A Criterion for Monotone Circuit Complexity

Stasys Jukna
Institute of Mathematics
Akademijos 4, 2600 Vilnius, Lithuania

September 10, 1991

Abstract

In this paper we study the lower bounds problem for monotone circuits. The main goal is to extend and simplify the well known method of approximations proposed by A. Razborov in 1985. The main result is the following combinatorial criterion for the monotone circuit complexity: a monotone Boolean function $f(X)$ of n variables $X = \{x_1, \ldots, x_n\}$ requires monotone circuits of size $\exp(\Omega(t/\log t))$ if there is a family $\mathcal{F} \subseteq 2^X$ such that: (i) each set in $\mathcal{F}$ is either a minterm or a maxterm of f, and (ii) $D_k(\mathcal{F})/D_{k+1}(\mathcal{F}) \geq t$ for every $k = 0, 1, \ldots, t-1$. Here $D_k(\mathcal{F})$ is the *k-th degree* of $\mathcal{F}$, i.e. maximum cardinality of a subfamily $\mathcal{H} \subseteq \mathcal{F}$ with $|\cap\, \mathcal{H}| \geq k$.

1 Introduction

The question of determining how much economy the universal non-monotone basis $\{\wedge, \vee, \neg\}$ provides over the monotone basis $\{\wedge, \vee\}$ has been a long standing open problem in Boolean circuit complexity. In 1985, Razborov [10, 11] achieved a major development in this direction. He worked out the, so-called, "method of approximations" and proved that Clique function requires super-polynomial circuits over $\{\wedge, \vee\}$. He then proved the same lower bound for Perfect Matching function which is known (see [7, 15]) to have polynomial size circuit over the complete basis $\{\wedge, \vee, \neg\}$. Using arguments similar to those in [10, 11], Andreev [3] and Alon and Boppana [2] proved that some Boolean functions in NP require exponential size monotone circuits.

After this progress the following problem to investigate as a "next step" arise naturally ([12]):

> Is there a uniform and tractable criterion for the monotone circuit complexity ? Or equivalently, what are the combinatorial properties of a Boolean function that imply (or even characterize) its hardness with respect to monotone circuits ?

Due to the observation made by Berkowitz in [4] that for some monotone functions their monotone and non-monotone circuits are of almost the same size, this last question deserves a special attention.

In this paper we prove some results answering this question. The main result is a simple and "purely combinatorial" criterion for monotone circuit complexity of Boolean functions. The criterion states that a monotone Boolean function $f(X)$ of n variables $X = \{x_1, \ldots, x_n\}$ requires monotone circuits of size $\exp(\Omega(t/\log t))$ if there is a family $\mathcal{F} \subseteq 2^X$ such that: (i) each set in $\mathcal{F}$ is either a minterm or a maxterm of f, and (ii) $D_k(\mathcal{F})/D_{k+1}(\mathcal{F}) \geq t$ for every $k = 0, 1, \ldots, t-1$. Here $D_k(\mathcal{F})$ is the maximum cardinality $|\mathcal{H}|$ of a subfamily $\mathcal{H} \subseteq \mathcal{F}$ with $|\cap\, \mathcal{H}| \geq k$. Using this criterion it becomes an easy task to re-prove known exponential lower bounds form [2, 3, 11, 12].

The criterion is derived via extension of Rarborov's approximations to more general class of circuits. Besides monotone circuits, this extension allows one to answer, for example, the following question about the size of *non-monotone* circuits. Each circuit over the basis $\{\wedge, \vee, \neg\}$ computes in a natural way not only the function f itself but also some DNF D of f. Thus, given a DNF D, one may ask what is the size of an optimal circuit over $\{\wedge, \vee, \neg\}$ computing D. The extension allows to show (see [8]) that some DNFs are 'incompressible', i.e. any circuit over $\{\wedge, \vee, \neg\}$ computing them must have almost the same number of gates as the number of monomials in them.

2 The main result

Before presenting the criterion in full generality, we will state a special case of our main theorem proved in Section 4. The definitions are easier to understand, and the criterion is more intuitive.

It is convenient for our purposes to look at Boolean functions $f : \{0,1\}^n \to \{0,1\}$ as mappings $f : 2^X \to \{0,1\}$ where X is a finite set of cardinality n. That is, we identify the arguments of f with subsets of X in the natural way: $f(A) = f(x_1, \ldots, x_n)$, where $x_i = 1$ if $i \in A$ and $x_i = 0$ otherwise. A Boolean function f is *monotone* if $A \subseteq B \implies f(A) \leq f(B)$. Throughout this section we will consider only monotone functions.

Given a set $Y \subseteq X$, we say that a set $A \subseteq X$ is *Y-critical* for f (or (f,Y)-critical) if $A \cap Y = \emptyset$, $f(A \cup Y) = 1$ and $f(X - A) = 0$. Let $CRIT(f,Y)$ denote the set of all such sets.

Informally, the criterion below states that f requires large monotone circuits if, for some $Y \subseteq X$, the family of (f,Y)-critical sets is "highly dispersed". To formalize this, let $\mathcal{F} \subseteq 2^X$ and $0 \leq k \leq n$. The k-th degree $D_k(\mathcal{F})$ of $\mathcal{F}$ is the maximum number of sets in $\mathcal{F}$ with at least k elements in common, i.e.

$$D_k(\mathcal{F}) = \max\, \{\, |\mathcal{F}'| \;:\; \mathcal{F}' \subseteq \mathcal{F} \text{ and } |\cap\, \mathcal{F}'| \geq k \,\}.$$

Notice that $D_0(\mathcal{F}) = |\mathcal{F}|$ and $D_k(\mathcal{F})$ decreases as k increases. The rate of this decrease is characterized by fractions $D_k(\mathcal{F})/D_{k+1}(\mathcal{F})$.

For a family $\mathcal{F} \subseteq 2^X$, let $\mu(\mathcal{F})$ be the maximum number $t \geq 0$ such that

$$D_k(\mathcal{F})/D_{k+1}(\mathcal{F}) \geq t \qquad \text{for all } k = 0, 1, \ldots, t.$$

For a Boolean function $f : 2^X \to \{0,1\}$, let

$$\mu(f) = \max_{Y \subseteq X} \max_{\mathcal{F}} \mu(\mathcal{F}).$$

where $\mathcal{F}$ ranges over all $\mathcal{F} \subseteq CRIT(f,Y)$.
Notice that for any function f of n arguments, it holds that

$$0 \leq \mu(f) \log \mu(f) \leq n.$$

Indeed, if $\mu(f) = t$ and $\mathcal{F} \subseteq 2^X$ then

$$1 \leq D_t(\mathcal{F}) \leq t^{-1} D_{t-1}(\mathcal{F}) \leq \cdots \leq t^{-t} D_0(\mathcal{F}) = t^{-t}|\mathcal{F}|,$$

and hence, $t \log t \leq \log |\mathcal{F}| \leq |X|$.
We consider usual monotone circuits with inputs $x_1, \ldots, x_n$, with AND and OR gates allowed but no NOT gates allowed. For a monotone Boolean function f, let $C_+(f)$ denote the minimum number of gates in a monotone circuit computing f.

THEOREM **2.1** *For any monotone Boolean function f, we have that*

$$C_+(f) \geq \exp\Big(\Omega(\mu(f)/\log \mu(f))\Big).$$

We postpone the proof of this criterion to Section 5. It is a direct consequence from general theorem proved in Section 4.

