Pseudorandom generators hard for propositional proof systems

Markus Latte

April 3 and 4, 2009
JASS 2009
Course 1: Propositional Proof Complexity
St. Petersburg

Abstract

Based on the concept of pseudorandom generators, the notation of a generator which is hard for a proof system is introduced. Such a generator admits a superpolynomial lower bound. For the resolution proof system a hard generator is constructed which bases on expanders.

1 Introduction

Informally speaking, in complexity theory, a *pseudorandom generator* [6, Chap. 8] is a computable function

$$G_n\colon \{0,1\}^n \to \{0,1\}^m \qquad\qquad (n < m)$$

which stretches a short (random) string x to a long (almost random) string $G_n(x)$ in a reasonable way. That is, a given computable function $f\colon \{0,1\}^m \to \{0,1\}$ can hardly distinguish them, i. e. the difference between

$$\underset{\mathbf{x}\in\{0,1\}^n}{\mathbf{Pr}}[f(G_n(\mathbf{x}))=1] \qquad\qquad \text{and}$$

$$\underset{\mathbf{y}\in\{0,1\}^m}{\mathbf{Pr}}[f(\mathbf{y})=1]$$

is small. Hence, a random generator for size m can be replaced by a random generator for size n together with G_n without affecting f essentially. On top of this concept, the strength of a whole class $\mathcal{C}$ of functions can be determined. A pseudorandom generator G is *secure* or *hard* for such a class if none of its functions can distinguish random inputs and stretched inputs in the previous

The primary source for the presentation is a publication by Alekhnovich, Ben-Sasson, Razborov and Wigderson [1].

sense. If this is the case then informally the class consists of “simple” functions only. Indeed, a function like

$$g\colon \{0,1\}^m \to \{0,1\}, \quad z \mapsto \text{if } (\exists y \in \{0,1\}^n . G_n(y) = z) \text{ then } 1 \text{ else } 0 \tag{1}$$

is a candidate to separate both distributions but might have a high run-time complexity.

The idea to expose a class as weak with respect to a generator motivates to adapt the concept of pseudorandom generators to the field of proof complexity.

Definition 1. *A* generator *is a family $(G_n)_{n\in\mathbb{N}}$ of functions such that $G_n : \{0,1\}^n \to \{0,1\}^{m(n)}$ for some* stretch function $m\colon \mathbb{N} \to \mathbb{N}$*. Such a generator is* hard *for a propositional proof system P iff for all $n \in \mathbb{N}$ and for any string $b \in \{0,1\}^{m(n)} \setminus \mathsf{Image}(G_n)$ the size of any P-proof of the propositional formula*

$$\varphi_{n,b} := \ulcorner G_n(x_1, \ldots, x_n) \neq b \urcorner \tag{2}$$

is superpolynomial. The notation $\ulcorner \textvisiblespace \urcorner$ denotes some reasonable encoding of its argument where $x_1, \ldots, x_n$ are propositional variables. ⌟

To get a superpolynomial lower bound for the proof system P infinite many instances of the formulas $\varphi_{n,b}$ have to be tautologies. In particular, this is the case if G_n is not surjective, because the propositional variables $x_1, \ldots, x_n$ are implicitly universally quantified. Being complete a proof system can prove the tautologies among $\varphi_{n,b}$ but might require superpolynomial or even exponential many steps—in analogy to (1). In other words, assume that there exists a hard generator G for a proof system P and that G consists of non-surjective functions only. Then P cannot efficiently prove even the most basic things about the generator, namely that it is not surjective mapping [1, Sec. 1].

Moreover, some principles which provide high lower bounds can be rephrased as families of non-surjective functions [1, Examples 1–3].

2 Preliminaries

The set $\{1, \ldots, n\}$ is denoted by $[n]$ for any natural number n. The logarithms, $\log(\textvisiblespace)$, refers to the base 2 always.

As for resolution, variables are written as x, y, ... A literal is a variable x or its negation $\neg x$. For convenience, we write x^1 for x, and x^0 for $\neg x$. The exponent is called *switch*. A clause is a (finite) disjunction of literals, often written as a set or as a list. The empty clause is denoted by $\bot$. The width of a clause C, written as $\mathsf{width}(C)$, is the number of its literals. A CNF is a (finite) conjunction of clauses. The size of a CNF is the sum of the widths of its clauses. The considered resolution calculus [4, Chap. 2][1] comprises the *resolution rule* $\frac{C,x \qquad D,\neg x}{C,D}$. If a resolution proof derived clause is empty, we say that it is a *resolution refutation*. For a resolution π, we write $|\pi|$ for its *size*, that is the number of rule applications. The *width*, $\mathsf{width}(\pi)$, is the biggest width of all its clauses. For the sake of convenience, the *weakening rule*, $\frac{C}{D}$ for $C \subseteq D$, is assumed to be admissible with respect to the width, that is that the elimination of instances of the weakening rule does not increase the width of a refutation.

The boolean values are 0 and 1 called “false” and “true”, respectively. A *restriction*, or *assignment* synonymously, is a partial function ρ from the variable to $\{0,1\}$. Its domain is written as $|\rho|$. A restriction is called *total* if its domain contains all variables. Just as for arbitrary functions, if the domain is $\{x_1,\dots,x_n\}$ we may also write $[x_1 \mapsto \rho(x_1),\dots x_n \mapsto \rho(x_n)]$ instead of ρ. The application of a restriction ρ to a boolean formula φ is written as $\varphi{\upharpoonright}\rho$. In that process the formula is also simplified as far as the laws of the neutral elements are applicable [4, page 5]. Similarly for a boolean function f, its restriction to ρ is denoted by $f{\upharpoonright}\rho$. If a restriction ρ to a formula φ or a function f yields 1, we may write $\rho \models \varphi$ or $\rho \models f$, respectively. In this case we say that the restriction *satisfies* the respective object. However, if the restriction yields 0 the restriction is said to *falsify*. For two boolean functions f and g we write $f \models g$ if every restriction ρ satisfies g whenever it does f.

For a boolean function $f(x_1,\dots,x_n)$ the variable x_i is *essential*, or, synonymously, *f depends on x_i*, if $f{\upharpoonright}[x_i \mapsto 0] \neq f{\upharpoonright}[x_i \mapsto 1]$. The set of all essential variables is $\mathsf{Vars}(f)$.

The usual asymptotic notations $\mathcal{O}(\textvisiblespace)$ and $\Omega(\textvisiblespace)$ are used [7]. The class of all exponential functions to a base greater than one is written as *exp*.

3 Generators for Resolution

To construct a generator hard for resolution, Definition 1 can be rephrased.

> A generator $(G_n\colon \{0,1\}^n \to \{0,1\}^{m(n)})_{n\in\mathbb{N}}$ is hard for resolution iff for all $n \in \mathbb{N}$ and for any string $\vec{b} \in \{0,1\}^{m(n)} \setminus \mathsf{Image}(G_n)$ the size of any resolution refutation of a CNF, stating that
>
> $$G_n(x_1,\dots,x_n) = \vec{b}, \tag{3}$$
>
> is superpolynomial.

