

Complexity of Random Operators in Circuits With Arbitrary Gates

Stasys Jukna^{a,b,1}, Georg Schnitger^{b,1}

^a*Institute of Mathematics and Computer Science, Akademijos 4, Vilnius, Lithuania*
^b*University of Frankfurt, Institute of Computer Science, Frankfurt am Main, Germany*

Abstract

We consider boolean circuits computing n -operators $f : \{0, 1\}^n \rightarrow \{0, 1\}^n$. As gates we allow *arbitrary* boolean functions; neither fanin nor fanout of gates is restricted. An operator is linear if it computes n linear forms, that is, computes a matrix-vector product Ax over $GF(2)$. We prove the existence of n -operators requiring about n^2 wires in any circuit, and *linear* n -operators requiring about $n^2 / \log n$ wires in depth-2 circuits, if either all output gates or all gates on the middle layer are linear.

Key words: Computational complexity; Boolean circuits; Matrix rigidity; Random operators; Kolmogorov complexity; Lower bounds

1. Introduction and results

We consider general circuits computing boolean operators n -operators $f : \{0, 1\}^n \rightarrow \{0, 1\}^n$. As gates we allow *arbitrary* boolean functions of their inputs; there is no restriction on their fanin or fanout. Thus, the phenomenon which causes complexity of such circuits is *information transfer* rather than *information processing* as in the case of single functions. Such a circuit is a directed acyclic graph with n input nodes x_1, \dots, x_n and n output nodes y_1, \dots, y_n . Each non-input node computes some boolean function of its predecessors. A circuit computes f if, for all $i = 1, \dots, n$ the boolean function computed at the i th output node y_i is the i th component f_i of the operator $f = (f_1, \dots, f_n)$. The *depth* of a circuit is the largest number of wires in a path from an input to an output node.

The *size* of a circuit is the total number of wires in it. We will denote by $s_d(f)$ the smallest number of wires in a general circuit of depth at most d computing f . If there are no restrictions on the depth, the corresponding measure is denoted by $s(f)$. Note that $s(f) \leq s_1(f) \leq n^2$ holds for any n -operator, so quadratic lower bounds are the highest ones.

Circuits of depth 2 constitute the first non-trivial model. Interest in depth-2 circuits comes from the following important result of Valiant [16]: if in every depth-2 circuit, computing f with $r = O(n / \ln \ln n)$ gates on the middle layer, at least $n^{1+\Omega(1)}$ wires must enter output gates, then f cannot be computed by log-depth circuit over $\{\&, \vee, \neg\}$ of linear size. To prove a non-linear lower bound for log-depth circuits is an old and well known problem in circuit complexity.

Super-linear lower bounds $s_2(f) = \Omega(n \log^2 n)$ were proved using graph-theoretic arguments by analyzing some super-concentration properties of the circuit as a graph [5, 8, 9, 11, 10, 1, 12, 13, 14]. Higher lower bounds of the form $s_2(f) = \Omega(n^{3/2})$ were proved using relatively simple information theoretical arguments [4, 6]. For larger depth d known lower bounds are only slightly non-linear. All these bounds, however, are on the *total* number of wires, so they still have no consequences for log-depth circuits.

In fact, in the class of general circuits, even the question about the complexity of a *random* operator remained unclear. In particular, it was unclear whether operators requiring a quadratic number of wires (even in depth 2) exist at all?

Email addresses: jukna@thi.informatik.uni-frankfurt.de (Stasys Jukna), georg@thi.informatik.uni-frankfurt.de (Georg Schnitger)

¹Research supported by a DFG grant SCHN 503/4-1.

Note that a direct counting argument, as in the case of constant fanin circuits, does not work for general circuits: already for $d > n + \log n$, the number 2^{2^d} of possible boolean functions that may be assigned to a node of fanin d may be larger than the total number 2^{n2^n} of n -operators.