Due to an extreme simplicity of the measure $\mu(f)$, it is an easy task to prove non-trivial lower bounds for the monotone circuit complexity of many natural functions: it is enough to estimate the degree of minterms and maxterms of f. It seems somewhat surprising that such a simple criterion yields the same lower bounds as those proved in [2, 3, 10, 11] via complicated combinatorical and probabilistic arguments.

3 The machinery

Here we recall from [8] the basic machinery which we will use in the proof of our criterion. The machinery is actually an extension of the method proposed by Razborov in [10, 11] (see also, [12, 13]) and known as the "method of approximations". Let us notice instantly that the machinery described below is somewhat too powerful for our purposes. The more general form is provided in order to avoid repetitions when applying the machinery to other purposes.

Let $\langle P, \leq \rangle$ be a finite partially ordered set (poset) with the least element 0. *Atoms* of P are its minimal elements $\neq 0$. A *predicate* over P is a function $f : P \to \{0,1\}$. We will often

identify a predicate f with the set $f^{-1}(1)$; thus, for example, $|f|$ is the cardinality of $f^{-1}(1)$ and $f-g$ stands for the predicate $f\wedge\neg g$. A predicate f is *monotone* if $x\le y \Longrightarrow f(x)\le f(y)$. We say that x *covers* y if $x\ge y$. Each subset $A\subseteq P$ defines the unique monotone predicate $\lceil A\rceil$ which given $y\in P$, computes 1 iff y covers some $x\in A$. We say that a predicate g *covers* a predicate f if $f\le g$, i.e. if $f(x)\le g(x)$ for all $x\in P$. For a singleton set $\{x\}$ we write $\lceil x\rceil$ instead of $\lceil\{x\}\rceil$, A *star* is a predicate $f_x=\lceil x\rceil$ where $x\in P$; f_x is a *prime star* if x is an atom of P.

Throughout, a *circuit* over a poset P will mean the circuit with gates $\{\wedge,\vee\}$ and all prime stars over P given as inputs. Each such circuit computes some monotone predicate $f: P\to\{0,1\}$. Let $C(f,P)$ (or simply $C(f)$ if the poset is clear from the context) denote the minimum number of gates in a circuit over P computing f.

EXAMPLE. A monomial over the set of Boolean variables $\{x_1,\ldots,x_n\}$ is a subset $\mathsf{m}\subseteq\{x_1,\ldots,x_n,\neg x_1,\ldots,\neg x_n\}$ which has no contrary pair, i.e. $|\mathsf{m}\cap\{x_i,\neg x_i\}|\ \le\ 1$. Let M be the set of all monomials, and M_+ be the set of all positive monomials (i.e. monomials without negations) ordered by $\subseteq$. Predicates over M correspond to disjunctive normal forms (DNF's for short). Atoms of M are all singleton monomials $\{x_1\},\ldots,\{x_n\},\{\neg x_1\},\ldots,\{\neg x_n\}$. Thus, circuits over M compute DNF's and $C(D,\mathsf{M})$ is the minimum size of a circuit over the universal basis $\{\wedge,\vee,\neg\}$ computing the DNF D. Moreover, the minimum of $C(D,\mathsf{M})$ over all DNF's D of a Boolean function f is the *non-monotone* circuit complexity of f. On the other hand, $C(f,\mathsf{M}_+)$ is *exactly* the monotone circuit complexity of f.

QUESTION: What combinatorial properties of a monotone predicate $f: P\to\{0,1\}$ are essential for $C(f)$ being large ?

Razborov [11] reduced this problem to an instance of MINIMAL COVER problem. He calls a family of monotone predicates $\mathcal{M}$ supplied with two binary operations $\bar\wedge$ and $\bar\vee$ to be a *legitimate model* if $\mathcal{M}$ contains 0,1 and all prime predicates over P. Let Δ_+ (Δ_-) denote the set of all predicates of the form $(g*h)-(g\bar{*}h)$ (respectively, of the form $(g\bar{*}h)-(g*h)$) where $*\in\{\wedge,\vee\}$ and $g,h\in\mathcal{M}$; these predicates are called *δ-predicates*. The distance $\rho(f,\bar f)$ from a predicate f to a predicate $\bar f$ within $\mathcal{M}$ is the minimum number t for which there exist predicates $\delta_1,\ldots,\delta_t$ in δ_+ and predicates $\delta'_1,\ldots,\delta'_t$ in Δ_- such that $f\le\bar f\vee\delta_1\vee\cdots\vee\delta_t$ and $\bar f\le f\vee\delta'_1\vee\cdots\vee\delta'_t$. The distance from f to a model $\mathcal{M}$ is defined by $\rho(f,\mathcal{M})=\inf\limits_{\bar f\in\mathcal{M}}\rho(f,\bar f)$.

THEOREM **3.1** ([10]) *Let f be a monotone predicate over a poset P. Then for any legitimate model $\mathcal{M}$, we have $C(f,P)\ge\rho(f,\mathcal{M})$.*

Thus, a monotone predicate f requires large circuits over P if f cannot be covered by a small number of δ-predicates. This theorem is the heart of so-called method of approximations and has been successfully employed in [10, 11, 2]. In all these works bounding $\rho(f,\mathcal{M})$ is the most involved part using complicated combinatorial and probabilistic arguments. This is stipulated by a complicated character of δ-predicates. So, one may ask if these predicates can be replaced by more natural ones. Theorem 3.2 below claims that δ-predicates may be actually replaced by more tractable ones, called "stars" and "costars".

To formulate the result in full generality we need some auxiliary lattice-theoretic notions.

A *lower semilattice* is a poset $\langle L, \preceq \rangle$ with 0, in which each two elements $x, y \in L$ have their meet $\inf_L(x, y)$ in L (whereas the join $\sup_L(x, y)$ may be undefined). A semilattice L is *Boolean* if each its interval $[0, x]$ is a Boolean lattice. The *height* $h(x)$ of an element $x \in L$ is the maximum length of a chain in L from 0 to x.

EXAMPLE. The poset $\langle \mathsf{M}, \subseteq \rangle$ of monomials is not a lattice since the join $\mathsf{m} \cup \mathsf{n}$ of two monomials m and n may contain a contrary pair, and hence, may be not a monomial. Newertheless, $\langle \mathsf{M}, \subseteq \rangle$ is a Boolean semilattice with the height function $h(\mathsf{m}) = |\mathsf{m}|$.

Let $\langle P, \leq \rangle$ be a finite poset with the least element 0. A *skeleton* for P is a Boolean semilattice $\langle L, \preceq \rangle$ with $L \subseteq P$ such that (i) L contains 0 and all the atoms of P, and (ii) for all $x, y \in L$, $x \preceq y \implies x \leq y$.

A *costar* over L is a predicate $f_Z = \theta_Z - \lceil Z \rceil$ where $Z \subseteq L$, $|Z| \geq 2$ and

$$\theta_Z(x) = 1 \iff x \geq \inf_L(y, z) \text{ for all } y, z \in Z, y \neq z.$$

Thus, $f_Z(x) = 1$ iff x covers the meets of all pairs of elements in Z but covers no element in Z. Recall that stars are predicates of the form $f_x = \lceil x \rceil$, and hence, are monotone. Costars are not monotone predicates and are, in some sense, the "complements" of stars. For example, if $\inf_L(y, z) = 0$ for all $y \neq z \in L$, then $f_Z = \neg\left(\bigvee_{x \in L} f_x\right)$.

