

# Proof Complexity of Relativized Statements

Albert Atserias\*

Universitat Politècnica de Catalunya  
Barcelona, Spain

Moritz Müller

Kurt Gödel Research Center for Mathematical Logic  
Vienna, Austria

Sergi Oliva†

Universitat Politècnica de Catalunya  
Barcelona, Spain

January 17, 2012

## Abstract

The first-order formulas that do not have finite models give rise to uniform families of propositional unsatisfiable formulas, one for each finite cardinality. Motivated by the question of characterizing the class of such formulas whose propositional translations have short resolution refutations, we study their closure properties. Our starting point is the observation that the class of relativized formulas that have polynomial-length refutations in  $R(c)$  is closed under quantifier-free interpretations. The analogous question for unrelativized formulas remains open and has an intriguing status relating to the existence of formulas without finite models but with well-behaved infinite models. Then we characterize the class of relativized formulas without finite models whose binary-translation into propositional logic have polynomial-length refutations in  $R(c)$ . We show that they agree with those that have no models at all, thus deviating from the unary-translation for which the analogous result is known to fail. Our lower bound generalizes some previously known lower bounds for  $R(c)$  to the minimal condition of having some infinite model.

## 1 Introduction

Understanding what makes a propositional contradiction hard to refute is an important problem, even for concrete refutation systems. From a practical side, for example, such an understanding is a potential source for improvements on satisfiability algorithms, and in this respect resolution-based refutation systems are of special interest. By this we mean, besides resolution, also more robust extensions of it, such as the  $R(k)$  refutation systems that operate on general  $k$ -DNF formulas instead of clauses (i.e. 1-DNF formulas). On the other hand, contradictions of special interest

---

\*Research partially supported by CICYT TIN2010-20967-C04-04 (TASSAT).

†Research partially supported by CICYT TIN- (FORMALISM).

are those expressing general combinatorial principles, of which the pigeonhole principle is a prime example, with the hope that these may appear in many diverse application contexts.

More formally, given a principle formalized as a first-order sentence  $\phi$  and a positive natural number  $n$ , it is straightforward to write down a propositional formula  $\langle\phi\rangle_n$  that is satisfiable if and only if  $\phi$  has a model of size  $n$ , i.e.  $n$  is in the spectrum of  $\phi$ . The question central to this paper is to understand the refutation complexity of the formulas  $\langle\phi\rangle_n$ , for  $n$  outside the spectrum of  $\phi$ . Assuming  $\text{NE} \neq \text{coNE}$ , one can show that there exists a single first-order sentence  $\phi$  such that no refutation system at all has short refutations of all  $\langle\phi\rangle_n$ s with  $n$  outside the spectrum of  $\phi$ . Abstractly, a refutation system is defined to be a polynomial-time computable surjection onto the set of propositional contradictions (cf. [5]).

In general, the refutation complexity of  $\langle\phi\rangle_n$  depends on the choice of the translation  $\langle\cdot\rangle_n$ . A very natural such translation, that we call the *unary encoding* and denote  $\langle\cdot\rangle_n^u$ , produces formulas in CNF provided it is applied to a universal formula in CNF without nested function symbols. Note that, by Skolemization, for every first-order sentence there is such a sentence with the same spectrum. Riis' celebrated gap theorem [17], partially improving on earlier results of Krajčček (cf. [12, 13]), completely answers our central question for the important case of tree-like resolution: for sufficiently large  $n$ , either  $\langle\phi\rangle_n^u$  has polynomial size refutations in tree-like resolution, or needs exponential size (i.e. size  $2^{\Omega(n)}$ ); moreover, the latter happens precisely when  $\phi$  has an infinite model.

What about stronger systems? Riis' gap extends to tree-like  $\text{R}(c)$  (cf. [8, 13]): if  $\phi$  has an infinite model, then not only tree-like  $\text{R}(1)$  requires exponential refutations, but even tree-like  $\text{R}(k)$  does for any constant  $k$ . However, the gap cannot be extended to dag-like resolution which is known to be stronger than every tree-like  $\text{R}(k)$  (cf. [10, 9]). A counterexample is the Least Number Principle (LNP) that has an infinite model but nevertheless its translations have polynomial size dag-like resolution refutations.

A further step has been taken by Dantchev and Riis [8] who showed that if a *relativized* sentence has an infinite model, then even dag-like resolution needs exponential size (i.e. size  $2^{n^{\Omega(1)}}$ ) to refute the corresponding contradictions (this answered a question from [13]). Roughly, Dantchev and Riis' argument runs as follows: on one hand they use the relativizing predicate to define suitable random restrictions that probably kill clauses of large "width" for a carefully defined notion of width. On the other hand, the infinite model enables an adversary argument (cf. [15]) to establish a width lower bound for dag-like resolution proofs. This implies a size lower bound on proofs in the usual manner.

**Our results** The argument by Dantchev and Riis does not work for dag-like  $\text{R}(k)$  where  $k \geq 2$ . Indeed, recently Dantchev and Martin [7] (see also [6]) proved that  $\text{R}(c)$  has short refutations for every sentence that allows a quantifier-free interpretation of the LNP. This implies that the relativized LNP has short  $\text{R}(c)$  refutations and puts forward the question of characterizing the principles that have short dag-like  $\text{R}(c)$ -refutations. Towards this goal, our first contribution is to show that the class of relativized principles that have short proofs in  $\text{R}(c)$  is closed under quantifier-free interpretations. For unrelativized principles, the corresponding statement remains open and has an intriguing status relating to the existence of principles with short  $\text{R}(1)$  proofs that do not interpret LNP.

As witnessed by the relativized LNP, Dantchev and Riis' [8] gap cannot be extended to  $\text{R}(c)$ . We do so nevertheless by considering a different natural translation. We introduce the so-called

binary encoding  $\langle \cdot \rangle_n^b$  and prove a gap theorem for relativized formulas and (dag-like)  $R(c)$ . More precisely, our main result reads as follows:

**Theorem 1 (Main).** *Let  $\phi$  be a standardized universal first-order sentence.*

- (a) *Assume  $\phi$  does not have infinite models. Then there is  $d \geq 1$  such that for every large enough  $n$ , there is an  $R(1)$ -refutation of  $\langle \phi^R \rangle_n^b$  of length at most  $n^d$ .*
- (b) *Assume  $\phi$  has infinite models. Then for every  $k \geq 1$  there is  $d \geq 1$  such that for every large enough  $n$ , every  $R(k)$ -refutation of  $\langle \phi^R \rangle_n^b$  has length at least  $2^{n^{1/d}}$ .*

For sentences in a unary vocabulary the lower bound holds for  $k$  up to  $o(\sqrt{\log n})$  thus matching the best known lower bound on  $k$  from [18] (see below). For general vocabularies we are able to exclude quasipolynomial size proofs for values of  $k$  up to  $\epsilon \cdot \log n$ . We also prove that our gap does not extend to  $R(\log)$ : this system has polynomial size refutations of  $\langle \text{LNP}^R \rangle_n^b$ .

**Comparison with related work** Previously, no such general lower bound, say by a model-theoretic criterion as in [17, 8, 13] for tree-like  $R(k)$  and in [8] for dag-like resolution, has been known for (dag-like)  $R(k)$ . However, strong lower bounds for the  $R(k)$  systems have been found earlier. The first appeared in [1] and states an exponential lower bound for the  $2n$  to  $n$  weak pigeonhole principle in  $R(2)$ . This has been improved by Segerlind et al. [18] to  $k = \sqrt{\log n / \log \log n}$ , and later by Razborov [20] even to  $k = \epsilon \cdot \log n / \log \log n$ .

Roughly, Segerlind et al.’s [18] argument runs as follows: they show that random restrictions that probably kill  $k$ -DNFs of large “covering number”, also probably simplify arbitrary  $k$ -DNFs to formulas with a shallow decision tree; in a second step they show how  $R(k)$  proofs, whose lines are computed by shallow decision trees, can be translated to small width (in the normal sense) resolution proofs – and hence one is back in Ben-Sasson and Wigderson’s [3] setting to establish width lower bounds. For example, to establish the mentioned lower bound for the weak pigeonhole principle, it is restricted to suitable small degree expander graphs where a width lower bound is known. As a second example, [18] contains an exponential lower bound on  $R(k)$  refutations of a version of the LNP. The needed width lower bound for resolution is again established by restricting the principle to graphs “with a certain expansion-like property” [18, Section 8.3]. In fact, the principle witnesses an exponential separation of  $R(k)$  and  $R(k+1)$ . Dantchev found a more natural such version of the restricted ordering principle that also serves as such a witness [6]. Roughly, while in [18] atoms are replaced in the restricted LNP by certain  $k$ -terms, in [6] it is relativized for  $k$  times. Both proofs are tied to special properties of the LNP.

Our binary encoding makes the methods from [8] and [18] nicely combine: we show that Dantchev and Riis’ random restrictions [8] probably kill  $k$ -DNFs of large “dimension”, and prove a switching lemma as in [18] where the height of the decision tree is replaced by a suitable notion of “rank”: the random restrictions from [8] probably transform  $k$ -DNFs to formulas with small rank decision trees. Instead of the second step in [18], we generalize the width lower bound for resolution from [8], and establish, also using some kind of adversary argument, directly a lower bound on the rank in any sound proof system. This should be seen as a generalization of known width lower bound techniques (as abstractly explained e.g. in [2]).

## 2 Preliminaries

We write  $[n] := \{0, \dots, n-1\}$  and  $|n| := \lceil \log(n+1) \rceil$ . All logarithms are base two. Note  $|n|$  is the length of the binary encoding of  $n$ . If  $\bar{a}$  is a  $k$ -tuple, its  $i$ -th component is  $a_i$ .

For the rest of the section we fix a first-order vocabulary  $\sigma$  split into  $\sigma_R$  and  $\sigma_F$ , where  $\sigma_R$  is the set of relation symbols and  $\sigma_F$  is the set of function symbols. We view constant symbols as 0-ary function symbols. It will be convenient to assume that every vocabulary has at least one constant symbol that we denote by 0. For each symbol  $S$  in  $\sigma$ , let  $r_S$  denote its arity. If  $\bar{r}$  and  $\bar{s}$  are  $k$ -tuples of first-order terms, sometimes we write  $\bar{r} = \bar{s}$  instead of  $r_1 = s_1 \wedge \dots \wedge r_k = s_k$ . Similarly, we write  $\forall \bar{x}$  and  $\exists \bar{x}$  instead of  $\forall x_1 \dots \forall x_k$  and  $\exists x_1 \dots \exists x_k$ , and  $\psi[\bar{x}/\bar{a}]$  instead of  $\psi[x_1/a_1, \dots, x_k/a_k]$ .

**Basic propositional logic** In writing logical formulas we identify  $\neg \bigwedge_i F_i$  with  $\bigvee_i \neg F_i$  and  $\neg \bigvee_i F_i$  with  $\bigwedge_i \neg F_i$  without mention. Similarly  $\neg \neg F$  and  $F$  are viewed as the same formula. Also we apply commutativity, associativity and idempotency of the propositional connectives without mention. Sometimes we write  $\bar{A}$  instead of  $\neg A$ .

We define four rules of inference. These are axiom (AXM), weakening (WKG), introduction of conjunction (IOC), and cut (CUT):

$$\frac{}{F \vee \neg F} \quad \frac{\Delta}{\Delta \vee G} \quad \frac{\Delta \vee F \quad \Delta' \vee G}{\Delta \vee \Delta' \vee (F \wedge G)} \quad \frac{\Delta \vee F \quad \Delta' \vee \neg F}{\Delta \vee \Delta'}$$

where  $F$  and  $G$  denote formulas, and  $\Delta$  and  $\Delta'$  denote either formulas or the special empty formula which we denote by  $\square$ . If  $\Delta$  is the special empty formula, then  $\Delta \vee \Delta'$  is simply  $\Delta'$ .

A *proof (of  $G$  from  $F_1, \dots, F_m$ )* takes assumptions  $F_1, \dots, F_m$  and produces a *conclusion*  $G$  through the application of these rules. A *refutation* of  $F_1, \dots, F_m$  is a proof of  $\square$  from  $F_1, \dots, F_m$ . The *size* of a proof is the number of symbols it contains. A *resolution proof* is one where all formulas are clauses and the only allowed rule is CUT. If  $k \geq 1$  is an integer, an  $R(k)$ -*proof* is one where all formulas are  $k$ -DNF formulas.