For simplicity, we may also write m instead of $m(n)$. Later on, in Corollary 29 et sequentes, we will take a function m which satisfies $m(n) > n$, finally. Hence there will be for (almost) all $n \in \mathbb{N}$ a suitable $\vec{b} \in \{0,1\}^{m(n)}$. As a byproduct we also get a superpolynomial lower bound. It remains, for one thing, to find a candidate for a generator and, for another thing, to prove the requested lower bound. The generator G_n to be constructed can be decomposed into functions $g_1(x_1,\dots,x_n)$, …, $g_m(x_1,\dots,x_n)$ such that (3) is equivalent to the following system of equations.

$$\left.\begin{aligned} g_1(x_1,\dots,x_n) &= 1 \\ &\vdots \\ g_m(x_1,\dots,x_n) &= 1 \end{aligned}\right\} \tag{4}$$

The functions g_is are called *base functions*. Beside trivial settings, the base functions are satisfiable. Hence, to prove this system of equations unsatisfiable any two base functions should share some of their essential variables. Intuitively, the more variables a set of function share the simpler it is prove (4) unsatisfiable. To state this property more precisely, we utilize the notation of an *expander*, c. f. Definition 6. To provide a size lower bound as required by Definition 1, we seize the following observation.

Theorem 2. *Let τ be an unsatisfiable CNF in n variable and clauses the width of which is at most w. Then every refutation of τ of size S has a clause of width $w + \mathcal{O}(\sqrt{n \log S})$.*

Proof. See [5, Theorem 3.2]. □

Hence, it suffices to provide a lower bound for the width of a refutation. The proof of this lower bound is given in Section 6.

4 Nisan-Wigderson Generators

We presume that there is an infinite support of variables $(x_i)_{i\in\mathbb{N}}$.

Definition 3. *Let $A = (a_{i,j})$ be a matrix of dimension $m \times n$ over $\{0, 1\}$. For any row number $i \in [m]$ let*

$$J_i(A) := \{j \in [n] \mid a_{i,j} = 1\} \text{ and}$$
$$\mathsf{X}_i(A) := \{x_j \mid j \in J_i(A)\}.$$

And additionally

$$\mathsf{X}(A) := \{x_i, \dots, x_n\}.$$

Unless stated otherwise, the parameters n, m, and A are used implicitly. ⌟

Definition 4. *Let $g_1(x_1, \dots, x_n)$, …, $g_m(x_1, \dots, x_n)$ be boolean functions such that $\mathsf{Vars}(g_i) \subseteq \mathsf{X}_i(A)$ for all $i \in [m]$. The function*

$$G : \{0, 1\}^n \to \{0, 1\}^m, \quad (x_1, \dots, x_n) \mapsto \Big(g_1(x_1, \dots, x_n), \dots, g_m(x_1, \dots, x_n)\Big)$$

is called as Nisan-Wigderson generator [1]. ⌟

From now on, the symbols g_1, …, g_m refer to such functions unless otherwise stated. Synonymously, we write $\vec{g}(\vec{x})$ or $\vec{g}$ for these functions. Using such kind of base functions, the construction of a hard generator can be decomposed into four aspects: (a) combinatorial properties of matrix A, (b) hardness conditions for the base functions $\vec{g}$, (c) encoding of the equation system $\vec{g}(\vec{x}) = \vec{1}$, and (d) a lower bound. As a first step, we will detail a combinatorial property which meets the idea sketched at the end of Section 3. An encoding is given later on, in Section 5, followed up by the proof of a lower bound.

Definition 5 (Boundary)**.** *Let A be an $(m \times n)$-matrix over $\{0,1\}$. For a set of rows $I \subseteq [m]$, its* boundary *is the set*

$$\partial_A(I) := \{j \in [n] \mid \exists! i \in I . a_{i,j} = 1\} \ .$$

The members of this set are called boundary elements (of the rows I)*.* ⌟

Remark. *The concept of a boundary also admits an unique function $\partial_A(I) \to I$ with $j \mapsto i$ such that $a_{i,j} = 1$.* ⌟

Definition 6 (Expander)**.** *An $(m \times n)$-matrix A over $\{0,1\}$ is an (r,s,c)-*expander *iff*

(a) for all $i \in [m]$: $|J_i(A)| \leq s$, and

(b) for all $I \subseteq [m]$: $|I| \leq r$ implies $|\partial_A(I)| \geq c\,|I|$. ⌟

The last item can be understood as that each row in I has at least c boundary elements on average and implements the idea explained at the end of Section 3.

Remark. *The concept of an expander can be related to the* edge expansion [2]

$$c_E(G) := \underset{1 \leq |U| \leq |V|/2}{min} \frac{|E(U, V \setminus U)|}{|U|}$$

of an undirected graph $G = (V,E)$. The notation $E(A,B)$ is an abbreviation for $\{\{a,b\} \in E \mid a \in A, b \in B\}$. Let $A_G = (a_{v,e})_{v \in V, e \in E}$ be the incidence matrix of G, that is $a_{v,e} = 1$ iff $v \in e$. Note that $\partial_{A_G}(I)$ is just $E(I, V \setminus I)$. Hence for any c, the matrix A_G is a $(|V|/2, d, c)$-expander iff $c_E(G) \geq c$, provided that d is an upper bound on the degree of the vertices in G. Moreover, the set $\partial_{A_G}(I)$ is the edge boundary of the nodes in I, that is the set of edges connecting the sets I and $V \setminus I$. ⌟

Theorem 7 (Existence of Expanders)**.** *For any positive parameters s and n there exists an $(r, s, \frac{3}{4}s)$-expander of size $n^2 \times n$ where*

$$r = \frac{\epsilon n}{s} \ n^{-\frac{1}{s\epsilon}}$$

for some constant ϵ.

Proof. We use a probabilistic argumentation. If the probability that a random object does not meet a specification is strictly less than one then there exists an object with the specification. Here, we construct a random matrix $A = (a_{i,j})$ as follows. The initial matrix has the size $(n^2 \times n)$ and contains zeros only. For each row i, we pick s times a column number j from $[n]$ independently and at random, and set the entry at (i,j) to one. Obviously, the matrix A meets the first expansion properties as mentioned in Definition 6. Next, we estimate the probability that A is *not* an $(r, s, \frac{3}{4}s)$-expander:

$$\begin{aligned}\mathbf{Pr}\left[A \text{ is not an } (r, s, \tfrac{3}{4}s)\text{-expander}\right] &\leq \sum_{\ell=1}^{r} \binom{n^2}{\ell} p_\ell \\ &\leq \sum_{\ell=1}^{r} n^{2\ell} p_\ell \ , \end{aligned} \tag{5}$$

where p_ℓ is the probability that *any* given ℓ rows violate the second expansion property. To estimate p_ℓ, we fix a set I of rows such that $\ell = |I| \leq r$. For each column $j \in \bigcup_{i\in I} J_i(A) \setminus \partial_A(I)$ there are at least two rows j such that $a_{i,j} = 1$. Since $\partial_A(I) \subseteq \bigcup_{i\in I} J_i(A)$,

$$\begin{aligned}\left|\bigcup_{i\in I} J_i(A)\right| &\leq |\partial_A(I)| + \frac{1}{2}\left(\sum_{i\in I} |J_i(A)| - |\partial_A(I)|\right) \\ &= \frac{1}{2}\left(\sum_{i\in I} |J_i(A)| + |\partial_A(I)|\right) \ .\end{aligned}$$

