# 6. Depth Reduction for Algebraic Circuits

Thursday, September 7, 2023 12:42 PM

**Thm** (Depth reduction for circuits) [Valiant-Skyum-Berkowitz-Rackoff '83]

If $f$ is computed by a circuit of size $s$ and $\deg(f)=d$, then $f$ is computed by a circuit of size $\mathrm{poly}(s,d)$ and depth $O(\log d\cdot(\log d+\log s))$

Let $C$ be a circuit computing $f$. $s=\mathrm{size}(C)$, $d=\deg(f)$.

We'll construct a new circuit of depth $O(\log d(\log d+\log s))$ computing $f$.

- We may assume $f$ is homogeneous since computing $f$ from $\mathrm{Hom}_i(f)$, $i=0,\dots,d$ only takes extra depth $O(\log d)$.
- By homogenization, we may further assume $C$ is a homogeneous circuit.

For a gate $v$, we abuse the notation and denote by $v$ the polynomial output by $v$.

Notation: for two gates $v, w$ in $C$, consider a new circuit where $w$ is replaced by a new variable $Y$.

Let $f_{v,w}\in\mathbb{F}[X_1,\dots,X_n,Y]$ be the polynomial output by $v$ in this new circuit.

Finally, define $\partial_w(v) := \left(\frac{\partial f_{v,w}}{\partial Y}\right)\Big|_{Y=w} = \left(\frac{\partial f_{v,w}}{\partial Y}\right)(X_1,\dots,X_n,w)\in\mathbb{F}[X_1,\dots,X_n]$.

Lemma: Let $v,w$ be two gates in a homogeneous circuit. Then:

(1) $\partial_w(v)$ is either zero or homogeneous of degree $\deg v-\deg w$.

(2) Suppose $v$ is a product gate with children $v_1$ & $v_2$, where $\deg(v_1)\geq\deg(v_2)$.

Assume $v\neq w$

If $\deg(w)>\frac{1}{2}\cdot\deg(v)$, then $\partial_w(v)=\partial_w(v_1)\cdot v_2$.

(3) Suppose $v$ is a sum gate with children $v_1$ & $v_2$. Then $\partial_w(v)=\partial_w(v_1)+\partial_w(v_2)$.

Pf: (1) Write $f_{v,w}=\sum_i f_i Y^i$. Then $\deg(f_i)+(\deg(w))^i=\deg(v)$ for all $i$.

(or $f_{v,w}=0$)

Pf: (1) Write $f_{v,w} = \sum_i f_i Y^i$. Then $\deg(f_i) + (\deg(w))^i = \deg(v)$ for all $i$. (or $f_{v,w} = 0$)

$$\partial_w(v) = \frac{\partial f_{v,w}}{\partial Y}\bigg|_{Y=w} = \sum_i f_i w^{i-1},$$ which is homogeneous of degree $\deg v - \deg w$ (or is zero).

(2). In the circuit where $w$ is replaced by $Y$, suppose $v, v_1, v_2$ computes $v', v_1', v_2' \in \mathbb{F}[x_1, \cdots, x_n, Y]$ respectively.

Then $v' = v_1' \cdot v_2'$. (where we use $v \neq w$).

So $\frac{\partial v'}{\partial Y} = \frac{\partial v_1'}{\partial Y} \cdot v_2' + \frac{\partial v_2'}{\partial Y} v_1'$.

So $\partial_w(v) = \partial_w(v_1) \cdot v_2 + \partial_w(v_2) \cdot v_1$

note: $\deg(v) = \deg(v_1) + \deg(v_2)$, $\deg(v_1) \geq \deg(v_2)$

But $\partial_w(v_2) = 0$ since $\deg(v_2) \leq \deg(v)/2 < \deg(w)$

(3). Use the above notations $v', v_1', v_2'$.

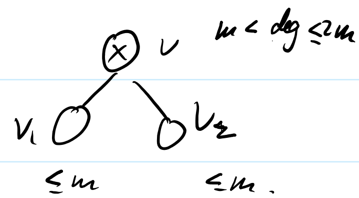
Then $v' = v_1' + v_2'$. So $\frac{\partial v'}{\partial Y} = \frac{\partial v_1'}{\partial Y} + \frac{\partial v_2'}{\partial Y}$.

So $\partial_w(v) = \partial_w(v_1) + \partial_w(v_2)$. $\square$

For an integer $m \geq 0$, denote by $G_m$ the set of product gates $v \in C$ satisfying:

$m < \deg(v) \leq 2m$

$\deg(v_1), \deg(v_2) \leq m$, where $v_1, v_2$ are the children of $v$.




Claim: Let $m \geq 0$. Let $v, w$ be gates such that

$$\deg(w) \leq m < \deg(v) \leq 2\deg(w).$$

Then: (1) $v = \sum_{t \in G_m} t \cdot \partial_t(v)$.

