A nonlinear bound on the number of wires in bounded depth circuits

(preliminary version)

Pavel Pudlák *

October 26, 2005

We shall prove a lower bound on the number of edges in some bounded depth graphs. This theorem is stronger than lower bounds proved on bounded depth superconcentrators [1] and enables us to prove a lower bound on certain bounded depth circuits for which we cannot use superconcentrators: we prove that conjunction cannot be computed by bounded depth circuits with modular gates that have linear number of wires. The proof of the theorem is based on the same ideas as the proof for bounded depth superconcentrators in [2]. It should be noted that another combinatorial lemma related to superconcentrators was proved by Raz and Shpilka [3] (also using the approach of [2]) in order to show nonlinear lower bounds on the number of edges in bounded depth arithmetic circuits computing matrix multiplication and some other functions. Some related results appeared recently also in [1].

Let G be a finite directed acyclic graph with a distinguished set of indegree zero vertices V_0, which will be called *input vertices*. Let X be a subset of input vertices. We shall say that a subset of vertices S *separates* X, if for every two different input vertices $x, y \in X$, every vertex v and every pair of directed paths p, q starting in x and y respectively and ending in v, at least one of the paths must contain a vertex from S. S may contain input vertices.

We shall say that X is *ε-separable,* if there exists an S such that S separates X and $|S| \leq \varepsilon|X|$.

We shall say that G is *ε-inseparable*, if for every subset of input vertices X, if $|X| \geq 2$, then X is not ε-separable. ($\varepsilon < 1$, as X separates itself.)

Define, for $d = 1, 2, \ldots,$

$$\lambda_1(n) = \lceil \log_2 n \rceil,$$

$$\lambda_{d+1}(n) = \min\{i \in \mathbb{N};\ \lambda_d^{(i)}(n) \leq 1\},$$

where the superscript i denotes the i-times iterated function.[2]

Theorem 1 *For every $\varepsilon > 0$ and every integer $d \geq 1$, there exists $\delta > 0$ such that for all n, if G has depth d, n inputs and it is ε-inseparable, then it has at least $\delta n \lambda_d(n)$ edges.*

We shall prove a stronger version of this theorem. For a set of inputs X of G, define

$$s(X) = \min\{|S|;\ S \text{ separates } X\}.$$

*Supported by grant A1019401 of the Academy of Sciences, Czech Republic

[1] when the depth of superconcentrators is even

[2] Note that the functions λ_i defined in [3] are different.

Let n be the number of input vertices, let $2 \le t \le n$, and $\varepsilon > 0$. We shall say that G is *weakly t, ε-inseparable*, if for all k, $t \le k \le n$,

$$\mathop{\mathbf{E}}_{|X|=k}(s(X)) > \varepsilon k.$$

The greater generality (in particular, the bound on the expectation, instead of an absolute bound) is needed for the proof.

Theorem 2 *For every $\varepsilon > 0$ and every integer $d \ge 1$, there exists $\delta > 0$ such that for every $2 \le t \le n$, every weakly t, ε-inseparable G of depth d with n input vertices has at least $\delta n \lambda_d(\frac{n}{t})$ edges.*

This theorem is proved by induction on the depth d. We shall assume w.l.o.g. that G is stratified into levels $V_0, V_1, \cdots, V_d$ and edges are only between consecutive levels. The following two lemmas formalize the induction base and the induction step.

Lemma 3 *For every $\varepsilon > 0$, there exists $\delta > 0$ such that if G has depth 1, has n input vertices and it is weakly t, ε-inseparable, where $2 \le t \le n$, then it has more than $\delta n \log \frac{n}{t}$ edges.*

Proof. Suppose G is weakly t, ε-inseparable. Let $v_1, v_2, \ldots$ be all vertices on the level 1 (the level 0 being the input vertices) ordered by the decreasing indegrees $d_1 \ge d_2 \ge \ldots$. For $t \le q \le \frac{\varepsilon n}{2}$ consider the undirected graph H_q with the set of vertices being the input vertices of G and edges (x, y) such that $x \to v_i, y \to v_i$ in G for some $i > q$. Thus H_q has $m \le \sum_{i>q} \binom{d_i}{2}$ edges. Let X be a random subset of inputs of cardinality $k = \lceil \frac{2q}{\varepsilon} \rceil$ (thus $t \le k \le n$). The expected number of edges on X is $\frac{m}{\binom{n}{2}} \binom{k}{2}$.