Let ξ be a (not necessarily uniformly distributed) random element of P. We will make use of the following two functions describing the behavior of ξ on the skeleton L :

$$\mathsf{F}_\xi(s) = \max_x \Pr[f_x(\xi) = 1]$$

and

$$\mathsf{D}_\xi(s, r) = \max_Z \Pr[f_Z(\xi) = 1]$$

where the first maximum is over all $x \in L$ with $h(x) = s$ and the second is over all subsets $Z \subseteq L$ containing $r + 1$ elements of height at most s each. (In order to indicate the skeleton used we shall also write $\mathsf{F}_\xi(s : L)$ and $\mathsf{D}_\xi(s, r : L)$.) Note that $\mathsf{F}_\xi(0) = 1$ and $\mathsf{F}_\xi(k + 1) \leq \mathsf{F}_\xi(k)$ for all $k \geq 0$. The rate to which $\mathsf{F}_\xi(k)$ decreases (as k increases) characterizes the "density" of ξ : intuitively, $\mathsf{F}_\xi(k + 1) \ll \mathsf{F}_\xi(k)$ if ξ is not concentrated on a particular star f_x with $h(x) = k + 1$.

To be more specific, say that ξ is (s, r)-*sparse* (with respect to L) if for all $k = 0, 1, \ldots, s - 1$

$$\mathsf{F}_\xi(k + 1) \leq \frac{1}{3^k \lambda(r, k)}.$$

Here $\lambda(r, k)$ is the *growth function* of the skeleton L defined inductively as follows: $\lambda(r, 0) = \lambda(1, k) = 1$, and for $r \geq 2$, $\lambda(r, k) = \sum_{i=0}^{k} \nu(k, i)\lambda(r - 1, i)$ where $\nu(k, i)$ is the maximum over all $x \in L$ of height k, of the number of elements in the i-th slice of the interval $[0, x] \subseteq L$. For example, if L is the Boolean lattice of all subsets of a finite set, then $\nu(k, i) = \binom{k}{i}$ and hence, $\lambda(r, k) = r^k$.

THEOREM **3.2** *Let f be a monotone predicate over a poset P, $L \subseteq P$ be a skeleton for P; $s, r \geq 1$ be integers and $l = \lceil (s+1)/2 \rceil$. Then for any (s,r)-sparse random element ξ of L and for any random element η of P we have that*

$$C(f,P) \geq \min \left\{ \frac{\Pr[f(\xi)=1] - \frac{1}{2}}{4\,\mathsf{F}_\xi(l:L)\lambda(r,l)},\ \ \frac{\Pr[f(\eta)=0]}{\mathsf{D}_\eta(s,r:L)\lambda(r+1,s)} \right\}. \tag{1}$$

Let us instantly notice that a very special case of this general theorem will be sufficient for the purposes of this paper. Namely, Theorem 2.1 will be deduced from this theorem in case when P is the Boolean lattice of all subsets of a finite set and $L = P$. However, the proof of Theorem 3.2 in this special case is not simpler than that in the general case. So, the more general form is provided in order to avoid repetitions when applying the machinery to other purposes.

This theorem is a slightly modified version of the main theorem from [8]. To make the paper self-contained, the full proof of Theorem 3.2 is given in the appendix. We just say a few words about it here. The arguments are similar to those appearing in [10, 11, 2] for the special poset P, namely, for the Boolean lattice $\langle\, 2^X, \subseteq \rangle$ of all subsets of a finite set X. The main step in [10, 11, 2] was to define an appropriate notion of "closure" and to prove that closed families of sets contain only small number of minimal sets. The key lemma of our proof is the extension of this result from the lattice $\langle\, 2^X, \subseteq \rangle$ to an arbitrary Boolean semilattice $\langle\, L, \preceq \,\rangle$. Theorem 3.2 was used in [8] to prove lower bounds for some classes of circuits over the complete basis $\{\wedge, \vee, \neg\}$. Here we will use the theorem to obtain a uniform criterion for the complexity of circuits over the monotone basis $\{\wedge, \vee\}$. The criterion is obtained by proving that in this case the function $\mathsf{D}_\eta(\cdot)$ in (1) can be replaced by more natural function $\mathsf{F}_\eta(\cdot)$.

4 The criterion

Recall that Boolean functions f are predicates over the Boolean lattice $P(X) = \langle\, 2^X, \subseteq \rangle$ of all subsets of a finite set X, $|X| = n$. Monotone circuits are circuits over the lattice $P(X)$.

A set $A \subseteq X$ is a 1-*term* (0-*term*) of f if setting all variables in A to 1 (0), forces the value of f to 1 (0). Throughout, let $T_1(f)$ ($T_0(f)$) denote the set of all 1-terms (0-terms) of f. *Minterms* (*maxterms*) are minimal 1-terms (0-terms). Notice that, although the families $f^{-1}(1)$ and $f^{-1}(0)$ are disjoint, the families $T_1(f)$ and $T_0(f)$ may be not, i.e. it may be that some subset $A \subseteq X$ is both a 1-term and a 0-term of f. In particular, if f is non-constant function then $X \in T_1(f) \cap T_0(f)$.

We say that a poset $\langle\, \mathcal{L}, \subseteq \,\rangle$ with $\mathcal{L} \subseteq 2^X$ is a *legitimate skeleton* for the lattice $P(X) = \langle\, 2^X, \subseteq \,\rangle$ if

(i) $\mathcal{L}$ contains all one-element sets $\{x\}$, $x \in X$; and

(ii) for any $A \in \mathcal{L}$, $\langle\, 2^A \cap \mathcal{L}, \subseteq \,\rangle$ is a Boolean lattice.

Notice that any legitimate skeleton $\langle\, \mathcal{L}, \subseteq \,\rangle$ is a skeleton for $P(X) = \langle\, 2^X, \subseteq \,\rangle$ with the height function h satisfying $0 = h(\emptyset) \leq h(A) \leq |A|$, and the growth function $\lambda(r,s) \leq r^s$. In particular, the lattice $P(X)$ is a legitimate skeleton for itself.

Razborov asked in [12] if there exists a unique and tractable criterion for the monotone complexity of Boolean functions. The theorem below gives such a criterion. Informally, the criterion states that a monotone Boolean function f requires large monotone circuits if for some legitimate skeleton $\mathcal{L}$, both $T_0(f)$ and $T_1(f)$ cannot be covered by a small number of "stars", i.e. by families of the form $\lceil\{A_0\}\rceil = \{A : A \supseteq A_0\}$.

Convention: To make our notations simpler, assume from now that: $L(X)$ stands for an arbitrary legitimate skeleton for $P(X)$; $\boldsymbol{A}$ is any (s,r)-sparse random set in $L(X)$; $\boldsymbol{B}$ is a random set in $P(X)$; $s \geq 1$ and $r \geq 1$ are fixed but arbitrary integers; $l = \lceil (s+1)/2 \rceil$, $\epsilon_+ = \frac{1}{4} r^{-l}$, and $\epsilon_- = (r+1)^{-s} c(s)^{-(r+1)}$ where $c(s)$ is the maximum cardinality of a set in $L(X)$ of height s.

