Circuit Lower Bounds à la Kolmogorov

Lance Fortnow\textsuperscript{1}  
\texttt{fortnow@cs.uchicago.edu}

Sophie Laplante\textsuperscript{1}  
\texttt{sophie@cs.uchicago.edu}

University of Chicago  
Department of Computer Science  
1100 East 58th Street  
Chicago, IL 60637

\textsuperscript{1}Partially supported by NSF grant CCR 92–53582
Proposed running header: Circuit Lower Bounds à la Kolmogorov

Send proofs to:

Sophie Laplante
The University of Chicago
Department of Computer Science
1100 E 58th Street, Ryerson 152
Chicago, IL 60637-1504
e-mail: sophie@cs.uchicago.edu
Abstract

In a recent paper, Razborov [Raz93] gave a new combinatorial proof of Håstad’s switching lemma [Hås89], eliminating the probabilistic argument altogether. In this paper we adapt his proof and propose a Kolmogorov complexity-style switching lemma, from which we derive the probabilistic switching lemma as well as a Kolmogorov complexity-style proof of circuit lower bounds for parity.
Håstad’s switching lemma [Hås89] is a prime example of the so-called “restriction” or “bottom-up” method used to find lower bounds in circuit complexity. This approach consists of considering circuits from the bottom level (inputs), and showing that restricting the function by fixing some of the inputs does not always force the function to zero or one.

In a recent paper [Raz93], Razborov presented a new proof of Håstad’s lemma, using a simpler counting argument instead of the usual probabilistic techniques. (See also the paper by Beame [Bea94] for a presentation of this proof, as well as extensions of the lemma and further applications.) The main part of the counting argument can be expressed in terms of elementary Kolmogorov complexity. We present a Kolmogorov complexity-style switching lemma, from which we derive both a more conventional probabilistic switching lemma and a Kolmogorov complexity-style proof of circuit lower bounds for the parity function.

1 Preliminaries

In order to talk about boolean functions and circuits, we start by introducing a few basic terms and definitions.

**Definition 1.1** A $t$-disjunct is a disjunction of $t$ variables or negations of variables. Similarly, a $t$-conjunct is a conjunction of $t$ variables or negations of variables.

**Definition 1.2** A boolean function is $t$-closed if it is a conjunction of $t$-disjuncts.

**Definition 1.3** A boolean function is $t$-open if it is a disjunction of $t$-conjuncts.

**Definition 1.4** A restriction $\rho$ is a function from a set of variables to the set $\{0, 1, \ast\}$. Given a boolean function $f$, $f|_\rho$ is the restriction of $f$ in the natural way, i.e., $x_i$ is free if $\rho(x_i) = \ast$ and $x_i$ takes on the value $\rho(x_i)$ otherwise.
The *domain* of a restriction $\rho$ is the set of variables mapped to 0 or 1 by $\rho$, and is written $\text{dom}(\rho)$. The juxtaposition of restrictions $\alpha\beta$ with disjoint domains is the restriction that takes on the value of $\alpha$ on $\text{dom}(\alpha)$ and that of $\beta$ everywhere else.

**Definition 1.5** A term (also called implicant) of a boolean function $f$ is restriction $\tau$ such that $f|_{\tau}$ is the constant function 1. A minterm (also called prime implicant) is a restriction such that no proper subset of the variables set by the restriction forms a term.

Let $\mathcal{R}^l$ be the set of restrictions that leave $l$ variables free. In the course of the proof, we will use the fact that the cardinality of $\mathcal{R}^l$ is

$$|\mathcal{R}^l| = \binom{n}{l}2^{n-l}.$$

## 2 Elements of Kolmogorov Complexity

The proof of the switching lemma requires only the most basic facts and definitions from Kolmogorov complexity. We present them here, following the notation of Li and Vitányi [LV93].

