Depth reduction of arithmetic ckts to depth three

A chasm at depth three

Mohith Raju N

April 2022

Depth reduction to depth three

- Constant depth circuits are circuits whose depth is bounded by a constant

Depth reduction to depth three

- Constant depth circuits are circuits whose depth is bounded by a constant
- We usually allow unbounded fanin and due to which we can assume alternating sum and product gates

Depth reduction to depth three

- Constant depth circuits are circuits whose depth is bounded by a constant
- We usually allow unbounded fanin and due to which we can assume alternating sum and product gates
- $\sum\prod\sum$ represents a depth three circuit

Depth reduction to depth three

- Constant depth circuits are circuits whose depth is bounded by a constant
- We usually allow unbounded fanin and due to which we can assume alternating sum and product gates
- $\sum\prod\sum$ represents a depth three circuit

Main Theorem. Let $f(x) \in \mathbb{Q}[x]$ be an n-variate polynomial of degree $d = n^{\mathcal{O}(1)}$ computed by an arithmetic circuit of size s. Then it can also be computed by a $\sum\prod\sum$ circuit of size $2^{\mathcal{O}(\sqrt{d \log n \log d \log s})}$

Remarks:

- By size we mean the number of edges in the circuit
- The intermediate polynomials have degree much higher than d

Why do we care about depth reductions

- Circuits with low depth correspond to computations which are highly parallelizable and therefore it is natural to try to minimize the depth of a circuit while allowing the size to increase somewhat
- Lower bounds for constant depth circuits imply lower bounds for general circuits thanks to depth reduction results

Example.

Given an explicit family of polynomials f_n, a $2^{\Omega(d \log n)}$ lower bound for $\sum\prod\sum$ circuits computing f_n implies a $2^{\Omega\left(\frac{d \log n}{\log d}\right)}$ lower bound for general arithmetic circuits computing f_n

How to depth reduce

Preliminaries

- Powering circuits are those which contain exponentiation gates, denoted by $\wedge$. Such a gate has all incoming edges coming from a single input x and computes x^n where n is the number of incoming nodes from x
- Exponentiation gate is just a product gate with n incoming edges all coming from the same input
- Exponentiation gate is a "weaker" product gate as it can only compute a specific type of product
- One can think of an Algebraic Branching Program (ABP) as a special type of a circuit
- Small lemma. For any n, k

$$\binom{n+k}{k} = \mathcal{O}\left(e \cdot \frac{n+k}{k}\right)^k = 2^{\mathcal{O}(k \log n)}$$

Overview

- **Step 0:** General ckts $\longrightarrow$ ABPs
- **Step 1:** ABPs $\longrightarrow \sum\prod^{[a]}\sum\prod^{[d/a]}$ ckts
- **Step 2:** $\sum\prod^{[a]}\sum\prod^{[d/a]}$ ckts $\longrightarrow \sum\bigwedge^{[a]}\sum\bigwedge^{[d/a]}\sum$ ckts
- **Step 3:** $\sum\bigwedge^{[a]}\sum\bigwedge^{[d/a]}\sum$ ckts $\longrightarrow \sum\prod\sum$ $\mathbb{C}$-ckts
- **Step 4:** $\sum\prod\sum$ $\mathbb{C}$-ckts $\longrightarrow \sum\prod\sum$ $\mathbb{Q}$-ckts

Overview

- **Step 0:** General ckts $\longrightarrow$ ABPs
- **Step 1:** ABPs $\longrightarrow \sum\prod^{[a]}\sum\prod^{[d/a]}$ ckts
- **Step 2:** $\sum\prod^{[a]}\sum\prod^{[d/a]}$ ckts $\longrightarrow \sum\bigwedge^{[a]}\sum\bigwedge^{[d/a]}\sum$ ckts
- **Step 3:** $\sum\bigwedge^{[a]}\sum\bigwedge^{[d/a]}\sum$ ckts $\longrightarrow \sum\prod\sum$ $\mathbb{C}$-ckts
- **Step 4:** $\sum\prod\sum$ $\mathbb{C}$-ckts $\longrightarrow \sum\prod\sum$ $\mathbb{Q}$-ckts