Our first result is an observation that this bad situation can be excluded by just turning the power of circuits against themselves to ensure that, in an optimal circuit, no gate can have fanin larger than n . This leads us to

Theorem 1. *For almost all n -operators f , $s(f) = \Omega(n^2)$.*

An important class of operators are *linear* ones. Each such operator computes n linear forms, that is, a matrix-vector product $f_A(\mathbf{x}) = A\mathbf{x}$ over $GF(2)$. In the class of *linear* circuits—where we only allow linear boolean functions (parities and their negations) as gates—easy counting shows that most such operators require $\Omega(n^2/\log n)$ wires. It is also known that $O(n^2/\log n)$ of wires is also sufficient to compute any linear operator, even with linear depth-2 circuits [15, 3, 2].

But what if we allow arbitrary (non-linear) boolean functions as gates—can we then compute linear operators f_A more efficiently? The largest known lower bound for an *explicit* linear operator f_A has the form $s_2(f_A) = \Omega(n \log n)$ [10]. This raises the following question: Do *linear* n -operators with $s(f_A) = \Omega(n^2/\log n)$ exist at all? This question remains open even for depth-2 circuits.

The next two theorems partially answer this question. By a *middle-linear* (resp., *output-linear*) circuit we will mean a depth-2 circuit whose all gates on the middle (resp., output) layer are linear boolean functions.

It turns out that the non-linearity of *middle* gates cannot help to compute linear operators by depth-2 circuits more efficiently, and hence, linear n -operators requiring about $n^2/\log n$ in such circuits exist.

Theorem 2. *Every output-linear circuit of size L computing a linear n -operator can be transformed to an equivalent linear depth-2 circuit of size $L + n$.*

The second case—when only gates on the middle layer are required to be linear—is more delicate. That such circuits *can* be more powerful than linear ones, was shown in [7]. Given a boolean $n \times n$ matrix A , say that a circuit *weakly computes* the operator $f_A(\mathbf{x}) = A\mathbf{x}$ if it correctly computes it on all n unit vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$. Note that, for *linear* circuits, this is no relaxation: such a circuit weakly computes f_A iff it correctly computes the f_A on all inputs. Hence, some linear operators cannot be weakly computed by *linear* depth-2 circuits using fewer than $\Omega(n^2/\log n)$ wires. It is however shown in [7] that the situation changes drastically if we allow *non-linear* gates: then *any* linear n -operator can be weakly computed using only $O(n \log n)$ wires.

Still, using Kolmogorov complexity arguments, we can prove

Theorem 3. *If middle gates are required to be linear, then linear n -operators f_A with $s_2(f_A) = \Omega(n^2/\log n)$ exist.*

2. Proof of Theorem 1

Let $\mu(L)$ be the number of different n -operators $f : \{0, 1\}^n \rightarrow \{0, 1\}^n$ computable by boolean circuits with at most L wires. Our goal is to upper bound this number in terms of n and L , and compare this bound with the total number 2^{n2^n} of n -operators.

Take an optimal circuit with $\ell \leq L$ wires computing some n -operator; hence, $\ell \leq n^2$. Then $\ell = \sum_{i=1}^m d_i$, where d_1, \dots, d_m are the fanins of its gates. It is clear that we need $m \geq n$ gates, since we must have n input gates. On the other hand, $m \leq \ell + n + 2 \leq 2n^2$ gates are always enough since every non-input gate, besides two possible constant gates, must have nonzero fanin.

We now make use of the fact that the gates in our circuits may be *arbitrary* boolean functions: This allows us to assume that $d_i \leq n$ for all i . Indeed, if $d_i > n$, then we can replace the i th gate by the boolean function computed at this gate and join it to all n input variables; when doing this, the total number of wires in the circuit can only decrease.