THEOREM **4.1** *For any monotone Boolean function $f : P(X) \to \{0,1\}$ we have that*

$$C_+(f) \;\geq\; \min\{\, \epsilon_+ \cdot \Phi^+(\boldsymbol{A}),\; \epsilon_- \cdot \Phi^-(\boldsymbol{B})\}$$

where

$$\Phi^+(\boldsymbol{A}) = \frac{\Pr[\boldsymbol{A} \in T_1(f)] - 1/2}{\mathsf{F}_{\boldsymbol{A}}(l : L(X))}, \tag{2}$$

$$\Phi^-(\boldsymbol{B}) = \frac{\Pr[\boldsymbol{B} \in T_0(f)]}{\mathsf{F}_{\boldsymbol{B}}(r+1 : P(X))}. \tag{3}$$

REMARK. Notice that the density function $\mathsf{F}_{\boldsymbol{A}}(\cdot)$ of $\boldsymbol{A}$ in (2) is defined with respect to the skeleton $L(X) \subseteq P(X)$ whereas that of $\boldsymbol{B}$ in (3) is defined with respect to the whole lattice $P(X)$.

Proof. We will apply Theorem 3.2. Since $\lambda(r,s) = r^s$, we have that $\epsilon_+ \cdot \Phi^+(\boldsymbol{A})$ is exactly the first term in (1). Thus, it suffices to prove that the second term in (1) is bounded from below by $\epsilon_- \cdot \Phi^+(\boldsymbol{B})$.

Let us consider the random set $\boldsymbol{C} = X - \boldsymbol{B}$. Since $\Pr[f(\boldsymbol{C}) = 0] = \Pr[\boldsymbol{B} \in T_0(f)]$, it is enough to prove that

$$\mathsf{D}_{\boldsymbol{C}}(s, r : L(X)) \leq c(s)^{r+1}\, \mathsf{F}_{\boldsymbol{B}}(r+1 : P(X)). \tag{4}$$

Call a family $\mathcal{Z} \subseteq L(X)$ to be an (r,s)-family if $|\mathcal{Z}| = r+1$ and $h(A) \leq s$ for all $A \in \mathcal{Z}$.

By the definition, $\mathsf{D}_{\boldsymbol{C}}(s, r : L(X))$ is the maximum probability of an event that $f_{\mathcal{Z}}(\boldsymbol{C}) = 1$ where $f_{\mathcal{Z}}$ is the costar induced by an (r,s)-family $\mathcal{Z}$.

Fix an (r,s)-family $\mathcal{Z}$ maximizing the probability $\Pr[f_{\mathcal{Z}}(\boldsymbol{C}) = 1]$, and let $\Theta(\mathcal{Z})$ denote the family of all subsets $A \subseteq X$ such that for all $B \in \mathcal{Z}$, $A \not\supseteq B$ but $A \supseteq \widehat{B}$ where $\widehat{B} = \bigcup\{B \cap C \;:\; B \neq C \in \mathcal{Z}\}$. Since the meet in $\mathcal{L}$ coincides with $\cap$, we have that $\Theta(\mathcal{Z}) = f_{\mathcal{Z}}^{-1}(1)$. Thus, $\mathsf{D}_{\boldsymbol{C}}(s, r : L(X)) = \Pr[\boldsymbol{C} \in \Theta(\mathcal{Z})]$, and hence, (4) is equivalent to

$$\Pr[\boldsymbol{C} \in \Theta(\mathcal{Z})] \;\leq\; c(s)^{r+1}\, \mathsf{F}_{\boldsymbol{B}}(r+1 : P).$$

Say that a subset $A \subseteq X$ is a *transversal* for a family $\mathcal{Z}$ if $A \cap (B - \widehat{B}) \neq \emptyset$ for all $B \in \mathcal{Z}$. Let $\mathtt{tr}(\mathcal{Z})$ denote the set of all transversals for $\mathcal{Z}$. For a family $\mathcal{F}$, denote co-$\mathcal{F} = \{X - A \mid A \in \mathcal{F}\}$.

It is easy to see that for any family $\mathcal{Z}$,

$$\text{co-}\Theta(\mathcal{Z}) \;\subseteq\; \mathtt{tr}(\mathcal{Z}). \tag{5}$$

Indeed, let $A \in \text{co-}\Theta(\mathcal{Z})$. Then $X - A \in \Theta(\mathcal{Z})$, and hence $\forall B \in \mathcal{Z} \;:\; X - A \not\supseteq B$ but $X - A \supseteq \widehat{B}$. This means that $A \cap (B - \widehat{B}) \neq \emptyset$ for all $B \in \mathcal{Z}$, and thus, $A \in \mathtt{tr}(\mathcal{Z})$.

Thus, by (5),

$$\Pr[\boldsymbol{C} \in \Theta(\mathcal{Z})] \;\leq\; \Pr[\boldsymbol{B} \in \mathtt{tr}(\mathcal{Z})]. \tag{6}$$

It remains to estimate the right-hand side of (6). Recall that $\boldsymbol{B} \in \mathtt{tr}(\mathcal{Z})$ iff $\widehat{\mathcal{Z}} \neq \emptyset$ and $\boldsymbol{B}$ intersects every set in $\widehat{\mathcal{Z}}$ where $\widehat{\mathcal{Z}} = \{B - \widehat{B} \;:\; B \in \mathcal{Z}\}$. Let $\mathcal{T}$ be the family of all minimal sets in $\mathtt{tr}(\mathcal{Z})$. Then the event "$\boldsymbol{B} \in \mathtt{tr}(\mathcal{Z})$" is the union of the events "$\boldsymbol{B} \supseteq T$" where $T \in \mathcal{T}$. Since $|\mathcal{Z}| = r+1$ and every set in $\widehat{\mathcal{Z}}$ is of cardinality at most $c(s)$, we have that $\mathcal{T}$ consists of at most $c(s)^{r+1}$ sets, each of cardinality exactly $r+1$. Hence,

$$\mathsf{D}_{\boldsymbol{C}}(s, r : L(X)) \;\leq\; \Pr[\boldsymbol{B} \in \mathtt{tr}(\mathcal{Z})] \;\leq\; |\mathcal{T}|\, \mathsf{F}_{\boldsymbol{B}}(r+1 : P(X))$$

where $|\mathcal{T}| \leq c(s)^{r+1}$. This completes the proof of (4), and thus, the proof of Theorem 4.1. □

For most applications of Theorem 4.1, including the proof of Theorem 2.1, it is enough to consider trivial skeletons, namely, one may take $L(X) = P(X)$. To illustrate how non-trivial legitimate skeletons may arise, let us consider the following NP-complete problem.

Let V be a finite set, $|V| = m$, and $X = \binom{V}{2}$; elements of X are edges, i.e pairs $\{u, v\}$ with $u, v \in V$, and $P(X) = \langle\, 2^X, \subseteq \,\rangle$ is the lattice of all graphs on V. A graph $A \subseteq X$ is a clique if $A = \binom{U}{2}$ for some $U \subseteq V$; A is a q-clique ($2 \leq q \leq m$) if $|U| = q$. In particular, edges are minimal non-empty cliques. Let $\mathcal{A}$ be the family of all cliques. Since the union of two (incomparable) cliques is not a clique, the poset $L(X) = \langle\, \mathcal{A}, \subseteq \,\rangle$ is not a lattice. Nevertheless, this poset is a Boolean semilattice since for any clique $A = \binom{U}{2}$, the poset $\langle\, 2^A \cap \mathcal{A}, \subseteq \,\rangle$ is a Boolean lattice isomorphic to the lattice $P(U)$. Thus, $L(X)$ is a legitimate skeleton for $P(X)$ with the height function $h\left(\binom{U}{2}\right) = |U|$ and the growth function $\lambda(r,k) = r^k$. Since a clique with s vertices has $\binom{s}{2}$ edges, we have also that for this skeleton $c(s) = \binom{s}{2}$.

