On Small Size Approximation Models

Alexander A. Razborov

Steklov Mathematical Institute, Vavilova 42, 117966, GSP-1, Moscow, Russia

Summary. In this paper we continue the study of the method of approximations in Boolean complexity. We introduce a framework which naturally generalizes previously known ones. The main result says that in this framework there exist approximation models providing in principle exponential lower bounds for almost all Boolean functions, and such that the number of testing functionals *is only singly exponential in the number of variables.*

1. Introduction

Proving superpolynomial lower bounds on the complexity of explicitly given Boolean functions is one of the most challenging tasks of the modern complexity theory. Its importance stems from the fact that such bounds could be easily translated into similar bounds for Turing models and, thus, would lead to resolving central open questions in Complexity Theory like $P \overset{?}{=} NP$ or $NC \overset{?}{=} P$.

At the moment, however, we have succeeded in proving desired bounds only for rather restrictive models. A substantial part of these bounds was obtained via a general scheme originally proposed in [16, 17] and called afterwards *the method of approximations*:

- on the monotone circuit size – [16, 17, 14, 1, 12, 15];
- for bounded-depth circuits with modular gates – [18, 11, 2];
- for switching-and-rectifier networks (= nondeterministic branching programs) – [19];
- for $\oplus$-branching programs (see [7] for definitions) – [6].

The reader willing to learn more about these and related results or about the general perspective of the field is referred to the survey paper [3].

Concrete approximation models appeared in the literature can be naturally subdivided into two large groups.

Models from the first group use inputs of the original function as their error tests. Such are models from [16, 17, 14, 1, 12, 15, 18, 11, 2]. We will call the method based on models of this kind the *pure approximation method.*

Other models use as error tests specially designed functionals, every functional being attached to a single input. These models were studied, and sometimes actually used in [9, 19, 4, 5, 6, 8]. See [13] for an extended survey; following this source, we will call the corresponding method the *fusion method.* The same word "fusion" will be also used for functionals and models.

The most interesting question is, of course, to which extent the approximation method might be useful in proving lower bounds for unrestricted circuits. To that end, it was shown in [9] that the pure approximations can not prove bounds greater than $O(n^2)$ or, more precisely, $O(n \cdot n_0)$, where $n_0 \leq n$ is the number of essential variables of our function. Since every fusion model with N functionals can always be considered as a pure approximation model with $n_0 := n$ and $n := n + \lceil \log_2 N \rceil$ (see [9, Claim 2.5]), it follows that lower bounds provable by any such fusion model never exceed $O(n^2 + n \log N)$.

R. L. Graham et al. (eds.), *The Mathematics of Paul Erdös I*

On the other hand, in [9] for every function f a fusion model was exhibited which, at least in principle, provides tight, up to a polynomial, lower bounds on the circuit size of f. The number of fusing functionals involved in that model was triply exponential in n, and it was also remarked in [9] that it can be decreased to doubly exponential. The resulting model, however, is somewhat artificial.

More natural fusion model with the number of fusing functionals being only doubly exponential in n has been found in [5]. Their model is universal for nondeterministic circuit size, hence it still can prove exponential lower bounds for almost all Boolean functions. Note that, in view of the above-mentioned limitation $O(n^2 + n \log N)$, this is roughly optimal for fusion models.

In this paper we study the question whether there exists a natural version of the method of approximations in which proving exponential lower bounds is possible (again, in principle) with the number of fusion functionals being only singly exponential in the number of variables. We indeed find such a framework generalizing both pure approximations and fusion models. In fact, our framework is obtained by cleaning the underlying idea of approximations from the prejudice of attaching error tests to particular input strings which is characteristic for previous models.

More exactly, we show that for every integer-valued function $t = t(n)$ in the range $n \leq t(n) \leq \frac{2^n}{3n}$ there exists an approximation model $\mathfrak{M}$ (in our framework) with $O(t^3 \log^2 t)$ error tests such that for almost all Boolean functions f, $\rho(f, \mathfrak{M}) \geq$ t, where $\rho(f, \mathfrak{M})$ is the distance between f and $\mathfrak{M}$ (Theorem 3.1).

