CSC2429, MAT1304: Circuit Complexity November 15, 2016

Lecture 9: AC^0 Circuit Size of SUB(G), Continued

Instructor: Benjamin Rossman

Throughout today's lecture, we fix a connected graph G and a strict threshold weighting $\theta : E(G) \to [0,2]$ (such that $\Delta_\theta(H) := |V(H)| - \sum_{e \in E(H)} \theta(e) > 0$ for all $\emptyset \subset H \subset G$).

Also throughout today's lecture, let $\boldsymbol{\alpha}$ be a uniform random element of $[n]^{V(G)}$ (independent of $\mathbf{X}_\theta$). We think of $\boldsymbol{\alpha}$ as a random choice of vertices from each color class of $G^{\uparrow n}$. For each subgraph $H \subseteq G$, we have the isomorphic subgraph $H^{(\boldsymbol{\alpha})} \subseteq G^{\uparrow n}$ given by

$$E(H^{(\boldsymbol{\alpha})}) := \{\{v^{(i)}, w^{(j)}\} \in E(G^{\uparrow n}) : \{v,w\} \in E(H) \text{ and } i = \alpha_v \text{ and } j = \alpha_w\}.$$

In this lecture, we will be interested in the random ensemble $\{H^{(\boldsymbol{\alpha})}\}_{H \subseteq G}$ given by a single $\boldsymbol{\alpha}$. (In the last lecture, we were concerned with a single $H \subseteq G$ and wrote $\mathbf{H}$ for $H^{(\boldsymbol{\alpha})}$.)

Last time: We introduced *patterns* ("binary union trees") and the dual notion of *hitting sets* (a family $\mathcal{H} \subseteq \{\text{subgraphs of } G\}$ such that every $A \in \text{Pattern}(G)$ contains a graph in $\mathcal{H}$). We also introduced the parameter $\kappa_\theta(G)$ and its dual expression:

$$\kappa_\theta(G) := \min_{A \in \text{Pattern}(G)} \max_{B \preceq A} \Delta_\theta(B) = \max_{\text{hitting set } \mathcal{H} \text{ for } G} \min_{H \in \mathcal{H}} \Delta_\theta(H).$$

For subgraphs $X, S \subseteq G^{\uparrow n}$, we defined the restriction $R_{X,S} : E(G^{\uparrow n}) \to \{0,1,*\}$ by

$$R_{X,S} : e \mapsto \begin{cases} * & \text{if } e \in E(S), \\ 1 & \text{if } e \in E(X) \setminus E(S), \\ 0 & \text{otherwise.} \end{cases}$$

In particular, we are interested in random restrictions of the form $R_{\mathbf{X}_\theta, H^{(\boldsymbol{\alpha})}}$ for independent $\mathbf{X}_\theta$ and $\boldsymbol{\alpha} \in [n]^{V(G)}$.

Today: We prove a lower bound of $n^{\kappa_\theta(G) - o(1)}$ on the average-case AC^0 circuit size of SUB(G) on $\mathbf{X}_\theta$. (Recall that we measure *size* of an AC^0 by the number of AND and OR gates.) We first prove a weaker version of this lower bound with respect to *wire-size* (i.e. the number of wires in a circuits); note that wire-size$(\cdot) \leq$ size$(\cdot)^2$, so this preliminary result implies a size lower bound of $n^{\frac{1}{2}\kappa_\theta(G) - o(1)}$. We then present an additional argument that yields the $n^{\kappa_\theta(G) - o(1)}$ lower bound on number of gates.

1 Review of Main Technical Lemma and Multi-Output Version

The main result of the last lecture was a lemma on the "H-shaped sensitivity" of AC^0 functions $f : \{0,1\}^{E(G^{\uparrow n})} \to \{0,1\}$. This lemma bounded the probability that the restricted function

$$f \restriction R_{\mathbf{X}_\theta, H^{(\boldsymbol{\alpha})}} : \{0,1\}^{E(H^{(\boldsymbol{\alpha})})} \to \{0,1\}$$

depends on all of its input variables (i.e. all coordinates in the set $E(H^{(\boldsymbol{\alpha})})$). (Recall that we identify $\{0,1\}^{E(G^{\uparrow n})}$ with the set of subgraphs of $G^{\uparrow n}$, and we identify $\{0,1\}^{E(H^{(\boldsymbol{\alpha})})}$ with the set of subgraphs of $H^{(\boldsymbol{\alpha})}$.)

Below, we restate this lemma and review the key steps in its proof. In fact, we give an extended version of this lemma to *multi-output* AC^0 functions $f : \{0,1\}^{E(G^{\uparrow n})} \to \{0,1\}^m$; we highlight the multi-output version in red.

Lemma 1 (Main Technical Lemma, multi-output version). *Suppose $f : \{0,1\}^{E(G^{\uparrow n})} \to \{0,1\}^m$ is computed by an m-output circuit of depth $o(\frac{\log n}{\log\log n})$ and size $n^{O(\log\log n)}$. Then for all $H \subseteq G$,*

$$\mathop{\mathbb{P}}_{\mathbf{X}_\theta, \boldsymbol{\alpha}}[\ \mathrm{ALL}(f\restriction R_{\mathbf{X}_\theta, H^{(\boldsymbol{\alpha})}})\ ] \leq m^{|V(H)|} \cdot n^{-\Delta_\theta(H)+o(1)} \tag{1}$$

where $\mathrm{ALL}(f\restriction R_{\mathbf{X}_\theta, H^{(\boldsymbol{\alpha})}})$ is the event that the restricted m-output function $f\restriction R_{\mathbf{X}_\theta, H^{(\boldsymbol{\alpha})}} : \{0,1\}^{E(H^{(\boldsymbol{\alpha})})} \to \{0,1\}^m$ depends on all of its input variables.

