SIAM J. COMPUT.
Vol. 19, No. 6, pp. 1064–1067, December 1990

008

LINEAR CIRCUITS OVER GF(2)*

NOGA ALON†, MAURICIO KARCHMER‡, AND AVI WIGDERSON§

Abstract. For $n = 2^k$, let S be an $n \times n$ matrix whose rows and columns are indexed by GF(2)^k and, for $i, j \in$ GF(2)^k, $S_{i,j} = \langle i, j \rangle$, the standard inner product. Size-depth trade-offs are investigated for computing $S\mathbf{x}$ with circuits using only linear operations. In particular, linear size circuits with depth bounded by the inverse of an Ackerman function are constructed, and it is shown that depth two circuits require $\Omega(n \log n)$ size. The lower bound applies to any Hadamard matrix.

Key words. size-depth trade-offs, Boolean circuits, linear circuits, graph covers, Hadamard matrices

AMS(MOS) subject classifications. 05B20, 68R05, 94C15

1. Introduction. Let F be a field, and A a fixed $n \times n$ matrix with entries in F. There are no nontrivial lower bounds for computing the linear transformation $A\mathbf{x}$ where $\mathbf{x} \in F^n$ is an input, even if the circuit uses only linear operations.

When F is GF(2), linear operations are not more than exclusive-or gates. Counting arguments show that for a random matrix A, circuits of size $\Omega(n^2/\log n)$ are needed. In fact, $O(n^2/\log n)$ is an upper bound on the size needed for every matrix [B]. However, no explicit matrix A is known which requires superlinear size, even if the depth is restricted to be $O(\log n)$. (Valiant [V] has given an algebraic condition on matrices that would imply such a lower bound, but no matrix satisfying this condition has been constructed.)

In this note we consider H, the Boolean Hadamard matrix, and investigate size-depth trade-offs for computing $H\mathbf{x}$.

2. Definitions. A *Boolean Hadamard matrix* H is a matrix with entries in GF(2) and such that every two rows have Hamming distance $n/2$. Note that a Boolean Hadamard matrix can be constructed from a Hadamard matrix H' by $H = \frac{1}{2}(J + H')$ where J is the matrix of all ones. For $n = 2^k$, the *Sylvester Boolean matrix*, S, is one whose rows and columns are indexed by GF(2)^k and, for $i, j \in$ GF(2)^k, $S_{i,j} = \langle i, j \rangle$, the inner product of i and j. It is easy to show that S is a Boolean Hadamard matrix.

A *circuit* for $y = B\mathbf{x}$ where B is an $m \times n$ Boolean matrix is a DAG with n input nodes $x_1, \cdots, x_n$, m output nodes $y_1, \cdots, y_m$ and every noninput node computing the sum mod 2 of its inputs. (There is no bound on the fanin or on the fanout.) The *size* of the circuit is its number of edges. The *depth* is the length of the longest directed input-output path. Let $s(B)$ denote the size of the smallest circuit for $B\mathbf{x}$, and let $s_d(B)$ be the smallest size when the depth of the circuit is restricted to d.

The following lemma is important in understanding size-depth trade-offs.

LEMMA 2.1. *Let A, B be any two Boolean matrices. Then*:

(1) $s(B) = s(B^T)$, *where B^T is the transpose of B.*

(2) $s(AB) \leqq s(A) + s(B)$ *if A and B can be multiplied together.*

(3) $s(B) = s(B^\pi)$, *where B^π is the matrix B with rows permuted according to π.*

Furthermore, the same is true for depth restricted circuits where (2) *is replaced by* $s_{d_1+d_2}(AB) \leqq s_{d_1}(A) + s_{d_2}(B)$.

* Received by the editors February 14, 1989; accepted for publication (in revised form) April 18, 1990.
† Institute of Mathematics and Computer Science, Tel-Aviv University, Tel Aviv, Israel.
‡ Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139.
§ Institute of Mathematics and Computer Science, Hebrew University, Jerusalem 91904, Israel.