So, the violation of the the second expansion property, i. e. $|\partial_A(I)| < \frac{3}{4}s\ell$, implies $\left|\bigcup_{i\in I} J_i(A)\right| \leq \frac{7}{8}s\ell$ because $\sum_{i\in I} |J_i(A)|$ is bounded by $s\ell$. Hence, we have

$$p_\ell \leq \mathbf{Pr}\left[\left|\bigcup_{i\in I} J_i(A)\right| \leq \tfrac{7}{8}s\ell\right] . \tag{6}$$

To calculate the right hand side, we record the construction process of A as a list starting with the empty list. For each entry (i, j) to be set one, we append i if $j \in \bigcup_{i\in I} J_i(A)$. The resulting list x contains exactly $s\ell$ elements, each in $[n]$. There are $n^{s\ell}$ forms of such lists. To count the lists which have at most $\frac{7}{8}s\ell$ different elements, we consider those lists equivalently as lists of length $s\ell$ over $[n]$ but where *exactly* $\frac{1}{8}s\ell$ positions have pointers. Each pointer addresses a previous position rather than a value in $[n]$. However, such a pointer *means* that the value of the source is that of the target. Transitivity applies. There are at most

$$\binom{s\ell}{s\ell/8} \cdot n^{7/8s\ell} \cdot (s\ell)^{s\ell/8}$$

such lists with pointer. Notice that the preceding formula also comprises those lists with a cyclic pointer structure which are useless for us. Finally, we have

$$\begin{aligned}\mathbf{Pr}\left[\left|\bigcup_{i\in I} J_i(A)\right| \leq \tfrac{7}{8}s\ell\right] &\leq \frac{\binom{s\ell}{s\ell/8} \cdot n^{7/8s\ell} \cdot (s\ell)^{s\ell/8}}{n^{s\ell}} \\ &\leq \binom{s\ell}{s\ell/8}\left(\frac{s\ell}{n}\right)^{s\ell/8} \\ &\leq \left(\frac{2^8 \cdot s\ell}{n}\right)^{s\ell/8} \end{aligned} \tag{7}$$

using that $\binom{s\ell}{s\ell/8} \le \sum_{k=0}^{s\ell} \binom{s\ell}{k} = 2^{s\ell}$. Sticking together the inequality (5), (6), and (7), one gets

$$\begin{aligned}\mathbf{Pr}\left[A \text{ is not an } (r, s, \tfrac{3}{4}s)\text{-expander}\right] &\le \sum_{\ell=1}^{r} n^{2\ell} \left(\frac{2^8 \cdot s\ell}{n}\right)^{s\ell/8} \\ &\le \sum_{\ell=1}^{r} n^{2\ell} \left(\frac{2^8 \cdot sr}{n}\right)^{s\ell/8} \quad .\end{aligned}$$

since $\ell \le r$ by the choice of ℓ. The geometric progression is bounded by $\frac{1}{2}$ if

$$n^2 \left(\frac{2^8 \cdot sr}{n}\right)^{s/8} < \frac{1}{2} \ ,$$

or, equivalently,

$$r < \frac{n}{s} 2^{-8} \left(\frac{1}{2n^2}\right)^{8/s} = \frac{n}{s} \cdot 2^{-8(1+1/s)} \cdot n^{-16/s} \ .$$

This inequality is satisfied for $r := \frac{\epsilon n}{s} n^{-\frac{1}{s\epsilon}}$ and $\epsilon := 2^{-16}$. Therefore the probability that A is not an $(r, s, \frac{3}{4}s)$-expander is strictly less than one. Hence there exists such an expander. □

Remark. *The fraction* 3/4 *in Theorem 7 is rather arbitrary. The same statement but for any other fraction strictly between* 0 *and* 1 *holds. However, the corresponding constant* ϵ *depends on the respective fraction.* $\lrcorner$

5 Encoding

5.1 The Functional Encoding

There are many possible encodings of the equation system (4) [1, Sec. 2.3]. Here, we detail the *functional encoding* only. In any case, the proof system should benefit from the encoding as less as possible to avoid short refutations. Therefore, the relationship among the base functions are hidden. The encoding says only which other functions a base function implies or supports. To this end, for every boolean function f satisfying $\mathsf{Vars}(f) \subseteq \mathsf{X}_i(A)$ for some $i \in [m]$, an *extension variable* y_f is presumed. The set of all such extension variables for a given matrix A is denoted by $\mathsf{Y}(A)$.

Definition 8 (Functional Encoding)**.** *The* functional encoding $\tau(A, \vec{g})$ *is a* CNF *over the variables* $\mathsf{Y}(A)$ *consisting of clauses*

$$y_{f_1}^{\varepsilon_1} \vee \ldots \vee y_{f_w}^{\varepsilon_w}$$

for which a row $i \in [m]$ *exists such that*

- $\mathsf{Vars}(f_1) \cup \ldots \cup \mathsf{Vars}(f_w) \subseteq \mathsf{X}_i(A)$, *and* (fun-var)
- $g_i \models f_1^{\varepsilon_1} \vee \ldots \vee f_w^{\varepsilon_w}$. (fun-sem)

Every clause in $\tau(A, \vec{g})$ *is called an* axiom. $\lrcorner$

Note that, from now on, we are faced with two kinds of variables. First, the variables $x_1, \ldots \in \mathsf{X}(A)$, and, secondly, the extension variable, namely $y_f \in \mathsf{Y}(A)$.

Example 9. *For every row* $i \in [m]$ *the clause* y_{g_i} *is an axiom.* $\lrcorner$

Example 10. *For any boolean function* $f(x, \vec{x})$ *the* Shannon expansion *holds, that is*

$$f(x, \vec{x}) \Leftrightarrow \Big(\neg x \wedge f(0, \vec{x})\Big) \ \vee\ \Big(x \wedge f(1, \vec{x})\Big) . \tag{8}$$

If $x, \vec{x} \in \mathsf{X}_i(A)$ *for some row* $i \in [m]$ *then the following clauses are axioms.*

$$y_{\neg f(x,\vec{x})} \vee y_{\neg x \wedge f(0,\vec{x})} \vee y_{x \wedge f(1,\vec{x})}$$

$$\neg y_{f(x,\vec{x})} \vee y_{\neg x \wedge f(0,\vec{x})} \vee y_{x \wedge f(1,\vec{x})}$$

$$y_{\neg(\neg x \wedge f(0,\vec{x}))} \vee y_{f(x,\vec{x})}$$

$$y_{\neg(x \wedge f(1,\vec{x}))} \vee y_{f(x,\vec{x})}$$

For instance, the first two axioms correspond to the implication from left to right in (8)*. Moreover, let* f *and* g *be two boolean functions the essential variables of which correspond to a certain row. Since* $\neg(f \wedge g) \vee f$, $\neg(f \wedge g) \vee g$, *and* $\neg f \vee \neg g \vee (f \wedge g)$ *are tautologies the following clauses are axioms as well.*

$$\neg y_{f \wedge g} \vee y_f$$

$$\neg y_{f \wedge g} \vee y_g$$

$$\neg y_f \vee \neg y_g \vee y_{f \wedge g}$$

$\lrcorner$

In other words, the last example shows that an *induction on the complexity of a function* is admissible. For instance, we have for any assignment α that $\alpha \models y_{f(x,\vec{x})}$ if and only if, $\alpha \models y_{f(0,\vec{x})}$ and $\alpha \not\models y_x$, or $\alpha \models y_{f(1,\vec{x})}$ and $\alpha \models y_x$.