A few remarks before reviewing the proof: First, note that the single-output case $m = 1$ is the precisely the version of Lemma 1 given in the last lecture (with all the red text erased). Second, note that Lemma 1 only gives a non-trivial bound when $\Delta_\theta(H) > 0$, that is, for nonempty proper subgraphs $\emptyset \subset H \subset G$. Finally, (as we will soon see) our preliminary wire-size lower bound only requires the single-output case $m = 1$ of Lemma 1, while our stronger gate-size lower bound requires us to consider $m \leq \log(n^{|V(G)|})$ (since G is fixed, we have $\log(n^{|V(G)|})^{|V(G)|} = n^{o(1)}$, so the bound (1) remains $n^{-\Delta_\theta(H)+o(1)}$).

We now review the steps in the proof of Lemma 1 and highlight the changes that yield the multi-output version.

- Without loss of generality, we assume that $\emptyset \subset H \subset G$ (since otherwise the bound (1) is trivial as $\Delta_\theta(\emptyset) = \Delta_\theta(G) = 0$).
- We generate an auxiliary random $\mathbf{S} \subseteq G^{\uparrow n}$ with $\mathbb{P}[\ \{v^{(i)}, w^{(j)}\} \in E(\mathbf{S})\ ] = n^{-\theta(\{v,w\})-\delta}$ for a fixed, arbitrarily small constant $\delta > 0$. We then generate $\boldsymbol{\alpha} \in [n]^{V(H)}$ uniformly at random conditioned on $H^{(\boldsymbol{\alpha})} \subseteq \mathbf{S}$; in case $\mathrm{sub}_H(\mathbf{S})$ is empty, generate $\boldsymbol{\alpha} \in [n]^{V(H)}$ independently.
- (Claim A) Using a tail bound for $\mathrm{sub}_H(\mathbf{S})$, we get
$$\mathbb{P}[\ \mathrm{sub}_H(\mathbf{S}) < \tfrac{1}{2} n^{\Delta_\theta(H)-\delta|E(H)|}\ ] \leq \exp(-n^{\Omega(1)}).$$
- (Claim B) Using the switching lemma, we show that for each $i \in [m]$
$$\mathbb{P}[\ \mathcal{D}(f_i\restriction R_{\mathbf{X}_\theta, \mathbf{S}}) > \delta \log n\ ] \leq n^{-\omega(\log\log n)}.$$
- (Claim C) For any fixed $X, S \subseteq G^{\uparrow n}$ with $\mathrm{sub}_H(S) \geq 1$, we have
$$\mathop{\mathbb{P}}_{\boldsymbol{\alpha}}[\ \mathrm{ALL}(f\restriction R_{X, H^{(\boldsymbol{\alpha})}}) \mid H^{(\boldsymbol{\alpha})} \subseteq S\ ] \leq \frac{(\sum_{i=1}^m 2^{\mathcal{D}(f_i\restriction R_{X,S})})^{|V(H)|}}{\mathrm{sub}_H(S)}.$$

  Proof of Claim C (multi-output version): Let $J_i \subseteq E(S)$ be the set of variables on which $f_i\restriction R_{X,S} : \{0,1\}^{E(S)} \to \{0,1\}$ depends, and let $J = J_1 \cup \cdots \cup J_m$. Note that $|J| \leq |J_1| + \cdots + |J_m|$

and $|J_i| \le 2^{\mathcal{D}(f_i \restriction R_{X,S})}$ (since every depth-k decision tree depends on $\le 2^k$ variables). For each $v \in V(H)$, define $I_v \subseteq [n]$ by

$$I_v := \big\{ i \in [n] : \{v^{(i)}, w^{(j)}\} \in J \text{ for some } w \in V(H) \text{ and } j \in [n] \big\}.$$

Now observe that, for all $\alpha \in [n]^{V(H)}$ with $H^{(\alpha)} \subseteq S$, we have

$$\begin{aligned} \mathrm{ALL}(f \restriction R_{X,H^{(\alpha)}}) &\iff f \restriction R_{X,H^{(\alpha)}} \text{ depends on all variables in } E(H^{(\alpha)}) \\ &\implies E(H^{(\alpha)}) \subseteq J \\ &\implies \textstyle\bigvee_{v \in V(H)} (\alpha_v \in I_v). \end{aligned}$$

It follows that

$$\begin{aligned} |\{\alpha \in [n]^{V(H)} : H^{(\alpha)} \subseteq S \text{ and } \mathrm{ALL}(f \restriction R_{X,H^{(\alpha)}})\}| &\le |\{\alpha \in [n]^{V(H)} : \textstyle\bigvee_{v \in V(H)} (\alpha_v \in I_v)\}| \\ &\le \textstyle\prod_{v \in V(H)} |I_v| \\ &\le |J|^{|V(H)|} \quad (\text{since } |I_v| \le |J| \text{ for all } v \in V(H)) \\ &\le \big(\textstyle\sum_{i=1}^m 2^{\mathcal{D}(f_i \restriction R_{X,S})}\big)^{|V(H)|}. \end{aligned}$$

