Pseudorandom Generators Hard for k-DNF Resolution and Polynomial Calculus Resolution

Alexander A. Razborov *

July 25, 2003

Abstract

A pseudorandom generator $G_n : \{0,1\}^n \to \{0,1\}^m$ is *hard* for a propositional proof system P if (roughly speaking) P can not efficiently prove the statement $G_n(x_1, \ldots, x_n) \neq b$ for *any* string $b \in \{0,1\}^m$. We present a function ($m \geq 2^{n^{\Omega(1)}}$) generator which is hard for $\mathrm{Res}(\epsilon \log n)$; here $\mathrm{Res}(k)$ is the propositional proof system that extends Resolution by allowing k-DNFs instead of clauses.

As a direct consequence of this result, we show that whenever $t \geq n^2$, every $\mathrm{Res}(\epsilon \log t)$ proof of the principle $\neg Circuit_t(f_n)$ (asserting that the circuit size of the Boolean function f_n in n variables is greater than t) must have size $\exp(t^{\Omega(1)})$. In particular, $\mathrm{Res}(\log n)$ does not possess efficient proofs of $\mathbf{NP} \not\subseteq \mathbf{P}/poly$. Similar results hold also for the system PCR (the natural common extension of Polynomial Calculus and Resolution) when the characteristic of the ground field is different from 2.

As a by-product, we also improve on the small restriction switching lemma from [SBI02] by removing a square root from the final bound. This in particular implies that the (moderately) weak pigeonhole principle PHP_n^{2n} is hard for $\mathrm{Res}(\epsilon \log n / \log \log n)$.

*Institute for Advanced Study, Princeton, US, on leave from Steklov Mathematical Institute, Moscow, Russia, razborov@ias.edu. Supported by The State of New Jersey and by the RFBR grant 02-02-01290.

1. Introduction

Propositional proof complexity is an area of study that has seen a rapid development over the last decade. It plays as important a role in the theory of feasible proofs as the role played by the complexity of Boolean circuits in the theory of efficient computations. And in most cases the basic question of propositional proof complexity boils down to this. Given a mathematical statement encoded as a propositional tautology ϕ and a class of admissible mathematical proofs formalized as a propositional proof system P, what is the minimal possible complexity of a P-proof of ϕ?

1.1. General overview

For most "interesting" propositional proof systems P one can easily define the accompanying (non-uniform) complexity class $\mathcal{C}_P$ typically consisting of functions computable by lines allowed in efficient P-proofs. This correspondence leads to a classification of propositional proof systems that is somewhat imprecise and potentially (and hopefully) time-dependent but nonetheless very instructive. Namely, we call a propositional proof system P *weak* if we (currently) know how to prove super-polynomial lower bounds for the accompanying circuit class $\mathcal{C}_P$ and *strong* otherwise.

There is a steady progress in studying the complexity of proofs in weak proof systems surveyed e.g. in [Urq95, Kra95, Raz96, BP98, Pud98, Raz02a].

For strong proof systems the current situation is by far more miserable. Although there are no rigorous results along these lines (and, moreover, this feeling is not universal – see e.g. [Kra02]), the empirical evidence strongly suggests that lower bounds for a proof system P are even harder to attain than computational lower bounds for the companion class $\mathcal{C}_P$. Therefore, with our current understanding, we can not apparently even hope to show lower bounds for systems like Frege or Extended Frege without first making a major breakthrough in the complexity theory.

A more accessible task which (in the author's opinion) is almost as interesting would be to show at least that proof complexity lower bounds are *at most* as hard as comparable problems in the computational world. Let us (informally) identify this task as

PROVING LOWER BOUNDS FOR STRONG PROOF SYSTEMS P LIKE FREGE OR EXTENDED FREGE MODULO ANY HARDNESS ASSUMPTION IN THE PURELY COMPUTATIONAL WORLD, HOWEVER STRONG BUT STILL NATURAL AND BELIEVABLE

This task will be referred to as *proving conditional lower bounds* (for the proof system P). It should be remarked in this respect that $\mathbf{NP} \neq co-\mathbf{NP}$ implies lower bounds for *any* propositional proof system whatsoever. Therefore, *purely computational* above refers to the demand that the assumption itself should speak only about computations and should not attempt to restrict the power of proofs even in a disguised form.

One extremely exciting and, in a sense, model approach to this task was gradually developed in the sequence of papers [Kra94, Raz95b, BPR97, Kra97a, Pud97] and finally became known as *Efficient Interpolation Property* (EIP in what follows). EIP was shown to be true for some weak proof systems and it was also remarked that for every proof system (be it weak or strong) EIP implies conditional lower bounds. Unfortunately, it turned out rather soon [KP98, BPR00] that neither Frege nor Extended Frege have Efficient Interpolation modulo (somewhat ironically) hardness assumptions of the same sort that are needed to prove conditional lower bounds for proof systems *with* EIP.

This omnipresent hardness assumption is nothing else as the existence of pseudorandom generators (arbitrary or specific) which also turns out to be the main primitive of the modern cryptography. After the (apparent) failure of the efficient interpolation approach, it was independently proposed in [Kra01a, ABRW00] to employ pseudorandom generators for proving conditional lower bounds in a more direct manner. On the conceptual level, a mapping $G_n : \{0,1\}^n \to \{0,1\}^m$, where $m > n$, is called *hard* for a propositional proof system P if P can not efficiently prove the (properly encoded) statement $G_n(x_1, \ldots, x_n) \neq b$ for *any* string $b \in \{0,1\}^m$. Since $m > n$, at least for half of all b's this statement is a tautology. Therefore, conditional lower bounds for a proof system P follow from the following task:

PROVE THAT FOR A REASONABLE CLASS OF MAPPINGS $G_n : \{0,1\}^n \to \{0,1\}^m$ WITH $m > n$, THEIR HARDNESS IN ANY REASONABLE COMPUTATIONAL OR COMBINATORIAL SENSE IMPLIES HARDNESS FOR P.

that will be referred to as the *generator approach* (to conditional lower bounds for P).

The generator approach certainly does not work for *arbitrary* mappings G_n believed to be pseudorandom generators in the standard sense of [Yao82], and specific counterexamples almost immediately follow from the results in [KP98] on the limitations of EIP. On the positive side, it was observed in [ABRW00] that for proof systems P *with* EIP there is an easy and general way of *converting any* pseudorandom generator that is *computationally* hard (in the standard sense) into a pseudorandom generator that is *hard for P*. No such *general* transformation is known for a single non-trivial proof system without EIP.

Thus, it is vital for the generator approach that we somehow restrict the class of mappings for which one hopes to trade computational hardness for the proof complexity hardness. Along these lines [ABRW00] specifically proposed to consider the class of Nisan-Wigderson generators (concrete results from that paper as well as from later improvements [AR01, Kra02] will be reviewed after we are done with this general overview). Some arguments advocating this choice (as opposed to other classical cryptographic constructions) were presented in [ABRW00] and in the introduction to [Raz02b]: the principle expressing hardness of NW-generators is tightly related to such familiar personages in proof complexity as the pigeonhole principle and Tseitin tautologies.

Elaborating on these arguments, we can further remark that the NW-generator is in a sense the quintessence of the very idea of "local consistency". Namely, the information contained in the output bits of a NW-generator is "local" (in the sense that every output bit depends only on a "small" subset of input bits which are "nearly independent" for different output bits), and it is locally consistent to such an extent that no interesting conclusion about the global behaviour of the generator can be obtained by an "easy" analysis of this local knowledge. This simple methodology is behind a great deal of lower bounds existing in proof complexity for weak proof systems, and it is also behind the efficient interpolation property.

This methodology also has a very clean mathematical meaning. A (non-existent) falsifying assignment to the tautology ϕ corresponds to a manifold with given local properties. The proof system P tries to argue that no such manifold may exist using tools at its disposal. And we (lower bounds provers) try to fool it by feeding into the potential proof something which looks like a (non-existing) manifold to such an extent that P can not discern the difference. See [Raz98] for (apparently) the cleanest implementation of this intuitive scheme.

Anyway, the moving forces that make the Nisan-Wigderson generator work in the computational world are of so general nature that we are ready to spell out the formal conjectures that the generator approach *always* works for Nisan-Wigderson generators. More specifically (assuming that the constructions are based on combinatorial designs with the same parameters as in the seminal paper [NW94]),

Conjecture 1. *Any* NW-generator based on *any* poly-time function that is hard on average for $\mathbf{NC^1}/poly$, is hard for the Frege proof system.

Conjecture 2. *Any* NW-generator based on *any* function in $\mathbf{NP} \cap co-\mathbf{NP}$ that is hard on average for $\mathbf{P}/poly$ (e.g. $B(f^{-1}(r))$, where $f(x)$ is any one-way permutation and $B(x)$ its hard-core bit), is hard for Extended Frege.

The suggestion to use Nisan-Wigderson generators for lower bounds in proof complexity has been recently re-iterated in [Kra02]. That paper also proposes a paradigm similar in spirit to the construction from [Gol00] in the context of computational complexity: hardness of the resulting mapping should depend on the randomness of the base functions rather than their complexity. So far all known results on the hardness of NW-generators for weak proof systems have not directly appealed to the randomness and used instead specific combinatorial properties of the base functions. But of course it remains to be seen yet which of the two paradigms (if any) will turn out more fruitful in the long run.

The task of proving lower bounds (even conditional) for strong proof systems is, in the author's opinion, extremely interesting and well justified in its own right (whereas the popular motivation that this should be regarded as an intermediate step in approaching the $\mathbf{NP}$ vs. $co-\mathbf{NP}$ problem looks, again in the author's opinion, more of a speculation). One venerable way, however, to make this study even more interesting is to look at the proof complexity of statements whose *validity* is also not known, and whose importance stretches well beyond any proof-theoretical studies.

To that end, [Raz95a] proposed[1] to study the proof complexity of the principle $\neg Circuit_t(f_n)$ expressing that the circuit size of the Boolean function f_n in n variables, given as its truth-table, is lower bounded by $t = t(n)$

[1][Raz95a] dealt with provability of Σ_0^b-statements in the theories of Bounded Arith-

(thus, e.g. $\mathbf{NP} \not\subseteq \mathbf{P}/poly$ is equivalent to the validity of $\neg Circuit_{t(n)}(SAT_n)$ for some $t(n) \geq n^{\omega(1)}$). [Raz95a] put forward the thesis (so far not refuted) that *all existing* proofs of lower bounds for restricted classes of circuits and for explicit functions translate to Extended Frege proofs (often to much weaker proof systems) of size $2^{O(n)}$. This makes the question of the efficient provability of the original principle $\neg Circuit_t(f_n)$ for *general* circuits even more intriguing.

The connection of this question to the generator approach above is the same as in the context of Natural Proofs [RR97]. Namely, if we have a function pseudorandom generator $G_n : \{0,1\}^{t_0} \to \{0,1\}^{2^n}$ that is hard for a proof system P, and such that the associated predicate $G(x)_y$ ($x \in \{0,1\}^{t_0}$, $y \in \{0,1\}^n$) can be computed by a size t circuit, then for every fixed seed $x \in \{0,1\}^{t_0}$, the Boolean function with the truth-table $G(x)$ is also computed by a size t circuit. Since for any given f_n the system P can not efficiently refute that f_n is different even from the functions $G(x)$ in the image of the generator G, the more it is not capable of efficient proofs of $\neg Circuit_t(f_n)$. In plain words,

TO SHOW THAT P DOES NOT HAVE EFFICIENT PROOFS OF THE FORMULA $\neg Circuit_t(f_n)$, WE NEED TO DESIGN A SUFFICIENTLY CONSTRUCTIVE PSEUDORANDOM GENERATOR HARD FOR P AND SUCH THAT THE NUMBER OF OUTPUT BITS, AS THE FUNCTION OF THE NUMBER OF INPUT BITS, IS AS LARGE AS POSSIBLE.

The larger number of output bits we can manage, the smaller are the parameters t_0, t (relatively to n) and the stronger is the result. In particular, in order to conclude the efficient unprovability of $\mathbf{NP} \not\subseteq \mathbf{P/poly}$, one needs a generator that stretches n bits to 2^{n^ϵ} bits; such generators are commonly known as *function pseudorandom generators*.

The tight connection between pseudorandom generators and the tautologies $\neg Circuit_t(f_n)$ has been also fruitfully exploited from more structural point of view. We already remarked above that $\mathbf{NP} \neq co - \mathbf{NP}$ implies the existence of hard tautologies for any propositional proof system, but this does not give any clue as to what these hard tautologies actually are.

metic, which is the uniform counterpart of propositional proof complexity. At the suggestion of Jan Krajíček, in later papers it was re-casted in more convenient framework of propositional proof complexity.

[IKW02] proved that under the assumption $\mathbf{NEXP} \subseteq \mathbf{P}/poly$, it is the specific tautologies $\neg Circuit_t(f_n)$ (for any f_n whatsoever) that are hard for any proof system. R. Impagliazzo (see a footnote in [Kra02, Section 1]) and independently M. Alekhnovich recently observed that the same conclusion holds under the assumption $\mathbf{BPP} \not\subseteq \mathbf{NP}$ (stronger than $\mathbf{NP} \neq co-\mathbf{NP}$). Although none of these two assumptions looks particularly plausible (and none of them is "purely computational" either), this still serves as another indication of the "distinguished" character of the tautologies $\neg Circuit_t(f_n)$.

The project of proving lower bounds for stronger and stronger classes of circuits until we arrive at $\mathbf{P} \neq \mathbf{NP}$ has met a solid obstacle in the form of Natural Proofs [RR97]. The project of proving lower bounds for stronger and stronger proof systems until we arrive at the (Extended) Frege proof system is restricted by the empirical observation that this task is even harder than the previous one. Given this gloomy background, fulfilling the generator approach for strong proof systems is certainly not an easy task, and the progress in this direction is much slower than originally hoped. However, it seems that we currently do not know of any general reasons (like those for the two previous projects formulated above) making us suspect that this task is unfeasible and/or requires an entirely different view of the subject.

1.2. Previous results and our contributions

In all known partial results along the generator approach, Nisan-Wigderson generators play the central role. Hardness results for generators of this kind are determined by combinatorial properties of the underlying set system, conditions imposed on the base functions, and by the specific way their computation is encoded as a propositional tautology. Disregarding for the moment all these technical issues, on the conceptual level [ABRW00] proved that the NW-generator is hard for Resolution but only when the complexity is measured by *width*. Another result from [ABRW00] says that the *Nisan generator* (that is, the partial case of the NW-generator in which all base functions must be linear mod 2) is hard for Polynomial Calculus (PC in what follows) over fields $\mathbb{F}$ with $char(\mathbb{F}) \neq 2$. The latter result was further extended and generalized in [AR01] to show that the Nisan-Wigderson generator is hard for PC over any field.

Results for resolution width and polynomial calculus degree are applicable to function generators stretching n bits to 2^{n^ϵ} bits. As long as the *size* complexity measures are concerned, [ABRW00] exhibited hard Nisan-Wigderson generators for the system PCR (that is a natural common extension of Polynomial Calculus and Resolution) but only when $m \leq o(n^2)$. This poor input/output ratio hindered their potential application to proving that $\mathbf{NP} \not\subseteq \mathbf{P}/poly$ is hard even for Resolution, and this was established by somewhat different methods in [RanRaz02, Raz02b].

A prominent *general* way for enhancing the I/O performance of pseudorandom generators in proof complexity has been recently proposed in [Kra02]. Like the classical constructions in computational complexity [Yao82, GGM86], it is very natural to try to achieve this goal by composing the given generator with itself. Unfortunately, it is far from clear whether hardness in the context of proof complexity is preserved under composition. [Kra02] proposed a way around this difficulty by showing that it is indeed the case if hardness is replaced by a stronger notion of *s-iterability* (the latter, in turn, being a variant of a similar notion of freeness earlier introduced in [Kra01b]). As a first application of this approach, [Kra02] showed that one particular construction of the Nisan generator from [ABRW00] can be iterated with itself once, thus giving a pseudorandom generator with $m = n^{3-\epsilon}$ output bits that is hard for Resolution.

In the current paper we continue this line of research. Let Res(k) be the propositional proof system that extends Resolution by allowing as its lines arbitrary k-DNFs. Our first main result (Theorem 2.7) exhibits Nisan generators that are hard for Res(k) and stretch n input bits to as many as $n^{(\epsilon \log n)/k}$ output bits. A relatively easy modification of this argument for $k = 1$ shows that this generator (from n to $n^{\epsilon \log n}$ bits) is hard not only for Resolution but also for its extension PCR (Theorem 2.19)[2]. These results were proved independently of [Kra02]; the proof method uses the resolution width/PC degree bounds from [ABRW00] cited above in combination with the machinery from the recent paper [SBI02] based upon the so-called *small restriction switching lemma*. In order to bring these two together, we also introduce a special kind of random restrictions specifically tailored to deal with Nisan generators.

Then, using a very simple reduction, we show that our generator is not only hard for Res(k) and PCR but it is in fact $\exp(n^{\Omega(1)})$-iterable for these systems (Theorems 2.10, 2.20). According to the paradigm from [Kra02], this

[2]Here and in the rest of inroduction we implicitly assume that $char(\mathbb{F}) \neq 2$.

implies that if we compose this generator with itself as many as $\exp(n^{\Omega(1)})$ times, the resulting mapping will still be hard for Res(k)/PCR. Applying in particular the classical GGM-construction [GGM86], we get in this way a function generator $G_n : \{0,1\}^n \longrightarrow \{0,1\}^{2^{n^\epsilon}}$ that is hard for Res($\epsilon \log n$), $\epsilon > 0$ a sufficiently small constant (Theorem 2.12), and for PCR. As we discussed in Section 1.1, this implies that neither Res($\log n$) nor PCR possess efficient proofs of $\mathbf{NP} \not\subseteq \mathbf{P}/poly$ (Theorems 2.13, 2.21), which are the first results of this kind for *any* propositional proof system that (most likely) does not have the Efficient Interpolation Property.

In the course of this work we were able to get a quadratic (in k) improvement of the small restriction switching lemma from [SBI02] based on Janson inequality (Lemma 4.4). This in particular implies that the moderately weak pigeonhole principle PHP^{2n}_n is exponentially hard for Res(k) when $k \leq (\epsilon \log n)/\log\log n$, as opposed to $k \leq \epsilon\sqrt{\log n/\log\log n}$ in [SBI02] (Theorem 2.15).

Finally we prove two miscellaneous results about the complexity of Nisan generators themselves; both of them underscore the sensitivity of their proof complexity behaviour with respect to the choice of encoding. The first result is somewhat unusual and demonstrates an interesting phenomenon that we call (by analogy with a similar notion in computational complexity [BGI+01]) *obfuscation*. Suppose that we do not impose any restrictions on the $\mathbb{F}_2$-matrix A defining the Nisan generator, except by requiring that it is sufficiently dense (thus, in particular, A may be the all-ones matrix so that every output bit computes the same sum of all input bits). It turns out that if the order in which input variables are processed is chosen independently and at random for all output bits then with overwhelming probability this in itself will make the resulting generator hard for Res(k) and PCR (Theorems 2.16, 2.22). In other words, every linear mapping defined by a sufficiently dense matrix A can be "obfuscated" for Res(k) and PCR by choosing the way it is computed and, respectively, encoded as a propositional tautology.

Our last result is specific for PCR. We show how to modify the (natural and reasonable) encoding of the Nisan generator used in the rest of the paper in such a way that this generator becomes hard for PCR even in the functional case $m = 2^{n^\epsilon}$ (Theorem 2.23). This encoding, however, is in a sense very bad: it does not seem to allow any reduction to $\neg Circuit_t(f_n)$, and the proof method apparently fails already for Res(2).

2. Definitions and statements of our results

In this section we typically confine ourselves to defining only those notions that are needed for stating our main results. Auxiliary concepts needed for their proofs will normally appear in respective places.

Let x be a *propositional variable*, i.e. a variable that ranges over the set $\{0,1\}$. A *literal* of x is either x (denoted sometimes as x^1) or $\bar{x}$ (denoted sometimes as x^0). A *clause* [*term*] is either a constant 0 or 1 (corresponding to FALSE and TRUE, respectively) or a disjunction [conjunction, respectively] of literals in which no variable appears along with its negation. A *CNF* [*DNF*] is a conjunction of clauses [disjunction of terms], often specified as the set of all participating clauses [terms, respectively]. Accordingly, a clause/term/CNF/DNF is a sub-clause/sub-term/sub-CNF/sub-DNF of another clause/term/CNF/DNF if every literal/literal/clause/term appearing in the first, appears also in the second. A clause/term/CNF/DNF is *monotone* if it does not contain occurrences of negated literals $\bar{x}$.

For a Boolean function f [a propositional formula F], let $Vars(f)$ [$Vars(F)$] be the set of its essential variables [the set of variables explicitly occurring in F, respectively]. The *width* of a clause C [of a term t] is defined as $w(C) \stackrel{\text{def}}{=} |Vars(C)|$ [$w(t) \stackrel{\text{def}}{=} |Vars(t)|$, respectively]. A k-*CNF* [k-DNF] is a CNF [DNF] in which all clauses [terms, respectively] are of width at most k. $|F|$ is the number of terms in a DNF F.

An *assignment* to a Boolean function f [a propositional formula F] is a mapping $\alpha : Vars(f) \to \{0,1\}$ [$\alpha : Vars(F) \to \{0,1\}$, respectively]. A *restriction of* f [of F] that, depending on the context, will be sometimes called *a partial assignment* is a mapping $\rho : Vars(f) \to \{0,1,\star\}$ [$\rho : Vars(F) \to \{0,1,\star\}$, respectively]. We let $\sup(\rho) \stackrel{\text{def}}{=} \rho^{-1}(\{0,1\})$ denote the set of assigned variables. The *restriction of a function f [of a formula F] by ρ*, denoted $f|_\rho$ [$F|_\rho$], is the Boolean function [propositional formula] obtained from f [from F, respectively] by setting the value of each $x_i \in \sup(\rho)$ to $\rho(x_i)$, and leaving each $x_i \notin \sup(\rho)$ unassigned. In the case of propositional formulas we as usual assume that simplifications are performed only when a subformula has become explicitly constant. A *variable substitution* of variables in V_1 by variables in V_2 is a mapping ρ that takes variables in V_1 to either propositional constants or *literals* of variables in V_2 (thus, restrictions are viewed as a special case of variable substitutions with $V_2 = V_1$). Variable substitutions ρ of variables in $Vars(f)$ or $Vars(F)$ also naturally act on the

Boolean function f or propositional formula F, and, as before, we denote the result of this action by $f|_\rho$, $F|_\rho$.

For an integer n, let $[n] \stackrel{\rm def}{=} \{1, 2, \ldots, n\}$. Whenever we use probabilistic methods, random variables will always appear in the bold face. We extensively use the Ω-notation (customary in complexity theory) that is opposite to the ordinary O-notation. For example, given two functions f, g with values in the set of non-negative reals, $f \geq \Omega(g)$ means that there exists an absolute constant $\epsilon > 0$ such that $f \geq \epsilon g$ for any specification of the parameters occurring in f, g.

Definition 2.1 ([Kra01a, ABRW00]) Let $m > n$, C be a Boolean circuit with n inputs $x_1, \ldots, x_n$ and m outputs, and $b \in \{0,1\}^m$ be an arbitrary Boolean vector. For every computational gate v of the circuit C we introduce a special *extension variable* y_v, and when v is the jth input gate, we identify y_v with the corresponding propositional variable x_j. Let $Vars_C \stackrel{\rm def}{=} \{x_1, \ldots, x_n\} \cup \{y_v \mid v \text{ a computational gate of } C\}$[3].

By $\tau(C, b)$ we denote the CNF in the variables $Vars_C$ that expresses the fact "$C(x_1, \ldots, x_n) = b$" and consists of the following clauses:

1. $y_{v_1}^{\bar{\epsilon}_1} \vee \ldots \vee y_{v_d}^{\bar{\epsilon}_d} \vee y_v^{\pi(\epsilon_1, \ldots, \epsilon_d)}$ whenever $v := \pi(v_1, \ldots, v_d)$ is an instruction of C of arity d and $\epsilon \in \{0,1\}^d$ is an arbitrary vector;

2. $y_{v_i}^{b_i}$ when v_i is the ith output gate of C, $i \in [m]$.

Remark 1 In this paper we will be mostly interested in the case when the circuit C is *linear*, that is consists only of linear instructions $v := v_1 \oplus v_2$ of arity 2.

For A an $(m \times n)$ 0-1 matrix and $i \in [m]$, let $J_i(A) \stackrel{\rm def}{=} \{j \in [n] \mid a_{ij} = 1\}$ and $X_i(A) \stackrel{\rm def}{=} \{x_j \mid a_{ij} = 1\}$ be the corresponding set of propositional variables. The set system $J_i(A)$ provides an alternative (and often more convenient) way to represent the matrix A.

