Lower Bounds

# Depth-3 over finite fields

- Reduction to depth-4 works for *any* $\mathbb{F}$.

- The one to depth-3, however, requires $\text{char}\,\mathbb{F} = \Omega(\sqrt{d})$ (in Ryser-Fischer's formula).

- Can we do reduction to depth-3 for small $\text{char}\,\mathbb{F} =: p$ ? NO:

Theorem (Grigoriev, Karpinski '98): Over the field $\mathbb{F}_q$, $\det_d$ (or $\text{per}_d$) requires depth-3 circuits of size $2^{\Omega_q(d)}$.

Rmk: If there was a reduction for $\det_d$ to depth-3, over $\mathbb{F}_q$, then the size would have been $d^{O(\sqrt{d})}$.

Proof: • Idea – $\mathbb{F}_q$ has $q$ elements. We will think of $q$ as fixed (i.e. constant wrt $d$).

• Let $C = \sum_{i\in[s]} T_i$ be a $\Sigma\Pi\Sigma$ circuit.

- Define $rk(T_i)$ to be the rank of the set of linear factors of $T_i$.
- Let $n := d^2$ & $\tau := \Theta_q(d)$ to be fixed later.
- A "low" rank $T_i$ (say $rk(T_i) \leq \frac{\tau}{10q}$) has low rank partial derivatives.

  A "high" rank $T_i$ $(rk(T_i) > \tau)$ we would like to zero out by picking a random evaluation in $\mathbb{F}_q^n$.
- These two together give us a matrix corresponding to the polynomial $C$.

$$M_k(C, \mathcal{A}) := \begin{array}{c} \\ \partial_\alpha \\ \\ \end{array} \underbrace{\begin{array}{c} \bar{a} \\ \left( \begin{array}{ccc} & & \\ & \partial_\alpha C(\bar{a}) & \\ & & \end{array} \right) \end{array}}_{\mathcal{A} \subseteq \mathbb{F}_q^n} \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} \partial^{=k}$$

where, $k := \tau/10q$

& $\mathcal{A}$ shall be the set of evaluations on which each derivative $\partial^{=k} T_i$, for high $rk(T_i)$, vanishes.

- Once $k, A$ are fixed we say that $\Gamma_{k,A}(f) := rk\, M_k(f, A)$ is a complexity measure (of polynomials).
- Obviously, we want to show $\Gamma_{k,A}(C)$ small & $\Gamma_{k,A}(\det_d)$ large.

Lemma 1 (Upper bound): $\forall \tau > 0$, $k \le \tau/10q$, there is a subset $\Sigma \subseteq \mathbb{F}_q^n$ of size $s \cdot e^{-\tau/8q} \cdot q^n$ s.t. for $A := \mathbb{F}_q^n \setminus \Sigma$, $\Gamma_{k,A}(C) < s \cdot q^\tau$.

Proof:

- To upper bound $\Gamma_{k,A}$ for $C$, it suffices to do it for $T_1$; because of subadditivity:

$$\Gamma(f+g) \le \Gamma(f) + \Gamma(g).$$ (Exercise)

- Let us now work with $T = \ell_1 \cdots \ell_D$.
- Case $[rk(T) \le \tau]$: Let $\{\ell_1, \ldots, \ell_r\}$ form a basis for $\{\ell_1, \ldots, \ell_D\}$.

Then $T$ is a $\mathbb{F}_q$-lr. combination of $M := \{\ell_1^{e_1} \cdots \ell_r^{e_r} \mid e_i < q,\ i \in [r]\}$, as long as we evaluate it over $\mathbb{F}_q^n$.

$\Rightarrow \forall A \subseteq \mathbb{F}_q^n,\ \Gamma_{k,A}(T) \leq |M| \leq q^r \leq q^\tau$.

• Case [$rk(T) > \tau$]: Now $r > \tau$ & $\ell_1, \ldots, \ell_r$ span $\{\ell_1, \ldots, \ell_D\}$.