In addition to this convenient property, the encoding also meets the required property.

Lemma 11. *The equation system* $\vec{g}(\vec{x}) = \vec{1}$ *is satisfiable if and only if* $\tau(A, \vec{g})$ *is satisfiable.*

Proof. $\Longrightarrow$**:** Let α be a total assignment to $\mathsf{X}(A)$ satisfying the equation system. Let β be the total assignment β on $\mathsf{Y}(A)$ which maps y_f to $f{\restriction}\alpha$. Thus, β satisfies every clause in $\tau(A, \vec{g})$ due to (fun-sem). Indeed, let C be a clause in $\tau(A, \vec{g})$. Using the notation in Definition 8, we know that $\alpha \models g_i$. By (fun-sem) there exists a $j \in [w]$ such that $\alpha \models f_j^{\varepsilon_j}$. That is $f_j{\restriction}\alpha = \varepsilon_j$. Hence $y_{f_j}{\restriction}\beta = \varepsilon_j$, i. e. $\beta \models y_{f_j}^{\varepsilon_j}$, and therefore $\beta \models C$.

$\Longleftarrow$: Let α be a total assignment on $\mathsf{Y}(A)$ which satisfies $\tau(A, \vec{g})$. We construct an assignment β on $\mathsf{X}(A)$ by setting

$$x_j \mapsto \alpha\left(y_{\vec{x} \mapsto x_j}\right).$$

Let f be an arbitrary boolean function with $\mathsf{Vars}(f) \subseteq \mathsf{X}_i(A)$ for some row $i \in [m]$. A simple induction on its complexity yields that $\alpha \models y_f$ if and only if $\beta \models f$. In particular, for any $i \in [m]$ we have that $\beta \models g_i$ because y_{g_i} is an axiom satisfied by α. $\square$

5.2 Size of the Functional Encoding

Lemma 12. *The* CNF *$\tau(A, \vec{g})$ contains at most $m \cdot 2^{2^s}$ (extension) variables provided that $|J_i(A)| \leq s$ for all $i \in [m]$ for some s.*

Proof. The matrix A has m rows. For each row $i \in [m]$ and for each function f on $\mathsf{X}_i(A)$ there is a extension variable y_f. Every variable in $\tau(A, \vec{g})$ has this form. $\square$

In the worst case, the CNF does not only contain many variable but it also shows a large clause.

Lemma 13. *If the matrix A contains a row containing s times an 1, then $\tau(A, \vec{g})$ contains a clause of size 2^{2^s}.*

Proof. Let i be a row in A which contains s times an 1. There are 2^{2^s} boolean functions on $\mathsf{X}_i(A)$. Since $\bigvee\{f \mid f \text{ is a function on } \mathsf{X}_i(A)\} \equiv 1$, the corresponding clause also has the size 2^{2^s} and is in $\tau(A, \vec{g})$. $\square$

5.3 Refinements and Transformations

To provide a size lower bound we like to apply Theorem 2 as it reduces this task to a search for a width lower bound. Thereto, the width of the encoding has to be reasonably bounded. In the remaining part of this section, we try to find a refinement of the previous encoding, Definition 8, which has only sufficiently short axioms.

Definition 14. *For any k, the expression $\tau^{\leq k}(A, \vec{g})$ denotes the* CNF *of all those clauses in $\tau(A, \vec{g})$ the width of each is at most k.* $\lrcorner$

Lemma 15. *Let π be a resolution refutation of $\tau(A, \vec{g})$. There is a resolution refutation π' of $\tau^{\leq 2^s}(A, \vec{g})$ such that*

(a) $\mathsf{width}(\pi') \leq \mathsf{width}(\pi)$*, and*

(b) $|\pi'| \leq 2 \cdot |\pi|$*,*

provided that $|J_i(A)| \leq s$ for all $i \in [m]$ for some s.

Proof. Informally, we have to get rid of large axioms. Let C be an arbitrary axiom in $\tau(A, \vec{g})$. By Definition 8, this axiom has the shape $\bigvee_k y_{f_k}^{\varepsilon_k}$ for some functions f_k and switches $\varepsilon_k \in \{0, 1\}$ such that

- $\bigcup_k \mathsf{Vars}(f_k) \subseteq \mathsf{X}_i(A)$, and
- $g_i \models \bigvee_k f_k^{\varepsilon_k} =: f$,

both for some row $i \in [m]$. The function g_i depends on at most s variables, namely $\mathsf{X}_i(A)$. Hence, on at most 2^s inputs the function g_i becomes 1. By (fun-sem) each such input must be trapped by one of the functions $f_k^{\varepsilon_k}$. Therefore, we can choose for each input such a function. Let K be the set of indices to all those chosen functions. Obviously, $C' := \bigvee_{k \in K} y_{f_k}^{\varepsilon_k}$ is both a subset of C and an axiom in $\tau(A, \vec{g})$ because $f = \bigvee_{k \in K} y_{f_k}^{\varepsilon_k}$. Moreover, $\mathsf{width}(C') \leq |K| \leq 2^s$. Hence, in π we can replace any large axiom C with a small axiom C' and with an instance of the weakening rule. □

Lemma 16. *Let π be a resolution refutation of $\tau^{\leq 2^s}(A, \vec{g})$. There exists a resolution refutation π' of $\tau^{\leq 3}(A, \vec{g})$ such that*

(a) $\mathsf{width}(\pi') \leq \mathsf{width}(\pi)$, *and*

(b) $|\pi'| \leq 2 \cdot 2^s \cdot |\pi|$.