The main motivation for this work comes from [10], where I put forward the thesis that the right theory capturing the kind of machinery existing in Boolean complexity at the moment is the second-order system V_1^1. This system can freely talk of those approximation models in which the number of error tests is bounded by $2^{O(n)}$ (and thus error tests can be represented by first order objects). Hence, unlike previous models, the models considered in this paper are within the reach of V_1^1 in terms of size. It should be noted, however, that gaining in size we lose in the constractibility. Indeed, our proof heavily relies upon Erdös probabilistic argument, and in order to carry it over in V_1^1 we need an explicit construction.

2. Definition of approximation models

Throughout the paper, F_n stands for the set of all Boolean functions in n variables. Let $P_n \rightleftharpoons \{x_1, \ldots, x_n, \neg x_1, \ldots, \neg x_n\} \subseteq F_n$ be the set of input functions.

Let $\mathcal{F}$ be a finite set of arbitrary nature. We define an *approximation model* $\mathfrak{M}$ as a subset $\mathfrak{M} \subseteq \mathfrak{F}_n \times \mathcal{P}(\mathcal{F})$ such that

$$P_n \times \{\emptyset\} \subseteq \mathfrak{M} \tag{2.1}$$

supplied with two binary operations $\wedge, \vee$ which are consistent with the projection onto F_n. In other words, we require

$$f(m_1 * m_2) = f(m_1) * f(m_2); \ m_1, m_2 \in \mathfrak{M}, \tag{2.2}$$

where $* \in \{\wedge, \vee\}$, and we once and for all have fixed notation $f(m)$ for denoting the projection of $m \in \mathfrak{M}$ onto the first coordinate F_n. Similarly, we will denote the projection onto $\mathcal{P}(\mathcal{F})$ by $\mathcal{F}(m)$ so that $m = \langle f(m), \mathcal{F}(m) \rangle$.

Now we give a set of definitions which is routine for the method of approximations. Namely, let

$$\delta_*(m_1,m_2) \rightleftharpoons \mathcal{F}(m_1 * m_2)\setminus(\mathcal{F}(m_1)\cup\mathcal{F}(m_2))\,;\ m_1,m_2\in\mathfrak{M},$$

$$\Delta \rightleftharpoons \{\delta_*(m_1,m_2)\mid *\in\{\wedge,\vee\};\ m_1,m_2\in\mathfrak{M}\},$$

$$\rho(f,\mathfrak{M}) \rightleftharpoons \min\left\{ t \,\middle|\, \exists m\in\mathfrak{M}\;\exists\delta_1,\dots,\delta_t\in\Delta\left(f(m)=f \;\&\; \mathcal{F}(m)\subseteq\bigcup_{i=1}^{t}\delta_i\right)\right\}. \tag{2.3}$$

The intuitive idea behind this is that if the real circuit computes some function f at a node u, then the approximating circuit must compute at u some $m \in \mathfrak{M}$ with $f(m) = f$ (due to (2.2)). Now, all tests $F \in \mathcal{F}(m)$ have already found "an error", that is $\mathcal{F}(m) \subseteq \bigcup_v \delta_v$, where the union is extended over all nodes v lying below u, and $\delta_v \in \Delta$ naturally corresponds to the node v. For the reader familiar with previous analogous statements, this should serve as a self-sufficient proof of the following

Theorem 2.1. *For every $f \in F_n$ and every approximation model $\mathfrak{M}$, we have $\rho(f,\mathfrak{M}) \leq \mathfrak{C}(f)$, where $C(f)$ is the minimal possible size of a circuit over the basis $\wedge, \vee$ with inputs from P_n computing f.*

We conclude this section by showing that our new framework generalizes both pure approximations and the fusion method.

Example 2.1. Let $(\mathcal{M}, \bar{\wedge}, \bar{\vee})$ be a legitimate model [9, Section 2]. Here $P_n \subseteq \mathcal{M} \subseteq F_n$ and $\bar{\wedge}, \bar{\vee}$ are arbitrary binary operations on $\mathcal{M}$. Recall that for $\bar{g}, \bar{h} \in \mathcal{M}$ and $* \in \{\wedge, \vee\}$, the subsets $\delta_*^+(\bar{g},\bar{h}), \delta_*^-(\bar{g},\bar{h})$ of $\{0,1\}^n$ are defined as follows:

$$\begin{aligned}
\delta_*^+(\bar{g},\bar{h}) &\rightleftharpoons (\bar{g} * \bar{h}) \setminus (\bar{g}\,\bar{*}\,\bar{h}),\\
\delta_*^-(\bar{g},\bar{h}) &\rightleftharpoons (\bar{g}\,\bar{*}\,\bar{h}) \setminus (\bar{g} * \bar{h})
\end{aligned}$$