Observe that if there are ℓ edges of H_q on X, then $s(X) \le \ell + q$ (take the vertices $v_1, \cdots, v_q$ and one vertex from each edge). Thus we have

$$\frac{m}{\binom{n}{2}} \binom{k}{2} + q \ge \mathbf{E}(s(x)) > \varepsilon k.$$

Since $q \le \varepsilon k/2$, we have

$$\frac{m}{\binom{n}{2}} \binom{k}{2} > \frac{\varepsilon k}{2}.$$

Substituting for m and simplifying we get

$$\sum_{i>q} \frac{\binom{d_i}{2}}{\binom{n}{2}} > \frac{\varepsilon}{k-1}.$$

Since $d_i \le n$, we can estimate $\frac{\binom{d_i}{2}}{\binom{n}{2}} \le \frac{d_i^2}{n^2}$. Thus we get

$$\sum_{i>q} \frac{d_i^2}{n^2} > \frac{\varepsilon}{k-1} = \frac{\varepsilon}{\lceil \frac{2q}{\varepsilon} \rceil - 1} \ge \frac{\varepsilon^2}{2q}.$$

By Lemma 4 of [2], this implies

$$\sum_i \frac{d_i}{n} \ge \delta_1 \log \frac{\lfloor \frac{\varepsilon n}{2} \rfloor}{t},$$

for some $\delta_1 > 0$ depending only on ε. Hence if $t = o(n)$, we get

$$\sum_i d_i \geq \delta n \log \frac{n}{t}.$$

Otherwise use the trivial lower bound εt on the number of edges. ∎

Lemma 4 *For every integer $d \geq 1$, reals $\varepsilon > 0$, and $\gamma > 0$, there exists $\delta > 0$ such that for every n, if*

> *(i) for every $2 \leq t \leq n$, every weakly $t, \frac{\varepsilon}{2}$-inseparable G of depth d with n input vertices has at least $\gamma n \lambda_d(\frac{n}{t})$ edges,*

then

> *(ii) for every $2 \leq t \leq n$, every weakly t, ε-inseparable G of depth $d+1$ with n input vertices has at least $\delta n \lambda_{d+1}(\frac{n}{t})$ edges.*

Proof. Suppose (i) holds true. Let G be weakly t, ε-inseparable directed graph with depth $d+1$ and n input vertices.

Let me briefly sketch the idea of the proof before doing detailed computations. We would like to distinguish two cases: either there are a lot of vertices of high degree on the first level, or not. In the first case there are, clearly, many edges. In the second case we can delete the vertices on the first level that have large degrees, connect inputs directly to the second level and then we can apply (i) to the resulting depth d graph. However, this does not quite work, as after deleting the vertices with high degree, the degrees of the remaining vertices on level 1 are still too large. Therefore we have to consider also vertices with intermediate degrees. If the number of those vertices would be small, then a random set of inputs would meet only a few edges connected to them.

Let $deg(v)$ denote the indegree of a vertex v. Let t be given, $2 \leq t \leq n$. Put $r = \frac{n}{t}$,

$$A_0 = \{v \in V_1;\ deg(v) > \lambda_d(r)\},$$

$$A_i = \{v \in V_1;\ \lambda_d^{(i+1)}(r) < deg(v) \leq \lambda_d^{(i)}(r)\},\ \text{for } i \geq 1.$$

Let E denote the set of edges of G.

Claim. For every i, $1 \leq i \leq \lambda_{d+1}(r)/2 - 3$, at least one of the following three inequalities is satisfied:

1. $|A_0 \cup \cdots \cup A_{i-1}| \geq \frac{\varepsilon}{4} \frac{n}{\lambda_d^{(i+1)}(r)}$;
2. $|\{(u,v) \in E;\ u \in V_0, v \in A_i \cup A_{i+1} \cup A_{i+2}\}| \geq \frac{\varepsilon}{4} n$;
3. $|\{(u,v) \in E;\ u, v \notin A_0 \cup \cdots \cup A_{i+2}\}| \geq \gamma n \frac{\lambda_d^{(i+2)}(r)}{\lambda_d^{(i+3)}(r)}$.