Proof. Informally, large axioms have be to replaced with derivations from small axioms. Let C be an axiom with $\mathsf{width}(C) > 3$. By Definition 8, it emerges from a row $i \in [m]$, functions $f_1, \ldots, f_w$, and from switches $\varepsilon_1, \ldots, \varepsilon_w$ for some w. The construction of an axiom relies on the composed function $\bigvee_{k \in [w]} f_k^{\varepsilon_k}$. However, it can be also understood as be built from just two functions. Those are $f_1^{\varepsilon_1}$ and $f_2^{\varepsilon_2} \vee \ldots \vee f_w^{\varepsilon_w}$, for instance. Hence,

$$C_1 := y_{f_1}^{\varepsilon_1} \vee y_{f_2^{\varepsilon_2} \vee \ldots \vee f_w^{\varepsilon_w}}$$

is also an axiom. Using that

$$\neg \left(f_i^{\varepsilon_i} \vee \ldots \vee f_w^{\varepsilon_w}\right) \;\vee\; f_i^{\varepsilon_i} \;\vee\; \left(f_{i+1}^{\varepsilon_{i+1}} \vee \ldots \vee f_w^{\varepsilon_w}\right)$$

is always true for any $i = 2, \ldots, w-1$, even the corresponding clauses

$$C_i := \neg y_{f_i^{\varepsilon_i} \vee \ldots \vee f_w^{\varepsilon_w}} \;\vee\; y_{f_i}^{\varepsilon_i} \qquad \vee\; y_{f_{i+1}^{\varepsilon_{i+1}} \vee \ldots \vee f_w^{\varepsilon_w}} \qquad (i = 2, \ldots, w-2)$$

and

$$C_{w-1} := \neg y_{f_{w-1}^{\varepsilon_{w-1}} \vee f_w^{\varepsilon_w}} \;\vee\; y_{f_{w-1}}^{\varepsilon_{w-1}} \;\vee\; y_{f_w}^{\varepsilon_w}$$

are axioms. Therefore, the clause C can be resolved from $C_1, \ldots, C_{w-1}$ successively in at most $2w$ steps. As the parameter w is bounded by 2^s due to the assumption, the refutation can be transformed as required. □

Although the transformation in the previous proof enlarges a refutation, the growth can be kept within a limit if s is bounded by the logarithm of the parameters n and m, for instance.

Example 17. *Let us carry out the preceding replacement of a large axiom for the functions f_1, ..., f_4. For simplicity we omit the switches.*

$$\begin{array}{lllll}
y_{f_1} \vee & y_{f_2 \vee f_3 \vee f_4} & & & \\
 & \neg y_{f_2 \vee f_3 \vee f_4} \vee & y_{f_2} \vee & y_{f_3 \vee f_4} & \\
 & & & \neg y_{f_3 \vee f_4} \vee & y_{f_3} \vee y_{f_4} \\
\hline
y_{f_1} \vee & & y_{f_2} \vee & & y_{f_3} \vee y_{f_4}
\end{array}$$

The clause below the line can be derived in five steps—also counting the axioms. ⌟

Remark. *In the literature, large clause are often replaced with small ones by introducing some auxiliary variable for each clause. However, in our setting, these auxiliary variables are already there and hence for free. In other words, the earlier transformation moves the syntactical complexity to a semantic one.* ⌟

6 Width Lower Bound

6.1 The Concept of a Measure

To prove a lower bound on the width of a refutation, we introduce a measure for clauses to determinate their degree of inconsistency [3, Proof of Lemma 1][5, Sec. 5]. This measure should be small for axioms but high for the empty clause. In addition, if a clause has a medium measure it should be large.

Definition 18. *Let Γ be a set of clauses called axioms. A* measure *is a function mapping a clause to an integer such that*

- *μ is* subadditive, *i. e.*

  $$\mu(C) \leq \mu(C_0) + \mu(C_1) \qquad (\mu\text{-subadd})$$

  holds for any resolution step $\frac{C_0 \quad C_1}{C}$,

- *$\mu(C) = 1$ for any axiom C, and* (μ-ax)

- *$\mu(\bot) > r$ holds for some $r > 1$.* (μ-$\bot$)

⌟

Lemma 19. *Let μ be a measure as above for a set Γ of clauses. Any refutation of Γ contains a* medium-measured clause, *i. e. a clause C with $r/2 < \mu(C) \leq r$.*

Proof. Starting at the root, we walk through the resolution refutation towards the leaves as long as the measure of the considered clause is greater than r. Obviously, this walk cannot stop at a leaf since $r > 1$. Assume that we have stopped at a clause C, that is $\mu(C) > r$ and $\mu(C_0), \mu(C_1) \leq r$ where C_0 and C_1 are the premises to C. We claim that C_0 or C_1 is medium measured. So, if not, then $\mu(C_0), \mu(C_1) \leq r/2$. But by ($\mu$-subadd), one gets $\mu(C) \leq r$ contradicting the assumption. □

6.2 The Measure

In general, the function μ shall measure how contradictory a clause is. In our setting, a contradiction is given by the set $G{:=}\{g_i \mid i \in [m]\}$ because it is unsatisfiable as long as $\tau(A, \vec{g})$ is, c. f. Lemma 11. Informally, the proportion of G needed to support a clause determinate this measure.

Definition 20. *For a clause C over $\mathsf{Y}(A)$, its semantic $[\![C]\!]$ is the function*

$$[\![C]\!] {:=} \bigvee_{y_f \in C} f \vee \bigvee_{\neg y_f \in C} \neg f \ .$$

⌟

Definition 21. *The measure $\mu(C)$ for a clause C is the* size *of a minimal $I \subseteq [m]$ such that*

- $\forall y_f^{\varepsilon} \in C \; \exists i \in I.\ \mathsf{Vars}(f) \subseteq \mathsf{X}_i(A)$, *and* ($\mu$-cover)
- $\{g_i \mid i \in I\} \models [\![C]\!]$. ($\mu$-sem)

By "a witness for $\mu(C)$" we refer to such a set I of minimal cardinality. ⌟

In other words, (μ-cover) just says that the essential variables of the denoted function $[\![C]\!]$ are covered by a row in I.

Lemma 22. *The measure μ meets the properties* (μ-subadd) *and* (μ-ax)*.*

Proof. As for (μ-ax), every axiom C comes from a row $i \in [n]$ by Definition 8. The set $\{i\}$ as I satisfies (μ-cover) and (μ-sem). Because of the existential quantifier in (μ-cover) the measure cannot be zero as C is not empty.

For (μ-subadd), consider the resolution step $\frac{C_0 \quad C_1}{C}$. Let I_0 and I_1, respectively, be the witnesses for $\mu(C_0)$ and $\mu(C_1)$. Then $I := I_0 \cup I_1$ is a superset of a witness for $\mu(C)$. Indeed, $C \subseteq C_0 \cup C_1$. Therefore (μ-cover) is satisfied for C by I as it is for C_0 by I_0 and for C_1 by I_1, respectively. The soundness of resolution and the definition of $[\![\textvisiblespace]\!]$ yield that $[\![C_0]\!] \wedge [\![C_1]\!] \models [\![C]\!]$. Since $\{g_i \mid i \in I_k\} \models [\![C_k]\!]$ holds for $k \in \{0,1\}$, so also $\{g_i \mid i \in I_0 \cup I_1\} \models [\![C]\!]$ does. □

6.3 Width Lower Bound

Definition 23. *A boolean function f is* ℓ-robust *if for every restriction ρ holds: if $f{\restriction}\rho$ is a constant function then $|\rho| \geq \ell$.* ⌟

Note that almost all boolean functions are robust in a certain sense, c. f. Appendix A.

Theorem 24. *Let A be an (r, s, c)-expander of size $m \times n$ and let $g_1, \ldots, g_m$ be ℓ-robust functions such that $\mathsf{Vars}(g_i) \subseteq \mathsf{X}_i(A)$. Then every resolution refutation of $\tau(A, \vec{g})$ must have a width which is at least*

$$\frac{r(c+\ell-s)}{2\ell}$$

provided that $c + \ell \geq s + 1$.

