Notations and conventions: We will be dealing with $\{\land, \lor, \neg\}$ formulas with the convention that the negations are pushed down to the leaves (variables).

We begin with the definition of restrictions

**Definition 1 (Restrictions).** A restriction $p$ is a mapping $\{1, 2, \ldots, n\} \rightarrow \{0, 1, \ast\}$. Given a function $\phi$ and a restriction $p$, the function $\phi$ restricted by $p$, denoted by $\phi|_p$, is defined as $\phi|_p(\vec{x}) = \phi(\vec{x})$ where

$$x_i = \begin{cases} x_i & \text{if } p(i) = \ast \\ p(i) & \text{otherwise} \end{cases}$$

Such a restriction simplifies the formulas. We can also exploit the fact that the gates are either $\lor$ or $\land$.

**Definition 2 (Constant Simplification).** A constant simplification is one in which a single literal $Z = X_i$ or $\neg X_i$ is replaced by either a 0 or 1.

Consider sub-formulas of the kind $Z \lor g$ and $Z \land g$. The following cases are possible

$$\begin{align*}
0 \land g & \rightarrow 0 \\
1 \land g & \rightarrow g \\
0 \lor g & \rightarrow g \\
1 \land g & \rightarrow 1
\end{align*}$$

We can exploit cases (1) and (4) to give further simplification.

The following fact leads to another kind of simplification.

**Fact 1.** In a minimal size $\{\land, \lor, \neg\}$ formula, any sub-formula $Z \lor \psi$ or $Z \land \psi$ where $Z$ is a literal ($X_i$ or $\neg X_i$) has no occurrence of $Z$ in $\psi$.
Proof. $Z \land \psi$ is equivalent to $Z \land \psi(Z|_1)$ and $Z \lor \psi$ is equivalent to $Z \lor \psi(Z|_0)$ where $\psi(Z|_i)$ is $\psi$ restricted to $Z = i$.

**Definition 3** (One Variable Simplification). A **one variable simplification** of a formula $\phi$ is where all occurrences of sub-formulas of form $Z \land \psi$ are replaced by $Z \land \psi(Z|_1)$ and all sub-formulas of form $Z \lor \psi$ are replaced by $Z \lor \psi(Z|_0)$.

**Question:** By how much does a formula size decrease by constant simplification?

**Theorem 1** (Subotovskaya’s Theorem). Let $\phi$ be a $\{\land, \lor, \neg\}$ formula, then $\exists$ a literal $Z$ such that the formula $\phi'$ on letting $\phi$ restricted by $Z = i$ (constant simplification) has leaf size bounded by

$$\text{leaf size}(\phi') \leq \left(1 - \frac{1}{n}\right)^3 \text{leaf size}(\phi) \tag{5}$$

**Proof.** We choose $X_i$ than appears more than $\frac{m}{n}$ times where $m$ is leaf size($\phi$). Choose either $X_i$ or $\overline{X}_i$ depending on which occurs more in the "Critical" cases which removes their neighboring sub-formulas. Without loss of generality, $X_i$ and $\overline{X}_i$ do not occur in any neighborhood of these "critical occurrences" (by Fact 1.). Now apply the constant substitution that causes most collapse. Without loss of generality, assume it is $Z \rightarrow 1$.

This removes $\frac{m}{n}$ gates where $Z$ occurs. Another $\frac{m}{2n}$ gates are removed because half of these occurrences are critical and remove the neighbor as well and as the neighbor do not include $Z$, no over counting occurs. There may also be addition removals as constant simplifications can iterate, but there are at least $\frac{3m}{2n}$ removals. Thus we get

$$\text{leaf size}(\phi') \leq m - \frac{3m}{2n} \tag{6}$$

$$= \left(1 - \frac{3}{2n}\right) m \tag{7}$$

$$\leq \left(1 - \frac{1}{n}\right)^{\frac{3}{2}} m \tag{8}$$

We will call the exponent $\frac{3}{2}$, the shrinkage factor $\Gamma$.