The a in $\prod^{[a]}$ denotes the maximum fanin of any gate in this layer of multiplication gates

Step 0: General ckts $\longrightarrow$ ABPs

Lemma III.1. Let f be a polynomial of degree d computed by a circuit of size s. Then there is a homogeneous ABP of depth d and size $2^{\mathcal{O}(\log s \cdot \log d)}$ computing f

What we shall prove next

Theorem I.1 Let $f(x) \in \mathbb{Q}[x]$ be an n-variate polynomial of degree $d = n^{\mathcal{O}(1)}$ computed by an ABP of size s. Then it can also be computed by a $\sum\prod\sum$ circuit of size $2^{\mathcal{O}(\sqrt{d \log n \log s})}$

Step 1: ABPs $\longrightarrow \sum\prod^{[a]}\sum\prod^{[d/a]}$ ckts

Theorem IV.1 ([Koi12]). Let f be an n-variate polynomial of degree d computed by an ABP of size s. Then, for all a there is an equivalent homogeneous $\sum\prod^{[a]}\sum\prod^{[d/a]}$ circuit computing f of size $s^a + s^2 d \cdot \binom{n+d/a}{d/a}$

Step 1: ABPs $\longrightarrow \sum\prod^{[a]}\sum\prod^{[d/a]}$ ckts

Theorem IV.1 ([Koi12]). Let f be an n-variate polynomial of degree d computed by an ABP of size s. Then, for all a there is an equivalent homogeneous $\sum\prod^{[a]}\sum\prod^{[d/a]}$ circuit computing f of size $s^a + s^2 d \cdot \binom{n+d/a}{d/a}$

- After applying the small lemma, the above size becomes $2^{a\log s} + s^2 d \cdot 2^{d/a \log n}$
- To minimize the quantity we choose $\sqrt{\dfrac{d\log n}{\log s}}$
- $2^{a\log s} + s^2 d \cdot 2^{d/a \log n} = 2^{\mathcal{O}(\sqrt{d\log n \log s})}$

Overview

Progress so far

$$(\mathrm{ABP}, s) \longrightarrow \left(\sum\prod^{[a]}\sum\prod^{[d/a]}, s_1 = 2^{\mathcal{O}(\sqrt{d \log n \log s})}\right)$$

Overview of steps

- **Step 0:** General ckts $\longrightarrow$ ABPs
- **Step 1:** ABPs $\longrightarrow \sum\prod^{[a]}\sum\prod^{[d/a]}$ ckts
- **Step 2:** $\sum\prod^{[a]}\sum\prod^{[d/a]}$ ckts $\longrightarrow \sum\bigwedge^{[a]}\sum\bigwedge^{[d/a]}\sum$ ckts
- **Step 3:** $\sum\bigwedge^{[a]}\sum\bigwedge^{[d/a]}\sum$ ckts $\longrightarrow \sum\prod\sum$ $\mathbb{C}$-ckts
- **Step 4:** $\sum\prod\sum$ $\mathbb{C}$-ckts $\longrightarrow \sum\prod\sum$ $\mathbb{Q}$-ckts

Step 2: $\sum\prod^{[a]}\sum\prod^{[d/a]}$ ckts $\longrightarrow \sum\bigwedge^{[a]}\sum\bigwedge^{[d/a]}\sum$ ckts

What we shall do is,

$$\sum\prod^{[a]}\sum\prod^{[d/a]} \longrightarrow \sum\left(\sum\bigwedge^{[a]}\sum\right)\sum\left(\sum\bigwedge^{[d/a]}\sum\right)$$

