Bounds on Certain Multiplications of Affine Combinations

Joan Boyar [*]
Loyola University Chicago
Odense University

Faith Fich [†]
University of Toronto

Kim S. Larsen [‡]
Aarhus University

Abstract

Let A and B be $n \times n$ matrices the entries of which are affine combinations of the variables $a_1, \ldots, a_m, b_1, \ldots, b_m$ over GF(2). Suppose that, for each i, $1 \leq i \leq m$, the term $a_i b_i$ is an element of the product matrix $C = A \cdot B$. What is the maximum value that m can have as a function of n? This question arises from a recent technique for improving the communication complexity of zero-knowledge proofs.

The obvious upper bound of n^2 is improved to $n^2/\sqrt[3]{3} + O(n)$. Tighter bounds are obtained for smaller values of n. The bounds for $n = 2$, $n = 3$, and $n = 4$ are tight.

1 Introduction

The problem described in the abstract and discussed in this paper is motivated by recent results in cryptography. A new technique for improving

[*] Supported in part by NSA Grant Number MDA90-H-4016. Some of this work was done while visiting the University of Toronto and Aarhus University.

[†] Supported in part by Natural Science and Engineering Research Council of Canada grant A-9176 and the Information Technology Research Centre of Ontario

[‡] Some of this work was done while visiting the University of Toronto.

the communication complexity of zero-knowledge proofs for circuit satisfiability was presented in [1]. The key idea is that the Prover shows that all the inputs and outputs to the AND gates are correct by showing that a matrix multiplication is correct. Suppose that the inputs to m AND gates are (a_1, b_1), $(a_2, b_2), ..., (a_m, b_m)$, and that the outputs are $c_1, c_2, ..., c_m$, respectively. Given encryptions for the a_i's, b_i's, and c_i's, the Prover is trying to show that the following equalities hold in GF(2): $a_1b_1 = c_1, \quad a_2b_2 = c_2, ..., \quad a_mb_m = c_m$. The variables $a_1, a_2, ..., a_m$ are put in an $n \times n$ matrix A which has zeros as its remaining elements. The variables $b_1, b_2, ..., b_m$ are put in an $n \times n$ matrix B which also has zeros as its remaining elements. These variables and zeros are placed so that every one of the c_i's is contained somewhere in the product matrix $C = A \cdot B$. For example, if the a_i's and the b_i's are on the diagonals of their respective matrices, and if the other entries of these matrices are 0, the c_i's will be on the diagonal of the product. The usefulness of the technique in [1], however, depends on m being significantly larger than n; the larger, the better.

The smallest interesting example has $m = 6$ and $n = 3$:

$$\begin{pmatrix} a_1 & a_2 & 0 \\ a_3 & 0 & a_4 \\ 0 & a_5 & a_6 \end{pmatrix} \cdot \begin{pmatrix} b_3 & b_1 & 0 \\ b_5 & 0 & b_2 \\ 0 & b_6 & b_4 \end{pmatrix} = \begin{pmatrix} a_1b_3 + a_2b_5 & a_1b_1 & a_2b_2 \\ a_3b_3 & a_3b_1 + a_4b_6 & a_4b_4 \\ a_5b_5 & a_6b_6 & a_5b_2 + a_6b_4 \end{pmatrix}.$$

A construction in [1] gives the values $m = 32^t$ and $n = 8^t$ for any positive integer t. Thus, it is possible to put $m = n^{5/3}$ entries in an $n \times n$ matrix if n is a power of 8. Although this is the best known result in the practical range, an asymptotic improvement of theoretical interest, also described in [1], has been discovered by Szemerédi [3], using a result of [2]. It is possible to put m entries in matrices of size $n \times n$, where $n \leq (\sqrt{m})^{1+\varepsilon_m}$ and $\varepsilon_m = 4\sqrt{2}/\sqrt{\lg m}$, which is better than the other construction, provided that $m \geq 2^{128}$. Since ε_m approaches zero as m approaches infinity, m is nearly linear in n^2, the number of entries in the matrix.