Proof. (1) Let $\mathscr{C}$ be a
nodes $\mathcal{O} = \{O_1, \cdots, O_m\}$. I
if $B_{i,j} = 1$. Hence, by switch
get a circuit for $B^T\mathbf{x}$.

(2) Let $\mathscr{C}_A(\mathscr{C}_B)$ be a
nodes $\mathcal{O}_A$ ($\mathcal{O}_B$). It is easy
computes $AB\mathbf{x}$.

(3) Let $\mathscr{C}$ be a circuit
according to π and redirec
The depth restricted c

3. Known results. The
THEOREM 3.1 [B]. *Fo*
THEOREM 3.2 [B]. *Fo*
FACT 3.1. *For every*
number of ones in B.

The only specific matr
matrix P, where $P_{i,j} = 1$ if a
functions as in [CFL]. Let
the Ackerman function. L
$\min\{j: A(j, j) \geqq n\}$.
THEOREM 3.3 [CFL].
In particular, this thec
COROLLARY 3.1. $s_{O(1)}$
COROLLARY 3.2. $s_{\alpha(n)}$

4. New results. We pr
THEOREM 4.1. $s_{2d}(S)$
In particular, S can be
from the behaviour of simil
to have size $\Omega(n \log n)$ reg
Though we would dar
we can only prove them fo
THEOREM 4.2. *If H is*
As far as we know, thi
in the restricted case of de
with a nonmonotone versio
graphs. The monotone qu
monotone, depth-three fo
combinatorial question bel
formulae. Let $G = ([n], [n]$
bipartite graphs $A_i \times B_i$, wh
only if (i, j) appears in an
define the cover number of
A Boolean matrix B
graph G_B. Given a depth-t
can associate a complete bi
nodes with edges to (from
left for the reader.

[1] $[n] = \{1, \cdots, n\}$.

Proof. (1) Let $\mathscr{C}$ be a circuit for $B\mathbf{x}$ with input nodes $\mathscr{I} = \{I_1, \cdots, I_n\}$ and output nodes $\mathscr{O} = \{O_1, \cdots, O_m\}$. Note that I_i has an odd number of paths to O_j if and only if $B_{i,j} = 1$. Hence, by switching the roles of $\mathscr{I}$ and $\mathscr{O}$ and reversing all edges in $\mathscr{C}$, we get a circuit for $B^T\mathbf{x}$.

(2) Let $\mathscr{C}_A(\mathscr{C}_B)$ be a circuit for $A\mathbf{x}$ ($B\mathbf{x}$) with input nodes $\mathscr{I}_A$ ($\mathscr{I}_B$) and output nodes $\mathscr{O}_A$ ($\mathscr{O}_B$). It is easy to see that the circuit obtained by identifying $\mathscr{O}_A$ with $\mathscr{I}_B$ computes $AB\mathbf{x}$.

(3) Let $\mathscr{C}$ be a circuit for $B\mathbf{x}$ with input nodes $\mathscr{I} = \{I_1, \cdots, I_n\}$. By permuting $\mathscr{I}$ according to π and redirecting the edges going out of $\mathscr{I}$ we get a circuit for $B^\pi\mathbf{x}$.

The depth restricted claims can be proved similarly. □

3. Known results. The following results are known (see [B]):

THEOREM 3.1 [B]. *For most $n \times n$ Boolean matrices B, $s(B) = \Omega(n^2/\log n)$.*

THEOREM 3.2 [B]. *For every $n \times n$ Boolean matrix B, $s_2(B) = O(n^2/\log n)$.*

FACT 3.1. *For every $n \times n$ Boolean matrix B, $s_1(B) = \omega(B)$, where $\omega(B)$ is the number of ones in B.*

The only specific matrix that has been studied in some detail is the parallel prefix matrix P, where $P_{i,j} = 1$ if and only if $i \leqq j$. We will first define some very slowly growing functions as in [CFL]. Let $A(0, j) = 2j$; $A(i, 1) = 2$; and $A(i, j) = A(i-1, A(i, j-1))$ be the Ackerman function. Let $\alpha(n, d) = \min\{j: A(d, j) \geqq n\}$. Furthermore, let $\alpha(n) = \min\{j: A(j, j) \geqq n\}$.