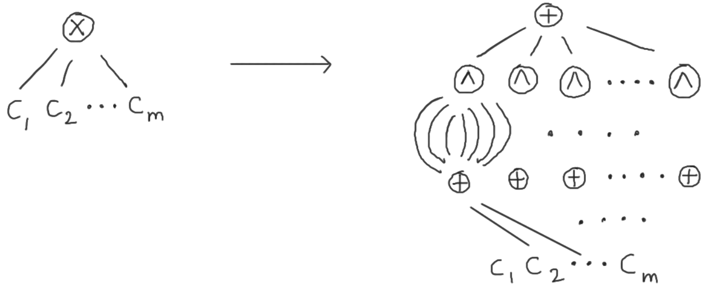
Lemma IV.3 (Fischer's trick). For any n, the monomial $x_1\cdots x_n$ can be expressed as a linear combination of 2^{n-1} powers of linear forms through the following:

$$n!\cdot x_1\cdots x_n = \sum_{S\subseteq[n]}(-1)^{n-|S|}\left(\sum_{i\in S}x_i\right)^n$$

Step 2: $\sum\prod^{[a]}\sum\prod^{[d/a]}$ ckts $\longrightarrow \sum\bigwedge^{[a]}\sum\bigwedge^{[d/a]}\sum$ ckts

Every multiplication gate computes $\prod_{i=1}^{m} C_i$. Using Fischer's trick we replace it as follows,

$$C_1 \cdots C_m = \sum_{S\subseteq[m]} \frac{(-1)^{m-|S|}}{m!}\left(\sum_{i\in S} C_i\right)^m$$



Step 2: $\sum\prod^{[a]}\sum\prod^{[d/a]}$ ckts $\longrightarrow \sum\bigwedge^{[a]}\sum\bigwedge^{[d/a]}\sum$ ckts

Observe:

- A product gate with fanin a is replaced with an exponentiation gate of fanin a. Thus $\prod^{[a]} \longrightarrow \sum\bigwedge^{[a]}\sum$
- Replacing one product gate as shown will increase size of ckt by $2^m + m \cdot 2^m + m \cdot 2^m = 2^{\mathcal{O}(m)}$

Step 2: $\sum\prod^{[a]}\sum\prod^{[d/a]}$ ckts $\longrightarrow \sum\bigwedge^{[a]}\sum\bigwedge^{[d/a]}\sum$ ckts

Observe:

- A product gate with fanin a is replaced with an exponentiation gate of fanin a. Thus $\prod^{[a]} \longrightarrow \sum\bigwedge^{[a]}\sum$
- Replacing one product gate as shown will increase size of ckt by $2^m + m\cdot 2^m + m\cdot 2^m = 2^{\mathcal{O}(m)}$
- Replacing every product gate in $\prod^{[a]}$ layer will increase size of circuit by $s_1\cdot 2^{\mathcal{O}(a)}$
- Replacing every product gate in $\prod^{[d/a]}$ layer will increase size of circuit by $s_1\cdot 2^{\mathcal{O}(d/a)}$
- Total increase is

$$\begin{aligned} &s_1\cdot(2^{\mathcal{O}(a)}+2^{\mathcal{O}(d/a)}) \\ &\qquad = 2^{\mathcal{O}(\sqrt{d\log n\log s})}\cdot\left(2^{\mathcal{O}\left(\sqrt{\frac{d\log n}{\log s}}\right)}+2^{\mathcal{O}\left(\sqrt{\frac{d\log s}{\log n}}\right)}\right) \\ &\qquad = 2^{\mathcal{O}(\sqrt{d\log n\log s})} \end{aligned}$$

Overview

Progress so far

$$(\mathrm{ABP}, s) \longrightarrow \left(\sum\prod^{[a]}\sum\prod^{[d/a]}, s_1 = 2^{\mathcal{O}(\sqrt{d\log n\log s})}\right)$$
$$\longrightarrow \left(\sum\bigwedge^{[a]}\sum\bigwedge^{[d/a]}\sum, s_2 = 2^{\mathcal{O}(\sqrt{d\log n\log s})}\right)$$