In all these examples, the matrix A contains only a_i's and zeros and the matrix B contains only b_i's and zeros. This restriction is neither stated nor necessary for the technique described in [1]. In fact, because of various properties of the encryption scheme used, the entries in both A and B could also have the form $\sum_{j=1}^{k} x_j$ where each $x_j \in \{a_1, a_2, ..., a_m, b_1, b_2, ..., b_m, 1\}$.

Thus, these entries can be affine combinations of the variables. For example, in 2 × 2 matrices, one could have:

$$\left(\begin{array}{cc} a_1+a_2+b_1+1 & b_1 \\ a_1+a_2 & a_1 \end{array}\right) \cdot \left(\begin{array}{cc} b_1 & b_2 \\ a_2 & b_2 \end{array}\right) = \left(\begin{array}{cc} a_1b_1 & a_1b_2+a_2b_2+b_2 \\ a_1b_1+a_2b_1+a_1a_2 & a_2b_2 \end{array}\right).$$

This example just gives $m = n = 2$, which is no improvement over what can be done without using affine combinations. In fact, there are no known examples where removing the original restrictions does give any improvement.

For a given n, let $M(n)$ denote the maximum value that m can have. Then $M(n) \leq n^2$. In this paper, we improve this bound to $n^2/\sqrt[3]{3} + O(n)$. This bound is definitely not tight for small n. We prove other results which give tighter bounds when n is small, and exact bounds for $n = 2$, $n = 3$, and $n = 4$.

2 Asymptotic Bounds

Given a matrix C, choose $k \geq 2$ rows and $\lfloor n/k \rfloor + 1$ columns and consider the $k \times (\lfloor n/k \rfloor + 1)$ submatrix of C consisting of the intersection of these rows and columns. In this section, we show that no such submatrix can consist entirely of distinct c_i's and use this fact to obtain an upper bound on $M(n)$.

In order to prove this, we use a result from the study of straight-line programs over fields. These are programs in which the ith statement has the form $V_i \leftarrow U_j$ or the form $V_i \leftarrow U_j \odot U_k$, where each of U_j and U_k is either an input to the program, some variable V_l with $l < i$, or a field constant, and $\odot$ is addition or multiplication. The next lemma follows from results of [5]. The proof given here is more direct and it is included for the sake of completeness.

Lemma 1 *Let $a_1, a_2, ..., a_k$ and $b_1, b_2, ..., b_k$ be independent variables over GF(2). Then any straight-line program for computing the inner product $\sum_{i=1}^{k} a_ib_i$ requires at least k (nonscalar) multiplications.*

Proof Suppose the claim is false. Consider the smallest value k for which there is a straight-line program P computing the inner product $\sum_{i=1}^{k} a_ib_i$ using less than k multiplications. Since even the product a_1b_1 cannot be computed without any multiplications, P must contain at least one (non-scalar) multiplication. Consider the first statement $z \leftarrow x \cdot y$ in P that involves a multiplication. Both x and y are affine combinations of one or more of the variables. Without loss of generality, say $x = a_k + x'$ where x' is a constant or an affine combination of other variables.

Construct a straight-line program P' from P by prepending the statements $b_k \leftarrow 0$ and $a_k \leftarrow x'$ and replacing the statement $z \leftarrow x \cdot y$ by $z \leftarrow 0$. Then P' computes $\sum_{i=1}^{k-1} a_i b_i$. None of the new statements in P' involve any multiplications, so P' uses fewer than $k-1$ multiplications. This contradicts the minimality of k. □

Lemma 2 *Let $a_1, a_2, ..., a_m$ and $b_1, b_2, ..., b_m$ be distinct variables over GF(2), and suppose $c_i = a_i b_i$ for $1 \leq i \leq m$. Let A, B, and C be $n \times n$ matrices such that $A \cdot B = C$. Suppose that the entries of A and B are affine combinations of the variables. If there exists an $s \times t$ submatrix of C in which every element is distinct and is one of the c_i's, then $st \leq n$. Furthermore, if $st = n$, then no other element in any of those s rows or t columns of C is a different c_i.*