We write  $F_1, \dots, F_m \vdash_k^s G$  if there is a  $R(k)$ -proof of size  $s$  that takes the assumptions  $F_1, \dots, F_m$  and produces  $G$ . We write  $\vdash_{k,*}^s$  for tree-like such proofs. Observe that  $F_1, \dots, F_m \vdash_1^s G$  (resp.  $\vdash_{1,*}^s$ ) if and only if there is a (resp. tree-like) resolution proof of  $G$  from  $F_1, \dots, F_m$ .

An  $R(c)$ -*proof* is one in which all formulas are  $k$ -DNFs, where  $k$  is some fixed constant. An  $R(\log)$ -*proof* is one in which all formulas are  $(\log s)$ -DNFs, where  $s$  is the size of the proof.

**Universal, flattened, function-negative formulas** Let  $\phi$  be a universal first-order formula over  $\sigma$ , which by standard manipulation we may assume has the form

$$\forall x_1 \dots \forall x_k (C_1 \wedge \dots \wedge C_s), \tag{1}$$

where each  $C_i$  is a clause made of literals on atomic formulas of one of the following forms:

1.  $x_i = x_j$  for  $i, j \in \{1, \dots, k\}$ ,
2.  $R(x_{i_1}, \dots, x_{i_r})$  for some  $R$  in  $\sigma_R$  of arity  $r$  and  $i_1, \dots, i_r \in \{1, \dots, k\}$ ,
3.  $F(x_{i_1}, \dots, x_{i_r}) = x_{i_0}$  for some  $F$  in  $\sigma_F$  of arity  $r$  and  $i_0, i_1, \dots, i_r \in \{1, \dots, k\}$ .

We do not allow nested terms. This is no loss of generality since nested terms can be *flattened* by noting that  $\psi(t)$  is logically equivalent to  $\forall z (t = z \rightarrow \psi(z))$ . If we apply this transformation to an

already flattened clause we end up with all the atoms of the form  $F(\bar{x}) = y$  occurring negatively. If all terms in  $\phi$  are flattened, we call it a *flattened universal formula*. If in addition atoms of the form  $F(\bar{x}) = y$  occur negatively, we call it a *flattened function-negative universal formula*. Note finally that, by standard Skolemization, every first-order sentence can be brought into one in this form while preserving the satisfiability at each cardinality. More formally:

**Fact 1.** *For every first-order sentence  $\phi$  there is a flattened function-negative universal sentence  $\phi'$  that has the same spectrum; that is, for every finite or infinite cardinal  $\kappa$  the sentence  $\phi$  has a model of cardinality  $\kappa$  if and only if so does  $\phi'$ .*

From now on we use the term *standardized universal formula* instead of “flattened function-negative universal formula”.

**Propositional encodings** Let  $\phi$  be a flattened universal sentence as in (1) and let  $n \geq 1$  be a natural number. We will define two propositional CNF-formulas  $\langle \phi \rangle_n^u$  and  $\langle \phi \rangle_n^b$ . In both cases the satisfying assignments of the propositional formula will be in one-to-one correspondence to the models of  $\phi$  with universe  $[n]$ . We start with  $\langle \phi \rangle_n^u$  (the  $u$  stands for *unary*). The variables are the following:

1.  $R_{\bar{a}}$  for each  $R$  in  $\sigma_R$  of arity  $r$  and each  $\bar{a} \in [n]^r$ ,
2.  $F_{\bar{a};a}$  for each  $F$  in  $\sigma_F$  of arity  $r$  and each  $\bar{a} \in [n]^r$  and  $a \in [n]$ .

These variables correspond in an obvious way to the ground first-order atoms of  $\phi$  through the translation  $\langle R(\bar{a}) \rangle_n := R_{\bar{a}}$  and  $\langle F(\bar{a}) = a \rangle_n := F_{\bar{a};a}$ . Once the translation is defined for atoms it extends to arbitrary formulas through the usual recurrence:

1.  $\langle \neg \psi \rangle_n := \neg \langle \psi \rangle_n$ ,
2.  $\langle \psi \wedge \theta \rangle_n := \langle \psi \rangle_n \wedge \langle \theta \rangle_n$ ,
3.  $\langle \psi \vee \theta \rangle_n := \langle \psi \rangle_n \vee \langle \theta \rangle_n$ ,
4.  $\langle \forall x \psi \rangle_n := \bigwedge_{a \in [n]} \langle \psi[x/a] \rangle_n$ ,
5.  $\langle \exists x \psi \rangle_n := \bigvee_{a \in [n]} \langle \psi[x/a] \rangle_n$ .

Now, if  $\phi$  is a flattened universal sentence, then the clauses of  $\langle \phi \rangle_n^u$  are:

1.  $\langle C_i[\bar{x}/\bar{a}] \rangle_n$  for each  $i \in \{1, \dots, s\}$  and each  $\bar{a} \in [n]^k$ ,
2.  $F_{\bar{a};1} \vee \dots \vee F_{\bar{a};n}$  for each  $F$  in  $\sigma_F$  of arity  $r$  and each  $\bar{a} \in [n]^r$ ,
3.  $\overline{F_{\bar{a};b}} \vee \overline{F_{\bar{a};c}}$  for each  $F$  in  $\sigma_F$  of arity  $r$  and each  $\bar{a} \in [n]^r$  and  $b, c \in [n]$  with  $b \neq c$ .

Clauses of type 1. are called *matrix clauses*, clauses of type 2. are called *long functional clauses*, and those of type 3. are called *short functional clauses*. Note that the size and the number of variables of  $\langle \phi \rangle_n^u$  are bounded by a fixed polynomial in  $n$  that depends only on  $\phi$ .

Next we define  $\langle \phi \rangle_n^b$  (the  $b$  stands for *binary*). The variables are:

1.  $R_{\bar{a}}$  for each  $R$  in  $\sigma_R$  of arity  $r$  and each  $\bar{a} \in [n]^r$ ,
2.  $F_{\bar{a};b}$  for each  $F$  in  $\sigma_F$  of arity  $r$  and each  $\bar{a} \in [n]^r$  and  $b \in \{0, \dots, |n| - 1\}$ .

The  $R_{\bar{a}}$  still have an obvious correspondence with the relational ground atoms. However, the intended meaning of  $F_{\bar{a};b}$  is different. Its meaning is that the  $b$ -th less significant bit in the binary encoding of  $F(\bar{a})$  is 1. Thus, in this case the base cases of the translation are  $\langle R(\bar{a}) \rangle'_n := R_{\bar{a}}$  and

$$\langle F(\bar{a}) = a \rangle'_n := \bigwedge_{b=0}^{|n|-1} F_{\bar{a};b}^{(\text{bit}(b,a))}$$

where  $\text{bit}(b, a)$  is the  $b$ -th less significant bit in the binary encoding of  $a$ , and  $X^{(1)}$  stands for  $X$  and  $X^{(0)}$  stands for  $\bar{X}$ . Recall that we identify  $\neg \bigwedge_i F_i$  with  $\bigvee_i \neg F_i$ . Therefore, if  $C$  is a ground clause in which all function atoms  $F(\bar{a}) = a$  appear negatively, the translation  $\langle C \rangle'_n$  is still a clause.

Now, if  $\phi$  is a standardized universal sentence, then the clauses of  $\langle \phi \rangle_n^b$  are:

1.  $\langle C_i(\bar{x}/\bar{a}) \rangle'_n$  for each  $i \in \{1, \dots, s\}$  and  $\bar{a} \in [n]^k$ ,
2.  $\bigvee_{b=0}^{|n|-1} F_{\bar{a};b}^{(1-\text{bit}(b,a))}$  for each  $F \in \sigma_F$  and  $a \in [2^n] \setminus [n]$ .

The second type of clauses are saying that  $F$  has its range inside  $[n]$ . Note that the long and short functional clauses for  $F$  are morally implicit.

**Fact 2.** *Let  $\phi$  be a standardized sentence. For every integer  $n \geq 1$ , there is a bijection between the set of models of  $\phi$  with universe  $[n]$  and the set of satisfying assignments for both  $\langle \phi \rangle_n^u$  and  $\langle \phi \rangle_n^b$ .*

If  $T$  is a theory given by a (finite or infinite) collection of standardized sentences, let  $\langle T \rangle_n^u := \bigcup_{\phi \in T} \langle \phi \rangle_n^u$  and  $\langle T \rangle_n^b := \bigcup_{\phi \in T} \langle \phi \rangle_n^b$ .

### 3 General Facts about $\mathbf{R}(k)$

This section contains some elementary, and mostly known facts about the power of the proof system  $\mathbf{R}(k)$ . The first is the following quantitative version of completeness:

**Lemma 1.** *Let  $\Gamma \cup \{F\}$  be a set of propositional formulas each of size at most  $s$  and mentioning  $n$  variables in total. If  $\Gamma \models F$ , then  $F$  has a tree-like proof from  $\Gamma$  of size  $O(s^2 n 2^n)$ ; moreover, the proof is an  $\mathbf{R}(k)$ -proof if each formula in  $\Gamma \cup \{F\}$  is a  $k$ -DNF.*

*Proof.* Fix a set of  $n$  variables. For an assignment  $\alpha$  to these variables let  $C_\alpha$  be the disjunction of all literals falsified by  $\alpha$ . Let  $G$  be a formula in the fixed variables and  $\alpha$  an assignment. We claim that there is a size  $O(|G|)$ , tree-like, cut free proof of  $G \vee C_\alpha$  or  $\neg G \vee C_\alpha$  depending on whether  $\alpha \models G$  or not. This can be verified by a straightforward induction on  $G$ : e.g. assume that  $G = H \wedge I$  and that the claim holds for  $H$  and  $I$ . If  $\alpha \not\models G$ , say  $\alpha \not\models H$ , we know there is proof as desired for  $\neg H \vee C_\alpha$ ; weakening gives  $\neg H \vee \neg I \vee C_\alpha$  and this is the same as  $\neg G \vee C_\alpha$  (recall our conventions on how to write formulas). If otherwise  $\alpha \models G$ , then we know there are proofs as desired of both  $H \vee C_\alpha$  and  $I \vee C_\alpha$ ; these can be joined by an introduction of conjunction. Because these proofs are cut free they are  $\mathbf{R}(k)$  proofs whenever  $G$  is a  $k$ -DNF.

Now assume  $\Gamma \models F$ . For all assignments  $\alpha$  that satisfy all  $G \in \Gamma$ , and hence  $F$ , derive  $F \vee C_\alpha$  with no assumptions. For all assignments  $\alpha$  that do not satisfy all  $G \in \Gamma$ , say  $\alpha \models \bar{G}$ , derive  $F \vee C_\alpha$  from  $\Gamma$  by first deriving  $\bar{G} \vee C_\alpha$  with no assumptions, then cutting on  $G$ , and finally adding  $F$  by weakening. Take a tree-like  $\mathbf{R}(1)$  proof of size  $O(n 2^n)$  that refutes the set of clauses  $C_\alpha$  where  $\alpha$  ranges over all assignments. Adding  $F$  to all formulas occurring in this refutation gives a derivation of  $F$  from the already derived  $F \vee C_\alpha$ . The overall size of this proof is  $O(|F|^2 n 2^n)$ , as claimed.

To make it an  $R(k)$ -proof when all formulas in  $\Gamma \cup \{F\}$  are  $k$ -DNFs argue as follows. In the paragraph above, instead of deriving  $\overline{G} \vee C_\alpha$ , derive  $C \vee C_\alpha$  for each clause  $C$  of  $\overline{G}$ . Then use these to cut all the terms of  $G$  one by one until we get  $C_\alpha$ . Finally add  $F$  by weakening and proceed as before.  $\square$

The following clarifies the relationship between the tree-like and the dag-like versions. It goes back to [10] and appears in the form stated here in [9, Theorem 16].