The number of sequences d_1, \dots, d_m of fanins with $0 \leq d_i \leq n$ does not exceed $(n+1)^m$. For each such sequence and for each $i = 1, \dots, m$, there are at most $\binom{m}{d_i} \leq m^{d_i}$ possibilities to choose the set of inputs for the i th node and at most $2^{2^{d_i}}$ possibilities to assign a boolean function to this node. Hence,

$$\mu(L) \leq (n+1)^m \prod_{i=1}^m m^{d_i} \prod_{i=1}^m 2^{2^{d_i}} = (n+1)^m m^{\sum_{i=1}^m d_i} 2^{\sum_{i=1}^m 2^{d_i}}.$$

Since $\sum_{i=1}^m d_i \leq L \leq n^2$ and $m \leq 2n^2$, this yields

$$\log_2 \mu(L) \leq \sum_{i=1}^m 2^{d_i} + O(n^2 \log_2 n).$$

We now observe that at most $n/2$ nodes can have fanin larger than $2L/n$, for otherwise we would have more than $(2L/n) \cdot (n/2) = L$ wires in total. Since $m \leq 2n^2$ and since the fanin of each gate does not exceed n , we obtain that

$$\sum_{i=1}^m 2^{d_i} \leq (m - n/2)2^{2L/n} + (n/2)2^n \leq 2n^2 4^{L/n} + 2^{n-1}.$$

Hence,

$$\log_2 \mu(L) \leq 2n^2 4^{L/n} + n2^{n-1} + O(n^2 \log_2 n). \quad (1)$$

Since the total number of operators $f : \{0, 1\}^n \rightarrow \{0, 1\}^n$ is 2^{n^2} , the smallest number L of wires sufficient to compute all of them must satisfy $\log_2 \mu(L) \geq n2^n$. By (1), this implies

$$2n^2 4^{L/n} \geq n2^{n-1} - O(n^2 \log_2 n).$$

Dividing both sides by $2n^2$, we obtain that $4^{L/n} = \Omega(2^n/n)$, and hence, $L = \Omega(n^2)$. \square

3. Proof of Theorem 2

Let A be an m -by- n $(0, 1)$ -matrix, and let F be a depth-2 circuit computing Ax . We may assume, for simplicity, that there are no direct wires from inputs to outputs: this can be easily achieved by adding n new wires.

Let $h = (h_1, \dots, h_r)$ be the operator $\{0, 1\}^n \rightarrow \{0, 1\}^r$ computed by the gates on the middle layer. Let also B be the m -by- r adjacency $(0, 1)$ -matrix of the bipartite graph formed by the wires joining the gates on the middle layer with those on the output layer.

Assume that all output gates of F are linear boolean functions. Then $Ax = B \cdot h(x)$ for all $x \in \{0, 1\}^n$. Write each vector $x = (x_1, \dots, x_n)$ as the linear combination $x = \sum_{i=1}^n x_i e_i$ of unit vectors $e_1, \dots, e_n \in \{0, 1\}^n$, and replace the operator H computed on the middle layer by a linear operator $h'(x) := \sum_{i=1}^n x_i h(e_i) \pmod{2}$. Hence, $h'(x) = x^\top M$, where M is an $n \times r$ matrix with rows $h(e_1), \dots, h(e_n)$. Using the linearity of the matrix-vector product, we obtain that (with all sums mod 2):

$$B \cdot h(x) = A \cdot \left(\sum x_i e_i \right) = \sum x_i A e_i = \sum x_i B \cdot h(e_i) = B \cdot h'(x).$$

Hence, the new (linear) circuit F' computes Ax as well. It remains to show that the number of wires in F' does not exceed the number of wires in F .