**Definition 2.1** Fix a universal Turing Machine $\Phi$. For any pair of strings $x, y \in \{0,1\}^*$, the Kolmogorov complexity of $x$ relative to $y$ is defined as

$$C_\Phi(x|y) = \text{Min}\{|p| : \Phi(p,y) = x\}.$$

For the remainder of this discussion, we fix $\Phi$ and write $C(x|y)$ instead of $C_\Phi(x|y)$.

**Proposition 2.1** Let $A \subseteq \{0,1\}^* \times \{0,1\}^*$ be a recursively enumerable set, and define $X_y = \{x \in \{0,1\}^* : (x,y) \in A\}$ for some $y \in \{0,1\}^*$. If $X_y$ is finite, then for every $x \in X_y$ $C(x|y) \leq \log |X_y| + c_A$ for a constant $c_A$ depending only on $A$. 

2
Proposition 2.2 For every \( y \in \{0,1\}^* \) and every \( A \subseteq \{0,1\}^* \) with \( |A| = m \), there is a string \( x \in A \) such that \( C(x|y) \geq \log m \).

In fact, the following more general statement also holds.

Proposition 2.3 For any positive integer \( c \), any \( y \), and any \( A \subseteq \{0,1\}^* \) with \( |A| = m \), the number of strings \( x \in A \) with \( C(x|y) \geq \log m - c \) is at least \( m(1 - 2^{-c}) \).

The reader is referred to the textbook by Li and Vitányi [LV93] for the proofs of these propositions and further background on Kolmogorov complexity.

3 A Kolmogorov switching lemma

As in Håstad’s paper [Hås89], we start by proving a switching lemma. The goal of this lemma is to provide conditions for when a “random” restriction of some of the variables of a \( t \)-closed formula \( \text{AND} \) of \( \text{ORs} \) yields an \( s \)-open formula \( \text{OR} \) of \( \text{ANDs} \). Here, the idea is to look at random restrictions from the point of view of Kolmogorov complexity: if the restriction is sufficiently random in the sense that it does not have a short representation, then the function once the restriction is applied will be \( s \)-open.

Lemma 3.1 Fix a \( t \)-closed function \( f \) on \( n \) variables, \( s < l < n, n \geq 2l - s \), and a restriction \( \rho \in \mathcal{R}^l \). If

\[
C(\rho|f,n,l,s,t) \geq \log \left( \frac{n}{l-s} \right) + n - l + s \log 8t + c
\]

then \( f|_\rho \) is \( s \)-open (where \( c \) is an absolute constant).

Proof:

We prove the contrapositive. Fix \( f, n, s, l, t \) and \( \rho \) as above and assume \( f|_\rho \) is not \( s \)-open. We will show a bound on the conditional complexity of \( \rho \) by exhibiting an extension \( \rho' \in \mathcal{R}^{l-s} \) of \( \rho \) as well as a string \( \sigma \) which together suffice to describe \( \rho \) and such that
1. $\sigma$ can be written in “blocks” as $\sigma = \sigma^{(1)} \sigma^{(2)} \cdots \sigma^{(k)}$ where each block has length $|\sigma^{(i)}| = t$;

2. Every block $\sigma^{(i)}$ for $i < k$ contains at least one non-$*$;

3. $\sigma$ has at most $s$ positions that are not $*$.

First we establish that this will suffice to get the desired upper bound on the conditional complexity of $\rho$.

The fact that $\rho'$ and $\sigma$ suffice to reconstruct $\rho$ gives us an upper bound on the complexity of $\rho$:

$$C(\rho|f, n, l, s, t) \leq C(\rho'|f, n, l, s, t) + C(\sigma|f, n, l, s, t) + c_\rho \quad (1)$$