(2) $\partial_w v = \sum_{t \in G_m} \partial_w(t) \cdot \partial_t(v)$.

Proof of Thm (assuming the claim):

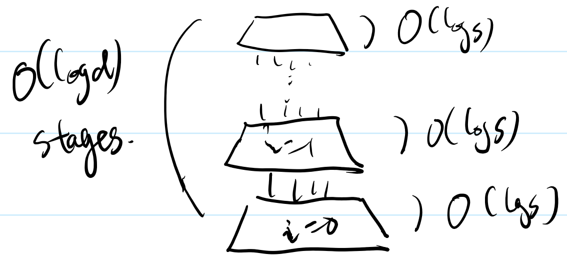
For $i = 0, 1, 2, \cdots, \lceil \log_2 d \rceil$,

(1) compute all $v \in C$ s.t. $2^{i-1} < \deg(v) \leq 2^i$ ← depth $O(\log s)$ /.

For $i = 0, 1, 2, \ldots, \log_2 d - 1$ ← depth $O(\log s)$

(1) compute all $v \in C$ s.t. $2^{i-1} < \deg(v) \le 2^i$

(2) compute all $\partial_w(v)$ for $v, w \in C$ s.t. $2^{i-1} < \deg(v) - \deg(w) \le 2^i$, $\deg(v) < 2\deg(w)$.




$s$ is actually $\overset{s'}{O}(d^2 s)$ due to homogenization.

$\Rightarrow$ depth $O((\log d)(\log s')) = O(\log d(\log d + \log s))$.

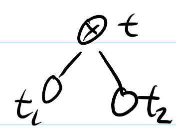
Base case: Compute $v$ and $\partial_w(v)$ directly if $\deg(v) \le 1$

(resp. $\deg(v) - \deg(w) \le 1$).

These polynomials have degree $\le 1$.

At stage $i+1$, let $m = 2^i$, $\forall$ for $v \in C$ s.t. $m = 2^i < \deg(v) \le 2^{i+1}$

part 1: by the claim, $v = \sum_{t \in G_m} t \cdot \partial_t(v) = \sum_{t \in G_m} t_1 \cdot t_2 \cdot \partial_t(v)$




where $\deg(t_1), \deg(t_2) \le m$ (since $t \in G_m$)

so $t_1$ and $t_2$ are already computed.

$\deg v - \deg t \le 2^{i+1} - 2^i \le 2^i$ $\quad$ $\deg v \le 2^{i+1} = 2 \cdot 2^i < 2 \cdot \deg(t)$ (since $t \in G_m$)

so $\partial_t(v)$ is also computed already.

Computing $v = \sum_{t \in G_m} t_1 \cdot t_2 \, \partial_t(v)$ takes depth $O(\log s)$.

Part 2: consider $v, w$ s.t. $2^i < \deg(v) - \deg(w) \le 2^{i+1}$ and $\deg(v) \le 2\deg(w)$.

let $m = 2^i + \deg(w)$.

By the claim, $\partial_w(v) = \sum_{t \in G_m} \partial_w(t)\, \partial_t(v)$.

where $\deg(t_2) \le \deg(t_1) \le m$

$m < \deg(t) = \deg(t_1) + \deg(t_2) \le 2m$

We may only consider $t$ s.t. $\partial_w(t) \ne 0$ and $\partial_t(v) \ne 0$

This means $\deg(w) \le \deg(t) \le \deg(v)$. ($< 2\deg(w)$ by assumption)

[illegible] $\exists$ a path $w \to t_1$ )

This means $\deg(w) \le \deg(t) \le \deg(v)$. ($< 2\deg(w)$ by assumption)

We may also assume $t_1$ or $t_2$ depends on $w$ (i.e. $\exists$ a path $w \to t_1$ or $w \to t_2$)

otherwise, we would have $\partial_w(t) = 0$.

So $\deg(w) \le \deg(t_1)$.

As $2\deg(w) > \deg(t)$, by (2) of the Lemma, $\partial_w(t) = \partial_w(t_1) \cdot t_2$.

So $\partial_w(v) = \sum_{t \in G'_m} \partial_w(t_1)\, \partial_t(v)\, t_2$. where $G'_m \subseteq G_m$.

check that $t_2$, $\partial_w(t_1)$, $\partial_t(v)$ are computed:

$$\deg(v) \le 2^{i+1} + \deg(w) \le 2^{i+1} + \deg(t_1) = 2^{i+1} + \deg(t) - \deg(t_2)$$

$\Rightarrow \deg(t_2) \le 2^{i+1} + \deg(t) - \deg(v) \le 2^{i+1}$. So $t_2$ is computed.