THEOREM 3.3 [CFL]. $s_d(P) = O(n\alpha(n, d))$.

In particular, this theorem implies the following corollaries.

COROLLARY 3.1. $s_{O(1)}(P) = O(n\alpha(n))$.

COROLLARY 3.2. $s_{\alpha(n)}(P) = O(n)$.

4. New results. We present the following results.

THEOREM 4.1. *$s_{2d}(S) = O(n\alpha(n, d))$ where S is the Sylvester Boolean matrix.*

In particular, S can be computed in linear size and $\alpha(n)$ depth. This seems different from the behaviour of similar matrices (say FFT) over other fields which are conjectured to have size $\Omega(n \log n)$ regardless of the depth.

Though we would dare to conjecture that the above bounds are the best possible, we can only prove them for $d = 2$, namely, Theorem 4.2.

THEOREM 4.2. *If H is a Boolean Hadamard matrix, then $s_2(H) = \Omega(n \log n)$.*

As far as we know, this is the first nontrivial lower bound for these circuits, even in the restricted case of depth two. Another interpretation of Theorem 4.2 has to do with a nonmonotone version of the problem of covering a graph with complete bipartite graphs. The monotone question was studied in connection to lower bounds on monotone, depth-three formulae (Hansel, Krichevskii [S]). In the same vein, the combinatorial question below relates to lower bounds on nonmonotone, depth-three formulae. Let $G = ([n], [n], E)$ be a bipartite graph.[1] Let $\mathscr{K}$ be a collection of complete bipartite graphs $A_i \times B_i$, where $A_i, B_i \subseteq [n]$. We say that $\mathscr{K}$ covers G if $(i, j) \in E$ if and only if (i, j) appears in an odd number of graphs in $\mathscr{K}$. Let $|\mathscr{K}| = \sum_i (|A_i| + |B_i|)$ and define the cover number of G, $\beta(G)$ as the minimum of $|\mathscr{K}|$, where $\mathscr{K}$ covers G.

A Boolean matrix B can be considered as the adjacency matrix of a bipartite graph G_B. Given a depth-two circuit $\mathscr{C}$ for B, and a node v in the middle layer, we can associate a complete bipartite graph $A(v) \times B(v)$ where $A(v)$ ($B(v)$) is the set of nodes with edges to (from) v. With this as a hint, the proof of the following fact is left for the reader.

[1] $[n] = \{1, \cdots, n\}$.

FACT 4.1. $s_2(B) = \beta(G_B)$.

We thus get the following corollary.

COROLLARY 4.1. $\beta(G_S) = \Theta(n \log n)$ *where* S *is the Sylvester Boolean matrix.*

It is an open problem to construct a graph G with $\beta(G) = \Omega(n^{1+\varepsilon})$.

5. Upper bounds. We now prove Theorem 4.1. Recall that we are working over GF (2). Let $n = 2^k$. Then, by definition, $S = DD^T$, where D is the $2^k \times k$ matrix whose rows are all the vectors in GF $(2)^k$. By Lemma 2.1, we have that $s(S) \leqq s(D) + s(D^T) = 2s(D)$.

A *Grey code* is an ordering $v_1, \cdots, v_{2^k}$ of GF $(2)^k$ such that $\omega(v_i \oplus v_{i+1}) = 1$ for all i (i.e., the Hamming distance of every two consecutive vectors is one). Once again, by Lemma 2.1 we can assume, without loss of generality, that the rows of D are $v_1, \cdots, v_{2^k}$ as above.