Since $\rho' \in \mathcal{R}^{l-s}$, we get (by Proposition 2.1)

$$C(\rho'|f, n, l, s, t) \leq \log |\mathcal{R}^{l-s}| + c_{\rho'}$$

$$= \log \left( \frac{n}{l-s} \right) + n - l + s + c_{\rho'} \quad (2)$$

Furthermore, $\sigma$'s structure allows us to give an upper bound on $C(\sigma)$ as follows. Property 3 tells us that $\sigma$ contains mostly $*$s; in fact, we can think of it as at most $s + 1$ substrings of consecutive $*$s, “interrupted” by at most $s$ non-$*$ values. Properties 1 and 2 ensure that each substring of consecutive $*$s has length bounded by $2t$, so the length of each of these substrings can be encoded with $\log 2t$ bits. To encode all of $\sigma$, we encode the length of each of these (possibly empty) substrings, followed by the value of the non-$*$ position that follows it. Note that we need not encode the length of the very last substring since it is implicit given $k$ and $t$. This gives us the following upper bound on the complexity of $\sigma$:

$$C(\sigma|f, n, l, s, t) \leq s \log 2t + s + c_{\sigma}$$

$$= s \log 4t + c_{\sigma} \quad (3)$$

Combining equations (1), (2), (3), we get precisely the bound claimed, i.e.,

$$C(\rho|f, n, l, s, t) \leq \log \left( \frac{n}{l-s} \right) + n - l + s + s \log 4t + c_{\rho} + c_{\sigma} + c_{\rho'}$$
\[ \log \left( \frac{n}{l-s} \right) + n - l + s \log lt + c \]

where \( c = c_{\rho} + c_{\sigma} + c_{\rho'} + 1 \).

We now describe how to obtain \( \sigma \) and \( \rho' \).

Since \( f \) is \( t \)-closed, we can write it as the conjunction

\[ f = \bigwedge_i D_i \]

of disjuncts \( D_i \), each of which has size at most \( t \).

On the other hand, \( f|_{\rho} \) can be written as the disjunction of conjuncts

\[ f|_{\rho} = \bigvee_j C_j, \]

where each \( C_j \) corresponds to a minterm of \( f|_{\rho} \). Notice that since \( f|_{\rho} \) is assumed not to be \( s \)-open, there must be at least one minterm which sets at least \( s + 1 \) variables. Fix \( \pi \) to be one such minterm. \( \rho' \) will restrict \( s \) of the variables of \( \pi \), as described now.

First we split \( \pi \) into “subrestrictions” \( \pi_i \) by considering how its variables lie within the disjuncts \( D_j \) of \( f \) subject also to the restriction \( \rho \). To define \( \pi_i \), assume \( \pi_1, \ldots, \pi_{i-1} \) have already been defined, and that there remain variables in \( \text{dom}(\pi) \setminus (\text{dom}(\pi_1 \cdots \pi_{i-1})) \). Consider the disjuncts \( D_j \) in increasing order of subscripts. Find the least index \( j \) for which there is at least one variable which appears in \( D_j \) and \( \text{dom}(\pi) \setminus \text{dom}(\pi_1 \cdots \pi_{i-1}) \) and which, subject to \( \rho \) and \( \pi_1 \cdots \pi_{i-1} \), does not already force this clause to 1. Let \( S \) be a maximal subset for \( D_j \). Define \( \pi_i \) as

\[
\pi_i(x) = \begin{cases} 
\pi(x) & \text{if } x \in S \\
\ast & \text{otherwise}.
\end{cases}
\]

Note that such a disjunct must exist because \( \pi \) is a minterm of \( f|_{\rho} \), so \( \pi \) must force each disjunct of \( f \) subject to \( \rho \) to 1, but no subrestriction of \( \pi \) (namely \( \pi_1 \cdots \pi_{i-1} \)) will.

Let \( k \) be the least integer such that \( \pi_1 \cdots \pi_k \) sets at least \( s \) variables, and “trim” \( \pi_k \) so that \( \pi_1, \ldots, \pi_k \) sets exactly \( s \) variables.
We are now ready to define the restriction $\rho'$. Recall that the restriction $\pi$ corresponds to a minterm of $f|_\rho$, hence each $\pi_i$ is in a sense *trying* to set its associated clause $D_j$ to 1. To produce $\rho'$, we will take each restriction $\pi_i$ and change it to $\overline{\pi_i}$ so that it tries to set its associated clause $D_j$ to zero. We do this by setting each variable $x$ which appears in the clause to 0 if it appears as $x$ and to 1 if it appears as $\overline{x}$. As we will see shortly, this will be key to reconstructing $\rho$ from $\rho'$.

