zero refutation. $\square$

Substituting Sipser functions into the refutations shown to exist in (the proof of) Lemma 11 gives the following bounds on proof size of the Sipser-ized ordering principles.

Corollary 12. *Let $d \in \frac{1}{2}\mathbb{N}$. The sets $\neg\,\mathrm{OP}^d(n)$ have polynomial sequence-size d-LK refutations, i.e., $\Theta^{S(n)}_d$-refutations of sequence-size $\le S(n)$ for some polynomial $S(n)$.*

Corollary 13. *Let $d \in \frac{1}{2}\mathbb{N}$. The sets $\neg\,\mathrm{OP}^d(n)$ have polynomial tree-size $(d+1)$-LK refutations, i.e., $\Theta^{S(n)}_{d+1}$-refutations of tree-size $\le S(n)$ for some polynomial $S(n)$.*

4 Exponential lower bounds and separations

The main goal for this section will be the following

Theorem 14. *Let $d \in \frac{1}{2}\mathbb{N}$ and $0 < \epsilon < \frac{1}{2}$. For n sufficiently large, any $(d+\frac{1}{2})$-LK refutation of $\neg\,\mathrm{OP}^d(n)$ must have tree-size $\geq 2^{n^\epsilon}$.*

This lower bound combined with Corollary 13 gives the desired exponential separation:

Corollary 15. *Fix $d \in \frac{1}{2}\mathbb{N}$. For all sufficiently large S, there are sets of clauses of depth d formulas (expressing the negations of tautologies of depth $(d+2)$ related to the ordering principle) which have $(d+1)$-LK refutations of tree-size $\leq S$, but every $(d+\frac{1}{2})$-LK refutation of them requires tree-size $\geq 2^{S^{\Omega(1)}}$.*

Before we can prove Theorem 14 we need a technical lemma which allows us to reduce the complexity of formulas in a refutation. We study this in the next subsection.

4.1 Cut-reduction by switching

Cut-elimination procedures are a classic way of reducing the complexity of formulas in proofs. The usual Gentzen- or Tait-style cut elimination methods work by eliminating the outermost connectives of cut formulas first. In general, these traditional cut-elimination procedures have the cost of an exponential increase (or worse) in the size of proofs. We shall instead use a cut-reduction method based on propositional restrictions and the Håstad switching lemma to reduce the complexity of formulas in the proof. The idea is to find a restriction (i.e., a partial substitution of propositional variables by truth values) and simplify a proof by applying the restriction to all formulas in the proof. This has the advantage that proof size is not increased and the structure of the proof can only be simplified. However, the disadvantage is that all formulas in the proof are reduced. That is, not only is the complexity of cut formulas reduced, but also the complexity of the derived formula. Fortunately, careful use of Sipser functions allows us to control the amount of reduction of complexity of formulas.

Our cut-reduction with switching lemmas will be based closely on techniques used by [9] to reduce the complexity of oracle computations related to definable functions of bounded arithmetic. This same approach has been used by [11] to reduce the complexity of d-LK proofs where

$d \in \mathbb{N} + \frac{1}{2}$, and by [7] to separate height restricted derivation systems using the order induction principle for the natural ordering $<$ on $\mathbb{N}$.

Before stating the switching theorems, we need some notation. The reader may wish to consult [9] for more details.

Fix $k, \ell, m \in \mathbb{N}$ with $m, k \geq 1$, $\ell \geq 0$. As before let $[m]$ denote the set $\{0, \ldots, m-1\}$. For $x_1, \ldots, x_k, y_1, \ldots, y_\ell \in \mathbb{N}$ let $p_{x_1,\ldots,x_k,y_1,\ldots,y_\ell}$ be a Boolean variable, and let

$$B_{k,\ell}(m) = \{p_{x_1,\ldots,x_k,y_1,\ldots,y_\ell} : x_1, \ldots, x_k, y_1, \ldots, y_\ell < m\}.$$

The cardinality of $B_{k,l}(m)$ is $m^{k+\ell}$. We shall henceforth use $\vec{x}$ as an abbreviation of $x_1, \ldots, x_k$, and $\vec{y}$ as an abbreviation of $y_1, \ldots, y_\ell$ or $y_1, \ldots, y_{\ell-1}$, depending on the context. Note that $B_{k,0}(m)$ is the set of variables $p_{\vec{x}}$ with $\vec{x} \in [m]^k$.

A propositional formula is $s\Sigma_1^t$ iff it is a disjunction of conjunctions of at most t literals. The $s\Pi_1^t$-formulas are the negations of $s\Sigma_1^t$-formulas. (Other authors use the terms "t-DNF" and "t-CNF" for $s\Sigma_1^t$- and $s\Pi_1^t$-formulas.) A formula is Δ_1^t iff it is equivalent to both a $s\Sigma_1^t$-formula and a $s\Pi_1^t$-formula. We call a formula Σ_1^t iff is a disjunction of Δ_1^t-formulas. The Π_1^t-formulas are negations of Σ_1^t-formulas. Observe that a Σ_1^t formula is equivalent to some $s\Sigma_1^t$ formula. Also observe that if all proper subformulas of a formula φ are in Δ_1^t, then φ is in $\Sigma_1^t \cup \Pi_1^t$.

Let $\ell \in \mathbb{N}$ and $d \in \frac{1}{2}\mathbb{N}$. For $\vec{x} \in [m]^k$, let $\mathrm{S}_{d,m}(\vec{x})$ denote one of the previously defined Sipser functions, i.e., either $\mathrm{S}^{\wedge}_{d,m}(\vec{x})$ or $\mathrm{S}^{\vee}_{d,m}(\vec{x})$. Note that $\mathrm{S}_{\ell,m}(\vec{x})$ involves the m^{k+l} variables from $B_{k,\ell}(m)$, and $\mathrm{S}_{\ell+0.5,m}(x_1, \ldots, x_k)$ involves $m^{k+\ell} \cdot (\log m)^2$ variables from $B_{k,\ell+1}(m)$.

We are now ready to formulate cut-reduction by switching. The notation $\mathcal{A}[p_{\vec{x}} \leftarrow \varphi_{\vec{x}} : \vec{x} \in M]$ denotes the result of simultaneously replacing the variable $p_{\vec{x}}$ by the formula $\varphi_{\vec{x}}$, for all $\vec{x} \in M$.

