On $\Sigma \wedge \Sigma \wedge \Sigma$ Circuits: The Role of Middle Σ Fan-in, Homogeneity and Bottom Degree

Chrisitan Engels Raghavendra Rao B V Karteek Sreenivasaiah

FCT 2017

Definitions

Definition

Arithmetic Circuit over $\langle \mathbb{K}, +, \times \rangle$

Directed acyclic graph C where nodes are labelled with $\{+, \times, x_1, \ldots, x_n\} \cup \mathbb{K}$.

Definitions

Definition

Arithmetic Circuit over $\langle \mathbb{K}, +, \times \rangle$

Directed acyclic graph C where nodes are labelled with $\{+, \times, x_1, \ldots, x_n\} \cup \mathbb{K}$.

- A node of out-degree zero, called output node of the circuit

Definitions

Definition

Arithmetic Circuit over $\langle \mathbb{K}, +, \times \rangle$

Directed acyclic graph C where nodes are labelled with $\{+, \times, x_1, \ldots, x_n\} \cup \mathbb{K}$.

- A node of out-degree zero, called output node of the circuit
- $\{x_1, \cdots, x_n\}$ are the inputs for the circuit, where $x_i \in \mathbb{K}$

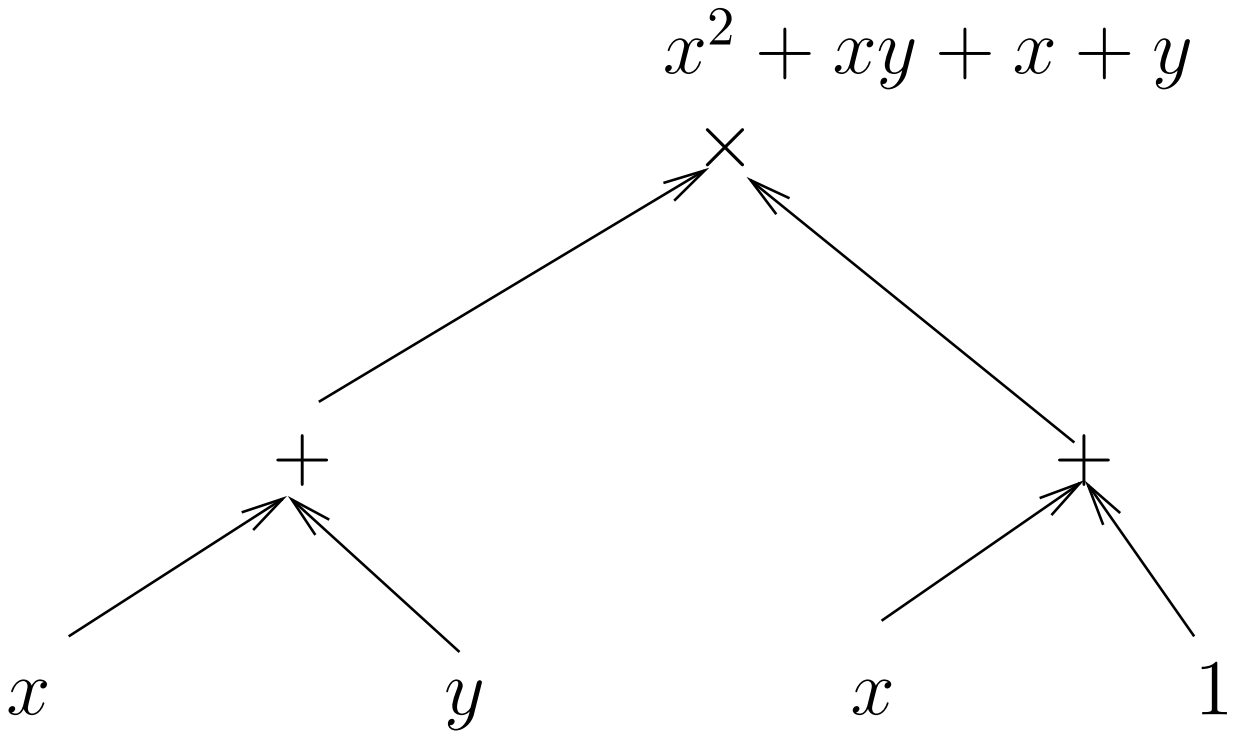


Figure: An arithmetic circuit, computing the polynomial $x^2 + xy + x + y$

Resource Measures

- size: Number of nodes and edges in the circuit.

Resource Measures

- size: Number of nodes and edges in the circuit.
- depth - length of longest path from an input node to the output node

Resource Measures

- size: Number of nodes and edges in the circuit.
- depth - length of longest path from an input node to the output node

These parameters are generally measured in terms of the number of variables.