Proof of Claim. Let i be given and suppose that conditions (1) and (2) are false. Let $n/\lambda_d^{(i+1)}(r) \leq k \leq n$. Observe that $n/\lambda_d^{(i+1)}(r) = n/\lambda_d^{(i+1)}(n/t) \geq t$, since $\lambda_d(x) \leq x$ for all x. Let $X \subseteq V_1$ be a random subset of size k. We shall show that if we remove from G all edges incident with $A_0 \cup \cdots \cup A_{i+2}$, then

$$\mathbf{E}(s'(X)) > \frac{\varepsilon}{2} k,$$

where $s'(X)$ denotes $s(X)$ in the modified graph, which we shall denote by G'.

Indeed, let $a = |A_0 \cup \cdots \cup A_{i-1}|$, $b(X) = |\{(u,v) \in E;\ u \in X, v \in A_i \cup A_{i+1} \cup A_{i+2}\}|$. Then

$$s(X) \leq a + b(X) + s'(X).$$

Hence

$$\mathbf{E}(s'(X)) \geq \mathbf{E}(s(X) - b(X) - a) = \mathbf{E}(s(X)) - \mathbf{E}(b(X)) - a.$$

By non-1, $a < \frac{\varepsilon}{4}\frac{n}{\lambda_d^{(i+1)}(r)} \leq \frac{\varepsilon}{4}k$. By non-2, we have $\mathbf{E}(b(X)) < \frac{\varepsilon}{4}k$, (each edge from $\{(u,v) \in E;\ u \in V_0, v \in A_i \cup A_{i+1} \cup A_{i+2}\}$ is chosen with probability k/n; use the linearity of expectation).

Thus G' is weakly $n/\lambda_d^{(i+1)}(r), \frac{\varepsilon}{2}$-inseparable.

We shall further modify G' by removing all edges between V_1 and V_2 and adding, for every path (u,v,w) in G' with $u \in V_0, v \in V_1, w \in V_2$, the edge (u,w). The resulting graph will be denoted by G''. It has depth d (the first level being $V_1 \cup V_2$, the second level being V_3 etc.) and at most $\lambda_d^{(i+3)}(r)$-times more edges.

Furthermore, G'' is also weakly $n/\lambda_d^{(i+1)}(r), \frac{\varepsilon}{2}$-inseparable. To see that, observe that if X is a set of inputs (in G' and G'') and S is a separating set for X in G'', then S is a separating set for X also in G'. Indeed, let S be a separating set for X in G'' and let $(v_0, \cdots, v_j)$ and $(u_0, \cdots, u_j)$ be two paths in G', $v_0, u_0 \in X$, $v_0 \neq u_0$ and $v_j = u_j$. Then if $j = 1$, these paths are also paths in G'', and if $j > 1$, $(v_0, v_2, \cdots, v_j)$ and $(u_0, u_2, \cdots, u_j)$ are paths in G''. In both cases they contain an element from S, whence the original pair of paths also contains an element from S. Thus separating sets are at least as large in G'' as in G'.

By the assumption (i), G'' must have at least $\gamma n \lambda_d(\lambda_d^{(i+1)}(r)) = \gamma n \lambda_d^{(i+2)}(r)$ edges. Hence G' has at least $\gamma n \lambda_d^{(i+2)}(r)/\lambda_d^{(i+3)}(r)$ edges, which proves 3. This finishes the proof of the Claim.

To finish the proof of Lemma 4, we shall use the inequality

$$\frac{\lambda_d^{(i)}(r)}{\lambda_d^{(i+1)}(r)} \geq \tfrac{1}{2}\lambda_{d+1}(r),$$

for every $i \leq \lambda_{d+1}(r)/2 - 1$, which was proved in [2] as Lemma 5. By the Claim it suffices to consider the following three cases.