Let $\rho' = \rho \pi_1 \cdots \pi_k$. Notice that to recover $\rho$ from $\rho'$, we will need to isolate and remove each $\overline{\pi_i}$ from $\rho'$.

With this in mind, we will encode the restrictions $\pi_i$ *without specifying* to which clause each $\pi_i$ corresponds. We will argue later that the clauses $D_j$ corresponding to each $\pi_i$ can be recovered, so that the variables set by $\pi_i$ (and $\overline{\pi_i}$) in that clause can be stripped from $\rho'$.

The string $\sigma$ will be this encoding of each $\pi_i$. For each $i$, let $T_i$ represent the ordered set of variables that appear in the disjunct associated with $\pi_i$. Recall that $f$ is $t$-closed by our assumption, so $|T_i|$ is bounded by $t$. For each $x$ in $T_i$, we let

$$\sigma^{(i)}(x) = \begin{cases} 
\pi_i(x) & \text{if } x \in \text{dom}(\pi_i) = \text{dom}(\overline{\pi_i}) \\
\ast & \text{otherwise.}
\end{cases}$$

To complete the proof, two points remain to be argued, namely that the three claimed properties of $\sigma$ are verified, and that $\sigma$ and $\rho'$ suffice to reconstruct $\rho$.

First we check the three properties of $\sigma$.

1. Each $\sigma^{(i)}$ has length bounded by $t$ because $|T_i|$ is bounded by $t$; if $|T_i| < t$, then $\sigma^{(i)}$ can be padded with the appropriate number of $\ast$s.

2. For each $i < k$, we claim that $\sigma^{(i)}$ must contain at least one non-$\ast$. This is clear because by definition of $\pi_i$, $\text{dom}(\pi_i) \neq \emptyset$.

3. Since the restriction $\pi$ was trimmed to set exactly $s$ variables, $\sigma$ must have at most $s$ positions that are not $\ast$s.
Finally we show that \( \sigma \) and \( \rho' \) suffice to recover the original restriction \( \rho \). For each \( i \), we will proceed as follows: find the clause \( D_j \) corresponding to \( \pi_i \), then deduce \( \tilde{\pi}_i \) from \( \rho' \) and \( \sigma^{(i)} \). The important step is to see that we can find \( D_j \).

Assume that we have already recovered \( \tilde{\pi}_1, \ldots, \tilde{\pi}_{i-1} \) so that we also know \( \rho \pi_1 \cdots \pi_{i-1} \tilde{\pi}_{i-1} \cdots \tilde{\pi}_{k} \). Recall that \( \pi_i \) was defined by considering the clauses of \( f \) in increasing order, and choosing the least clause not already forced to 1 by \( \rho \pi_1 \cdots \pi_{i-1} \). It is crucial to see that the additional restriction \( \tilde{\pi}_{i-1} \cdots \tilde{\pi}_{k} \) does not make the clauses behave any differently than they would have under \( \rho \) and \( \pi_1, \ldots, \pi_{i-1} \). Therefore the least clause not forced to 1 under the currently known restriction \( \rho \pi_1 \cdots \pi_{i-1} \tilde{\pi}_{i-1} \cdots \tilde{\pi}_{k} \) is also the least clause that was chosen to correspond to \( \pi_i \).

Given the index to this clause, it suffices to find the actual variables set by \( \pi_i \) by looking at all the non-\( \ast \) positions in \( \sigma^{(i)} \); the assignments to \( \tilde{\pi}_i \) can then be read off in \( \rho' \).

\[ \square \]

4 Håstad’s lemma

Before proceeding to the proof of circuit lower bounds using the Kolmogorov switching lemma, we observe that we can derive Håstad’s switching lemma as a corollary of Lemma 3.1.