Definition 2.2 An *ordering* $\leq$ of an $(m \times n)$ 0-1 matrix A is a tuple $(\leq_1, \ldots, \leq_m)$, where $\leq_i$ is a linear ordering of the set $J_i(A)$. Given any ordering $\leq$, let $C_1, \ldots, C_m$ be the (single-output) circuits naturally computing the parity functions $\bigoplus \{x_j \mid x_j \in X_i(A)\}$ according to this ordering. That is:

[3] although $Vars_C$ is in a one-to-one correspondence with the set of all gates, we prefer to draw a clear distinction between inputs and computational gates

- gates of C_i have the form v^i_Σ, where Σ is a non-empty $\leq_i$-initial segment of $J_i(A)$;
- when x_j is the minimal element of $J_i(A)$, $v^i_{\{x_j\}}$ is the input gate x_j;
- instructions have the form $v^i_{\Sigma\cup\{x_j\}} := v^i_\Sigma \oplus x_j$, where Σ is a proper initial segment of $J_i(A)$ and j is the minimal element in $J_i(A) \setminus \Sigma$;
- $v^i_{J_i(A)}$ is the output gate.

Denote by $C_{A,\leq}$ the m-output linear circuit which is a disconnected union of $C_1, \ldots, C_m$ (that is, these circuits do not have any gates in common except for input gates). Let $Vars_{\leq}(A) \stackrel{\text{def}}{=} Vars_{C_{A,\leq}}$ and $\tau_{\leq}(A, b) \stackrel{\text{def}}{=} \tau(C_{A,\leq}, b)$.

Definition 2.3 For a set of rows $I \subseteq [m]$ in the matrix A, we define its *boundary* $\partial_A(I)$ as the set of all $j \in [n]$ (called *boundary elements*) such that $\{\, a_{ij} \mid i \in I \,\}$ contains exactly one 1. We say that A is an (r, d)-*lossless expander* if

$$\forall I \subseteq [m](|I| \leq r \Rightarrow \sum_{i \in I} |J_i(A)| - |\partial_A(I)| \leq d \cdot |I|). \tag{1}$$

Remark 2 Another useful way to interpret the expansion property (1) in this definition is to say that rows in I have at most d *non-boundary* elements on average. In the regular case, that is when all $J_i(A)$ have the same cardinality s, (r, d)-lossless expanders are exactly (r, s, c)-expanders in the terminology of [ABRW00] for $c = s - d$. We introduce this new definition mainly to stress that it is in fact the difference between s and c that matters, and for most applications we need not know s (and in fact even do not need regularity). Also, the "ordinary" lossless expanders recently constructed in [CRVW02] correspond to the case $d = \epsilon s$ for a small constant $\epsilon > 0$, whence our choice of the name.

Despite recent progress, no explicit constructions of (r, d)-lossless expanders with the parameters sufficient for our purposes are currently known. Fortunately, we will be satisfied with the following simple non-constructive bound.

Definition 2.4 $\boldsymbol{A_{m,n}}$ is a random $(m \times n)$ 0-1 matrix in which all entries are independent and $\mathbf{P}[\boldsymbol{a_{ij}} = 1] = n^{-2/3}$.

Theorem 2.5 *Let $m \le 2^{n^\epsilon}$, where $\epsilon > 0$ is a sufficiently small constant. Then $\boldsymbol{A_{m,n}}$ is an $\left(n^{\Omega(1)}, O\left(\frac{\log m}{\log n}\right)\right)$-lossless expander with probability $\ge 1 - O(1/m)$.*

Definition 2.6 Res(k) is the propositional proof system whose lines are k-DNFs and whose inference rules are given below (F, G are k-DNFs, $1 \le w \le k$ and ℓ, ℓ_i are literals).

$$\frac{F}{F \vee \ell}\ (\textsc{Weakening}) \qquad\qquad \frac{F \vee \ell_1\ \ldots\ F \vee \ell_w}{F \vee (\bigwedge_{i=1}^{w} \ell_i)}\ (\text{AND-}\textsc{introduction})$$

$$\frac{F \vee (\bigwedge\limits_{i=1}^{w} \ell_i)}{F \vee \ell_i}\ (\text{AND-}\textsc{elimination}) \qquad\qquad \frac{F \vee (\bigwedge\limits_{i=1}^{w} \ell_i) \qquad G \vee \bigvee\limits_{i=1}^{w} \bar{\ell}_i}{F \vee G}\ (\textsc{Cut})$$

A *Res(k) refutation of a set of k-DNFs* is a Res(k) proof of 0 from this set. In particular, a *Res(k) refutation of a CNF τ* is a Res(k) refutation of the set of clauses τ consists of.

We define the *size* of a Res(k) proof as the number of *lines* in it. Note that since in this paper we deal exclusively with lower bounds, this makes our results only stronger (as opposed to measuring the complexity by the number of bits).

Note that any variable substitution ρ takes a Res(k) refutation of a CNF τ into a Res(k) refutation of $\tau|_\rho$. This in particular implies that the minimal size of a Res(k) refutation of τ is at least the same size for $\tau|_\rho$.

Finally, the case $k = 1$ corresponds to *Resolution*: this system operates with clauses and has the only inference rule

$$\frac{F \vee x \qquad G \vee \bar{x}}{F \vee G}$$

called *resolution rule* (the weakening rule does not change the power of Resolution and is usually, often implicitly, also assumed for convenience). The *width* of a resolution proof is the maximal width of a clause occurring in this proof.

Now we are ready to formulate our main result for Nisan generators.

Theorem 2.7 *Let A be an $(m \times n)$ (r,d)-lossless expander, and assume that*

$$\min_{i \in [m]} |J_i(A)| \geq Cd(k + \log_2 m) \tag{2}$$

for a sufficiently large constant $C > 0$. Let $\leq$ be an arbitrary ordering of A and $b \in \{0,1\}^m$ be an arbitrary vector. Then every Res(k) refutation of $\tau_{\leq}(A, b)$ must be of size $\geq \exp(r/2^{O(kd)})$.

Combining Theorems 2.7 and 2.5, we in particular get

Corollary 2.8 *Let m, n, k be parameters such that*

$$m \leq n^{(\epsilon \log n)/k}, \tag{3}$$

where $\epsilon > 0$ is a sufficiently small constant. Then with probability $1 - O(1/m)$ the following holds. For every ordering $\leq$ of $\boldsymbol{A_{m,n}}$ and every $b \in \{0,1\}^m$, every Res(k)-refutation of $\tau_{\leq}(\boldsymbol{A_{m,n}}, b)$ must have size $\geq \exp(n^{\Omega(1)})$.

Definition 2.9 ([Kra02]) For an n-input m-output circuit C and a vector $z_1, \ldots, z_m$ of propositional variables, let $\tau_C(x_1, \ldots, x_n, \vec{y}, z_1, \ldots, z_m)$ be defined in the same way as $\tau(C, b)$ (see Definition 2.1) with the difference that the "output axioms" $y_{v_i}^{b_i}$ get replaced by the two clauses $y_{v_i} \vee \bar{z}_i,\ \bar{y}_{v_i} \vee z_i$ (thus, $\tau(C, b)$ is the same as $\tau_C(\vec{x}, \vec{y}, b)$).

For a proof system P and an integer S, the circuit C is *S-iterable in P* if for every CNF of the form

$$\bigwedge_{\nu=1}^{H} \tau_C(x_1^{(\nu)}, \ldots, x_n^{(\nu)}, \vec{y}^{(\nu)}, q_1^{(\nu)}, \ldots, q_m^{(\nu)}), \tag{4}$$

every its refutation in the system P must have size $\geq S$. Here $x_j^{(\nu)}, \vec{y}^{(\nu)}$ are pairwise disjoint tuples of the variables, and every $q_i^{(\nu)}$ is either a Boolean constant or belongs to the set $\left\{ x_j^{(\mu)} \mid j \in [n],\ \mu < \nu \right\}$.

Since $\tau(C, b)$ itself has the form (4) (with $H = 1$), iterability is stronger than just hardness. As the proof of the following theorem will show, for the particular case of NW-generators they are not very much apart from each other.

Theorem 2.10 *Let A be an $(m \times n)$ (r,d)-lossless expander, and assume that*

$$s \stackrel{\mathrm{def}}{=} \min_{i\in[m]} |J_i(A)| \geq Cd(k+\log_2 m) \tag{5}$$

for a sufficiently large constant $C > 0$. Then for every ordering $\leq$ of the matrix A, the circuit $C_{A,\leq}$ from Definition 2.2 *is* $\exp(\min\{\Omega(s/d),\ r/2^{O(kd)}\})$-*iterable in Res*$(k)$.

As in [Kra02], this in particular implies that composing the generator $C_{A,\leq}$ with itself preserves hardness. We will note here only one particular iteration protocol corresponding to the classical construction from [GGM86].

Definition 2.11 Let C be a Boolean circuit (over an arbitrary basis) with n inputs and $2n$ outputs, and $h \geq 1$ be an integer. The circuit C^h with n inputs and $2^h n$ outputs is constructed as follows. For every binary string u with $|u| \leq h-1$ we prepare an isomorphic copy C_u of C. The inputs of C^h are the inputs of C_Λ (Λ the empty string), and the outputs of C^h are the outputs of the circuits C_u with $|u| = h-1$. Finally, for every u with $|u| < h-1$, the first n outputs of C_u are identified with the inputs of C_{u*0}, and the last n output bits are identified with inputs of C_{u*1} (in everything else these circuits are completely independent).

Theorem 2.12 *Let $\leq$ be an arbitrary ordering of an $(2n \times n)$ (r,d)-lossless expander A, and assume that*

$$s \stackrel{\mathrm{def}}{=} \min_{i\in[2n]} |J_i(A)| \geq Cd(k+\log_2 n)$$

for a sufficiently large constant $C > 0$. Let $h \geq 1$ be an arbitrary integer, and $b \in \{0,1\}^{2^h \cdot n}$ be an arbitrary vector. Then every Res(k) refutation of $\tau(C^h_{A,\leq}, b)$ must have size $\geq \exp(\min\{\Omega(s/d),\ r/2^{O(kd)}\})$.

For a Boolean function f_n in n variables[4] and $t \leq 2^n$, denote by $Circuit_t(f_n)$ an $O(1)$-CNF of size $2^{O(n)}$ encoding the description of a size-t fan-in 2 Boolean circuit over the standard basis $\{\neg, \wedge, \vee\}$ for computing f_n. We will recall an exact definition in Section 8; for the time being let us just observe that proving that the circuit size of f_n is greater than t is tantamount to showing that $Circuit_t(f_n)$ is unsatisfiable.

[4]Note that the "intended meaning" of n in Theorems 2.13, 2.21 is completely different from all other results

Theorem 2.13 *Let f_n be any Boolean function in n variables, and $n^2 \leq t \leq 2^n$. Then every Res$(\epsilon \log t)$ refutation of Circuit$_t(f_n)$ ($\epsilon > 0$ a sufficiently small constant) must have size* $\exp(t^{\Omega(1)})$.

Let us now consider the pigeonhole principle.

Definition 2.14 Assume $m > n$. The (negation of the) *onto pigeonhole principle* is the unsatisfiable CNF in the variables $\{x_{ij} \mid i \in [m],\ j \in [n]\}$ denoted by $\neg onto - PHP_n^m$ that is the conjunction of the following clauses:

$$Q_i \stackrel{\text{def}}{=} \bigvee_{j=1}^{n} x_{ij} \ (i \in [m]);$$
$$Q_{i_1,i_2;j} \stackrel{\text{def}}{=} (\bar{x}_{i_1 j} \vee \bar{x}_{i_2 j}) \ (i_1 \neq i_2 \in [m],\ j \in [n]);$$
$$Q_j \stackrel{\text{def}}{=} \bigvee_{i=1}^{m} x_{ij} \ (j \in [n])$$

(the prefix "onto" refers to the presence of the last group of axioms).

[SBI02] proved that $\neg onto - PHP_n^{2n}$ is exponentially hard to refute in Res(k) as long as $k \leq \sqrt{\log n / \log\log n}$. We improve this as follows:

Theorem 2.15 *Every Res(k) refutation of $\neg onto - PHP_n^{2n}$ must have size* $\exp(n/(\log n)^{O(k)})$.

This bound is exponential up to $k = (\epsilon \log n)/\log\log n$.

Now we return to pseudorandom generators and show that expansion properties of A can be traded for the properties of the ordering $\leq$. Here we formulate this result only in style of Corollary 2.8, that is for random ordering. For the combinatorial property of the ordered matrix $(A, \leq)$ that implies its hardness see Section 10.

Theorem 2.16 *Let A be any fixed $(m \times n)$ matrix A such that $|J_i(A)| \geq n^{2/3}$ for all $i \in [m]$, and assume that* (3) *holds. Let $\leq$ be the random ordering of this matrix. Then with probability $1 - O((mn)^{-1})$, for every $b \in \{0,1\}^m$ every Res(k) refutation of $\tau_{\leq}(A, b)$ must have size* $\geq \exp(n^{\Omega(1)})$.

Let us now recall another extension of Resolution that is of more algebraic flavour.

Definition 2.17 ([CEI96]) Let $\mathbb{F}$ be a fixed field. *Polynomial Calculus* (PC for short) is the proof system whose lines are polynomials $f \in \mathbb{F}[x_1, \ldots, x_n]$ (a polynomial f is interpreted as the polynomial equation $f = 0$). It has polynomials $x_i^2 - x_i$ ($i \in [n]$) as *default axioms* and has two inference rules:

$$\frac{f \qquad g}{\alpha f + \beta g}; \ \alpha, \beta \in \mathbb{F} \ (\textsc{Scalar Addition});$$

$$\frac{f}{x_i \cdot f} \ (\textsc{Variable Multiplication}).$$

Definition 2.18 ([ABRW02]) Let again $\mathbb{F}$ be a fixed field. *Polynomial Calculus with Resolution* (PCR) is the proof system whose lines are polynomials from $F[x_1, \ldots, x_n, \bar{x}_1, \ldots, \bar{x}_n]$, where $\bar{x}_1, \ldots, \bar{x}_n$ are treated as new formal variables. PCR has all default axioms and inference rules of PC (including, of course, those that involve new variables $\bar{x}_i$), plus additional default axioms $x_i + \bar{x}_i = 1$ ($i \in [n]$).

For a clause C, denote by Γ_C the monomial

$$\Gamma_C \stackrel{\mathrm{def}}{=} \prod_{\bar{x} \in C} x \cdot \prod_{x \in C} \bar{x}, \tag{6}$$

and for a CNF τ, let $\Gamma_\tau \stackrel{\mathrm{def}}{=} \{\Gamma_C \mid C \in \tau\}$ (note that τ is unsatisfiable if and only if the polynomials Γ_τ have no common root in $\mathbb{F}$ satisfying all default axioms of PCR). A *PCR refutation of a CNF* τ is a PCR proof of the contradiction 1=0 from Γ_τ.

The *degree* of a PCR proof is defined as the maximal degree of a polynomial appearing in it, and its *size* is the number of different *monomials* in this proof.

Remark 3 PC and PCR are equivalent with respect to the degree measure (via the linear transformation $\bar{x}_i \mapsto 1 - x_i$). Also note that we measure the size of PCR proofs differently from Definition 2.6; namely, by the number of monomials. We do not know if our results still hold if the size is measured by the number of lines.

All our lower bounds for Res(1) (= Resolution) generalize to PCR over any field $\mathbb{F}$ with $char(\mathbb{F}) \neq 2$. That is,

Theorem 2.19 *Let A be an $(m \times n)$ (r,d)-lossless expander, and assume that*

$$\min_{i\in[m]} |J_i(A)| \geq Cd\log_2 m \tag{7}$$

for a sufficiently large constant $C > 0$. Let $\leq$ be an arbitrary ordering of A, $b \in \{0,1\}^m$ be an arbitrary vector, and $\mathbb{F}$ be an arbitrary field with $char(\mathbb{F}) \neq 2$. Then every PCR refutation of $\tau_{\leq}(A,b)$ over the field $\mathbb{F}$ must be of size $\geq \exp(r/2^{O(d)})$.

Theorem 2.20 *Let A be an $(m \times n)$ (r,d)-lossless expander such that (7) holds, and let $s \stackrel{\text{def}}{=} \min_{i\in[m]} |J_i(A)|$. Then for every ordering $\leq$ of the matrix A and for every field $\mathbb{F}$ with $char(\mathbb{F}) \neq 2$, the circuit $C_{A,\leq}$ is* $\exp(\min\{\Omega(s/d),\ r/2^{O(d)}\})$*-iterable in PCR over $\mathbb{F}$.*

Let $Circuit_t^{\oplus}(f_n)$ be defined in the same way as $Circuit_t(f_n)$, with the exception that besides the standard connectives $\{\neg, \wedge, \vee\}$, the encoded circuit also allows PARITY gates of fan-in 2.

Theorem 2.21 *Let f_n be any Boolean function in n variables, $n^2 \leq t \leq 2^n$, and $\mathbb{F}$ be any field with $char(\mathbb{F}) \neq 2$. Then every PCR refutation of $Circuit_t^{\oplus}(f_n)$ over $\mathbb{F}$ must have size* $\exp(t^{\Omega(1)})$.

Theorem 2.22 *Let A be any fixed $(m \times n)$ matrix such that $|J_i(A)| \geq n^{2/3}$ for all $i \in [m]$, and assume that $m \leq n^{\epsilon \log n}$ for a sufficiently small constant $\epsilon > 0$. Let $\underline{\leq}$ be the random ordering of this matrix, and $\mathbb{F}$ be any field with $char(\mathbb{F}) \neq 2$. Then with probability $1 - O((mn)^{-1})$, for every $b \in \{0,1\}^m$ every PCR refutation of $\tau_{\underline{\leq}}(A,b)$ over $\mathbb{F}$ must have size* $\exp(n^{\Omega(1)})$.

We conclude with one miscellaneous result about the hardness of Nisan generators that is specific to PCR. For these generators we still do not know how to get more than $n^{\epsilon \log n}$ output bits even in case of Resolution. We now show how to make a *function* (that is, with $m = 2^{n^\epsilon}$ bits) Nisan generator hard for PCR at the expense of spoiling its encoding.

Let now $\leq\ = (\leq_1, \ldots, \leq_m)$ be a tuple of *cyclic* orderings of the sets $(J_1(A), \ldots, J_m(A))$ and

$$Vars_{\leq}^{Cycl}(A) \stackrel{\text{def}}{=} \{x_1, \ldots, x_n\} \cup \left\{ y_\Delta^i \,\middle|\, i \in [m],\ \Delta \text{ a } \textit{cyclic} \text{ interval in } J_i(A) \right.$$
$$\left. \text{such that } \Delta = J_i(A) \text{ or } \frac{|J_i(A)|}{3} - 1 \leq |\Delta| \leq \frac{2|J_i(A)|}{3} + 1 \right\}.$$

Let $\tau_{\leq}^{Cycl}(A,b)$ be the 3-CNF consisting of the clauses that result from expanding those constraints of the form

$$y^i_{\Delta\cup\{j\}} \equiv y^i_{\Delta} \oplus x_j \ (i \in [m],\ j \text{ adjacent to } \Delta); \tag{8}$$
$$y^i_{\Delta_1\cup\Delta_2} \equiv y^i_{\Delta_1} \oplus y^i_{\Delta_2} \ (i \in [m],\ \Delta_1, \Delta_2 \text{ disjoint and adjacent}); \tag{9}$$
$$y^i_{J_i(A)} \equiv b_i. \tag{10}$$

in which all variables belong to $Vars_{\leq}^{Cycl}(A)$ (that is, $\Delta, \Delta \cup \{j\}, \Delta_1$ and Δ_2 must obey the size bound in its definition).

Let us verify that $\tau_{\leq}^{Cycl}(A,b)$ is indeed a complete encoding of the linear system $AX = b$ (this is not immediately clear due to the presence of the constraints on $|\Delta|$).

Fact 1 *The system of $\mathbb{F}_2$-linear equations $AX = b$ is consistent if and only if $\tau_{\leq}^{Cycl}(A,b)$ is satisfiable.*

Proof. The "only if" part is obvious. For the opposite direction, we only have to show that the x-part of every satisfying assignment for $\tau_{\leq}^{Cycl}(A,b)$ also satisfies the system $AX = b$.

For any given $i \in [m]$ fix an arbitrary partition $J_i(A) = \Delta_{i,1} \dot\cup \Delta_{i,2} \dot\cup \Delta_{i,3}$ into three cyclic intervals of almost equal sizes: $\frac{|J_i(A)|}{3} - 1 \leq |\Delta_{i\nu}| \leq \frac{|J_i(A)|}{3} + 1$. Then (9) (applied twice) and (10) imply that in every satisfying assignment we have $y^i_{\Delta_{i,1}} \oplus y^i_{\Delta_{i,2}} \oplus y^i_{\Delta_{i,3}} = b_i$. On the other hand, by summing up appropriate axioms (8) (with $j \in \Delta_{i\nu}$), we obtain $y^i_{\Delta_{i\nu'}} \oplus \bigoplus_{j\in\Delta_{i\nu}} x_j = y^i_{\Delta_{i\nu}\cup\Delta_{i\nu'}}$ $(\nu' \neq \nu)$ and then, using again (9), $\bigoplus_{j\in\Delta_{i\nu}} x_j = y^i_{\Delta_{i\nu}}$ for every $\nu \in [3]$. The statement follows.■

Theorem 2.23 *Let A be an $(m\times n)$ (r,d)-lossless expander such that* $\min_{i\in[m]} |J_i(A)| \geq Cd\log_2 m$ *for a sufficiently large constant $C > 0$, $\leq$ be an arbitrary cyclic ordering of A, $b \in \{0,1\}^m$ be an arbitrary vector, and $\mathbb{F}$ any field with $char(\mathbb{F}) \neq 2$. Then every PCR refutation of $\tau_{\leq}^{Cycl}(A,b)$ over the field $\mathbb{F}$ must have size $\geq \exp(r/d^{O(1)})$.*

The bound of this theorem is as good as we might hope. The encoding itself, however, is not very useful: it does not correspond to any circuit, and, moreover, the proof of this theorem seems to completely break apart for Res(2).

3. Preliminaries

In this section we begin the proof of Theorem 2.7 by presenting some known results in the form adapted to our purposes.

Recall several definitions from [ABRW00].

Definition 3.1 A Boolean function f is *ℓ-robust* if every restriction ρ such that $f|_\rho = const$, satisfies $|\sup(\rho)| \geq \ell$.

Definition 3.2 Let A be an $(m \times n)$ 0-1 matrix. For every Boolean function f with the property $\exists i \in [m](Vars(f) \subseteq X_i(A))$ we introduce a new *extension variable* y_f. Let $Vars(A)$ be the set of all these variables.

Given Boolean functions $\vec{g} = (g_1, \ldots, g_m)$ such that $Vars(g_i) \subseteq X_i(A)$, we denote by $\tau(A, \vec{g})$ the CNF in the variables $Vars(A)$ that consists of those clauses $y_{f_1}^{\epsilon_1} \vee \ldots \vee y_{f_w}^{\epsilon_w}$ for which there exists $i \in [m]$ such that

$$Vars(f_1) \cup \ldots \cup Vars(f_w) \subseteq X_i(A)$$

and

$$g_i \leq f_1^{\epsilon_1} \vee \ldots \vee f_w^{\epsilon_w}.$$

We now have the following (minor) generalization of [ABRW00, Theorem 3.1].

Theorem 3.3 *Let A be an (r, d)-lossless expander of size $(m \times n)$, and $g_1, \ldots, g_m$ be ℓ-robust functions with $Vars(g_i) \subseteq X_i(A)$, where $\ell \geq d + 1$. Then every resolution refutation of $\tau(A, \vec{g})$ must have width $> \frac{r(\ell-d)}{2\ell}$.*

Proof. The only difference from [ABRW00, Theorem 3.1] is that we now use (r, d)-lossless expanders instead of (r, s, c)-expanders with $s - c = d$. The proof goes the same way, and we only remark on the changes to be made to its text. All these changes pertain to the proof of [ABRW00, Claim 3.3].

First, the bound on the size of $\partial_A(I)$ now becomes $|\partial_A(I)| \geq \sum_{i \in I} |J_i(A)| - d \cdot |I|$ (as opposed to $|\partial_A(I)| \geq c \cdot |I|$ in [ABRW00]). Next, the proof of the crucial inequality $|J_{i_1}(A) \cap \partial_A(I)| \leq s - \ell$ in [ABRW00] actually shows $|J_{i_1}(A) \cap \partial_A(I)| \leq |J_{i_1}(A)| - \ell$. Last, the final calculation now looks like

$$\begin{aligned} \sum_{i \in I} |J_i(A)| - d \cdot |I| &\leq |\partial_A(I)| \leq \sum_{i \in I_0} |J_i(A)| + \sum_{i \in I_1} (|J_i(A)| - \ell) \\ &= \sum_{i \in I} |J_i(A)| - \ell \cdot |I_1| \leq \sum_{i \in I} |J_i(A)| - \ell \cdot (|I| - w(C)) \end{aligned}$$

which implies our bound $w(C) \geq \frac{r(\ell-d)}{2\ell}$ since $|I| > r/2$.$\blacksquare$

We will only need the following special case of this.