One can also check $\partial_w(t_1)$ and $\partial_t(v)$ have been computed.

Reference: Shpilka - Yehudayoff '10. $\square$

Proof of the Claim: We first verify $v = \sum_{t \in G_m} \partial_t(v) \cdot t$

Induction on the length of the longest path from $G_m$ to $v$.

Base case: $v \in G_m$. Then $v = v_1 \cdot v_2$. $\deg(v_1), \deg(v_2) \le m$.

For $t = v$, $\partial_t(v) = 1$

For $t \ne v$, $t \in G_m$: $v_1$ and $v_2$ does not depend on $t$ since $\deg(t) > m$. So $\partial_t(v) = 0$.

So $v = \sum_{t \in G_m} \partial_t(v) \cdot t$

Induction Step: Suppose $\overset{\notin G_m}{v} = v_1 + v_2$. May assume $\deg(v_1) = \deg(v_2) = \deg(v)$.

Then claim follows by induction.

Now suppose $v = v_1 \cdot v_2$, $v \notin G_m$, $\deg(v_1) \ge \deg(v_2)$

By the induction hypothesis on $v_1$,

$v_1 = \sum_{t \in G_m} \partial_t(v_1) \cdot t$.

$\deg(v) \le 2m < 2\deg(t)$. Use Lemma (2).

$$v_1 = \sum_{t \in G_m} \partial_t(v_1) \cdot t.$$

$\deg(v) \le 2m < 2\deg(t)$. use Lemma (2).

$$v = v_1 \cdot v_2 = \sum_{t \in G_m} \left(\partial_t(v_1) v_2\right) t = \sum_{t \in G_m} \partial_t(v) \cdot t$$

---

Next, we verify $\partial_w(v) = \sum_{t \in G_m} \partial_t(v) \cdot \partial_w(t)$. $(*)$

Base case: $v \in G_m$: for $t = v$, $\partial_t(v) = 1$, $\partial_w(t) = \partial_w(v)$. So $(*)$ holds.

for $t \ne v$, $t \in G_m$, $\partial_t(v) = 0$.

Induction step: $v \notin G_m$. $v = v_1 + v_2$. follows since $\partial_w(\cdot)$ and $\partial_t(\cdot)$ are linear.

$v = v_1 \cdot v_2$. $\deg(v_1) \ge \deg(v_2)$

Then $\partial_w(v_1) = \sum_{t \in G_m} \partial_t(v_1) \cdot \partial_w(t)$,

$\partial_w(v) = \partial_w(v_1) \cdot v_2$

use (2) of the Lemma.

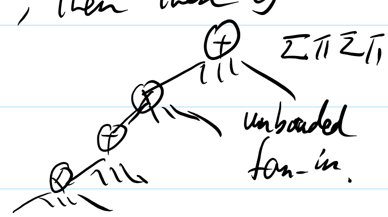
$$= \sum_{t \in G_m} \left(\partial_t(v_1)\, v_2\right) \cdot \partial_w(t) = \sum_{t \in G_m} \partial_t(v) \cdot \partial_w(t) \qquad \square$$

---

Other depth reductions:

Thm [Agrawal-Vinay'08]: Suppose $f \in \mathbb{F}[x_1, \dots, x_n]$ has degree $d = O(n)$.

If there is a circuit of size $2^{o(d + d\log(n/d))}$ computing $f$, then there is a $\Sigma\Pi\Sigma\Pi$-circuit of size $2^{o(d + d\log(n/d))}$ computing $f$.




In particular, $\exp(n)$-lower bound for $\Sigma\Pi\Sigma\Pi$ circuits

$\Rightarrow$ $\exp(n)$-lower bound for general circuits.

Thm [Koiran'10]: Suppose $f \in \mathbb{F}[x_1, \dots, x_n]$ is computed by a polynomial-size circuit, $d = \deg(f)$.

Then $f$'s computed by a $\Sigma\Pi\Sigma\Pi$ circuit of size $n^{O(\sqrt{d}\log d)}$.

In particular, $2^{n^{1/2+\varepsilon}}$ - lower bound for $\mathrm{PERM} = \sum_{\sigma \in S_n} \prod_{i=1}^{n} x_{i\sigma(i)}$ (in $n^2$ variables)

for any constant $\varepsilon > 0$

$\Rightarrow$ VP $\ne$ VNP. (We will show PERM $\in$ VNP)

$\sigma \in S_n \; i=1$

for any constant $\varepsilon > 0$ ) $\Rightarrow VP \neq VNP$. (We will show PERM $\in$ VNP).

Thm [Gupta - Kamath - Kayal - Saptharishi '14] : $2^{\Omega(n^{1/2})}$ - lower bound for PERM.

However, the same lower bound applies to DET ...