Theorem 16 (Cut-Reduction by Switching). *Let $d \in \frac{1}{2}\mathbb{N}$ and $\epsilon \in \mathbb{R}$ with $0 < \epsilon < \frac{1}{2}$. Let $M \subseteq \mathbb{N}$ be some infinite set. For $m \in M$ let η_m be in $\mathbb{N}$ with $\eta_m \leq m^\epsilon$, and let $\mathcal{A}_m$ be a set of formulas with variables in $B_{k,0}(m)$. Furthermore, assume that $\mathcal{A}'_m := \mathcal{A}_m\big[p_{\vec{x}} \leftarrow \mathrm{S}_{d+1,m}(\vec{x}) : \vec{x} \in [m]^k\big]$ has $\Theta^{2^{\eta_m}}_{d+1.5}$-*LK *refutations of height η_m .*

Then, for all $m \in M$ which are sufficiently large, $\mathcal{A}_m$ has LK *refutations of height $\leq \eta_m$ in which every occurring* proper *subformula is in $\Delta_1^{\eta_m}$, and hence, in which every occurring formula is in $\Sigma_1^{\eta_m} \cup \Pi_1^{\eta_m}$.*

Theorem 16 can be viewed as being weaker than the corresponding Main Theorem in [9]. Here we start with a principle, where a depth-$(d+1)$ Sipser-function is plugged in, but we switch only d times, which in the end allows us to recover the original principle by applying a further transformation. In [9], principles with depth-d Sipser-function plugged in are switched d times, so in the end only a modified principle is left (the last switching modifies the principle too). As we will see, for certain principles like the ordering principle, our weaker theorem even produces stronger separation results than previously known.

The proof of Theorem 16 will take up the rest of this section. For $d \in \mathbb{N}$, the proof is based primarily on the computations in [9]. For $d \in \mathbb{N} + \frac{1}{2}$ it needs some additional observations. To make these differences clear we have to repeat some technical definitions and statements from [10, 9].

Let $\ell \geq 1$. The sets $B_{k,\ell}(m)$ are partitioned into blocks of variables

$$(B_{k,\ell}(m))_{(x_1,\ldots,x_k,y_1,\ldots,y_{\ell-1})} \quad := \quad \{p_{x_1,\ldots,x_k,y_1,\ldots,y_{\ell-1},z} \,:\, z < m\},$$

for $(x_1,\ldots,x_k,y_1,\ldots,y_{\ell-1}) \in [m]^{k+\ell-1}$. Each block contains m variables.

A restriction ρ on $B_{k,\ell}(m)$ is a map going from $B_{k,l}(m)$ to $\{0,1,*\}$:

$$\rho \,:\, B_{k,\ell}(m) \rightarrow \{0,1,*\} \ .$$

We think of $\rho(p) = 0$ or $\rho(p) = 1$ as p *is replaced by* 0 or 1 respectively, and of $\rho(p) = *$ as p *is left untouched.* Alternatively, we can think of ρ as a partial map going from $B_{k,\ell}(m)$ to $\{0,1\}$.

The probability space $\mathbb{R}^+_{k,\ell,m}(q)$ of restrictions ρ for $0 < q < 1$ is defined by the following: Let $\vec{x} \in [m]^k$, $\vec{y} \in [m]^{\ell-1}$ and $y_\ell < m$.

$$\begin{tikzcd}
 & 1 & \\
\rho \in \mathbb{R}^+_{k,\ell,m}(q)\colon p = p_{\vec{x},\vec{y},y_\ell} \arrow[ur, mapsto, "1-q"] \arrow[dr, mapsto, "q"'] & & 0 \\
 & s_{\vec{x},\vec{y}}(\rho) \arrow[ur, mapsto, "1-q"] \arrow[dr, mapsto, "q"'] & \\
 & & *
\end{tikzcd}$$

This means that the value of $\rho(p)$, where $p = p_{\vec{x},\vec{y},y_\ell}$, is determined by: first choosing $s_{\vec{x},\vec{y}}$ such that $s_{\vec{x},\vec{y}} = *$ with probability q and $s_{\vec{x},\vec{y}} = 0$ with probability $1-q$; then choosing $\rho(p)$ such that $\rho(p) = s_{\vec{x},\vec{y}}$ with probability q and $\rho(p) = 1$ with probability $1-q$.

Define $\mathbb{R}^-_{k,\ell,m}(q)$ by interchanging 0 and 1:

$$
\begin{tikzcd}[column sep=large]
 & 0 & \\
\rho \in \mathbb{R}^{-}_{k,\ell,m}(q)\colon p = p_{\vec{x},\vec{y},y_\ell} \arrow[ur, mapsto, "1-q"] \arrow[r, mapsto, "q"'] & s_{\vec{x},\vec{y}}(\rho) \arrow[r, mapsto, "1-q"] \arrow[dr, mapsto, "q"'] & 1 \\
 & & *
\end{tikzcd}
$$

Let $\rho \in \mathbb{R}^{+}_{k,\ell,m}(q)$. We define $g(\rho)$ by further restricting variables so there is at most one $*$ in each block of variables, and sending that variable to a variable in $B_{k,\ell-1}(m)$. Formally, the transformation $\restriction_{\mathrm{g}(\rho)}$ maps formulas with variables in $B_{k,\ell}(m)$ to formulas with variables in $B_{k,\ell-1}(m)$ by the following procedure:

1. Apply ρ.
2. Assign 1 to every $p_{\vec{x},\vec{y},z}$ with $\rho(p_{\vec{x},\vec{y},z}) = *$ such that there is some $z' < z$ with $\rho(p_{\vec{x},\vec{y},z'}) = *$. I.e., at most one variable in a block remains untouched.
3. Rename each untouched $p_{\vec{x},\vec{y},z}$ to $p_{\vec{x},\vec{y}}$.

For $\rho \in \mathbb{R}^{-}_{k,\ell,m}(q)$, define $\restriction_{\mathrm{g}(\rho)}$ similarly, replacing 1 by 0. See [10] for a proof of the next theorem:

Theorem 17 (Håstad's Switching Lemma [10]). *Let $\nu \in \{+,-\}$. Let φ be a Σ^t_1-formula with variables from $B_{k,\ell}(m)$ and $0 < q < 1$. Then, the probability of a randomly chosen ρ from $\mathbb{R}^{\nu}_{k,\ell,m}(q)$ that the formula $\varphi\restriction_{\mathrm{g}(\rho)}$ is not equivalent to some Π^s_1-formula is at most $(6qt)^s$.*

For the rest of this subsection we assume w.l.o.g. that $\mathrm{S}_{\ell,m}(\vec{x})$ denotes $\mathrm{S}^{\wedge}_{\ell,m}(\vec{x})$. For the following inductive proof, the previously defined Sipser functions $\mathrm{S}_{\ell,m}(\vec{x})$ have to be modified. We define $\overline{\mathrm{S}}_{\ell,m}(\vec{x})$ for every $\vec{x} \in [m]^k$ with variables from $B_{k,\ell}(m)$. They compute modified Sipser functions (cf. [10, 9]) and are defined by

$$
\overline{\mathrm{S}}_{\ell,m}(\vec{x}) \quad = \quad \bigwedge_{y_1<m} \quad \bigvee_{y_2<m} \quad \cdots \quad \underset{y_{\ell-1}<m}{Q^{\ell-1}} \quad \underset{y_\ell<\sqrt{\frac{1}{2}\,(k+\ell+1)\,m\log m}}{Q^{\ell}} \quad p_{\vec{x},\vec{y}}
$$

where either $Q^{\ell-1}$ or Q^{ℓ} is $\bigwedge$, depending on whether ℓ is even or odd, respectively, and the other is $\bigvee$. Note that for distinct $\vec{x}$, the formulas $\overline{\mathrm{S}}_{\ell,m}(\vec{x})$ contain distinct propositional variables.