Corollary 3.4 *Let A be an (r,d)-lossless expander of size $(m \times n)$ such that $|J_i(A)| \geq 2d$ for all $i \in [m]$. Then for every ordering $\leq$ and every $b \in \{0,1\}^m$, every resolution refutation of $\tau_{\leq}(A,b)$ must have width $> r/4$.*

Proof. In Theorem 3.3, let g_i be the function $\bigoplus_{j \in J_i(A)} = b_i$ and $\ell := 2d$. Since $\tau_{\leq}(A,b)$ is a sub-CNF of $\tau(A,\vec{g})$ (for any ordering $\leq$), the result follows.$\blacksquare$

Definition 3.5 ([SBI02]) A *decision tree* is a rooted binary tree such that every internal node is labelled with a variable, the edges leaving this node correspond to whether the variable is set to 0 or 1, and the leaves are labelled with either 0 or 1. As usual, we assume that on every given path no variable appears more than once. Then every path from the root to a leaf may be viewed as a partial assignment, and this assignment, in turn, will be sometimes identified with the corresponding leaf. For a decision tree T and $\epsilon \in \{0,1\}$, we write the set of paths (partial assignments) that lead from the root to a leaf labelled ϵ as $Br_\epsilon(T)$. We say that a decision tree T *strongly represents a DNF F* if for every $\pi \in Br_0(T)$ and for all $t \in F$, $t|_\pi = 0$ and for every $\pi \in Br_1(T)$, there exists $t \in F$ such that $t|_\pi = 1$. Let the *representation height* of F, $h(F)$ be the minimum height of a decision tree strongly representing F.

Proposition 3.6 ([SBI02]) *Let τ be an h-CNF, P be a $Res(k)$ refutation of τ, and let ρ be a partial assignment so that for every line F of P, $h(F|_\rho) \leq h$. Then $\tau|_\rho$ has a resolution refutation of width $\leq kh$.*

Finally, we recall a combinatorial inequality originally proved in [Jan90] and further generalized in [AS00, Section 8.1].

Definition 3.7 For propositional variables $x_1, \ldots, x_n$ and probabilities $p_1, \ldots, p_n \in [0,1]$, denote by $\boldsymbol{a}_{\vec{p}}$ a random assignment that independently assigns every variable x_i to 1 with probability p_i, and to 0 with probability $(1-p_i)$.

Proposition 3.8 *Let $t_1, \ldots, t_H$ be monotone terms in the variables $x_1, \ldots, x_n$ (not necessarily distinct), and $p_1, \ldots, p_n \in [0, 1]$. Let*

$$\mu \stackrel{\text{def}}{=} \sum_{\alpha \in [H]} \mathbf{P}[t_\alpha(\boldsymbol{a}_{\vec{\boldsymbol{p}}}) = 1] \tag{11}$$

be the expectation of the number of terms satisfied by $\boldsymbol{a}_{\vec{\boldsymbol{p}}}$, and let

$$\Delta \stackrel{\text{def}}{=} \sum_{\substack{\alpha \neq \beta \in [H] \\ Vars(t_\alpha) \cap Vars(t_\beta) \neq \emptyset}} \mathbf{P}[t_\alpha(\boldsymbol{a}_{\vec{\boldsymbol{p}}}) = t_\beta(\boldsymbol{a}_{\vec{\boldsymbol{p}}}) = 1]\,.$$

Then

$$\mathbf{P}\left[\bigvee_{\alpha \in [H]} t_\alpha(\boldsymbol{a}_{\vec{\boldsymbol{p}}}) = 0\right] \leq e^{-\mu + \frac{\Delta}{2}}.$$

Proof. The notation of [AS00, Section 8.1] corresponds to ours as follows: Ω is $\{x_1, \ldots, x_n\}$, I is $[H]$, $A_i := Vars(t_i)$ ($i \in [H]$), and B_i is the event $t_i(\boldsymbol{a}_{\vec{\boldsymbol{p}}}) = 1$. After this translation, our proposition is exactly the second inequality in [AS00, Theorem 8.1.1].■

Remark 4 We will actually need another version of this inequality ([AS00, Theorem 8.1.2]), useful when $\Delta \gg \mu$. We, however, will need to dig into the proof of that theorem rather deeply. For this reason we do not formulate the corresponding general statement here, but rather incorporate it as an ad hoc argument in the proof of Lemma 4.4.

4. Small restriction switching lemma

In this section we give a quadratic improvement on the original version of this lemma from [SBI02]. In our main application, the underlying random restriction will not act totally independently on different variables, but at least it will have some "weak local independence" property. We will be able to capture this property in the main statement so that the proof will not become more complicated than for truly independent restrictions, but this will require several auxiliary definitions.

Definition 4.1 We say that a restriction ρ' is a *sub-restriction* of another restriction ρ if $\sup(\rho') \subseteq \sup(\rho)$ and ρ, ρ' coincide on $\sup(\rho')$. For random restrictions $\boldsymbol{\rho}, \boldsymbol{\rho}'$, $\boldsymbol{\rho}'$ is a *sub-restriction* of $\boldsymbol{\rho}$ if there exists a set Ω, a random variable $\boldsymbol{\omega} \in \Omega$, and two random functions $\boldsymbol{\rho}, \boldsymbol{\rho}'$ from Ω to the set of all restrictions such that:

1. $\boldsymbol{\rho}$ has the same distribution as $\boldsymbol{\rho}(\boldsymbol{\omega})$, and $\boldsymbol{\rho}'$ has the same distribution as $\boldsymbol{\rho}'(\boldsymbol{\omega})$;

2. $\boldsymbol{\rho}'(\boldsymbol{\omega})$ is a sub-restriction of $\boldsymbol{\rho}(\boldsymbol{\omega})$ (with probability 1).

Clearly, if $\boldsymbol{\rho}'$ is a sub-restriction of $\boldsymbol{\rho}$ then for every terms $t_1, \ldots, t_H$ we have the inequality $\mathbf{P}\left[\bigvee_{\alpha\in[H]}(t_\alpha|_{\boldsymbol{\rho}'} \equiv 1)\right] \le \mathbf{P}\left[\bigvee_{\alpha\in[H]}(t_\alpha|_{\boldsymbol{\rho}} \equiv 1)\right]$.

Definition 4.2 A *weight function* is any function $\mu : \{x_1, \ldots, x_n\} \longrightarrow \mathbb{N}^+$ from variables to strictly positive integers. The *weight* $\mu(V)$ *of a set of variables* V is defined as $\mu(V) \stackrel{\text{def}}{=} \sum_{x\in V}\mu(x)$, and the *weight of a term* t is $\mu(t) \stackrel{\text{def}}{=} \mu(Vars(t))$. We will denote by μ_{triv} the *trivial weight function* identically equal to 1; in this case the weight of a term is equal to its width. For arbitrary weight function μ we have this in one direction: $\mu(t) \ge w(t)$. A DNF F is a *weighted K-DNF* (w.r.t. a weight function μ) if all terms $t \in F$ have weight $\le K$.

Finally, we define the amount of "weak local independence" needed to carry out the proof of our switching lemma.

Definition 4.3 Let μ be a weight function, and let $p \in [0,1]$. Let $\boldsymbol{\rho_{\mu,p}}$ be the random restriction that independently assigns every variable x to 0,1 with probabilities $p^{\mu(x)}/2$, and leaves it unassigned with probability $(1-p^{\mu(x)})$. Given an integer r, say that a random restriction $\boldsymbol{\rho}$ is *(r,μ,p)-independent* if for every subset X of variables with $|X| \le r$, $\boldsymbol{\rho_{\mu,p}}|_X$ is a sub-restriction of $\boldsymbol{\rho}|_X$.

Thus, if $\boldsymbol{\rho}$ is (r,μ,p)-independent, then for any set of terms $t_1, \ldots, t_H$ we have $\mathbf{P}\left[\bigvee_{\alpha\in[H]}(t|_{\boldsymbol{\rho_{\mu,p}}} \equiv 1)\right] \le \mathbf{P}\left[\bigvee_{\alpha\in[H]}(t|_{\boldsymbol{\rho}} \equiv 1)\right]$ (and, as a matter of fact, this is the only property of (r,μ,p)-independent restrictions we will need).

The rest of this section is devoted to the proof of the following lemma.

Lemma 4.4 *Let μ be a weight function, and F be a weighted K-DNF w.r.t. μ. Suppose that $\boldsymbol{\rho}$ is a random (r,μ,p)-independent restriction. Then for every $h \le r$,*

$$\mathbf{P}[H(F|_{\boldsymbol{\rho}}) > h] \le \exp(-h(p/2)^{O(K)}). \tag{12}$$

Proof. We begin with the case when F is monotone, and in that case we are going to prove by induction on K that

$$\mathbf{P}[H(F|_{\boldsymbol{\rho}}) > h] \leq e^{-h(p\epsilon)^{2K}}, \tag{13}$$

where $\epsilon > 0$ is a sufficiently small constant[5]. This clearly implies (12).

Base $K = 0$ is obvious since F is a constant and $h(F|_{\boldsymbol{\rho}}) = 0$ with probability 1.

Inductive step. Let F be a non-trivial monotone weighted K-DNF, and assume that (13) is established for all weighted K'-DNF with $K' < K$ (and for all integers h). We define a numerical invariant $\delta(F)$ that represents, up to a scaling factor, the optimal value of the parameter Δ in Proposition 3.8 when we attempt to apply it to the formula F. Further analysis will be sharply divided according to whether $\delta(F)$ is small or large (cf. [SBI02, Theorem 3]).

First, for a set of variables V we let

$$\delta(V) \stackrel{\text{def}}{=} \begin{cases} 0 & \text{if } V = \emptyset \\ (2/p)^{\mu(V)} & \text{otherwise.} \end{cases}$$

Next, given a random term $\boldsymbol{t} \in F$, we define

$$\delta(\boldsymbol{t}) \stackrel{\text{def}}{=} \mathbf{E}[\delta(Vars(\boldsymbol{t}) \cap Vars(\boldsymbol{t'}))]\,,$$

where $\boldsymbol{t'}$ is an independent copy of $\boldsymbol{t}$. Finally, let

$$\delta(F) \stackrel{\text{def}}{=} \min_{\boldsymbol{t}} \delta(\boldsymbol{t}),$$

where the minimum is taken over all random terms $\boldsymbol{t} \in F$. Let

$$s \stackrel{\text{def}}{=} \lceil 2h(p\epsilon)^{2K} \rceil \tag{14}$$

(so that the right-hand side of (13) is roughly $e^{-s/2}$). Consider two cases.

Case 1: $\delta(F) \leq s^{-1}$. Let $\boldsymbol{t} \in F$ be the random term for which $\delta(\boldsymbol{t}) \leq s^{-1}$. Arguing as in the proof of [AS00, Theorem 8.1.2], let $\boldsymbol{t_1}, \ldots, \boldsymbol{t_H}$ be independent samples from F according to the distribution of $\boldsymbol{t}$, where

$$H \stackrel{\text{def}}{=} \lceil s(2/p)^K \rceil. \tag{15}$$

[5]The factor 2 in (13) can be removed by using a slightly more sophisticated analysis analysis in Case 2 below, but we do not need that precision in what follows.

Then

$$\mathbf{E}\left[\sum_{\alpha\neq\beta\in[H]}\delta(Vars(\boldsymbol{t_\alpha})\cap Vars(\boldsymbol{t_\beta}))\right]=H(H-1)\delta(\boldsymbol{t})\leq H(H-1)s^{-1},$$

and we fix any particular sampling $t_1,\ldots,t_H$ for which

$$\sum_{\alpha\neq\beta\in[H]}\delta(Vars(t_\alpha)\cap Vars(t_\beta))\leq H(H-1)s^{-1}. \tag{16}$$

The terms $t_1,\ldots,t_H$ altogether contain at most HK variables, and $HK\leq h\leq r$ as long as the constant ϵ is small enough. Therefore, since $\boldsymbol{\rho}$ is (r,μ,p)-independent, we have

$$\mathbf{P}[h(F|_{\boldsymbol{\rho}})>h]\leq\mathbf{P}\left[\bigwedge_{\alpha\in[H]}(t_\alpha|_{\boldsymbol{\rho}}\not\equiv 1)\right]\leq\mathbf{P}\left[\bigwedge_{\alpha\in[H]}(t_\alpha|_{\boldsymbol{\rho_{\mu,p}}}\not\equiv 1)\right]. \tag{17}$$

Furthermore, since the terms t_α are monotone, $t_\alpha|_{\boldsymbol{\rho_{\mu,p}}}\equiv 1$ if and only if $t_\alpha(\boldsymbol{a})=1$, where $\boldsymbol{a}$ is the random (total) assignment obtained from $\boldsymbol{\rho_{\mu,p}}$ by additionally assigning to 0 all unassigned variables. Now, $\boldsymbol{a}$ has the same distribution as the assignment $\boldsymbol{a_{\tilde{\vec{p}}}}$ from Definition 3.7 for the vector of probabilities $\tilde{p}$ given by $\tilde{p}_i\stackrel{\text{def}}{=}p^{\mu(x_i)}/2$. Let $p_i\stackrel{\text{def}}{=}(p/2)^{\mu(x_i)}$; then $p_i\leq\tilde{p}_i$ and, since the event $\bigvee_{\alpha\in[H]}t_\alpha(a)=0$ is anti-monotone in a, we get

$$\mathbf{P}\left[\bigwedge_{\alpha\in[H]}(t_\alpha|_{\boldsymbol{\rho_{\mu,p}}}\not\equiv 1)\right]=\mathbf{P}\left[\bigvee_{\alpha\in[H]}t_\alpha(\boldsymbol{a_{\tilde{\vec{p}}}})=0\right]\leq\mathbf{P}\left[\bigvee_{\alpha\in[H]}t_\alpha(\boldsymbol{a_{\vec{p}}})=0\right]. \tag{18}$$

We are going to upper bound $\mathbf{P}\left[\bigvee_{\alpha\in[H]}t_\alpha(\boldsymbol{a_{\vec{p}}})=0\right]$, given (16).

For every $\alpha\in[H]$ such that $\mu(t_\alpha)<K$ introduce a new auxiliary variable y_α with weight $K-\mu(t_\alpha)$, and replace t_α with $t_\alpha\wedge y_\alpha$. This operation does not change the value of the sum $\sum_{\alpha\neq\beta\in[H]}\delta(Vars(t_\alpha)\cap Vars(t_\beta))$ in (16), and $\mathbf{P}\left[\bigvee_{\alpha\in[H]}t_\alpha(\boldsymbol{a_{\vec{p}}})=0\right]$ may only increase. Therefore, we may assume w.l.o.g. that all terms in F have weight *exactly* K.

And now we apply Proposition 3.8. All events $\mathbf{P}[t_\alpha(\boldsymbol{a_{\vec{p}}})=1]$ have the same probability $(p/2)^K$; therefore, the quantity μ given by (11) is equal to $H(p/2)^K$. Also, whenever $Vars(t_\alpha)\cap Vars(t_\beta)\neq\emptyset$, we have

$$\mathbf{P}[t_\alpha(\boldsymbol{a_{\vec{p}}})=t_\beta(\boldsymbol{a_{\vec{p}}})=1]=(p/2)^{\mu(t_\alpha\wedge t_\beta)}=(p/2)^{2K}\delta(Vars(t_\alpha)\cap Vars(t_\beta));$$

therefore,

$$\Delta = (p/2)^{2K} \sum_{\alpha\neq\beta\in[H]} \delta(Vars(t_\alpha)\cap Vars(t_\beta)) \leq (p/2)^{2K} H(H-1)s^{-1} \leq H(p/2)^K = \mu$$

by (16) and (15). Applying Proposition 3.8, we get

$$\mathbf{P}\left[\bigvee_{\alpha\in[H]} t_\alpha(\boldsymbol{a}_{\vec{\boldsymbol{p}}}) = 0\right] \leq e^{-\frac{H(p/2)^K}{2}} \leq e^{-s/2}$$

which, along with (17) and (18), implies (13).

Case 2: $\delta(F) \geq s^{-1}$**.** Note first that

$$\begin{aligned} \delta(\boldsymbol{t}) &= \mathbf{E}[\delta(Vars(\boldsymbol{t}) \cap Vars(\boldsymbol{t'}))] = \sum_V \delta(V)\cdot \mathbf{P}[Vars(\boldsymbol{t}) \cap Vars(\boldsymbol{t'}) = V] \\ &\leq \sum_V \delta(V)\cdot \mathbf{P}[V \subseteq Vars(\boldsymbol{t}) \wedge V \subseteq Vars(\boldsymbol{t'})] = \sum_{V\neq\emptyset} (2/p)^{\mu(V)} \mathbf{P}[V \subseteq Vars(\boldsymbol{t})]^2 ; \end{aligned}$$

therefore, we also have

$$\min_{\boldsymbol{t}} \sum_{V\neq\emptyset} (2/p)^{\mu(V)} \mathbf{P}[V \subseteq Vars(\boldsymbol{t})]^2 \geq s^{-1}.$$

Let $\boldsymbol{t}$ be the random term on which the quadratic form $\sum_{V\neq\emptyset}(2/p)^{\mu(V)}\mathbf{P}[V \subseteq Vars(\boldsymbol{t})]^2$ in the variables $p_t \stackrel{\text{def}}{=} \mathbf{P}[\boldsymbol{t} = t]$ attains its minimal value $\delta \geq s^{-1}$. Denoting further $\mathbf{P}[V \subseteq Vars(\boldsymbol{t})]$ by p_V, we have

$$\sum_{\substack{t\in F\\ p_t\neq 0}} \frac{\partial(\sum_{V\neq\emptyset}(2/p)^{\mu(V)}p_V^2)}{\partial p_t} p_t = 2\sum_{V\neq\emptyset}(2/p)^{\mu(V)}p_V^2 = 2\delta.$$

Since $\sum_{\substack{t\in F\\ p_t\neq 0}} p_t = 1$, this implies the existence of some $t_0 \in F$ with $p_{t_0} \neq 0$ and such that

$$\frac{\partial(\sum_{V\neq\emptyset}(2/p)^{\mu(V)}p_V^2)}{\partial p_{t_0}} = 2 \sum_{\substack{V\neq\emptyset\\ V\subseteq Vars(t_0)}} (2/p)^{\mu(V)} p_V \geq 2\delta.$$

Therefore, for any term $t \in F$ (even when $p_t = 0$!) we also have $\sum_{\substack{V\neq\emptyset\\ V\subseteq Vars(t)}} (2/p)^{\mu(V)} p_V \geq \delta$ since otherwise, by slightly increasing p_t and decreasing p_{t_0}, we would have obtained contradiction with the assumption that $\boldsymbol{t}$ minimizes $\sum_{V\neq\emptyset}(2/p)^{\mu(V)}\mathbf{P}[V \subseteq Vars(\boldsymbol{t})]^2$.

Summarizing the above argument, we have found coefficients $p_V \geq 0$ (V a non-empty set of variables) and $\delta \geq s^{-1}$ such that

$$\sum_{V \neq \emptyset} (2/p)^{\mu(V)} p_V^2 = \delta; \tag{19}$$

$$\forall t \in F\Big(\sum_{\substack{V \neq \emptyset \\ V \subseteq Vars(t)}} (2/p)^{\mu(V)} p_V \geq \delta \Big). \tag{20}$$

These are the only properties of p_Vs we will need in the rest of the proof; in particular, at this point we can forget their interpretation as certain probabilities.

Let us order all non-empty sets of variables V in the decreasing order with respect to the coefficient p_V: $p_{V_1} \geq p_{V_2} \geq p_{V_3} \geq \ldots$ Assume w.l.o.g. that the variables $x_1, \ldots, x_n$ are enumerated in the order in which they appear for the first time in the sequence $V_1, V_2, V_3, \ldots$ (equivalently, we require that every union $V_1 \cup V_2 \cup \ldots \cup V_H$ should be an initial segment in the set of variables ordered according to their indices). For a non-empty V, denote by $i(V)$ the *maximal* index i for which $x_i \in V$.

Now we classify terms $t \in F$ as follows. Represent a given term $t \in F$ in the form $t = x_{i_1} \wedge x_{i_2} \wedge \ldots \wedge x_{i_w}$ ($i_1 < i_2 < \ldots < i_w$). For a non-empty $V \subseteq Vars(t)$, let $\mu_t(V) \stackrel{\text{def}}{=} \mu(x_{i_1} \wedge x_{i_2} \wedge \ldots \wedge x_{i(V)})$ (in other words, this is the weight of the minimal initial segment in $Vars(t)$ containing V). We split the sum in (20) with respect to the value of $\mu_t(V)$, and classify t according to which part is large enough. Formally, let

$$F_\mu \stackrel{\text{def}}{=} \left\{ t \in F \;\middle|\; \sum_{\substack{\mu_t(V) = \mu \\ V \subseteq Vars(t)}} p_V \geq \delta (p/4)^\mu \right\}.$$

Then, noting that $\mu_t(V) \geq \mu(V)$ for all t, V, we have $F = \bigvee_{\mu=1}^K F_\mu$. This clearly implies $h(F|_\rho) \leq \sum_{\mu=1}^K h(F_\mu|_\rho)$ for every restriction ρ (the concatenation of decision trees strongly representing $F_\mu|_\rho$ for $\mu = 1, \ldots, K$, strongly represents $F|_\rho$). Thus, we obtain

$$\mathbf{P}[H(F|_{\boldsymbol{\rho}}) > h] \leq \sum_{\mu=1}^K \mathbf{P}\Big[H(F_\mu|_{\boldsymbol{\rho}}) > h 2^{-\mu}\Big], \tag{21}$$

and we treat every term in the right-hand side separately.

So, let us fix μ, $1 \le \mu \le K$, and let us fix $t \in F_\mu$; $t = x_{i_1} \wedge x_{i_2} \wedge \ldots \wedge x_{i_w}$ $(i_1 < i_2 < \ldots < i_w)$. Let d be the index for which $\mu(x_{i_1} \wedge x_{i_2} \wedge \ldots \wedge x_{i_d}) = \mu$. Then $d \le \mu$, which implies that there are at most 2^μ subsets $V \subseteq t$ with $\mu_t(V) = \mu$. Consulting the definition of F_μ, we see that there exists a particular $V \subseteq Vars(t)$ with $\mu_t(V) = \mu$ such that $p_V \ge \delta(p/8)^\mu$.

Let this particular V have rank ℓ in our enumeration of all non-empty subsets. We are going to upper bound $\sum_{\nu=1}^{\ell} |V_\nu|$. Notice that, according to the choice of this enumeration, $p_{V_\nu} \ge p_V$ for all $\nu \in [\ell]$. Therefore,

$$\sum_{\nu=1}^{\ell} (2/p)^{\mu(V_\nu)} p_{V_\nu}^2 \ge p_V^2 \cdot \sum_{\nu=1}^{\ell} (2/p)^{\mu(V_\nu)} \ge \delta^2 (p^2/64)^\mu \sum_{\nu=1}^{\ell} \mu(V_\nu) \ge \delta^2 (p^2/64)^\mu \sum_{\nu=1}^{\ell} |V_\nu|.$$

Comparing this with (19), we get

$$\sum_{\nu=1}^{\ell} |V_\nu| \le \delta^{-1} (64/p^2)^\mu \le s(64/p^2)^\mu. \tag{22}$$

The conclusion (22) holds for every $t \in F_\mu$. Therefore, if we define ℓ_μ as the maximal value ℓ for which the bound (22) holds, then for *every* $t \in F_\mu$, $V_1 \cup \ldots \cup V_{\ell_\mu}$ will contain some $V \subseteq Vars(t)$ with $\mu_t(V) = \mu$. Since $V_1 \cup \ldots \cup V_{\ell_\mu}$ is an initial segment in $\{x_1, \ldots, x_n\}$, this implies that $\mu((V_1 \cup \ldots \cup V_{\ell_\mu}) \cap Vars(t)) \ge \mu$.