The wires on the second level haven't changed at all. To show that the number of wires on the first level has not increased as well, let $\text{fanout}(x_i)$ be the fanout of the i th input node x_i , and $\text{fanin}(h_j)$ the fanin of the j th gate h_j on the middle layer. Then $\sum_{i=1}^n \text{fanout}(x_i) = \sum_{j=1}^r \text{fanin}(h_j)$ is the total number L of wires on the first level. Since $A\mathbf{0} = \mathbf{0}$, we can assume that $h(\mathbf{0}) = \mathbf{0}$, that is, $h_j(\mathbf{0}) = 0$ for all $j = 1, \dots, r$. Now we make a simple (but crucial) observation: if there is no wire from x_i to h_j , then $h_j(e_i) = h_j(\mathbf{0}) = 0$. This implies that the j th column of M can have at most $\text{fanin}(h_j)$ ones. Since the number of wires on the first level of F' is just the total number of 1's in M , we are done. \square

4. Proof of Theorem 3

We use Kolmogorov complexity argument. Let A be a boolean $n \times n$ matrix of Kolmogorov complexity $\Omega(n^2)$. Hence, the linear operator $f_A(x) = Ax$ cannot be described using fewer than $\Omega(n^2)$ bits.

Fix an arbitrary depth-2 circuit F computing f_A , and assume that all its gates on the middle layer are linear. Let L be the number of wires in F . As in the previous section, we may assume that there are no direct wires from

inputs to outputs. Our goal is to show that, using the circuit F , the operator f_A can be described using $O(L \log n)$ bits. This will imply the desired lower bound $L = \Omega(n^2 / \log n)$ on the number of wires.

Let v_1, \dots, v_r nodes on the middle layer. Since at these nodes only linear functions are computed, the first level (between inputs and middle layer) computes some linear operator $y = Bx$, where B is a boolean $r \times n$ matrix such that $B[i, j] = 0$ if there is no wire from input variable x_i to the node v_j .

Let C be an incidence $n \times r$ matrix of the second level of the circuit. Using these two matrices B and C as well as the fact that the operator computed by the circuit F is linear, we can encode the entire work of the circuit using $O(L \log n)$ bits as follows. Let L_1 be the number of wires on the first level (between input and middle layer), and L_2 be the number of wires on the second level (between the middle and output layers).

1. The matrix B can be described using $L_1 \lceil \log_2 n \rceil$ bits. Indeed, if b_j denotes the number of 1's in the j th row of B , then $\sum_{j=1}^r b_j$ does not exceed the number L_1 of wires on the first level. Since the i th row of B can be described by $b_i \lceil \log_2 n \rceil$ bits, the entire matrix B can be described using $\sum_{j=1}^r b_j \lceil \log_2 n \rceil = L_1 \lceil \log_2 n \rceil$ bits.
2. By the same argument, the matrix C can be described using $L_2 \lceil \log_2 n \rceil$ bits.
3. For each output gate g_i , let B_i be the submatrix of B whose rows correspond to the d_i nodes on the middle layer seen by this gate. Let $\text{Im}(B_i) = \{B_i x : x \in \{0, 1\}^n\}$ be the column space of B_i . If this space has dimension t then any t linearly independent columns of B form its basis. Take the set $B'_i = \{u_1, \dots, u_t\}$ of the first t linearly independent columns of B_i , and call it the *first basis* of $\text{Im}(B_i)$.
4. Encode the behavior of g_i on this basis B'_i by a string of $t \leq d_i$ bits $g_i(u_1), \dots, g_i(u_t)$. The entire string, for all n output gates g_1, \dots, g_n , has length at most $\sum_{i=1}^n d_i = L_2$.

Having this encoding, we can recover the value $g_i(x)$ of the i th output gate on a given input $x \in \{0, 1\}^n$ as follows.

1. Compute $y = B_i x$. We can do this since the i th row of C tells us what rows of B appear in B_i , and we know the entire matrix B .
2. Take the first basis B'_i of $\text{Im}(B_i)$ and write y as a linear combination $y = \sum_{k=1}^t \lambda_k u_k$ of basis vectors.
3. Give $z_i = \sum_{k=1}^t \lambda_k g_i(u_k)$ as an output. We can compute this number since we know the values $g_i(u_1), \dots, g_i(u_t)$.