The switching lemma is at the core of the proof of lower bound theorems in circuit complexity. It provides a way reducing the height of a circuit by switching adjacent levels of AND gates and OR gates, hence its name. It does so at the price of reducing the number of free variables in the circuit. The lemma gives the probability that a random restriction will result in a switch, for given bounds on initial and final fan-in for the gates on the lower level.

The following combinatorial fact will be used in the proof.

**Proposition 4.1** For any \( s < l < n, n \geq 2l - s \),

\[
\frac{\binom{n}{i}}{\binom{n}{l-s}} \geq \left( \frac{n-l+s}{l} \right)^s
\]
Corollary 4.1: Fix a \( t \)-closed function \( f \) on \( n \) variables, \( s < l < n, n \geq 2l - s \). If \( \rho \) is chosen from the uniform distribution over the set \( \mathcal{R}^l \) then the probability that \( f|\rho \) is \( s \)-open is at least

\[
1 - 2^e \left( \frac{8tl}{n - l + s} \right)^s
\]

(where \( e \) is the constant from Lemma 3.1)

Proof: By Proposition 2.3, the number of restrictions in \( \mathcal{R}^l \) such that

\[
C(\rho|f,n,l,s,t) \geq \log \left( \frac{n}{l - s} \right) + n - l + s \log st + c
\]

is at least

\[
|R^l|(1 - 2^{\log |R^l| + (\log (\frac{n}{l-s}) + n - l + s \log st + c)}).
\]

Proposition 4.1 and simple algebraic manipulations yield the claimed result.

Our version of this lemma differs slightly from that of Håstad [Hås89], which we cite here for the sake of comparison. The probability distribution on the random restrictions is different: whereas we are choosing restrictions uniformly from \( \mathcal{R}^l \), in Håstad’s lemma, each variable (independently) is restricted with probability \( 1 - p \), where \( p = l/n \), so that the expected number of variables that are left free is \( l \). Call this distribution on the restrictions \( R_p \).

Lemma 4.1 (Håstad): If \( f \) is \( t \)-closed on \( n \) variables then the probability that a restriction \( \rho \) chosen from the distribution \( R_p \) (\( p = l/n \)) is such that \( f|\rho \) is \( s \)-open is

\[
1 - \left( \frac{\gamma tl}{n} \right)^s
\]

for a fixed constant \( \gamma < 5 \).

Although in the statement of Corollary 4.1 the constant we use is \( 8 \), with a more involved analysis of \( C(\sigma) \) it can be brought down to \( 2 \log 2e \approx 5.44 \).
5 Circuit Lower Bounds

In this section we show how to obtain circuit lower bounds similar to those of Håstad [Hås89] by applying the Kolmogorov switching lemma. The proof here is similar to Håstad’s [Hås89], but without any probability arguments.

We use the standard model for circuits with AND and OR gates, and variables and their negations at the inputs. We assume that each level has only OR gates or only AND gates. We refer the reader to Boppana and Sipser’s survey paper [BS90] for a more thorough description of the model.

As in Håstad’s paper [Hås89], we provide lower bounds for the parity function. To be precise, we are interested in parity or its negation, which is why we can safely assume that the gates closest to the inputs are all OR gates: if it is not the case, we simply take the negation of the circuit and proceed with the result.

We start by extending our Kolmogorov switching lemma from depth-2 circuits to depth-\(k\) circuits. Lemma 5.1 will provide us with sufficient conditions for the switching lemma (Lemma 3.1) to apply to all of the bottom depth-2 subcircuits. This is what will allow us to reduce the depth of the circuit by 1. In the incompressibility lemma (Lemma 5.2), we derive sufficient conditions for when an incompressible restriction is used. Finally, we follow the proof of Håstad and prove the final lower bound theorem (Theorem 5.1) by way of an auxiliary induction lemma (Lemma 5.3).