(we identify a Boolean function with its set of ones). For $f \in F_n$, the *distance* $\rho(f,\mathcal{M})$ is the minimal t for which there exist $\bar{f}, \bar{g}_1, \dots, \bar{g}_t, \bar{h}_1, \dots, \bar{h}_t \in \mathcal{M}$ and $*_1, \dots, *_t \in \{\wedge, \vee\}$ such that

$$\begin{aligned}
f \setminus \bar{f} &\subseteq \bigcup_{i=1}^{t} \delta_{*_i}^+(\bar{g}_i,\bar{h}_i),\\
\bar{f} \setminus f &\subseteq \bigcup_{i=1}^{t} \delta_{*_i}^-(\bar{g}_i,\bar{h}_i).
\end{aligned}$$

The quantity $\rho(f,\mathcal{M})$ provides a lower bound on the circuit size of f.

Take now two disjoint copies B_+^n, B_-^n of $\{0,1\}^n$, and let $\mathcal{F} \rightleftharpoons B_+^n \cup B_-^n$. Consider the product $F_n \times \mathcal{M}$ of two $\{\wedge,\vee\}$-algebras, and embed it into $F_n \times \mathcal{P}(\mathcal{F}) \approx F_n \times \mathcal{P}(B_+^n) \times \mathcal{P}(B_-^n)$ as follows:

$$\begin{aligned}
\pi : F_n \times \mathcal{M} &\longrightarrow F_n \times \mathcal{P}(B_+^n) \times \mathcal{P}(B_-^n),\\
< f, \bar{f} > &\longmapsto < f,\ f \setminus \bar{f},\ \bar{f} \setminus f > .
\end{aligned}$$

We let $\mathfrak{M} \rightleftharpoons \mathrm{im}(\pi)$ and endow $\mathfrak{M}$ with the structure of $\{\wedge,\vee\}$-algebra induced from $F_n \times \mathcal{M}$. Note that (2.1) is implied by $P_n \subseteq \mathcal{M}$.

Assume that $m_1, m_2 \in \mathfrak{M}$; $m_1 = \pi(g,\bar{g})$, $m_2 = \pi(h,\bar{h})$. Representing $\delta_*(m_1,m_2)$ in the form $\delta_*^+(m_1,m_2) \cup \delta_*^-(m_1,m_2)$, where $\delta_*^\circ(m_1,m_2) \subseteq B_\circ^n$, we have:

$$\begin{aligned}\delta_*^+(m_1,m_2) &= \left((g*h)\backslash(\bar{g}\bar{*}\bar{h})\right)\backslash\left(g\backslash\bar{g}\cup h\backslash\bar{h}\right) = \left((g*h)\backslash(g\backslash\bar{g}\cup h\backslash\bar{h})\right)\backslash(\bar{g}\bar{*}\bar{h})\\ &\subseteq (\bar{g}*\bar{h})\backslash(\bar{g}\bar{*}\bar{h}) = \delta_*^+(\bar{g},\bar{h}),\end{aligned}$$

and similarly for $\delta_*^-(m_1,m_2)$. Noting that the condition

$$\exists m\in\mathfrak{M}\ \exists\delta_1,\ldots,\delta_{\mathfrak{t}}\in\Delta\left(\mathfrak{f}(\mathfrak{m})=\mathfrak{f}\ \&\ \mathcal{F}(\mathfrak{m})\subseteq\bigcup_{\mathfrak{i}=1}^{\mathfrak{t}}\delta_{\mathfrak{i}}\right)\tag{2.4}$$

from (2.3) can be rewritten in the form

$$\exists\bar{f}\in\mathcal{M}\ \exists *_i\in\{\wedge,\vee\}\ \exists m_i^{(1)},m_i^{(2)}\in\mathfrak{M}$$
$$\left(f\backslash\bar{f}\subseteq\bigcup_{i=1}^{t}\delta_{*_i}^+\left(m_i^{(1)},m_i^{(2)}\right)\ \&\ \bar{f}\backslash f\subseteq\bigcup_{i=1}^{t}\delta_{*_i}^-\left(m_i^{(1)};m_i^{(2)}\right)\right),$$

we immediately see that $\rho(f,\mathcal{M})\le\rho(f,\mathfrak{M})$. In other words, every legitimate model in the sense of [9] can be simulated in our framework.

