Multilinear models

- Multilinearity is a natural restriction on arithmetic circuits.

  Eg. det & per are multilinear polynomials. Can they be computed without computing monomials like x_i^j, $j>1$?

Defn: A circuit C is multilinear if every gate computes a multilinear polynomial.

- We will first focus on multilinear depth-3 & prove exponential lower bounds.

- The measure, on polynomials, that will succeed is as follows.

  Partition the variables X as $Y \sqcup Z$. For a monomial m_y (resp. m_z) in the Y (resp. Z) variables, let $coef(m_y \cdot m_z)(f)$ denote the coefficient of $m_y m_z$ in $f(X)$.

Define matrix $M_{Y,Z}(f)$ as:

$$\begin{array}{cc} & m_Z \\ m_Y & \left.\underbrace{\begin{pmatrix} & & \\ & \mathrm{coef}(m_Y m_Z)(f) & \\ & & \end{pmatrix}}_{\text{monomials in } Z}\right\} \text{monomials in } Y \end{array}$$

<u>Defn</u>: $\Gamma_{Y,Z}(f) := \mathrm{rk}\, M_{Y,Z}(f)$.

For multilinear f, $\mathrm{coef}(m_Y m_Z)(f) = (\partial_{m_Y m_Z} f)(\overline{0})$. So, $M_{Y,Z}(f)$ is also called the partial derivative matrix of f.

– It behaves well under ring operations:

<u>Lemma</u> (Sub-additivity): $\Gamma_{Y,Z}(f_1+f_2) \leq \Gamma(f_1) + \Gamma(f_2)$.

<u>Pf</u>: Follows from the rank property of $A+B$ for matrices. $\square$

<u>Lemma</u> (Multiplicativity): For $f_1 \in \mathbb{F}[Y_1, Z_1]$, $f_2 \in \mathbb{F}[Y_2, Z_2]$ with $Y = Y_1 \sqcup Y_2$ & $Z = Z_1 \sqcup Z_2$,

we have $\Gamma_{Y,Z}(f_1 \cdot f_2) = \Gamma_{Y_1,Z_1}(f_1) \cdot \Gamma_{Y_2,Z_2}(f_2)$.

Proof:

- Note that $M_{Y,Z}(f_1 \cdot f_2)$ equals the tensor product $M_{Y_1,Z_1}(f_1) \otimes M_{Y_2,Z_2}(f_2)$.

- This follows from the disjointness of these subsets, which allows $\mathrm{coef}(m_Y m_Z)(f) = \mathrm{coef}(m_{Y_1} m_{Z_1})(f) \cdot \mathrm{coef}(m_{Y_2} m_{Z_2})(f)$.

- Rank property of a tensor product gives $\Gamma_{Y,Z}(f) = \prod_{i \in [2]} \Gamma_{Y_i,Z_i}(f_i)$. $\square$

Lemma (Mult. by Z-free): For any $g \in \mathbb{F}[Y]^*$, $\Gamma_{Y,Z}(g \cdot f) = \Gamma_{Y,Z}(f)$.

Proof:

- If we consider the function field $\mathbb{F}(Z)$, then g, f can be considered as polynomials in $\mathbb{F}(Z)[Y]$.

- By "coefficient" extraction we can prove:

$$\Gamma_{Y,Z}(gf) = rk_{\mathbb{F}}\{(\partial_m(gf))_{Z=0} \mid m \text{ is a monomial in } Z\}$$

result in $\mathbb{F}[Y]$.

[Pf idea: column m_Z in $M_{Y,Z}(gf)$ exactly represents the polynomial in $\mathbb{F}[Y]$ which is the "coefficient" of m_Z in (gf).

$\Rightarrow$ the rank of these Y-polynomials equals the column rank of $M_{Y,Z}(gf)$.]

- Now, using the Z-freeness of g, we get:

$$\Gamma_{Y,Z}(gf) = rk_{\mathbb{F}}\{g\cdot(\partial_m f)_{Z=0} \mid m \text{ is a monomial in } Z\}$$

$$= \Gamma_{Y,Z}(f). \qquad \square$$

<u>Lemma</u>: For any multilinear f, we have

$$\Gamma_{Y,Z}(f) \leq 2^{\min(|Y|,|Z|)}.$$

<u>Proof</u>:

- Follows from the size of $M_{Y,Z}(f)$. $\square$

- Eg. $f(Y,Z) = \prod_{i\in[n]} (y_i + z_i)$ proves the optimality of the above upper bound as:
$$\Gamma_{Y,Z}(f) = \prod_{i\in[n]} \Gamma_{y_i,z_i}(y_i + z_i) = 2^n.$$

- The above example is merely depth-2! Raz showed that the measure for $\Pi\Sigma$ can be significantly reduced if we consider a random partition of X.

Theorem (Upper bound) [Raz'09]: Let $f(X) = \ell_1 \cdots \ell_d$ be an n-variate multilinear. For a random partition $X = Y \sqcup Z$, $|Y| = |Z| = n/2$, we have whp:
$$\Gamma_{Y,Z}(f) \leq 2^{n/2 - n/32}.$$

Proof: • Wlog each ℓ_i has support-size ≥ 2, as univariate ℓ_i's do not change the measure wrt <u>any</u> partition.

(by multiplicativity)

- Clearly, $T_{Y,Z}(f) \leq 2^d$. Hence, we are done if $d < n/3$.

Assume $d \geqslant n/3$.

- Since l_i's are disjoint-support & many ($\geqslant n/3$) we get by an averaging argument that:

l_i's with support-size 2 or 3 is $\geqslant d/4$.

[Otherwise, $> 3d/4$ l_i's have support-size $\geqslant 4 \Rightarrow n > 3d$, a #.]

- We call these l_i's <u>small</u>.
- Now for a small l_i, we have

$$\Pr_{Y,Z}[\text{support}(l_i) \subseteq Y \text{ or } Z] \geqslant 2 \cdot \frac{1}{2^3} = \frac{1}{4}.$$

$$\Rightarrow \operatorname{Exp}_{Y,Z}[\#i \mid \text{small } l_i \text{ is in } \mathbb{F}[Y], \mathbb{F}[Z]] \geqslant d/16.$$

- These l_i's stop contributing to $T_{Y,Z}(f)$.