Overview of steps

- **Step 0:** General ckts $\longrightarrow$ ABPs
- **Step 1:** ABPs $\longrightarrow \sum\prod^{[a]}\sum\prod^{[d/a]}$ ckts
- **Step 2:** $\sum\prod^{[a]}\sum\prod^{[d/a]}$ ckts $\longrightarrow \sum\bigwedge^{[a]}\sum\bigwedge^{[d/a]}\sum$ ckts
- **Step 3:** $\sum\bigwedge^{[a]}\sum\bigwedge^{[d/a]}\sum$ ckts $\longrightarrow \sum\prod\sum$ $\mathbb{C}$-ckts
- **Step 4:** $\sum\prod\sum$ $\mathbb{C}$-ckts $\longrightarrow \sum\prod\sum$ $\mathbb{Q}$-ckts

Step 3: $\sum \bigwedge^{[a]} \sum \bigwedge^{[d/a]} \sum$ ckts $\longrightarrow \sum \prod \sum$ $\mathbb{C}$-ckts

What we shall do is

$$\bigwedge^{[a]} \sum \longrightarrow \sum \prod^{[s_2]} E$$

$$\sum \bigwedge^{[a]} \sum \bigwedge^{[d/a]} \sum \longrightarrow \sum \left(\sum \prod^{[s_2]} E \right) \bigwedge^{[d/a]} \sum$$

Step 3: $\sum \bigwedge^{[a]} \sum \bigwedge^{[d/a]} \sum$ ckts $\longrightarrow \sum \prod \sum$ $\mathbb{C}$-ckts

What we shall do is

$$\bigwedge^{[a]} \sum \longrightarrow \sum \prod^{[s_2]} E$$

$$\sum \bigwedge^{[a]} \sum \bigwedge^{[d/a]} \sum \longrightarrow \sum \left(\sum \prod^{[s_2]} E \right) \bigwedge^{[d/a]} \sum$$

$$E \bigwedge^{[d/a]} \longrightarrow \prod \sum\nolimits_{\mathbb{C}}$$

$$\sum \prod^{[s_2]} E \bigwedge^{[d/a]} \sum \longrightarrow \sum \prod^{[s_2]} \left(\prod \sum\right) \sum\nolimits_{\mathbb{C}}$$

(informal)

Step 3: $\sum \bigwedge^{[a]} \sum \bigwedge^{[d/a]} \sum$ ckts $\longrightarrow \sum \prod \sum$ $\mathbb{C}$-ckts

Lemma IV.6 (Saxena's duality trick). For every $m, d > 0$ and distinct $\alpha_1, \ldots, \alpha_{md+1} \in \mathbb{Q}$, there exists $\beta_1, \ldots, \beta_{md+1} \in \mathbb{Q}$ such that

$$(u_1 + \cdots + u_m)^d = \sum_{i=1}^{md+1} \beta_i \prod_{j=1}^{m} E_d(\alpha_i \cdot u_j)$$

where $E_d(u) := 1 + \frac{u}{1!} + \cdots + \frac{u^d}{d!}$

Step 3: $\sum \bigwedge^{[a]} \sum \bigwedge^{[d/a]} \sum$ ckts $\longrightarrow \sum \prod \sum$ $\mathbb{C}$-ckts

Lemma IV.6 (Saxena's duality trick). For every $m, d > 0$ and distinct $\alpha_1, \ldots, \alpha_{md+1} \in \mathbb{Q}$, there exists $\beta_1, \ldots, \beta_{md+1} \in \mathbb{Q}$ such that

$$(u_1 + \cdots + u_m)^d = \sum_{i=1}^{md+1} \beta_i \prod_{j=1}^{m} E_d(\alpha_i \cdot u_j)$$

where $E_d(u) := 1 + \frac{u}{1!} + \cdots + \frac{u^d}{d!}$

Proof: Let $l := (u_1 + \cdots + u_m)$

Note, $e^{lz} = 1 + \frac{l}{1!}z + \cdots + \frac{l^d}{d!}z^d + \cdots$

Hence,

$$\begin{aligned} l^d &= d! \cdot (\text{coeff of } z^d \text{ in } e^{lz}) \\ &= d! \cdot (\text{coeff of } z^d \text{ in } e^{u_1 z} \cdot e^{u_2 z} \cdots e^{u_m z}) \\ &= d! \cdot (\text{coeff of } z^d \text{ in } E_d(u_1 z) \cdot E_d(u_2 z) \cdots E_d(u_m z)) \end{aligned}$$