Since the circuit computes Ax , the i th output gate must compute the scalar product $\langle a_i, x \rangle$ of input vector x with the i th row a_i of A . Hence, $g_i(Bx) = \langle a_i, x \rangle$, meaning that g_i must be *linear* on $\text{Im}(B)$, and hence, also on $\text{Im}(B_i)$. Thus,

$$z_i = \sum_{k=1}^t \lambda_k g_i(u_k) = g_i \left(\sum_{k=1}^t \lambda_k u_k \right) = g_i(y) = g_i(B_i x) = g_i(Bx),$$

that is, z_i is a scalar product of x with the i th row of A , as desired. □

5. Conclusion

We have shown that linear operators requiring $\Omega(n^2 / \log n)$ wires in any depth-2 circuit exists, if we require that all gates on the middle layer or all gates on the output layer must be linear. We conjecture, however, that this also holds without this requirement. An even more important question is to prove that some *explicit* linear operator requires $n^{1+\epsilon}$ wires in general depth-2 circuits. The highest known explicit lower bounds, even for linear circuits, have the form $\Omega(n \ln^{3/2} n)$ [2, 10, 1].

References

- [1] N. Alon, P. Pudlák, Superconcentrators of depth 2 and 3; odd levels help (rarely), J. Comp. Sys. Sci. 48 (1994) 194–202.
- [2] N. Alon, M. Karchmer, A. Wigderson, Linear circuits over $\text{GF}(2)$, SIAM J. Comput. 19(6) (1990) 1064–1067.
- [3] S. Bublitz, Decomposition of graphs and monotone size of homogeneous functions, Acta Inform. 23 (1986) 689–696.
- [4] D. Y. Cherukhin, The lower estimate of complexity in the class of schemes of depth 2 without restrictions on a basis, Moscow Univ. Math. Bull. 60(4) (2005) 42–44.

- [5] D. Dolev, C. Dwork, N. Pippenger, A. Wigderson, Superconcentrators, generalizer and generalized connectors with limited depth, in: Proc. 15th STOC 1983 pp. 42–51.
- [6] S. Jukna, Entropy of operators or why matrix multiplication is hard for depth-two circuits, Theory of Comp. Syst. (2008) doi: 10.1007/s00224-008-9133-y.
- [7] S. Jukna, Representing $(0,1)$ -matrices by depth-2 circuits with arbitrary gates, Discrete Mathematics (submitted).
- [8] N. Pippenger, Superconcentrators, SIAM J. Comput. 6 (1977) 298–304.
- [9] N. Pippenger, Superconcentrators of depth 2, J. Comput. Syst. Sci. 24 (1982) 82–90.
- [10] P. Pudlák, Communication in bounded depth circuits, Combinatorica 14 (2) (1994) 203–216.
- [11] P. Pudlák and P. Savický. On shifting networks, Theoret. Comput. Sci. 116 (1993) 415–419.
- [12] P. Pudlák, V. Rödl, J. Sgall, Boolean circuits, tensor ranks, and communication complexity, SIAM J. Comput. 26(3) (1997) 605–633.
- [13] J. Radhakrishnan, A. Ta-Shma, Bounds for dispersers, extractors, and depth-two superconcentrators, SIAM J. Discrete Math. 13(1) (2000) 2–24.
- [14] R. Raz, A. Shpilka, Lower bounds for matrix product in bounded depth circuits with arbitrary gates, SIAM J. Comput. 32(2) (2003) 488–513.
- [15] Zs. Tuza, Coverings of graphs by complete bipartite subgraphs, complexity of 0-1 matrices, Combinatorica 4(1) (1984) 111–116.
- [16] L. Valiant, Graph-theoretic methods in low-level complexity, in Proc. 6th MFCS, Springer Lect. Notes in Comput. Sci. 53 (1977), pp. 162–176.