We now have Claim C as follows:

$$\begin{aligned} \mathop{\mathbb{P}}_{\boldsymbol{\alpha}}[\, \mathrm{ALL}(f \restriction R_{X,H^{(\boldsymbol{\alpha})}}) \mid H^{(\boldsymbol{\alpha})} \subseteq S \,] &= \frac{|\{\alpha \in [n]^{V(H)} : H^{(\alpha)} \subseteq S \text{ and } \mathrm{ALL}(f \restriction R_{X,H^{(\alpha)}})\}|}{\mathrm{sub}_G(S)} \\ &\le \frac{\big(\sum_{i=1}^m 2^{\mathcal{D}(f_i \restriction R_{X,S})}\big)^{|V(H)|}}{\mathrm{sub}_G(S)}. \end{aligned}$$

- Concluding the proof of Lemma 1, we have

$$\begin{aligned} \mathbb{P}[\, \mathrm{ALL}(f \restriction R_{\mathbf{X}_\theta,H^{(\boldsymbol{\alpha})}}) \,] &\le \mathbb{P}[\, \mathrm{sub}_H(\mathbf{S}) < \tfrac{1}{2} n^{\Delta_\theta(H) - \delta|E(H)|} \,] \\ &\quad + \mathbb{P}[\, \textstyle\bigvee_{i=1}^m \mathcal{D}(f_i \restriction R_{\mathbf{X}_\theta,\mathbf{S}}) > \delta \log n \,] \\ &\quad + \mathbb{P}[\, \mathrm{ALL}(f \restriction R_{\mathbf{X}_\theta,H^{(\boldsymbol{\alpha})}}) \mid \mathrm{sub}_H(\mathbf{S}) \ge \tfrac{1}{2} n^{\Delta_\theta(H) - \delta|E(H)|} \;\&\; \textstyle\bigwedge_{i=1}^m \mathcal{D}(f_i \restriction R_{\mathbf{X}_\theta,\mathbf{S}}) \le \delta \log n \,] \\ &\le \exp(-n^{\Omega(1)}) \qquad (\text{by Claim A}) \\ &\quad + m \cdot n^{-\omega(\log\log n)} \qquad (\text{by Claim B and a union bound}) \\ &\quad + m^{|V(H)|} \cdot 2n^{-\Delta_\theta(H) + \delta(|E(H)| + |V(H)|)} \qquad (\text{by Claim C}) \\ &= n^{-\omega(\log\log n)} + O(m^{|V(H)|} \cdot n^{-\Delta_\theta(H) + \delta(|E(H)| + |V(H)|)}) \quad (\text{using } m \le n^{O(\log\log n)}) \\ &= m^{|V(H)|} \cdot n^{-\Delta_\theta(H) + o(1)} \qquad (\text{taking } \delta > 0 \text{ arbitrarily small}). \end{aligned}$$

2 Planting a G-subgraph in $\mathbf{X}_\theta$

Definition 2. We say that a function $f : \{0,1\}^{E(G^{\uparrow n})} \to \{0,1\}$ computes $\mathrm{SUB}(G)$ a.a.s. (asymptotically almost surely) on $\mathbf{X}_\theta$ if $\mathbb{P}[\, f(\mathbf{X}_\theta) = 1 \iff \mathrm{sub}_G(\mathbf{X}_\theta) \ge 1 \,] = 1 - o(1)$.

In order to prove our lower bounds, we require a second technical lemma concerning the random sub-function $f \restriction R_{\mathbf{X}_\theta, G^{(\boldsymbol{\alpha})}}$ in the setting where f computes $\mathrm{SUB}(G)$ a.a.s. on $\mathbf{X}_\theta$. Recall that $\mathrm{sub}_G(\mathbf{X}_\theta)$ is asymptotically Poisson(1), since θ is strict (Theorem 13 in Lecture 7). In particular, we will use the fact

$$\mathbb{P}[\ \mathrm{sub}_G(\mathbf{X}_\theta) = 0\ ] = \frac{1}{e} \pm o(1). \tag{2}$$

(Note: Even if θ is not strict, we have $\lim_{n\to\infty} \mathbb{P}[\ \mathrm{sub}_G(\mathbf{X}_\theta) = 0\ ] \in (0,1)$. We consider the strict case for simplicity sake and w.l.o.g. since strict threshold weightings are dense in the polytope of all threshold weightings for any connected G.)

Lemma 3. *Suppose $f : \{0,1\}^{E(G^{\uparrow n})} \to \{0,1\}$ computes $\mathrm{SUB}(G)$ a.a.s. on $\mathbf{X}_\theta$. Then the randomly restricted function $f \restriction R_{\mathbf{X}_\theta, G^{(\boldsymbol{\alpha})}} : \{0,1\}^{E(G^{(\boldsymbol{\alpha})})} \to \{0,1\}$ is a.a.s. either the* AND *function (on variables $E(G^{(\boldsymbol{\alpha})})$) or the constant* 1 *function, according to whether or not $\mathrm{sub}_G(\mathbf{X}_\theta) = 0$. That is,*

$$\mathbb{P}[\ f \restriction R_{\mathbf{X}_\theta, G^{(\boldsymbol{\alpha})}} \text{ is identically 1} \mid \mathrm{sub}_G(\mathbf{X}_\theta) \geq 1\ ] = 1 - o(1), \tag{3}$$

$$\mathbb{P}[\ f \restriction R_{\mathbf{X}_\theta, G^{(\boldsymbol{\alpha})}} \text{ is the AND function} \mid \mathrm{sub}_G(\mathbf{X}_\theta) = 0\ ] = 1 - o(1) \tag{4}$$

As a corollary of (2) *and* (4)*, we have*

$$\mathbb{P}[\ \mathrm{ALL}(f \restriction R_{\mathbf{X}_\theta, G^{(\boldsymbol{\alpha})}})\ ] \geq \frac{1}{e} - o(1). \tag{5}$$

[Exercise: Make sure you understand why (5) follows from (2) and (4).]