For each nonconstant $\ell_i$, $i \in [r]$, we have $\Pr_{\bar{a} \in \mathbb{F}_q^n} [\ell_i(\bar{a}) = 0] = 1/q$.

$$\Rightarrow \mathop{E}_{\bar{a}} [\#i \in [r],\ \ell_i(\bar{a}) = 0] = r/q > \tau/q$$

$$\Rightarrow \Pr_{\bar{a}} \left[ \#\{i \mid \ell_i(\bar{a}) = 0\} < k = \frac{\tau}{10q} \right] < e^{-\tau/8q}$$

{ Exercise: Chernoff bounds
$$\Pr[X \gtrless (1 \pm \delta)\mu] < \left( \frac{e^{\pm\delta}}{(1\pm\delta)^{1\pm\delta}} \right)^{\mu}.$$ }

• Let $\mathcal{E}_T$ be the $\bar{a}$'s in the above "low" probability event. Then, $\bar{a} \notin \mathcal{E}_T$ makes $> k$ $\ell_i$'s zero in $T$.

$$\Rightarrow \forall \bar{a} \in \bigcup_{rk(T) > \tau} \mathcal{E}_T, \text{ every } \partial^{=k} T(\bar{a}) = 0.$$

$rk(T) > \tau$

$\Rightarrow \mathcal{E} := \bigcup_{rk(T)>\tau} \mathcal{E}_T$ has size $< s \cdot \varepsilon^{-\tau/8q} \cdot q^n$

& $A := \mathbb{F}_q^n \setminus \mathcal{E}$ zeroes out every function in $\partial^{=k} T$, for $T \in \{T_i \mid i \in [s], rk(T_i) > \tau\}$.

$\Rightarrow \Gamma_{k,A}(c)$ is contributed by only $T_i$'s with $rk(T_i) < \tau$

$$\Rightarrow \Gamma_{k,A}(c) < s \cdot q^{\tau}. \qquad \square$$

- Next, we understand the measure for $\det_d$ & $\mathrm{per}_d$, $n := d^2$.

<u>Lemma 2</u> (Lower bound): For any $A \subseteq \mathbb{F}_q^n$ of size $(1-o(1))q^n$, we have $\Gamma_{k,A}(\det_d) = \binom{d}{k}^2$.

<u>Proof:</u> (from Saptharishi's survey)

- We consider the rank of the matrix $M_k(\det_d, A)$.
- An order-$k$ derivative (partial) of $\det_d$ is, either zero, or an order-$(d-k)$ <u>minor</u>.

• Since $\det_d$ has $\binom{d}{d-k}^2$ many order-$(d-k)$ minors, it can be seen that the rank of $M_k(\det_d, \mathbb{F}_q^n) = \binom{d}{k}^2$.

[ We can pick a point $\bar{\alpha} \in \mathbb{F}_q^n$ s.t. the column $\bar{\alpha}$ has exactly one nonzero entry in $M_k(\det_d, \mathbb{F}_q^n)$ : a desired order-$(d-k)$ minor. Thus, we identify a "diagonal" matrix inside $M_k(\cdot,\cdot)$; lower bounding its rank. ]

• However, $M_k(\det_d, A)$ has, possibly, many columns missing. How do we lower bound its rank?

<u>Idea</u>- We study arbitrary linear combinations of its rows.

<u>Claim 1</u>: Let $f(\bar{x})$ be a $\mathbb{F}_q$-linear combination of $r \times r$ minors of $X = (x_{ij})$. Then,

$$\Pr_{\bar{\alpha} \in \mathbb{F}_q^n} [f(\bar{\alpha}) \neq 0] \geq \frac{1}{4}.$$

- This claim immediately implies that the rows of $M_k(\det_d, A)$, corresponding to the minors, are linearly independent;

Since the #zeros, of a linear combination of minors of $X$, is $\leq \frac{3}{4}q^n$ & so $|A| - \frac{3}{4}q^n = \frac{1}{4}q^n - o(1)q^n > 0$.

$\Rightarrow$ We only need to prove Claim 1.
First, we prove a base case:

<u>Claim 2:</u> $\Pr_{\bar{\alpha} \in \mathbb{F}_q^n}[\det_d(\bar{\alpha}) \neq 0] \geq 1/4$.

<u>Proof:</u> • The number of <u>invertible</u> matrices in $\mathbb{F}_q^{d\times d}$ is $(q^d-1)\cdot(q^d-q)\cdot\dots\cdot(q^d-q^{d-1})$.

$\Rightarrow$
$$\Pr_{\bar{\alpha}}[\det_d(\bar{\alpha}) \neq 0] = \left(1-\frac{1}{q}\right)\cdot\left(1-\frac{1}{q^2}\right)\cdots\left(1-\frac{1}{q^d}\right)$$
$$\geq \left(1-\frac{1}{2}\right)\left(1-\frac{1}{2^2}\right)\cdots\left(1-\frac{1}{2^d}\right) \geq \frac{1}{4}.$$
$\square$

<u>Exercise:</u> Prove Claim 2 for $\mathrm{per}_d$.