The first part of the next theorem is also due to Håstad [10], who used it to control the collapse of Sipser functions. Part 1 of the theorem is essentially the same as in Buss and Krajíček [9]. Part 2 is a slight modification useful for Sipser-functions with small bottom fanin. The proofs of both parts follows the same pattern as in [9], so we omit the proofs.

We say that a formula φ contains formula ψ, written as $\psi \subseteq \varphi$, if we can transform φ into ψ by renaming some variables, by removing some 1's from conjunctions and some 0's from disjunctions, and by deleting some $\bigvee$ or some $\bigwedge$ in front of singleton sets, i.e. sets which consist of exactly one element.

Theorem 18. *Let m be sufficiently large, $\ell \in \mathbb{N}$ with $\ell \geq 1$. Let $v(\ell) = +$ or $v(\ell) = -$ if ℓ is odd or even respectively.*

1. *Let $q_1 := \sqrt{\frac{2\,(k+\ell+1)\,\log m}{m}}$ and assume $q_1 \leq \frac{1}{5}$. Then, the probability for a randomly chosen ρ from $\mathbb{R}^{\nu(\ell+1)}_{k,\ell+1,m}(q_1)$ that the formula $\overline{\mathrm{S}}_{\ell+1,m}(\vec{x})\lceil_{\mathrm{g}(\rho)}$ does not contain $\overline{\mathrm{S}}_{\ell,m}(\vec{x})$ is at most $m^{-(k+1)}$.*

2. *Let $q_2 := \frac{k+\ell+1}{\log m}$. Then, the probability for a randomly chosen ρ from $\mathbb{R}^{\nu(\ell+1)}_{k,\ell+1,m}(q_2)$ that the formula $\mathrm{S}_{\ell+0.5,m}(\vec{x})\lceil_{\mathrm{g}(\rho)}$ does not contain $\overline{\mathrm{S}}_{\ell,m}(\vec{x})$ is at most $m^{-(k+1)}$.* $\square$

With the last two theorems, we are ready to start the proof Theorem 16. We start by proving a lemma about how formulas appearing in $(d+1.5)$-LK refutation can be reduced by restrictions.

Lemma 19. *Let $d \in \frac{1}{2}\mathbb{N}$, and further let ϵ, M, $\eta_m \leq m^{\epsilon}$, $\mathcal{A}_m$, $\mathcal{A}'_m$ satisfy the hypotheses of Theorem 16. For $m \in M$, fix a $\Theta^{2^{\eta_m}}_{d+1.5}$-LK refutation of $\mathcal{A}'_m$ and let Γ_m be the smallest set of formulas which includes all formulas occurring in that refutation, and which is closed under subformulas.*

Then for each $r = \lfloor d \rfloor, \ldots, 2, 1, 0$, letting $j = d - r$, there exist transformations κ_j such that

(a) *For every $\varphi \in \Gamma_m$ of depth $(j + \frac{1}{2})$, $\varphi\lceil_{\kappa_j}$ is in $\Delta^{\eta_m}_1$; and*

(b) *For all $\vec{x} \in [m]^k$, $\mathrm{S}_{d+1,m}(\vec{x})\lceil_{\kappa_j}$ includes $\overline{\mathrm{S}}_{d+1-j,m}(\vec{x})$.*

Observe that the cardinality of Γ_m is bounded by $2^{(d+O(1))\cdot\eta_m}$, because all formulas in the refutation are in $\Theta^{2\eta_m}_{d+1.5}$.

Proof. The lemma will be proved separately for $d \in \mathbb{N}$ and for $d \in \mathbb{N} + \frac{1}{2}$. First, assume $d \in \mathbb{N}$. In this case, the relevant values for j are $j = 0, 1, 2, \ldots, d$, and we will prove the existence of κ_j satisfying conditions (a) and (b) by induction on j. The induction basis of $j=0$ is trivial: let κ_0 be the identity transformation.

For the induction step $j \rightsquigarrow j+1$ we first observe that we have at most $2^{c\cdot\eta_m}$ many formulas of depth at most $(j+1.5)$ in Γ_m, with $c = d + O(1)$. Let φ be such a formula. W.l.o.g. it has the form $\varphi = \bigwedge_{i<2^{\eta_m}} \varphi_i$. By the induction hypothesis, each $\varphi_i\!\restriction_{\kappa_j}$ is in $\Delta_1^{\eta_m}$, hence $\varphi\!\restriction_{\kappa_j}$ is in $\Pi_1^{\eta_m}$. Let q be $\sqrt{\frac{2(k+d+1-j)\log m}{m}}$. By Theorem 17, the probability that $\varphi\!\restriction_{\kappa_j}\restriction_{\mathrm{g}(\rho)}$ is not equivalent to some $\Sigma_1^{\eta_m}$-formula (and hence is not in $\Delta_1^{\eta_m}$) for a randomly chosen $\rho \in \mathbb{R}^{\nu}_{k,d+1-j,m}(q)$ is at most $(6q\eta_m)^{\eta_m}$. Hence, the probability that some $\varphi \in \Gamma_m$ of depth at most $(j+1.5)$ does not have $\varphi\!\restriction_{\kappa_j}\restriction_{\mathrm{g}(\rho)}$ in $\Delta_1^{\eta_m}$ is bounded by

$$\begin{aligned} & 2^{c\cdot\eta_m} \cdot \left(6\sqrt{\frac{2(k+d+1-j)\log m}{m}} \cdot m^{\epsilon}\right)^{\eta_m} \\ = & \quad 2^{O(\log\log m)\cdot\eta_m} \cdot 2^{(-0.5+\epsilon)\cdot(\log m)\cdot\eta_m} \end{aligned}$$

which will be arbitrarily small for big m.

On the other hand, for all $\vec{x} \in [m]^k$, $\overline{\mathrm{S}}_{d+1-j,m}(\vec{x}) \subseteq \mathrm{S}_{d+1,m}(\vec{x})\restriction_{\kappa_j}$ by induction hypothesis, and the probability for a randomly chosen $\rho \in \mathbb{R}^{\nu(d+1-j)}_{k,d+1-j,m}(q)$ that $\overline{\mathrm{S}}_{d-j,m}(\vec{x}) \not\subseteq \overline{\mathrm{S}}_{d+1-j,m}(\vec{x})\restriction_{\mathrm{g}(\rho)}$ is at most $m^{-(k+1)}$. Therefore, the probability for a randomly chosen $\rho \in \mathbb{R}^{\nu(d+1-j)}_{k,d+1-j,m}(q)$ that all depth at most $(j+1.5)$ formulas φ in Γ_m are in $\Delta_1^{\eta_m}$ after transforming them according to $\restriction_{\kappa_j}\restriction_{\mathrm{g}(\rho)}$, and that for all $\vec{x} \in [m]^k$, $\overline{\mathrm{S}}_{d-j,m}(\vec{x})$ is included in $\mathrm{S}_{d+1,m}(\vec{x})\restriction_{\kappa_j}\restriction_{\mathrm{g}(\rho)}$, is bigger than 0. Let ρ be such a restriction, and define κ_{j+1} as $\kappa_j\!\restriction_{\mathrm{g}(\rho)}$. That completes the proof of the lemma for $d \in \mathbb{N}$.