**Theorem 2.** *Let  $\Gamma$  be a set of clauses. If  $\Gamma \vdash_{k,*}^s \square$ , then  $\Gamma \vdash_1^{2s} \square$ .*

The next states the Deduction Theorem for  $R(k)$ :

**Lemma 2.** *Let  $\Gamma \cup \{F\}$  be a set of  $k$ -DNFs and let  $C_1, \dots, C_n$  be clauses with at most  $k$  literals each. If  $\Gamma, C_1, \dots, C_n \vdash_k^s F$ , then  $\Gamma \vdash_k^{s'} \neg C_1 \vee \dots \vee \neg C_n \vee F$  for  $s' = O(s \cdot \sum_{i=1}^n |C_i|)$ .*

*Proof.* Let  $\Gamma = \{F_1, \dots, F_m\}$ , assume  $\Gamma, C_1, \dots, C_n \vdash_k^s F$ , and let  $\Pi$  be the proof witnessing it. Let  $H := \neg C_1 \vee \dots \vee \neg C_n$ . Add  $H$  to every formula in the proof to get a proof of  $H \vee F$  from axioms  $F_i \vee H$  and  $C_i \vee H$ . Observe that  $F_i \vee H$  can be obtained from  $F_i$  by weakening and  $H \vee C_i$  is a weakening of an axiom  $C_i \vee \neg C_i$ .  $\square$

In the next lemma we show that we can replace small formulas for variables in proofs and distribute out. Let  $D$  be the distributivity operator on propositional formulas that recursively distributes conjunctions over disjunctions. In other words, the operator converts an arbitrary propositional formula into a DNF by the naive method. The operator is defined inductively by cases on the outermost connective of the formula. Formally:

1. If  $A$  is a literal or a conjunction of literals, then  $D(A) := A$ ,
2. If  $A = \bigvee_{i=1}^r A_i$  where each  $A_i$  is not a disjunction, then  $D(A) := \bigvee_{i=1}^r D(A_i)$ ,
3. If  $A = \bigwedge_{i=1}^r \bigvee_{j=1}^{s_i} A_{i,j}$  where each  $A_{i,j}$  is not a disjunction, then

$$D(A) := \bigvee_{j_1=1}^{s_1} \dots \bigvee_{j_r=1}^{s_r} D\left(\bigwedge_{i=1}^r A_{i,j_i}\right).$$

The next we call the distributivity lemma for substituted instances:

**Lemma 3.** *Let  $\Gamma \cup \{F\}$  be a set of  $k$ -DNFs. For every variable  $x$ , let  $G_x$  and  $H_x$  be equivalent  $t$ -term- $c$ -DNF and  $t$ -clause- $c$ -CNF formulas, respectively, with  $v$  variables. For every formula  $A$  let  $A'$  be the result of replacing each positive occurrence of a variable  $x$  by  $G_x$  and each negative occurrence of a variable  $x$  by  $H_x$ . If  $\Gamma \vdash_k^s F$ , then  $D(\Gamma') \vdash_{k'}^{s'} D(F')$  where  $k' = kc$  and  $s' = O(s(kct^k)^2 kv 2^{kv})$ .*

*Proof.* We may assume that the given proof applies AXM only to atoms since every axiom  $A \vee \overline{A}$  has a short derivation from its underlying atom-axioms. It suffices to show that, whenever  $C$  is derived from  $A$  and  $B$  by a single  $R(k)$ -rule, there is a short  $R(k')$ -proof of  $D(C')$  from  $D(A')$  and  $D(B')$ . We distinguish by cases on the type of rule.

(AXM): We want a proof of  $D(x' \vee \overline{x'})$ , which is  $G_x \vee \overline{H_x}$ . Since  $G_x$  and  $H_x$  are equivalent, this is a tautology, and by completeness it has a proof. Since it is a  $c$ -DNF of size at most  $ct$  on at most  $v$  variables, by Lemma 1 the proof is in  $R(c)$  and has size at most  $O((ct)^2 v 2^v)$  times larger than the size of  $x \vee \overline{x}$ .

(WKG): Suppose  $A \vee B$  is derived from  $A$  by weakening. From  $D(A')$  we derive  $D(A') \vee D(B') = D(A' \vee B')$  in one weakening step. The derived formula is a  $kc$ -DNF and its size is at most  $kc t^k$  times larger than the size of  $A \vee B$ .

(CUT): Suppose  $A \vee B$  is derived through the cut rule on  $A \vee T$  and  $B \vee \bar{T}$ , where  $T$  is a term of at most  $k$  literals. We want to obtain  $D(A' \vee B')$  from  $D(A' \vee T')$  and  $D(A' \vee \bar{T}')$ . To that end it suffices to derive the empty clause from  $D(T')$  and  $D(\bar{T}')$ , which are contradictory since  $G_x$  and  $H_x$  are equivalent. By completeness, such a refutation exists. Since both are  $kc$ -DNFs of size at most  $kc t^k$  on at most  $kv$  variables, by Lemma 1 the refutation is in  $R(kc)$  and has size  $O((kc t^k)^2 kv 2^{kv})$ . Overall the corresponding proof of  $D(A' \vee B')$  is also this many times larger than the size of  $A \vee B$ .

(IOC): Suppose  $A \vee B \vee (S \wedge T)$  is derived by introduction of conjunction from  $A \vee S$  and  $B \vee T$ , where  $S \wedge T$  has at most  $k$  literals. We want to derive  $D(A' \vee B' \vee (S \wedge T)')$  from  $D(A' \vee S')$  and  $D(B' \vee T')$ . To that end, it suffices to derive  $D((S \wedge T)')$  from  $D(S')$  and  $D(T')$ , which can be done by completeness. These formulas are  $kc$ -DNFs of size at most  $kc t^k$  on at most  $kv$  variables. Therefore, by Lemma 1, the proof is in  $R(kc)$  and has size  $O((kc t^k)^2 kv 2^{kv})$ . Overall the corresponding proof of  $D(A' \vee B' \vee (S \wedge T)')$  is also this many times larger than the size of  $A \vee B \vee (S \wedge T)$ .

To complete the proof, note that each simulation step is a proof in  $R(kc)$  and that its size is at most  $O((kc t^k)^2 kv 2^{kv})$  times larger than the corresponding formula in the original proof.  $\square$

As a first application of the distributivity lemma we show how to translate refutations of the unary encoding to the binary encoding.

**Lemma 4.** *Let  $k \geq 1$  and  $\phi$  be a standardized universal formula. There exists a polynomial  $p$  such that, for every  $n$ , if  $\langle \phi \rangle_n^u \vdash_k^s \square$ , then  $\langle \phi \rangle_n^b \vdash_{\log}^{s'} \square$  for  $s'$  polynomial in  $s$  and  $n^k$ .*

*Proof.* For a formula  $G$  let  $G'$  be obtained by replacing every atom  $F_{\bar{a};a}$  by  $\bigwedge_{b=0}^{|n|-1} F_{\bar{a};b}^{(\text{bit}(b,a))}$ . In a first step we derive from the initial clauses of  $\langle \phi \rangle_n^b$ , the  $|n|$ -DNF's of the primed axioms of  $\langle \phi \rangle_n^u$ . The matrix clauses of  $\langle \phi \rangle_n^b$  are exactly the primed matrix clauses of  $\langle \phi \rangle_n^u$  because  $\phi$  is in function-negative form. We show how to derive the primed functional clauses of  $\langle \phi \rangle_n^u$ .

The primed long functional clauses of  $\langle \phi \rangle_n^u$  read  $\bigvee_{a \in [n]} \bigwedge_{b=0}^{|n|-1} F_{\bar{a};b}^{(\text{bit}(b,a))}$  for  $F \in \sigma$  and  $\bar{a} \in [n]^{r_F}$ . Note that if  $n$  is a power of 2, then this formula is a tautology in  $\log n$  many variables, and therefore has a short  $R(\log)$ -proof by Lemma 1. If  $n$  is not a power of 2, then first give a proof of the formula for  $2^{\lceil \log n \rceil}$  and then arrive at the formula for  $n$  by cuts with the axioms  $\bigvee_{b=0}^{|n|-1} F_{\bar{a};b}^{1-\text{bit}(b,a)}$  for  $a \in [2^{\lceil \log n \rceil}] \setminus [n]$ .

Primed short functional clauses of  $\langle \phi \rangle_n^u$  read  $\bigvee_{b=0}^{|n|-1} (F_{\bar{a};b}^{(1-\text{bit}(b,a))} \vee F_{\bar{a};b}^{(1-\text{bit}(b,a'))})$  for  $F \in \sigma$ ,  $\bar{a} \in [n]^{r_F}$  and  $a, a' \in [n]$  with  $a \neq a'$ . But such a formula is a weakening of an axiom since the binary representations of  $a$  and  $a'$  differ in some bit.

Let  $\Gamma := \langle \phi \rangle_n^u$ . We now have derived  $\Gamma'$ , the primed version of  $\Gamma$ , which is obtained from replacing each occurrence of an  $F$ -atom by a size- $|n|$  conjunction. Note that a size- $|n|$  conjunction is both a 1-term- $|n|$ -DNF and an  $|n|$ -clause-1-CNF with  $|n|$  variables. By assumption  $\Gamma \vdash_k^s \square$ , so by Lemma 3 we get  $D(\Gamma') \vdash_{\log}^{s'} D(\square)$  for some  $s'$  that is polynomial in  $s$  and  $n^k$ . It now suffices to note that  $D(\Gamma') = \Gamma'$ , and obviously  $D(\square) = \square$ .  $\square$

*Remark 1.* Note that the derivation of the substitution instances of the long functional clauses is tree-like  $R(\log)$ .  $\diamond$

## 4 Closure under Interpretability

Here we show that the class of relativized-sentences that have short R(c)-refutations is closed under quantifier-free interpretability when encoded in unary. Before we start we need to define relativization and interpretability.

### 4.1 Relativization and interpretability

Let  $\sigma$  be a vocabulary, split into  $\sigma_R$  and  $\sigma_F$  for relation and function symbols, respectively. Recall that we assume that  $\sigma$  has at least one constant symbol denoted by 0.

Let  $\phi$  be a flattened first-order formula over  $\sigma$ . The *relativization of  $\phi$* , denoted by  $\phi^U$ , is the result of adding one unary relation symbol  $U$  and replacing each universal quantifier  $\forall x\psi$  by  $\forall x(U(x) \rightarrow \psi)$  and each existential quantifier  $\exists x\psi$  by  $\exists x(U(x) \wedge \psi)$ . If  $\phi$  is a sentence, then the sentence *relativized- $\phi$* , denoted by  $\phi^R$ , is the result of adding a new unary relation symbol  $U$  and forming the formula

$$\phi^U \wedge \bigwedge_{F \in \sigma_F} \forall \bar{x} \forall y (U(x_1) \wedge \cdots \wedge U(x_{r_F}) \wedge F(\bar{x}) = y \rightarrow U(y)).$$

Recall that  $\sigma$  contains the constant 0 and thus  $U$  is forced non-empty by the second conjunct on  $F = 0$ . Note that  $\phi^R$  holds in a structure  $M$  if and only if  $\phi$  holds on some (non-empty) substructure of  $M$ . Note also that if  $\phi$  is a standardized sentence, then  $\phi^R$  is a conjunction of such sentences. If  $T$  is a theory, define  $T^R := \bigcup_{\phi \in T} \phi^R$ .

*Example 1.* The Least Number Principle (LNP) is the sentence saying that  $<$  is a strict partial order and that  $F$  is a function mapping each element  $x$  to some  $y$  with  $y < x$ . It was shown in [19] (see also [4]) that  $\langle \text{LNP} \rangle_n^u$  has polynomial-size resolution refutations. Its relativized version  $\text{LNP}^R$  was considered in [6, 7] and it was argued that its unary encoding  $\langle \text{LNP}^R \rangle_n^u$  has polynomial-size R(2)-refutations. Let us also note that, by Lemma 4, it follows that its binary encoding  $\langle \text{LNP}^R \rangle_n^b$  has polynomial-size R(log)-refutation.  $\diamond$

Let  $\tau$  be a first-order vocabulary other than  $\sigma$ . Let  $S$  and  $T$  be theories over  $\sigma$  and  $\tau$ , respectively. Let  $\Theta$  be a collection of flattened first-order formulas. A  $\Theta$ -translation from  $\sigma$  to  $\tau$  consists of  $\Theta$ -formulas  $\theta_U(\bar{x}; \bar{p})$ ,  $\theta_R(\bar{x}_1, \dots, \bar{x}_{r_R}; \bar{p})$  and  $\theta_F(\bar{x}_1, \dots, \bar{x}_{r_F}, \bar{y}; \bar{p})$ , where  $R$  and  $F$  range over  $\sigma_R$  and  $\sigma_F$  respectively, in which all the depicted tuples except  $\bar{p}$  have the same length  $s \geq 1$ . The variables in  $\bar{p}$  are called *parameters*, and  $s$  is the *arity* of the translation. We often omit writing down the parameters.