Example 2.2. Let us now turn to the fusion method. In fact, we might first apply the construction from [9] to get a pure approximation model, and then the construction from Example 2.1. Things become much more transparent, however, if we combine the two steps into one. Recall some necessary definitions [9, 13].

Let $f\in F_n$ be a *fixed* function, $U\rightleftharpoons f^{-1}(0)$, and $V\rightleftharpoons f^{-1}(1)$. Let $\Omega_f\subseteq\{0,1\}^{\mathcal{P}(U)}$ consist of those functionals on $\mathcal{P}(U)$ which satisfy the following two conditions:

1. F is monotone,
2. there exists an (uniquely determined) $z(F)\in V$ such that for all $x_i^\epsilon\in P_n$,

$$F(x_i^\epsilon|_U)=x_i^\epsilon(z(F)).$$

Here, as usual, $x_i^1\rightleftharpoons x_i$ and $x_i^0\rightleftharpoons(\neg x_i)$.

Note that these two conditions imply $F(\emptyset)=0$ and $F(U)=1$.

For $\bar{g},\bar{h}\in\{0,1\}^U$ we say that the pair $(\bar{g},\bar{h})$ *covers* $F\in\Omega_f$ if $F(\bar{g})=F(\bar{h})=1$ and $F(\bar{g}\wedge\bar{h})=0$. The minimal number of pairs needed to cover the whole Ω_f is denoted by $\rho(f)$ and provides a lower bound on the circuit size of f which is tight up to a polynomial.

Define now the mapping

$$\begin{aligned}\pi:F_n &\longrightarrow \mathcal{P}(\Omega_f),\\ g &\longmapsto \{F\,|\,g(z(F))=1\ \&\ F(g|_U)=0\}.\end{aligned}$$

Note that $\pi(g)=\emptyset$ when $g\in P_n$, and $\pi(f)=\Omega_f$.

Let $\mathcal{F}\rightleftharpoons\Omega_f$. We take the diagonal mapping $\theta:F_n\longrightarrow F_n\times\mathcal{P}(\Omega_f)$; $g\longmapsto<g,\pi(g)>$, denote $\mathfrak{M}\rightleftharpoons\mathrm{im}(\theta)$ and endow $\mathfrak{M}$ with the induced structure of $\{\wedge,\vee\}$-algebra. (2.1) is implied by the remark above.

Now, $\delta_\vee(\theta(g),\theta(h))=\emptyset$ due to the monotonicity of every $F\in\Omega_f$. If $F\in\delta_\wedge(\theta(g),\theta(h))$ then $g(z(F))=h(z(F))=1$ and $F((g\wedge h)|_U)=0$. Since $F\not\in\pi(g)$ and $F\not\in\pi(h)$, we have $F(g|_U)=F(h|_U)=1$. Hence the pair $(g|_U,h|_U)$ covers F. As (2.4) in our case simplifies to $\exists\delta_1,\ldots,\delta_t\in\Delta\left(\mathcal{F}\subseteq\bigcup_{i=1}^t\delta_i\right)$, we see that $\rho(f)\le\rho(f,\mathfrak{M})$.

Another version of the fusion method using $GF(2)$-affine functionals instead of monotone functionals was proposed in [5, 6]. It can also be embedded into our framework if we consider approximation models over the basis $\{\wedge,\oplus\}$ rather than over $\{\wedge,\vee\}$.

3. Main result

In this section we prove the following:

Theorem 3.1. *Let $t = t(n)$ be an integer-valued function in the range $n \leq t(n) \leq \frac{2^n}{3n}$. Then there exists an approximation model $\mathfrak{M} \subseteq \mathfrak{F}_{\mathfrak{n}} \times \mathcal{P}(\mathcal{F})$, where $|\mathcal{F}| \leq O(t^3 \log^2 t)$, such that $\rho(f, \mathfrak{M}) \geq \mathrm{t}(\mathrm{n})$ for almost all functions $f \in F_n$.*

Proof. Let $\ell \rightleftharpoons \lfloor 20t^3 \ln^2 t \rfloor$ and $S \rightleftharpoons \left(\begin{array}{c} \ell \\ t \end{array} \right)$. Fix a set $\mathcal{F}$ of cardinality ℓ.