Let now T_μ be the (oblivious) decision tree that queries all variables in $V_1 \cup \ldots \cup V_{\ell_\mu}$ (in an arbitrary order). Our analysis implies that for every leaf π of T_μ, $(F_\mu)|_\pi$ is a monotone weighted $(K - \mu)$-DNF. Therefore, we may apply the inductive assumption and conclude that

$$\mathbf{P}\left[h((F_\mu|_\pi)|_{\boldsymbol{\rho}}) > \frac{1}{2} h 2^{-\mu}\right] \le e^{-\frac{1}{2} h 2^{-\mu} (p\epsilon)^{2(K-\mu)}}.$$

By the definition of ℓ_μ and (14), the height of T_μ does not exceed $h(p\epsilon)^{2K} \cdot \{O(1/p^2)\}^\mu$, which is at most $\frac{1}{4} h 2^{-\mu} (p\epsilon)^{2(K-\mu)}$, as long as ϵ is small enough. Summing over all leaves of T_μ,

$$\mathbf{P}\left[\exists \pi \in \{V_1 \cup \ldots \cup V_{\ell_\mu}\} \left(h((F_\mu|_\pi)|_{\boldsymbol{\rho}}) > \frac{1}{2} h 2^{-\mu}\right)\right] \le e^{-\frac{1}{4} h 2^{-\mu} (p\epsilon)^{2(K-\mu)}}.$$

Noting that the height of the tree $(T_\mu|_{\boldsymbol{\rho}})$ never exceeds $\frac{1}{2} h 2^{-\mu}$, the latter event is logically implied by $h(F_\mu|_{\boldsymbol{\rho}}) > h 2^{-\mu}$. Comparing with (21),

$$\mathbf{P}[H(F|_{\boldsymbol{\rho}}) > h] \le \sum_{\mu=1}^{K} e^{-\frac{1}{4} h 2^{-\mu} (p\epsilon)^{2(K-\mu)}} \le \sum_{\mu=1}^{K} e^{-h(p\epsilon)^{2K} \cdot \frac{1}{4} \left(\frac{1}{2\epsilon^2}\right)^\mu},$$

which is at most $e^{-h(p\epsilon)^{2K}}$, as long as ϵ is small enough.

This proves (13) in Case 2, completes the inductive step, and the proof of Lemma 4.4 for monotone F.

In order to extend this lemma to the general case, let us note first that, by symmetry, we automatically have it when the DNF F is *pseudo-monotone*, defined as those DNF that can be turned into monotone by negating some variables. Now, every DNF F (interpreted as a set of terms) has a straightforward *fractional* cover by pseudo-monotone DNFs of acceptable size. If it were an ordinary cover, we would have been trivially done. The following lemma (that will also be repeatedly used in both our applications of Lemma 4.4) shows how to handle the more general fractional case.

Lemma 4.5 *Let F be a k-DNF, and $\boldsymbol{G}$ be its random sub-DNF such that*

$$\forall t \in F(\mathbf{P}[t \in \boldsymbol{G}] \geq \epsilon), \tag{23}$$

ϵ an arbitrary parameter. Let $\boldsymbol{\rho}$ be a random restriction such that for every fixed G from the support of $\boldsymbol{G}$,

$$\mathbf{P}[h(G|_{\boldsymbol{\rho}}) > h] \leq \delta,$$

δ another parameter. Then

$$\mathbf{P}\left[h(F|_{\boldsymbol{\rho}}) > h\left(\frac{2k}{\epsilon} + 1\right)\right] \leq 2\delta/\epsilon.$$

Proof of Lemma 4.5. Averaging over the distribution of $\boldsymbol{G}$, we get $\mathbf{P}[h(\boldsymbol{G}|_{\boldsymbol{\rho}}) > h] \leq \delta$, where $\boldsymbol{G}$ and $\boldsymbol{\rho}$ are assumed independent. By Markov's inequality,

$$\mathbf{P}_{\boldsymbol{\rho}}[\mathbf{P}_{\boldsymbol{G}}[h(\boldsymbol{G}|_{\boldsymbol{\rho}}) > h] \geq \epsilon/2] \leq 2\delta/\epsilon.$$

Therefore, we only have to show that for every individual restriction ρ,

$$\mathbf{P}[h(\boldsymbol{G}|_{\rho}) > h] < \epsilon/2 \tag{24}$$

logically implies $h(F|_{\rho}) \leq h\left(\frac{2k}{\epsilon} + 1\right)$.

This is done by an adaptation of the "block sensitivity" method [Nis91, BBC+98] to our setting. Assume that (24) holds. We want to construct a decision tree of height $\leq h\left(\frac{2k}{\epsilon} + 1\right)$ strongly representing $F|_{\rho}$. We begin with a recursive construction of a sequence of decision trees $T_0, T_1, \ldots, T_\ell, \ldots$ that goes as follows.

T_0 is the trivial tree of height 0. In order to construct $T_{\ell+1}$ from T_ℓ, examine one by one all leaves π of T_ℓ. If either $t|_\pi = 0$ for all $t \in F|_\rho$, or $t|_\pi = 1$ for some $t \in F|_\rho$, we leave π alone. Otherwise, pick up an arbitrary non-trivial term $t_{\ell+1} \in (F|_\rho)|_\pi$, and append to the leaf π the oblivious decision tree querying all variables in $Vars(t_{\ell+1}) \setminus (\sup(\rho) \cup \sup(\pi))$. Repeating this procedure for all leaves π of T_ℓ, we get $T_{\ell+1}$.

We terminate this construction after $(2h/\epsilon)$ steps. The only leaves π of $T_{(2h/\epsilon)}$ that still may violate Definition 3.5 are those for which the procedure of appending a new tree was repeated all $(2h/\epsilon)$ times before we arrived at π. Let us fix any such leaf π, and let $t_1, t_2, \ldots, t_{(2h/\epsilon)}$ be the terms picked by our algorithm along the path to π.

By (23),

$$\mathbf{E}\Big[\Big|\{t_1, t_2, \ldots, t_{(2h/\epsilon)}\} \cap \boldsymbol{G}\Big|\Big] \geq 2h.$$

On the other hand, denoting by $\boldsymbol{Bad}$ the indicator function of the event $h(\boldsymbol{G}|_\rho) > h$, (24) implies

$$\mathbf{E}\Big[\Big|\{t_1, t_2, \ldots, t_{(2h/\epsilon)}\} \cap \boldsymbol{G}\Big| \cdot \boldsymbol{Bad}\Big] \leq \frac{2h}{\epsilon} \cdot \mathbf{E}[\boldsymbol{Bad}] < h.$$

Therefore, $\mathbf{E}\Big[\Big|\{t_1, t_2, \ldots, t_{(2h/\epsilon)}\} \cap \boldsymbol{G}\Big| \cdot (1 - \boldsymbol{Bad})\Big] > h$; pick up a particular sub-DNF G that contains $(h+1)$ terms $t_{\alpha_1}, \ldots, t_{\alpha_{h+1}}$ from the list $\{t_1, t_2, \ldots, t_{(2h/\epsilon)}\}$ such that $h(G|_\rho) \leq h$. Let T_π be the decision tree of height $\leq h$ strongly representing $G|_\rho$. We complete out construction by appending the tree T_π to the leaf π (all vertices asking questions from $\sup(\pi)$ are contracted in the process), and repeating this for all leaves π that still violate Definition 3.5.

Clearly, the final tree constructed in this way has height at most $h\left(\frac{2k}{\epsilon} + 1\right)$, and we claim that it strongly represents $F|_\rho$. For that we only have to check the leaves of the form $\pi * \sigma$, where π is a problematic leaf of $T_{(2h/\epsilon)}$, and σ is a leaf of T_π such that π and σ are consistent.

For every $\nu \in [h+1]$, let $V_\nu \stackrel{\text{def}}{=} Vars(t_{\alpha_\nu}) \setminus \bigcup_{\alpha < \alpha_\nu} Vars(t_\alpha)$. Then V_ν are disjoint and non-empty; therefore, since $|\sup(\sigma)| \leq h$, there exists at least one $\nu_0 \in [h+1]$ such that $V_{\nu_0} \cap |\sup(\sigma)| = \emptyset$. Now, the values of π at the variables V_{ν_0} can be changed in such a way that for the resulting partial assignment π' we have $t_{\alpha_{\nu_0}}|_{\pi'} = 1$. $V_{\nu_0} \cap |\sup(\sigma)| = \emptyset$ implies that π' and σ are still consistent, hence $t_{\alpha_{\nu_0}}|_\sigma \not\equiv 0$ and, since $t_{\alpha_{\nu_0}} \in G|_\rho$, $\sigma \in Br_0(T_\pi)$

is impossible. Thus, $\sigma \in Br_1(T_\pi)$ which implies that there exists a term $t \in G|_\rho \subseteq F|_\rho$ such that $t|_\sigma = 1$, and the more $t|_{\pi * \sigma} = 1$.

We have constructed (under the assumption (24)) a tree of height $\leq h\left(\frac{2k}{\epsilon}+1\right)$ strongly representing $F|_\rho$. This completes the proof of Lemma 4.5.■

Let us now finish the proof of Lemma 4.4 for arbitrary weighted K-DNF F. For every (total) assignment a to F, let F_a be the set of all terms in F satisfied by a. Then F_a is a pseudo-monotone sub-DNF of F; therefore, we already have for it the required bound (12). Next, if $\boldsymbol{a}$ is picked completely at random, then for every $t \in F$, $\mathbf{P}[t \in F_{\boldsymbol{a}}] = 2^{-w(t)} \geq 2^{-K}$. Thus, we may apply Lemma 4.5 with $k := K$, $\boldsymbol{G} := F_{\boldsymbol{a}}$, $\epsilon := 2^{-K}$, $h := \frac{h}{2K2^K+1}$ and $\delta := \exp\left(-\frac{h}{2K2^K+1}(p/2)^{O(K)}\right) \leq \exp(-h(p/2)^{O(K)})$, and this completes the proof of Lemma 4.4 in the general situation.■

5. Hardness of Nisan generator for Res(k)

In this section we prove Theorem 2.7. The proof is technically involved, and for that reason is split into a chain of auxiliary claims. They will be assembled together only at the end of the section, although some informal intuition as to what we are doing will be provided as we go along.

Let us fix for the rest of the section an $(m \times n)$ (r, d)-lossless expander A, its ordering $\leq$ and an integer $k \geq 1$ such that (2) holds. The overall strategy of our proof for the first time appeared in [BP96] and since that has become a standard tool in proof complexity. Namely, we want to design a random partial assignment $\boldsymbol{\rho}$ of the variables $Vars_{\leq}(A)$ that has the following two properties:

Height-reduction for every fixed k-DNF F, $h(F|_{\boldsymbol{\rho}})$ is small with high probability;

Width-preservation with high probability every *resolution* refutation of $\tau_{\leq}(A, b)|_{\boldsymbol{\rho}}$ must have large width.

Now, if a small size Res(k) refutation P of $\tau_{\leq}(A, b)$ existed, then with high probability $h(F|_{\boldsymbol{\rho}})$ would be small for *every* $F \in P$, and we could apply Proposition 3.6. Its conclusion, however, would be in the immediate contradiction with the **Width-preservation** property.

We begin realizing this plan with **Width-preservation** (mostly because this part is by far easier), and we will show that large width of resolution refutations of $\tau_{\leq}(A,b)|_{\boldsymbol{\rho}}$ is implied by a simple combinatorial property of $\boldsymbol{\rho}$.

Some conventions on notation. In the rest of the paper we will be abbreviating the extension variables $y^i_{v^i_\Sigma}$ (where Σ is a non-empty $\leq_i$-initial segment of $J_i(A)$) by y^i_Σ. For technical reasons it will be also convenient to introduce the variables $y^i_\emptyset$, along with the axioms $\bar{y}^i_\emptyset$. Likewise, it will be convenient *not* to identify $y^i_{\{j\}}$ with x_j and with each other (as required in the general Definition 2.1), but introduce instead new axioms $y^i_{\{j\}} \vee \bar{x}_j,\ \bar{y}^i_{\{j\}} \vee x_j$. These conventions imply that all variables in the vectors $\vec{x}, \vec{y}^i$ ($i \in [m]$) are pairwise distinct, and the axioms of $\tau_{\leq}(A,b)$ become of particularly symmetric form

$$y^i_\emptyset = 0, \quad y^i_{\Sigma\cup\{j\}} = y^i_\Sigma \oplus x_j, \quad y^i_{J_i(A)} = b_i \quad (i \in [m]) \tag{25}$$

(more precisely, they are clauses of width ≤ 3 resulting from the straightforward CNF expansion of these linear equations mod 2).

Definition 5.1 A restriction ρ of the variables $Vars_{\leq}(A)$ is *closed* if $\rho(y^i_\emptyset) = 0$ and $\rho(y^i_{J_i(A)}) = b_i$ for all $i \in [m]$. Let $J_x(\rho) \stackrel{\text{def}}{=} \{j \in [n] \mid x_j \in \mathrm{sup}(\rho)\}$. ρ is *sparse* if it is closed and for every $i \in [m]$ and every two different segments $\Sigma \subset \Sigma'$ in $J_i(A)$ such that $y^i_\Sigma, y^i_{\Sigma'} \in \mathrm{sup}(\rho)$, we have $|(\Sigma' \setminus \Sigma) \setminus J_x(\rho)| \geq 2d$.

Claim 5.2 *If ρ is sparse then every resolution refutation of $\tau_{\leq}(A,b)|_\rho$ must be of width $> r/4$.*

Proof. Given a closed restriction ρ of the variables $Vars_{\leq}(A)$, define the matrix $A|_\rho$ as follows. For every row i of the original matrix A let $\emptyset = \Sigma^i_0 \subset \Sigma^i_1 \subset \ldots \subset \Sigma^i_{s_i} = J_i(A)$ be the complete list of those initial segments Σ in $J_i(A)$ for which $y^i_\Sigma \in \mathrm{sup}(\rho)$. Then the rows of the matrix $A|_\rho$ are, by definition, indexed by the pairs (i,ν) ($i \in [m], \nu \in [s_i]$), its columns are indexed by $[n] \setminus J_x(\rho)$ (i.e., by those x-variables that are left unassigned by ρ), and the underlying set system is described as $J_{i,\nu}(A|_\rho) \stackrel{\text{def}}{=} (\Sigma^i_\nu \setminus \Sigma^i_{\nu-1}) \setminus J_x(\rho)$.

Note that $A|_\rho$ satisfies all assumptions of Corollary 3.4. Indeed, the bound $|J_{i,\nu}(A|_\rho)| \geq 2d$ is exactly the definition of sparseness. Next, given any set $\{(i_1,\nu_1), \ldots, (i_\ell,\nu_\ell)\}$ of rows in the matrix $A|_\rho$ with $\ell \leq r$, applying the expansion property (1) for the original matrix A to the set $\{i_1, \ldots, i_\ell\}$ gives us

$$\sum_{i \in \{i_1,\ldots,i_\ell\}} |J_i(A)| - |\partial_A(\{i_1,\ldots,i_\ell\})| \leq d \cdot |\{i_1,\ldots,i_\ell\}| \leq d \cdot \ell. \tag{26}$$

On the other hand, it is easy to see that

$$\left.\begin{aligned}\sum_{\alpha=1}^{\ell} |J_{i_\alpha,\nu_\alpha}(A|_\rho)| - |\partial_{A|_\rho}(\{(i_1,\nu_1),\ldots,(i_\ell,\nu_\ell)\})|\\ \le \sum_{i\in\{i_1,\ldots,i_\ell\}} |J_i(A)| - |\partial_A(\{i_1,\ldots,i_\ell\})|;\end{aligned}\right\} \quad (27)$$

the key observation being that the sets $J_{i_\alpha,\nu_\alpha}(A|_\rho)$ are mutually disjoint for a fixed i. Indeed, if for every $j\in[n]$ we let $\ell_j \stackrel{\text{def}}{=} |\{i\in\{i_1,\ldots,i_\ell\} \mid j\in J_i(A)\}|$ then the right-hand side of (27) is exactly equal to $\sum_{j=1}^n(\ell_j \dot{-} 1)$, whereas every $j\in[n]$ contributes at most $\ell_j\dot{-}1$ to the sum on the left-hand side. (26) and (27) imply that $A|_\rho$ is an (r,d)-lossless expander and therefore all assumptions of Corollary 3.4 are fulfilled.

We now define (in a natural way) the assignment $b|_\rho$ of the rows of the matrix $A|_\rho$ by letting

$$(b|_\rho)|_{i,\nu} \stackrel{\text{def}}{=} \rho(y^i_{\Sigma^i_{\nu-1}}) \oplus \rho(y^i_{\Sigma^i_\nu}) \oplus \bigoplus_{j\in(\Sigma^i_\nu\setminus\Sigma^i_{\nu-1})\cap J_x(\rho)} \rho(x_j).$$

Finally, let $\le|_\rho$ be the ordering of the matrix $A|_\rho$, where $(\le|_\rho)|_{i,\nu}$ is the restriction of $\le_i$ onto $J_{i,\nu}(A|_\rho)$.

We extend the restriction ρ to a variable substitution of variables in $Vars_\le(A)$ by variables in $Vars_{\le|_\rho}(A|_\rho)$ defining it outside of $\sup(\rho)$ as follows. All $x_j\notin\sup(\rho)$ are simply left alone: $\rho(x_j)\stackrel{\text{def}}{=}x_j$. For every $y^i_\Sigma\notin\sup(\rho)$ we identify the index ν such that $\Sigma^i_{\nu-1}\subset\Sigma\subset\Sigma^i_\nu$, and let

$$\rho(y^i_\Sigma) \stackrel{\text{def}}{=} y^{(i,\nu)}_{(\Sigma\setminus\Sigma^i_{\nu-1})\setminus J_x(\rho)} \oplus \rho(y^i_{\Sigma^i_{\nu-1}}) \oplus \bigoplus_{j\in(\Sigma\setminus\Sigma^i_{\nu-1})\cap J_x(\rho)} \rho(x_j)$$

(cf. the definition of $b|_\rho$). It is straightforward to check that this extension of the original restriction ρ takes every equation in (25) either to 0=0 or to an equation of the same form corresponding to $\tau_{\le|_\rho}(A|_\rho,b|_\rho)$, and therefore maps every resolution refutation P of $\tau_\le(A,b)$ to a resolution refutation $P|_\rho$ of $\tau_{\le|_\rho}(A|_\rho,b|_\rho)$. By Corollary 3.4 (applied to $(A|_\rho,\le|_\rho,b|_\rho)$), $P|_\rho$ must contain a clause of width $> r/4$. Even more, its pre-image in P has width $> r/4$. Claim 5.2 is proved.■

Our second (and much more complicated) goal is to design a distribution on sparse restrictions that fulfils the **Height-reduction** property. After a short thinking it becomes clear that in full generality this is impossible. Namely, if a term t contains a couple of variables $y^i_\Sigma, y^i_{\Sigma'}$ with Σ, Σ' close to each other than *no* sparse restriction whatsoever can set t to 1, which completely ruins the whole argument. Thus, our most immediate task is to get rid of such "nasty" terms.

More notation. For a term t in the variables $Vars_\leq(A)$, $J_x(t) \stackrel{\text{def}}{=} \{j \in [n] \mid x_j \in Vars(t)\}$ and $\text{dom}(t) \stackrel{\text{def}}{=} \{i \in [m] \mid \exists \Sigma(y^i_\Sigma \in Vars(t))\}$. For a DNF F, $\text{dom}(F) \stackrel{\text{def}}{=} \bigcup_{t\in F} \text{dom}(t)$. For a non-empty $\leq_i$-initial segment Σ, let $r(\Sigma)$ be its right end (maximal element). For uniformity of notation, we let $r(\emptyset) \stackrel{\text{def}}{=} \text{nil}_i$, where nil_i is an imaginary element with $\text{nil}_i <_i j$ for all $j \in J_i(A)$. In this notation, the difference $\Sigma' \setminus \Sigma$ of two segments $\Sigma \subseteq \Sigma'$ coincides with the interval $(r(\Sigma), r(\Sigma')]$. For intervals L, R, $Conv(L, R)$ is the minimal interval containing them both.

Definition 5.3 Let t be a term in the variables $Vars_\leq(A)$. For every $i \in \text{dom}(t)$ list in the form $\Sigma^i_1 \subset \Sigma^i_2 \subset \ldots \subset \Sigma^i_{k_i}$ all segments Σ such that $y^i_\Sigma \in Vars(t)$. Say that t is *protected* if there exists a system of subsets $L_{i\nu}, R_{i\nu} \subseteq J_i(A)$ ($i \in \text{dom}(t)$, $\nu \in [k_i]$) such that:

1. $L_{i,1} <_i r(\Sigma^i_1) <_i R_{i,1} <_i L_{i,2} <_i r(\Sigma^i_2) <_i \ldots <_i r(\Sigma^i_{k_i}) <_i R_{i,k_i}$;
2. $|L_{i\nu}|, |R_{i\nu}| = 3d$;
3. $\sum_{i\in\text{dom}(t)} \sum_{\nu\in[k_i]} |Conv(L_{i\nu}, R_{i\nu})| \leq 20kd$;
4. all $L_{i\nu}, R_{i\nu}$ are disjoint with $J_x(t)$.

The pair $(L_{i\nu}, R_{i\nu})$ (often also written in the form $(L_{i\Sigma}, R_{i\Sigma})$, where $\Sigma = \Sigma^i_\nu$), will be called a *protection of the variable* $y^i_{\Sigma^i_\nu} \in Vars(t)$, and the entire set system $\{L_{i\nu}, R_{i\nu}\}$ will be called a *protection of the term* t (protections are, of course, not necessarily uniquely defined). A DNF F is *protected* if all terms $t \in F$ have this property.

We now show that every term t with $w(t) \leq k$ can be replaced by a protected DNF that will be denoted by $R(t)$.

Definition 5.4 Let t be a term of width $\leq k$ in the variables $Vars_\leq(A)$. Denote by t' the term obtained from t by appending to it the literals $\bar{y}^i_\emptyset \wedge$

$(y^i_{J_i(A)})^{b_i}$ for every $i \in \mathrm{dom}(t)$ (if t contains either $y^i_\emptyset$ or $(y^i_{J_i(A)})^{\bar{b}_i}$ for some $i \in [m]$, we immediately abort the construction and let $R(t) \stackrel{\mathrm{def}}{=} 0$).

For a fixed $i \in \mathrm{dom}(t)$, list all segments Σ with $y^i_\Sigma \in Vars(t')$: $\emptyset = \Sigma^i_0 \subset \Sigma^i_1 \subset \ldots \subset \Sigma^i_{k_i+1} = J_i(A)$, and let $r_{i\nu} \stackrel{\mathrm{def}}{=} r(\Sigma^i_\nu)$. Let

$$\mathrm{Ker}_i(t) \stackrel{\mathrm{def}}{=} J_i(A) \cap (J_x(t) \cup \bigcup_{i' \in \mathrm{dom}(t)\setminus\{i\}} J_{i'}(A)). \tag{28}$$

Construct the graph G_i on $\{0, 1, \ldots, k_i + 1\}$ by connecting $(\nu - 1)$ with ν if $|(r_{i,\nu-1}, r_{i\nu}] \setminus \mathrm{Ker}_i(t)| < 6d$, and let $\Gamma_{i,0}, \Gamma_{i,1}, \ldots, \Gamma_{i\ell_i}$ be the connected components of this graph. $\Gamma_{i\alpha}$ is an interval in $\{0, 1, \ldots, k_i + 1\}$; let $\Gamma_{i\alpha} = [\nu^\ell_{i\alpha}, \nu^r_{i\alpha}]$, and $\hat{\Gamma}_{i\alpha} \stackrel{\mathrm{def}}{=} (r_{i,\nu^\ell_{i\alpha}}, r_{i,\nu^r_{i\alpha}}]$ be the corresponding interval in $J_i(A)$.