We merely sketch the proof of Lemma 3, omitting tedious details that involve calculations of total-variation distance.

First, let us justify (3). For simplicity sake, let's assume that f computes $\mathrm{SUB}(G)$ exactly (i.e. for every $X \subseteq G^{\uparrow n}$, we have $f(X) = 1 \iff \mathrm{sub}_G(X) \geq 1$). Consider any $X \subseteq G^{\uparrow n}$ with $\mathrm{sub}_G(X) \geq 1$, and consider any $G^{(\beta)} \in \mathrm{Sub}_G(X)$. A.a.s. for random $\boldsymbol{\alpha} \in [n]^{V(G)}$, we have $\boldsymbol{\alpha}_v \neq \beta_v$ for all $v \in V(G)$; whenever this happens, the G-subgraph $G^{(\beta)}$ will persist in X under any modifying with respect to the edges in $G^{(\boldsymbol{\alpha})}$ (in particular, $\mathrm{sub}_G(X - G^{(\boldsymbol{\alpha})}) \geq 1$), hence $f \restriction R_{\mathbf{X}_\theta, G^{(\boldsymbol{\alpha})}}$ is identically 1. This argument establishes (3) in the case where f computes $\mathrm{SUB}(G)$ exactly. Rigorously proving (3) under the hypothesis that f computes $\mathrm{SUB}(G)$ a.a.s. on $\mathbf{X}_\theta$ requires some straightforward but tedious calculations concerning total-variation distance (a.k.a. statistical distance) between random graphs $\mathbf{X}_\theta$ and $\mathbf{X}_\theta \cup H^{(\boldsymbol{\alpha})}$. In particular, one shows

$$d_{\mathrm{TV}}(\mathbf{X}_\theta, \mathbf{X}_\theta \cup H^{(\boldsymbol{\alpha})}) = o(1) \text{ for all proper } H \subset G, \tag{6}$$

$$d_{\mathrm{TV}}(\mathbf{X}_\theta, \mathbf{X}_\theta \cup G^{(\boldsymbol{\alpha})}) \xrightarrow{n\to\infty} d_{\mathrm{TV}}(\mathrm{Poisson}(1), \mathrm{Poisson}(1) + 1) < 1. \tag{7}$$

[Exercise: Show that (2) & (6) & (7) $\implies$ (3).]

Equation (4) is established by showing that for all proper $H \subset G$,

$$\mathbb{P}[\ \mathrm{sub}_G(\mathbf{X}_\theta) = \mathrm{sub}_G(\mathbf{X}_\theta \cup H^{(\boldsymbol{\alpha})})\ ] = 1 - o(1). \tag{8}$$

In other words, the probability that adding a random H-subgraph to $\mathbf{X}_\theta$ creates an additional G-subgraph is $o(1)$. Equation (8) is shown by a straightforward union bound:

$$\begin{aligned}
\mathbb{P}[\ \mathrm{sub}_G(\mathbf{X}_\theta) \neq \mathrm{sub}_G(\mathbf{X}_\theta \cup H^{(\boldsymbol{\alpha})})\ ] &\leq \mathbb{P}[\ \bigvee_{\emptyset \subset H' \subseteq H} \ \bigvee_{\beta \in [n]^{V(G)-V(H')}} G^{(\boldsymbol{\alpha}_{V(H')} \cup \beta)} \subseteq \mathbf{X}_\theta \cup H^{(\boldsymbol{\alpha})}\ ] \\
&\leq \sum_{\emptyset \subset H' \subseteq H} \ \sum_{\beta \in [n]^{V(G)-V(H')}} \mathbb{P}[\ G^{(\boldsymbol{\alpha}_{V(H')} \cup \beta)} \subseteq \mathbf{X}_\theta \cup H^{(\boldsymbol{\alpha})}\ ] \\
&\leq \sum_{\emptyset \subset H' \subseteq H} n^{|V(G)|-|V(H')|} \prod_{e \in E(G-H')} n^{-\theta(e)} \\
&= \sum_{\emptyset \subset H' \subseteq H} n^{|V(G)|-|V(H')|} \Big(\prod_{e \in E(G)} n^{-\theta(e)} \Big) \Big(\prod_{e \in E(H')} n^{\theta(e)} \Big) \\
&= \sum_{\emptyset \subset H' \subseteq H} n^{-\Delta_\theta(H')} \\
&\leq 2^{|V(G)|} \cdot n^{-\Omega(1)} \quad \text{(by strictness of } \theta) \\
&\leq o(1).
\end{aligned}$$

Using (2) and (8), by a further union bound over proper $H \subset G$, we get

$$\mathbb{P}[\ \bigwedge_{H \subset G} \mathrm{sub}_G(\mathbf{X}_\theta \cup H^{(\boldsymbol{\alpha})}) = 0 \mid \mathrm{sub}_G(\mathbf{X}_\theta) = 0\ ] = 1 - o(1). \tag{9}$$

[Exercise: Show that (9) & (6) & (7) $\Longrightarrow$ (4).]