For a subset $\mathfrak{M}$ of $F_n \times \mathcal{P}(\mathcal{F})$ and $\mathcal{F}_0 \subseteq \mathcal{F}$, we let

$$\mathfrak{M}(\mathcal{F}_{\mathrm{o}}) \rightleftharpoons \{ \mathrm{m} \in \mathfrak{M} \mid \mathcal{F}(\mathrm{m}) \subseteq \mathcal{F}_{\mathrm{o}} \}$$

and

$$w_{\mathfrak{M}}(\mathcal{F}_0) \rightleftharpoons |\mathfrak{M}(\mathcal{F}_{\mathrm{o}})| \, .$$

Let also

$$w_{\mathfrak{M}} \rightleftharpoons \ln \left(\mathbf{E}\left[e^{w_{\mathfrak{M}}(\mathcal{F}_0)} \right] \right) ,$$

where $\mathcal{F}_0 \subseteq \mathcal{F}$ is a random subset of cardinality t.

We are going to define by induction on k a sequence

$$P_n \times \{\emptyset\} = \mathfrak{M}_{\mathrm{o}} \subseteq \mathfrak{M}_1 \subseteq \ldots \subseteq \mathfrak{M}_{\mathfrak{k}} \subseteq \ldots \subseteq \mathfrak{F}_{\mathfrak{n}} \times \mathcal{P}(\mathcal{F}) \tag{3.1}$$

along with binary operations $\wedge_k, \vee_k : \mathfrak{M}_{\mathfrak{k}-1} \times \mathfrak{M}_{\mathfrak{k}-1} \longrightarrow \mathfrak{M}_{\mathfrak{k}}$ maintaining the following properties:

1. *if $k \leq k'$ then $\wedge_{k'}|_{\mathfrak{M}_{\mathfrak{k}-1} \times \mathfrak{M}_{\mathfrak{k}-1}} = \wedge_k$ and $\vee_{k'}|_{\mathfrak{M}_{\mathfrak{k}-1} \times \mathfrak{M}_{\mathfrak{k}-1}} = \vee_k$;*
2. *$f\,(m_1 *_k m_2) = f(m_1) * f(m_2)$ for $m_1, m_2 \in \mathfrak{M}_{\mathfrak{k}-1}$;*
3. *for every $m_1, m_2 \in \mathfrak{M}_{\mathfrak{k}-1}$ and $* \in \{\wedge, \vee\}$,*
$$|\mathcal{F}(m_1 *_k m_2) \backslash \left(\mathcal{F}(m_1) \cup \mathcal{F}(m_2) \right)| \leq \mathbf{1};$$
4. *for every $m \in \mathfrak{M}_{\mathfrak{k}} \backslash \mathfrak{M}_{\mathfrak{k}-1}$, $|\mathcal{F}(m)| \geq \min(k, \ell)$;*
5. *$w_{\mathfrak{M}_{\mathfrak{k}}} \leq 2(n + k)$.*

Base $k = 0$ is obvious.

Inductive step. Assume that $\mathfrak{M}_{\mathrm{o}}, \mathfrak{M}_1, \ldots, \mathfrak{M}_{\mathfrak{k}-1}, \mathfrak{M}_{\mathfrak{k}}$ and $\wedge_k, \vee_k : \mathfrak{M}_{\mathfrak{k}-1} \times \mathfrak{M}_{\mathfrak{k}-1} \longrightarrow \mathfrak{M}_{\mathfrak{k}}$ are already defined. Then we randomly extend $\wedge_k, \vee_k$ to $\wedge_{\mathbf{k+1}}, \vee_{\mathbf{k+1}} : \mathfrak{M}_{\mathfrak{k}} \times \mathfrak{M}_{\mathfrak{k}} \longrightarrow \mathfrak{F}_{\mathfrak{n}} \times \mathcal{P}(\mathcal{F})$ as follows. For $(m_1, m_2) \in (\mathfrak{M}_{\mathfrak{k}} \times \mathfrak{M}_{\mathfrak{k}}) \setminus (\mathfrak{M}_{\mathfrak{k}-1} \times \mathfrak{M}_{\mathfrak{k}-1})$ we let

$$m_1 *_{\mathbf{k+1}} m_2 \rightleftharpoons \left(f(m_1) * f(m_2),\ \mathcal{F}(m_1) \cup \mathcal{F}(m_2) \cup \{ \mathbf{F}_*(m_1, m_2) \} \right),$$

where $\mathbf{F}_*(m_1, m_2)$ is chosen at random from $\mathcal{F} \backslash \left(\mathcal{F}(m_1) \cup \mathcal{F}(m_2) \right)$ if $\mathcal{F}(m_1) \cup \mathcal{F}(m_2) \neq \mathcal{F}$ and arbitrarily otherwise. All $\mathbf{F}_*(m_1, m_2)$ are assumed to be independent.