Let  $\phi$  be a formula over  $\sigma$  and let  $I$  be a translation from  $\sigma$  to  $\tau$ . Choose the number  $s \geq 1$  as above. The *translation of  $\phi$  through  $I$* , denoted by  $I(\phi)$ , is the formula over  $\tau$  that is obtained from  $\phi^U$  as follows: fix for every variable  $x$  a new  $s$ -tuple  $\bar{x} = (x_1, \dots, x_s)$  of pairwise distinct variables; replace in  $\phi^U$  every occurrence of a quantifier  $\forall x$  or  $\exists x$  by  $\forall \bar{x}$  and  $\exists \bar{x}$  respectively; replace in  $\phi^U$  every atom  $x = y$  by  $\bar{x} = \bar{y}$ , every atom  $U(x)$  by  $\theta_U(\bar{x})$  every atom  $R(y_1, \dots, y_{r_R})$  for  $R \in \sigma_R$  by  $\theta_R(\bar{y}_1 \cdots \bar{y}_{r_R})$  and every atom  $F(y_1, \dots, y_{r_F}) = z$  for  $F \in \sigma_F$  by  $\theta_F(\bar{y}_1 \cdots \bar{y}_{r_F}, \bar{z})$ .

We say that  $I$  *interprets  $S$  in  $T$*  if the following conditions are satisfied. Let  $\delta = \delta(\bar{p})$  be the formula  $\bigwedge_{i \neq j} \neg p_i = p_j$ . We agree that this formula is  $\top$  for the empty parameter tuple. Then:

1.  $T \models \delta \rightarrow \forall \bar{x} \exists \bar{y} \theta_F(\bar{x}, \bar{y})$ ,
2.  $T \models \delta \rightarrow \forall \bar{x} \forall \bar{y} \forall \bar{z} (\theta_F(\bar{x}, \bar{y}) \wedge \theta_F(\bar{x}, \bar{z}) \rightarrow \bar{y} = \bar{z})$ ,

3.  $T \models \delta \rightarrow \forall \bar{x} \forall \bar{y} (\theta_U(\bar{x}_1) \wedge \cdots \wedge \theta_U(\bar{x}_{r_F}) \wedge \theta_F(\bar{x}, \bar{y}) \rightarrow \theta_U(\bar{y}))$ ,
4.  $T \models \delta \rightarrow I(\phi)$ , for every  $\phi \in S$ .

Say that  $S$  is  $\Theta$ -interpretable in  $T$ , or that  $T$   $\Theta$ -interprets  $S$ , if there is a  $\Theta$ -translation that interprets  $S$  in  $T$ .

*Example 2.* Recall the Least Number Principle (LNP) from Example 4.1. The Dense Linear-Order Principle (DLOP) is the sentence saying that  $<$  is a strict linear order, and that  $G$  is a function mapping any two different elements  $x$  and  $y$  to some  $z$  that lies strictly between  $x$  and  $y$ . Both LNP and DLOP are written as standardized universal sentences. We show that LNP is quantifier-free interpretable in DLOP. The interpretation uses two parameter variables  $p$  and  $q$  and goes as follows. Define:

1.  $\theta_U(x) := (p < q \wedge p < x \leq q) \vee (q < p \wedge q < x \leq p)$ ,
2.  $\theta_{<}(x, y) := x < y$ ,
3.  $\theta_0(y) := (p < q \wedge y = q) \vee (q < p \wedge y = p)$ ,
4.  $\theta_F(x, y) := (p < q \wedge G(p, x) = y) \vee (q < p \wedge G(q, x) = y)$ .

The correctness relies on the fact that the interpretation is required to work only when  $p \neq q$ . Note that this formula is flattened.  $\diamond$

## 4.2 Proof translations

Our next goal is to show that under the unary encoding, small refutations in  $R(c)$  transfer through interpretability, provided the interpreted sentence has a short refutation of its relativization.

We will make use of the *upper-bound half* of Riis' Gap Theorem for tree-like resolution as stated below:

**Theorem 3** (Riis, 2001). *Let  $\phi_1(\bar{x}), \dots, \phi_t(\bar{x})$  be flattened universal formulas with  $r$  free variables. If  $\bigwedge_{i=1}^t \phi_i(\bar{x})$  is unsatisfiable, then there exists a polynomial  $p$  such that for every natural  $n$  and every  $\bar{a} \in [n]^r$  we have  $\langle \phi_1[\bar{x}/\bar{a}] \rangle_n^u, \dots, \langle \phi_t[\bar{x}/\bar{a}] \rangle_n^u \vdash_{1,*}^{p(n)}$ .  $\square$ .*

The following consequence of this result will be used several times:

**Lemma 5.** *Let  $\psi$  be a flattened universal sentence and let  $\phi(\bar{x}, \bar{y})$  be a flattened quantifier-free formula in DNF form, where  $\bar{x}$  has length  $r$  and  $\bar{y}$  has length  $s$ . If  $\psi \models \forall \bar{x} \exists \bar{y} \phi(\bar{x}, \bar{y})$ , then for every natural  $n$  and every  $\bar{a} \in [n]^r$  there is a natural  $k$  and a polynomial  $p$  such that  $\langle \psi \rangle_n^u \vdash_k^{p(n)} \bigvee_{\bar{b} \in [n]^s} \langle \phi[\bar{x}/\bar{a}, \bar{y}/\bar{b}] \rangle_n$ .*

*Proof.* Write  $\phi$  as  $\bigvee_i t_i$  where each  $t_i$  is a term. Then,  $\psi \wedge \forall \bar{y} \bigwedge_i \neg t_i$  is unsatisfiable. Note that  $\bar{x}$  is free in this unsatisfiable formula. Since  $\psi$  is a flattened universal sentence and  $\forall \bar{y} \bigwedge_i \neg t_i$  is a flattened universal formula with free variables, we may apply Theorem 3. We get

$$\langle \psi \rangle_n^u, \langle \forall \bar{y} \bigwedge_i \neg t_i[\bar{x}/\bar{a}] \rangle_n^u \vdash_1^{\text{poly}} \square$$

for every  $\bar{a} \in [n]^r$ . Here,  $\vdash^{\text{poly}}$  refers to a polynomial-size (in  $n$ ) proof. Now, since the vocabulary of both premises is the same, the functional clauses of both encoded formulas are already present in  $\langle \psi \rangle_n^u$ . We get

$$\langle \psi \rangle_n^u, \{ \langle \neg t_i[\bar{x}/\bar{a}, \bar{y}/\bar{b}] \rangle_n : \bar{b} \in [n]^s, i \} \vdash_1^{\text{poly}} \square$$

Now, by Lemma 2 we obtain

$$\langle \psi \rangle_n^u \vdash_k^{\text{poly}} \bigvee_{\bar{b} \in [n]^s} \bigvee_i \langle t_i[\bar{x}/\bar{a}, \bar{y}/\bar{b}] \rangle_n = \bigvee_{\bar{b} \in [n]^s} \langle \phi[\bar{x}/\bar{a}, \bar{y}/\bar{b}] \rangle_n$$

for some natural  $k$ . □

**Lemma 6.** *Let  $\phi$  and  $\psi$  be standardized universal sentences. If  $\phi$  is quantifier-free interpretable in  $\psi$ , then for every natural  $k$ , there exists a natural  $k'$  and a polynomial  $p$ , such that for all naturals  $n$  and  $\ell$ , if  $\langle \phi^R \rangle_n^u \vdash_k^s \square$ , then  $\langle \psi \rangle_n^u \vdash_{k'}^{p(n,s)} \square$ .*

*Proof.* Let  $\sigma$  be the vocabulary of  $\phi$  split into  $\sigma_R$  and  $\sigma_F$  as usual. Let  $\tau$  be the vocabulary of  $\psi$ . Let  $I$  be a quantifier-free translation that interprets  $\phi$  in  $\psi$ . To simplify notation, we will assume that the arity  $s$  of the translation  $I$  is 1, and further, that the parameter tuple  $\bar{p}$  is empty. Note that  $\delta = \top$  in this case. It is easy to adapt the following argument to the general case.

By the definition of interpretability we have  $\psi \models I(\phi)$ . The underlying idea of the proof is simple: from the short propositional refutation of the encoding of  $\phi^R$ , get a propositional refutation of the encoding of  $I(\phi)$ , and combine it with a propositional proof of  $\psi \models I(\phi)$  to get a propositional refutation of the encoding of  $\psi$ . However the details are no so easy because  $I(\phi)$  does not have the right form and cannot be converted directly into sets of clauses. We need to do more work.

The conditions 1., 2., 3. and 4. of the definition of interpretability hold for  $\psi$  and  $\phi$  and the  $\theta$ -formulas that compose  $I$ . The formulas for conditions 1., 2., and 3. on the right of  $\delta$  have the form  $\forall \bar{x} \exists \bar{y} \gamma$ , where  $\gamma$  is a single clause on the  $\theta$ -formulas. For 4. this is not necessarily true, but we can write the formula as a conjunction of such formulas. Since the  $\theta$ -formulas are flattened quantifier-free, each can be written equivalently as a flattened quantifier-free DNF and a flattened quantifier-free CNF. If we replace each positive occurrence of a  $\theta$ -formula by its DNF and each negative occurrence of a  $\theta$ -formula by its CNF, what we obtain is a collection of formulas of the form  $\forall \bar{x} \exists \bar{y} \gamma$  where  $\gamma$  is a quantifier-free flattened DNF formula.

For  $q \in \{1, 2, 3\}$  and  $F \in \sigma_F$ , let  $\gamma_{q,F}$  be the quantifier-free part in the formula for condition  $q$ . Finally let  $\gamma_{4,i}$  be the quantifier-free part in the  $i$ -th conjunct for condition 4. By Lemma 5,

1.  $\langle \psi \rangle_n^u \vdash_c^{\text{poly}} \bigvee_{a \in [n]} \langle \gamma_{1,F}[\bar{x}/\bar{a}, x/a] \rangle_n$  for every  $F \in \sigma_F$  and  $\bar{a} \in [n]^{r_F}$ ,
2.  $\langle \psi \rangle_n^u \vdash_c^{\text{poly}} \langle \gamma_{2,F}[\bar{x}/\bar{a}, y/b, z/c] \rangle_n$  for every  $F \in \sigma_F$ , every  $\bar{a} \in [n]^{r_F}$ , and every  $b, c \in [n]$ ,
3.  $\langle \psi \rangle_n^u \vdash_c^{\text{poly}} \langle \gamma_{3,F}[\bar{x}/\bar{a}, y/b] \rangle_n$  for every  $F \in \sigma_F$ , every  $\bar{a} \in [n]^{r_F}$ , and  $b \in [n]$ ,
4.  $\langle \psi \rangle_n^u \vdash_c^{\text{poly}} \langle \gamma_{4,i}[\bar{x}/\bar{a}] \rangle_n$  for every  $i$  and  $\bar{a} \in [n]^k$ .

Here, by  $\vdash_c^{\text{poly}}$  we refer to a polynomial (in  $n$ ) size  $R(k)$  proof for a suitable constant  $k$ .

Let  $\Gamma := \langle \phi^R \rangle_n^u$ . Our next goal is to show that there is a substitution into  $\Gamma$  that takes every clause in  $\Gamma$  into one of the DNF formulas on the right-hand sides above. The clauses in  $\Gamma$  can be split into four types; namely:

1.  $C_{1,F}(\bar{a})$ : long functional clause for  $F \in \sigma_F$  and  $\bar{a} \in [n]^{r_F}$ ,

2.  $C_{2,F}(\bar{a}, b, c)$ : short functional clause for  $F \in \sigma_F$ ,  $\bar{a} \in [n]^{r_F}$ , and  $b, c \in [n]$  with  $b \neq c$ ,
3.  $C_{3,F}(\bar{a}, b)$ : clause enforcing closure of  $U$  under  $F \in \sigma_F$  for  $\bar{a} \in [n]^{r_F}$  and  $b \in [n]$ ,
4.  $C_{4,i}(\bar{a})$ : clause encoding the  $i$ -th conjunct of  $\phi^U$  for tuple  $\bar{a} \in [n]^k$ .

For each clause  $C \in \Gamma$ , let  $C'$  be the result of replacing each positive occurrence of an atom by the propositional translation of the DNF formula of the corresponding  $\theta$ -formula, and each negative occurrence of an atom by the propositional translation of the CNF formula of the corresponding  $\theta$ -formula. Direct inspection shows that:

1. if  $C = C_{4,i}(\bar{a})$ , then  $C' = \langle \gamma_{4,i}[\bar{x}/\bar{a}] \rangle_n$ ,
2. if  $C = C_{3,F}(\bar{a}, b)$ , then  $C' = \langle \gamma_{3,F}[\bar{x}/\bar{a}, y/b] \rangle_n$ ,
3. if  $C = C_{2,F}(\bar{a}, b, c)$ , then  $C' = \langle \gamma_{2,F}[\bar{x}/\bar{a}, y/b, z/c] \rangle_n$ .
4. if  $C = C_{1,F}(\bar{a})$ , then  $C' = \bigvee_{a \in [n]} \langle \gamma_{1,F}[\bar{x}/\bar{a}, x/a] \rangle_n$ .