Proof. By Lemma 19, Lemma 22, and Lemma 25, following. □

Lemma 25. *The following holds.*

(a) If $r/2 < \mu(C) \leq r$ then $\mathsf{width}(C) \geq \frac{r(c+\ell-s)}{2\ell}$.

(b) $\mu(\bot) > r$ provided that $c + \ell \geq s + 1$.

Proof. Since the proofs for each claim are rather similar, we start with their common part. As for the second claim, its contrapositive is shown, and we identify C as the empty clause, just for convenience. Let I be a witness for $\mu(C)$. The set I can be partitioned in I_0 and I_1 such that I_0 is minimal for (μ-cover). As an intermediate step, we claim that

$$|J_{i_1} \cap \partial_A(I)| \leq s - \ell \text{ for all } i_1 \in I_1, \tag{9}$$

that is, the intersection is quite small.

Proof of claim (9)*.* Let $i_1 \in I_1$. The minimality of I, and the choice of I_0 and I_1 ensure that the removal of any element of I_1 ruins the property (μ-sem) for I. So, $\{g_i \mid i \in I \setminus \{i_1\}\} \not\models [\![C]\!]$. Now let α be an arbitrary assignment to the variables $\mathsf{X}(A)$ which models g_i for all $i \in I \setminus \{i_1\}$ but falsifies $[\![C]\!]$.

Let ρ be the restriction α but additionally undefined on all variables x_j for $j \in J_{i_1} \cap \partial_A(I)$. In other words, ρ is undefined for those variables on which only the function g_{i_1} among $\{g_i \mid i \in I\}$ depends. Therefore, ρ is still total for g_i for all $i \in I \setminus \{i_1\}$. All variables in $[\![C]\!]$ are mentioned in $\bigcup_{i \in I_0} \mathsf{X}_i(A)$ by (μ-cover) and by the choice for I_0. Since $i_1 \notin I_0$, as I_0 and I_1 are disjoint, the restriction ρ is also defined on all variables in $[\![C]\!]$. Therefore, $g_i{\restriction}\rho = 1$ for all $i \in I \setminus \{i_1\}$, and $[\![C]\!]{\restriction}\rho = 0$. By ($\mu$-sem), the function $g_{i_1}{\restriction}\rho$ must be 0.

Let σ be the restriction ρ but only defined on the variables g_{i_1} depending on. All these variables are listed in $\mathsf{X}_{i_1}(A)$. Obviously, $g_{i_1}{\restriction}\sigma = 0$. As ρ is undefined on $J_{i_1}(A) \cap \partial_A(I)$, the domain of σ is just $J_{i_1}(A) \setminus \partial_A(I)$. Since the function $g_{i_1}(A)$ is ℓ-robust, the domain of σ is at least ℓ. In other words, $|J_{i_1}(A) \setminus \partial_A(I)| \geq \ell$. Since the matrix A is an $(\llcorner\!\lrcorner, s, \llcorner\!\lrcorner)$-expander, $J_{i_1}(A) \leq s$. Hence $|J_{i_1}(A) \cap \partial_A(I)| = |J_{i_1}(A)| - |(J_{i_1}(A) \setminus \partial_A(I))| \leq s - \ell$. □

We continue with the proof of Lemma 25. For both parts, we have $|I| \leq r$. Notice that we also prove the contrapositive of the second part. Since A is an (r, s, c)-expander and $|I| \leq r$, we have

$$c\,|I| \leq |\partial_A(I)|\ .$$

Each boundary element dates from an element either of I_0 or of I_1. In the first case, each element of I_0 can contribute at most s boundary elements. As for I_1, every element of I_1 produce at most $s - \ell$ boundary elements due to (9).

$$\begin{aligned} c\,|I| \leq \ldots &\leq s\,|I_0| \quad + \quad (s-\ell)\,|I_1| \\ &= (s-\ell)\,|I| \quad + \quad \ell\,|I_0| \end{aligned} \tag{10}$$

Since I_0 is chosen as minimal, each of its elements is justified by a literal in C.

$$c\,|I| \leq \ldots \leq (s-\ell)\,|I| \quad + \quad \ell \cdot \mathsf{width}(C) \tag{11}$$

Now we can address the two parts to be proven.

(a) The assumption of the first part yields $|I| > r/2$. Hence the inequality (11) results in
$$\mathsf{width}(C) \geq \frac{(c+\ell-s)\,|I|}{\ell} > \frac{(c+\ell-s)r}{2\ell}\ .$$

(b) For the second claim, we have $I_0 = \emptyset$ as C is the empty clause. Therefore the inequality (10) is just $c\,|I| \leq (s-\ell)\,|I|$. By ($\mu$-sem) the set I cannot be empty. Hence the last inequality is the negation of $c + \ell \geq s + 1$. □

7 Size Lower Bound

7.1 From a Width Lower Bound to a Size Lower Bound

Corollary 26. *Let $\epsilon > 0$ be an arbitrary constant, let A be a $(r, s, \epsilon s)$-expander of dimension $m \times n$, and let $g_1, \ldots, g_m$ be $(1 - \epsilon/2) \cdot s$-robust functions such that $\mathsf{Vars}(g_i) \subseteq X_i(A)$ for all $i \in [m]$. Then the size of any resolution refutation of $\tau(A, \vec{g})$ is at least*
$$exp\left(\Omega\left(\frac{r^2}{m\ 2^{2^s}}\right)\right)/2^s.$$

Proof. Let π be a resolution refutation of $\tau(A, \vec{g})$. Apply the transformation described in Lemma 15 and then the one in Lemma 16. Let π' be the obtained resolution refutation of $\tau^{\leq 3}(A, \vec{g})$ such that

(a) $\mathsf{width}(\pi') \leq \mathsf{width}(\pi)$ and

(b) $|\pi'| \leq 2^{s+2} \cdot |\pi|$.

Moreover, $\tau(A, \vec{g})$ contains at most $m \cdot 2^{2^s}$ variables, c. f. Lemma 12. The same amount of variables occur in $\tau^{\leq 3}(A, \vec{g})$. In all, we get the following inequality.

$$\begin{aligned}
&3 + \mathcal{O}\left(\sqrt{(m \cdot 2^{2^s}) \cdot \log(2^{s+2} \cdot |\pi|)}\right) \\
&\geq \mathsf{width}(\pi') && \text{(by Theorem 2)} \\
&\geq \frac{r \cdot \big(\epsilon s \,+\, (1-\epsilon/2)s - s\big)}{2(1-\epsilon/2)s} && \text{(by Theorem 24)} \\
&= r \frac{\epsilon}{2(2-\epsilon)}
\end{aligned}$$

Taking the fraction on the right hand side as a constant, we can solve the inequality and get the claimed lower bound for $|\pi|$. □

7.2 Some Corollaries

The lower bound presented in Corollary 26 is conditional. As a next step, we eliminate these conditions.