5.1 Kolmogorov switching lemma for circuits

The following lemma extends the reach of the Kolmogorov switching lemma (Lemma 3.1) from depth-2 circuits to any boolean circuit. It provides the conditions under which a single restriction will collapse all the depth-2 subcircuits at the bottom level of a circuit.

**Lemma 5.1** Fix a circuit\( C \) on \( n \) variables, with bottom fan-in bounded by \( t \). Let \( \mathcal{F} \) denote the set of \( t \)-closed functions corresponding to the subcircuits of depth 2 in \( C \).
Fix \( s < l < n, n \geq 2l - s \) and a restriction \( \rho \in \mathcal{R}^l \). If
\[
C(\rho|C, n, l, s, t) \geq \log \left( \frac{n}{l-s} \right) + n - l + s \log 8l + \log |\mathcal{F}| + c
\]
where \( c \) is the constant from Lemma 3.1, then for each \( f \in \mathcal{F} \) \( f|_\rho \) is \( s \)-open.

**Proof:** Assume there is some \( f \in \mathcal{F} \) such that \( f|_\rho \) is not \( s \)-open.

Then by Lemma 3.1, it must be the case that
\[
C(\rho|f, n, l, s, t) < \log \left( \frac{n}{l-s} \right) + n - l + s \log 8l + c
\]
Furthermore,
\[
C(\rho C, n, l, s, t) < \log |\mathcal{F}| + \log \left( \frac{n}{l-s} \right) + n - l + s \log 8l + c.
\]
This is because given the circuit, \( \rho \) can be described by giving both its description given a function \( f \) and a description of this \( f \). A pointer to the subcircuit in \( C \) corresponding to \( f \), which has length \( \log |\mathcal{F}| \), suffices to encode \( f \). This completes the proof of the contrapositive. \( \square \)

## 5.2 Incompressibility lemma

**Lemma 5.2** If \( \rho \) is incompressible, i.e., \( C(\rho|C, n, l, s, t) \geq \log |\mathcal{R}^l| \) and
\[
\left( \frac{n - l + s}{8lt} \right)^t \geq 2^t |\mathcal{F}|
\]
where \( c \) is the constant from Lemma 3.1, then for each \( f \in \mathcal{F} \) \( f|_\rho \) is \( s \)-open.

**Proof:** It suffices to show that under the two hypotheses,
\[
C(\rho C, n, l, s, t) \geq \log \left( \frac{n}{l-s} \right) + n - l + s \log 8l + \log |\mathcal{F}| + c.
\]

10
\[ C(\rho|C, n, l, s, t) \geq \log |\mathcal{R}'| \]
\[ \geq s \log \left( \frac{n-l+s}{l} \right) + \log \left( \frac{n}{l-s} \right) + n - l \quad \text{By Prop. 4.1} \]
\[ \geq c + \log |\mathcal{F}| + s \log 8l + \log \left( \frac{n}{l-s} \right) + n - l \quad \text{By Eq. (4)} \]

\[ \square \]

This lemma is used in two key ways. First, it is at the center of the induction step in the induction lemma which follows. Second, it is used for different values of \( s, t, l \) in the final step of the proof, which follows the induction lemma.

### 5.3 Induction lemma

For the induction to run smoothly, we keep track only of the bottom fan-in and the number of gates of height 2 or higher. This is because in applying the switching lemma, we have no control over the fan-in of gates at the second level in the resulting circuit. This is why the switching lemma is applied again in the proof of the lower bound theorem, in order to obtain a bound on the total circuit size.

**Lemma 5.3** If a depth-\( k \) circuit \( C \) on \( n \) variables has

- bottom fan-in bounded by \( \frac{1}{17}n^{1/k-1} \)
- at most \( 2\frac{1}{17}n^{1/k-1} \) gates of height 2 or higher

then \( C \) cannot correctly compute parity.

**Proof:** Following Håstad’s original proof, we proceed by induction on the depth of the circuit.
In the base case, \( k = 2 \), we assume the circuit has bottom fan-in bounded by \( \frac{1}{17}n \). Such a circuit cannot correctly compute parity since it is well-known that a depth-2 circuit computing parity must have bottom fan-in at least \( n \).