This covers all cases. Let  $\Gamma' := \{C' : C \in \Gamma\}$  and note that since each  $C'$  is a DNF we have  $D(\Gamma') = \Gamma'$ . Therefore by the distributivity lemma applied to the hypothesis that  $\Gamma \vdash_k^s \square$  we get

$$\Gamma' = D(\Gamma') \vdash_c^{\text{poly}} D(\square) = \square.$$

Again  $\vdash^{\text{poly}}$  refers to polynomial-size proof. Composing proofs we get  $\langle \psi \rangle_n^u \vdash_c^{\text{poly}} \square$  as was to be proved.  $\square$

*Remark 2.* The previous lemma fails if one requires the implication 4. in the definition of interpretability to hold only with respect to finite structures. Indeed, with this weaker notion of interpretability  $\phi = \perp$  would become quantifier-free interpretable in any  $\psi$  without finite models. Then  $\square \in \langle \phi^R \rangle_n^u$ , but the clauses  $\langle \psi \rangle_n^u$  may not admit polynomial size  $R(c)$  refutations (e.g.  $\psi$  could be the pigeonhole principle).  $\diamond$

*Remark 3.* One might be tempted to try to prove Lemma 6 by the following sort of argument. Assume that  $\phi$  has no finite models and has a quantifier-free interpretation  $I$  in  $\psi$ , say for simplicity, with empty parameter tuple  $\bar{p}$ . Further assume that the contradictions  $\Gamma := \langle \phi^R \rangle_n^u$  have short  $R(k)$ -refutations. Then, using the refutation as a strategy, a prover playing against an adversary must be able to construct a partial assignment falsifying a clause of  $\Gamma$  by querying only some few  $k$ -terms on the atoms of  $\Gamma$ . If instead of querying these terms prover's queries were about the truth value of the formulas in which the atoms are replaced by their definitions through  $I$ , then the position reached by the prover would hold a partial assignment falsifying a DNF from  $\Gamma'$ , where  $\Gamma'$  is the propositional translation of  $I(\phi)$  written as a set of DNF formulas. Now, since  $\psi \models I(\phi)$ , we have  $\langle \psi \rangle_n^u \models \Gamma'$ , so some clause of  $\langle \psi \rangle_n^u$  must be false under any assignment extending the partial one constructed by the prover. It is thus tempting to interpret prover's position as winning in the game for  $\langle \psi \rangle_n^u$ . The problem this has is that the falsified clause in  $\langle \psi \rangle_n^u$  may vary with the extension chosen. In other words, the partial assignment constructed has no need to falsify a particular clause of  $\langle \psi \rangle_n^u$ . To find such a clause one would like to use a short proof of  $\langle \psi \rangle_n^u \models \Gamma'$  but that might not be available.

Note that the wrong argument above needs  $\psi \models I(\phi)$  to hold only in the finite, and the previous remark showed that this cannot work. What makes our argument work is that it exploits the fact

that  $\psi \models I(\phi)$  is much stronger than  $\langle \psi \rangle_n^u \models \Gamma'$ , which is equivalent to  $\psi \models I(\phi)$  holding in the finite, so much stronger that the latter becomes provable through short proofs.  $\diamond$

As an application of Lemma 6, we show that the class of relativized-formulas whose propositional translations have short R(c) refutations is closed under quantifier-free interpretability. This hinges on the fact that if  $\phi$  is a standardized universal formula, then  $\phi$  and  $\phi^R$  are mutually interpretable by quantifier-free translations.

**Lemma 7.** *Let  $\phi$  be a standardized universal sentence. Then  $\phi$  is quantifier-free interpretable in  $\phi^R$  and  $\phi^R$  is quantifier-free interpretable in  $\phi$ .*

*Proof.* Let  $\sigma$  be the vocabulary of  $\phi$  split into  $\sigma_R$  and  $\sigma_F$  as usual. Recall that we assume that  $0 \in \sigma_F$ . Define  $\theta_U(x) := U(x)$ ,  $\theta_R(\bar{x}) := R(\bar{x})$  for every  $R \in \sigma_R$ , and  $\theta_F(\bar{x}, y) := F(\bar{x}) = y$  for every  $F \in \sigma_F$ . It is straightforward to check that this is an interpretation of  $\phi$  in  $\phi^R$ .

In the reverse direction, let us first rename the relativizing predicate from  $\phi^R$  to  $V$  so avoid the overloading of names. Then define  $\theta_U(x) := x = x$ ,  $\theta_V(x) := x = x$ ,  $\theta_R(\bar{x}) := R(\bar{x})$  for every  $R \in \sigma_R$ , and  $\theta_F(\bar{x}, y) := F(\bar{x}) = y$  for every  $F \in \sigma_F$ . It is straightforward to check that this is an interpretation of  $\phi^R$  in  $\phi$ .  $\square$

**Theorem 4.** *Let  $\phi^R$  and  $\psi^R$  be relativized, standardized universal sentences. If  $\phi^R$  has a quantifier-free interpretation in  $\psi^R$ , then for every natural  $k$ , there exists a natural  $k'$  and a polynomial  $p$ , such that for all naturals  $n$  and  $\ell$ , if  $\langle \phi^R \rangle_n^u \vdash_k^\ell \square$ , then  $\langle \psi^R \rangle_n^u \vdash_{k'}^{p(n,\ell)}$ .*

*Proof.* Assume  $\langle \phi^R \rangle_n^u \vdash_k^\ell \square$ . First note that  $\phi$  is quantifier-free interpretable in  $\phi^{RR}$  by two applications of Lemma 7 and transitivity of interpretability. Therefore, by Lemma 6, we have  $\langle \phi^{RR} \rangle_n^u \vdash_{k''}^{p'(n,\ell)} \square$  for some natural  $k''$  and some polynomial  $p'$ . Now, if  $\phi^R$  is quantifier-free interpretable in  $\psi^R$ , then again by Lemma 6 we get  $\langle \psi^R \rangle_n^u \vdash_{k'}^{p(n,\ell)} \square$  for some natural  $k'$  and some polynomial  $p$ .  $\square$

## 5 Lower Bound and Dichotomy Theorem

For this section, fix a standardized universal sentence  $\phi$  that has some infinite model  $M$ . Let  $\sigma$  be the vocabulary of  $\phi$ , split into its relation symbols  $\sigma_R$  and its function symbols  $\sigma_F$ . Recall that we assumed  $0 \in \sigma_F$ . Let  $A_n$  denote the atoms of  $\langle \phi \rangle_n^b$  and let  $A_n^R$  denote the atoms of  $\langle \phi^R \rangle_n^b$ . For later reference, let us note that  $\langle \phi^R \rangle_n^b$  has three types of clauses:

1.  $\langle C_i(\bar{x}/\bar{a}) \rangle_n'$  for each  $i \in \{1, \dots, s\}$  and  $\bar{a} \in [n]^k$ ,
2.  $\langle \neg U(a_1) \vee \dots \vee \neg U(a_{r_F}) \vee \neg F(\bar{a}) = a \vee U(a) \rangle_n'$  for each  $F \in \sigma_F$ ,  $\bar{a} \in [n]^{r_F}$ , and  $a \in [n]$ ,
3.  $\bigvee_{b=0}^{|n|-1} F_{\bar{a};a}^{(1-\text{bit}(b,a))}$  for each  $F \in \sigma_F$ ,  $\bar{a} \in [n]^{r_F}$ , and  $a \in [2^{|n|}] \setminus [n]$ .

We need one technical definition (following [8]). Both propositional atoms  $R_{\bar{a}}$  and  $F_{\bar{a};b}$  are said to *mention* the elements in  $\bar{a}$ . Note that e.g.  $F_{2123;4}$  is *not* said to mention 4. More generally, a propositional formula in the atoms  $A_n$  *mentions* an element of  $[n]$  if so does an atom occurring in it. The following is easy to see but quite crucial for our argument:

**Fact 3.** *Let  $r_\phi$  be the maximal arity of a symbol from  $\sigma$  and at least 1. There is  $d_\phi \geq 1$  such that for every  $n \geq 2$  and every  $B \subseteq [n]$  there are at most  $d_\phi \cdot |B|^{r_\phi} \cdot \log n$  many atoms that mention only elements from  $B$ .*

For the rest of this section, let  $r_\phi$  and  $d_\phi$  denote the integers from this fact.

## 5.1 Formula dimension

We introduce the combinatorial objects underlying our argument, that we call 2-hypergraphs. Intuitively, these are families of hypergraphs on the same vertex set. Formally, a (finite) *2-hypergraph*  $\mathcal{H}$  is a pair  $(V, E)$  for a finite nonempty set  $V$  and  $E \subseteq P(P(V))$  is a set of *2-hyperedges*. We say  $D \subseteq E$  *touches*  $e \in E$  if there is  $x \in e$  such that  $x \subseteq \bigcup_{e \in D} \bigcup_{y \in e} y$ . The *dimension* of  $\mathcal{H}$ , denoted by  $\dim(\mathcal{H})$ , is the maximal  $\ell \in \mathbb{N}$  such that there is a sequence  $(e_0, \dots, e_{\ell-1})$  of 2-hyperedges such that no 2-hyperedge in it is touched by the preceding ones, that is, for all  $i < \ell$  the set  $\{e_j \mid j < i\}$  does not touch  $e_i$ ; in other words, for all  $x \in e_i$  we have

$$x \setminus \bigcup_{j < i} \bigcup_{y \in e_j} y \neq \emptyset.$$

Obviously, 2-hyperedges in such a sequence are pairwise distinct, so  $\ell \leq |E|$  and the maximum is well-defined.

Let  $n \geq 1$  and let  $F$  be a (propositional) DNF formula in the atoms  $A_n^R$ . The 2-hypergraph  $\mathcal{H}(F)$  has vertices  $[n]$  and for each term  $t$  of  $F$  a 2-hyperedge

$$e_t := \{x \subseteq [n] \mid x \text{ is the set of elements mentioned by some literal in } t\}.$$

We write  $\dim(F) := \dim(\mathcal{H}(F))$ .

## 5.2 Restrictions

Let  $n \geq 1$ . A *restriction (on  $A_n^R$ )* is a partial assignment  $\rho$  to  $A_n^R$ . For a formula  $F$  in the atoms  $A_n^R$  let  $F \upharpoonright \rho$  be obtained from  $F$  by replacing every atom in  $F$  and in  $\text{Dom}(\rho)$  by its truth value; then replace subformulas according to the rules  $\neg 0 \mapsto 1, \neg 1 \mapsto 0$  and  $G \vee 1 \mapsto 1, G \vee 0 \mapsto G$  and similarly for conjunctions involving some Boolean constant.

For a real number  $p \in (0, 1)$ , let  $\mathcal{D}_p$  be the probability distribution on restrictions defined by the following random experiment:

1. independently for each  $a \in [n]$ , call it *black* with probability  $p$  and *white* with probability  $(1 - p)$ ; set  $U_a$  to 1 if  $a$  is black, and to 0 if  $a$  is white.
2. independently for each atom mentioning some white element, set it to 1 with probability  $1/2$  and to 0 with probability  $1/2$ .

The *support* of  $\mathcal{D}_p$  is the set of restrictions with positive probability under  $\mathcal{D}_p$ .

## 5.3 Killing large disjunctions

Let  $n \geq 1$ . We write  $\exp(x)$  for  $e^x$ .

**Lemma 8.** *Let  $k \geq 1$  and let  $F$  be a  $k$ -DNF in the atoms  $A_n^R$ . Then*

$$\Pr_{\rho \sim \mathcal{D}_p} [F \upharpoonright \rho \neq 1] \leq \exp\left(-\dim(F) \cdot (1/2 - p/2)^k\right).$$

*Proof.* Let  $\ell := \dim(F)$  and  $(e_0, \dots, e_{\ell-1})$  be a sequence of 2-hyperedges (in  $\mathcal{H}(F)$ ) witnessing this. For  $i < \ell$  choose a  $k$ -term  $t_i$  in  $F$  such that  $e_{t_i} = e_i$ .