Now consider the case where $d = d' + \frac{1}{2} \in \mathbb{N} + \frac{1}{2}$; the relevant values of j are $j = \frac{1}{2}, \frac{3}{2}, \ldots, d$.

This time the induction basis of $j = \frac{1}{2}$ needs some proof. First, we observe that we have at most $2^{c\cdot\eta_m}$ many formulas of depth 1 in Γ_m, with $c = d + O(1)$. Let φ be such a formula. W.l.o.g. it has the form $\varphi = \bigwedge_{i<2^{\eta_m}} \ell_i$ with ℓ_i being literals, i.e., φ is in Π_1^1. Let q be $\frac{k+d+1.5}{\log m}$. By Theorem 17, the probability that $\varphi\!\restriction_{\mathrm{g}(\rho)}$ is not equivalent to some $\Sigma_1^{\eta_m}$-formula (and hence is not in $\Delta_1^{\eta_m}$) for a randomly chosen $\rho \in \mathbb{R}^{\nu}_{k,d+0.5-j,m}(q)$ is at most $(6q1)^{\eta_m}$. Hence, the probability that there is some $\varphi \in \Gamma_m$ of depth 1 such that $\varphi\!\restriction_{\mathrm{g}(\rho)}$ is not in $\Delta_1^{\eta_m}$ is bounded by

$$2^{c\cdot\eta_m} \cdot \left(6 \cdot \frac{k+d+1.5}{\log m}\right)^{\eta_m} = 2^{O(\eta_m)} \cdot 2^{-(\log\log m)\cdot\eta_m}$$

which will be arbitrarily small for big m.

On the other hand, for all $\vec{x} \in [m]^k$, part 2 of Theorem 18 implies that the probability that a randomly chosen $\rho \in \mathbb{R}^{\nu(d)}_{k,d+1.5,m}(q)$ satisfies $\overline{\mathrm{S}}_{d+0.5,m}(\vec{x}) \nsubseteq \mathrm{S}_{d+1,m}(\vec{x})\!\upharpoonright_{\mathrm{g}(\rho)}$ is at most $m^{-(k+1)}$. Therefore, the probability for a randomly chosen $\rho \in \mathbb{R}^{\nu(d)}_{k,d+1.5,m}(q)$ that all depth 1 formulas φ in Γ_m are in $\Delta_1^{\eta_m}$ after transforming them according to $\upharpoonright_{\mathrm{g}(\rho)}$, and that for all $\vec{x} \in [m]^k$, $\overline{\mathrm{S}}_{d+0.5,m}(\vec{x})$ is included in $\mathrm{S}_{d+1,m}(\vec{x})\!\upharpoonright_{\mathrm{g}(\rho)}$, is bigger than 0. Let ρ be such a restriction, and define κ_0 as $\mathrm{g}(\rho)$. That completes the proof of the base case with $j = \frac{1}{2}$. The induction step is proved by the same method used for the case $d \in \mathbb{N}$. □

We now can finish the proof of Theorem 16. Given a refutation R of $\mathcal{A}'_m$, apply Lemma 19 with $r = 0$ and thus $j = d$; this gives a κ_d such that, for every depth $d + \frac{1}{2}$ (sub)formula φ in R, the formula $\varphi\!\upharpoonright_{\kappa_d}$ is in $\Delta_1^{\eta_m}$. Now, every formula φ in R either has depth $\leq d + \frac{1}{2}$, and or is a conjunction or disjunction of formulas of depth $\leq d + \frac{1}{2}$. It follows that each $\varphi\!\upharpoonright_{\kappa_d}$ is either $\Delta_1^{\eta_m}$ or is a conjunction or disjunction of $\Delta_1^{\eta_m}$ formulas.

Theorem 16 is almost done, but we still need to argue that the 0's and 1's can be eliminated from the refutation as restricted by $\upharpoonright_{\kappa_j}$ to get a valid LK refutation. For this, after κ_j is applied, we transform formulas in R repeatedly by the following operations:

(1) Any "1" (resp., "0") in a conjunction (resp., disjunction) is removed.

(2) If any conjunction (resp, disjunction) contains a "0" (resp., a "1"), then replace that conjunction (resp., disjunction) by "0" (resp., "1").

(3) Replace any empty conjunction (resp., disjunction) by "1" (resp., "0").

This transformation of formulas clearly preserves the property of a formula being in $\Sigma_1^{\eta_m}$, in $\Pi_1^{\eta_m}$, or in $\Delta_1^{\eta_m}$.

Cedents in the refutation P are then transformed by (i) removing any "0" appearing in a cedent, and (ii) eliminating any cedent containing a "1".

It is straightforward to prove that the refutation R is transformed by this process into a new refutation R' on the transformed cedents. Every inference in R' corresponds to some inference in R. Thus, the size and height of P' are at most the size and height of P.

Now consider the formula $\mathcal{A}'_m$. By condition (b) of the lemma, its subformulas $S_{d+1}(\vec{x})$ have the property that $S_{d+1}(\vec{x})\!\upharpoonright_{\kappa_d}$ includes $\overline{\mathrm{S}}_{1,m}(\vec{x})$. By further restricting and renaming variables, and further transformations to eliminate 0's and 1's and $\bigvee$ and $\bigwedge$ in front of singleton sets, the formulas

$\overline{\mathrm{S}}_{1,m}(\vec{x})$ are replaced by just the variable $p_{\vec{x}}$. This transforms the refutation into a refutation of $\mathcal{A}_m$ and Theorem 16 is proved. $\square$

4.2 The proof of Theorem 14

Fix $d \in \frac{1}{2}\mathbb{N}$ and $\epsilon < \frac{1}{2}$, and let n be sufficiently large. We shall prove there is no $(d+\frac{1}{2})$-LK refutation of $\neg\,\mathrm{OP}^d(n)$ of tree-size $S < 2^{n^\epsilon}$. Suppose, for sake of contradiction, that such a refutation exists. Corollary 9 then implies that $\neg\,\mathrm{OP}^d(n)$ has $(d+1.5)$-LK refutation of height $< \eta$ where $\eta = c \cdot n^\epsilon$ for some constant $c > 0$. The clauses $\neg\,\mathrm{OP}^d(n)$ are equivalent to

$$\neg\,\mathrm{OP}(n)\Big[p_{x,y} \leftarrow \mathrm{S}^{\vee}_{d+1,n}(x,y) : x,y \in [n]\Big].$$

By applying Theorem 16, it follows that $\neg\,\mathrm{OP}(n)$ has an LK refutation of height $< \eta$ in which every proper subformula is in Δ_1^η, and hence every formula is in $\Sigma_1^\eta \cup \Pi_1^\eta$. We shall prove below that in this case $n \le 2\eta^2$; this is a contradiction since $\eta = O(n^\epsilon)$ and thus will complete the proof of Theorem 14.