Step 3: $\sum \bigwedge^{[a]} \sum \bigwedge^{[d/a]} \sum$ ckts $\longrightarrow \sum \prod \sum$ $\mathbb{C}$-ckts

Proof cont'd: Let $l := (u_1 + \cdots + u_m)$

$$l^d = d! \cdot (\text{coeff of } z^d \text{ in } E_d(u_1 z) \cdot E_d(u_2 z) \cdots E_d(u_m z))$$

- Now define $F(z) := E_d(u_1 z) \cdot E_d(u_2 z) \cdots E_d(u_m z)$ to be a univariate poly of degree (md)
- By interpolation, given $md + 1$ distinct points $\alpha_1, \ldots, \alpha_{md+1}$, we can write the coeff of z^d in $F(z)$ as a linear combination of $F(\alpha_1), \ldots, F(\alpha_{md+1})$

$$\text{coeff of } z^d \text{ in } F(z) = \sum_{i=1}^{md+1} \delta_i F(\alpha_i)$$

$$\Rightarrow d! \cdot (\text{coeff of } z^d \text{ in } E_d(u_1 z) \cdots E_d(u_m z)) = \sum_{i=1}^{md+1} d! \delta_i \prod_{j=1}^{m} E_d(\alpha_i \cdot u_j)$$

□

Step 3: $\sum \bigwedge^{[a]} \sum \bigwedge^{[b]} \sum$ ckts $\longrightarrow \sum \prod \sum$ $\mathbb{C}$-ckts

Lemma IV.7. Let f be a polynomial computed by a $\sum \bigwedge^{[a]} \sum \bigwedge^{[d/a]} \sum$ circuit of size s_2 over $\mathbb{Q}$. Then, there is an equivalent $\sum \prod \sum$ circuit over $\mathbb{C}$ of size $s_3 = \mathcal{O}(s_2^3 a^2 bn)$ computing f. The circuit has formal degree at most $\mathcal{O}(s_2 ab)$

Step 3: $\sum \bigwedge^{[a]} \sum \bigwedge^{[d/a]} \sum$ ckts $\longrightarrow \sum \prod \sum$ $\mathbb{C}$-ckts

Proof:

- A $\sum \bigwedge^{[a]} \sum \bigwedge^{[b]} \sum$ circuit C computes a polynomial of the form $C = T_1 + \cdots + T_{s_2}$ where each $T_i = ({l_{i_1}}^b + \cdots + {l_{i_{s_2}}}^b)^a$ for some linear forms l_{i_j}'s
- Applying Saxena's trick to each $T = (l_1^b + \cdots + l_{s_2}^b)^a$ we get

$$T = \sum_{i=1}^{s_2 a+1} \beta_i \prod_{j=1}^{s_2} E_a(\alpha_i \cdot l_j^b)$$

$$= \sum_{i=1}^{s_2 a+1} \beta_i \prod_{j=1}^{s_2} f_i(l_j) \qquad \text{where } f_i(t) = E_a(\alpha_i \cdot t^b)$$

Step 3: $\sum \bigwedge^{[a]} \sum \bigwedge^{[d/a]} \sum$ ckts $\longrightarrow \sum \prod \sum$ $\mathbb{C}$-ckts

Proof:

- A $\sum \bigwedge^{[a]} \sum \bigwedge^{[b]} \sum$ circuit C computes a polynomial of the form $C = T_1 + \cdots + T_{s_2}$ where each $T_i = ({l_{i_1}}^b + \cdots + {l_{i_{s_2}}}^b)^a$ for some linear forms l_{i_j}'s
- Applying Saxena's trick to each $T = (l_1^b + \cdots + l_{s_2}^b)^a$ we get

$$\begin{aligned} T &= \sum_{i=1}^{s_2 a+1} \beta_i \prod_{j=1}^{s_2} E_a(\alpha_i \cdot l_j^b) \\ &= \sum_{i=1}^{s_2 a+1} \beta_i \prod_{j=1}^{s_2} f_i(l_j) \qquad \text{where } f_i(t) = E_a(\alpha_i \cdot t^b) \\ &= \sum_{i=1}^{s_2 a+1} \beta_i \prod_{j=1}^{s_2} \prod_{k=1}^{ab} (l_j - \gamma_{ik}) \end{aligned}$$

- f can be computed by a $\sum \prod \sum$ ckt having intermediate degree at most $s_2 ab$
- The final size of the ckt is $s_2 \cdot (s_2 a + 1) \cdot (s_2 ab) \cdot (n+1)$

$\square$

Overview

Progress so far

$$
\begin{aligned}
(\text{ABP}, \quad s) &\longrightarrow (\textstyle\sum\prod^{[a]}\sum\prod^{[d/a]}, & s_1 &= 2^{\mathcal{O}(\sqrt{d\log n\log s})}) \\
&\longrightarrow (\textstyle\sum\bigwedge^{[a]}\sum\bigwedge^{[d/a]}\sum, & s_2 &= 2^{\mathcal{O}(\sqrt{d\log n\log s})}) \\
&\longrightarrow (\textstyle\sum\prod\sum_{\mathbb{C}}, & s_3 &= 2^{\mathcal{O}(\sqrt{d\log n\log s})})
\end{aligned}
$$

Overview of steps

- **Step 0:** General ckts $\longrightarrow$ ABPs
- **Step 1:** ABPs $\longrightarrow \sum\prod^{[a]}\sum\prod^{[d/a]}$ ckts
- **Step 2:** $\sum\prod^{[a]}\sum\prod^{[d/a]}$ ckts $\longrightarrow \sum\bigwedge^{[a]}\sum\bigwedge^{[d/a]}\sum$ ckts
- **Step 3:** $\sum\bigwedge^{[a]}\sum\bigwedge^{[d/a]}\sum$ ckts $\longrightarrow \sum\prod\sum$ $\mathbb{C}$-ckts
- **Step 4:** $\sum\prod\sum$ $\mathbb{C}$-ckts $\longrightarrow \sum\prod\sum$ $\mathbb{Q}$-ckts

Step 4: $\sum\prod\sum$ $\mathbb{C}$-ckts $\longrightarrow$ $\sum\prod\sum$ $\mathbb{Q}$-ckts

Lemma IV.8 (An algebraic observation). Let $\gamma_1, \ldots, \gamma_a$ be roots of $E_a(t)$, and let ω be a primitive b-th root of unity. Then, the field $\mathbb{Q}(\gamma_1^{1/b}, \ldots, \gamma_a^{1/b}, \omega)$ contains the roots of $E_a(\alpha \cdot t^b)$ for every $\alpha \in \mathbb{Q}$ such that $\alpha^{1/b} \in \mathbb{Q}$

Proof:

- The roots of $E_a(\alpha t^b)$ are exactly $\left(\frac{\gamma_i}{\alpha}\right)^{\frac{1}{b}} \omega^j$ for $i \in [a]$ and $j \in [b]$
- As $\alpha^{1/b} \in \mathbb{Q}$, each root is in $\mathbb{Q}(\gamma_1^{1/b}, \ldots, \gamma_a^{1/b}, \omega)$

$\square$

Step 4: $\sum\prod\sum$ $\mathbb{C}$-ckts $\longrightarrow$ $\sum\prod\sum$ $\mathbb{Q}$-ckts