*Claim.* There exists a family  $(C_i)_{i < \ell}$  of subsets of  $[n]$  such that

1.  $|C_i| \leq k$  for all  $i < \ell$ ,
2.  $C_i \cap C_j = \emptyset$  for all  $i < j < \ell$ ,
3. every literal in  $t_i$  mentions an element of  $C_i$ .

*Proof of the claim.* To see this, let  $C_0$  contain for each literal in  $t_0$  an element mentioned by the literal. Because  $e_{t_1}$  is not touched by  $e_{t_0}$ , every literal in  $t_1$  mentions an element outside  $C_0$ ; let  $C_1$  contain such an element for each literal in  $t_1$ . And so on.  $\dashv$

For  $i < \ell$  define the following event  $E_i$ : all elements in  $C_i$  are white under  $\rho$  and the random choice in (2) is such that every literal in  $t_i$  becomes true. Note

$$\Pr_{\rho \sim \mathcal{D}_p} [E_i] \geq (1/2 \cdot (1-p))^k.$$

These events are mutually independent. Furthermore, for every  $i < \ell$  we have  $E_i \subseteq \{(\bigvee_{i < \ell} t_i) \mid \rho = 1\}$ , that is,  $\{(\bigvee_{i < \ell} t_i) \mid \rho \neq 1\} \subseteq \bigcap_{i < \ell} E_i^c$ , where  $E_i^c$  denotes the complement of  $E_i$ . Thus

$$\begin{aligned} \Pr_{\rho \sim \mathcal{D}_p} [F \mid \rho \neq 1] &\leq \Pr_{\rho \sim \mathcal{D}_p} \left[ \left( \bigvee_{i < \ell} t_i \right) \mid \rho \neq 1 \right] \leq \prod_{i < \ell} \Pr_{\rho \sim \mathcal{D}_p} [E_i^c] \\ &\leq \left( 1 - (1/2 \cdot (1-p))^k \right)^\ell \leq \exp \left( -\ell \cdot (1/2 - p/2)^k \right). \end{aligned}$$

□

## 5.4 Tree ranks

Decision trees represent formulas in the atoms of  $A_n$  for some  $n$ , that is, formulas without the atoms  $U_a$ .

A *decision tree* is a finite, rooted, ordered tree whose inner vertices are labeled by propositional atoms from  $A_n$  and whose leafs are labeled by 0 or 1. Each inner vertex is labeled with an atom from  $A_n$  and has two sons; by a *son* we mean an immediate successor on a branch (i.e. a path from the root to some leaf); further, we demand that no atom from  $A_n$  occurs twice on any branch. Since the tree is ordered we can distinguish between a *left* and a *right* son of an inner vertex with two sons. By a *0-branch* (*1-branch*) we mean a branch leading to a leaf labeled 0 (labeled 1).

Every path  $\pi$  from the root to some vertex corresponds to the following restriction that we also denote by  $\pi$ : if an atom from  $A_n$  occurs as a label of a vertex  $p$  in the path  $\pi$ , then the restriction sets this atom to 0 if the left son of  $p$  is in  $\pi$  and to 1 if the right son of  $p$  is in  $\pi$ ; if  $\pi$  contains no son of  $p$ , then the restriction does not evaluate the atom.

Let  $F$  be a formula in the atoms  $A_n$ . A decision tree  $T$  (*strongly*) *represents*  $F$  if  $F \upharpoonright \pi \equiv b$  (resp.  $F \upharpoonright \pi = b$ ) for every  $b \in \{0, 1\}$  and every  $b$ -branch  $\pi$  of  $T$ ; we say  $F$  *evaluates to*  $b$  *under*  $\pi$  if  $F \upharpoonright \pi \equiv b$ .

The *rank* of a decision tree is the maximum natural number  $c$  such that one of its branches mentions  $c$  elements (of  $[n]$ ). Here, a branch is said to *mention*  $a \in [n]$  if  $a$  is mentioned by some of its labels. For a formula  $F$  in the atoms  $A_n$  we let  $r^*(F)$  (resp.  $r(F)$ ) denote the minimal  $c$  such that there is a decision tree of rank  $c$  that (strongly) represents  $F$ .

The following lemmas are easy to verify.

**Lemma 9.** *Let  $s, t \geq 1$  be integers,  $T$  a decision tree of rank  $t$ , and  $F$  be a formula in the atoms  $A_n$ . If  $r^*(F \upharpoonright \pi) \leq s$  for every branch  $\pi$  of  $T$ , then  $r^*(F) \leq s + t$ .*

**Lemma 10.** *Let  $F$  and  $G$  be formulas in the atoms  $A_n$  and let  $T_F$  and  $T_G$  be decision trees of rank  $s_F$  and  $s_G$  that represent  $F$  and  $G$ , respectively. Then there exists a decision tree  $T$  of rank at most  $s_F + s_G$  that represents  $(F \wedge G)$  and such that every 0-branch of  $T$  extends some 0-branch of  $T_F$  or some 0-branch of  $T_G$ .*

## 5.5 Switching lemma

Let  $n \geq 2$  be a natural,  $p \in (0, 1)$  a real, and  $F$  be a formula in the atoms  $A_n$ . For every  $\rho$  in the support of  $\mathcal{D}_p$  the formula  $F \upharpoonright \rho$  has atoms in  $A_n$ , so the notation  $r(F \upharpoonright \rho)$  is explained.

Recall the constants  $r_\phi$  and  $d_\phi$  from Fact 3.

**Lemma 11.** *Let  $n \geq 2$  and  $(s_{k-1})_{k \geq 1}$  be a sequence of positive naturals. Set*

$$b_k := \sup_F \Pr_{\rho \sim \mathcal{D}_p} \left[ r(F \upharpoonright \rho) > r_\phi k \sum_{i < k} s_i \right],$$

where  $F$  ranges over  $k$ -DNFs in the atoms  $A_n^R$ . Then

- (a)  $b_1 < \exp(-s_0 \cdot (1/2 - p/2))$ ;
- (b)  $b_{k+1} < n^{d_\phi \cdot ((k+1)r_\phi s_k)^{r_\phi}} \cdot b_k + \exp\left(-s_k \cdot (1/2 - p/2)^{k+1}\right)$ .

*Proof.* We show (a) by distinguishing two cases. Let  $F$  be a 1-DNF, i.e. a clause. In case  $\dim(F) > s_0$ , then  $F \upharpoonright \rho \neq 1$  with probability  $< e^{-s_0(1/2-p/2)}$ ; if  $F \upharpoonright \rho = 1$  there is a rank 0 decision tree for  $F \upharpoonright \rho$ .

Otherwise,  $\dim(F) =: \ell \leq s_0$  and we find  $(e_0, \dots, e_{\ell-1})$  witnessing this. For  $e_i$  choose a term  $t_i$  in  $F$  such that  $e_{t_i} = e_i$ .

*Claim.*  $F$  mentions at most  $r_\phi \ell$  many elements.

*Proof of the claim.* For every term  $t$  of  $F$  the 2-hyperedge  $e_t$  is touched by  $\{e_0, \dots, e_\ell\}$ . But any  $e_t$  is a singleton containing the set of elements mentioned by  $t$ . Hence all these elements are contained in the set of elements mentioned by some  $t_i, i < \ell$ , and this set has cardinality at most  $r_\phi \ell$ .  $\dashv$

Thus  $F \upharpoonright \rho$  mentions at most  $r_\phi s_0$  many elements and we get (with probability 1) a trivial decision tree of this rank that queries all atoms in  $F \upharpoonright \rho$ .

In both cases, we get a decision tree of rank at most  $r_\phi s_0 = r_\phi k \sum_{i < k} s_i$  with probability larger than  $1 - \exp(-s_0(1/2 - p/2))$ .

We show (b) again by distinguishing two cases. Let  $F$  be a  $(k+1)$ -DNF. In case  $\dim(F) > s_k$ , then  $F \upharpoonright \rho \neq 1$  with probability less than  $\exp(-s_k \cdot (1/2 - p/2)^{k+1}) < b_{k+1}$ ; if  $F \upharpoonright \rho = 1$  we find a decision tree of rank 0.

Otherwise,  $\dim(F) =: \ell \leq s_k$  and we find  $(e_0, \dots, e_{\ell-1})$  witnessing this. For  $e_i$  choose a term  $t_i$  in  $F$  such that  $e_{t_i} = e_i$ . Let  $C$  be the set of elements mentioned by some  $t_i$  and note

$$|C| \leq (k+1)r_\phi s_k.$$

Define  $T_C$  to be the decision tree that successively queries all atoms that mention only elements of  $C$  (and, say, with all leafs labeled 0).

*Claim:* Let  $\rho$  be an arbitrary restriction. Assume that  $r((F \upharpoonright \pi) \upharpoonright \rho) \leq r_\phi k \sum_{i < k} s_i$  for every branch  $\pi$  of  $T_C$ . Then  $r(F \upharpoonright \rho) \leq r_\phi (k+1) \sum_{i < k+1} s_i$ .

*Proof of the claim.* For every  $\pi$  that is compatible with  $\rho$  (i.e.  $\rho \cup \pi$  is a restriction) we have

$$(F \upharpoonright \pi) \upharpoonright \rho = (F \upharpoonright \rho) \upharpoonright \pi = (F \upharpoonright \rho) \upharpoonright (\pi \setminus \rho),$$

and hence, by assumption,

$$r((F \upharpoonright \rho) \upharpoonright (\pi \setminus \rho)) \leq r_\phi k \sum_{i < k} s_i.$$

Let  $T_C^\rho$  be a decision tree with branches  $\pi \setminus \rho$  for  $\pi$  compatible with  $\rho$  and note that this tree has rank at most  $|C| \leq (k+1)r_\phi s_k$ . Now apply Lemma 9 with  $T := T_C^\rho$ ,  $t := (k+1)r_\phi s_k$ ,  $s := r_\phi k \sum_{i < k} s_i$  on the formula  $(F \upharpoonright \rho)$ . –

By the claim it suffices to show

$$\Pr_{\rho \sim \mathcal{D}_p} \left[ \exists \pi : r((F \upharpoonright \pi) \upharpoonright \rho) > r_\phi k \sum_{i < k} s_i \right] \leq n^{d_\phi \cdot ((k+1)r_\phi s_k)^{r_\phi}} \cdot b_k, \quad (2)$$

where  $\pi$  ranges over the branches of  $T_C$ .

Let  $\pi$  be an arbitrary branch of  $T_C$ . Every term  $t$  of  $F$  is touched by  $\{e_0, \dots, e_{\ell-1}\}$ . This implies that  $t$  contains a literal that only mentions elements from  $C$ . This literal is evaluated by  $\pi$ . Hence  $(F \upharpoonright \pi)$  is a  $k$ -DNF and thus, by definition of  $b_k$ ,

$$\Pr_{\rho \sim \mathcal{D}_p} \left[ r((F \upharpoonright \pi) \upharpoonright \rho) > r_\phi k \sum_{i < k} s_i \right] \leq b_k. \quad (3)$$

By Fact 3 the tree  $T_C$  has at most

$$2^{d_\phi \cdot |C|^{r_\phi} \cdot \log n} \leq n^{d_\phi \cdot ((k+1)r_\phi s_k)^{r_\phi}}$$

many branches  $\pi$ . Thus (2) follows from (3) and the union bound. □

**Corollary 1.** *Let  $n \geq 2, k \geq 1$  and  $s_0, \dots, s_{k-1} \geq 1$  be naturals, and let  $F$  be a  $k$ -DNF in the atoms  $A_n^R$ . Then*

$$\Pr_{\rho \sim \mathcal{D}_p} \left[ r(F \upharpoonright \rho) > r_\phi k \sum_{i < k} s_i \right] < \sum_{i < k} \exp \left( -s_i \cdot (1/2 - p/2)^{i+1} + \ln(n) \cdot d_\phi \cdot \sum_{j=i+1}^{k-1} ((j+1)r_\phi s_j)^{r_\phi} \right). \quad (4)$$

## 5.6 Rank lower bound

A *semantic refutation* of a CNF formula  $F$  is a sequence of formulas  $F_1, F_2, \dots, F_t$  in the variables of  $F$  such that  $F_t$  is unsatisfiable and every  $F_i$  is either logically equivalent to a clause from  $F$ , or a logical consequence of the conjunction of exactly two previous formulas in the sequence.