In the following let φ always be an LK-formula with variables $p_{x,y}$ in $B_{2,0}(n)$. The intent is that the variables $p_{x,y}$ define a binary relation on $[n]$, in particular, a total order. Fix a subset D of $[n]$, and let $\prec$ be a total order on D. Note $\prec$ determines D if $|D| \ge 2$, since $D = \mathit{domain}(\prec)$. A truth assignment is said to extend $\prec$ provided that whenever $x,y \in D$, then $\tau(p_{x,y}) = \mathit{True}$ iff $x \prec y$. A truth assignment is called a *total order* if it extends (or, is) some total order on all of $[n]$. We say that $\prec$ *fixes* φ *to false* if every truth assignment which is a total order and extends $\prec$ assigns the value *False* to φ.

A basic property of Π_1^t formulas is that if there is a total ordering truth assignment that assigns a Π_1^t formula φ the value *False*, then there is some small domain ordering $\prec$ that fixes φ to false.

Lemma 20. *Suppose that φ is a Π_1^t formula, and $\prec_1$ is a total ordering on D_1. Further suppose there is a truth assignment τ which is a total order and extends $\prec_1$, and which gives φ the value* False. *Then there is a total order $\prec_2$ on domain D_2 such that (a) $\prec_2 \supseteq \prec_1$, and (b) $|D_2 \setminus D_1| \le 2t$, and (c) $\prec_2$ fixes φ to false.*

Proof. Let φ be equivalent to $\bigwedge_{i<s}\bigvee_{j<t} z_{i,j}$ for some s and literals $z_{i,j}$. If τ makes φ false, then it makes $\bigvee_{j<t} z_{i_0,j}$ false, for some i_0. Set D_2 to be D_1 plus all values $x \in [n]$ such that $p_{x,y}$, $\overline{p}_{x,y}$, $p_{y,x}$ or $\overline{p}_{y,x}$ occurs among

the literals $z_{i_0,j}$. Define $\prec_2$ to be the ordering defined by τ restricted to the domain D_2. It is clear that $\prec_2$ satisfies the conditions of the lemma. $\square$

Theorem 14 now follows from the following boundedness theorem.

Theorem 21 (Boundedness). *Suppose* $\neg\,\mathrm{OP}(n)$ *has an* LK*-refutation* R *of height* η *in which every occurring formula is in* $\Sigma_1^t \cup \Pi_1^t$ *and every proper subformula in the proof is in* Δ_1^t. *Then* $n \leq 2 \cdot \eta \cdot t$.

Proof. We shall find cedents Δ_i in R for $i = 0, 1, 2, \ldots, k$, and total orders $\prec_i$ with domains $D_i \subset [n]$. The first cedent Δ_0 will be the end cedent of R, i.e., the empty cedent. Likewise, D_0 is the empty set, so $\prec_0$ is the total ordering of the empty set. Each cedent Δ_{i+1} will be one of the hypotheses to the inference that derives Δ_i. Each D_i will have cardinality $\leq 2 \cdot i \cdot t$. Furthermore, each $\prec_i$ will fix the cedent Δ_i to be false. The process stops when $i = k$ where Δ_k is an initial cedent of R, i.e, Δ_k is a clause from $\neg\,\mathrm{OP}(n)$.

By examination, the only way an initial cedent Δ_k from $\neg\,\mathrm{OP}(n)$ can be fixed to false by a total ordering $\prec_k$ is for the clause to be one of the M_x clauses and for the domain D_k of $\prec_k$ to equal all of $[n]$, from whence, $n \leq 2 \cdot k \cdot t$. Thus, if the clauses Δ_k can be constructed, then n must be less than $2t$ times the maximum number of cedents along any path in the refutation.

To complete the proof, it remains to show how to define Δ_{i+1} from Δ_i. The proof splits into cases depending on the type of inference used to derive Δ_i. If the inference that derives Δ_i is a cut inference,

$$\frac{\Delta_i, \neg\varphi \qquad \Delta_i, \varphi}{\Delta_i}$$

with a Σ_1^t-formula φ, then we consider the following two subcases: First assume that all total orderings $\prec$ on $[n]$ which extend $\prec_i$ do not satisfy φ. In this case, let $\prec_{i+1}$ be the same as $\prec_i$, and let Δ_{i+1} be the right upper cedent Δ_i, φ. Otherwise, by Lemma 20, there exists a total ordering $\prec_{i+1}$ which extends $\prec_i$, has domain D_{i+1} with $|D_{i+1} \setminus D_i| \leq 2t$, and fixes $\neg\varphi$ to false. In this case, let Δ_{i+1} be the cedent $\Delta_i, \neg\varphi$.

If the inference that derives Δ_i is a $\bigvee$ inference,

$$\frac{\Delta_i', \varphi}{\Delta_i', \bigvee\Phi}$$

where $\varphi \in \Phi$, then let Δ_{i+1} be the upper cedent, and let $\prec_{i+1}$ be $\prec_i$. This works since any truth assignment that falsifies the lower cedent also falsifies the upper cedent.

Assume now that the inference that derives Δ_i is a $\bigwedge$ inference,

$$\frac{\Delta_i', \varphi \qquad \text{for all } \varphi \in \Phi}{\Delta_i', \bigwedge\Phi}$$

By assumption, all formulas in Φ are in Δ_1^t. Take any truth assignment τ which corresponds to a total order on $[n]$ extending $\prec_i$. By the induction hypothesis, it does not satisfy the lower cedent, so there must be some $\varphi \in \Phi$ which is not satisfied by τ. By Lemma 20 again, there is a $\prec_{i+1}$ such that $\prec_i \subseteq \prec_{i+1}$ and such that the domain D_{i+1} has at most $2t$ new elements and such that $\prec_{i+1}$ fixes φ to false. Then let Δ_{i+1} be Δ_i', φ and this case is finished.

The case of weakening is even easier and we omit its proof. □

Careful examination of the proof shows we actually proved a strengthened form of Theorem 14. Since we used only total orderings $\prec_i$, we could also allow as initial clauses $\{p_{x,y}, p_{y,x}\}$ for distinct $x, y \in [n]$ without affecting the validity of the theorem. These clauses make the ordering total rather than just partial. Thus, the lower bounds we have obtained on constant depth proofs of the ordering principles also apply to the constant depth proofs of the "total ordering principle."

5 Exponential lower bound for a Ramsey principle

This section describes a method which lifts principles requiring exponential tree-size $(d + \frac{1}{2})$-LK-refutations for $d \in \mathbb{N}$ to principles requiring exponential sequence-size d-LK-refutations. The basic proof technique is to replace variables in the principle by small Sipser functions of depth two and then distributing to reduce formula depth. Rather than explore this method in full generality, we will describe this lifting for the Ramsey principle and $d = 0$.