1. Suppose for some $i \leq \lambda_{d+1}(r)/2 - 3$ the condition (i) of Claim is satisfied. Then, since every $v \in A_0 \cup \cdots \cup A_{i-1}$ has degree $> \lambda_d^{(i)}(r)$, the number of edges in G is at least

$$\frac{\varepsilon}{4}\frac{n}{\lambda_d^{(i+1)}(r)}\lambda_d^{(i)}(r) \geq \frac{\varepsilon}{8} n \lambda_{d+1}(r).$$

2. Suppose for *all* $i \leq \lambda_{d+1}(r)/2 - 3$ the condition (ii) of Claim is satisfied. Then the number of edges of G is at least

$$\tfrac{1}{3}(\lambda_{d+1}(r)/2 - 3)\frac{\varepsilon}{4}n = \Omega(n\lambda_{d+1}(r)).$$

3. Suppose for some $i \leq \lambda_{d+1}(r)/2 - 3$ the condition (iii) of Claim is satisfied. Then the number of edges of G is at least

$$\gamma n \frac{\lambda_d^{(i+2)}(r)}{\lambda_d^{(i+3)}(r)} \geq \tfrac{1}{2}\gamma n \lambda_{d+1}(r).$$

■

Corollary 5 *For every q and d, there exists $\delta > 0$ such that every circuit computing $x_1 \wedge \cdots \wedge x_n$ that has depth $d+1$ and uses only MOD_q (with arbitrary coefficients), has at least $\delta n \lambda_d(n)$ edges.*

Proof. By Theorem 2.1 of Thérien [4] and a theorem of Euler about primes (stated therein), for a given q there exists a $\gamma > 0$ such that for every m and s and every linear mapping $\Theta : \mathbb{Z}_q^m \to \mathbb{Z}_q^s$, if $s \leq \gamma m$, then the kernel of Θ contains a nonzero element of $\{0,1\}^m$.

W.l.o.g. suppose that a circuit C with MOD_q gates computes $\neg x_1 \wedge \cdots \wedge \neg x_n$. Let $0 < \varepsilon < \gamma$, let $\delta > 0$ be given by Theorem 1 for these ε and d. Suppose that the circuit has $< \delta n \lambda_d(n)$ edges. Then, by Theorem 1, there exists a set of inputs X which is ε-separated in the depth d graph obtained by removing the output gate from the circuit. Let S be the separating set augmented with the output gate. Then S is a separating set in the whole circuit and $|S| \leq \varepsilon|X| + 1$. We may moreover require that $|X| \geq \log n$, thus if n is sufficiently large, $|S| \leq \gamma|X|$.

Fix all inputs that are not in X to 0. Furthermore, for every $v \in S$, disconnect v from its inputs and set it to be the constant equal to the boolean value computed at v when all inputs are 0. Let C' be the resulting circuit. Let $v \in S$ and let w be an input gate of v in C. Then in C', the gate w only depends on at most one input from X, because S is a separating set. Thus if we put back the original MOD_q gate on v, the function computed at v will be some MOD_q function f_v.

Thus in order to get a contradiction with the assumption that C computes $\neg x_1 \wedge \cdots \wedge \neg x_n$, we need only to find an assignment $\bar{a}$, $\bar{a} \neq \bar{0}$, such that for every $v \in S$, $f_v(\bar{a}) = f_v(\bar{0})$, because then C will compute the same output on the all-zero input as on the input with zeros on variables outside of X and $\bar{a}$ on X. Since every f_v is a MOD_q function, the set of these equations is equivalent to a set of linear equations over $\mathbb{Z}_q$ plus the condition that $\bar{a}$ is a 0-1 vector. Hence the existence of such a vector follows from Thérien's theorem. ∎

References

[1] M Koucký, P. Pudlák and D. Thérien, Bounded depth circuits: Separating wires from gates. *37th ACM STOC,* 2005, pp.257-265.

[2] P. Pudlák, Communication in bounded depth circuits. *Combinatorica* 14(2), (1994), pp.203-216.

[3] R. Raz and A. Shpilka, Lower bounds for matrix product in bounded depth circuits with arbitrary gates. *SIAM J. on Computing* 32(2), 2003, pp.488-513.

[4] D. Thérien, Circuits constructed with MOD_q gates cannot compute "AND" in sublinear size. *Computational Complexity* 4 (1994), pp.383-388.