Conjecture (Valiant's Hypothesis)

For infonitely many $n \geq 0$ the polynomial

$$perm_n = \sum_{\sigma \in S_n} \prod_i x_{i,\sigma(i)}$$

does not have polynomial size arithmetic circuits.

Depth : Are shallow circuits powerful?

- Poly size circuits computing polynomials of poly degree $=$ log depth circuits with unbounded Σ fan in [Valiant Skyum Berkowitz Rackoff 1981]

Depth : Are shallow circuits powerful?

- Poly size circuits computing polynomials of poly degree = log depth circuits with unbounded Σ fan in [Valiant Skyum Berkowitz Rackoff 1981]
- Poly size circuits computing polynomials of degree $d \subseteq$ Depth 4 $\Sigma\Pi\Sigma\Pi$ circuits of size $n^{\sqrt{d}}$ [Agrawal-Vinay 2008, the best bound by [Tavenas 2013].

Depth : Are shallow circuits powerful?

- Poly size circuits computing polynomials of poly degree = log depth circuits with unbounded Σ fan in [Valiant Skyum Berkowitz Rackoff 1981]
- Poly size circuits computing polynomials of degree $d \subseteq$ Depth 4 $\Sigma\Pi\Sigma\Pi$ circuits of size $n^{\sqrt{d}}$ [Agrawal-Vinay 2008, the best bound by [Tavenas 2013].
- Poly size circuits computing polynomials of degree $d \subseteq$ Depth 4 $\Sigma\Pi\Sigma$ circuits of size $n^{\sqrt{d}}$ over large fields.[Gupta Kamat Kayal Saptharishi 2013]

Constant depth circuits with powering gates

- Powering gate $\wedge^i g$ computes the polynomial g^i.

Constant depth circuits with powering gates

- Powering gate $\wedge^i g$ computes the polynomial g^i.
- Bounded fain-in $\times$ gates can be replaced with $\wedge$ gates: $f \cdot g = ((f+g)^2 - (f-g)^2)/4$.

Question

Convert Π *gates of unbounded fan-in to circuit with only* $\wedge$ *and* Σ *gates?*

Constant depth circuits with powering gates

- Powering gate $\wedge^i g$ computes the polynomial g^i.
- Bounded fain-in $\times$ gates can be replaced with $\wedge$ gates: $f \cdot g = ((f+g)^2 - (f-g)^2)/4$.

Question

Convert Π *gates of unbounded fan-in to circuit with only* $\wedge$ *and* Σ *gates?*

Fischer's Identity

Theorem (Fischer 94)

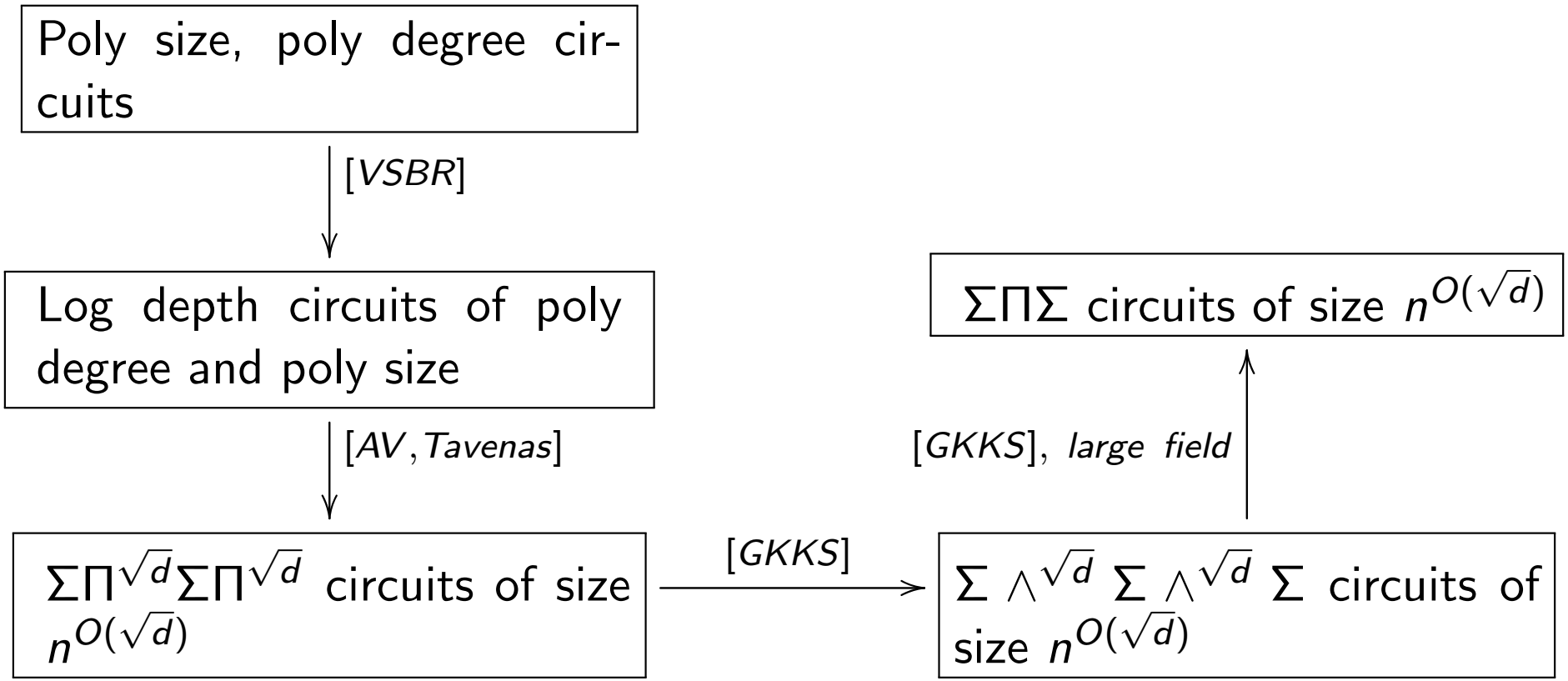
There are homogeneous linear forms $\ell_1, \ell_2, \ldots, \ell_{2^n}$ such that

