An exposition of the complexity of partial derivatives for constant
deepth arithmetic circuits and its relation to Raz’s lower bound.

Suryajith Chillara∗
May 22, 2018

The aim of this note is to point out the subtlety in the proof of Corollary 4.6 of [Raz10] and its relation
to the main result1.

In this context, size refers to the number of edges in the underlying graph of the circuit. Baur and
Strassen [BS83] (cf. [Mor85]) proved that if a polynomial \( f \) is computed by an arithmetic circuit of size \( s \) and depth \( d \), then there is another circuit of size at most \( 5s \) and depth at most \( 3d \) that simultaneously
computes the polynomial \( f \) and all of its first order derivatives. This proof of Baur and Strassen
[BS83], and Morgenstern [Mor85] starts with an assumption that the fan-in of the product gates is bounded by
2 or instead refers to the number of nodes as the size. In the context of Raz’s result [Raz10], we need a
bound on the number of edges. In case of constant depth arithmetic circuits that assumption on fan-in is
not feasible as converting a given constant depth arithmetic circuit with unbounded fan-in, converting that
into a circuit of bounded fan-in product gates may blow up the depth2. Reworking their proof for constant
depth multilinear circuits tells us that we can not bound the number of edges going in to a product gate
by \( O(s) \) if we were to keep the depth as \( O(d) \), using this technique.

However, we observe that it is sufficient to prove a bound on the total number of edges going into the
sum gates of the new circuit that computes the partial derivatives. We believe that this was implicit in the
proof of Raz [Raz10].

1 Complexity of Partial Derivatives

**Theorem 1.1 ([BS83, Mor85]).** Let \( C \) be an arithmetic circuit, with unbounded fan-in, that computes \( f \in \mathbb{F}[x_1, \ldots, x_n] \), with \( s \) many edges and depth \( d = O(1) \). Then there is another arithmetic circuit \( C' \) of depth
\( O(d) \) that computes \( \{ \frac{\partial f}{\partial x_i} | i \in [n] \} \cup \{ f \} \) simultaneously. Further, the total number of edges going into the
sum gates in \( C' \) is at most \( O(s) \).

**Proof.** This proof follows the arguments presented in the proof of Theorem 9.10 in [CKW11]. Let
\( g_1, \ldots, g_s \) be the variables corresponding to the polynomials computed at the non-input nodes3. Consider
a sequence of polynomials \( f_s, f_{s-1}, \ldots, f_0 \) over the variables \( \{ x_1, \ldots, x_n, g_1, \ldots, g_s \} \) such that \( f_s = g_s \) corresponding to the output node and \( f_0 = f \). Further, for \( k \in [0, \ldots, s-1] \), \( f_k \) is obtained from \( f_{k+1} \) by substituting for \( g_{k+1} \) in it. It is now clear that \( f_k \in \mathbb{F}[x_1, \ldots, x_n, g_1, \ldots, g_k] \).

∗Department of CSE, IIT Bombay, India. suryajith@cse.iitb.ac.in
†The same could also be found towards the end of Section 3.5 in [SY10].
‡However, we currently do not know if this blow up of depth is inevitable for general constant depth circuits.
§The number of nodes is at most the number of edges.
We shall now prove the theorem by induction. We shall inductively compute $\frac{\partial f_k}{\partial g_k}$ using $\frac{\partial f_{\ell}}{\partial g_{\ell}}$ for all $\ell > k$.

**Base case:** $\frac{\partial f_k}{\partial g_k} = 1$.

**Induction step:** Given $\frac{\partial f_{\ell}}{\partial g_{\ell}}$ for all $\ell > k$, we need to compute $\frac{\partial f_k}{\partial g_k}$. Using a telescopic sum, we get

$$\frac{\partial f_k}{\partial g_k} = \sum_{\ell=k}^{s-1} \left( \frac{\partial f_{\ell}}{\partial g_{\ell}} - \frac{\partial f_{\ell+1}}{\partial g_{\ell+1}} \right) + \frac{\partial f_s}{\partial g_k}. \quad (1.2)$$

Now consider a summand of the form $\left( \frac{\partial f_{\ell}}{\partial g_{\ell}} - \frac{\partial f_{\ell+1}}{\partial g_{\ell+1}} \right)$. Let us recall that $f_{\ell}$ is obtained from $f_{\ell+1}$ by substituting for $g_{\ell+1}$ in it. Thus, by using chain rule, we get

$$\frac{\partial f_{\ell}}{\partial g_k} = \frac{\partial f_{\ell+1}}{\partial g_k} + \frac{\partial g_{\ell+1}}{\partial g_k} \cdot \frac{\partial f_{\ell+1}}{\partial g_{\ell+1}}. \quad (1.3)$$

Substituting Equation 1.3 into Equation 1.2, we get

$$\frac{\partial f_k}{\partial g_k} = \sum_{\ell=k}^{s-1} \frac{\partial g_{\ell+1}}{\partial g_k} \cdot \frac{\partial f_{\ell+1}}{\partial g_{\ell+1}} = \sum_{g_{\ell}\in\text{parent}(g_k)} \frac{\partial g_{\ell}}{\partial g_k} \cdot \frac{\partial f_{\ell}}{\partial g_{\ell}}. \quad (1.4)$$

This shows that the number of edges in to the sum gates in the new circuit are at most the number of parents for a node (or the outgoing edges). Summing it over all the nodes, we get that there are at most $s$ many new edges to the new sum gates are added. Here, we make no such claims about the edges going to the product gates. We can not bound that by $O(s)$ as $\frac{\partial g_{\ell}}{\partial g_k}$ could add as many edges as the number of siblings of $g_k$.

A careful observation would reveal that the depth of the circuit constructed through the above process is at most $3d$ and the number of nodes is at most $O(s)$.

2 Application to Raz’s result

The following observation can be found in section 3.2 of [Raz10].

**Lemma 2.1.** Let $C'$ be an arithmetic circuit of size $s'$ and depth $d'$ with edge weights from some field $\mathbb{G}$. Then there is a circuit $C''$ that computes the same polynomial(s) as $C$ and is such that the weight of all edges that feed into any product gate is 1, size is $s'$ and depth is $d'$.

This transformation can be achieved by multiplying the edge weights of each of the outgoing edges from a product gate with the edge weights of each of the incoming edges into a product gate. Now, the edge weights of each of those incoming edges can be set to 1.

Raz then considers the constant depth circuits with edge weights just on the edges in to the sum gates in new circuit $C''$, say $\{y_1, \cdots, y_s\}$ and obtains a lower bound of $\Omega(n^{1+1/2d})$ on $s'$ in Corollary 4.5, [Raz10]. Combining this with Theorem 1.1, we get the main result.

2
Acknowledgement

We thank Sébastien Tavenas for asking the question at WACT 2018, that resulted in this note.

References