Let $CLIQUE(m,q) : 2^X \to \{0,1\}$ be the function (of $n = \binom{m}{2}$ Boolean variables) which, given a graph $G \subseteq X$, computes 1 iff G contains a q-clique. In [10] Razborov showed that for $q = \lceil (\ln m)/4 \rceil$, this function requires monotone circuits of super-polynomial $m^{\Omega \log m}$ size. Modifying his arguments, Alon and Boppana improved this bound to $\exp\left(\Omega((m/\log m)^{1/3})\right)$ for $q \approx (m/\log m)^{2/3}$.

Direct application of Theorem 4.1 yields the following lower bound.

LEMMA **4.2** *Suppose* $2 \leq q \leq m^{2/3}$, *and let* $f = CLIQUE(m,q)$. *Then for any* $1 \leq s \leq q-1$ *and* $1 \leq r \leq m/3q$ *we have that* $C_+(f) \geq \min\{\phi^+, \phi^-\}$ *where*

$$\phi^+ = \frac{1}{8}\left(\frac{m}{qr}\right)^{\lceil (s+1)/2 \rceil} \qquad \textit{and} \qquad \phi^- = \frac{1}{2}\left(\frac{q}{s^2 \ln(m/q)}\right)^{r+1} (r+1)^{-s}.$$

Proof. Let $L(X)$ be the family of all cliques. We have seen above that $L(X)$ is a legitimate skeleton for the lattice of all graphs $P(X)$. To prove the desired lower bound on $C_+(f)$ it is enough by Theorem 4.1 to choose an (s,r)-sparse (for any $s \leq q$ and $r \leq m/3q$) random clique $\boldsymbol{A}$ in $L(X)$ and a random graph $\boldsymbol{B}$ in $P(X)$ so that $\epsilon_+ \cdot \Phi^+(\boldsymbol{A}) \geq \phi^+$ and $\epsilon_- \cdot \Phi^-(\boldsymbol{B}) \geq \phi^-$.

Choose $\boldsymbol{A}$ randomly with uniform distribution on the set of all q-cliques. Then $\Pr[\boldsymbol{A} \in T_1(f)] = 1$ and for any $k \leq m-1$,

$$\mathsf{F}_{\boldsymbol{A}}(k : L(X)) = \binom{m-k}{q-k} \Big/ \binom{m}{q} \leq \left(\frac{q}{m}\right)^k. \tag{7}$$

So, $\boldsymbol{A}$ is (s,r)-sparse for any s and r satisfying $(q-s)/(m-s) \leq 1/3r$, and hence, for $s \leq q$ and $r \leq m/3q$. Putting $\epsilon_+ = \frac{1}{4} r^{\lceil (s+1)/2 \rceil}$ and the estimate (7) in 2 we get the desired upper bound for ϕ^+.

To obtain the desired upper bound for ϕ^-, let $\boldsymbol{B}_p$ be a random graph, with each edge appearing independently with probability $p = q^{-1}\ln(m/q)$. Then

$$\begin{aligned}
\Pr[\boldsymbol{B}_p \in T_0(f)] &= = 1 - \Pr[\boldsymbol{B}_{1-p} \text{ has a } q\text{-clique}] \\
&\geq 1 - \sum_{\substack{U \subseteq V \\ |U| = k}} \Pr\left[\boldsymbol{B}_{1-p} \supseteq \binom{U}{2}\right] \\
&\geq 1 - \binom{m}{q}(1-p)^{\binom{q}{2}} \geq \frac{1}{2}.
\end{aligned}$$

Since $c(s) = \binom{s}{2} < s^2$ and $\mathsf{F}_{\boldsymbol{B}_p}(r+1 : P(X)) = p^{r+1}$, (3) gets the desired upper bound for ϕ^-, which completes the proof of Lemma 4.2. □

5 The criterion (symmetric version)

Although general, Theorem 4.1 in its present form may be hard to apply for some functions directly. Suppose, for example, that f has the form $f = \lceil \{Y\} \rceil \wedge f'$ where f' is a function defined on $P(\overline{Y})$. Then, for any random set $\boldsymbol{A}$ on $P(X)$ and for any integer $k \leq |Y|$, we have that $\mathsf{F}_{\boldsymbol{A}}(k : P(X)) \geq \Pr[\boldsymbol{A} \in T_1(f)]$ since $\boldsymbol{A} \in T_1(f)$ yields $\boldsymbol{A} \supseteq Y$, and $\mathsf{F}_{\boldsymbol{A}}(k : P(X))$ is the maximum of $\Pr[\boldsymbol{A} \supseteq A]$ over *all* sets $A \in P(X)$ with $|A| = k$. There is a simple way to overcome this difficulty. Namely, one may apply Theorem 4.1 to the subfunction f' rather than to the function f itself. Since f' is defined on the smaller lattice $P(\overline{Y})$, one may replace $\mathsf{F}_{\boldsymbol{A}}(k : P(X))$ by $\mathsf{F}_{\boldsymbol{A}}(k : P(\overline{Y}))$ which may be much smaller than $\mathsf{F}_{\boldsymbol{A}}(k : P(X))$ since now the maximum is taken over sets A in $P(\overline{Y})$ only. Let us state this amplification more precisely.

We start by recalling the following well known fact concerning subfunctions. Given a function $f : P(X) \to \{0,1\}$, its subfunctions are defined as follows. For $Y, Z \subseteq X$ with $Y \cap Z = \emptyset$, let

$$P_{Y,Z}(X) = \{A \,:\, A \subseteq \overline{Y \cup Z}\}.$$

The (Y,Z)-*subfunction* of a Boolean function $f : P(X) \to \{0,1\}$ is the Boolean function $f_{Y,Z} : P_{Y,Z}(X) \to \{0,1\}$ defined by

$$f_{Y,Z}(A) = 1 \iff f(A \cup Y) = 1.$$

In other words, the subfunction $f_{Y,Z}$ is the function of $|X| - |Y \cup Z|$ variables and is obtained from f by setting all the variables in Y (Z) to 1(0).

FACT **5.1** *For any function $f : P(X) \to \{0,1\}$ and any $Y, Z \subseteq X$ with $Y \cap Z = \emptyset$, it holds that*

$$C_+(f) \geq C\Big(f_{Y,Z}, P_{Y,Z}(X)\Big).$$

Proof. Given a circuit over $P(X)$ computing f, replace all its inputs $\lceil x \rceil$ with $x \in Y$ ($x \in Z$) by 1 (0). The resulting circuit is a circuit over the lattice $P_{Y,Z}(X)$ and computes the projection $f_{Y,Z}$. □

Recall that a subset $A \subseteq \overline{Y}$ is Y-*critical* for f if $f(A \cup Y) = 1$ and $f(\overline{A}) = 0$; $CRIT(f,Y)$ denotes the family of all subsets $A \subseteq \overline{Y}$ which are Y-critical for f. The main property of critical sets is that they are both 1-terms and 0-terms of the $(Y,\emptyset)$-subfunction of f.