After this we let

$$\mathfrak{M}_{\mathfrak{k}+1} \rightleftharpoons \mathfrak{M}_{\mathfrak{k}} \cup \mathrm{im}(\wedge_{\mathbf{k+1}}) \cup \mathrm{im}(\vee_{\mathbf{k+1}}). \tag{3.2}$$

Properties 1-3 readily follow from definitions, and 4 follows from the inductive assumption. We are going to show that 5 (with $k := k + 1$) also takes place with a non-zero probability.

We may assume that $k+1 \leq t$ since otherwise property 5 follows from 4 and the inductive assumption. For simplicity we will abbreviate $w_{\mathfrak{M}_i}(\mathcal{F}_0)$ and $w_{\mathfrak{M}_i}$ to $w_i(\mathcal{F}_0), w_i$ respectively.

Let us first fix some $\mathcal{F}_0 \subseteq \mathcal{F}$ of cardinality t and estimate $\mathbf{E}\left[e^{\mathbf{w}_{\mathbf{k}+\mathbf{1}}(\mathcal{F}_0)}\right]$ for this particular $\mathcal{F}_0$. Denote the set

$$\{(m_1, m_2, *) \mid (m_1, m_2) \in (\mathfrak{M}_{\mathfrak{k}}(\mathcal{F}_0) \times \mathfrak{M}_{\mathfrak{k}}(\mathcal{F}_0)) \setminus (\mathfrak{M}_{\mathfrak{k}-1} \times \mathfrak{M}_{\mathfrak{k}-1}),\ * \in \{\wedge, \vee\}\}$$

by A. Then $|A| \leq 2w_k^2(\mathcal{F}_0)$ and

$$\mathfrak{M}_{\mathfrak{k}+1}(\mathcal{F}_0) = \mathfrak{M}_{\mathfrak{k}}(\mathcal{F}_0) \cup \{\mathfrak{m}_1 *_{\mathbf{k}+\mathbf{1}} \mathfrak{m}_2 \mid (\mathfrak{m}_1, \mathfrak{m}_2, *) \in \mathfrak{A} \ \&\ \mathbf{F}_*(\mathfrak{m}_1, \mathfrak{m}_2) \in \mathcal{F}_0\}.$$

Hence

$$\mathbf{w}_{\mathbf{k}+\mathbf{1}}(\mathcal{F}_0) \leq w_k(\mathcal{F}_0) + \sum_{(m_1, m_2, *) \in A} \xi_*(m_1, m_2), \tag{3.3}$$

where $\xi_*(m_1, m_2)$ is the indicator function of the event $\mathbf{F}_*(m_1, m_2) \in \mathcal{F}_0$.

All $\xi_*(m_1, m_2)$ are, however, independent. Therefore (3.3) gives us the estimate

$$\begin{aligned} \mathbf{E}\left[e^{\mathbf{w}_{\mathbf{k}+\mathbf{1}}(\mathcal{F}_0)}\right] &\leq e^{w_k(\mathcal{F}_0)} \cdot \prod_{(m_1, m_2, *) \in A} \mathbf{E}\left[e^{\xi_*(m_1, m_2)}\right] \leq e^{w_k(\mathcal{F}_0)} \cdot \left(1 + \frac{t(e-1)}{\ell}\right)^{2w_k^2(\mathcal{F}} \\ &\leq e^{w_k(\mathcal{F}_0) + \frac{4t}{\ell} w_k^2(\mathcal{F}_0)}. \end{aligned}$$

Averaging this inequality over $\mathcal{F}_0$, we have

$$\mathbf{E}\left[e^{\mathbf{w}_{\mathbf{k}+\mathbf{1}}(\mathcal{F}_0)}\right] \leq \mathbf{E}\left[e^{w_k(\mathcal{F}_0) + \frac{4t}{\ell} w_k^2(\mathcal{F}_0)}\right].$$

Now we fix a particular choice of $\mathfrak{M}_{\mathfrak{k}+1}$ with the property

$$e^{w_{k+1}} = \mathbf{E}\left[e^{w_{k+1}(\mathcal{F}_0)}\right] \leq \mathbf{E}\left[e^{w_k(\mathcal{F}_0) + \frac{4t}{\ell} w_k^2(\mathcal{F}_0)}\right]. \tag{3.4}$$

We will show that this implies the desired inequality $w_{k+1} \leq 2n + 2k + 2$ if $k+1 \leq t$.