3 Wire-Size Lower Bound

Onto our main results! We begin with our preliminary lower bound for wire-size.

Theorem 4. *The average-case* AC^0 *circuit wire-size of* $\mathrm{SUB}(G)$ *on* $\mathbf{X}_\theta$ *is at least* $n^{\kappa_\theta(G)-o(1)}$. *That is, suppose* $f : \{0,1\}^{E(G^{\uparrow n})} \to \{0,1\}$ *satisfies*

$$\mathbb{P}[\ f(\mathbf{X}_\theta) = 1 \iff \mathrm{sub}_G(\mathbf{X}_\theta) \geq 1\ ] = 1 - o(1)$$

and f *is computed by an* AC^0 *circuit* C_0 *of depth* $o(\frac{\log n}{\log \log n})$. *Then* C_0 *has at least* $n^{\kappa_\theta(G)-o(1)}$ *wires.*

We give the proof in $0 + 3$ steps.

Step 0: Toward a contradiction, assume that f computes $\mathrm{SUB}(G)$ a.a.s. on $\mathbf{X}_\theta$, but yet is computed by an AC^0 circuit C_0 of depth $d = o(\frac{\log n}{\log \log n})$ with $s \leq n^{\kappa_\theta(G)-\Omega(1)}$ wires. (As usual, w.l.o.g. we assume that C_0 has negations on inputs.)

Step 1: Replace each unbounded fan-in AND/OR gate in C_0 with a binary tree of fan-in 2 AND/OR gates. This results in an equivalent circuit C (computing the same function f) where

- C has $\leq s$ gates,

- C has unbounded depth (i.e. we don't care about the depth of C), but every input-to-output path in C has at most d *alternations* between ANDs and ORs.

Crucially, the function $g : \{0,1\}^{E(G^{\uparrow n})} \to \{0,1\}$ computed at any gate g of C is *computable* by a standard AC^0 circuit of depth $\leq d$ and size $\leq s$ (i.e. collapsing the sub-circuit C_g back down to depth $\leq d$ by combining adjacent fan-in 2 gates of the same type into a single AND/OR gate). Since $s \leq n^{\kappa_\theta(G)-\Omega(1)} \leq n^{O(\log\log n)}$, we may apply Lemma 1 to the function g: for all $H \subseteq G$, we have

$$\mathbb{P}[\ \mathrm{ALL}(g\restriction R_{\mathbf{X}_\theta, H^{(\boldsymbol{\alpha})}})\ ] \leq n^{-\Delta_\theta(H)+o(1)}.$$

Step 2: By the dual expression for $\kappa_\theta(G)$, there exists a hitting set $\mathcal{H} \subseteq \{\text{subgraphs of } G\}$ such that $\kappa_\theta(G) = \min_{H\in\mathcal{H}}\ \Delta_\theta(H)$. Fix any such $\mathcal{H}$.

Taking a union bound over gates in C and graphs in $\mathcal{H}$, we have

$$\begin{aligned}
\mathbb{P}[\bigvee_{\text{gates } g \text{ in } C}\ \bigvee_{H\in\mathcal{H}} \mathrm{ALL}(g\restriction R_{\mathbf{X}_\theta, H^{(\boldsymbol{\alpha})}})\ ] &\leq \sum_{\text{gates } g \text{ in } C}\ \sum_{H\in\mathcal{H}} \mathbb{P}[\ \mathrm{ALL}(g\restriction R_{\mathbf{X}_\theta, H^{(\boldsymbol{\alpha})}})\ ] \\
&\leq \sum_{\text{gates } g \text{ in } C}\ \sum_{H\in\mathcal{H}} n^{-\Delta_\theta(H)+o(1)} \quad \text{(by Step 1)} \\
&\leq s \cdot |\mathcal{H}| \cdot n^{-\kappa_\theta(G)+o(1)} \\
&\leq n^{-\Omega(1)} \quad \text{(since } s \leq n^{\kappa_\theta(G)-\Omega(1)} \text{ and } |\mathcal{H}| \leq n^{o(1)}\text{)} \\
&\leq o(1).
\end{aligned}$$

Step 3: We claim that the following implication holds (with probability 1):

$$\mathrm{ALL}(f\restriction R_{\mathbf{X}_\theta, G^{(\boldsymbol{\alpha})}}) \implies \bigvee_{\text{gates } g \text{ of } C}\ \bigvee_{H\,\in\,\mathcal{H}} \mathrm{ALL}(g\restriction R_{\mathbf{X}_\theta, H^{(\boldsymbol{\alpha})}}).$$

Before proving this claim, let's first see that it suffices to complete the proof of Theorem 4. On the one hand, Lemma 3 implies $\mathbb{P}[\ \mathrm{ALL}(f\restriction R_{\mathbf{X}_\theta, G^{(\boldsymbol{\alpha})}})\ ] \geq \frac{1}{e} - o(1)$. On the other hand, Steps 2 and 3 imply $\mathbb{P}[\ \mathrm{ALL}(f\restriction R_{\mathbf{X}_\theta, G^{(\boldsymbol{\alpha})}})\ ] \leq o(1)$. These two inequalities are inconsistent, which gives the desired contradiction (showing that the assumption in Step 0 is inconsistent, which proves the theorem).