**Lemma 12.** *Let  $n, m \geq 1$  and  $0 < p < 1$ , and let  $\rho$  be a restriction in the support of  $\mathcal{D}_p$  such that exactly  $m$  elements of  $[n]$  are black under  $\rho$ . If  $F_1, F_2, \dots, F_t$  is a semantic refutation of  $\langle \phi^R \rangle_n^b \upharpoonright \rho$ , then there exists  $i \in \{1, \dots, t\}$  such that  $r^*(F_i) > (m/11d_\phi)^{1/r_\phi}$ .*

*Proof.* Write  $F := \langle \phi^R \rangle_n^b \upharpoonright \rho$  and  $C := \lfloor (m/11d_\phi)^{1/r_\phi} \rfloor$  and let  $F_1, F_2, \dots, F_t$  be a semantic refutation of  $F$ . Note each  $F_i$  is a formula in the variables  $A_n$ . Let  $T_i$  be a decision tree that represents  $F_i$  and assume, for the sake of contradiction, that every  $T_i$  has rank at most  $C$ .

Let  $X$  be the set of black elements of  $\rho$  and note  $|X| = m$ . Recall that  $M$  is an infinite model of  $\phi$ . For  $B \subseteq M$  let  $\partial B$  be the set of all values of functions in  $M$  taken on  $B$ , that is,

$$\partial B := \bigcup_{F \in \sigma_F} \text{Im}(F^M \upharpoonright B).$$

The following is easy to see.

*Claim 0.*  $|\partial B| \leq d_\phi |B|^{r_\phi}$  for every  $B \subseteq M$ .

A pair  $(f, g)$  of partial injections from  $X$  into  $M$  is *good* if  $f \subseteq g$  and  $\text{Im}(g) = \text{Im}(f) \cup \partial \text{Im}(f)$ . For example,  $(\emptyset, \emptyset)$  is good. For every good pair  $(f, g)$ , let  $\rho(f, g)$  be the minimal restriction on  $A_n^R$  that extends  $\rho$  and is such that

1.  $\rho(f, g)(R_{\bar{a}}) = 1$  if  $f(\bar{a}) \in R^M$ , for each  $R \in \sigma_R$  and  $\bar{a} \in \text{Dom}(f)^{r_R}$ ,
2.  $\rho(f, g)(R_{\bar{a}}) = 0$  if  $f(\bar{a}) \notin R^M$ , for each  $R \in \sigma_R$  and  $\bar{a} \in \text{Dom}(f)^{r_R}$ ,
3.  $\rho(f, g)(F_{\bar{a};b}) = \text{bit}(b, g^{-1}(F^M(f(\bar{a}))))$ , for each  $F \in \sigma_F$ ,  $\bar{a} \in \text{Dom}(f)^{r_F}$ , and  $b \in [|n|]$ .

Observe that the third case is well-defined as  $F^M(f(\bar{a}))$  belongs to  $\partial \text{Im}(f)$ , which is included in  $\text{Im}(g)$ . For example,  $\rho(\emptyset, \emptyset) = \rho$ .

*Claim 1.* If  $(f, g)$  is good, then  $\rho(f, g)$  does not evaluate any clause of  $F$  to 0.

*Proof of Claim 1.* Let  $(f, g)$  be good. The CNF formula  $F$  has three kinds of non-trivial clauses (see the beginning of this section):

1.  $\langle C_i(\bar{x}/\bar{a}) \rangle'_n$  for black  $\bar{a}$ ,
2.  $\langle \neg F(\bar{a}) = a \rangle'_n = \bigvee_{b=0}^{|n|-1} F_{\bar{a};a}^{(1-\text{bit}(b,a))}$  for black  $\bar{a}$  and white  $a$ .
3.  $\bigvee_{b=0}^{|n|-1} F_{\bar{a};a}^{(1-\text{bit}(b,a))}$  for black  $\bar{a}$  and  $a \in [2^{|n|}] \setminus [n]$

Observe that if  $\rho(f, g)$  evaluates a clause of the form  $\bigvee_{b=0}^{|n|-1} F_{\bar{a};a}^{(1-\text{bit}(b,a))}$  to 0, then  $\bar{a}$  is from  $\text{Dom}(f)$  and  $F^M(f(\bar{a})) = g(a)$ ; in particular  $a$  belongs to  $\text{Dom}(g) \subseteq X \subseteq [n]$  and is black. This already implies that  $\rho(f, g)$  cannot falsify a clause of type 2. or 3..

Assume  $\rho(f, g)$  evaluates a clause  $\langle C_i(\bar{x}/\bar{a}) \rangle'_n$  to 0 where  $\bar{a}$  is black. We claim that there is a (first-order) assignment falsifying  $C_i(\bar{x})$  in  $M$ , contradicting  $M \models \phi$ . In fact, we can take any assignment  $\bar{x} \mapsto g'(\bar{a})$  where  $g'$  is an arbitrary injection that extends  $g$  and contains  $\bar{a}$  in its domain. By assumption  $\langle C_i(\bar{x}/\bar{a}) \rangle'_n$  cannot be tautological, so all equality literals are false under any assignment realizing the same equality type as  $\bar{a}$ , and in particular this is true for  $\bar{x} \mapsto g'(\bar{a})$  (since  $g'$  is injective). A literal of the form  $R(\bar{y})$  or  $\neg R(\bar{y})$  evaluates under  $\bar{x} \mapsto g'(\bar{a})$  according to  $\rho(f, g)$  and hence is false. Finally, the claim that literals of the form  $\neg F(\bar{y}) = z$  are false follows from the above.  $\dashv$

*Claim 2.* If  $(f, g)$  is good,  $|\text{Dom}(f)| \leq 3C$  and  $I \subseteq X$  with  $|I| \leq C$ , then there is a good  $(f', g')$  such that  $f' \supseteq f$  and  $g' \supseteq g$  and  $\text{Dom}(f') = \text{Dom}(f) \cup I$ .

*Proof of Claim 2.* Write  $I = I_0 \cup I_1$  with  $I_0 := I \cap \text{Dom}(g)$  and  $I_1 := I \setminus \text{Dom}(g)$ . Define  $f' := f \cup (g \upharpoonright I_0) \cup f^*$  where  $f^*$  is some injection from  $I_1$  into  $A \setminus \text{Im}(g)$ . Note that  $|\text{Dom}(f')| \leq |\text{Dom}(f)| + |I| \leq 4C$  and thus by Claim 0 we have

$$|\partial \text{Im}(f')| \leq d_\phi(4C)^{r_\phi}. \quad (5)$$

By Claim 0,  $|\partial \text{Im}(f)| \leq d_\phi(3C)^{r_\phi}$  and thus  $|\text{Dom}(g)| \leq 3C + d_\phi(3C)^{r_\phi}$ . Hence

$$|\text{Dom}(g) \cup \text{Dom}(f')| \leq |\text{Dom}(g)| + |I_1| \leq d_\phi(3C)^{r_\phi} + 4C. \quad (6)$$

By definition of  $C$  we have  $m \geq d_\phi 11^{r_\phi} C^{r_\phi} \geq d_\phi 3^{r_\phi} C^{r_\phi} + d_\phi 4^{r_\phi} C^{r_\phi} + 4C$  and thus

$$m - d_\phi(3C)^{r_\phi} - 4C \geq d_\phi(4C)^{r_\phi}.$$

With (5) and (6) it follows that there exists a partial bijection  $g^*$  from  $X \setminus (\text{Dom}(g) \cup \text{Dom}(f'))$  onto  $\partial \text{Im}(f') \setminus \text{Im}(g)$ . We define  $g' := g \cup g^*$ . +

*Claim 3.* If there is a good  $(f, g)$  such that  $F_i \upharpoonright \rho(f, g) \equiv 0$  and  $|\text{Dom}(f)| \leq C$ , then there are  $j < i$  and a good  $(f', g')$  such that  $F_j \upharpoonright \rho(f', g') \equiv 0$  and  $|\text{Dom}(f')| \leq C$ .

*Proof of Claim 3.* By Claim 1,  $F_i$  is not logically equivalent to a clause of  $F$ . Hence there are  $j, j' < i$  such that  $F_i$  is a logical consequence of  $(F_j \wedge F_{j'})$ . Choose trees  $T_j$  and  $T_{j'}$  of rank at most  $C$  for  $F_j$  and  $F_{j'}$  respectively. Choose  $T$  according to Lemma 10, and note that  $T$  has rank at most  $2C$ .

*Subclaim.* There is a branch  $\pi$  in  $T$  and a good  $(f_\pi, g_\pi)$  such that

- (a)  $f_\pi \supseteq f$  and  $g_\pi \supseteq g$ ,
- (b)  $\rho(f_\pi, g_\pi)$  extends  $\pi$ .
- (c)  $\text{Dom}(f_\pi) \setminus \text{Dom}(f) \subseteq I_\pi$ ,

where  $I_\pi$  is the set of elements mentioned by a label of some vertex in  $\pi$ .

The subclaim implies Claim 3: choose  $\pi$  and  $(f_\pi, g_\pi)$  according to the Subclaim. If  $\pi$  were a 1-branch, then  $(F_j \wedge F_{j'}) \upharpoonright \rho(f_\pi, g_\pi) \equiv 1$  since  $\rho(f_\pi, g_\pi)$  extends  $\pi$ . Since  $\rho(f_\pi, g_\pi)$  extends  $\rho(f, g)$  we have  $F_i \upharpoonright \rho(f_\pi, g_\pi) \equiv 0$  and this contradicts the fact that  $(F_j \wedge F_{j'})$  logically implies  $F_i$ . Thus  $\pi$  is a 0-branch. By choice of  $T$ , then  $\pi$  extends a 0-branch  $\pi'$  of  $T_j$  or of  $T_{j'}$ ; say  $\pi'$  mentions the elements  $I_{\pi'}$ . Then  $|I_{\pi'}| \leq C$  and  $(f_\pi \upharpoonright I_{\pi'}, g_\pi \upharpoonright \partial \text{Im}(f_\pi \upharpoonright I_{\pi'}))$  is a good pair as desired.

We are left to prove the subclaim. Starting at the root of  $T$  and the pair  $(f, g)$ , we show how to extend a path in  $T$  by one more vertex and a new good pair until we reach a leaf of  $T$ . Assume we have a path  $\pi$  in  $T$  and a good pair  $(f_\pi, g_\pi)$  such that the conditions (a), (b), (c) hold true. Note that the trivial path that is formed by the root only together with the pair  $(f, g)$  satisfies the conditions. Further assume that  $\pi$  leads to a vertex  $p$  that is not a leaf of  $T$ . Then  $p$  is labeled by an atom from  $A_n$ . We choose  $(f', g')$  according to Claim 2 for  $(f_\pi, g_\pi)$  and  $I$  the set of elements mentioned in the label of  $p$ . Then the label of  $p$  evaluates under  $\rho(f', g')$  and we enlarge  $\pi$  by the son  $p'$  of  $p$  corresponding to this truth value. We take  $(f', g')$  for  $(f_{\pi p'}, g_{\pi p'})$ . Observe that Claim 2 is applicable because  $(f_\pi, g_\pi)$  satisfies condition (c) and thus  $|\text{Dom}(f_\pi)| \leq |\text{Dom}(f)| + |I_\pi| \leq 3C$  since  $|I_\pi| \leq 2C$ . +

Note that  $(\emptyset, \emptyset)$  is a good pair such that  $F_i \upharpoonright \rho(\emptyset, \emptyset) \equiv 0$  and  $|\text{Dom}(\emptyset)| = 0 \leq C$ . Thus Claim 1 and Claim 3 are contradictory. □

## 5.7 Proof-size lower bounds

Remind that we fixed a standardized universal first-order formula  $\phi$  that has some infinite model. We are ready to prove our lower bound:

**Theorem 5.** *For every  $k \geq 1$  there is  $d_k \geq 1$  such that for every large enough  $n$ , every  $R(k)$  refutation of  $\langle \phi^R \rangle_n^b$  has length at least  $2^{n^{1/d_k}}$ .*

*Proof.* Apply Corollary 1 with  $p = 1/2$  and the sequence  $s_0, s_1, \dots$  defined by

$$s_0 := \frac{n^{1/r_\phi}}{(44d_\phi)^{1/r_\phi} \cdot r_\phi k^2} \quad (7)$$

$$s_{i+1} := \frac{(1/2 \cdot 1/4^{i+1})^{1/r_\phi} \cdot s_i^{1/r_\phi}}{(d_\phi \ln(n)(k-1-i))^{1/r_\phi} \cdot kr_\phi}. \quad (8)$$