Definition 27. *Let A be a matrix over $\{0, 1\}$ of dimension $m \times n$. A sequence of functions $g_1, \ldots, g_m$ is* good *for A if and only if for each $i \in [m]$ the following holds.*

(a) g_i is $\frac{5}{16}\log(\log n)$-robust and

(b) $\mathsf{Vars}(g_i) \subseteq \mathsf{X}_i(A)$. ⌟

Corollary 28 (First Version)**.** *There exists a family $\left(A^{(m,n)}\right)_{n,m \in \mathbb{N}}$ of $m \times n$-sized matrices such that for any sequence of functions $\vec{g} \equiv g_1, \ldots, g_m$ which is good for $A^{(m,n)}$, and for any resolution refutation π of $\tau(A^{(m,n)}, \vec{g})$ the size of π is (at least)*

$$exp\left(\frac{n^{2-\mathcal{O}(1/\log(\log n))}}{m}\right) / \sqrt{\log n}\,.$$

Moreover, the number of (extension) variables in $\tau(A^{(m,n)}, \vec{g})$ is at most $m \cdot n$.

Proof. Let m and n be given. With loss of generality, $m \leq n^2$, as otherwise the exponent is decreasing and gets zero in the limit. The expander construction in the proof of Theorem 7 yields an $(r, s, \frac{3}{4}s)$-expander A for some s—to be fixed later on—and $r = \frac{\epsilon n}{s}\ \ n^{-\frac{1}{s\epsilon}}$ and for a constant ϵ. The matrix A has the dimension $n^2 \times n$. We cross out all rows but m rows arbitrarily. The resulting matrix, say $A^{(m,n)}$, is still an $(r, s, \frac{3}{4}s)$-expander.

In our setting, Corollary 26 states that any resolution refutation of $\tau(A, \vec{g})$ requires

$$exp\left(\Omega\left(\frac{r^2}{m \cdot 2^{2^s}}\right)\right)\ /\ 2^s \tag{12}$$

steps if it is applicable at all. However, we first try to get the claimed lower bound. The exponent in (12) can be simplified to

$$\begin{aligned}\frac{r^2}{m \cdot 2^{2^s}} &= \frac{\epsilon^2 n^2 n^{-\frac{2}{s\epsilon}}}{m \cdot 2^{2^s}} && \text{(choice for } r\text{)} \\ &= \frac{\epsilon^2 n^{2-\frac{2}{s\epsilon}-\frac{2^s}{\log n}}}{m}. && \text{(using that } 2^{2^s} = n^{2^s/\log n}\text{)} \qquad (13)\end{aligned}$$

To get a reasonable lower bound, we try to find an s (as a function in n and m) such that the exponent to n in (13) is as close to 2 as possible. Therefore, we have to ensure that in the limit $\frac{2^s}{\log n}$ is bounded by a constant $c \ll 1$. Hence, $s \leq \log(c) + \log(\log n)$ should hold in the limit. To this end, we choose $s := \frac{1}{2}\log(\log n)$. Thus, the exponent to n can be bounded by

$$\begin{aligned}2 - \frac{2}{s\epsilon} - \frac{2^s}{\log n} &= 2 - \frac{4}{\epsilon \log(\log n)} - \frac{1}{\sqrt{\log n}} \\ &\geq 2 - \frac{4}{\epsilon \log(\log n)} - \frac{1}{\log(\log n)} && \text{for } n \geq 2^{16} \\ &= 2 - \mathcal{O}(1/\log(\log n)).\end{aligned}$$

To apply Corollary 26, the functions $\vec{g}$ need to be $(1 - \frac{3/4}{2}) \cdot s$-robust, that is $\frac{5}{16}\log(\log n)$-robust. Indeed, this is our assumption. Finally, we obtain the claimed lower bound using that $2^s = \sqrt{\log(n)}$ for the denominator in (12). As $s \leq \log(\log(n))$, there are at most $m \cdot n$ variables in the considered CNF due to Lemma 12. □

Having a reasonable lower bound for $\tau(A^{(m,n)}, \vec{g})$, we have to ensure that the formula is unsatisfiable at all. For this purpose, let $\sqcup \oplus \sqcup$ denote the *exclusive or*. In the context of vector notation, $\oplus$ is meant componentwise. As the operation is associative and commutative, we set $\oplus\{x_1, \ldots, x_k\}{:=}x_1 \oplus \ldots \oplus x_k$.

Corollary 29 (Second Version)**.** *There exists a family of $m \times n$ matrices, $A^{(m,n)}$, such that for any sequence of functions $\vec{g}$ good for $A^{(m,n)}$ it is true that*

(a) $\tau(A^{(m,n)}, \vec{g} \oplus \vec{b})$ is unsatisfiable for some $\vec{b} \in \{0,1\}^m$ if $m > n$, and

(b) for any $\vec{b} \in \{0,1\}^m$, any resolution refutation of $\tau(A^{(m,n)}, \vec{g} \oplus \vec{b})$ has a size at least

$$exp\left(\frac{n^{2-\mathcal{O}(1/\log(\log n))}}{m}\right) / \sqrt{\log n}.$$

Proof. For any $\vec{b} \in \{0,1\}^m$ the following is true.

$$\begin{aligned}
&\tau(A^{(m,n)}, \vec{g} \oplus \vec{b}) \text{ unsatisfiable w. r. t. } \mathsf{Y}(A) \\
\Longleftrightarrow &\vec{g}(\vec{x}) \oplus \vec{b} = 1 \text{ is unsatisfiable w. r. t. } \vec{x} \equiv x_1, \ldots, x_n \qquad \text{(by Lemma 11)} \\
\Longleftrightarrow &\vec{g}(\vec{x}) = \neg\vec{b} \text{ is unsatisfiable w. r. t. } \vec{x} \\
\Longleftrightarrow &\vec{g}(\vec{x}) \neq \neg\vec{b} \text{ for all } \vec{x} \in \{0,1\}^n \\
\Longleftrightarrow &\neg\vec{b} \text{ is not in the image of } \vec{g}
\end{aligned}$$

Indeed, $\vec{g} : \{0,1\}^n \to \{0,1\}^m$ is not surjective, since $m > n$. As the robustness is invariant under negation, Corollary 28 yields the claim. □

Consequently, it remains to find candidates for the base functions. Although, by Lemma 32, almost any function is suitable, we consider particular ones.

Definition 30. *Let A be a matrix over $\{0,1\}$ of dimension $m \times n$. The* characteristic function *of the row $i \in [m]$ is*

$$\chi_i^{\oplus}(A) \colon \vec{x} \mapsto \oplus \mathsf{X}_i(A).$$

Notice that the functions $\chi_i^{\oplus}(A)$ are $|\mathsf{X}_i(A)|$-robust. Additionally, for $b \in \{0,1\}^m$ we set

$$\tau_\chi(A, \vec{b}) := \tau(A, \overrightarrow{\chi^{\oplus}(A)} \oplus \vec{b}\,). \qquad ⌟$$

Corollary 31 (Third Version)**.** *There exists a family of $m \times n$ matrices, $A^{(m,n)}$, such that:*

(a) $\tau_\chi(A^{(m,n)}, \vec{b}\,)$ is unsatisfiable for some $\vec{b} \in \{0,1\}^m$ if $m > n$, and