It remains to prove the implication in Step 3. Since the statement is non-probabilistic, we consider *any* fixed $X \subseteq G^{\uparrow n}$ and $\alpha \in [n]^{V(G)}$. By induction on gates/nodes g in C, for each subgraph $H \subseteq G$ such that $\mathrm{ALL}(g\restriction R_{X, H^{(\alpha)}})$, we define a pattern $A_{g,H} \in \mathrm{Pattern}(H)$ inductively:

- If g is an input labeled by a variable $x_{\{v^{(i)}, w^{(j)}\}}$ or its negation $\neg x_{\{v^{(i)}, w^{(j)}\}}$. If $\mathrm{ALL}(g\restriction R_{X, H^{(\alpha)}})$ holds, then either H is empty (in which case $A_{g,H} := \emptyset$, or H is the single-edge graph $E(H) = \{\{v, w\}\}$ and $\alpha_v = i$ and $\alpha_w = j$). In this case, let $A_{g,H}$ be the atomic pattern labeled by H.

- Suppose g is a (fan-in 2) gate with children g_1 and g_2 (that is, $g = g_1 \wedge g_2$ or $g = g_1 \vee g_2$). For $i \in \{1,2\}$, let H_i be the subgraph of H defined by

$$E(H_i) := \Big\{\{v, w\} \in E(H) : g_i\restriction R_{g, H^{(\alpha)}} \text{ depends on variable } \{v^{(\alpha_v)}, w^{(\alpha_w)}\}\Big\}.$$

Note that $H = H_1 \cup H_2$, since each sensitive variable of $g\restriction R_{g,H^{(\alpha)}}$ must be sensitive for $g_1\restriction R_{g,H^{(\alpha)}}$ or $g_2\restriction R_{g,H^{(\alpha)}}$ (or both). Next, note that $\mathrm{ALL}(g_i\restriction R_{X,H_i^{(\alpha)}})$ [exercise: convince yourself!]; therefore, the pattern A_{g_i,H_i} is well-defined by the induction hypothesis. Finally, define $A_{g,H} := \langle A_{g_1,H_1}, A_{g_2,H_2}\rangle$. (Recall our notation from Lecture 8: $\langle A, B\rangle$ is the pattern consisting of a root with sub-patterns A and B as children.)

If we now assume that $\mathrm{ALL}(f\restriction R_{X,G^{(\alpha)}})$, then we may consider the pattern $A_{f,G}$. By definition of "hitting set", there exists a sub-pattern $A_{g,H} \preceq A_{f,G}$ such that $A_{g,H} \in \mathcal{H}$. (Obs: All sub-patterns of $A_{f,G}$ have the form $A_{g,H}$ for some gate $g \in C$ and subgraph $H \subseteq G$.) This completes the proof of Step 3 and hence of the entire theorem.

4 Improving Wire-Size to (Gate-)Size

As remarked earlier, Theorem 4 implies a lower bound of $n^{\frac{1}{2}\kappa_\theta(G)-o(1)}$ on the average-case AC^0 circuit size of $\mathrm{SUB}(G)$ on $\mathbf{X}_\theta$ (via the fact that wire-size$(\cdot) \le$ size$(\cdot)^2$). We will now improve this to $n^{\kappa_\theta(G)-o(1)}$ using a similar argument to Theorem 4, but with an extra trick. As we will see, this trick requires the multi-output version of Lemma 1 with $m = \mathrm{polylog}(n)$.

Theorem 5. *The average-case* AC^0 *circuit size of* $\mathrm{SUB}(G)$ *on* $\mathbf{X}_\theta$ *is at least* $n^{\kappa_\theta(G)-o(1)}$.

We present the proof along similar steps as Theorem 4.

Step 0: Toward a contradiction, assume that f computes $\mathrm{SUB}(G)$ a.a.s. on $\mathbf{X}_\theta$, but yet is computed by an AC^0 circuit C of depth $d = o(\frac{\log n}{\log\log n})$ with $s \le n^{\kappa_\theta(G)-\Omega(1)}$ gates.

Step 1: Consider any gate g in C. Suppose g is an AND gate, that is, $g = \mathrm{AND}(g_1, \dots, g_m)$ for some $m \ge 2$. Let $\widetilde{g} : \{0,1\}^{E(G^{\uparrow n})} \to \{0,1\}^m$ be the m-output function where

$$\widetilde{g}_i := \mathrm{AND}(g_1, \dots, g_i).$$

That is, $\widetilde{g}$ is the m-tuple of ANDs of initial segments of children $g_1, \dots, g_m$ of g. Note that $\widetilde{g}$ takes at most $m+1$ possible values, namely strings 1^i0^{m-i} where $i \in \{0, \dots, m\}$.

Let $\widetilde{g}_{\mathrm{concise}} : \{0,1\}^{E(G^{\uparrow n})} \to \{0,1\}^{\lceil \log(m+1)\rceil}$ be the $\lceil \log(m+1)\rceil$-output function that gives the binary representation of $i \in \{0, \dots, m\}$ such that $\widetilde{g}(\cdot) = 1^i0^{m-i}$. We define $\widetilde{g}$ and $\widetilde{g}_{\mathrm{concise}}$ analogously when g is an OR gate (with the roles of 0 and 1 reversed).