The Ramsey principle has been shown by Pudlák [21] to have polynomial size Frege proofs, and Krajíček [14] has shown that it requires exponential tree-size depth $\frac{1}{2}$ LK-proofs. In addition, Krajíček showed that if the weak pigeonhole principle $WPHP_n^{n^4}$ requires exponential sequence-size *Res*(2) proofs, then the Ramsey principle requires exponential sequence-size for depth 0 LK proof (i.e., requires exponential sequence-size resolution proofs). The system *Res*(2) is resolution extended to have clauses containing terms of two literals. Although several sets of researchers have found lower bounds on the sequence-size of proofs $Res(f(n))$ of weak pigeonhole principles for $f(n)$ as large as $\epsilon \log n / \log\log n$, none of them apply to the $WPHP_n^{n^4}$ principle (see [2], [24], and especially [23]). Thus, it is still open whether the Ramsey principle has depth 0 LK proofs of subexponential sequence-size.

Theorem 22 below establishes exponential sequence-size lower bounds on depth 0 LK proofs of a Sipser-ized version of the Ramsey principle. A yet stronger lower bound is obtained in Theorem 23.

We start by defining the Ramsey tautologies. Let X be a finite set. $[X]^2$ denotes the set of all possible edges of a graph on X; $[X]^2 = \{ e \subseteq X \,;\, |e| = 2 \}$. We write $[n]^2$ instead of $[[n]]^2$, so that for this section, $[n]^2$ denotes the set of edges on the vertices $[n]$ rather than $[n] \times [n]$. For $e \in [n]^2$ let p^e be an edge variable. The *Ramsey principle* states that every undirected graph on n vertices contains a homogeneous set of size $\lceil \frac{\log n}{2} \rceil$, where a homogeneous set is either a clique or an independent set. The negation of the Ramsey principle can be written as follows:

$$\bigwedge_{X \subseteq [n], |X| = \lceil \frac{\log n}{2} \rceil} \Big(\bigvee_{e \in [X]^2} \overline{p}^e \wedge \bigvee_{e \in [X]^2} p^e \Big).$$

This is equivalently expressed by the following clauses, for every $X \subseteq [n]$ with $|X| = \lceil \frac{\log n}{2} \rceil$:

C_X	$\{\overline{p}^e : e \in [X]^2\}$	<u>C</u>lique clause
I_X	$\{p^e : e \in [X]^2\}$	<u>I</u>ndependent set clause.

Let $\overline{n}$ be $(\log n)^2$. A modified Ramsey principle, $\neg\,\overline{\mathrm{Ram}}(n)$, is defined by replacing each p^e by the "small" depth 2 Sipser function $\mathrm{S}^{\wedge}_{2,\overline{n}}(e) = \bigwedge_{i<\overline{n}} \bigvee_{j<\overline{n}} p^e_{i,j}$ and then distributing the connectives out:

$$\bigwedge_{X \subseteq [n], |X| = \lceil \frac{\log n}{2} \rceil} \Big(\bigwedge_{f : [X]^2 \times [\overline{n}] \to [\overline{n}]} \bigvee_{e \in [X]^2} \bigvee_{i \in [\overline{n}]} \overline{p}^e_{i,f(e,i)} \wedge \bigwedge_{g : [X]^2 \to [\overline{n}]} \bigvee_{e \in [X]^2} \bigvee_{j \in [\overline{n}]} p^e_{g(e),j} \Big).$$

This can be equivalently expressed by the following set of clauses, where $X \subseteq [n]$ with $|X| = \lceil \frac{\log n}{2} \rceil$ and $f \colon [X]^2 \times [\overline{n}] \to [\overline{n}]$ and $g \colon [X]^2 \to [\overline{n}]$:

$C_{X,f}$	$\{\overline{p}^e_{0,f(e,0)}, \ldots, \overline{p}^e_{\overline{n}-1,f(e,\overline{n}-1)} : e \in [X]^2\}$
$I_{X,g}$	$\{p^e_{g(e),0}, \ldots, p^e_{g(e),\overline{n}-1} : e \in [X]^2\}$.

These clauses constitute the set $\neg\,\overline{\mathrm{Ram}}(n)$. The size of each C- or I-clause is bounded by $(\log n)^4$, and there are $2^{O((\log n)^5)}$ many C clauses and $2^{O((\log n)^3)}$ many I clauses.

Theorem 22. *For n sufficiently large, if $\neg\overline{\mathrm{Ram}}(n)$ has a resolution refutation of sequence-size S, then S must be bigger than 2^{n^ϵ} for any ϵ with $0 < \epsilon < \frac{1}{4}$.*

Proof. The proof will use a variant of LK where all rules have at most two premises. This variant is called *binary* LK and is denoted LK^2. The binary LK system is defined like LK, but with the $\bigwedge$ inference rule replaced by

$$(\bigwedge{}^2)\,\frac{\Gamma, \bigwedge\Phi_1 \qquad \Gamma, \bigwedge\Phi_2}{\Gamma, \bigwedge(\Phi_1 \cup \Phi_2)}\ ,$$

where, for the purposes of this rule, we identify φ and $\bigwedge\{\varphi\}$. The following simulation is straightforward.

> If Γ has a Θ_d^σ-LK-derivation from $\mathcal{A}$ of cedent-size σ, then Γ has a Θ_d^σ-LK^2-derivation from $\mathcal{A}$ of cedent-size σ^2.

This simulation holds since any $\bigwedge$ inference

$$(\bigwedge)\,\frac{\Gamma, \varphi \qquad \text{for all } \varphi \in \Phi}{\Gamma, \bigwedge\Phi}$$

in a Θ_d^σ-LK-derivation has $|\Phi| \leq \sigma$ and therefore can be replaced by a binary tree of $\bigwedge^2$-inferences of height $\log\sigma$.

To begin the proof of Theorem 22, assume, for the sake of a contradiction, that there is some $0 < \epsilon < \frac{1}{4}$ and some large $n \in \mathbb{N}$ such that $\neg\overline{\mathrm{Ram}}(n)$ has a resolution refutation of sequence-size $\leq 2^{n^\epsilon}$. Using Lemma 4 and Lemma 5 to transform to a tree-size proof and converting the resulting LK refutation into an LK^2 refutation, we obtain that $\neg\overline{\mathrm{Ram}}(n)$ has a $\Theta_1^{2\eta}$-LK^2-refutation R of tree-cedent-size $\leq 2^\eta$, where $\eta = O(n^\epsilon)$.