Proof Consider an $s \times t$ submatrix of C consisting of the intersection of rows $r_1, r_2, ..., r_s$ and columns $q_1, q_2, ..., q_t$ and which contains the entries $c_1, \ldots, c_{st}$. Since $C = A \cdot B$,

$$\begin{aligned}\sum_{i=1}^{s\times t} a_i b_i = \sum_{i=1}^{s\times t} c_i &= \sum_{j=1}^{s}\sum_{l=1}^{t} C[r_j, q_l] \\ &= \sum_{j=1}^{s}\sum_{l=1}^{t}\sum_{k=1}^{n} A[r_j, k] \cdot B[k, q_l] \\ &= \sum_{j=1}^{s}\sum_{k=1}^{n} A[r_j, k] \cdot \left(\sum_{l=1}^{t} B[k, q_l]\right) \\ &= \sum_{k=1}^{n}\left(\sum_{j=1}^{s} A[r_j, k]\right) \cdot \left(\sum_{l=1}^{t} B[k, q_l]\right).\end{aligned}$$

Each of the terms $A[r_j, k]$ and $B[k, q_l]$ is an affine combination of the a_i's and b_i's, so the sums $\sum_{j=1}^{s} A[r_j, k]$ and $\sum_{l=1}^{t} B[k, q_l]$ can be computed without any multiplications. Thus the right hand side can be computed by a straight-line program with only n multiplications. By lemma 1, the left hand side requires at least st multiplications. Thus, $st \leq n$.

Now assume $st = n$. Then, for each $k \in \{1, \ldots, n\}$, $\sum_{j=1}^{s} A[r_j, k]$ is an affine combination of the variables $a_1, \ldots, a_n, b_1, \ldots, b_n$. To see why, suppose $\sum_{j=1}^{s} A[r_j, k'] = a_{n+1}+d$, where $k' \in \{1, \ldots, n\}$ and d is an affine combination of variables excluding a_{n+1}. Let A', B', and C' be the matrices obtained by replacing all occurrences of a_{n+1} by d in A, B, and C, respectively. Then $A' \cdot B' = C'$. Furthermore, since $C[r_j, q_l]$ does not contain a_{n+1}, $C'[r_j, q_l] = C[r_j, q_l]$ for all $j \in \{1, \ldots, s\}$, $l \in \{1, \ldots, t\}$. Thus

$$\begin{aligned}
\sum_{i=1}^{s\times t} a_i b_i &= \sum_{j=1}^{s}\sum_{l=1}^{t} C[r_j, q_l] \\
&= \sum_{j=1}^{s}\sum_{l=1}^{t} C'[r_j, q_l] \\
&= \sum_{k=1}^{n} \left(\sum_{j=1}^{s} A'[r_j, k]\right) \cdot \left(\sum_{l=1}^{t} B'[k, q_l]\right).
\end{aligned}$$

Since $\sum_{j=1}^{s} A[r_j, k'] = a_{n+1} + d$, it follows that $\sum_{j=1}^{s} A'[r_j, k'] = 0$. But this implies that the right hand side can be computed by a straight-line program using only $n-1$ multiplications, contradicting lemma 1.

Similarly, $\sum_{l=1}^{t} B[k, q_l]$ is an affine combination of the variables $a_1, \ldots, a_n, b_1, \ldots, b_n$, for each $k \in \{1, \ldots, n\}$.

In fact, for each $j \in \{1, \ldots, s\}$ and $k \in \{1, \ldots, n\}$, $A[r_j, k]$ is, itself, an affine combination of the variables $a_1, \ldots, a_n, b_1, \ldots, b_n$. Suppose, to the contrary, that $A[r, k'] = a_{n+1} + d$, where $r \in \{r_1, \ldots, r_s\}$, $k' \in \{1, \ldots, n\}$, and d is an affine combination of variables excluding a_{n+1}. Let $e = \sum_{j=1}^{s} A[r_j, k']$ and let A' and A'' be obtained by replacing all occurrences of a_{n+1} in A by 0 and e, respectively. Define B', B'', C', and C'' analogously. Then $A' \cdot B' = C'$ and $A'' \cdot B'' = C''$. Since $\sum_{j=1}^{s} A[r_j, k]$ and $\sum_{l=1}^{t} B[k, q_l]$ are not functions of a_{n+1}, for any $k \in \{1, \ldots, n\}$, and $C[r_i, q_l]$ is not a function of a_{n+1}, for any $j \in \{1, \ldots, s\}$ and $l \in \{1, \ldots, t\}$,