Claim 5.5 **a)**

$$\sum_{i \in \mathrm{dom}(t)} |Ker_i(t) \cup \bigcup_{\alpha=0}^{\ell_i} \hat{\Gamma}_{i\alpha}| \leq 14kd; \tag{29}$$

b) *$\ell_i \geq 1$ for every $i \in [m]$ (that is, $y^i_\emptyset$ and $y^i_{J_i(A)}$ are not in the same connected component).*

Proof. a). We have

$$\begin{aligned}
\mathrm{Ker}_i(t) \cup \bigcup_{\alpha=0}^{\ell_i} \hat{\Gamma}_{i\alpha} \quad &= \quad \mathrm{Ker}_i(t) \cup \bigcup_{(\nu-1,\nu)\in G_i} (r_{i,\nu-1}, r_{i\nu}] \\
&= \quad \mathrm{Ker}_i(t) \cup \bigcup_{(\nu-1,\nu)\in G_i} ((r_{i,\nu-1}, r_{i\nu}] \setminus \mathrm{Ker}_i(t)) \\
&= \quad (J_i(A) \cap \bigcup_{i' \in \mathrm{dom}(t)\setminus\{i\}} J_{i'}(A)) \\
&\quad \cup \quad (J_i(A) \cap (J_x(t) \setminus \bigcup_{i' \in \mathrm{dom}(t)\setminus\{i\}} J_{i'}(A))) \\
&\quad \cup \quad \bigcup_{(\nu-1,\nu)\in G_i} ((r_{i,\nu-1}, r_{i\nu}] \setminus \mathrm{Ker}_i(t)).
\end{aligned}$$

Accordingly,

$$\left.\begin{aligned} &\sum_{i\in \mathrm{dom}(t)} |\mathrm{Ker}_i(t) \cup \bigcup_{\alpha=0}^{\ell_i} \hat{\Gamma}_{i\alpha}| \\ &\quad \leq \sum_{i\in \mathrm{dom}(t)} |J_i(A) \cap \bigcup_{i'\in \mathrm{dom}(t)\setminus\{i\}} J_{i'}(A)| \\ &\qquad + \sum_{i\in \mathrm{dom}(t)} |J_i(A) \cap (J_x(t) \setminus \bigcup_{i'\in \mathrm{dom}(t)\setminus\{i\}} J_{i'}(A))| \\ &\qquad + \sum_{i\in \mathrm{dom}(t)} \sum_{(\nu-1,\nu)\in G_i} |(r_{i,\nu-1}, r_{i\nu}] \setminus \mathrm{Ker}_i(t)|. \end{aligned}\right\} \quad (30)$$

The first summand in the right-hand side is bounded by kd due to the expansion property (1). Sets appearing in the second summand are disjoint, therefore, it is bounded by $|J_x(t)| \leq k$. Finally, $|(r_{i,\nu-1}, r_{i\nu}]\setminus \mathrm{Ker}_i(t)| < 6d$ for every $i \in \mathrm{dom}(t)$ and $(\nu-1,\nu) \in G_i$, and also $\sum_{i\in dom(t)} |G_i| \leq \sum_{i\in \mathrm{dom}(t)} k_i \leq 2k$. Therefore, the third summand is bounded by $12kd$. (29) follows.

Part b) of this claim immediately follows from part a) and the bound (2).■

Definition 5.6 (construction of $R(t)$ continued) Fix now one representative $\nu_{i\alpha}$ in every connected component $\Gamma_{i\alpha}$ such that $\nu_{i,0} = 0$ and $\nu_{i\ell_i} = k_i$ (Claim 5.5 b) guarantees that this is possible). Let ϵ^i_α be the sign with which the corresponding variable appears in t'. Define the expression $\tilde{R}(t)$ as the result of replacing in the term t' all variables $y^i_{\Sigma^i_\nu}$ with $\nu \in \Gamma_{i\alpha}$ and $\nu \neq \nu_{i\alpha}$ by the following linear forms:

$$y^i_{\Sigma^i_\nu} \mapsto \bigoplus \left\{ x_j \,\middle|\, j \in \Sigma^i_\nu \triangle \Sigma^i_{\nu_{i\alpha}} \right\} \oplus \epsilon^i_\alpha. \quad (31)$$

Finally, we let $R(t)$ be the straightforward DNF expansion of $\tilde{R}(t)$, in which we also remove all "cosmetic" literals $\bar{y}^i_\emptyset$, $(y^i_{J_i(A)})^{b_i}$ inserted there at the beginning of the construction.

Claim 5.7 **a)** *$R(t)$ is a protected DNF such that $|Vars(R(t))| \leq O(kd)$.*

b) *There exist $Res(O(kd))$ inferences of $R(t)$ from $t, \tau_{\leq}(A,b)$ and, vice versa, of t from $R(t), \tau_{\leq}(A,b)$ that have size $\exp(O(kd))$ and contain at most $O(kd)$ variables.*

Proof. a). The bound on $|Vars(R(t))|$ follows from the construction and Claim 5.5 a). We protect terms in $R(t)$ as follows. Let $y^i_{\Sigma^i_{\nu_{i\alpha}}} \in Vars(R(t))$ (note that $\Sigma^i_{\nu_{i\alpha}} \not\in \{\emptyset, J_i(A)\}$). Protect this variable by the sets $L_{i\alpha}$, $R_{i\alpha}$, where $L_{i\alpha}$ consists of $(3d)$ right-most points in $(r_{\nu^\ell_{i\alpha}-1}, r_{\nu^\ell_{i\alpha}}] \setminus \mathrm{Ker}_i(t)$, and $R_{i\alpha}$ consists of $(3d)$ left-most points in $(r_{\nu^r_{i\alpha}}, r_{\nu^r_{i\alpha}+1}] \setminus \mathrm{Ker}_i(t)$. Let us check that all requirements of Definition 5.3 are fulfilled.

Since $(\nu^\ell_{i\alpha}-1, \nu^\ell_{i\alpha}), (\nu^r_{i\alpha}, \nu^r_{i\alpha}+1) \not\in G_i$, the sets $(r_{\nu^\ell_{i\alpha}-1}, r_{\nu^\ell_{i\alpha}}] \setminus \mathrm{Ker}_i(t)$ and $(r_{\nu^r_{i\alpha}}, r_{\nu^r_{i\alpha}+1}] \setminus \mathrm{Ker}_i(t)$ have cardinality $\geq 6d$ each, which implies 1).

2) is obvious.

For 3) note that the intervals $Conv(L_{i\alpha}, R_{i\alpha})$ are disjoint for every fixed i, and that $Conv(L_{i\alpha}, R_{i\alpha}) \subseteq \mathrm{Ker}_i(t) \cup L_{i\alpha} \cup \hat{\Gamma}_{i\alpha} \cup R_{i\alpha}$. Hence $\sum_{i\in\mathrm{dom}(t)}\sum_\alpha |Conv(L_{i\alpha}, R_{i\alpha})| = \sum_{i\in\mathrm{dom}(t)} |\bigcup_\alpha Conv(L_{i\alpha}, R_{i\alpha})|$; applying Claim 5.5 a) along with the obvious estimate $\sum_{i\in\mathrm{dom}(t)} |\bigcup_\alpha (L_{i\alpha} \cup R_{i\alpha})| \leq 6kd$, we get the required bound.

Finally, 4) immediately follows from the construction. Indeed, given any fixed protection $(L_{i_0\alpha_0}, R_{i_0\alpha_0})$, all new x-variables introduced in $J_{i_0}(A)$ by the substitution (31) either belong to $\mathrm{Ker}_{i_0}(t)$, or belong to $\bigcup_\alpha \hat{\Gamma}_{i_0\alpha}$. Both sets are disjoint with $L_{i_0\alpha_0}, R_{i_0\alpha_0}$.

b). Clearly, the axioms from (25) pertaining to the same connected component of G_i semantically imply the equivalence of both sides in (31) in the presence of the literal $(y^i_{\Sigma^i_{\nu_{i\alpha}}})^{\epsilon^i_\alpha}$, contain the axioms $\bar{y}^i_\emptyset$, $(y^i_{J_i(A)})^{b_i}$, and thus semantically imply $t \equiv R(t)$. Now we only have to refer to the implicational soundness of $\mathrm{Res}(k)$ and to the well-known fact that everything provable in this theory (and, in fact, even in Resolution) also has a proof of size at most exponential in the number of variables.■

Remark 5 It is worth noting that the construction of $R(t)$ in fact gives yet another property crucial for the argument: the protections $L_{i\alpha}, R_{i\alpha}$ are pairwise disjoint (even when i varies). This property, however, will be re-established in the proof of Claim 5.13 anyway, and for this reason is not included in Definition 5.3.

Claim 5.8 *Assume that $\tau_\leq(A, b)$ has a Res(k) refutation of size S. Then it also has an Res($O(kd)$) refutation of size $S\cdot\exp(O(kd))$ in which all lines are of the form $F \vee F'$, where F is a protected $O(kd)$-DNF, and $|Vars(F')| \leq O(kd)$.*

Proof. For a k-DNF F, let $R(F) \stackrel{\mathrm{def}}{=} \bigvee_{t\in F} R(t)$. Claim 5.7 b) implies that every axiom $C \in \tau_{\leq}(A, b)$ gets converted to a DNF $R(C)$ that has an inference P from $\tau_{\leq}(A, b)$ of size $\exp(O(kd))$ and with $|Vars(P)| \leq O(kd)$. Also, the image of any inference rule in Definition 2.6 is admissible which can be seen by decoding and then encoding back the principal formula. For example, we simulate the R-image of AND-introduction as follows:

$$\dfrac{\dfrac{\dfrac{R(F) \vee R(\ell_1)}{R(F) \vee \ell_1} \quad \cdots \quad \dfrac{R(F) \vee R(\ell_w)}{R(F) \vee \ell_w}}{R(F) \vee \bigwedge_{\nu=1}^{w} \ell_\nu}}{R(F) \vee R(\bigwedge_{\nu=1}^{w} \ell_\nu)}$$

In order to do this encoding/decoding, we simply use the inferences from Claim 5.7 b) weakened by $R(F)$. Since weakening does not change the number of lines, this induced inference has all the required properties.■

The claims proved so far will allow us to concentrate for the purpose of **Height-reduction** on protected DNFs, and we now define a distribution on sparse restrictions that, with an overwhelming probability, will reduce the representation height of any such DNF.

Definition 5.9 The random restriction $\boldsymbol{\rho}$ of the variables $Vars_{\leq}(A)$ is constructed as follows. Pick a subset $\boldsymbol{J} \subseteq [n]$ completely at random. Then for every $i \in [m]$ independently apply the following construction.

Pick a random set of variables $\tilde{\boldsymbol{Y}}^{\boldsymbol{i}} \subseteq \{y^i_\Sigma \mid \Sigma \text{ a segment of } J_i(A)\}$ by including there the cosmetic variables $y^i_\emptyset$ and $y^i_{J_i(A)}$ with probability 1, and every other y^i_Σ with probability $1/(2d)$, independently of each other. Say that $y^i_\Sigma, y^i_{\Sigma'} \in \tilde{\boldsymbol{Y}}^{\boldsymbol{i}}$ with $\Sigma \subset \Sigma'$ *collide* if $|(\Sigma' \setminus \Sigma) \setminus \boldsymbol{J}| < 2d$. Remove from $\tilde{\boldsymbol{Y}}^{\boldsymbol{i}}$ all variables $y^i_\Sigma \notin \{y^i_\emptyset, y^i_{J_i(A)}\}$ that collide with at least one other variable in $\tilde{\boldsymbol{Y}}^{\boldsymbol{i}}$ (possibly, with $y^i_\emptyset$ or $y^i_{J_i(A)}$). Let $\boldsymbol{Y}^{\boldsymbol{i}} \subseteq \tilde{\boldsymbol{Y}}^{\boldsymbol{i}}$ be the resulting set.

$\boldsymbol{\rho}$, by definition, assigns $y^i_\emptyset$ to 0, assigns $y^i_{J_i(A)}$ to b_i and assigns all other variables in $\{x_j \mid j \in \boldsymbol{J}\} \cup \bigcup_{i=1}^m \boldsymbol{Y}^{\boldsymbol{i}}$ at random (and independently of each other).

Claim 5.10 $\mathbf{P}[\boldsymbol{\rho}$ *is sparse*$] \geq 1/2$.

Proof. By inspecting definitions, $\boldsymbol{\rho}$ may be not sparse only in the "pathological" case when for some $i \in [m]$ we have $\boldsymbol{Y^i} = \{y^i_\emptyset, y^i_{J_i(A)}\}$. We bound the probability of this bad event separately for every $i \in [m]$.

By the bound (2), we may choose $s \geq \frac{C \log_2 m}{9}$ disjoint intervals $\Delta_1, \ldots, \Delta_s$ in $J_i(A)$, of length $9d$ each. Subdivide every Δ_ν into three sub-intervals $\Delta^\ell_\nu, \Delta^c_\nu, \Delta^r_\nu$, where $|\Delta^\ell_\nu| = |\Delta^r_\nu| = 4d$ and $|\Delta^c_\nu| = d$. Then the following events:

- $|\Delta^\ell_\nu \cap \boldsymbol{J}| \leq 2d$, $|\Delta^r_\nu \cap \boldsymbol{J}| \leq 2d$;
- $\boldsymbol{\tilde{Y}^i}$ contains exactly one variable y^i_Σ with $r(\Sigma) \in \Delta^c_\nu$ and no variables with $r(\Sigma) \in \Delta^\ell_\nu \cup \Delta^r_\nu$

have probability $\geq \Omega(1)$ each, are independent and logically imply that the variable $y^i_\Sigma \in \boldsymbol{\tilde{Y}^i}$ with $r(\Sigma) \in \Delta^c_\nu$ does not collide with any other variable and hence stays in $\boldsymbol{Y^i}$. Therefore, the probability that this happens for at least one $\nu \in [s]$ (and, in particular, $\boldsymbol{Y^i} \neq \{y^i_\emptyset, y^i_{J_i(A)}\}$) is at least $1 - 1/(2m)$, provided the constant C in (2) is large enough. By the union bound, this implies $\mathbf{P}\left[\forall i \in [m](\boldsymbol{Y^i} \neq \{y^i_\emptyset, y^i_{J_i(A)}\})\right] \geq 1/2$.∎

We are going to apply Lemma 4.4 to show that for any protected DNF, $h(F|_{\boldsymbol{\rho}})$ is small with high probability. For that we need to know that the restriction $\boldsymbol{\rho}$, when restricted to the set of variables $Vars(F)$, satisfies the weak independence property from Definition 4.3. The intuitive reason why it should be the case is already suggested by the proof of Claim 5.10 (the role of the "wings" $\Delta^\ell_\nu, \Delta^r_\nu$ in that proof will be played by protections). The major problem is, of course, that protections need not be disjoint, may be inconsistent for different occurrences of the same variable etc. We circumvent this by showing that F has a relatively small fractional cover by sub-DNFs for which these problems already do not occur, whereupon we will apply Lemma 4.5.

More notation. Given a protected DNF F, we fix once and for all protections $\{L_{i\Sigma}(t), R_{i\Sigma}(t)\}$ for every $t \in F$. Let also $P_{i\Sigma}(t) \stackrel{\text{def}}{=} L_{i\Sigma}(t) \cup R_{i\Sigma}(t)$.

Definition 5.11 A protected DNF F is *weakly regular* if the following three properties hold:

1. the sets $\bigcup_{t\in F} J_x(t)$ and $\bigcup_{t\in F} \bigcup_{i,\Sigma} P_{i\Sigma}(t)$ are disjoint;

2. for every variable y^i_Σ appearing in more than one term $t \in F$, its protection $(L_{i\Sigma}(t), R_{i\Sigma}(t))$ does not depend on t (in which case it will be denoted simply by $(L_{i\Sigma}, R_{i\Sigma})$);

3. for every **fixed** $i \in [m]$, the intervals $Conv(L_{i\Sigma}, R_{i\Sigma})$ ($y^i_\Sigma \in Vars(F)$) are pairwise disjoint.

Remark 6 The adjective "weakly" refers to the fact that we do not require the protections $P_{i\Sigma}$ to be disjoint for *different* i. Recall (Remark 5) that this disjointness property automatically follows from our construction for protections in any *fixed* term $t \in F$. This, however, seems to be of absolutely no help whatsoever for the uniform version, when t varies. The only way to enforce the disjointness of protections uniformly we know of additionally requires $|F|$ to be small, and for that reason will be incorporated in the proof of Claim 5.13.

Claim 5.12 *For every protected $O(kd)$-DNF F, there exists a random weakly regular sub-DNF $\boldsymbol{G} \subseteq F$ such that* $\min_{t\in F} \mathbf{P}[t \in \boldsymbol{G}] \geq \exp(-O(kd))$.

Proof. The proof consists of three independent steps, and at every step we enforce one property required in Definition 5.11.

Step 1. For a colouring $\chi : [n] \longrightarrow \{0,1\}$, let

$$G_\chi \stackrel{\mathrm{def}}{=} \left\{ t \in F \;\middle|\; \chi(J_x(t)) \equiv 0 \wedge \chi\Big(\bigcup_{t\in F} \bigcup_{i,\Sigma} P_{i\Sigma}(t)\Big) \equiv 1 \right\}.$$

Then for every χ, G_χ has property 1) in Definition 5.11. Let $\boldsymbol{\chi} : [n] \longrightarrow \{0,1\}$ be picked completely at random. Then, due to property 3) in Definition 5.3, $\mathbf{P}[t \in G_{\boldsymbol{\chi}}] \geq \exp(-O(kd))$ for every particular $t \in F$.

Step 2. The idea behind enforcing property 2) in Definition 5.11 is similar to Step 1, but calculations get substantially more involved. For any system of protections $\vec{P} = (P_{i\Sigma} \mid y^i_\Sigma \in Vars(F))$, the sub-DNF $G_{\vec{P}} \stackrel{\mathrm{def}}{=} \{ t \in F \mid \forall y^i_\Sigma \in Vars(t)(P_{i\Sigma}(t) = P_{i\Sigma}) \}$ has the required property 2). We only need to construct a random system $\vec{\boldsymbol{P}}$ in such a way that

$$\min_{t\in F} \mathbf{P}[t \in G_{\vec{\boldsymbol{P}}}] \geq \exp(-O(kd)). \tag{32}$$

All $\boldsymbol{P_{i\Sigma}}$ will be mutually independent. In order to construct $\boldsymbol{P_{i\Sigma}}$ for a fixed variable $y^i_\Sigma \in Vars(F)$, pick first an integer $\boldsymbol{\mu_{i\Sigma}} \geq 6d$ according to the distribution $\mathbf{P}[\boldsymbol{\mu_{i\Sigma}} = \mu] = 2^{6d-\mu-1}$ ($\mu \geq 6d$). Next, let $\boldsymbol{\Delta_{i\Sigma}}$ be the interval of length $2\boldsymbol{\mu_{i\Sigma}}$ centered at $r(\Sigma)$ (if $|\Sigma| < \boldsymbol{\mu_{i\Sigma}}$ or $|J_i(A) \setminus \Sigma| < \boldsymbol{\mu_{i\Sigma}}$, then we abort the construction and output $\boldsymbol{P_{i\Sigma}}$ arbitrarily). Finally, pick $\boldsymbol{P_{i\Sigma}}$ as a random subset of $\boldsymbol{\Delta_{i\Sigma}}$ that has cardinality $6d$.

Consider an arbitrary $t \in F$, and let $\mu_{i\Sigma}(t) \stackrel{\rm def}{=} |Conv(L_{i\Sigma}(t), R_{i\Sigma}(t))|$. Then $\mathbf{P}[\boldsymbol{\mu_{i\Sigma}} = \mu_{i\Sigma}(t)] \geq \exp(-O(\mu_{i\Sigma}(t)))$, and $\mathbf{P}[\boldsymbol{P_{i\Sigma}} = P_{i\Sigma}(t) \mid \boldsymbol{\mu_{i\Sigma}} = \mu_{i\Sigma}(t)] \geq \binom{2\mu_{i\Sigma}(t)}{6d}^{-1} \geq \left(\frac{\mu_{i\Sigma}(t)}{d}\right)^{-O(d)}$. Combining these bounds together, we get $\mathbf{P}[\boldsymbol{P_{i\Sigma}} = P_{i\Sigma}(t)] \geq \exp(-O(\mu_{i\Sigma}(t))) \cdot \left(\frac{\mu_{i\Sigma}(t)}{d}\right)^{-O(d)}$. Multiplying over all $y^i_\Sigma \in Vars(t)$,

$$\mathbf{P}[t \in G_{\vec{\boldsymbol{P}}}] \geq \prod_{y^i_\Sigma \in Vars(t)} \left\{ \exp(-O(\mu_{i\Sigma}(t))) \cdot \left(\frac{\mu_{i\Sigma}(t)}{d}\right)^{-O(d)} \right\}.$$

Since $\sum_{y^i_\Sigma \in Vars(t)} \mu_{i\Sigma}(t) \leq O(kd)$ by property 3) in Definition 5.3, the first term in this product is bounded from below by $\exp(-O(kd))$. In order to bound the second term, we apply the arithmetic-geometric mean inequality. Denoting the number of y-variables in t by w (which, due to properties 2), 3) in Definition 5.3, is bounded by $O(k)$), we have the calculation

$$\prod_{y^i_\Sigma \in Vars(t)} \left(\frac{\mu_{i\Sigma}(t)}{d}\right)^{-O(d)} \geq \left(\frac{\sum_{y^i_\Sigma \in Vars(t)} \mu_{i\Sigma}(t)}{dw}\right)^{-O(dw)} \geq (O(k/w))^{-O(dw)} \geq \exp(-O(kd)).$$

This proves (32) and concludes the analysis at Step 2.

Step 3. For this step we additionally assume that F already satisfies the consistency property 2) in Definition 5.11, and let $\vec{P}$ be the corresponding system of protections of the variables $y^i_\Sigma \in Vars(F)$. Choose $\tilde{\boldsymbol{Y}}^{\boldsymbol{i}}$ according to Definition 5.9. Say that $y^i_\Sigma, y^i_{\Sigma'} \in Vars(F) \cap \tilde{\boldsymbol{Y}}^{\boldsymbol{i}}$ *$\vec{P}$-collide* if $Conv(L_{i\Sigma}, R_{i\Sigma}) \cap Conv(L_{i\Sigma'}, R_{i\Sigma'}) \neq \emptyset$. For every $\vec{P}$-colliding pair $y^i_\Sigma, y^i_{\Sigma'}$ identify arbitrarily any one point j in the above intersection, and remove from $\tilde{\boldsymbol{Y}}^{\boldsymbol{i}}$ that variable $y^i_{\Sigma_0}$ ($\Sigma_0 \in \{\Sigma, \Sigma'\}$) *for which the $\leq_i$-distance between $r(\Sigma_0)$ and j is larger* (both y^i_Σ and $y^i_{\Sigma'}$ if these distances are equal). Let $\boldsymbol{Y}^{\boldsymbol{i}}(\vec{P})$ be the set of remaining variables.

Clearly, the sub-DNF $G_{\boldsymbol{Y}(\vec{P})} \stackrel{\rm def}{=} \left\{ t \in F \,\middle|\, Vars(t) \subseteq \{x_1, \ldots, x_n\} \cup \bigcup_{i \in \mathrm{dom}(t)} \boldsymbol{Y}^{\boldsymbol{i}}(\vec{P}) \right\}$ has the required property 3). Further, every variable y^i_Σ may be removed from $\tilde{\boldsymbol{Y}}^{\boldsymbol{i}}$ only if it $\vec{P}$-collides with some $y^i_{\Sigma'}$ such that the $\leq_i$-distance between $r(\Sigma)$ and $r(\Sigma')$ does not exceed $2|Conv(L_{i\Sigma}, R_{i\Sigma})|$.

Now, for every particular term $t \in F$ mark all those variables $y^i_{\Sigma'}$ in $Vars(F)$ for which the above bound on the $\leq_i$-distance holds for at least one $y^i_\Sigma \in Vars(t)$. Property 3) in Definition 5.3 implies that we altogether have marked at most $O(kd)$ variables. Therefore, the event "for every marked variable y^i_Σ, $y^i_\Sigma \in \bigcup_{i\in \mathrm{dom}(t)} \tilde{\boldsymbol{Y}}^{\boldsymbol{i}}$ if and only if $y^i_\Sigma \in Vars(t)$" has probability $\geq \exp(-O(kd))$. On the other hand, since variables in $Vars(t)$ never $\vec{P}$-collide with each other (due to property 1) in Definition 5.3), this event logically implies that $Vars(t) \subseteq \{x_1, \ldots, x_n\} \cup \bigcup_{i\in \mathrm{dom}(t)} \boldsymbol{Y}^{\boldsymbol{i}}(\vec{P})$. Thus, $\mathbf{P}\left[t \in G_{\boldsymbol{Y}(\vec{P})}\right] \geq \exp(-O(kd))$, which completes the analysis of Step 3.

Combining things together, the random sub-DNF $G_{\boldsymbol{\chi}} \cap G_{\vec{\boldsymbol{P}}} \cap G_{\boldsymbol{Y}(\vec{\boldsymbol{P}})}$ (where $\boldsymbol{\chi}$, $\vec{\boldsymbol{P}}$ and the auxiliary random variables $\tilde{\boldsymbol{Y}}^{\boldsymbol{i}}$ are independent) has all the required properties. This completes the proof of Claim 5.12.∎

We assume that any weakly regular DNF F is weighted according to the following weight function μ_F: $\mu_F(x_j) \stackrel{\text{def}}{=} 1$ and $\mu_F(y^i_\Sigma) \stackrel{\text{def}}{=} |Conv(L_{i\Sigma}, R_{i\Sigma})|$.