The key observation is that $\widetilde{g}_{\mathrm{concise}}$ is computable by a circuit of size $\le s + \mathrm{poly}(m) \le n^{O(\log\log n)}$ and depth $\le d + O(1) \le o(\frac{\log n}{\log\log n})$. [Exercise: Make sure you see why.] Since $\widetilde{g}$ and $\widetilde{g}_{\mathrm{consise}}$ have the same information content (i.e. there is a bijection π from Range$(\widetilde{g})$ to Range$(\widetilde{g}_{\mathrm{consise}})$ such that $\pi(\widetilde{g}(X)) = \widetilde{g}_{\mathrm{concise}}(X)$ for all $X \subseteq G^{\uparrow n}$), we have $\mathrm{ALL}(\widetilde{g}\restriction\varrho) \iff \mathrm{ALL}(\widetilde{g}_{\mathrm{concise}}\restriction\varrho)$ for any restriction ϱ. By the multi-output version of Lemma 1, for all $H \subseteq G$ we have

$$\begin{aligned}
\mathbb{P}[\ \mathrm{ALL}(\widetilde{g}\restriction R_{\mathbf{X}_\theta,H^{(\boldsymbol{\alpha})}})\ ] &= \mathbb{P}[\ \mathrm{ALL}(\widetilde{g}_{\mathrm{concise}}\restriction R_{\mathbf{X}_\theta,H^{(\boldsymbol{\alpha})}})\ ] \\
&\le \lceil \log(m+1)\rceil^{|V(H)|} \cdot n^{-\Delta_\theta(H)+o(1)} \\
&\le \lceil \log(n^{\Delta_\theta(H)-\Omega(1)})\rceil^{|V(H)|} \cdot n^{-\Delta_\theta(H)+o(1)} \\
&\le n^{-\Delta_\theta(H)+o(1)}.
\end{aligned}$$

Step 2: By the dual expression for $\kappa_\theta(G)$, there exists a hitting set $\mathcal{H} \subseteq \{\text{subgraphs of } G\}$ such that $\kappa_\theta(G) = \min_{H \in \mathcal{H}} \Delta_\theta(H)$. Fix any such $\mathcal{H}$.

Taking a union bound over gates in C and graphs in $\mathcal{H}$, we have

$$\begin{aligned}
\mathbb{P}[\bigvee_{\text{gates } g \text{ in } C} \bigvee_{H \in \mathcal{H}} \mathrm{ALL}(\widetilde{g}\restriction R_{\mathbf{X}_\theta, H^{(\boldsymbol{\alpha})}})\ ] &\le \sum_{\text{gates } g \text{ in } C} \sum_{H \in \mathcal{H}} \mathbb{P}[\ \mathrm{ALL}(\widetilde{g}\restriction R_{\mathbf{X}_\theta, H^{(\boldsymbol{\alpha})}})\ ] \\
&\le \sum_{\text{gates } g \text{ in } C} \sum_{H \in \mathcal{H}} n^{-\Delta_\theta(H)+o(1)} \quad \text{(by Step 1)} \\
&\le s \cdot |\mathcal{H}| \cdot n^{-\kappa_\theta(G)+o(1)} \\
&\le n^{-\Omega(1)} \quad (\text{since } s \le n^{\kappa_\theta(G)-\Omega(1)} \text{ and } |\mathcal{H}| \le n^{o(1)}) \\
&\le o(1).
\end{aligned}$$

(This looks exactly like Step 2 of Theorem 4, except that the event $\mathrm{ALL}(g\restriction R_{\mathbf{X}_\theta, H^{(\boldsymbol{\alpha})}})$ has been replaced by $\mathrm{ALL}(\widetilde{g}\restriction R_{\mathbf{X}_\theta, H^{(\boldsymbol{\alpha})}})$.)

Step 3: We claim that the following implication holds (with probability 1):

$$\mathrm{ALL}(f\restriction R_{\mathbf{X}_\theta, G^{(\boldsymbol{\alpha})}}) \implies \bigvee_{\text{gates } g \text{ of } C} \bigvee_{H \in \mathcal{H}} \mathrm{ALL}(\widetilde{g}\restriction R_{\mathbf{X}_\theta, H^{(\boldsymbol{\alpha})}}).$$

Similarly as in Theorem 4, Steps 2 and 3 imply that $\mathbb{P}[\ \mathrm{ALL}(f\restriction R_{\mathbf{X}_\theta, G^{(\boldsymbol{\alpha})}})\ ] \le o(1)$, while Lemma 3 implies $\mathbb{P}[\ \mathrm{ALL}(f\restriction R_{\mathbf{X}_\theta, G^{(\boldsymbol{\alpha})}})\ ] \ge \frac{1}{e} - o(1)$. This contradiction shows that the assumption in Step 0 is inconsistent, which proves the theorem.

To justify this claim, we consider the fan-in 2 circuit C^* where each AND gate of C of the form $g = \mathrm{AND}(g_1, \dots, g_m)$ is replaced by the binary tree of fan-in 2 gates

$$(\cdots((g_1 \wedge g_2) \wedge g_3) \cdots) \wedge g_m$$

(and similarly for each OR gate of C). By Step 3 of Theorem 4, we have the implication

$$\mathrm{ALL}(f\restriction R_{\mathbf{X}_\theta, G^{(\boldsymbol{\alpha})}}) \implies \bigvee_{\text{gates } g^* \text{ of } C^*} \bigvee_{H \in \mathcal{H}} \mathrm{ALL}(g^*\restriction R_{\mathbf{X}_\theta, H^{(\boldsymbol{\alpha})}}).$$

For each gate g^* in C^*, there is a gate g of fan-in m in C and an index $i \in \{1, \dots, m\}$ such that g^* computes the same function as $\widetilde{g}_i$; hence, for any restriction ϱ,

$$\mathrm{ALL}(g^*\restriction \varrho) \iff \mathrm{ALL}(\widetilde{g}_i\restriction \varrho) \implies \mathrm{ALL}(\widetilde{g}\restriction \varrho).$$