$$\begin{aligned}
&\textstyle\sum_{j=1}^{s} A[r_j, k] = \sum_{j=1}^{s} A'[r_j, k] = \sum_{j=1}^{s} A''[r_j, k], \\
&\textstyle\sum_{l=1}^{t} B[k, q_l] = \sum_{l=1}^{t} B'[k, q_l] = \sum_{l=1}^{t} B''[k, q_l], \text{ and} \\
&C[r_j, q_l] = C'[r_j, q_l] = C''[r_j, q_l].
\end{aligned}$$

Thus,

$$\begin{aligned}
\sum_{l=1}^{t}\sum_{k=1}^{n} A'[r,k]\cdot B'[k,q_l] &= \sum_{l=1}^{t} C'[r,q_l] \\
&= \sum_{l=1}^{t} C''[r,q_l] \\
&= \sum_{l=1}^{t}\sum_{k=1}^{n} A''[r,k]\cdot B''[k,q_l] \\
&= \sum_{k=1}^{n} A''[r,k]\cdot\left(\sum_{l=1}^{t} B''[k,q_l]\right) \\
&= \sum_{k=1}^{n} A''[r,k]\cdot\left(\sum_{l=1}^{t} B'[k,q_l]\right) \\
&= \sum_{l=1}^{t}\sum_{k=1}^{n} A''[r,k]\cdot B'[k,q_l]
\end{aligned}$$

$$\begin{aligned}
\text{and } A''[r,k'] + \sum_{\substack{j=1\\ r_j\neq r}}^{s} A'[r_j,k'] &= A''[r,k'] + A'[r,k'] + \sum_{j=1}^{s} A'[r_j,k'] \\
&= A''[r,k'] + A'[r,k'] + \sum_{j=1}^{s} A[r_j,k'] \\
&= (e+d)+d+e = 0.
\end{aligned}$$

From these facts, it follows that

$$\begin{aligned}
\sum_{i=1}^{s\times t} a_i b_i &= \sum_{j=1}^{s}\sum_{l=1}^{t}\sum_{k=1}^{n} A'[r_j,k]\cdot B'[k,q_l] \\
&= \sum_{l=1}^{t}\sum_{k=1}^{n} A'[r,k]\cdot B'[k,q_l] + \sum_{\substack{j=1\\ r_j\neq r}}^{s}\sum_{l=1}^{t}\sum_{k=1}^{n} A'[r_j,k]\cdot B'[k,q_l] \\
&= \sum_{l=1}^{t}\sum_{k=1}^{n} A''[r,k]\cdot B'[k,q_l] + \sum_{\substack{j=1\\ r_j\neq r}}^{s}\sum_{l=1}^{t}\sum_{k=1}^{n} A'[r_j,k]\cdot B'[k,q_l] \\
&= \sum_{k=1}^{n}\left(A''[r,k] + \sum_{\substack{j=1\\ r_j\neq r}}^{s} A'[r_j,k]\right)\cdot\left(\sum_{l=1}^{t} B'[k,q_l]\right)
\end{aligned}$$

$$= \sum_{\substack{k=1\\k\neq k'}}^{n} \left(A''[r,k] + \sum_{\substack{j=1\\r_j\neq r}}^{s} A'[r_j,k] \right) \cdot \left(\sum_{l=1}^{t} B'[k,q_l] \right).$$

But this contradicts lemma 1, since the right hand side can be computed by a straight-line program using only $n-1$ multiplications.