Claim 5.13 *There exists an absolute constant $p > 0$ such that for every weakly regular protected DNF F, the random restriction $\boldsymbol{\rho}$ restricted to the variables $Vars(F)$ is (r, μ_F, p)-independent.*

Proof. Let $Z \subseteq Vars(F)$ be a set of variables with $|Z| \leq r$. Denote by I the set of all $i \in [m]$ for which Z contains at least one variable y^i_Σ. Clearly, $|I| \leq r$. Therefore, by the expansion property (1), there exists $i_0 \in I$ such that

$$|J_{i_0}(A) \cap \bigcup_{i \in I \setminus \{i_0\}} J_i(A)| \leq d. \tag{33}$$

This implies that for every Σ with $y^{i_0}_\Sigma \in Vars(F)$ there exist $L'_{i_0\Sigma} \subseteq L_{i_0\Sigma}$ and $R'_{i_0\Sigma} \subseteq R_{i_0\Sigma}$ of size $2d$ each that are disjoint with $\bigcup_{i\in I\setminus\{i_0\}} J_i(A)$. In particular, these $L'_{i_0\Sigma}, R'_{i_0\Sigma}$ are disjoint not only between themselves (by property 3) in Definition 5.11), but also with any other $P_{i\Sigma}$ when $i \neq i_0$. Next, choose $i_1 \in I \setminus \{i_0\}$ such that $|J_{i_1}(A) \cap \bigcup_{i\in I\setminus\{i_0,i_1\}} J_i(A)| \leq d$ and repeat this procedure until we find pairwise disjoint subsets $L'_{i\Sigma} \subseteq L_{i\Sigma}$, $R'_{i\Sigma} \subseteq R_{i\Sigma}$ for all $y^i_\Sigma \in Z$. Let as before $P'_{i\Sigma} \stackrel{\text{def}}{=} L'_{i\Sigma} \cup R'_{i\Sigma}$.

Now, the space Ω required in Definition 4.1 consists of sets of variables, and the role of $\boldsymbol{\omega}$ is played by the random set $\{x_j \mid j \in \boldsymbol{J}\} \cup \bigcup_{i=1}^m \tilde{\boldsymbol{Y}}^{\boldsymbol{i}}$ from Definition 5.9. The mapping $\boldsymbol{\rho}(J, \tilde{Y}^i)$ is provided by the construction from that definition ($\boldsymbol{\rho}$ uses additional independent Bernoulli variables for assigning $\{x_j \mid j \in \boldsymbol{J}\} \cup \bigcup_{i=1}^m \boldsymbol{Y^i}$).

Next, we construct the second required mapping $\boldsymbol{\rho_{\mu_F,P}}(J, \tilde{Y^i})$ as follows. First, define from J and $\tilde{Y}^i$ random subsets $\boldsymbol{J'} \subseteq J$ and $\boldsymbol{(Y^i)'} \subseteq \tilde{Y}^i \cap Z$. Let $j \in \boldsymbol{J'}$ with probability $2p$, independently for all $j \in J \cap Z$. If $y^i_\Sigma \in \tilde{Y}^i \cap Z$ and either $J \cap P'_{i\Sigma} \neq \emptyset$ or $\tilde{Y}^i$ contains any other $y^i_{\Sigma'}$ with $r(\Sigma') \in Conv(L'_{i\Sigma}, R'_{i\Sigma})$, then $y^i_\Sigma \notin \boldsymbol{(Y^i)'}$. Otherwise, y^i_Σ is included into $\boldsymbol{(Y^i)'}$ with some probability $p_{i\Sigma}$ to be specified later, independently for all variables y^i_Σ (and independently of $\boldsymbol{J'}$). Finally, $\boldsymbol{\rho_{\mu_F,P}}(J, \tilde{Y^i})$ independently and at random assigns all variables in $\{x_j \in Z \mid j \in \boldsymbol{J'}\} \cup \bigcup_{i \in I} \boldsymbol{(Y^i)'}$ using the same Bernoulli variables as $\boldsymbol{\rho}$ above.

$y^i_\Sigma \in \boldsymbol{(Y^i)'}$ can take place only if $J \cap P'_{i\Sigma} = \emptyset$ and $\tilde{Y}^i$ does not contain any $y^i_{\Sigma'}$ with $r(\Sigma') \in Conv(L'_{i\Sigma}, R'_{i\Sigma})$ other than y^i_Σ. This implies that y^i_Σ can not collide with any other variable in $\tilde{Y}^i$. Therefore, $\boldsymbol{(Y^i)'} \subseteq \boldsymbol{Y^i}$, and $\boldsymbol{\rho_{\mu_F,P}}(J, \tilde{Y^i})$ is a sub-restriction of $\boldsymbol{\rho}(J, \tilde{Y^i})$.

Finally, the properties listed in Definition 5.11, along with the disjointness property of $L'_{i\Sigma}, R'_{i\Sigma}$ ensured at the beginning of our proof, imply that the facts $j \in \boldsymbol{J'}$, $y^i_\Sigma \in \boldsymbol{(Y^i)'}$ ($x_j, y^i_\Sigma \in Z$) depend on the behaviour of $J, \tilde{Y}^i$ on pairwise different variables. Therefore, if we randomize also over $\boldsymbol{J}, \boldsymbol{\tilde{Y}^i}$, all these events are independent of each other, and, as an easy calculation shows,

$$\left.\begin{aligned} \mathbf{P}\Big[y^i_\Sigma \in \boldsymbol{(Y^i)'}\Big] &= 2^{-4d} \cdot \frac{1}{2d}\left(1 - \frac{1}{2d}\right)^{|Conv(L'_{i\Sigma}, R'_{i\Sigma})|} p_{i\Sigma} \\ &\geq 2^{-4d} \cdot \frac{1}{2d}\left(1 - \frac{1}{2d}\right)^{\mu(y^i_\Sigma)} p_{i\Sigma} \\ &\geq \exp(-O(\mu(y^i_\Sigma)))p_{i\Sigma}, \end{aligned}\right\} \tag{34}$$

where for the last estimate we used the obvious bound $\mu(y^i_\Sigma) \geq 6d$. Therefore, if $p > 0$ is small enough, the probabilities $p_{i\Sigma} \leq 1$ can be chosen in such a way that $\mathbf{P}\Big[y^i_\Sigma \in \boldsymbol{(Y^i)'}\Big] = p^{\mu(y^i_\Sigma)}$, and, clearly, $\mathbf{P}[j \in \boldsymbol{J'}] = p$ for $x_j \in Z$. Thus, $\boldsymbol{\rho_{\mu_F,P}}(\boldsymbol{J}, \boldsymbol{\tilde{Y^i}})$ has the same distribution as $\boldsymbol{\rho_{\mu_F,P}}$, which completes the proof of Claim 5.13.■

Now we have at our disposal all tools necessary for proving the following switching lemma for protected DNFs.

Claim 5.14 *Let F be a protected $O(kd)$-DNF. Then for every $h \le r$,*

$$\mathbf{P}[h(F|_{\boldsymbol{\rho}}) \ge h] \le \exp(-h/2^{O(kd)}).$$

Proof. Assume first that F is weakly regular, and let $\mu = \mu_F$. Then property 3) in Definition 5.3 implies that F is even a *weighted* $O(kd)$-DNF. With this remark, our claim for weakly regular F immediately follows from Lemma 4.4 and Claim 5.13.

In order to generalize this to the case of arbitrary protected $O(kd)$-DNF, we use Lemma 4.5 in combination with Claim 5.12, just in the same way as at the end of Section 4.∎

Now we can easily finish the proof of Theorem 2.7. Let P be a Res(k) refutation of $\tau_{\le}(A, b)$ that has size S. We need to prove that $S \ge \exp(r/2^{O(kd)})$.

Applying Claim 5.8, for some $K = O(kd)$ we get a Res(K) refutation P' of $\tau_{\le}(A, b)$ that has size $\le S2^K$, and in which every line has the form $F \vee F'$, where F is a protected K-DNF and $|Vars(F')| \le K$. By Claims 5.10, 5.2,

$$\mathbf{P}[\text{every resolution refutation of } \tau_{\le}(A,b)|_{\boldsymbol{\rho}} \text{ has width } > r/4] \ge 1/2. \quad (35)$$

Comparing this with Proposition 3.6 ($k := K$ and $h := r/(4K)$), we get

$$\mathbf{P}[\exists G \in P'(h(G|_{\boldsymbol{\rho}}) > r/(4K))] \ge 1/2. \quad (36)$$

On the other hand,

$$\mathbf{P}[\exists G \in P'(h(G|_{\boldsymbol{\rho}}) > r/(4K))] \le S2^K \cdot \max_{G \in P'} \mathbf{P}[h(G|_{\boldsymbol{\rho}}) > r/(4K)], \quad (37)$$

and we treat every line $G \in P'$ individually. Let $G = F \vee F'$, where F is a protected K-DNF and $|Vars(F')| \le K$. Obviously, $h(F'|_{\boldsymbol{\rho}}) \le K$, and we can assume w.l.o.g. that $K < r/(8K)$ (otherwise, the bound we are proving becomes trivial). Thus,

$$\mathbf{P}[h(G|_{\boldsymbol{\rho}}) > r/(4K)] \le \mathbf{P}[h(F|_{\boldsymbol{\rho}}) > r/(8K)]. \quad (38)$$

Applying Claim 5.14 (with $h := r/(8K) - 1$), we find

$$\mathbf{P}[h(F|_{\boldsymbol{\rho}}) > r/(8K)] \le \exp(-r/2^{O(kd)}). \quad (39)$$

Theorem 2.7 follows by comparing (39), (38) and (37) with (36).

6. Stretching the number of output bits

In this section we prove Theorems 2.10, 2.12.

Proof of Theorem 2.10. Let A be an $(m \times n)$ (r,d)-lossless expander such that the bound (5) holds, and let $\leq$ be an arbitrary ordering of A. Let S be the minimal size of a Res(k) refutation P of any CNF having the form

$$\bigwedge_{\nu=1}^{H} \tau_{C_{\leq,A}}(x_1^{(\nu)},\ldots,x_n^{(\nu)},\vec{y}^{(\nu)},q_1^{(\nu)},\ldots,q_m^{(\nu)}), \tag{40}$$

from Definition 2.9. First we remark that we may assume w.l.o.g. that $H \leq S$.

Indeed, P may contain at most S axioms; therefore, there exist at most S indices $\nu \in [H]$ for which P contains an axiom from $\tau_{C_{\leq,A}}(x_1^{(\nu)},\ldots,x_n^{(\nu)},\vec{y}^{(\nu)},q_1^{(\nu)},\ldots,q_m^{(\nu)})$. Remove from (40) all other conjunctive terms; then P will still be a Res(k) refutation of the resulting sub-CNF. This sub-CNF itself has the form (40), with the only exception that some $q_i^{(\nu)}$ may be equal to variables not appearing in the lists $\vec{x}^{(\nu)},\vec{y}^{(\nu)}$. Substituting them arbitrarily with Boolean constants 0,1, we get a CNF of the form (40) with $H \leq S$ that still has a Res(k) refutation of size S.

Note that (40) encodes the system of linear equations

$$\bigoplus_{j\in J_i(A)} x_j^{(\nu)} = q_i^{(\nu)} \;\; (\nu \in [H],\ i \in [m]).$$

Let us transfer those $q_i^{(\nu)}$ which actually are the variables from the list $\left\{x_j^{(\mu)} \mid j \in [n],\ \mu < \nu\right\}$ to the left-hand side. Then we get a linear system with constant right-hand side, and it also has the form $\hat{A}X = \hat{b}$ for some matrix $\hat{A}$ and vector $\hat{b}$. It will turn out that $\hat{A}$ has almost as good expansion properties as A itself, and we will apply to it Theorem 2.7.

In order to formalize this intuition, let A^H be the *direct sum* of H copies of A. That is, rows of A^H are indexed by $[H] \times [m]$, columns are indexed by $[H] \times [n]$ and $a^H_{(\nu,i),(\mu,j)} \stackrel{\text{def}}{=} \begin{cases} a_{ij} & \text{if } \nu = \mu \\ 0 & \text{if } \nu \neq \mu \end{cases}$. Let $\hat{A}$ be obtained from A^H by setting additionally $\hat{a}_{(\nu,i),(\mu,j)} := 1$ whenever $q_i^{(\nu)} = x_j^{(\mu)}$. Let also $\hat{b}_{(\nu,i)} \stackrel{\text{def}}{=} \begin{cases} q_i^{(\nu)} & \text{if } q_i^{(\nu)} \text{ is a Boolean constant} \\ 0 & \text{if } q_i^{(\nu)} \text{ is a variable} \end{cases}$. Finally, we order the rows of

A^H according to the ordering $\leq$ of the matrix A. Whenever we add in $\hat{A}$ the new 1-entry (μ, j) to the row (i, ν) (that is, when $q_i^{(\nu)} = x_j^{(\mu)}$), we declare it to be the *largest* element in $J_{(i,\nu)}(\hat{A})$. Denote the resulting ordering of $\hat{A}$ by $\hat{\leq}$.

Given the above intuition, it is straightforward to check that there exists a (naturally defined) variable substitution that takes every clause in (40) into a clause having a resolution inference from $\tau_{\hat{\leq}}(\hat{A}, \hat{b})$ of size $O(1)$. This implies that $\tau_{\hat{\leq}}(\hat{A}, \hat{b})$ has a Res(k) refutation of size $O(S)$.

On the other hand, it is easy to see that the direct sum of (r, d)-lossless expanders is still an (r, d)-lossless expander. Also, if we append at most one 1-entry per row to an (r, d)-lossless expander, we come up with an $(r, d+1)$-lossless expander. Applying Theorem 2.7, we see that either the condition (2) is violated for $\hat{A}$ or we have $S \geq \exp(r/2^{O(kd)})$ (in which case we are done). It remains to note that in the first case we have $s \leq Cd(k + \log_2(mH))$ which implies (if the constant in (5) is twice as large as the constant in (2)) that $H \geq \exp(\Omega(s/d))$. Since $S \geq H$, Theorem 2.10 follows in this case as well.■

Proof of Theorem 2.12. (cf. [Kra02]) Given any Boolean circuit C with n inputs and $2n$ outputs and any $b \in \{0,1\}^{2^h \cdot n}$, $\tau(C^h, b)$ can be expressed in the form (4) as $\bigwedge_{|u| \leq h-1} \tau_C(x_1^{(u)}, \ldots, x_n^{(u)}, \vec{y}^{(u)}, q_1^{(u)}, \ldots, q_{2n}^{(u)})$, where

$$q_i^{(u)} \stackrel{\text{def}}{=} \begin{cases} x_i^{(u*0)} & \text{if } |u| < h-1,\ i \leq n \\ x_{n-i}^{(u*1)} & \text{if } |u| < h-1,\ i \geq n+1 \\ b_{(u*0),i} & \text{if } |u| = h-1,\ i \leq n \\ b_{(u*1),i} & \text{if } |u| = h-1,\ i \geq n+1. \end{cases}$$

Now Theorem 2.12 immediately follows from Theorem 2.10.■

7. Random matrices have good expansion properties

The statements of this sort have been re-appearing in the literature at a steady rate beginning from [Pin73]. We, however, have not been able to find any particular source handling the matter in the generality needed for our purposes. Thus, we prove Theorem 2.5 from the scratch (this is not hard anyway).

Let
$$r \stackrel{\text{def}}{=} n^{\delta}, \quad d \stackrel{\text{def}}{=} C\frac{\log_2 m}{\log_2 n},$$
where δ is a sufficiently small and C a sufficiently large constant. Note that $d \leq n^{\epsilon}$, where ϵ is the constant from the statement of the theorem. We claim that if a set I of rows with $|I| = \ell \leq r$ in the matrix $\boldsymbol{A_{m,n}}$ violates the expansion property (1), then there exists a set of t columns J, where $t \leq \ell d$, such that $\boldsymbol{A_{m,n}}$ contains at least $t + \ell d$ ones in the rectangle $I \times J$. Indeed, let $J \stackrel{\text{def}}{=} \{j \in [n] \mid |\{i \in I \mid j \in J_i(A)\}| \geq 2\}$ $(= \bigcup_{i \in I} J_i(A) \setminus \partial_A(I))$. If $|J| \leq \ell d$, then J itself has the required property; otherwise its arbitrary subset of cardinality exactly ℓd would do. Calling a rectangle $I \times J$ with such properties *dense*, it remains to show that the probability of existence of at least one dense rectangle is $O(1/m)$.

For any fixed values of ℓ, t, there exist at most m^{ℓ} choices of I, at most n^t choices of J, and at most $(\ell t)^{t+\ell d}$ choices of $t + \ell d$ positions in $I \times J$. Therefore, the probability that for given ℓ, t there exists at least one dense $(\ell \times t)$ rectangle does not exceed
$$\begin{aligned} m^{\ell} n^t (\ell t)^{t+\ell d} (n^{-2/3})^{t+\ell d} &= m^{\ell} n^{t/3} (\ell t)^{t+\ell d} n^{-2\ell d/3} \leq m^{\ell} n^{\ell d/3} (\ell^2 d)^{2\ell d} n^{-2\ell d/3} \\ &= m^{\ell} \left(\frac{(\ell d)^{O(1)}}{n}\right)^{\ell d/3} \leq m^{\ell} \left(\frac{(rd)^{O(1)}}{n}\right)^{\ell d/3}. \end{aligned}$$
Next, this is $\leq m^{\ell} n^{-\Omega(\ell d)}$ provided the constants ϵ, δ are small enough which, in turn, is $\leq m^{-2\ell}$ provided the constant C in the definition of d is large enough.

This is the bound on the probability that $\boldsymbol{A_{m,n}}$ contains a dense rectangle $I \times J$ with $|I| = \ell,\ |J| = t$. Therefore, since $t \leq \ell d$, the probability that $\boldsymbol{A_{m,n}}$ is not an (r, d)-lossless expander is bounded by
$$\sum_{\ell=1}^{\infty} (\ell d) m^{-2\ell} \leq C \log_2 m \cdot \sum_{\ell=1}^{\infty} \ell m^{-2\ell} \leq O(1/m).$$
Theorem 2.5 is proved.

8. Unprovability of circuit lower bounds by small Res(k) proofs

The general idea toward extracting Theorem 2.13 from a pseudorandom function generator was extensively discussed in Introduction. We are, however, dealing with a rather weak proof system and, moreover, Definition 2.2 severely restricts the choice of the encoding for the circuit $C^h_{A,\leq}$. Thus, we should be careful in checking that the natural reduction can be indeed carried over with the limited tools at our disposal (cf. the previous arguments of this sort in [Raz98, RanRaz02, Raz02b, Kra02]).

Let f_n be a Boolean function in n variables and $n^2 \leq t \leq 2^n$. We begin with reproducing the formal definition of the CNF $Circuit_t(f_n)$ from [Raz98, Raz02b].

First, we list all variables of $Circuit_t(f_n)$ (some of them have peculiar long names like $InputType'_\nu(v)$), along with their intended meaning:

y_{av} $(a \in \{0,1\}^n,\ v \in [t])$	–	the Boolean value computed at the computational node v on the input string a;
$y_{a\nu v}$ $(a \in \{0,1\}^n,\ \nu \in \{1,2\},\ v \in [t])$	–	the value brought to v by ν's wire on a;

$Fanin(v)$	–	this is 0 if v is NOT-gate and 1 if v is AND-gate or OR-gate;
$Type(v)$	–	when $Fanin(v) = 1$, this is 0 if v is AND-gate and 1 if v is OR-gate;
$InputType_\nu(v)$	–	this is 0 if ν's input to v is a constant or a variable and 1 if it is one of the previous computational gates;
$InputType'_\nu(v)$	–	when $InputType_\nu(v) = 0$, this is 0 if ν's input to v is a constant, and 1 if it is a variable;
$InputType''_\nu(v)$	–	when $InputType_\nu(v) = InputType'_\nu(v) = 0$, this equals the ν's input to v;

$InputVar_\nu(v,i)$ $(i \in [n])$	–	when $InputType_\nu(v) = 0$, $InputType'_\nu(v) = 1$, this is 1 iff ν's input to v is the ith variable;
$INPUTVAR_\nu(v,i)$	–	equals $\bigvee_{i' \leq i} InputVar(v,i')$, introduced to keep bottom fan-in bounded;
$InputNode_\nu(v,v')$ $(v' < v)$	–	when $InputType_\nu(v) = 1$, this is 1 iff ν's input to v is the previous gate v';
$INPUTNODE_\nu(v,v')$	–	analogously to $INPUTVAR_\nu(v,i)$.

$Circuit_t(f_n)$ is the conjunction of (conjunctive normal forms equivalent to) the following axioms:

$$\neg InputType_\nu(v) \wedge \neg InputType'_\nu(v) \longrightarrow (y_{a\nu v} \equiv InputType''_\nu(v));$$
$$\neg InputType_\nu(v) \wedge InputType'_\nu(v) \longrightarrow \neg(InputVar_\nu(v,i) \wedge InputVar_\nu(v,i'))\ (i \neq i');$$
$$\neg InputType_\nu(v) \wedge InputType'_\nu(v) \longrightarrow (INPUTVAR_\nu(v,i) \equiv (INPUTVAR_\nu(v,i-1) \vee InputVar_\nu(v,i)))\ (INPUTVAR_\nu(v,0) \stackrel{\text{def}}{=} 0);$$
$$\neg InputType_\nu(v) \wedge InputType'_\nu(v) \longrightarrow INPUTVAR_\nu(v,n);$$
$$\neg InputType_\nu(v) \wedge InputType'_\nu(v) \wedge InputVar_\nu(v,i) \longrightarrow (y_{a\nu v} \equiv a_i);$$

the analogous group of axioms for $InputNode$;

$$\neg Fanin(v) \longrightarrow (y_{av} \equiv \neg y_{a1v});$$
$$Fanin(v) \wedge \neg Type(v) \longrightarrow (y_{av} \equiv (y_{a1v} \wedge y_{a2v}));$$
$$Fanin(v) \wedge Type(v) \longrightarrow (y_{av} \equiv (y_{a1v} \vee y_{a2v}));$$
$$y_{at} \equiv f(a).$$

Let $t_0 \stackrel{\text{def}}{=} \delta\sqrt{t/n}$, δ a sufficiently small constant. Note that $t_0 \geq t^{\Omega(1)}$ (since $t \geq n^2$). Fix an arbitrary $(2t_0 \times t_0)$ $(t^{\Omega(1)}, O(1))$-lossless expander A such that $|J_i(A)| \geq t^{\Omega(1)}$ for all $i \in [2t_0]$ (its existence follows from Theorem 2.5). Fix an arbitrary ordering $\leq$ of A. Let $b \in \{0,1\}^{2^{n+1}t_0}$ be given by $b_{aj} \stackrel{\text{def}}{=} f_n(a_1, \ldots, a_n)$ $(a \in \{0,1\}^{n+1},\ j \in [t_0])$.

In order to prove Theorem 2.13, we are going to reduce $\tau(C^{n+1}_{A,\leq}, b)$ to $Circuit_t(f_n)$ (that is, substitute variables of $Circuit_t(f_n)$ by "simple" formulas in the variables of $\tau(C^{n+1}_{A,\leq}, b)$ so that "simple" refutations of $Circuit_t(f_n)$ get transformed into simple refutations of $\tau(C^{n+1}_{A,\leq}, b)$). For that purpose we will convert the circuit $C^{n+1}_{A,\leq}$ (with t_0 inputs and $2^{n+1}t_0$ outputs, of which we are interested only in 2^n) to the (single-output) Boolean circuit $D_{n,\vec{x}}$ in n Boolean variables $z_1, \ldots, z_n$ parameterized by t_0 variables $x_1, \ldots, x_{t_0}$.

The skeleton of $D_{n,\vec{x}}$ consists of the gates $v[\ell, i, \Sigma]$, where $0 \leq \ell \leq n$, $i \in [2t_0]$ and Σ an initial segment in $J_i(A)$ computing Boolean functions $f_{\vec{x}}[\ell, i, \Sigma]$ in the variables $z_1, \ldots, z_n$. The values $f_{\vec{x}}[\ell, i, \Sigma](a_1, \ldots, a_n)$ of these functions are defined as follows. We take the circuit $(C_{A,\leq})_{a_1 \ldots a_\ell}$ from Definition 2.11, look at the gate v^i_Σ in this circuit (that we will denote in what follows by $v^i_\Sigma(a_1 \ldots a_\ell)$) and output as $f_{\vec{x}}[\ell, i, \Sigma](a_1, \ldots, a_n)$ its value when $\tau(C^{n+1}_{A,\leq}, b)$ is fed with $x_1, \ldots, x_{t_0}$ (we also naturally let $f_{\vec{x}}[\ell, i, \emptyset](a_1, \ldots, a_n) \stackrel{\text{def}}{=} 0$). Then these functions have the following recursive definition

$$\left.\begin{array}{rcl} f_{\vec{x}}[\ell, i, \emptyset] & := & 0 \\ f_{\vec{x}}[0, i, \Sigma \cup \{j\}] & := & f_{\vec{x}}[0, i, \Sigma] \oplus x_j \\ f_{\vec{x}}[\ell, i, \Sigma \cup \{j\}] & := & f_{\vec{x}}[\ell, i, \Sigma] \oplus \{(\bar{z}_\ell \wedge f_{\vec{x}}[\ell-1, j, J_j(A)]) \\ & & \qquad \vee (z_\ell \wedge f_{\vec{x}}[\ell-1, j+t_0, J_{j+t_0}(A)])\},\ \ell \geq 1. \end{array}\right\} \quad (41)$$

We define $D_{n,\vec{x}}$ as the circuit resulting from expanding these recursive definitions in the standard basis $\{\neg, \wedge, \vee\}$. We let the output gate of $D_{n,\vec{x}}$ be $v[n, 1, J_i(A)]$ (cf. the definition of b above).