FACT **5.2** *Let $Y \subseteq X$ and f' be the $(Y,\emptyset)$-subfunction of $f : P(X) \to \{0,1\}$. Then*

$$CRIT(f,Y) \subseteq T_1(f') \cap T_0(f').$$

Proof. Let $A \subseteq \overline{Y}$ be a Y-critical set for f. We have to show that both $f'(A) = 1$ and $f'(A^c) = 0$ where $A^c = \overline{Y} - A$ is the complement of A in $P(\overline{Y})$.
By the definition of critical sets, $f'(A) = f(A \cup Y) = 1$. On the other hand, $A^c \cup Y = \overline{A}$, and hence, $f'(A^c) = f(A^c \cup Y) = f(\overline{A}) = 0$. □

Facts 4.2 and 4.3 lead to the following symmetrical version of Theorem 4.1.

THEOREM **5.3** *Let $f : P(X) \to \{0,1\}$ be a monotone Boolean function, $Y \subseteq X$ and $L(\overline{Y})$ be a legitimate skeleton for $P(\overline{Y})$. Then for any (s,r)-sparse random set $\boldsymbol{A}$ in $L(\overline{Y})$, we have*

$$C_+(f) \geq \alpha \min\Big\{ \epsilon_+ \cdot \mathsf{F}_{\boldsymbol{A}}(l : L(\overline{Y}))^{-1},\ \epsilon_- \cdot \mathsf{F}_{\boldsymbol{A}}(r+1 : P(\overline{Y}))^{-1} \Big\}$$

where $\alpha = \Pr[\boldsymbol{A}$ is Y-critical for $f] - \frac{1}{2}$

Proof. Let f' be the $(Y,\emptyset)$-subfunction of f. By 5.1 , $C_+(f) = C(f, P(X)) \geq C(f', P(\overline{Y}))$. On the other hand, by 5.2 we have that both $\Pr[\boldsymbol{A} \in T_1(f')]$ and $\Pr[\boldsymbol{A} \in T_0(f')]$ are at least α. With this in mind Theorem 4.1 yields the desired lover bound on $C(f', P(\overline{Y}))$, and hence, on $C_+(f)$. □

We conclude this section by the proof of Theorem 2.1 stated in Section 2.

Proof of Theorem 2.1. Set $t = \mu(f)$, and take $Y \subseteq X$ and $\mathcal{F} \subseteq CRIT(f,Y)$ for which $\mu(\mathcal{F}) = \mu(f) = t$. We will apply Theorem 5.3 with $L(\overline{Y}) = P(\overline{Y})$. Notice that in this case $\epsilon_+ = \frac{1}{4}r^{-\lceil (s+1)/2 \rceil}$ and $\epsilon_- = (r+1)^{-s}s^{-(r+1)}$ since $c(s) = s$.

Choose $\boldsymbol{A}$ randomly with uniform distribution on $\mathcal{F}$. Then $\boldsymbol{A}$ is Y-critical for f with probability 1. Next, by the definition of $\mu(\mathcal{F})$, we have that $\mathsf{F}_{\boldsymbol{A}}(k+1)/\mathsf{F}_{\boldsymbol{A}}(k) = D_{k+1}(\mathcal{F})/D_k(\mathcal{F}) \leq t^{-1}$ for all $k \leq t$, and hence, $\boldsymbol{A}$ is (s,r)-sparse for any $s, r \leq t/3$. Moreover, since $\mathsf{F}_{\boldsymbol{A}}(0) = 1$, we have that $\mathsf{F}_{\boldsymbol{A}}(k) \leq t^{-k}$ for all $k \leq t$. Thus, Theorem 5.3 yields the following lower bound

$$C_+(f) \;\geq\; \min\left\{ \frac{1}{8}\left(\frac{t}{r}\right)^{\lceil (s+1)/2 \rceil},\; \left(\frac{t}{s}\right)^{r+1}(r+1)^{-s} \right\}$$

which for $r = \lfloor t/3 \rfloor$ and $s = \lfloor r/\log r \rfloor$ yields the desired lower bound $C_+(f) \geq \exp\left(\Omega(t/\log t)\right)$. □

References

[1] A. Ajtai and Y. Gurevich, *Monotone versus positive,* Journal of ACM, 34:5 (1987), pp. 1004-1015

[2] N. Alon and R. Boppana, *The monotone circuit complexity of Boolean functions,* Combinatorica, 7:1 (1987), pp. 1-22.

[3] A. E. Andreev, *On a method for obtaining lower bounds for the complexity of individual monotone functions,* Doklady Akademii Nauk SSSR, 282:5 (1985), pp. 1033-1037. English translation in: Soviet Mathematics Doklady, 31:3, 539-534

[4] S.J. Berkowitz, *On some relations between monotone and non-monotone circuit complexity,* Technical Report, Computer Science Department, University of Toronto, 1982.

[5] Z. Füredi *On maximal intersecting families of finite sets,* J. Comb. Theory (A), 28 (1980), pp. 282-289

[6] M. Grötschler, L. Lovász and A. Schrijver, *The elipsoid method and its consequences in combinatorial optimization,* Combinatorica, 1 (1981), pp. 169-197

[7] J. E. Hopcroft and R.M. Karp, *An $n^{5/2}$ algorithm for maximum matching in bipartite graphs,* SIAM J. Comput., 4 (1973), pp. 225-231

[8] S. Jukna, *Functional approximations in the theory of lower bounds for circuit complexity,* Diskretnaja Matematika, 2:2 (1990), pp. 45-59 (in Russian)

[9] L. Lovász, *On the Shannon capacity of graphs,* IEEE Trans. on Information Theory, 25 (1979), pp. 1-7

[10] A. A. Razborov, *A lower bound on the monotone network complexity of the logical permanent,* Matematicheskie Zametki, 37:6 (1985) pp. 887-990 (in Russian); English translation in: Math. Notes Acad. of Sciences USSR, 37:6, pp. 485-493.

[11] A. A. Razborov, *Lower bounds on the monotone complexity of some Boolean functions,* Doklady Akademii Nauk SSSR, 281:4 (1985), pp. 798-801. English translation in: Soviet Mathematics Doklady, 31, pp. 354-357

[12] A. A. Razborov, *Lower bounds on the monotone complexity of Boolean functions,* Proceedings of the International Congress of Mathematicians, Berkeley, California (1986), pp. 1478-1487

[13] A. A. Razborov, *On the method of approximations* Proceedings of the 21th Annual ACM Symposium on Theory of Computing, (1989)

[14] É. Tardos, *The gap between monotone and non-monotone circuit complexity is exponential,* Combinatorica, 7:4 (1987), pp. 141-142

[15] L. G. Valiant, *The complexity of computing the permanent,* Theoret. Comput. Sci., 8 (1979), pp. 181-201

6 Appendix A: The proof of Theorem 3.2

This appendix contains the full proof of Theorem 3.2. A variant of this theorem was proved in [8].

By Theorem 3.1 it is enough to define an appropriate model $\mathcal{M}$ for which the distance $\rho(f, \mathcal{M})$ is bounded from the below by the right-hand side of (1). To achieve this goal, we make use of "generalized filters" or "closed set", a notion introduced in [11].