$$x_1 \cdot x_2 \cdots x_n = \sum_{i=1}^{2^n} \ell_i^n.$$

Corollary

A polynomial computable by a $\Sigma\Pi^k\Sigma\Pi^k\Sigma$ circuit of size s can be computed by as $\Sigma \wedge^k \Sigma \wedge^k \Sigma$ circuit of size $s \cdot 2^k$.

Depth five circuits with $\wedge$ gates



Lower bounds against shallow circuits

- Any homogeneous $\Sigma\Pi^{\sqrt{n}}\Sigma\Pi^{\sqrt{n}}$ circuit computing permanent requires size $2^{\Omega\sqrt{n}}$. [Gupta et al 13, extended to other polynomials later.]
- A $\omega(\log n)$ factor improvement in the above would resolve Valiant's hypothesis.
- Best known lower bound against $\Sigma\Pi\Sigma$ circuits over infinite fields is $\Omega(n^3/(\log n)^2)$ [Kayal - Saha - Tavenas]
- No known lower bounds against depoth five circuits with powering gates.

Our Results

Theorem (1)

Let $g = \sum_{i=1}^{s} f_i^{\alpha_i}$ *where* $f_i = \ell_{i_1}^{d_i} + \cdots + \ell_{i_n}^{d_i} + \beta_i$ *for some scalars* β_i *and for every* i*, either* $d_i = 1$ *or* $d_i \geq 21$ *and* $\ell_{i_1}, \ldots, \ell_{i_n}$ *are homogeneous linear forms. If* $g = x_1 \cdot x_2 \cdots x_n$ *then* $s = 2^{\Omega(n)}$.

Our Results

Theorem (1)

Let $g = \sum_{i=1}^{s} f_i^{\alpha_i}$ *where* $f_i = \ell_{i_1}^{d_i} + \cdots + \ell_{i_n}^{d_i} + \beta_i$ *for some scalars* β_i *and for every* i*, either* $d_i = 1$ *or* $d_i \geq 21$ *and* $\ell_{i_1}, \ldots, \ell_{i_n}$ *are homogeneous linear forms. If* $g = x_1 \cdot x_2 \cdots x_n$ *then* $s = 2^{\Omega(n)}$.

Theorem (2)

Let $g = \sum_{i=1}^{s} f_i^{\alpha_i}$ *where* $f_i = \sum_{j=1}^{N_i} \ell_{i_j}^{d_i} + \beta_i$*, for some scalars* β_i *and* $\sqrt{n} \leq d_i \leq n$*,* $N_i \leq 2^{\sqrt{n}/1000}$*, and* $\ell_{i_1}, \ldots, \ell_{i_{N_i}}$ *are homogeneous linear forms. If* $g = x_1 \cdot x_2 \cdots x_n$ *then* $s = 2^{\Omega(n)}$.