Lemma IV.8 (An algebraic observation). Let $\gamma_1, \ldots, \gamma_a$ be roots of $E_a(t)$, and let ω be a primitive b-th root of unity. Then, the field $\mathbb{Q}(\gamma_1^{1/b}, \ldots, \gamma_a^{1/b}, \omega)$ contains the roots of $E_a(\alpha \cdot t^b)$ for every $\alpha \in \mathbb{Q}$ such that $\alpha^{1/b} \in \mathbb{Q}$

Proof:

- The roots of $E_a(\alpha t^b)$ are exactly $\left(\frac{\gamma_i}{\alpha}\right)^{\frac{1}{b}} \omega^j$ for $i \in [a]$ and $j \in [b]$
- As $\alpha^{1/b} \in \mathbb{Q}$, each root is in $\mathbb{Q}(\gamma_1^{1/b}, \ldots, \gamma_a^{1/b}, \omega)$

Observe

- We want to apply the above to $E_a(\alpha_i \cdot t^b)$. As we can choose α_i to be any distinct rations we choose them s.t. $\alpha_i^{1/b} \in \mathbb{Q}$
- Thus the coefficients in step 3's $\sum\prod\sum$ $\mathbb{C}$-ckt come from $\mathbb{K} := \mathbb{Q}(\gamma_1^{1/b}, \ldots, \gamma_a^{1/b}, \omega)$
- Note $[\mathbb{Q}(\gamma^{1/b}) : \mathbb{Q}] \leq ab$
- $\Rightarrow [\mathbb{K} : \mathbb{Q}] \leq (ab)^a \cdot b$

Step 4: $\sum\prod\sum$ $\mathbb{C}$-ckts $\longrightarrow$ $\sum\prod\sum$ $\mathbb{Q}$-ckts

Lemma IV.9 ($\sum\prod\sum$ $\mathbb{K}$-ckt $\longrightarrow$ $\sum\prod\sum$ $\mathbb{Q}$-ckt).
Let $f(x) \in \mathbb{Q}[x]$ be computed by a $\sum\prod\sum$ circuit of formal degree D with coefficients coming from a finite extension field $\mathbb{K}/\mathbb{Q}$. Then, there is an equivalent $\sum\prod\sum$ circuit computing f of size $\mathrm{poly}(s_3, D, [\mathbb{K} : \mathbb{Q}])$ with coefficients coming from $\mathbb{Q}$

Rk: This along with Lemma IV.8 tells us that

$$(\mathrm{ABP}, s) \longrightarrow \left(\sum\prod\sum\nolimits_{\mathbb{Q}}, s_4 = 2^{\mathcal{O}(\sqrt{d\log n\log s})}\right)$$

Step 4: $\sum\prod\sum$ $\mathbb{C}$-ckts $\longrightarrow$ $\sum\prod\sum$ $\mathbb{Q}$-ckts

Lemma IV.9 ($\sum\prod\sum$ $\mathbb{K}$-ckt $\longrightarrow$ $\sum\prod\sum$ $\mathbb{Q}$-ckt).

Proof:

- Let $[\mathbb{K} : \mathbb{Q}] = m$
- There is $\theta \in \mathbb{K}$ s.t. $\mathbb{K} = \mathbb{Q}(\theta)$
- $\mathbb{K}$ is a vector space over $\mathbb{Q}$ with basis $\{\theta^0, \theta^1, \theta^2, \ldots, \theta^{m-1}\}$
- $\theta^j = \sum_{i=0}^{m-1} c_{ij}\theta^i$ for all $j \in \mathbb{N}$
- Thus any $g(x) \in \mathbb{K}[x]$ is uniquely written as $g^{[0]}\theta^0 + g^{[1]}\theta^1 + \cdots + g^{[m-1]}\theta^{m-1}$ where $g^{[r]} \in \mathbb{Q}[x]$

Step 4: $\sum\prod\sum$ $\mathbb{C}$-ckts $\longrightarrow$ $\sum\prod\sum$ $\mathbb{Q}$-ckts

Lemma IV.9 ($\sum\prod\sum$ $\mathbb{K}$-ckt $\longrightarrow$ $\sum\prod\sum$ $\mathbb{Q}$-ckt).