$D_{n,\vec{x}}$ has size $O(t_0^2 n)$ which is at most t if the constant δ in the definition of t_0 is small enough. We begin constructing the required substitution ρ by assigning first all structural variables $Fanin(v), Type(v)$ etc., except for $InputType''_\nu(v)$, to appropriate Boolean constants describing the topology of $D_{n,\vec{x}}$. We also let $\rho(InputType''_\nu(v)) \stackrel{\text{def}}{=} x_j$ whenever v is the gate of $D_{n,\vec{x}}$ (necessarily resulting from an instruction in (41) of the second type) whose ν's input is x_j. In the case $InputType_\nu(v) = 1$ we let $\rho(y_{a\nu v}) \stackrel{\text{def}}{=} \rho(y_{av'})$ ($\rho(y_{av'})$ themselves are yet to be defined), where (v', v) is the ν's input leading to v in $D_{n,\vec{x}}$. If $\neg InputType_\nu(v)$, then ν's input to v is either one of the variables z_i or a known Boolean constant $\epsilon \in \{0, 1\}$ or one of the unknown constants x_j, and we let $\rho(y_{a\nu v})$ be a_i, ϵ or x_j, respectively. Further, if the gate $v[\ell, i, \Sigma]$

explicitly appears in (41), then we let $\rho(y_{a,v[\ell,i,\Sigma]}) \stackrel{\mathrm{def}}{=} y^i_{v^i_\Sigma(a_1\ldots a_\ell)}$ (which is the variable of $\tau(C^{n+1}_{A,\leq}, b)$).

We, however, can not extend ρ as a variable substitution to the remaining variables y_{av}, where v is an auxiliary gate resulting from expanding instructions in (41) in the standard basis $\{\neg, \wedge, \vee\}$. But even in this case we still can let $\rho(y_{av})$ be a Boolean function of just two variables $y^i_{v^i_\Sigma(a_1\ldots a_\ell)}$ and $y^{j^*}_{v^{j^*}_{J_{j^*}(A)}(a_1\ldots a_{\ell-1})}$, where j^* is determined by j and a_ℓ.

Summarizing the above argument, we have constructed a substitution ρ that takes every variable of $Ciruit_t(f_n)$ to a Boolean function depending of at most two variables of $\tau(C^{n+1}_{A,\leq}, b)$. Let us extend this substitution to k-DNF formulas F by letting $\rho(F) \stackrel{\mathrm{def}}{=} \bigvee_{t\in F} \widehat{\rho(t)}$, where $\widehat{\rho(t)}$ is the straightforward DNF expansion of $\rho(t)$. Clearly, $\rho(F)$ is a $2k$-DNF, and the ρ-image of any inference rule of $\mathrm{Res}(k)$ can be simulated in $\mathrm{Res}(2k)$ by an inference of size $\exp(O(k))$. Further, given the “intended interpretation”, it is easy to check by inspection that for every axiom C of $Circuit_t(f_n)$, $\rho(C)$ has a $\mathrm{Res}(2)$ inference of size $O(1)$ from $\tau(C^{n+1}_{A,\leq}, b)$.

These remarks imply that every $\mathrm{Res}(k)$ refutation of $Circuit_t(f_n)$ of size S gives rise to a $\mathrm{Res}(2k)$ refutation of $\tau(C^{n+1}_{A,\leq}, b)$ of size $S\cdot\exp(O(k))$. The proof of Theorem 2.13 is now completed by applying Theorem 2.12 (with $n := t_0$; $r, s \geq t^{\Omega(1)}$; $d \leq O(1)$; $h := n+1$).

9. Pigeonhole principle

In this section we prove Theorem 2.15. Review first some more material from [SBI02].

Let $m = 2n$ (the same proof works with $m = (1+\epsilon)n$, for any absolute constant $\epsilon > 0$). For a bipartite graph $G = ([m] \cup [n], E)$ ($E \subseteq [m] \times [n]$), let $\neg onto - PHP(G)$ be the CNF in the variables $\{x_{ij} \mid (i,j) \in E\}$ which is obtained from $\neg onto - PHP^m_n$ by the restriction assigning to 0 all variables $\{x_{ij} \mid (i,j) \notin E\}$. Let $\boldsymbol{\rho}^{\boldsymbol{G}}$ be the random restriction constructed in the following way. Pick a random subset $\boldsymbol{J} \subseteq [n]$ by including there every $j \in [n]$ with probability $1/4$ independently of each other. For $j \in \boldsymbol{J}$ select uniformly (and independently for different j) one neighbour $\boldsymbol{i}_j$ of j in G, assign $x_{\boldsymbol{i}_j j}$ to 1, and assign to 0 all other x_{ij} with $(i,j) \in E$ and $i \neq \boldsymbol{i}_j$.

Proposition 9.1 ([SBI02]) *There exists a graph G of maximal degree $O(\log n)$ such that*

$$\mathbf{P}\Big[\textit{every resolution refutation of } \neg onto - PHP(G)|_{\rho^G} \textit{ has width } > n/24\Big] \geq 1/2.$$

Proof. This is essentially [SBI02, Lemma 18], with the only difference that we have additionally included the group of onto axioms Q_j. This does not affect its proof in any way.∎

Fix from now on any particular graph G with properties from Proposition 9.1, and let $\Delta \leq O(\log n)$ be its maximal degree. Since $\neg onto - PHP(G)$ is obtained from $\neg onto - PHP_n^m$ by a restriction, it is sufficient to prove the required bound for $\neg onto - PHP(G)$. The proof follows the pattern laid out in Section 5, although in the current case it is much simpler.

Definition 9.2 A term in the variables $\{x_{ij} \mid (i,j) \in E\}$ is *reduced* if it is monotone and does not contain any sub-term of the form $x_{ij} \wedge x_{i'j}$, $i \neq i'$. A DNF F is *reduced* if all terms $t \in F$ have this property.

Reduced DNFs will play the role of protected ones in Section 5.

Definition 9.3 For a term t, let us denote by $R(t)$ the reduced DNF that is constructed as follows. Let $\tilde{R}(t)$ be the result of replacing in the term t all negative literals $\bar{x}_{ij}$ with $\bigvee \{x_{i'j} \mid (i',j) \in E \wedge i' \neq i\}$. Let $R(t)$ be the straightforward DNF expansion of $\tilde{R}(t)$, in which we remove all terms containing at least one sub-term of the form $x_{ij} \wedge x_{i'j}$ $(i' \neq i)$.

Claim 9.4 *Let t be a term of width $\leq k$ in the variables $\{x_{ij} \mid (i,j) \in E\}$.*

- **a)** *$R(t)$ is a reduced DNF.*
- **b)** *There exist $Res(O(k))$-inferences of $R(t)$ from $t, \neg onto - PHP(G)$ and, vice versa, of t from $R(t), \neg onto - PHP(G)$ that have size $\Delta^{O(k)}$ and contain at most $k\Delta$ variables.*

Proof. By inspection.∎

Claim 9.5 *Assume that $\neg onto - PHP(G)$ has a $Res(k)$ refutation of size S. Then it also has an $Res(O(k))$ refutation of size $S \cdot \Delta^{O(k)}$ in which all lines are of the form $F \vee F'$, where F is a reduced k-DNF, and $|Vars(F')| \leq k\Delta$.*

Proof. As in the proof of Claim 5.8, for a DNF F let $R(F) \stackrel{\mathrm{def}}{=} \bigvee_{t\in F} R(t)$. Notice that R does not change the monotone axioms Q_i, Q_j, and that $R(Q_{i_1,i_2;j}) = Q_j$. With these remarks in mind, the rest of the proof is identical to the proof of Claim 5.8.∎

Finally, let us prove the following PHP-oriented switching lemma.

Claim 9.6 *For any reduced k-DNF F and any parameter h,*

$$\mathbf{P}[h(F|_{\boldsymbol{\rho}}) > h] \leq \exp(-h/\Delta^{O(k)}).$$

Proof. Call a reduced DNF F *regular* if $Vars(F)$ does not contain any pair of variables $x_{ij}, x_{i'j}$ with $i \neq i'$. Then for every regular F, $\boldsymbol{\rho}^{\boldsymbol{G}}$ acts independently on the variables from $Vars(F)$, and, moreover, $\mathbf{P}\left[\boldsymbol{\rho}^{\boldsymbol{G}}(x_{ij}) = 0\right], \mathbf{P}\left[\boldsymbol{\rho}^{\boldsymbol{G}}(x_{ij}) = 1\right] \geq \frac{1}{4\Delta}$. This implies that $\boldsymbol{\rho}^{\boldsymbol{G}}$ is $(\infty, \mu_{triv}, 1/(2\Delta))$-independent. Applying Lemma 4.4, we prove our claim in the case F is regular.

For the general case we apply the same trick as before. Let the function $\boldsymbol{\theta}$: $[2n] \longrightarrow [n]$ be picked completely at random, and $G_\theta \stackrel{\mathrm{def}}{=} \{t \in F \mid \forall x_{ij} \in Vars(t)(i = \theta(j))\}$. Then G_θ is regular, and $\forall t \in F(\mathbf{P}[t \in G_{\boldsymbol{\theta}}] \geq \Delta^{-k})$. Now the proof is completed by applying Lemma 4.5.∎

The proof of Theorem 2.15 is completed in exactly the same way as the proof of Theorem 2.7 at the end of Section 5.

10. Obfuscation

In this section we prove Theorem 2.16.

The proof of our basic result in Section 5 uses the expansion property of A three times. We need it for Claim 5.2 (since it refers to Corollary 3.4). In the proof of Claim 5.5 a) it is used for bounding the first summand in the right-hand side of (30). Finally, it is also crucial for constructing subprotections $L'_{i\Sigma}, R'_{i\Sigma}$ in the proof of Claim 5.13. In all three cases, however, the expansion property is de facto applied not to the entire rows $J_i(A)$, but rather to their subsets which are disjoint unions of not too many not too large intervals. This motivates the following definition.

Definition 10.1 An ordered $(m \times n)$ 0-1 matrix A is an *(r, H, d)-interval expander* if for every set system $\Delta_1, \ldots, \Delta_\ell \subseteq [n]$ such that $\ell \leq r$, every Δ_ν

is an interval in one of $J_i(A)$ of length $\leq H$ and intervals corresponding to the same i are mutually disjoint, we have

$$\sum_{\nu=1}^{\ell} |\Delta_\nu| - |\partial(\Delta_1, \ldots, \Delta_\ell)| \leq d\ell.$$

Here $\partial(\Delta_1, \ldots, \Delta_\ell)$ is the set of all $j \in [n]$ that are contained in exactly one of the intervals $\Delta_1, \ldots, \Delta_\ell$.

Note that (r, d)-lossless expanders are just (r, ∞, d)-interval expanders (independently of the ordering $\leq$). For this reason, the following lemma refines Theorem 2.7.

Lemma 10.2 *Let $(A, \leq)$ be an $(m \times n)$ (r, H, d)-interval expander, and assume that*

$$\min\{H, \min_{i\in[m]} |J_i(A)|\} \geq Cd(k + \log_2(mn)) \tag{42}$$

for a sufficiently large constant $C > 0$. Let $b \in \{0, 1\}^m$ be an arbitrary vector. Then every Res(k) refutation of $\tau_{\leq}(A, b)$ must be of size $\geq \exp(r/2^{O(kd)})$.

Proof. We proceed by adjusting the proof of Theorem 2.7; in particular, we utilize all the notation from Section 5 (except, of course, that now A is only assumed to be an (r, H, d)-interval expander). The following notion is in a sense dual to the notion of a sparse restriction.

Definition 10.3 A restriction ρ of the variables $Vars_{\leq}(A)$ is *bounded* if for every $i \in [m]$ and every two segments $\Sigma \subset \Sigma'$ in $J_i(A)$ such that $y^i_\Sigma, y^i_{\Sigma'} \in \sup(\rho)$ but $y^i_{\Sigma''} \notin \sup(\rho)$ for any $\Sigma \subset \Sigma'' \subset \Sigma'$, we have $|\Sigma' \setminus \Sigma| \leq H$.

Claim 10.4 *If ρ is sparse and bounded then every resolution refutation of $\tau_{\leq}(A, b)|_\rho$ must be of width $> r/4$.*

Proof of Claim 10.4. By inspecting the proof of Claim 5.2. Note that every row (i, ν) of $A|_\rho$ is contained in the interval $\Sigma^i_\nu \setminus \Sigma^i_{\nu-1}$ of $J_i(A)$; these intervals are disjoint for a fixed i and (since ρ is bounded) have length $\leq H$. Checking the proof that $A|_\rho$ is a (r, d)-lossless expander, we see that this actually follows from the weaker assumption that A is an (r, H, d)-expander. The rest of the proof is the same.■

Claim 10.5 $\mathbf{P}[\boldsymbol{\rho}$ *is not bounded*$] \leq 1/4$.

Proof of Claim 10.5. Similar to the proof of Claim 5.10. There are at most mn^2 choices of the "bad" pair (Σ, Σ') in the definition of bounded restriction. For every one of them, pick $s \geq \Omega(H/d)$ disjoint intervals $\Delta_1, \ldots, \Delta_s$ in $\Sigma' \setminus \Sigma$ of size $9d$ each. Arguing as in the proof of Claim 5.10, the probability that at least one of them contains a point in $\boldsymbol{Y^i}$ (and henceforth (Σ, Σ') is actually good), is $\geq 1 - \exp(-\Omega(H/d))$ which is $\geq 1 - (4mn^2)^{-1}$ provided the constant C in (42) is large enough.∎

Claims 10.4 and 10.5 take care of the first use of the expansion property in Section 5.

In the proof of Claim 5.5 a) it is used for bounding the first summand $\sum_{i \in \mathrm{dom}(t)} \left| J_i(A) \cap \bigcup_{i' \in \mathrm{dom}(t) \setminus \{i\}} J_{i'}(A) \right|$ in (30). Let P_i be the set of all points in $J_i(A)$ that either belong to $J_x(t)$ or are within distance $17kd$ from at least one of $r_{i\nu}$ ($\nu \in [k_i]$). Then P_i is the union of at most $2k$ intervals of length $O(kd)$ each, and this union can be made disjoint by shrinking some of the intervals. Now we can use the *interval* expansion property and get the bound $\sum_{i \in \mathrm{dom}(t)} \left| P_i \cap \bigcup_{i' \in \mathrm{dom}(t) \setminus \{i\}} P_{i'} \right| \leq kd$. Replace in the definition (28) of $\mathrm{Ker}_i(A)$ the sets $J_i(A)$ with P_i. Then the proof of Claim 5.5 a), as well as checking the first three requirements of Definition 5.3 in the proof of Claim 5.7 a), goes in exactly the same way as before. Claim 5.5 a) implies that P_i in fact contains $\hat{\Gamma}_{i\alpha}$ ($0 \leq \alpha \leq \ell_i$) and $L_{i\alpha}, R_{i\alpha}$ ($1 \leq \alpha \leq \ell_i - 1$) (all these points are within distance $17kd$ from one of the two end-points of $\hat{\Gamma}_{i\alpha}$). With this remark in mind, generalization of the proof of the remaining property 4) is straightforward.

Likewise, the proof of Claim 5.13 can be easily modified to work with the system of intervals $\{Conv(L_{i\alpha}, R_{i\alpha})\}$. That is, instead of a *row* i_0 with the property (33), we pick up an *individual protection* $L_{i_0\Sigma_0}, R_{i_0\Sigma_0}$ such that

$$|Conv(L_{i_0\Sigma_0}, R_{i_0\Sigma_0}) \cap \bigcup_{(i,\Sigma) \neq (i_0,\Sigma_0)} Conv(L_{i\Sigma}, R_{i\Sigma})| \leq d,$$

and proceed in a similar way one protection at a time.

We have shown how to replace the assumption that A is an (r, d)-lossless expander by the weaker assumption that it is an (r, H, d)-interval expander everywhere it is used in the proof of Theorem 2.7. This proves Lemma 10.2.∎

Now we only have to show that random orderings lead to good interval expanders; i.e., we need a (dual) analogue of Theorem 2.5.

Lemma 10.6 *Let $m \leq 2^{n^\epsilon}$, where ϵ is a sufficiently small constant, A be any fixed $(m \times n)$ matrix such that $s \stackrel{\rm def}{=} \max_{i\in[m]} J_i(A) \geq n^{2/3}$ and $\leq$ be its random ordering. Then $(A, \leq)$ is an $\left(n^{\Omega(1)}, n^{\Omega(1)}, O\left(1+\frac{\log m}{\log n}\right)\right)$-interval expander with probability $\geq 1 - O(1/(mn))$.*

Proof. We apply the same strategy as in Section 7. Let

$$r \stackrel{\rm def}{=} H \stackrel{\rm def}{=} n^\delta, \quad d \stackrel{\rm def}{=} C\left(1 + \frac{\log_2 m}{\log_2 n}\right),$$

where δ is a sufficiently small and C a sufficiently large constant. Similarly to that proof, violation of the interval expansion condition for intervals $\Delta_1, \ldots, \Delta_\ell$ ($\ell \leq r$, $|\Delta_\nu| \leq H$, intervals corresponding to the same i are mutually disjoint) implies the existence of a set of columns J with $|J| = t \leq \ell d$ such that $\sum_{\nu=1}^\ell |\Delta_\nu \cap J| \geq t + \ell d$. Let $\boldsymbol{\theta_i} : J_i(A) \longrightarrow [|J_i(A)|]$ be the random bijection that associates with $j \in J_i(A)$ its rank in the ordering $\leq_{\boldsymbol{i}}$. For $\nu \in [\ell]$, let $\boldsymbol{\mu_\nu} : \Delta_\nu \longrightarrow [|J_i(A)|]$ be the restriction of $\boldsymbol{\theta_i}$ onto Δ_ν, where i is the row for which Δ_ν is an interval in $J_i(A)$.

The number of choices for the ℓ integer intervals $\boldsymbol{\mu_1}(\Delta_1), \ldots, \boldsymbol{\mu_\ell}(\Delta_\ell)$ does not exceed $(mn^2)^\ell$. Like in the proof of Theorem 2.5, there are at most n^t choices of J, and (since $\sum_{\nu=1}^\ell |\Delta_\nu| \leq H\ell$) at most $(H\ell)^{t+\ell d}$ choices for the sets $T_\nu \subseteq \boldsymbol{\mu_\nu}(\Delta_\nu \cap J)$ such that $\sum_{\nu=1}^\ell |T_\nu| = t + \ell d$.

Having fixed $\boldsymbol{\mu_1}(\Delta_1), \ldots, \boldsymbol{\mu_\ell}(\Delta_\ell), J, T_1, \ldots, T_\ell$, let T^i be the (disjoint) union of those T_ν for which Δ_ν is an interval in $J_i(A)$. $T^i \subseteq [|J_i(A)|]$, and the events $T_\nu \subseteq \boldsymbol{\mu_\nu}(\Delta_\nu \cap J)$ mean exactly that $\boldsymbol{\theta_i}^{-1}(T^i) \subseteq J$. Since the bijection $\boldsymbol{\theta_i}$ is completely random, $\mathbf{P}\left[\boldsymbol{\theta_i}^{-1}(T^i) \subseteq J\right] \leq \frac{\binom{t}{|T^i|}}{\binom{s}{|T^i|}} \leq (t/s)^{|T^i|}$, and since $\boldsymbol{\theta_i}$s are also independent for different i, the probability that $\boldsymbol{\mu_1}(\Delta_1), \ldots, \boldsymbol{\mu_\ell}(\Delta_\ell), J, T_1, \ldots, T_\ell$ is bad is at most $(t/s)^{t+\ell d}$. Altogether, for fixed ℓ, t we have the bound $(mn^2)^\ell \cdot n^t \cdot (H\ell)^{t+\ell d} \cdot (t/s)^{t+\ell d}$ which is $\leq (mn^2)^{-2\ell}$ by a calculation similar to that in the proof of Theorem 2.5. The rest of the proof is the same.∎

Theorem 2.16 immediately follows from Lemma 10.6 and Lemma 10.2.

11. Polynomial Calculus with Resolution

Throughout this section we fix an arbitrary field $\mathbb{F}$ with $char(\mathbb{F}) \neq 2$. First we need the following generalization of Corollary 3.4.

Corollary 11.1 *Let A be an (r,d)-lossless expander of size $(m \times n)$ such that $|J_i(A)| \geq 2d$ for all $i \in [m]$. Then for every ordering $\leq$ and every $b \in \{0,1\}^m$, every PCR refutation of $\tau_{\leq}(A,b)$ must have degree $> r/8$.*

Proof. Although this follows by the technique of [BGIP99] (cf. remark in [ABRW00] before Theorem 3.10), it is easier to apply more general result from [BI99] that for our purposes can be stated as follows. If a CNF τ results from expanding a set of $\mathbb{F}_2$-linear equations, then every PCR refutation of τ over $\mathbb{F}$ (remember that $char(\mathbb{F}) \neq 2$) gives rise to a *resolution* refutation of τ whose width is at most twice as large as the degree of P. Since $\tau_{\leq}(A,b)$ always has this "linear" form, Corollary 11.1 follows from Corollary 3.4.■

Now the proofs of Theorems 2.19, 2.20, 2.21 and 2.22 are more or less straightforward adaptation of the corresponding results for Resolution (= Res(1)).

Proof of Theorem 2.19. Monomials in the variables $x_1, \ldots, x_n, \bar{x}_1, \ldots, \bar{x}_n$ can be identified, via the transformation (6) and up to a multiplicative constant $\alpha \in \mathbb{F}^*$, with respective clauses. In this way, variable substitutions naturally act on polynomials from $\mathbb{F}[x_1, \ldots, x_n, \bar{x}_1, \ldots, \bar{x}_n]$ and take PCR inferences to PCR inferences of the same (or lesser) size. This remark, along with Corollary 11.1, implies the following analogue of Claim 5.2 (proved in exactly the same way):

Claim 11.2 *If ρ is sparse then every PCR refutation of $\tau_{\leq}(A,b)|_\rho$ must have degree $> r/8$.*

The analogue of (35) will thus be

$$\mathbf{P}[\text{every PCR refutation of } \tau_{\leq}(A,b)|_{\boldsymbol{\rho}} \text{ has degree } > r/8] \geq 1/2. \qquad (43)$$

Consider the mapping R from Definitions 5.4, 5.6 restricted to literals (R, in fact, is almost always identical, except for literals of those variables y^i_Σ for which either Σ or $J_i(A) \setminus \Sigma$ is very small). Let $R_{\mathbb{F}}$ denote the corresponding polynomial homomorphism over the field $\mathbb{F}$. $R_{\mathbb{F}}$ takes any PCR refutation P of $\tau_{\leq}(A,b)$ into another PCR refutation in which every line has the form $R_{\mathbb{F}}(f) \cdot f'$, where $f \in P$ and $\deg(f') \leq O(d)$. Note that unlike Claim 5.8, we do *not* make any conclusions about the *size* of the refutation $R_{\mathbb{F}}(P)$ (and in fact it *may* grow out of control).

Next, we remark that the proof of Claim 5.14 actually allows a finer analysis for protected $O(d)$-DNF of the form $R(C)$, C a clause. Namely, assuming $k = 1$, we get

Claim 11.3 *For every clause C in the variables $Vars_{\leq}(A)$, and every $h \leq r$, either $|Vars(R(C))| \leq h$ or*

$$\mathbf{P}[R(C)|_{\boldsymbol{\rho}} \neq 1] \leq \exp(-h/2^{O(d)})$$

(*unlike all previous claims of this sort, $R(C)|_{\boldsymbol{\rho}} \neq 1$ here simply means that $R(C)|_{\boldsymbol{\rho}}$ is not* **semantically** *equal to 1*).

Proof of Claim 11.3. Assume that $|Vars(R(C))| > h$. Applying Claim 5.14, we get that with the required probability $1 - \exp(-h/2^{O(d)})$ there exists a decision tree $T_{\boldsymbol{\rho}}$ of height $< h$ strongly representing $R(C)|_{\boldsymbol{\rho}}$. Inspecting the proof of Lemma 4.4 for monotone F, we see that the tree T given by this construction actually has a stronger property. Namely, whenever $\pi \in Br_0(T)$, then not only $(t|_{\boldsymbol{\rho}})|_{\pi} = 0$ for all $t \in F$ (as required by Definition 3.5), but in fact even $t|_{\pi} = 0$. The proof of Lemma 4.5 can be easily modified to preserve this property: we only have to go over terms $t \in F$ (rather than $t \in F|_{\rho}$) in the construction of the sequence $T_0, T_1, \ldots, T_{\ell}, \ldots$ Therefore, we may also assume that the tree $T|_{\boldsymbol{\rho}}$ also has this stronger property.