Similarly, for each $l \in \{1,\ldots,t\}$ and $k \in \{1,\ldots,n\}$, $B[k,q_l]$ is an affine combination of the variables $a_1,\ldots,a_n,b_1,\ldots,b_n$.

If $a_{n+1}b_{n+1} = C[r,q] = \sum_{k=1}^{n} A[r,k]\cdot B[k,q]$, then there exists $k \in \{1,\ldots,n\}$ such that a_{n+1} is contained in $A[r,k]$ and b_{n+1} is contained in $B[k,q]$, or vice versa. This implies that $r \notin \{r_1,\ldots,r_s\}$ and $q \notin \{q_1,\ldots,q_l\}$. □

Given an $n\times n$ matrix C with m distinct c_i's, construct an $n \times n$ matrix D with m ones (corresponding to distinct c_i's) and n^2-m zeros. If C is the product of two matrices the entries of which are affine combinations of the variables $a_1,\ldots,a_m,b_1,\ldots,b_m$, we say that the zero-one matrix D is a *representative* matrix.

Corollary 1 *If an $n\times n$ representative matrix has an $s\times t$ submatrix containing only ones, then $st \leq n$. Furthermore, if $st = n$, then no other element in any of those s rows or t columns is one.*

To prove an upper bound on $M(n)$, it suffices to prove an upper bound on the maximum number of ones in any $n\times n$ representative matrix. This is a special case of the problem: determine the least positive integer $k_{i,j}(m,n)$ such that if a zero-one matrix of size $m\times n$ contains $k_{i,j}(m,n)$ ones, then it must have an $i\times j$ submatrix containing only ones. This is a generalization of a problem originally posed by Zarankiewicz [10]. The first upper bound on this problem,

$$k_{i,j}(m,n) \leq 1 + (i-1)n + \lfloor (j-1)^{1/i} n^{1-1/i} m \rfloor,$$

was given by Hyltén-Cavallius [9], using the methods of Kövari, Sós, and Turán [8]. This has been improved slightly by others, including Guy and Znám [6] and Roman [7]. Tighter results have been found for small values of i and j. In particular, Hyltén-Cavallius [9] has shown that

$$k_{2,j}(m,n) \leq 1 + \lfloor \tfrac{1}{2}n + \sqrt{(j-1)nm(m-1) + \tfrac{1}{4}n^2} \rfloor.$$

All of these upper bounds are obtained using Dirichlet's pigeonhole principle as the main tool, and we use the same techniques in lemma 5.

The following lemmas give upper bounds on the number of ones in an $n \times n$ representative matrix and thus upper bounds on $M(n)$.

Lemma 3 *If an $n \times n$ representative matrix D contains more than $(1 - 1/k)n^2 - (k-2)n$ ones, then it contains no $k \times \lceil n/k \rceil$ submatrix consisting entirely of ones.*

Proof If n is not divisible by k, then $k\lceil n/k \rceil > n$, so, by corollary 1, D does not contain a $k \times \lceil n/k \rceil$ submatrix consisting entirely of ones. Therefore suppose that n is divisible by k and D contains a $k \times n/k$ submatrix consisting entirely of ones. Then, by corollary 1, none of the $k(n - n/k)$ other elements in the same rows and none of the $(n-k)n/k$ other elements in the same columns are ones. Hence D contains at most $n^2 - nk + n - n^2/k + n = (1 - 1/k)n^2 - (k-2)n$ ones. □

Lemma 4 *Suppose D is an $n \times n$ representative matrix, with $n \geq 2$. Then D contains at most $\frac{n}{2}\left(1 + \sqrt{1 + 4(\lceil n/2 \rceil - 1)(n-1)}\right) = n^2/\sqrt{2} + O(n)$ ones.*

Proof By lemma 3, we may assume that D does not contain a $2 \times \lceil n/2 \rceil$ submatrix consisting entirely of ones. Thus, we can apply the result of Hyltén-Cavallius [9] on $k_{2,j}(m,n)$, setting $j = \lceil n/2 \rceil$ and $m = n$. Since $k_{2,\lceil n/2 \rceil}(n,n)$ is the number of ones necessary to ensure that a $2 \times \lceil n/2 \rceil$ submatrix containing only ones exists, the value we need is one less. □