Let $u_1 = (0, 0, \cdots, 0)$, and let $u_i = v_{i-1} \oplus v_i$ for $2 \leqq i \leqq 2^k$. Let U be the matrix whose rows are $u_1, \cdots, u_{2^k}$. Clearly, $s(U) \leqq \omega(U) \leqq n - 1$. Furthermore, $D = PU$ where P is the parallel prefix matrix as defined in § 3. We thus get $s(S) \leqq 2(s(P) + n - 1)$ and, by Theorem 3.3, $s_{2d}(S) = O(n\alpha(n, d))$.

6. Lower bounds. We give two different lower bounds for $s_2(H)$. The first is weaker than Theorem 4.2, but uses only the combinatorial structure of Hadamard matrices. The proof of Theorem 4.2 will use the algebraic structure of Hadamard matrices, together with results of Valiant [V] and Alon and Maass [AM].

Let $\mathscr{C}$ be a circuit for H and let $\mathscr{I}$ and $\mathscr{O}$ be the set of inputs and outputs of $\mathscr{C}$, respectively. Furthermore, let $\mathscr{M}$ be the set of nodes of $\mathscr{C}$ in the middle layer. Without loss of generality, we may assume that all edges of $\mathscr{C}$ are either in $\mathscr{I} \times \mathscr{M}$ or in $\mathscr{M} \times \mathscr{O}$.

The following combinatorial fact will be needed. A *sunflower* with k petals is a set system $\{R_1, \cdots, R_k\}$, where $R_i = C \cup Z_i$ and the Z_i's are pairwise disjoint. C is called the *center* of the sunflower and the Z_i's are called the *petals.* The following theorem is well known.

THEOREM 6.1 (Erdös–Rado). *Every family of* $r!k^r$ *sets each of which has cardinality less than* r *contains a sunflower with* k *petals.*

Now we prove the weak version of Theorem 4.2.

THEOREM 6.2. *For any Hadamard matrix* H, $s_2(H) = \Omega(n \log n / \log\log n)$.

Proof. Let E_1 be the set of edges of $\mathscr{C}$ in $\mathscr{I} \times \mathscr{M}$, and let E_2 be the set of edges in $\mathscr{M} \times \mathscr{O}$. Let $m = c \log n / \log\log n$, for some constant c to be determined later. We will show that if $|E_1| \leqq mn$ then $|E_2| \geqq mn/2$, proving the theorem.

Let $S \subseteq \mathscr{I}$ be the set of inputs with (out)degree at most $2m$. Clearly, $|S| \geqq n/2$. For a vertex $i \in S$, let $T_i \subseteq \mathscr{M}$ be its set of neighbours. The collection of T_i's for $i \in S$ forms a set system of many small sets. By the sunflower theorem, there exists $R \subseteq S$, $|R| = 2m$ such that $\{T_i \mid i \in R\}$ form a sunflower. Let C be the center of the sunflower and $\{Z_i \mid i \in R\}$ its petals. Let F_i denote the edges of E_2 emanating from Z_i. □

CLAIM 6.1. *For every* $i, j \in R$, $|F_i| + |F_j| \geqq n/2$.

The theorem follows from Claim 6.1 by pairing the elements of R into $\{(i_1, j_1), \cdots, (i_m, j_m)\}$ arbitrarily so that

$$|E_2| \geqq \left| \bigcup_{i \in R} F_i \right| = \sum_{i \in R} |F_i| = \sum_{l=1}^{m} (|F_{i_l}| + |F_{j_l}|) \geqq \frac{mn}{2}.$$