Note that this sequence is decreasing. By (8),

$$\begin{aligned} \ln(n) \cdot d_\phi \cdot \sum_{j=i+1}^{k-1} ((j+1)r_\phi s_j)^{r_\phi} &\leq \ln(n) \cdot d_\phi \cdot (k-1-i) \cdot (kr_\phi s_{i+1})^{r_\phi} \\ &\leq (1/2 \cdot 1/4^{i+1}) s_i. \end{aligned}$$

Then the probability bound in Corollary 1 is at most

$$\sum_{i < k} \exp(-(1/2 \cdot 1/4^{i+1}) s_i) \leq k e^{-s_k/4^{k+1}}.$$

Rewrite (8) by  $s_{i+1} = q_i \cdot s_i^{1/r_\phi}$  and note that for all  $i \leq k$  we have  $q_i \geq 4^{-(k+2)/r_\phi} \cdot (d_\phi k \ln(n))^{-1/r_\phi}$ .  $(kr_\phi)^{-1} \geq c^{-k/r_\phi} \cdot \ln(n)^{-1/r_\phi} =: p_k$  for some constant  $c$ . We can suppose that  $r_\phi \geq 2$ . Then

$$s_k \geq p_k \cdot s_{k-1}^{1/r_\phi} \geq p_k \cdot p_k^{1/r_\phi} \cdot s_{k-2}^{1/r_\phi^2} \geq \dots \geq p_k^2 \cdot s_0^{1/r_\phi^k}.$$

Further, by (7),

$$r_\phi k \sum_{i < k} s_i \leq r_\phi k^2 \cdot s_0 \leq n^{1/r_\phi} / (44d_\phi)^{1/r_\phi} = ((n/4) / 11d_\phi)^{1/r_\phi}.$$

In summary, Corollary 1 implies

$$\Pr_{\rho \sim \mathcal{D}_p} [r^*(\alpha \upharpoonright \rho) > ((n/4) / 11d_\phi)^{1/r_\phi}] < \exp(\ln(k) - p_k^2 \cdot s_0^{1/r_\phi^k} \cdot 4^{-k-1}). \quad (9)$$

By (7), we find for every  $k \geq 1$  some  $c_k \in \mathbb{N}$  such that for large enough  $n$  the above is bounded by  $e^{-n^{1/c_k}}$ .

Assume there is a  $R(k)$ -refutation of  $\langle \phi^R \rangle_n^b$  of length at most  $e^{n^{1/c_k-1}}$ . Chernoff bounds imply that with probability at least  $1 - e^{-n/16}$  there are at least  $n/4$  elements black under  $\rho \sim \mathcal{D}_{1/2}$ . Then  $e^{n^{1/c_k-1}} \cdot e^{-n^{1/c_k}} + e^{-n/16} = e^{-1} + e^{-n/16} < 1$ . Thus there exists some  $\rho$  in the support of  $\mathcal{D}_{1/2}$  that has  $m \geq n/4$  black elements and is such that for every line  $\beta$  of the refutation

$$r^*(\beta \upharpoonright \rho) \leq (m/11d_\phi)^{1/r_\phi}.$$

If one applies such a  $\rho$  to every line of the refutation, one gets a semantic refutation of  $\langle \phi^R \rangle_n^b \upharpoonright \rho$  that contradicts Lemma 12.  $\square$

We are ready to prove our main result:

*Proof.* (of Theorem 1) Statement (b) follows from the previous theorem. To prove (a) assume  $\phi$  does not have an infinite model. Then neither does  $\phi^R$ , so by compactness there exists  $n_0 \in \mathbb{N}$  such that for all  $n \geq n_0$ ,  $\phi^R$  does not have a model of size  $n$ . By Riis' Theorem 3 we get  $\langle \phi^R \rangle_n^u \vdash_{1,*}^{p(n)} \square$  for some polynomial  $p$ . As this proof is in tree-like  $R(1)$ , we can assume without loss of generality that it only uses cuts. Replace in this proof each occurrence of  $F_{\bar{a};a}$  by  $\bigwedge_{b=0}^{|n|-1} F_{\bar{a};b}^{\text{bit}(b,a)}$ . This results in a tree-like  $R(\log)$  refutation of the binary encoding using the substitution instances of the functional clauses as additional axioms. The short functional clauses are (as noted in the proof of Lemma 4) weakenings of axioms, and the long ones have short tree-like  $R(\log)$  refutations according Remark 3. We conclude that  $\langle \phi^R \rangle_n^b \vdash_{\log,*}^{p'(n)} \square$  for some polynomial  $p'$ . By Theorem 2 this implies  $\langle \phi^R \rangle_n^b \vdash_1^{p''(n)} \square$  for some polynomial  $p''$ .  $\square$

We finish by deriving two lower bounds on  $k$  in “short”  $R(k)$  refutations.

**Theorem 6.** *Let  $(k_n)_{n \geq 1}$  be a sequence of positive integers. There is  $\epsilon > 0$  such that for all  $d \geq 1$  and all sufficiently large  $n$  the following holds: if there is an  $R(k_n)$  refutation of  $\langle \phi^R \rangle_n^b$  of length at least  $2^{(\log n)^d}$ , then  $k_n \geq \epsilon \cdot \log \log n$ .*

*Proof.* As seen in the previous proof, there are no  $R(k)$ -refutations of  $\langle \phi^R \rangle_n^b$  length  $S$  in case  $k$  and  $S$  are such that

$$-\ln(k) + p_k^2 \cdot s_0^{-1/r_\phi^k} > \log S.$$

Recall the definition of  $p_k$  and suppose the constant  $c$  is chosen to be at least 4. Then the above inequality reads

$$-\ln k + \ln(n)^{-1} \cdot s_0^{1/r_\phi^k} \cdot c^{-2k} > \log S.$$

We verify this for  $S \leq 2^{(\log n)^d}$ , where  $d \geq 1$  is arbitrary, and  $k \leq (1/2 \log r_\phi) \cdot \log \log n$  (i.e. we take  $\epsilon := 1/2 \log r_\phi$ ). For the sake of contradiction assume the inequality fails. Then

$$\ln(n)^{-1} \cdot s_0^{1/r_\phi^k} \cdot c^{-2k} \leq (\log n)^{d+1}.$$

This implies  $\log s_0 \cdot r_\phi^{-k} \leq d' \cdot \log \log n$  for a suitable constant  $d' \in \mathbb{N}$ . But this in turn is false for sufficiently large  $n$ , because  $\log s_0 \geq \Omega(\log n)$  and  $r_\phi^k \leq (\log n)^{1/2}$ .  $\square$

Call the vocabulary  $\sigma$  *unary* if all symbols in  $\sigma$  have arity at most 1. For example, the pigeonhole principle can be formulated as a universal sentence in a (functional) unary vocabulary. Weak pigeonhole principles are naturally formulated with function symbols of higher arity. Segerlind et al. [18] proved that a subexponential size  $R(k)$  proof of the weak  $2n$  to  $n$  pigeonhole principle (in a relational formalization) needs  $k > \sqrt{\log n / \log \log n}$ . We match this lower bound on  $k$  in general for unary vocabularies.

**Theorem 7.** *Assume that the vocabulary of  $\phi$  is unary and let  $(k_n)_{n \geq 1}$  be a sequence of positive integers. There is  $\epsilon > 0$  such that for all sufficiently large  $n$  the following holds: if there is an  $R(k_n)$  refutation of  $\langle \phi^R \rangle_n^b$  of length at least  $2^{n^\epsilon}$ , then  $k_n \geq \epsilon \cdot \sqrt{\log n}$ .*

*Proof.* Observe that now we have  $r_\phi = 1$ . Following the proof of Theorem 5 we get

$$s_k \geq p_k^{-k} \cdot s_0 \geq c^{-k^2} \cdot (\ln n)^{-k}.$$

Then, for  $k \leq o(\sqrt{\log n})$ , we arrive at (9) in the form

$$\Pr_{\rho \sim \mathcal{D}_p} [r^*(\alpha \mid \rho) > ((n/4) / 11d_\phi)] < \exp(\ln(k) - c^{-k^2} \cdot (\ln n)^{-k} \cdot s_0 \cdot 4^{-k-1}).$$

As in the proof of Theorem 5 we can conclude that  $R(k)$  does not have size  $S$  refutations of  $\langle \phi^R \rangle_n^b$  in case  $k$  and  $S$  are such that  $-\ln(k) + c^{-k^2} \cdot (\ln n)^{-k} \cdot s_0 > \log S$ . It thus suffices to verify this for  $S \leq 2^{n^{o(1)}}$  and  $k \leq o(\sqrt{\log n})$  and sufficiently large  $n$ . Note that in this case  $\log S + \ln(k) \leq n^{o(1)}$ . Hence, it suffices to verify

$$\log s_0 > o(1) \cdot \log n + k^2 \log c + k \cdot \log \log n. \quad (10)$$

For the second term on the right hand side we have  $k^2 \log c \leq o(1) \cdot \log n$  and for the third we have  $k \cdot \log \log n \leq o(1) \cdot \sqrt{\log n} \cdot \log \log n \leq o(1) \cdot \log n$ ; in summary the right hand side is bounded by  $o(1) \cdot \log n$ . But for the left hand side we have  $\log s_0 \geq \Omega(\log n)$ . This implies (10) for sufficiently large  $n$ .  $\square$

## References

- [1] A. Atserias, M. L. Bonet, and J. L. Esteban. Lower bounds for the weak pigeonhole principle and random formulas beyond resolution. *Information and Computation* 176(2):136-152, 2002.
- [2] A. Atserias and V. Dalmau. A combinatorial characterization of resolution width, *Journal of Computer and System Sciences*, 74(3):323-334, 2008.
- [3] E. Ben-Sasson and A. Wigderson. Short Proofs are Narrow- Resolution made Simple. *Journal of the ACM* 48(2), 2001.
- [4] M. L. Bonet and N. Galesi. Optimality of Size-width Tradeoffs for Resolution. *Computational Complexity* 10(4), pp. 261-276, 2001.
- [5] S. Cook and R. Reckhow. The relative efficiency of propositional proof systems. *The Journal of Symbolic Logic* 44:36-50, 1979.
- [6] S. Dantchev. Relativisation provides natural separations for resolution-based proof systems. *Computer Science—Theory and Applications: First International Computer Science Symposium in Russia, Lecture Notes in Computer Science* 3967, pp. 147-158, 2006.
- [7] S. Dantchev and B. Martin. The limits of tractability in resolution-based propositional proof systems. 6th Conference on Computability in Europe. *Lecture Notes in Computer Science* 6158, pp 98-107, 2010.
- [8] S. Dantchev and S. Riis. On relativization and complexity gap for resolution-based proof systems. 12th Annual Conference of the EACSL Computer Science Logic, Vienna, Austria, Springer, pp. 142-154, 2003.

- [9] J. L. Esteban, N. Galesi and J. Messner. On the complexity of resolution with bounded conjunctions. *Theoretical Computer Science* 321: 347-370,2004.
- [10] J. Krajíček. Lower bounds to the size of constant-depth propositional proofs. *Journal of Symbolic Logic* 59(1):73-86, 1994.
- [11] J. Krajíček. *Bounded Arithmetic, Propositional Logic, and Complexity Theory*. Encyclopedia of Mathematics and its Applications, Cambridge University Press, 1995.
- [12] J. Krajíček. On the weak pigeonhole principle. *Fundamenta Mathematicae* 170:123-140, 2001.
- [13] J. Krajíček. Combinatorics of first order structures and propositional proof systems. *Archive for Mathematical Logic*, 43(4):427-441, 2004.
- [14] J. Krajíček. A note on propositional proof complexity of some Ramsey-type statements *Archive for Mathematical Logic* 50(1-2): 245-255, 2011.
- [15] P. Pudlák. Proofs as games. *American Mathematical Monthly*, pp. 541-550, June-July 2000.
- [16] S. Riis. Making infinite structures finite in models of second order bounded arithmetic. in: *Arithmetic, proof theory and computational complexity*, Oxford University Press, pp. 289-319, 1993.
- [17] S. Riis. A complexity gap for tree-resolution. *Computational Complexity* 10:179-209, 2001.
- [18] N. Segerlind, S. Buss and R. Impagliazzo. A Switching Lemma for Small Restrictions and Lower Bounds for k-DNF Resolution. *SIAM Journal on Computing* 33(5): 1171-1200, 2004.
- [19] G. Stålmarck. Short Resolution Proofs for a Sequence of Tricky Formulas. *Acta Informatica* 33(3), pp. 277-280, 1996.
- [20] A. Razborov. Pseudorandom generators hard for k-DNF resolution and polynomial calculus resolution. Manuscript, 2003.