For \( k > 2 \), fix circuit \( C \) of depth-\( k \) with bottom fan-in bounded by \( t = \frac{1}{17}n^{1/k-1} \) and at most \( 2^t n^{1/k-1} \) gates of height 2 or higher. Let \( \mathcal{F} \) be the set of \( t \)-closed functions corresponding to gates in \( C \) of height 2.

To apply the incompressibility lemma, let \( s = \frac{1}{17}n^{1/k-1} \) and \( l = n^{k-2/k-1} \). Choose an incompressible restriction \( \rho \in \mathcal{R}^l \).

With these values,
\[
\left( \frac{n - l + s}{8tl} \right)^s = \left( \frac{n - n^{1/k-1} + \frac{1}{17}n^{1/k-1}}{8 \frac{1}{17}n^{1/k-1}n^{k-2/k-1}} \right)^s
= \left( \frac{17}{8}(1 - n^{-1/k-1} + \frac{1}{17}n^{-k/k-1}) \right)^s
= 2^s \left( \frac{17}{16}(1 - n^{-1/k-1} + \frac{1}{17}n^{-k/k-1}) \right)^s
\]

Since \( n^{-1/k-1} - \frac{1}{17}n^{-k/k-1} \) goes to zero as \( n \) increases, we have
\[
2^s \left( \frac{17}{16}(1 - n^{-1/k-1} + \frac{1}{17}n^{-k/k-1}) \right)^s > 2^s 2^c \quad \text{(for large enough } n) \]
\[
= 2^{\frac{17}{16}n^{1/k-1} - 2^c}
\geq |\mathcal{F}| 2^c
\]

By the incompressibility lemma, all functions in \( \mathcal{F}|\rho \) are \( s \)-open, so the circuit \( C \) can be rewritten as a depth-\( k - 1 \) circuit by collapsing levels 2 and 3 (both levels of OR gates). Assume now for a contradiction that \( C \) correctly computed parity on \( n \) variables. Then this new circuit of depth \( k - 1 \) must also compute parity, but here for \( l \) variables. The bottom fan-in is now \( s = \frac{1}{17}n^{1/k-1} = \frac{1}{17}n^{1/k-2} \) and the number of gates at levels 2 and higher remains bounded by \( 2^t n^{1/k-1} = 2^t n^{1/k-2} \). But by induction, this circuit cannot correctly compute parity, a contradiction. \( \square \)
5.4 Lower bound theorem

Theorem 5.1 If a depth-$k$ circuit has at most $2^{\frac{1}{k^k\cdot n^{1/k-1}}}$ gates, then it cannot correctly compute parity.

Proof: Given a circuit $C$ of depth $k$, we consider it as a depth-$k+1$ circuit by adding dummy single-input OR gates at the input level. We apply the incompressibility lemma to this circuit, using values $t = 1$, $l = \frac{1}{17}n$ and $s = \frac{1}{17}l^{1/k-1}$. It is easy to check that the lemma goes through for these values, so applying an incompressible restriction $\rho$ to the circuit leaves us with a new circuit of depth $k$ on $l$ variables with at most $2^{\frac{1}{17}l^{1/k-1}} = 2^{\frac{1}{17}l^{1/k-1}}$ gates of height at most 2, and bottom fan-in $s = \frac{1}{17}l^{1/k-1}$. The induction lemma applies, therefore the new circuit cannot compute parity on $l$ variables, hence the original circuit cannot compute parity of $n$ variables.

6 Further work

We feel that the techniques used in this paper may also be applicable to other circuit complexity and lower bound arguments. It would be interesting to see other proofs in the literature simplified, if not improved, using similar techniques.

7 Acknowledgements

We are grateful to Martin Kummer, Ming Li, John Rogers and Amber Settle for helpful comments on earlier drafts. Thanks also go to the two anonymous referees for painstakingly combing through all the details and making many very useful corrections and suggestions to improve this paper.

References