This result implies that $M(2) \leq 2$, $M(3) \leq 6$, and $M(4) \leq 9$. The lower bounds, $M(2) \geq 2$ and $M(3) \geq 6$, follow from the examples in the introduction. The following example, in which each * represents some uninteresting bilinear form, gives that $M(4) \geq 9$.

$$\begin{pmatrix} a_1 & a_2 & a_3 & 0 \\ 0 & a_4 & 0 & a_5 \\ a_6 & 0 & 0 & a_7 \\ 0 & 0 & a_8 & a_9 \end{pmatrix} \cdot \begin{pmatrix} b_1 & 0 & 0 & b_6 \\ 0 & b_2 & 0 & b_4 \\ 0 & 0 & b_3 & b_8 \\ b_9 & b_7 & b_5 & 0 \end{pmatrix} = \begin{pmatrix} a_1b_1 & a_2b_2 & a_3b_3 & * \\ * & * & a_5b_5 & a_4b_4 \\ * & a_7b_7 & * & a_6b_6 \\ a_9b_9 & * & * & a_8b_8 \end{pmatrix}.$$

Thus $M(2) = 2$, $M(3) = 6$, and $M(4) = 9$.

The proof of lemma 4 only used corollary 1 for $s = 2$. The same technique can also be applied for other values of s. Using the standard pigeonhole technique, the value $s = 3$ gives the best result asymptotically. The results of [9], [6], and [7] all give the asymptotic result we obtain in the following lemma, but since our problem is less general, the result given here is slightly tighter.

Lemma 5 *Suppose D is an $n \times n$ representative matrix, with $n \geq 4$. Let*

$$u = \tfrac{1}{2}\left(\lceil n/3 \rceil - 1\right)(n-1)(n-2) \ \ \textit{and} \ \ v = \sqrt{u^2 - 1/27}.$$

Then D contains at most

$$n\left(1 + \sqrt[3]{u+v} + \sqrt[3]{u-v}\right) = n^2/\sqrt[3]{3} + O(n)$$

ones.

Proof By lemma 3, we may assume that D does not contain a $3 \times \lceil n/3 \rceil$ submatrix consisting entirely of ones. Consider any set of three rows. Then the number of columns in which all three rows have value one is no more than $\lceil n/3 \rceil - 1$. Let T be the sum of this quantity, taken over all $\binom{n}{3}$ sets of three rows. Then $T \leq \left(\lceil n/3 \rceil - 1\right)\binom{n}{3}$.

For $1 \leq i \leq n$, let k_i denote the number of ones in the ith column. Then $m = \sum_{i=1}^{n} k_i$ is the number of ones in the entire matrix and $T = \sum_{i=1}^{n} \binom{k_i}{3}$. By convexity, $T \geq n\binom{m/n}{3}$. This implies that $\left(\lceil n/3 \rceil - 1\right)(n-1)(n-2) \geq \frac{m}{n}\left(\frac{m}{n} - 1\right)\left(\frac{m}{n} - 2\right)$. Let $x = m/n - 1$. Then $x^3 - x - 2u \leq 0$. Since $u^2 - 1/27 > 0$ for $n \geq 4$, the formula for the roots of cubic equations implies that $x \leq \sqrt[3]{u+v} + \sqrt[3]{u-v}$ and, hence, $m \leq n\left(1 + \sqrt[3]{u+v} + \sqrt[3]{u-v}\right)$. $\square$

For some small values of n, the upper bound on $M(n)$ implied by the following result is better. Like lemma 4, it only uses corollary 1 for $s = 2$.

Lemma 6 *Suppose D is an $n \times n$ representative matrix, with $n \geq 2$. Then D contains at most $K = (n-1)\left(\lceil 3n/2 \rceil - 2\right) - (n-2)\left(\lceil 3n/4 \rceil - 1\right) + 3$ ones.*

Proof For $1 \leq i \leq n$, let k_i denote the number of ones in the ith row. Without loss of generality, assume $k_i \geq k_{i+1}$ for $1 \leq i < n$.