(Clearly if $\widetilde{g}_i\restriction\varrho$ depends on all variables, then so does $\widetilde{g}$, since $\widetilde{g}_i$ is just one coordinate of the output of $\widetilde{g}$.) Putting things together to complete the proof, we have

$$\begin{aligned}
\mathrm{ALL}(f\restriction R_{\mathbf{X}_\theta, G^{(\boldsymbol{\alpha})}}) &\implies \bigvee_{\text{gates } g^* \text{ of } C^*} \bigvee_{H \in \mathcal{H}} \mathrm{ALL}(g^*\restriction R_{\mathbf{X}_\theta, H^{(\boldsymbol{\alpha})}}) \\
&\iff \bigvee_{\text{gates } g = \mathrm{AND/OR}(g_1, \dots, g_m) \text{ of } C} \bigvee_{i \in \{1,\dots,m\}} \bigvee_{H \in \mathcal{H}} \mathrm{ALL}(\widetilde{g}_i\restriction R_{\mathbf{X}_\theta, H^{(\boldsymbol{\alpha})}}) \\
&\implies \bigvee_{\text{gates } g \text{ of } C} \bigvee_{H \in \mathcal{H}} \mathrm{ALL}(\widetilde{g}\restriction R_{\mathbf{X}_\theta, H^{(\boldsymbol{\alpha})}}).
\end{aligned}$$

5 Tree-width and $\max_\theta \kappa_\theta(G)$

Definition 6 (Tree decompositions, tree-width, path-width)**.**

- A *tree decomposition* of a graph G is a pair $(T, \mathcal{W})$ where T is an (unrooted) tree and $\mathcal{W} = \{W_t\}_{t \in V(T)}$ is a family of sets $W_t \subseteq V(G)$ such that
  - every edge of G has both ends in some W_t,
  - if a vertex $v \in V(G)$ lies in both W_t and $W_{t'}$, then $v \in W_{t''}$ for every node t'' that lie on the path in T between t and t'.
- The *width* of a tree decomposition $(T, \mathcal{W})$ is defined as $\max_{t \in V(T)} |W_t| - 1$.
- The *tree-width* of G, denoted $\mathbf{tw}(G)$, is the minimum width of a tree decomposition for G.

Exercise 7. Show that SUB(G) is solvable (in the worst-case) by AC^0 circuits of size $n^{\mathbf{tw}(G)+1+o(1)}$.

Together, Theorem 5 and Exercise 7 imply that $\max_\theta \kappa_\theta(G) \leq \mathbf{tw}(G) + 1$. Below, we give a direct combinatorial proof of this fact. In fact, we show that $\max_\theta \kappa_\theta(G)$ is upper-bounded by a quantity known as *branch-width*, which is at most $\mathbf{tw}(G) + 1$.

Definition 8. A *branch-decomposition* of G is a cubic tree T (where all non-leaves have degree exactly 3) together a bijection ε from Leaves(T) and $E(G)$. Each edge $\{t, t'\} \in E(T)$ induces a partition $G = H \uplus H'$ into edge-disjoint subgraphs, where $E(H)$ (resp. $E(H')$) is the set of edges $\varepsilon(\ell) \in E(G)$ for all $\ell \in$ Leaves(T) such that $\mathrm{dist}_T(\ell, t)$ is one less (resp. one more) than $\mathrm{dist}_T(\ell, t')$. The *width* of a branch-decomposition T is the maximum of $|V(H) \cap V(H')|$ over partitions $G = H \uplus H'$ given by edges of T. The *branch-width* of G, denoted $\mathbf{bw}(G)$, is the minimum width of a branch-decomposition of G.

Branch-width is known to be within a linear factor of tree-width.

Theorem 9 (Robertson and Seymour 1991)**.** $\mathbf{bw}(G) \leq \mathbf{tw}(G) + 1 \leq \frac{3}{2}\mathbf{bw}(G)$

We now directly show:

Proposition 10. *$\kappa_\theta(G) \leq \mathbf{bw}(G)$ for every threshold weighting θ.*

Proof. For every $H \subseteq G$, let $\overline{H} \subseteq G$ be the edge-complementary subgraph with $E(\overline{H}) = E(G) \setminus E(H)$.

Fix a minimum width branch-decomposition T of G. This induces a (read-once) pattern $A \in$ Pattern(G) as follows: viewed as a rooted binary tree, the vertices of A are the edges of T; the root of A is chosen to be any edge of T which has a leaf as an endpoint. The partitions $G = H \uplus H'$ induced by edges of T are precisely the partitions $G = G_B \uplus G_{B'}$ for sub-patterns $\langle B, B' \rangle \preceq A$ (note that $G_{B'} = \overline{G_B}$). For every threshold weighting θ, we have

$$\begin{aligned}
\max_{B \preceq A} \Delta_\theta(G_B) &\leq \max_{B \preceq A} \Delta_\theta(G_B) + \Delta_\theta(\overline{G_B}) \\
&= \max_{B \preceq A} |V(G_B)| + |V(\overline{G_B})| - \sum_{e \in E(G)} \theta(e) \\
&= \max_{B \preceq A} |V(G_B)| + |V(\overline{G_B})| - |V(G)| \\
&= \max_{B \preceq A} |V(G_B) \cap V(\overline{G_B})| \\
&\leq \mathrm{width}(T) = \mathbf{bw}(G).
\end{aligned}$$

□