$R(C)$, however, is the result of a DNF expansion of a disjunction of linear forms. Such DNFs can be set to 0 only by restrictions that assign *all* their variables. Therefore, since $|Vars(R(C))| > h$ and the height of $T|_{\boldsymbol{\rho}}$ is at most h, we have $Br_0(T|_{\boldsymbol{\rho}}) = \emptyset$. $R(C)|_{\boldsymbol{\rho}} = 1$ (in the semantical sense) follows.■

We now finish the proof of Theorem 2.19. Let P be a PCR refutation of $\tau_{\leq}(A, b)$. From (43) we in particular get

$$\mathbf{P}[\exists g \in R_{\mathbb{F}}(P)(\deg(g|_{\boldsymbol{\rho}}) > r/8)] \geq 1/2.$$

As we remarked above, every line in the refutation $R_{\mathbb{F}}(P)$ has the form $R_{\mathbb{F}}(f) \cdot f'$ with $f \in P$ and $\deg(f') \leq O(d) \leq r/16$. Next, $\deg(R_{\mathbb{F}}(f)|_{\boldsymbol{\rho}}) > r/16$ implies that there exists a monomial $\alpha \cdot \Gamma_C$ in f ($\alpha \in \mathbb{F}^*$) such that $|Vars(R(C))| > r/16$ and $R_{\mathbb{F}}(\Gamma_C)|_{\boldsymbol{\rho}} \neq 0$. Applying Claim 11.3 (with $h := r/16$), we see that the probability of this event for every particular $\alpha \cdot \Gamma_C \in P$ is bounded by $\exp(-r/2^{O(d)})$. Theorem 2.19 now follows by the same calculation as at the end of Section 5.■

All proofs in Section 6 hold for any proof system that is closed under variable substitutions. In particular, Theorem 2.20 (as well as the analogue of Theorem 2.12 for PCR not stated explicitly in Section 2) follows from Theorem 2.19 by the same proof.

The only problem with the proof of Theorem 2.13 is that the reduction ρ constructed in Section 8 is not a variable substitution (and, as we already remarked above in the proof of Theorem 2.19, the size of PCR proofs may blow up exponentially from applying such ρ). The only variables that create this problem are y_{av}, where v is an auxiliary gate resulting from expanding the instructions in (41). This time, however, our circuit $D_{n,\vec{x}}$ is also allowed to contain the parity gates; therefore, such auxiliary gates are confined to be one of $\bar{z}_\ell \wedge f[\ell-1, j, J_j(A)]$, $z_\ell \wedge f[\ell-1, j+t_0, J_{j+t_0}(A)]$, $(\bar{z}_\ell \wedge f[\ell-1, j, J_j(A)]) \vee (z_\ell \wedge f[\ell-1, j+t_0, J_{j+t_0}(A)])$ explicitly appearing in (41). It remains to notice that for every *fixed* $a \in \{0,1\}^n$, every one of these three gates computes either 0 or one of the two functions $f[\ell-1, j, J_j(A)]$, $f[\ell-1, j+t_0, J_{j+t_0}(A)]$. Thus, at the expense of allowing $\oplus$-gates, ρ can be turned into a variable substitution, and the rest of the proof carries over to PCR without any further changes.

For Theorem 2.22, it is sufficient to combine the adjustments to the proof of Theorem 2.7 made in Section 10 with those in the proof of Theorem 2.19 above.

We now turn to the proof of our last result, Theorem 2.23, that does not seem to have any obvious analogue for Res(k).

Let ρ be a restriction of the variables $Vars^{Cycl}_{\leq}(A)$. Denote by $E^i(\rho)$ the set of all endpoints of all cyclic intervals $\Delta \neq J_i(A)$ with $y^i_\Delta \in \mathrm{sup}(\rho)$. $E^i(\rho)$ defines a partition of $J_i(A)$ into cyclic intervals $\Delta^i_1(\rho), \ldots, \Delta^i_{\ell_i(\rho)}(\rho)$ such that whenever $y^i_\Delta \in \mathrm{sup}(\rho)$, Δ is a disjoint union of some of these intervals. Say that ρ is *consistent* if whenever $y^i_{\Delta(1)}, \ldots, y^i_{\Delta(w)} \in \mathrm{sup}(\rho)$, and $\Delta(1) \oplus \cdots \oplus \Delta(w) = 0$, we have $\rho(\Delta(1)) \oplus \cdots \oplus \rho(\Delta(w)) = 0$. Like in Section 5, let $J_x(\rho) \stackrel{\mathrm{def}}{=} \{j \in [n] \mid x_j \in \mathrm{sup}(\rho)\}$, and say that ρ is *sparse* if $|\Delta^i_\nu(\rho) \setminus J_x(\rho)| \geq 2d$ for every $i \in [m]$, $\nu \in [\ell_i(\rho)]$.

For a clause C in the variables $Vars^{Cycl}_{\leq}(A)$, let $J_x(C) \stackrel{\mathrm{def}}{=} \{j \in [n] \mid x_j \in Vars(C)\}$ and $\mathrm{dom}(C) \stackrel{\mathrm{def}}{=} \{i \in [m] \mid \exists \Delta \neq J_i(A)(y^i_\Delta \in Vars(C))\}$. Call the quantity $w_A(C) \stackrel{\mathrm{def}}{=} |J_x(C)| + |\mathrm{dom}(C)|$ *A-width of the clause C. A-degree of a monomial* $\alpha \cdot \Gamma_C$ ($\alpha \in \mathbb{F}^*$) is defined as $\deg_A(\alpha \cdot \Gamma_C) \stackrel{\mathrm{def}}{=} w_A(C)$, and *$A$-degree of* *a polynomial* is the maximal A-degree of a monomial occurring in it[6]. *A-width of a resolution proof* [*A-degree of a PCR proof*] is the maximal A-width [A-degree] of a clause [polynomial, respectively] occurring in it.

Claim 11.4 *If ρ is consistent and sparse then every PCR refutation P of $\tau_{\leq}^{Cycl}(A, b)$ must contain a monomial m such that* $\deg_A(m) > r/8$ *and* $m|_\rho \not\equiv 0$.

Proof. This is analogous to the proof of Claim 5.2, so we only remark the differences.

First, in the cyclic case the reduction ρ constructed in the proof of that claim is no longer a variable substitution: $\rho(y_\Delta^i)$ is in general the parity of two variables in $Vars_{\leq|_\rho}^{Cycl}(A|_\rho)$. If, however, we extend in a natural way the notions of A-width and A-degree to clauses/monomials in the variables $Vars_{\leq|_\rho}^{Cycl}(A|_\rho)$ (simply by ignoring the second superscript ν in $y_\Delta^{(i,\nu)}$), then it turns out that ρ still does not increase A-degree. Assuming (for the sake of contradiction) that no monomial m with the required properties exist, we conclude that $\deg_A(P|_\rho) \leq r/8$.

Next, we use the same argument from [BI99] as in the proof of Corollary 11.1 to convert $P|_\rho$ into a *resolution* refutation of $\tau_{\leq|_\rho}^{Cycl}(A|_\rho, b|_\rho)$. In more elaborate form the bound from [BI99] says that for every clause C in the resulting resolution refutation, $Vars(C) \subseteq Vars(m_1) \cup Vars(m_2)$, where m_1, m_2 are some monomials in the original PC refutation. Hence, the A-width of this resulting refutation of $\tau_{\leq|_\rho}^{Cycl}(A|_\rho, b|_\rho)$ is at most $r/4$.

Finally, when Corollary 3.4 is applied to the triple $(A|_\rho, \leq|_\rho, b|_\rho)$, it can be generalized in two ways. First, $\tau_{\leq|_\rho}(A|_\rho, b|_\rho)$ can be replaced by $\tau_{\leq|_\rho}^{Cycl}(A|_\rho, b|_\rho)$ (since, like $\tau_{\leq|_\rho}(A|_\rho, b|_\rho)$, this is also a sub-CNF of $\tau(A|_\rho, \vec{g})$ for the same $\vec{g}$). Second, and this is more crucial, width can be replaced by A-width. This is also done by an easy adjustment of [ABRW00, Theorem 3.1] and Theorem 3.3. The only non-trivial thing to be remarked in this respect is that in the matrix $A|_\rho$ the expansion property holds for every set of rows $\{(i_1, \nu_1), \ldots, (i_\ell, \nu_\ell)\}$ whose *projection* $\{i_1, \ldots, i_\ell\}$ onto the first coordinate has size $\leq r$ (cf. the bound (26) in the proof of Claim 5.2), that is to every set of the form $\mathrm{dom}(C)$ with $w_A(C) \leq r$.

[6]Our notion of A-degree is slightly different from the one used in [ABRW00], mainly since the variables we consider here are automatically stratified w.r.t. rows $i \in [m]$

The generalization of Corollary 3.4 obtained in this way gives the required contradiction with $\deg_A(P|_\rho) \leq r/8$ and completes the proof of Claim 11.4.$\blacksquare$

Next, we, similarly to Definition 5.9, define a consistent random restriction $\boldsymbol{\rho}$ of the variables $Vars^{Cycl}_{\leq}(A)$. As before, $\boldsymbol{\rho}$ assigns x-variables completely at random (with $\mathbf{P}[\boldsymbol{\rho}(x_j) = 0] = \mathbf{P}[\boldsymbol{\rho}(x_j) = 1] = 1/4$). y-variables are assigned as follows. Pick a random subset of endpoints $\tilde{\boldsymbol{E}}^{\boldsymbol{i}}$ by including there every endpoint with probability $(1/2d)$, independently of each other. $\tilde{\boldsymbol{E}}^{\boldsymbol{i}}$ induces a partition of $J_i(A)$ into cyclic intervals $\tilde{\boldsymbol{\Delta}}^{\boldsymbol{i}}_{\boldsymbol{1}}, \ldots, \tilde{\boldsymbol{\Delta}}^{\boldsymbol{i}}_{\tilde{\boldsymbol{\ell}}_{\boldsymbol{i}}}$. We take all those $\tilde{\boldsymbol{\Delta}}^{\boldsymbol{i}}_{\boldsymbol{\nu}}$ for which $|\tilde{\boldsymbol{\Delta}}^{\boldsymbol{i}}_{\boldsymbol{\nu}} \setminus J_x(\boldsymbol{\rho})| < 2d$, and remove from $\tilde{\boldsymbol{E}}^{\boldsymbol{i}}$ both endpoints of these intervals. Let $\boldsymbol{E}^{\boldsymbol{i}}$ be the resulting set of endpoints and $\boldsymbol{\Delta}^{\boldsymbol{i}}_{\boldsymbol{1}}, \ldots, \boldsymbol{\Delta}^{\boldsymbol{i}}_{\boldsymbol{\ell}_{\boldsymbol{i}}}$ be the corresponding partition into cyclic intervals. We pick at random Boolean values $\boldsymbol{b}_{\boldsymbol{i},\boldsymbol{1}}, \ldots, \boldsymbol{b}_{\boldsymbol{i},\boldsymbol{\ell}_{\boldsymbol{i}}}$ subjected to the only linear constraint $\bigoplus_{\nu=1}^{\ell_i} \boldsymbol{b}_{\boldsymbol{i},\boldsymbol{\nu}} = b_i$, and assign in the natural way all those variables $y^i_\Delta \in Vars^{Cycl}_{\leq}(A)$ for which both endpoints of Δ are in $\boldsymbol{E}^{\boldsymbol{i}}$. That is, if $\Delta = \dot{\bigcup}_{\nu\in\Gamma} \boldsymbol{\Delta}^{\boldsymbol{i}}_{\boldsymbol{\nu}}$ then $\boldsymbol{\rho}(y^i_\Delta) \stackrel{\text{def}}{=} \bigoplus_{\nu\in\Gamma} \boldsymbol{b}_{\boldsymbol{i},\boldsymbol{\nu}}$.

Clearly, $\boldsymbol{\rho}$ is consistent with probability 1.

Claim 11.5 $\mathbf{P}[\boldsymbol{\rho}$ *is sparse* $] \geq 1/2$.

Proof. By exactly the same analysis as in the proof of Claim 5.10.$\blacksquare$

Lastly, we need the following simple version of Claim 11.3 (note the polynomial dependence on d – this is where our choice of the encoding will be used).

Claim 11.6 *Let C be any clause in the variables $Vars^{Cycl}_{\leq}(A)$ with $w_A(C) \leq r$. Then*

$$\mathbf{P}[C|_{\boldsymbol{\rho}} \not\equiv 1] \leq \exp(-w_A(C)/d^{O(1)}).$$

Proof. We may assume w.l.o.g. that C consists either only of x-variables, or only of y-variables. In the first case the claim is obvious (since $\boldsymbol{\rho}$ acts on x-variables completely at random). In the second, for every $i \in \mathrm{dom}(C)$ choose arbitrarily a cyclic interval Δ^i with $\frac{|J_i(A)|}{3} - 1 \leq |\Delta^i| \leq \frac{2|J_i(A)|}{3} + 1$ such that $y^i_{\Delta^i} \in Vars(C)$. Let j^i_1, j^i_2 be its endpoints, and for $\alpha = 1, 2$, let $L_{i\alpha}$, $R_{i\alpha}$ be the two cyclic intervals, of length $5d$ each, with the endpoint j^i_α. Note that, due to the constraints on $|\Delta^i|$, we have $Conv(L_{i1}, R_{i1}) \cap Conv(L_{i2}, R_{i2}) = \emptyset$, and we use $L_{i\alpha}, R_{i\alpha}$ as we used protections in Section 5. The property of weak regularity is immediate (C does not contain x-variables, and for every

$i \in [m]$ contains at most one y^i_Δ), therefore we need neither the reduction operator R nor any analogue of Claim 5.12.

In the proof of Claim 5.13, the sub-protections $L'_{i\alpha}, R'_{i\alpha}$ will now have cardinalities $4d$ (as opposed to $2d$), and we can save by relaxing the requirement $J \cap P'_{i\alpha} \neq \emptyset$ in the definition of $\boldsymbol{(Y^i)'}$ to $|J \cap L'_{i\alpha}|, |J \cap R'_{i\alpha}| \leq 2d$. Then in (34) we have the better bound $\mathbf{P}\Big[y^i_\Sigma \in \boldsymbol{(Y^i)'}\Big] \geq \Omega(p_{i\Sigma}/d)$, and with the same argument we get that $\boldsymbol{\rho}$ restricted to $Vars(C)$ is $(r, \mu_{triv}, \Omega(1/d))$-independent. Claim 11.6 follows.∎

Theorem 2.23 is immediately implied by Claims 11.4, 11.5 and 11.6.

12. Open problems

The central open problem in this area is obvious: construct pseudo-random generators that would be hard for as strong proof systems as possible and get as many pseudo-random output bits as possible. We complement this with several other (more minor) questions.

Does there exist a *function* pseudorandom generator of Nisan-Wigderson type which is hard for Resolution? Say, do there exist any $A, \leq, b$ such that $m = 2^{n^\epsilon}$ and the minimal resolution refutation size of $\tau_{\leq}(A, b)$ is exponential in n? The importance of this problem is, of course, greatly undermined by the iterability trick which allowed us to turned around it in Theorem 2.13. Still, this problem might be interesting in its own right. Also, it is not clear at the moment how general this trick will turn out, so after all a better understanding of the hardness of NW-generators themselves still may be useful in further research.

The next problem is of similar flavour. Even $n^{\Omega(\log n)}$ output bits we were able to get only for Nisan generators, that is, when the base functions are parity functions. Get more than quadratic number of output bits for a wider class of base functions. Suppose for example that $m = O(n^2)$, and the functions $\vec{g}$ are picked at random. Is $\tau(\boldsymbol{A_{mn}}, \vec{g})$ hard for Resolution?

If the last problem is solved, then we might ask to extend the resulting bound to the system PCR over fields of characteristic 2 (this case is left completely open by the current paper).

Does PCR possess efficient proofs of $\mathbf{NP} \not\subseteq \mathbf{P/poly}$ when the latter class is defined by circuits over the standard basis $\{\neg, \wedge, \vee\}$? The natural attempt to simply ignore the difficulty occurred in the proof of Theorem 2.21 leads to the system PCRes(2) which is a natural hybrid of PC and Res(2). Now, lower bounds for this system are known (see much more general result in [Kra97b]), but what we really need is a pseudo-random generator hard for it.

Last, but not the least, construct explicit lossless expanders (ideally, expanders with parameters close to those in Theorem 2.5). The importance of this last problem stretches of course well beyond proof complexity (see e.g. the impressive list of potential applications of expanders in [CRVW02]).

Acknowledgement

I am grateful to Misha Alekhnovich, Jan Krajíček and Avi Wigderson for their interest, involvement and useful discussions at various stages of this ongoing project.

References

[ABRW02] M. Alekhnovich, E. Ben-Sasson, A. Razborov, and A. Wigderson. Space complexity in propositional calculus. *SIAM Journal on Computing*, 31(4):1184–1211, 2002.

[ABRW00] M. Alekhnovich, E. Ben-Sasson, A. Razborov, and A. Wigderson. Pseudorandom generators in propositional complexity. In *Proceedings of the 41st IEEE FOCS*, pages 43–53, 2000. Journal version to appear in *SIAM Journal on Computing*.

[AR01] M. Alekhnovich and A. Razborov. Lower bounds for the polynomial calculus: non-binomial case. In *Proceedings of the 42nd IEEE Symposium on Foundations of Computer Science*, pages 190–199, 2001.

[AS00] N. Alon and J. Spencer. *The probabilistic method.* Wiley-Interscience, 2 edition, 2000.

[BBC+98] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf. Quantum lower bounds by polynomials. In *Proceedings of the 39th IEEE Symposium on Foundations of Computer Science*, pages 352–361, 1998. Preliminary version available at `quant-ph/9802049`.

[BGI+01] B. Barak, O. Goldreich, R. Impagliazzo, S. Rudich, A. Sahai, S. Vadhan, and K. Yang. On the (im)possibility of obfuscating programs. In *Proceedings of the 21st Annual International Cryptology Conference*, pages 1–18, 2001.

[BGIP99] S. Buss, D. Grigoriev, R. Impagliazzo, and T. Pitassi. Linear gaps between degrees for the Polynomial Calculus modulo distinct primes. In *Proceedings of the 31st ACM STOC*, pages 547–556, 1999.

[BI99] E. Ben-Sasson and R. Impagliazzo. Random CNF"s are hard for the polynomial calculus. In *Proceedings of the 40th IEEE FOCS*, pages 415–421, 1999.

[BP96] P. Beame and T. Pitassi. Simplified and improved resolution lower bounds. In *Proceedings of the 37th IEEE FOCS*, pages 274–282, 1996.

[BP98] P. Beame and T. Pitassi. Propositional proof complexity: Past, present and future. Technical Report TR98-067, Electronic Colloquium on Computational Complexity, 1998.

[BPR97] M. Bonet, T. Pitassi, and R. Raz. Lower bounds for cutting planes proofs with small coefficients. *Journal of Symbolic Logic*, 62(3):708–728, 1997.

[BPR00] M. Bonet, T. Pitassi, and R. Raz. On interpolation and automatization for Frege systems. *SIAM Journal on Computing*, 29(6):1939–1967, 2000.

[CEI96] M. Clegg, J. Edmonds, and R. Impagliazzo. Using the Groebner basis algorithm to find proofs of unsatisfiability. In *Proceedings of the 28th ACM STOC*, pages 174–183, 1996.

[CRVW02] M. Capalbo, O. Reingold, S. Vadhan, and A. Wigderson. Randomness conductors and constant-degree expansion beyond the degree/2 barrier. In *Proceedings of the 34th ACM Symposium on the Theory of Computing*, pages 659–668, 2002.

[GGM86] O. Goldreich, S. Goldwasser, and S. Micali. How to construct random functions. *Journal of the ACM*, 33(4):792–807, 1986.

[Gol00] O. Goldreich. Candidate one-way functions based on expander graphs. Manuscript, 2000.

[IKW02] R. Impagliazzo, V. Kabanets, and A. Wigderson. In search of an easy witness: exponential time vs. probabilistic polynomial time. *Journal of Computer and System Sciences*, 65(4):672–694, 2002.

[Jan90] S. Janson. Poisson approximation for large deviations. *Random Structures and Algorithms*, 1:221–230, 1990.

[KP98] J. Krajíček and P. Pudlák. Some consequences of cryptographical conjectures for S^1_2 and EF. *Information and Computation*, 142:82–94, 1998.

[Kra94] J. Krajíček. Lower bounds to the size of constant-depth propositional proofs. *Journal of Symbolic Logic*, 59(1):73–86, 1994.

[Kra95] J. Krajíček. *Bounded arithmetic, propositional logic and complexity theory.* Cambridge University Press, 1995.

[Kra97a] J. Krajíček. Interpolation theorems, lower bounds for proof systems and independence results for bounded arithmetic. *Journal of Symbolic Logic*, 62(2):457–486, 1997.

[Kra97b] J. Krajíček. Lower bounds for a proof system with an exponential speed-up over constant-depth Frege systems and over polynomial calculus. In P. Ružička and I. Prívara, editor, *Proceedings of the 22nd International Symposium on the Mathematical Foundations of Computer Science (Bratislava, August '97), Lecture Notes in Computer Science* 1295, pages 85–90. Springer-Verlag, 1997.

[Kra01a] J. Krajíček. On the weak pigeonhole principle. *Fundamenta Mathematicae*, 170(1-3):123–140, 2001.

[Kra01b] J. Krajíček. Tautologies from pseudo-random generators. *Bulletin of Symbolic Logic*, 7(2):197–212, 2001.

[Kra02] J. Krajíček. Dual weak pigeonhole principle, pseudo-surjective functions, and provability of circuit lower bounds. Manuscript, 2002.

[Nis91] N. Nisan. CREW PRAMs and decision trees. *SIAM Journal on Computing*, 20(6):999–1007, 1991.

[NW94] N. Nisan and A. Wigderson. Hardness vs. randomness. *Journal of Computer and System Sciences*, 49(2):149–167, October 1994.

[Pin73] M. Pinsker. On the complexity of a concentrator. In *7th Annual Teletraffic Conference*, pages 318/1–318/4, Stockholm, 1973.

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

[Pud98] P. Pudlák. The lengths of proofs. In S. Buss, editor, *Handbook of Proof Theory*, pages 547–637. Elsevier, 1998.

[RanRaz02] R. Raz. Resolution lower bounds for the weak pigeonhole principle. In *Proceedings of the 34th ACM Symposium on the Theory of Computing*, pages 553–562, 2002.

[Raz95a] A. Razborov. Bounded Arithmetic and lower bounds in Boolean complexity. In P. Clote and J. Remmel, editors, *Feasible Mathematics II. Progress in Computer Science and Applied Logic, vol.* 13, pages 344–386. Birkhaüser, 1995.

[Raz95b] A. Razborov. Unprovability of lower bounds on the circuit size in certain fragments of bounded arithmetic. *Izvestiya of the RAN*, 59(1):201–224, 1995.

[Raz96] A. Razborov. Lower bounds for propositional proofs and independence results in Bounded Arithmetic. In F. Meyer auf der Heide and B. Monien, editors, *Proceedings of the 23rd ICALP, Lecture Notes in Computer Science,* 1099, pages 48–62, New York/Berlin, 1996. Springer-Verlag.

[Raz98] A. Razborov. Lower bounds for the polynomial calculus. *Computational Complexity*, 7:291–324, 1998.

[Raz02a] A. Razborov. Proof complexity of pigeonhole principles. In *Proceedings of the 5th DLT, Lecture Notes in Computer Science,* 2295, pages 100–116, New York/Berlin, 2002. Springer-Verlag.

[Raz02b] A. Razborov. Resolution lower bounds for perfect matching principles. In *Proceedings of the 17th IEEE Conference on Computational Complexity*, pages 29–38, 2002.

[RR97] A. Razborov and S. Rudich. Natural proofs. *Journal of Computer and System Sciences*, 55(1):24–35, 1997.

[SBI02] N. Segerlind, S. Buss, and R. Impagliazzo. A switching lemma for small restrictions and lower bounds for k-DNF resolution. In *Proceedings of the 43rd IEEE Symposium on Foundations of Computer Science*, pages 604–613, 2002.

[Urq95] A. Urquhart. The complexity of propositional proofs. *Bulletin of Symbolic Logic*, 1:425–467, 1995.

[Yao82] A. Yao. Theory and applications of trapdoor functions. In *Proceedings of the 23rd IEEE FOCS*, pages 92–99, 1982.