Let $\langle M, \preceq \rangle$ be a Boolean semilattice. Following [11] say that an element $x \in M$ is *derivable* from a subset $Z \subseteq M$ if $f_Z(x) = 1$, where f_Z is the costar over M defined by Z (see Section 3). A *semifilter* over M is a subset $F \subseteq M$ such that $F \ni x \preceq y \in M \implies y \in M$. An *r-filter* over M is a semifilter $F \subseteq M$ which together with any subset $Z \subseteq F$, $|Z| = r+1$ contains all the points of M derivable from Z, i.e.

$$Z \subseteq F \text{ and } |Z| = r+1 \text{ implies } f_Z^{-1}(1) \subseteq F.$$

Note that 1-filters are just usual filters over M. (This is why we prefer the term "filter" instead of "closed set").

Since the intersection of two r-filters is an r-filter, we may associate with any subset $A \subseteq M$ the minimal r-filter A^* containing A, namely $A^* = \bigcap\{B | A \subseteq B \text{ and } B \text{ is an } r\text{-filter }\}$. The nice property of filters is that they have small number of minimal elements.

Lemma **6.1** *Let M be a Boolean semilattice with the growth function λ. Then for any semifilter A over M and an integer $k \geq 0$, both A^* and $A^* - A$ have no more than $\lambda(r,k)$ minimal elements of height k.*

We postpone the proof of this combinatorial lemma to the appendix B . Here we just say that similar result for the number of minimal elements in A^* has been proved by Z. Füredi

[5] in case when M is the Boolean lattice $\langle\, 2^X, \subseteq \,\rangle$ of all subsets of a finite set X, and was one of the key lemmas in [2, 10, 11]; our lemma extends the result to $A^* - A$ and to arbitrary semilattices.

Proof of Theorem 3.2: Let M be the set of all elements in the skeleton L of height at most s, and take

$$\mathcal{M} = \{\lceil F \rceil : F \text{ is an } r\text{-filter over } M\}.$$

Define the operations $\bar{\wedge}$ and $\bar{\vee}$ on $\mathcal{M}$ by $\lceil F \rceil \bar{\wedge} \lceil G \rceil = \lceil F \cap G \rceil$ and $\lceil F \rceil \bar{\vee} \lceil G \rceil = \lceil (F \cup G)^* \rceil$ where, for $A \subseteq M$, A^* stands for the least r-filter over M containing A. Since $s \geq 1$, the semilattice M contains all the atoms of P, and hence $\mathcal{M}$ is a legitimate model. By Theorem 3.1, $C(f) \geq \rho(f, \mathcal{M})$. Put $t = \rho(f, \mathcal{M})$, and let $\bar{f}$ be a predicate in $\mathcal{M}$ for which $\rho(f, \bar{f}) = t$. As in [2] we consider two possible cases.

Case 1. $\bar{f} \neq 1$.

We prove that in this case

$$t \geq \frac{\Pr[f(\xi) = 1] - \frac{1}{2}}{4\, \mathsf{F}_\xi(l)\lambda(r, l)}.$$

By the definition of $\rho(f, \mathcal{M})$, there exists a family of δ-predicates $\Delta \subseteq \Delta_+$ such that $|\Delta| \leq t$ and $f \leq \bar{f} \vee \bigvee \Delta$. Thus, the desired lower bound on t is a direct consequence of the following two claims.

Claim 1. $\Pr[\bar{f}(\xi) = 1] \leq 1/2$.

Claim 2. *For any* $\delta \in \Delta_+$, $\Pr[\delta(\xi) = 1] \leq 4\, \mathsf{F}_\xi(l)\lambda(r, l)$.

Proof of Claim 1. Let F be an r-filter for which $\bar{f} = \lceil F \rceil$. Notice that since $\bar{f} \neq 1$, each element of F is of height at least 1. Each element $x \in P$ for which $\bar{f}(x) = 1$, covers some minimal element of F. (Here and throughout, the minimality means the minimality in $M \subseteq L$). By Lemma 6.1, for each $1 \leq k \leq s$, the filter F has at most $\lambda(r, k)$ minimal elements of height k. Since $\mathsf{F}_\xi(0)\lambda(r, 0) = 1$ and ξ is (s, r)-sparse, we conclude that

$$\Pr[\bar{f}(\xi) = 1] \;\leq\; \sum_{k=1}^{s} \mathsf{F}_\xi(k)\lambda(r, k) \;\leq \sum_{k=1}^{s} \left(\frac{1}{3}\right)^k < 1/2.$$

Proof of Claim 2. Let $\delta \in \Delta_+$. If $\delta = 0$ then there is nothing to prove. Assume that $\delta \neq 0$. Then $\delta = (\lceil F \rceil \cap \lceil G \rceil) - (\lceil F \cap G \rceil)$ for some r-filters F and G over M.

Take an element $x \in L$ for which $\delta(x) = 1$. Then x covers a minimal element $a \in F$ and a minimal element $b \in G$, but no element of $F \cap G$. Since L is a Boolean semilattice L, we have that elements a and b have their join $c = \sup_L(a, b)$ in the interval $[0, x] \subseteq L$. Moreover, at least one of a or b must have the height at least l since otherwise c would have the height $h(c) \leq h(a) + h(a) \leq s$, and hence, would belong to $F \cap G$ since both F and G are semifilters over M; but this is impossible since $\lceil F \cap G \rceil(x) = 0$. We therefore conclude that each element $x \in L$ for which $\delta(x) = 1$, covers a minimal element of height $k \geq l$ of either F or G (or both). Let A be the set of these minimal elements, and for $l \leq k \leq s$, put $A(k) = \{a \in A :\ h(a) = k\}$. By 6.1, $|A(k)| \leq 2\lambda(r, k)$. Since $\delta \leq \lceil A \rceil$ and ξ is (s, r)-sparse, we have that

$$\begin{aligned}\Pr[\delta(\xi)=1] &\le \sum_{k=l}^{s}\sum_{a\in A(k)}\Pr[\lceil a\rceil(\xi)=1] \le \sum_{k=l}^{s}\sum_{a\in A(k)}\mathsf{F}_\xi(k)\\ &\le 2\sum_{k=l}^{s}\mathsf{F}_\xi(k)\lambda(r,k) \le 2\,\mathsf{F}_\xi(l)\lambda(r,l)\sum_{k=0}^{s-l}\left(\frac{1}{3}\right)^k\\ &< 4\,\mathsf{F}_\xi(l)\lambda(r,l).\end{aligned}$$

This completes the proof of Theorem 3.2 in Case 1. □

Case 2 $\bar{f}=1$.

We prove that in this case

$$t \ge \frac{\Pr[f(\eta)=0]}{\mathsf{D}_\eta(s,r)\lambda(r+1,s)}.$$

By the definition of $\rho(f,\mathcal{M})$, there exists a family of δ-predicates $\Delta\subseteq\Delta_-$ such that $|\Delta|\le t$ and $\bar{f}\wedge\neg f\le\vee\Delta$ where $\bar{f}\wedge\neg f=\neg f$ since $\bar{f}=1$. Hence, the desired lower bound on t is a direct consequence of the following claim.

Claim 3. *For any* $\delta\in\Delta_-$, $\Pr[\delta(\eta)=1]\le\mathsf{D}_\eta(s,r)\lambda(r+1,s)$.