Let us denote $e^{w_k(\mathcal{F}_0)}$ by $\theta_k(\mathcal{F}_0)$. Then the inductive assumption can be rewritten in the form

$$\mathbf{E}[\theta_k(\mathcal{F}_0)] \leq e^{2(n+k)}, \tag{3.5}$$

and (3.4) – in the form

$$e^{w_{k+1}} \leq \mathbf{E}\left[\theta_k(\mathcal{F}_0) \cdot e^{\frac{4t}{\ell} \ln^2 \theta_k(\mathcal{F}_0)}\right]. \tag{3.6}$$

The function $x \cdot a^{\ln^2 x}$, where $a = e^{\frac{4t}{\ell}}$ is, however, convex on $[1, \infty)$. Hence, under the condition (3.5), the right-hand side of (3.6) achieves its maximal value when $\theta_k(\mathcal{F}_0)$ takes on $(S-1)$ times the value 1, and the remaining time – the value $S \cdot e^{2(n+k)} - S + 1 \leq S \cdot e^{2(n+k)}$. This gives us the estimate

$$e^{w_{k+1}} \leq \frac{S-1}{S} + e^{2(n+k)} \cdot e^{\frac{4t}{\ell}(\ln S + 2(n+k))^2}.$$

Finally,

$$w_{k+1} \leq \ln\left(1 + e^{2(n+k)} \cdot e^{\frac{4t}{\ell}(\ln S + 2(n+k))^2}\right) \leq 1 + 2(n+k) + \frac{4t}{\ell}(\ln S + 2(n+k))^2,$$

and it is easy to see that $\frac{4t}{\ell}(\ln S + 2(n+k))^2 \leq \frac{4t}{\ell}(\ln S + 2(n+t))^2 \leq 1$ due to our choice of parameters.

When we have the desired sequence (3.1), the rest is easy. We let $\mathfrak{M} \rightleftharpoons \bigcup_{\mathfrak{t}\geq o}\mathfrak{M}_{\mathfrak{t}}$. Property 1 ensures that we can glue together the partial operations $\wedge_k, \vee_k$ to endow $\mathfrak{M}$ with a natural structure of $\{\wedge, \vee\}$-algebra. Property 2 gives us (2.2), and property 3 lets us to conclude that $\forall\delta \in \Delta\ |\delta| \leq 1$. Hence, if $\rho(f, \mathfrak{M}) \leq \mathfrak{t}$ for some $f \in F_n$ then $\exists m \in \mathfrak{M}\,(\mathfrak{f}(\mathfrak{m}) = \mathfrak{f}\ \&\ |\mathcal{F}(\mathfrak{m})| \leq \mathfrak{t})$. Due to property 4, we may replace here $\mathfrak{M}$ by $\mathfrak{M}_{\mathfrak{t}}$.

However, the total number of $m \in \mathfrak{M}_{\mathfrak{t}}$ such that $|\mathcal{F}(m)| \leq t$ does not exceed

$$\sum_{\substack{\mathcal{F}_0 \subseteq \mathcal{F} \\ |\mathcal{F}_0| = t}} w_t(\mathcal{F}_0).$$

Since $e^{w_t(\mathcal{F}_0)} \leq S \cdot e^{w_t} \leq S \cdot e^{2(n+t)}$ by property 5, we have that this number is bounded from above by $S(\ln S + 2(n+t)) \leq o\left(2^{2^n}\right)$. The theorem follows.

4. Conclusion

The most interesting open question is, of course, whether the proof of Theorem 3.1 can be made constructive. The connection with V_1^1 mentioned in Introduction suggests the following specific form of this question.

Can we find a good approximation model $\mathfrak{M}$ such that, as a subset of $F_n \times \mathcal{P}(\mathcal{F})$, it is recognizable in polynomial time, and the operations $\wedge, \vee$ are polynomially time computable? Here "good" means "such that $\rho(f_n, \mathfrak{M}) \geq \mathrm{n}^{\omega(1)}$ for some choice of $f_n \in F_n$", and "polynomial" means "polynomial in 2^n".