We want to reduce the complexity of formulas in R by applying the switching lemma. Let $q = \frac{4}{\log n}$ and recall that $\overline{n} = (\log n)^2$. We consider the restrictions $\rho \in \mathbb{R}^-(q) := \mathbb{R}^-_{2,2,n}(q)$ and how they affect Σ_1^η-formulas and the Sipser-functions $\mathrm{S}_{2,\overline{n}}(e) := \mathrm{S}^\wedge_{2,\overline{n}}(e)$ for $e \in [n]^2$. The transformation $\restriction_{\mathrm{g}(\rho)}$ is defined to act as follows:

1. Apply the restriction ρ.
2. Assign 0 to every $p^e_{y,z}$ with $\rho(p^e_{y,z}) = *$ such that there is some $z' < z$ with $\rho(p^e_{y,z'}) = *$. This leaves at most one variable in a block untouched.
3. Rename each untouched $p^e_{y,z}$ to p^e.

We claim that, for all $e \in [n]^2$,

$$\mathrm{Pr}_{\rho\in\mathbb{R}^-(q)}\left[\mathrm{S}_{2,\overline{n}}(e)\lceil_{\mathrm{g}(\rho)} \text{ is not equivalent to } p^e\right] \leq n^{-3} \tag{4}$$

for n sufficient large. To prove this, first note that each $\bigvee$ of $\mathrm{S}_{2,\overline{n}}(e)$ is assigned the value 0 or the value $s_{e,y}(\rho) \in \{*, 1\}$ by a restriction $\rho \in \mathbb{R}^-(q)$. It receives the value 0 with probability at most

$$(1-q)^{\overline{n}} < e^{-q\overline{n}} = e^{-4\log n} < \frac{1}{2}n^{-4}.$$

Since there are $\overline{n}$ many $\bigvee$'s in $\mathrm{S}_{2,\overline{n}}(e)$, the probability that one or more are assigned the value 0 is less than $\frac{1}{2}n^{-4}\overline{n} < \frac{1}{2}n^{-3}$ for sufficiently large n.

If none of the $\bigvee$'s take value 0, then they all take value $s_{e,y}(\rho)$. Each $s_{e,y}(\rho)$ takes value 1 or $*$ under $\mathbb{R}^-(q)$. The probability that they all receive value 1 is at most $(1-q)^{\overline{n}} < \frac{1}{2}n^{-3}$. It follows that the overall probability that $\mathrm{S}_{2,\overline{n}}(e)$ is not transformed by $\lceil_{\mathrm{g}(\rho)}$ into p^e is bounded above by n^{-3}.

Now we consider how a $\rho \in \mathbb{R}^-(q)$ reduces formula complexity in R. Consider any formula φ in R of depth 1. It is, of course, either an $s\Sigma_1^1$-formula or an $s\Pi_1^1$-formula. Taking $t=1$ and $s=\eta$, and applying the Håstad switching lemma, shows that $\varphi\lceil_{\mathrm{g}(\rho)}$ is equivalent to some Π_1^s-formula or to some Σ_1^s-formula (respectively) with probability greater than $1-(6q)^\eta$. Actually, since $t=1$, the proof of the Håstad switching lemma [10] shows even more; namely, $\varphi\lceil_{\mathrm{g}(\rho)}$ is either a disjunction (or, a conjunction, respectively) of at most η literals with probability $\geq 1-(6q)^\eta$. (Remark: it is also simple to prove this fact from the statement of the switching lemma, and thus it is not really necessary to refer back to the proof of the switching lemma.) Since every formula in the refutation is in $\Theta_1^{2^\eta}$ and since there are at most 2^η different depth 1 formulas in the proof, the event that there is a formula of depth 1 which is not switched to a conjunction or disjunction of size $\leq \eta$ occurs with probability less than

$$2^\eta(6q)^\eta \;=\; 2^\eta\left(\frac{24}{\log n}\right)^\eta \;=\; o(1).$$

Thus, with probability approaching 1, applying the transformation $g(\rho)$ turns R into a refutation R' of the original Ramsey principle, $\neg\,\mathrm{Ram}(n)$, such that every depth 1 formula in R' is a conjunction or disjunction of at most η literals. Therefore, every formula in R' is either a literal or a conjunction or disjunction of at most η literals. This proof is essentially an $R^*(\log)$ proof in the sense of Krajíček, and by Krajíček [14, Thm 5.2], this is possible only if $\eta \geq c\cdot n^{1/4}$ for some constant c, i.e., only if $\epsilon \geq \frac{1}{4}$. ☐

Theorem 22 can be strengthened as follows.

Theorem 23. *Fix $\epsilon < \frac{1}{4}$ and $\delta < \frac{1}{48}$. Let n be sufficiently large. Let $t = \delta \log n$. Then there is no* LK *refutation of $\neg\,\overline{\mathrm{Ram}}(n)$ of sequence-size 2^{n^ϵ} in which every formula is a disjunction or conjunction of at most t literals.*

Proof. (Sketch.) Theorem 23 is proved by almost the same proof as Theorem 22. Assume, for sake of a contradiction, that there is such a refutation. By Lemma 4 and Lemma 5 and the conversion of LK refutations into LK^2 refutations, there is an LK^2 refutation R of $\neg\,\overline{\mathrm{Ram}}(n)$ of tree-cedent-size 2^η with $\eta = O(n^\epsilon)$ in which every formula has depth two or less. The connectives (conjunctions or disjunctions) at the top level have fanin at most 2^{n^ϵ}, and the connectives at the second level have fanin at most t.

Let $q = 4/\log n$ as before and choose the restriction $\rho \in \mathbb{R}^-(q)$ at random. Again take $s = \eta$, but now use $t = \delta \log n$. Then, with high probability, we obtain a refutation R' of $\neg\,\mathrm{Ram}(n)$ such that every formula in R' is in Δ_1^η. The rest of the proof is as before, based on a strengthened version of Theorem 5.2 of [14]. □

6 Width lower bounds

In this section we prove that lower bounds on the size of tautologies with Sipser functions substituted in for variables can imply lower bounds on the width of resolution refutations of the tautologies. Rather than prove this in full generality, we illustrate the technique by proving a lower bound on the width of resolution proofs of the Ramsey principle. This lower bound has already been obtained by Krajíček [14]: the novel part is that it follows directly from Theorem 22. Together with the result from the last section this gives a general method to deduce width lower bounds for resolution refutations from size lower bounds of $\frac{1}{2}$-LK-refutations.

Let $\neg\,\mathrm{Ram}(n)$ be the set of clauses expressing the negation of the Ramsey principle. Let $\overline{n} = (\log n)^2$, and let $\neg\,\overline{\mathrm{Ram}}(n)$ be as defined earlier.

The *width* of a resolution refutation is the maximum number of literals in any clause in the refutation. In order to obtain optimal results with respect to the width of a resolution refutation, for this section we replace LK's cut rule by the more common resolution rule:

$$\text{Res:}\ \frac{\Gamma_1, \neg p \qquad \Gamma_2, p}{\Gamma_1, \Gamma_2}$$

Resolution based on the resolution rule versus resolution based on the cut rule (plus weakening) are polynomially equivalent in terms of size, but the widths may differ strongly.