Proof approach

- Obtain a measure $\mu : \mathbb{F}[x_1, \ldots, x_n] \to \mathbb{R}$ such that

$$\mu(f_1 + \cdots + f_s) \leq \mu(f_1) + \cdots + \mu(f_s)$$

Proof approach

- Obtain a measure $\mu : \mathbb{F}[x_1, \ldots, x_n] \to \mathbb{R}$ such that

$$\mu(f_1 + \cdots + f_s) \leq \mu(f_1) + \cdots + \mu(f_s)$$

For a polynomial $f_i \in \wedge\Sigma \wedge \Sigma$, assume that $\mu(f_i) \leq t$.

Proof approach

- Obtain a measure $\mu : \mathbb{F}[x_1, \ldots, x_n] \to \mathbb{R}$ such that

$$\mu(f_1 + \cdots + f_s) \leq \mu(f_1) + \cdots + \mu(f_s)$$

For a polynomial $f_i \in \wedge\Sigma \wedge \Sigma$, assume that $\mu(f_i) \leq t$. Then

$$\mu(f_1 + \ldots + f_s) \leq s \cdot t.$$

Additionally, if $\mu(g) \geq R$ for some polynomial g we have,

$$s \geq R/t.$$

Our Measure: Projected Multilinear derivatives

Let $f \in \mathbb{F}[x_1, \ldots, x_n]$.

- $S \subseteq \{x_1, \ldots, x_n\}$, let $\pi_S : \mathbb{F}[x_1, \ldots, x_n] \to \mathbb{F}[x_1, \ldots, x_n]$ be the projection map that sets all variables in S to zero.
- Let $\pi_{\mathsf{m}}(f)$ denote the projection of f onto its multilinear monomials

Definition

For $S \subseteq \{1, \ldots, n\}$ and $0 < k \leq n$, the dimension of Projected Multilinear Derivatives (PMD) of a polynomial f is defined as:

$$\mathrm{PMD}_S^k(f) \triangleq \dim(\mathbb{F}\text{-Span}\left\{\pi_S(\pi_{\mathsf{m}}(\partial^{=k}_{\mathsf{ML}} f))\right\}).$$

Hard polynomial

Lemma

For any $S \subseteq \{x_1, \ldots, x_n\}$, $|S| = n/2 + 1$, *and* $k = 3n/4$

$$\mathrm{PMD}^k_S(x_1 \ldots x_n) \geq \binom{n/2-1}{n/4} = 2^{\Omega(n)}.$$

Structure of projected multilinear derivatives

Lemma

Suppose that $f = (\ell_1^d + \ldots + \ell_n^d + \beta)$.
Let $Y = \{\ell_i^{d-j} \mid 1 \leq i \leq n, 1 \leq j \leq d\}$ *and* $\lambda = 1/4 + \varepsilon$ *for some* $0 < \varepsilon < 1/4$. *Then, for* $k = 3n/4$ *and any* $S \subseteq \{1, \ldots, n\}$ *with* $|S| = n/2 + 1$, *we have:*

$$\pi_S(\pi_m(\partial_{ML}^{=k} f^\alpha) \subseteq \mathbb{F}\text{-Span}\left\{\pi_S(\pi_m(\mathcal{F} \odot \left(\mathcal{X}_{\lambda n}^{n/2-1}(\overline{S}) \cup \mathcal{M}_{\leq(1+\varepsilon)n/d}(Y)\right)))\right\}$$

where $\mathcal{F} = \cup_{i=1}^{k} f^{\alpha-i}$ *and* $\overline{S} = \{1, \ldots, n\} \setminus S$.

An upper bound for the measure

- By Lemma,
  $\mathrm{PMD}^k_S(f^\alpha) \leq k \cdot (|\mathcal{X}^{n/2-1}_{\lambda n}(\overline{S})| + |\mathcal{M}_{\leq(1+\varepsilon)n/d}(Y)|)$.
- For $1/4 < \lambda < 1/2$,

$$|\mathcal{X}^{n/2-1}_{\lambda n}(\overline{S})| \leq O(n/2 \cdot \binom{n/2}{\lambda n}) \leq 2^{.498n}.$$

- Also,

$$|\mathcal{M}_{\leq(1+\varepsilon)n/d}(Y)| = \binom{|Y| + (1+\epsilon)n/d}{(1+\epsilon)n/d} \leq 2^{.4995n} \text{ for } d \geq 21.$$

- Therefore, $\mathrm{PMD}^k_S(f^\alpha) \leq 2^{.4995n}$.

Further work

- Chillara and Saptharishi Simplified the arguments and generalized to non-homogeneous circuits.
- Theorem (1) holds for $d \geq 10$.

Future Directions

- Obtain lower bound for non-homogeneous $\Sigma \wedge \Sigma \wedge \Sigma$ circuits.
- Obtain a complexity measure μ for polynomial such that $\mu(f^{\alpha}) \leq poly(\mu(f))$.

Thank You!!