Proof:

- Let $[\mathbb{K} : \mathbb{Q}] = m$
- There is $\theta \in \mathbb{K}$ s.t. $\mathbb{K} = \mathbb{Q}(\theta)$
- $\mathbb{K}$ is a vector space over $\mathbb{Q}$ with basis $\{\theta^0, \theta^1, \theta^2, \ldots, \theta^{m-1}\}$
- $\theta^j = \sum_{i=0}^{m-1} c_{ij}\theta^i$ for all $j \in \mathbb{N}$
- Thus any $g(x) \in \mathbb{K}[x]$ is uniquely written as $g^{[0]}\theta^0 + g^{[1]}\theta^1 + \cdots + g^{[m-1]}\theta^{m-1}$ where $g^{[r]} \in \mathbb{Q}[x]$
- $f = T_1 + \cdots + T_{s_3}$ where each T_i is a product of linear polynomials over $\mathbb{K}$, then $f = T_1^{[0]} + \cdots + T_{s_3}^{[0]}$
- Hence it suffices to show that each $T_i^{[0]}$ can be expressed as a small depth-3 circuit over $\mathbb{Q}$

Step 4: $\sum\prod\sum$ $\mathbb{C}$-ckts $\longrightarrow$ $\sum\prod\sum$ $\mathbb{Q}$-ckts

Proof cont'd:

- Let $T = l_1 \cdots l_D \in \mathbb{K}[x]$

$$\begin{aligned} T &= \prod_{i\in[D]} (l_i^{[0]}\theta^0 + l_i^{[1]}\theta^1 + \cdots + l_i^{[m-1]}\theta^{m-1}) \\ &= T^{[0]}\theta^0 + T^{[1]}\theta^1 + \cdots + T^{[m-1]}\theta^{m-1} \end{aligned}$$

- Consider the polynomial obtained by replacing θ with a formal variable y

$$\begin{aligned} \tilde{T}(\underline{\boldsymbol{x}}, y) &= \prod_{i\in[D]} (l_i^{[0]}y^0 + l_i^{[1]}y^1 + \cdots + l_i^{[m-1]}y^{m-1}) \\ &= \tilde{T}_0 y^0 + \tilde{T}_1 y^1 + \cdots + \tilde{T}_{(m-1)D} y^{(m-1)D} \end{aligned}$$

- Using interpolation, $\tilde{T}_i$ can be written as a linear combination of $\left\{ \tilde{T}(x, \beta_j) : 1 \leq j \leq (m-1)D + 1 \right\}$
- Thus $\tilde{T}_i$ has a small depth-3 ckt

Step 4: $\sum\prod\sum$ $\mathbb{C}$-ckts $\longrightarrow$ $\sum\prod\sum$ $\mathbb{Q}$-ckts

Proof cont'd:

- To get $T^{[0]}$ from $T_1, T_2, \ldots, T_{(m-1)D}$, note $T^{[0]}\theta^0 + \cdots + T^{[m-1]}\theta^{m-1} = \tilde{T}_0\theta^0 + \cdots + \tilde{T}_{(m-1)D}\theta^{(m-1)D}$
- Use $\theta^j = \sum_{i=0}^{m-1} c_{ij}\theta^i$ to get
- $T^{[0]} = \sum_{j=0}^{(m-1)D} c_{0j}\tilde{T}_j$
- Thus $T^{[0]}$ is computable by a small depth-3 ckt and hence f is computable by a small depth-3 ckt
  $\blacksquare$

References

1. A. Gupta, P. Kamath, N. Kayal, R. Saptharishi. Arithmetic circuits: A chasm at depth three.