Lemma 24. *Suppose* $\neg\,\mathrm{Ram}(n)$ *has a resolution refutation of sequence-size* S *and width* w*. Then* $\neg\,\overline{\mathrm{Ram}}(n)$ *has a resolution refutation of sequence-size* $\leq S \cdot 2^{O(w(\log n)^2 \log\log n)}$*.*

Proof. (Sketch) A resolution refutation R of $\neg\,\mathrm{Ram}(n)$ can be converted into a resolution refutation of $\neg\,\overline{\mathrm{Ram}}(n)$ by the following procedure: Replace each p^e by $\mathrm{S}^{\wedge}_{2,\overline{n}}(e)$ in all the clauses in the refutation. Each clause C now contains at most w formulas of depth ≤ 2, and the conjunctions and disjunctions in these formulas have fanin $\overline{n}$. Viewing the clause as a formula F_C which is a disjunction of depth 2 formulas, apply the distributive law to the *top* disjunctions and conjunctions in F_C. This converts the clause into a conjunction of disjunctions (a CNF formula). The conjunction has fanin at most $\overline{n}^{w\overline{n}}$ (since there are at most $w\overline{n}$ conjunctions in F_C and since each conjunction has fanin $\overline{n}$. Then convert this CNF into a set of clauses. This process converted a clause C of R into at most $\overline{n}^{w\overline{n}} = (\log n)^{2w(\log n)^2}$ clauses.

We leave it to the reader to verify that it is possible to build a valid resolution refutation of $\overline{\mathrm{Ram}}(n)$ from these converted clauses, adding additional clauses to fill the gaps. This increases the number of clauses by a factor of at most $\overline{n}^{2\overline{n}} = (\log n)^{4(\log n)^2}$ (this factor is a not-quite-optimal upper bound). $\square$

Lemma 25. *If* $\neg\,\mathrm{Ram}(n)$ *has a resolution refutation of width* w*, then that refutation has sequence-size* $< n^{2w}$*.*

Proof. This is a simple consequence of the fact that there are $< n^2$ many literals for the $\neg\,\mathrm{Ram}(n)$ principle, and thus less than n^{2w} distinct clauses that can appear in a resolution refutation. $\square$

Theorem 26. *For any* $\epsilon < \frac{1}{4}$*, the* $\neg\,\mathrm{Ram}(n)$ *principles do not have resolution refutations of width* $< n^{\epsilon}$ *for sufficiently large* n*.*

Proof. This is an immediate consequence of Lemmas 24 and 25 and Theorem 22. $\square$

Using Theorem 23 instead of Theorem 22 we obtain a strengthening of the last Theorem.

Theorem 27. *Fix $\epsilon < \frac{1}{4}$ and $\delta < \frac{1}{48}$. Let n be sufficiently large. Let $t = \delta \log n$. Then there is no* LK *refutation of $\neg\,\mathrm{Ram}(n)$ of width $< n^\epsilon$ in which every formula is a disjunction or conjunction of at most t literals.*

References

[1] M. Ajtai. The complexity of the pigeonhole principle. In *Proceedings of the 29-th Annual IEEE Symposium on Foundations of Computer Science*, pages 346–355, 1988.

[2] Albert Atersias, Maria Luisa Bonet, and J.L. Esteban. Lower bounds for the weak pigeonhole principle and random formulas beyond resolution. *Information and Computation*, 176(2):27–39, 2002.

[3] Paul Beame, Russell Impagliazzo, Jan Krajíček, Toniann Pitassi, and Pavel Pudlák. Lower bounds on Hilbert's Nullstellensatz and propositional proofs. *Proceedings of the London Mathematical Society*, 73:1–26, 1996.

[4] Paul Beame and Toniann Pitassi. An exponential separation between the parity principle and the pigeonhole principle. *Annals of Pure and Applied Logic*, 80(3):195–228, 1996.

[5] Arnold Beckmann. *Separating Fragments of Bounded Arithmetic*. PhD thesis, Westfälishe Wilhelms Universität Münster, 1996.

[6] Arnold Beckmann. Dynamic ordinal analysis. *Archive for Mathematical Logic*, 42:303–334, 2003.

[7] Arnold Beckmann. Height restricted constant depth LK. Technical Report TR03-034, Electronic Colloquium in Computational Complexity (ECCC), 2003.

[8] Maria Luisa Bonet and Nicola Galesi. A study of proof search algorithms for resolution and polynomial calculus. In *40th Annual IEEE Symp. on Foundations of Computer Science*, pages 422–431. IEEE Computer Society, 1999.

[9] Samuel R. Buss and Jan Krajíček. An application of Boolean complexity to separation problems in bounded arithmetic. *Proc. London Math. Society*, 69:1–21, 1994.

[10] Johan Hastad. *Almost Optimal Lower Bounds for Small Depth Circuits*, volume 5 of *Advances in Computing Research*, pages 143–170. JAI Press, 1989.

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

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

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

[14] Jan Krajíček. On the weak pigeonhole principle. *Fundamenta Mathematica*, 170:123–140, 2001.

[15] Jan Krajíček and Pavel Pudlák. Quantified propositional calculi and fragments of bounded arithmetic. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 36:29–46, 1990.

[16] Jan Krajíček, Pavel Pudlák, and Alan Woods. Exponential lower bound to the size of bounded depth Frege proofs of the pigeonhole principle. *Random Structures and Algorithms*, 7:15–39, 1995.

[17] Balakrishnan Krishnamurthy. Short proofs for tricky formulas. *Acta Informatica*, 22(3):253–275, 1985.

[18] Alexis Maciel, Toniann Pitassi, and Alan R. Woods. A new proof of the weak pigeonhole principle. *Journal of Computer and System Sciences*, 64(4):843–872, 2002.

[19] J. B. Paris and A. J. Wilkie. Δ_0 sets and induction. In W. Guzicki, W. Marek, A. Pelc, and C. Rauszer, editors, *Open Days in Model Theory and Set Theory*, pages 237–248, 1981.

[20] Toniann Pitassi, Paul Beame, and Russell Impagliazzo. Exponential lower bounds for the pigeonhole principle. *Computational Complexity*, 3:97–140, 1993.

[21] Pavel Pudlák. Ramsey's theorem in bounded arithmetic. In *Computer Science Logic, Lecture Notes in Computer Science #553*, pages 308–312. Springer-Verlag, 1992.

[22] Alexander A. Razborov. On provably disjoint NP-pairs. Technical Report RS-94-36, Basic Research in Computer Science Center, Aarhus, Denmark, November 1994. http://www.brics.dk/index.html.

[23] Alexander A. Razborov. Pseudorandom generators hard for k-DNF resolution and polynomial calculus resolution. Manuscript, 2003.

[24] Nathan Segerlind, Samuel R. Buss, and Russell Impagliazzo. A switching lemma for small restrictions and lower bounds for k-DNF resolution. In *Proc. 43rd IEEE Symp. of Foundations of Computer Science (FOCS'02)*, pages 604–613, 2002.

[25] Gunnar Stålmarck. Short resolution proofs for a sequence of tricky formulas. *Acta Informatica*, 33(3):277–280, 1996.