If $k_1 \leq \lceil 3n/4 \rceil - 2$, then the total number of ones in D is

$$\sum_{i=1}^{n} k_i \leq n\left(\lceil 3n/4 \rceil - 2\right) \leq K.$$

Therefore, assume $k_1 \geq \lceil 3n/4 \rceil - 1$.

If any row, other than the first, contains $\lceil 3n/2 \rceil - k_1$ ones, then D contains a $2 \times \lceil n/2 \rceil$ submatrix consisting entirely of ones. Thus, by lemma 3, we may assume that no row, other than the first, contains more than $\lceil 3n/2 \rceil - k_1 - 1$ ones. Let s be the number of rows which contain exactly this many ones. Then the total number of ones in the matrix is bounded by $k_1 + s\left(\lceil 3n/2 \rceil - k_1 - 1\right) + (n - s - 1)\left(\lceil 3n/2 \rceil - k_1 - 2\right)$ which equals $s - (n-2)k_1 + (n-1)\left(\lceil 3n/2 \rceil - 2\right)$.

The s rows must have ones where row one has zeros. By corollary 1, we must have that $s(n - k_1) \leq n$, so the number of ones in the matrix is bounded by $\lfloor \frac{n}{n-k_1} \rfloor - (n-2)k_1 + (n-1)\left(\lceil 3n/2 \rceil - 2\right) \leq 3 - (n-2)\left(\lceil 3n/4 \rceil - 1\right) + (n-1)\left(\lceil 3n/2 \rceil - 2\right)$. □

The following examples show that lemma 4 gives a tight bound for the problem of putting as many ones as possible in a matrix without violating the conditions in corollary 1, for $n = 5$ and $n = 8$. Ad hoc arguments show that the second matrix, with 21 ones, has the largest possible number of ones for $n = 6$, and the third matrix, with 31 ones, has the largest possible number of ones for $n = 7$.

$$\begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

Perhaps tighter results could be obtained by considering the $k \times (\lfloor n/k \rfloor + 1)$ submatrices for all $k \geq 2$, simultaneously, but this seems to be a hard problem. However, even if we could exactly determine the maximum number of ones that can be in an $n \times n$ matrix that does not contain any $k \times (\lfloor n/k \rfloor + 1)$ submatrix consisting only of ones, for all $k \geq 2$, we might still not have tight upper bounds for our original problem. For example, the best known lower bound for $M(5)$ is 12, though it is possible to put 16 ones in a 5×5 matrix satisfying the conditions of corollary 1. It seems that other techniques might be necessary to prove exact bounds. In the next section, we demonstrate some other techniques which could be useful.

3 A tight bound for 2×2 matrices

In this section, we prove that $M(2) \leq 2$ using different techniques. The example in the introduction shows that this upper bound is tight.

Notice that a function that can be expressed as an affine combination of the variables $x_1, x_2, ..., x_k$ over GF(2) is either the constant 0 or 1 or a parity function of a subset of those variables. To prove the upper bound, we first need to develop some properties of the product of two such functions.

Let $f, g : \{0,1\}^k \to \{0,1\}$ be constant or parity functions. Then $f(x_1, \ldots, x_k) = f_0 + \sum_{i=1}^{k} f_i x_i$ and $g(x_1, \ldots, x_k) = g_0 + \sum_{i=1}^{k} g_i x_i$ for some $f_0, \ldots, f_k, g_0, \ldots, g_k \in \{0,1\}$. Let $\vec{0} \in \{0,1\}^k$ denote the all zero vector and, for any subset $S \subseteq \{1, \ldots, k\}$, let $\vec{0}^{(S)} \in \{0,1\}^k$ denote the vector such that

$$\vec{0}_i^{(S)} = \begin{cases} 1 & \text{if } i \in S \\ 0 & \text{if } i \notin S. \end{cases}$$

An assignment of a value in $\{0,1\}^k$ to $x_1, x_2, ..., x_k$ will be called an input.