Proof of claim. Let $K \subseteq \mathscr{O}$ be the set of outputs that depend on exactly one of the inputs x_i and x_j. By the definition of H, $|K| = n/2$. For every element $k \in K$, the number of paths from i to k must have a different parity than the number of paths from j to

k. But then, there must
every $k \in K$, $|F_i| + |F_j| \geqq$
Now we prove T
$L(S, T) \subseteq \mathscr{M}$ be the nod
$|S| \times |T|$ submatrix of H
[V] is the key to our pr
FACT 6.1. $|L(S, T)|$
Intuitively, this fac
values cannot be compr
[AM].
THEOREM 6.3. *If* f
$\varepsilon \log n$, *then* $\mathscr{C}$ *has at l*
Hence it will suffic
CLAIM 6.2. *Let* $|S|$
Proof. Assume this
so that one appears at le
this column has more o
of size $(n^{1/2+\varepsilon}) \times (n^{1/2+}$
monochromatic submat

[AM] N. ALON AND W. M
Proc. 27th Annu
Society, Washing
[B] S. BUBLITZ, *Decomp*
Inform., 23 (198
[CFL] A. K. CHANDRA, S. F
in Proc. 15th An
Machinery, New
[L] P. ERDÖS AND J. H. S
1974.
[S] J. SAVAGE, *The Com*
[V] L. VALIANT, *Graph-*
Science, Vol. 53,

k. But then, there must exist at least one edge from $Z_i \cup Z_j$ to k. Since this is true for every $k \in K$, $|F_i| + |F_j| \geq n/2$. □

Now we prove Theorem 4.2. $S_2(H) = \Omega(n \log n)$. For $S \subseteq \mathscr{I}$ and $T \subseteq \mathscr{O}$, let $L(S, T) \subseteq \mathscr{M}$ be the nodes in $\mathscr{M}$ connected both to nodes in S and T. Let $H_{S,T}$ be the $|S| \times |T|$ submatrix of H indexed by S and T. The following observation of Valiant [V] is the key to our proof.

FACT 6.1. $|L(S, T)| \geq \text{rank}\,(H_{S,T})$.

Intuitively, this fact says that the information contained in linearly independent values cannot be compressed. We will use the following theorem of Alon and Maass [AM].

THEOREM 6.3. *If for every $S \subseteq \mathscr{I}$ and $T \subseteq \mathscr{O}$ with $|S| = |T| = n^{1/2+2\varepsilon}$, $|L(S, T)| \geq \varepsilon \log n$, then $\mathscr{C}$ has at least $\Omega(n \log n)$ edges.*

Hence it will suffice to prove Claim 6.2.

CLAIM 6.2. *Let $|S| = |T| = n^{1/2+2\varepsilon}$, then* $\text{rank}\,(H_{S,T}) \geq \varepsilon \log n$.

Proof. Assume this is not so. Then there are at most n^{ε} different columns in $H_{S,T}$ so that one appears at least $n^{1/2+\varepsilon}$ many times. Without loss of generality, assume that this column has more ones than zeros. Then H contains a monochromatic submatrix of size $(n^{1/2+\varepsilon}) \times (n^{1/2+\varepsilon}/2)$, which contradicts the well-known fact [L] that every monochromatic submatrix of H has area at most $4n$. □

REFERENCES

[AM] N. ALON AND W. MAASS, *Meanders, Ramsey theory and lower bounds for branching programs*, in Proc. 27th Annual IEEE Symposium on Foundations of Computer Science, IEEE Computer Society, Washington, DC, 1986, pp. 410-417.

[B] S. BUBLITZ, *Decomposition of graphs and monotone formula size of homogeneous functions*, Acta Inform., 23 (1986), pp. 689-696.

[CFL] A. K. CHANDRA, S. FORTUNE, AND R. LIPTON, *Unbounded fanin circuits and associative functions*, in Proc. 15th Annual ACM Symposium on Theory of Computing, Association for Computing Machinery, New York, 1983, pp. 52-60.

[L] P. ERDÖS AND J. H. SPENCER, *Probabilistic Methods in Combinatorics*, Academic Press, New York, 1974.

[S] J. SAVAGE, *The Complexity of Computing*, John Wiley, New York, 1976.

[V] L. VALIANT, *Graph-theoretic arguments in low-level complexity*, in Lecture Notes in Computer Science, Vol. 53, Springer-Verlag, Berlin, New York, 1977.