Proof of Claim 3. By the definition of operations $\bar{\wedge}$ and $\bar{\vee}$, each predicate $\delta\in\Delta_-$ has the form $\delta=\lceil A^*\rceil-\lceil A\rceil$ for some semifilter $A\subseteq M$. By the definition of A^*, the predicate $\lceil A^*\rceil$ is the result of the following "closure algorithm" :

$$\lceil A\rceil = f_0\to f_1\to\cdots\to f_p=\lceil A^*\rceil$$

where $f_{i+1}=f_i\vee\lceil y_i\rceil$ and y_i is some minimal element derivable from some $Z_i\subseteq f_i^{-1}(1)$, with $|Z_i|=r+1$. (The algorithm terminates since the poset P is finite). For $i=1,\ldots,p$, let $\gamma_i=\lceil y_i\rceil-\lceil Z_i\rceil$. Then $\delta\le\gamma_1\vee\cdots\vee\gamma_p$. Indeed, if $\delta(x)=1$ then, since $\lceil A\rceil(x)=0$, there exists an i such that $f_{i-1}(x)=0$ but $f_i(x)=1$, and hence $\gamma_i(x)=1$. Let $I=\{i\ :\ y_i$ is a minimal element of $\delta\}$. By Lemma 6.1, $|I|\le\sum_{k=1}^{s}\lambda(r,k)\le\lambda(r+1,s)$. Since all the minimal elements of δ are among $\{y_1,\ldots,y_p\}$, we conclude that $\delta\le\bigvee_{i\in I}\gamma_i\le\bigvee_{i\in I}f_{Z_i}$ where f_Z is the costar induced by a subset Z. Hence, $\Pr[\delta(\eta)=1]\le|I|\,\mathsf{D}_\eta(s,r)\le\lambda(r,s+1)\,\mathsf{D}_\eta(s,r)$ which completes the proof of Claim 3, and thus, the proof of Theorem 3.2. □

7 Appendix B: The proof of Lemma 6.1

The goal of this appendix is to prove Lemma 6.1 about the number of minimal elements in (generalized) filters.

Let $\langle M,\preceq\rangle$ be a Boolean semilattice with the height function h and the growth function λ. For $x\in M$ and $Y\subseteq M$, we write $x\succ Y$ if $x\succ y$ for some $y\in Y$ where $x\succ y$ means that $x\succeq y$ and $x\ne y$.

We say that a set $A\subseteq M$ has property $P(r,k)$ if

(i) $h(x)=k$ for all $x\in A$, and

(ii) there are no $x \in A$ and $Z \subseteq A$ with $|Z| = r+1$, such that $x \succ \theta(Z)$ where

$$\theta(Z) = \{z \in M \ : \ z \succeq \inf_M(z_1, z_2) \text{ for all } z_1, z_2 \in Z, z_1 \neq z_2 \ \}.$$

Fix an integer $r \geq 1$, and for a subset $A \subseteq M$, let A^* denote the minimal r-filter over M containing A.

Lemma 6.1 is a direct consequence of the following two lemmas.

LEMMA **7.1** *Let A be a semifilter over M, and let F and G be sets of all minimal elements of height k in A^* and $A^* - A$, respectively. Then both F and G have the property $P(r,k)$.*

Proof. The set F has $P(r,k)$ by the definition of A^*. Let us check that G also has this property. Suppose the opposite, i.e. that $x \succ z$ for some $x \in G$ and $z \in \theta(Z)$ where $Z \subseteq G$ and $|Z| = r+1$. If $z \not\succeq Z$ then z is derivable from Z and hence belongs to A^*, a contradiction with the minimality of x. If $z \succeq Z$ then $h(z) \geq k$ since $Z \subseteq G$, and hence $x \succ z \Longrightarrow h(x) \geq h(z) + 1 \geq k+1$, a contradiction with $x \in G$. Thus, the set G also has the property $P(r,k)$. □

LEMMA **7.2** *If $A \subseteq M$ has the property $P(r,k)$ then $|A| \leq \lambda(r,k)$ where λ is the growth function of M.*

Proof. We argue by induction on $r \geq 1$. The basis $r = 1$ is trivial since $P(1,k) \implies |X| = 1 = \lambda(1,k)$.

Assuming the result for $r-1$, we prove it for r. Let A be a set having the property $P(r,k)$. Choose an arbitrary point $x_0 \in A$, and put

$$Y = \{x \wedge x_0 \ \mid \ x \in A - \{x_0\}\}.$$

Since M is a Boolean semilattice, each its interval $[0,x]$ is a Boolean lattice. Thus, each point $y = x \wedge x_0 \in Y$ has the complement in the interval $[0,x]$, i.e. there is an element y' such that $y' \wedge y = 0$ and $y' \vee y = x$ where $\wedge$ and $\vee$ are the meet and the join operation of the lattice $[0,x]$. Moreover, in Boolean lattices this complement is unique. Let $\partial_y(x)$ stand for the complement of y in $[0,x]$. For $y \in Y$, set

$$A_y = \{\partial_y(x) \ \mid \ x \in A \text{ and } x \wedge x_0 = y\}.$$

CLAIM *For each $y \in Y$, A_y has the property $P(r-1, k-h(y))$.*

Note that, by induction hypothesis, this claim directly yields the desired result:

$$|A| \ \leq \ \sum_{y \in Y} |A_y| \ \leq \ \sum_{y \preceq x_0} \lambda(r-1, k-h(y)) \ \leq \lambda(r,k).$$

To prove the claim, notice that all points in A_y are of height $k - h(y)$ since the height function in Boolean lattices is modular: $h(a \vee b) + h(a \wedge b) = h(a) + h(b)$.

Now, suppose that some set A_y does not have the property $P(r-1, k-h(y))$, i.e. that

$$(*) \qquad y' \succ z'$$

for some $y' \in A_y$ and $z' \in \theta(Z')$ where $Z' \subseteq A_y$ and $|Z'| = r$. Let X be the set of those $x \in A$ for which $\partial_y(x) \in Z'$. Take a point $x \in A$ for which $y' = \partial_y(x)$, and let $\vee$ and $\wedge$ be the join and the meet operations in the Boolean lattice $[0, x]$. Since $z' \in \theta(Z')$, we have by $(*)$ that for all $x' \neq x''$ in X,

$$\begin{aligned} x = y \vee y' \succeq y \vee z' &\succeq y \vee (\partial_y(x') \wedge \partial_y(x'')) \\ &= (y \vee \partial_y(x')) \wedge (y \vee \partial_y(x'')) \\ &= x' \wedge x''. \end{aligned}$$

Since $x_0 \wedge x = y$ for all $x \in X$, we obtain that $y \vee z' \succeq x' \wedge x''$ for all $x' \neq x''$ in $Z = \{x_0\} \cup X$, and hence the point $y \vee z'$ belongs to $\theta(Z)$. Since $Z \subseteq A$, $|Z| = |X| + 1 = r + 1$ and A has the property $P(r, k)$, the situation $x \succ y \vee z'$ is impossible. Hence, $x = y \vee z'$. On the other hand, (5) implies that $y \wedge z' \preceq y \wedge y' = 0$. Thus, the point z' is also a complement of y in the interval $[0, x]$, and hence $z' = \partial_y(x) = y'$, a contradiction with $(*)$. This completes the proof of the claim, and thus, the proof of 7.2. $\square$