We can now iterate this process which gives the following lemma.

**Lemma 1.** Let $\phi$ be as before. Let $k < n$, then one can choose $n - k$ variables $X_{i_1}, \ldots, X_{i_{n-k}}$ and values $a_1, \ldots, a_{n-k} \in \{0, 1\}$ such that setting $X_{i_j} = a_j$ gives

$$\text{leaf size}(\phi') \leq \left(\frac{k}{n}\right)^{\frac{3}{2}} \text{leaf size}(\phi) \tag{9}$$

where $\phi'$ is $\phi$ with constant simplification.
Proof. We can iterate the previous construction to get

\[
\text{leaf size}(\phi') \leq \left(1 - \frac{1}{n}\right)^{\frac{3}{2}} \left(1 - \frac{1}{n-1}\right)^{\frac{3}{2}} \ldots \left(1 - \frac{1}{k+1}\right)^{\frac{3}{2}} \text{leaf size}(\phi) \tag{10}
\]

\[
= \left(\frac{n-1}{n}\right)^{\frac{3}{2}} \left(\frac{n-2}{n-1}\right)^{\frac{3}{2}} \ldots \left(\frac{k}{k+1}\right)^{\frac{3}{2}} \text{leaf size}(\phi) \tag{11}
\]

\[
= \left(\frac{k}{n}\right)^{\frac{3}{2}} \text{leaf size}(\phi) \tag{12}
\]

Another modification is possible when we choose the restrictions at random. Let \( R_k \) be the distribution on restrictions of the form where \( \rho \in R_k \) has property \( \rho(X_1, \ldots, X_n) \rightarrow \{0, 1, \ast\} \) and \( |\rho^{-1}(\ast)| = k \) and we choose these restrictions with equal probability (from a uniform distribution).

**Theorem 2** (Subotovskaya). Let \( \phi, k \) and \( n \) be as above and \( \rho_k \in R_k \) is chosen at random, then

\[
\mathbb{E}[\text{leaf size}(\phi')] \leq \left(\frac{k}{n}\right)^{\frac{3}{2}} \text{leaf size}(\phi) \tag{13}
\]

and thus

\[
\mathbb{P}\left[\text{leaf size}(\phi') \geq 4 \left(\frac{k}{n}\right)^{\frac{3}{2}} \text{leaf size}(\phi)\right] \leq \frac{1}{4} \quad [\text{by Markov’s Inequality}] \tag{14}
\]

**Proof.** Same as above except numbers are replaced by expectations everywhere. \qed

Thus from Subotovskaya, we have the shrinkage exponent \( \Gamma = \frac{3}{2} \). Using the shrinkage factors allows us to get a lower bound on the formula size of functions. Progressive improvements on the shrinkage factor have been made by using one variable simplification in addition to constant simplification. Impagliazzo-Nisan gave a shrinkage exponent of 1.55. Paterson-Zande gave a value 1.65 and Hasto gave a value of 2 for the shrinkage factor.

Using the shrinkage factor, Andreev gave a much better lower bound of \( \frac{5}{2} \) (almost) using \( \Gamma = \frac{3}{2} \) and a bound of 3 (almost) using \( \Gamma = 2 \)

We will now prove Andreev’s lower bound result.

Let \( u_{ij} \) be new variable for \( i = 1, \ldots, k \) and \( j = 1, \ldots, n/k \). Define the function

\[
f(y_0, \ldots, y_{m-1}, u_{11}, \ldots, u_{k, \frac{n}{k}}) = \text{SA}_n(y, \oplus_{j=1}^{n/k} u_{1j}, \ldots, \oplus_{j=1}^{n/k} u_{kj}).
\]

Here \( \text{SA}_n \) is the storage access function as defined in the last class.