Pf of Clm 1: • Let the linear combination of the $r\times r$ minors of $\det_d(X)$ be

$$f(\bar{x}) = \sum_{\text{row-1 in } M_i} c_i \cdot M_i + \sum_{\substack{\text{row-1 not}\\ \text{in } M_j}} c_j \cdot M_j \,.$$

• We now want to further expand each $M_i$ by row-1 of $X$ & rearrange the first part of $f(\bar{x})$ above:

$$f(\bar{x}) = \sum_{i\in[d]} x_{1i} \cdot M_i' + M'' \,.$$

Now $M_i'$ are $\mathbb{F}_q$-linear combinations of certain order-$(r+1)$ minors of $\det_d(X)$.

$M''$ is "free" of $x_{1j}$ variables.

• Wlog we can assume that at least two distinct order-$r$ minors participated in defining $f(\bar{x})$, and that at least one of the $M_i'$ above is nonzero.

• We would like to pick a random $\bar{\alpha}$ by first picking the rows $\{2,..,d\}$ & picking the first row in the end (from $\mathbb{F}_q^d$).

- From this viewpoint it is clear that:

$$LHS = \Pr_{\bar{\alpha}}\left[ \sum_{i=1}^{d} \alpha_{1i} \cdot M_i'|_{\bar{\alpha}} + M''|_{\bar{\alpha}} \neq 0 \right]$$

$$\geq \Pr_{\bar{\alpha}}\left[ \sum_{i=1}^{d} \alpha_{1i} \cdot M_i'|_{\bar{\alpha}} \neq 0 \right]$$ (koutis' trick).

- The latter involves only the minors that have row-1 of X.
- Repeating this, several times, we end up with the probability estimate for a single minor (as in Clm 2).

$\Rightarrow LHS \geq 1/4$. $\square$

- As discussed before Clm 1 implies that $\Gamma_{k,A}(\det_d) = \binom{d}{k}^2$, finishing Lem 2. $\square$

Exercise: Prove the same for $per_d$.

- Assuming that $\det_d$ has a depth-3 circuit, we compare the bounds in Lemmas 1 & 2: let $\tau = \alpha d$, $k = \tau/10q$,

$$\binom{d}{k}^2 = \Gamma_{k,A}(\det_d) < s \cdot q^{\alpha d}$$

[Stirling's approx. gives: $\lg \binom{n}{\varepsilon n} = H_2(\varepsilon) \cdot n - O(\lg n)$,
where $H_2(\varepsilon) := -\varepsilon \lg \varepsilon - (1-\varepsilon) \lg (1-\varepsilon)$. ]
[Eg. it follows that $\binom{n}{\varepsilon n} = 2^{\Omega_\varepsilon(n)}$. ]

$$\Rightarrow \lg \binom{d}{k}^2 = \Omega\left(d \cdot H_2(k/d)\right) = \Omega\left(d \cdot H_2(\alpha/10q)\right)$$

$$\Rightarrow \lg s = \Omega\left(d H_2\left(\frac{\alpha}{10q}\right)\right) - \alpha d \lg q$$

$$\Rightarrow \lg s/d = \Omega\left(\frac{\alpha}{10q} \lg \frac{10q}{\alpha} + \left(1 - \frac{\alpha}{10q}\right) \lg \frac{10q}{10q-\alpha}\right) - \alpha \cdot \lg q$$

- Thus, there is some constant $c > 0$ s.t. it suffices to pick $\alpha$ satisfying

$$\lg \frac{10q}{\alpha} > cq \cdot \lg q$$

$\Leftrightarrow \alpha < 10q/q^{cq}$. Thus, $\tau = O(d/q^{cq-1})$.

- For constant $q$, $\tau$ makes sense & we get a lower bound on the top fanin:

$$\lg s = \Omega_q(d)$$

finishing the theorem. $\square$

- The lower bound can be improved by considering a sum of elementary symmetric polynomials on $n = d^2$ variables & $\deg \le d$.

Define $sym_{\le d} := \sum_{S \in \binom{[n]}{\le d}} x_S$ .

- It can be shown that the rank of the matrix $M_k(sym_{\le d}, \mathbb{F}_q^n) \geqslant \binom{n}{d/2}$, for $k = d/2$.

- This gives $s = n^{\Omega_q(d)}$.