References

1. N. Alon and R. Boppana. The monotone circuit complexity of Boolean functions. *Combinatorica*, 7(1):1–22, 1987.
2. D. A. Barrington. A note on a theorem of Razborov. Technical report, University of Massachusetts, 1986.
3. R. B. Boppana and M. Sipser. The complexity of finite functions. In Jan van Leeuwen, editor, *Handbook of Theoretical Computer Science, vol. A (Algorithms and Complexity)*, chapter 14, pages 757–804. Elsevier Science Publishers B.V. and The MIT Press, 1990.
4. M. Karchmer. On proving lower bounds for circuit size. In *Proceedings of the 8th Structure in Complexity Theory Annual Conference*, pages 112–118, 1993.
5. M. Karchmer and A. Wigderson. Characterizing non-deterministic circuit size. In *Proceedings of the 25th Annual ACM Symposium on the Theory of Computing*, pages 532–540, 1993.
6. M. Karchmer and A. Wigderson. On span programs. In *Proceedings of the 8th Structure in Complexity Theory Annual Conference*, pages 102–111, 1993.
7. C. Meinel. *Modified Branching Programs and Their Computational Power, Lecture Notes in Computer Science,* 370. Springer-Verlag, New York/Berlin, 1989.
8. K. Nakayama and A. Maruoka. Loop circuits and their relation to Razborov's approximation model. Manuscript, 1992.

9. A. Razborov. On the method of approximation. In *Proceedings of the 21st ACM Symposium on Theory of Computing*, pages 167–176, 1989.
10. A. Razborov. Bounded Arithmetic and lower bounds in Boolean complexity. To appear in the volume *Feasible Mathematics II*, 1993.
11. R. Smolensky. Algebraic methods in the theory of lower bounds for Boolean circuit complexity. In *Proceedings of the 19th ACM Symposium on Theory of Computing*, pages 77–82, 1987.
12. É. Tardos. The gap between monotone and nonmonotone circuit complexity is exponential. *Combinatorica*, 8:141–142, 1988.
13. A. Wigderson. The fusion method for lower bounds in circuit complexity. In *Combinatorics, Paul Erdos is Eighty*. 1993.
14. А. Е. Андреев. Об одном методе получения нижних оценок сложности индивидуальных монотонных функций. *ДАН СССР*, 282(5):1033–1037, 1985. A.E. Andreev, On a method for obtaining lower bounds for the complexity of individual monotone functions. *Soviet Math. Dokl.* 31(3):530-534, 1985.
15. А. Е. Андреев. Об одном методе получения эффективных нижних оценок монотонной сложности. *Алгебра й логика*, 26(1):3–21, 1987. A.E. Andreev, On one method of obtaining effective lower bounds of monotone complexity. *Algebra i logika*, 26(1):3-21, 1987. In Russian.
16. А. А. Разборов. Нижние оценки монотонной сложности некоторых булевых функций. *ДАН СССР*, 281(4):798–801, 1985. A. A. Razborov, Lower bounds for the monotone complexity of some Boolean functions, *Soviet Math. Dokl.*, 31:354-357, 1985.
17. А. А. Разборов. Нижние оценки монотонной сложности логического перманента. *Матем. Зам.*, 37(6):887–900, 1985. A. A. Razborov, Lower bounds of monotone complexity of the logical permanent function, *Mathem. Notes of the Academy of Sci. of the USSR*, 37:485-493, 1985.
18. А. А. Разборов. Нижние оценки размера схем ограниченной глубины в полном базисе, содержащем функцию логического сложения. *Матем. Зам.*, 41(4):598–607, 1987. A. A. Razborov, Lower bounds on the size of bounded-depth networks over a complete basis with logical addition, *Mathem. Notes of the Academy of Sci. of the USSR*, 41(4):333-338, 1987.
19. А. А. Разборов. Нижние оценки сложности реализации симметрических булевых функций контактно-вентильными схемами. *Матем. Зам.*, 48(6):79–91, 1990. A. A. Razborov, Lower bounds on the size of switching-and-rectifier networks for symmetric Boolean functions, *Mathem. Notes of the Academy of Sci. of the USSR*.