(b) for any $\vec{b} \in \{0,1\}^m$, any resolution refutation of $\tau_\chi(A^{(m,n)}, \vec{b}\,)$ has a size at least

$$exp\left(\frac{n^{2-\mathcal{O}(1/\log(\log n))}}{m}\right) / \sqrt{\log n}.$$

Proof (as a patch to the proofs of previous corollaries). It remains to show that the functions $\chi_i^{\oplus}(A)$ are good for A. During the *construction* of the expander, the 1s in each rows are chosen randomly. The cancellation of its rows is also at random. Hence, any $\chi_i^{\oplus}(A)$ is a function on at most $\frac{1}{2}\log(\log n)$ variables. Similarly to the argument in the proof to Theorem 7 leading from (6) to (7), these are $\frac{5}{8} \cdot \frac{1}{2}\log(\log n)$ robust, therefore also good for A, with a high probability. Thus, there is a matrix such that each of its characteristic functions is robust enough. □

Let $C > 0$ and $b > 1$ be the witnesses in the lower bound of Corollary 31, i. e. the size of any refutation of the considered CNF is at least

$$b^{\left(\frac{n^{2-C/\log(\log n)}}{m}\right)} / \sqrt{\log n}.$$

If we take $m{:=}n^{1+C/\log(\log n)}$ then the lower bound gets a superpolynomial lower bound, namely

$$b^{n^{1-2C/\log(\log n)}}/\sqrt{\log n}, \tag{14}$$

as $m > n$. However, this lower bound is measured in the parameter n but not in the size of the respective CNF. Nevertheless, if $\tau_\chi(A,\vec{b}\,)$ is unsatisfiable then so $\tau_\chi^{\leq 3}(A,\vec{b}\,){:=}\tau^{\leq 3}(A,\overrightarrow{\chi^\oplus(A)}\oplus\vec{b}\,)$ is by Lemma 15 and Lemma 16. On the other hand, any resolution refutation of $\tau_\chi^{\leq 3}(A,\vec{b}\,)$ is also a refutation of $\tau_\chi(A,\vec{b}\,)$ as the former is a sub-CNF of the latter. Since the amount of variables in $\tau_\chi(A,\vec{b}\,)$ is bounded by $n\cdot m$ because of Corollary 28, the size of $\tau_\chi^{\leq 3}(A,\vec{b}\,)$ is $\mathcal{O}\left((n\cdot m)^3\right)$. With respect to the foregoing setting, this size is also polynomially bounded in n. All in all, the lower bound in (14) turns into a proper lower bound which measures the size of a refutation in terms of the size of the refuted formula. Hence, we got a hard generator for resolution.

Remark (On Weakening)**.** *The weakening rule is used in the proof of Lemma 15 only. To incorporate the weakening rule explicitly in the proofs for the lower bounds, at least the measure should be aware of this rule. Definition 18 can be adapted easily to weakening, just by requiring that*

$$\mu(C)/2 < \mu(C') \leq \mu(C) \tag{μ-weakening}$$

holds for every instance $\frac{C'}{C}$ *of the weakening rule. This setting keeps Lemma 19 valid. However, our choice for* μ *in Definition 21 does not meet this requirement. Indeed, consider the instance*

$$\frac{C'}{C}$$

where C' *is an axiom of* $\tau(A,\vec{g})$ *and*

$$C := C' \cup \{y_{\pi_i^n} \mid i \in [n]\}.$$

The functions π_i^n *are the projections* $\pi_i^n(x_1,\ldots,x_n) = x_i$*. Following Lemma 22,* $\mu(C') = 1$*. But* $\mu(C) \geq \lfloor \frac{n}{s} \rfloor$ *because the witnesses for* (μ-cover) *must cover all columns and each row is responsible for at most* s *columns, in our setting. The parameter can be chosen in such way that this instance violates* (μ-weakening)*. However, the might be another solution to this weakness.* ⌟

References

[1] Michael Alekhnovich, Eli Ben-Sasson, Alexander Razborov, and Avi Wigderson. Pseudorandom Generators In Propositional Proof Complexity. *Foundations of Computer Science*, pages 43–53, 2000.

[2] Noga Alon. Spectral Techniques in Graph Algorithms. In *Lecture Notes In Computer Science*, volume 1380, pages 206–215. Springer-Verlag, 1998.

[3] Paul Beame and Toniann Pitassi. Simplified and Improved Resolution Lower Bounds. In *Proceedings of the 37th IEEE Foundations of Computer Science*, pages 274–282. IEEE, 1996.

[4] Arnold Beckmann and Jan Johannsen. *Bounded Arithmetic and Resolution-Based Proof Systems.* Kurt Gödel Society, Vienna, 2004.

[5] Eli Ben-Sasson and Avi Wigderson. Short proofs are narrow – resolution made simple. *Journal of the Association for Computing Machinery*, 48(2):149 – 169, March 2001.

[6] Oded Goldreich. *Computational Complexity: A Conceptual Perspective.* Cambridge University Press, 2008.

[7] Donald E. Knuth. Big Omicron and Big Omega and Big Theta. *SIGACT News*, 8(2):18–24, 1976.

[8] Alexander Razborov. Pseudorandom Generators hard for k-DNF Resolution and Polynomial Calculus. Availalble at the author's web page, 2003.

[9] Grigori S. Tseitin. On the complexity of derivations in propositional calculus. In A. O. Slissenko, editor, *Studies in constructive mathematics and mathematical logic, Part II, Seminars in Mathematics (translated from Russian)*, pages 115–125. Consultants Bureau, New-York-London, 1968.

A Almost all Functions are Robust

Lemma 32. *Let $0 < \epsilon < 1$. For any sufficiently large k, any random function over k variables is ϵk-robust which a probability $\geq \frac{1}{2}$. Here "random" means that the value of a function is chosen randomly and independently for each input.*

Proof. A function f is not ϵk-robust if and only if there exists a restriction ρ such that $|\rho| < \epsilon k$ and $f\lceil\rho$ is a constant function. In particular, then there also is a restriction ρ such that $|\rho| = \epsilon k$ and $f\lceil\rho$ is a constant function. Informally, the truth table of such a function f contains a "block" of $|\rho|$ columns and $2^{k-|\rho|}$ rows such that the result values are constant. This block might be distributed. In total, there are 2^{2^k} boolean functions over k variables. On the other hand, there are $\binom{n}{\ell}2^\ell$ restrictions each fixes ℓ variables. Presuming that a given restriction of size ϵk makes a function constant, there are $2^{2^k-2^{k-\epsilon k}+1}$ many ways to choose such a function.

$$\begin{aligned}\mathbf{Pr}\left[f \text{ is not } \epsilon k\text{-robust}\right] &\leq \frac{\binom{k}{\epsilon k}2^{\epsilon k}\ \ 2^{2^k-2^{k-\epsilon k}+1}}{2^{2^k}} \\ &= \underbrace{\binom{k}{\epsilon k}}_{\leq 2^k} 2^{\epsilon k-2^{(1-\epsilon)k}+1} \\ &\leq 2^{(1+\epsilon)k-2^{(1-\epsilon)k}+1}\end{aligned}$$

The right hand side can be strictly bounded by 2^{-1}, if we require that $(1+\epsilon)k+2 < 2^{(1-\epsilon)k}$ holds additionally. As $\epsilon < 1$ and the exponential function dominates the linear function eventually, the considered probability is less than a half as long as k exceeds a certain value. □