**Claim 1.** There are constants \( \{a_0, \ldots, a_{m-1}\} \) such that \( \text{SA}_n(a_0, \ldots, a_{m-1}, z_1, \ldots, z_k) = \text{SA}_n^2(z_1, \ldots, z_k) \) requires formula size greater than \( \frac{1}{2k} \frac{2k}{\log k} \).
Proof. Immediate by Riordon-Shannon Theorem \(\square\)

Now let’s fix such a value of \(\vec{a}\). Let \(g(\vec{u}) = f(\vec{a}, \vec{u})\)

Claim 2. If \(\rho\) is a restriction such that \(\forall i = 1, \ldots, k\) there is a \(j\) such that \(\rho(u_{ij}) = *\), then \(g|_{\rho}\) requires formulas of leaf size greater than \(\frac{2^{k-1}}{\log k}\)

Proof. \(g = SA_n(\vec{y}, \oplus_{j=1}^{n/k} u_{1j}, \ldots, \oplus_{j=1}^{n/k} u_{kj})\). Now consider a restriction \(\rho' \supseteq \rho\) such that \(\rho'\) sets exactly one \(u_{ij}\) equal to * for each \(i\). Then \(g|_{\rho'} = g\) because we can flip the free variables in \(g|_{\rho'}\) to get any value of \(g\) we want (possibly with some variables negated). \(\square\)

Consider \(R_s\), the set of restrictions which leave exactly \(s\) literals unset, where \(s = k \ln(4k)\).

Claim 3. If \(\rho \in R_s\) chosen at random, then \(\Pr[\forall i \exists j \rho(u_{ij}) = \ast] \geq \frac{3}{4}\)

Proof. Each \(u_{ij} = \ast\) with probability \(\frac{s}{n} = \frac{k \ln(4k)}{n}\). So for fixed \(i\), we have

\[
\Pr[\exists j \rho(u_{ij}) \neq \ast] \leq \left( 1 - \frac{s}{n} \right)^{\frac{k}{n}} \quad (15)
\]

\[
= \left( 1 - \frac{k \ln(4k)}{n} \right)^{\frac{k}{n}} \quad (16)
\]

\[
\leq e^{-\ln(4k)} = \frac{1}{4k} \quad (17)
\]

\[
\implies \Pr[\forall i \exists j \rho(u_{ij}) \neq \ast] < k \frac{1}{4k} = \frac{1}{4} \quad (18)
\]

\(\square\)

From Subotovskaya (Theorem 2) we have

\[
\Pr \left[ \text{leaf size}(g|_{\rho}) \leq 4 \left( \frac{s}{n} \right)^{\frac{3}{2}} \text{leaf size}(g) \right] \geq \frac{3}{4}
\]

There is at least one \(\rho \in R_s\) such that \(g|_{\rho}\) requires formula of leaf size greater than \(\frac{2^{k-1}}{\log k}\) and \(g|_{\rho}\) has leafsize less than \(4 \left( \frac{s}{n} \right)^{\frac{3}{2}} \text{leaf size}(g)\). Thus

\[
\text{leaf size}(g) \geq \frac{1}{4} \left( \frac{n}{s} \right)^{\frac{3}{2}} 2^{k-1} \log k = \Omega \left( n^{\frac{s}{2}} \frac{k}{(\log n)^{\frac{3}{2}}} \log \log n \right) \quad (19)
\]

\[
= \Omega(n^{\frac{s}{2} - o(1)}) \quad (20)
\]

Thus we have proved Andreev’s lower bound.

**Corollary 1** (Andreev’s Theorem). For \(\{\land, \lor, \neg\}\) basis, size of any formula \(f\) is greater than \(n^{\frac{s}{2} - o(1)}\)
Proof. If $f$ has smaller size then $g$ does too which is not possible. □

Using Hastad’s value of 2 for the shrinkage exponent, it is possible to show that

$$L_{\{\land, \lor, \neg\}} = \Omega\left(\frac{n^3}{(\log n)^{\frac{3}{2}} (\log \log n)^{\frac{3}{2}}}\right)$$

This lower bound is tight as there are examples that achieve this bound.