Lemma 7 *If $f \cdot g = 1$, then $f = g = 1$.*

Proof Suppose $f \cdot g = 1$. Since $1 = (f \cdot g)(\vec{0}) = f(\vec{0}) \cdot g(\vec{0}) = f_0 g_0$, it follows that $f_0 = g_0 = 1$.

Now consider $i \in \{1, \ldots, k\}$. Since $1 = (f \cdot g)(\vec{0}^{(\{i\})}) = (f_i + f_0)(g_i + g_0) = (f_i + 1)(g_i + 1)$, it follows that $f_i = g_i = 0$. Thus $f = g = 1$. □

Lemma 8 *If f, g, and h are parity functions and $f \cdot g = h$, then $f = g = h$.*

Proof Since f, g and h are parity functions, they are satisfied by (i.e. have value 1 for) exactly half the inputs. But the inputs that satisfy h are the inputs that satisfy both f and g. Thus, the inputs that satisfy h are contained in the set of inputs that satisfy f, and in the set of inputs that satisfy g. Therefore, $f = h$ and $g = h$. □

Lemma 9 *If f and f' are parity functions and g and g' are either constant or parity functions such that $f \cdot g + f' \cdot g' = 1$, then*

$$\begin{aligned} &f = f' + 1, \\ &g = f \text{ or } g = 1, \text{ and} \\ &g' = f' \text{ or } g' = 1. \end{aligned}$$

Proof Since f and f' are parity functions, they are satisfied by exactly half the inputs. The inputs that satisfy $f \cdot g$ are a subset of those that satisfy f; thus $f \cdot g$ is satisfied by at most half the inputs. This is also true for $f' \cdot g'$. But $f \cdot g + f' \cdot g' = 1$, so every input satisfies either $f \cdot g$ or $f' \cdot g'$. Therefore, $f \cdot g$ and $f' \cdot g'$ are each satisfied by exactly half the inputs.

For $f \cdot g$ to be satisfied for exactly half the inputs, it must be the case that g is satisfied by all inputs that satisfy f. This implies that $f \cdot g = f$. If $g \neq 1$, then by lemma 8, $f = g$. Similarly, $f' \cdot g' = f'$ and either $g' = f'$ or $g' = 1$. Hence, $1 = f \cdot g + f' \cdot g' = f + f'$, so $f = f' + 1$. □

Theorem 1 $M(2) \leq 2$.

Proof Let

$$A = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix},$$

where f_{11}, f_{12}, f_{21}, f_{22}, g_{11}, g_{12}, g_{21}, and g_{22} are constant or parity functions of the variables $a_1, a_2, a_3, b_1, b_2, b_3$. Suppose, to obtain a contradiction, that a_1b_1, a_2b_2, and a_3b_3 are three of the four entries in the product matrix $C = A \cdot B$. Without loss of generality, we may assume that

$$\begin{aligned} f_{11} \cdot g_{11} + f_{12} \cdot g_{21} &= a_1b_1, \\ f_{11} \cdot g_{12} + f_{12} \cdot g_{22} &= a_2b_2, \text{ and} \\ f_{21} \cdot g_{12} + f_{22} \cdot g_{22} &= a_3b_3. \end{aligned}$$

Consider the functions f'_{11}, f'_{12}, f'_{21}, f'_{22}, g'_{11}, g'_{12}, g'_{21}, and g'_{22} that result from setting $a_1 = b_1 = a_3 = b_3 = 1$. These functions are also constant or parity functions. Now

$$\begin{aligned} f'_{11} \cdot g'_{11} + f'_{12} \cdot g'_{21} &= 1 \\ f'_{11} \cdot g'_{12} + f'_{12} \cdot g'_{22} &= a_2b_2, \text{ and} \\ f'_{21} \cdot g'_{12} + f'_{22} \cdot g'_{22} &= 1. \end{aligned}$$

If $f'_{11} = 0$, then $f'_{12} \cdot g'_{21} = 1$; so by lemma 7, $f'_{12} = g'_{21} = 1$. This implies $a_2b_2 = g'_{22}$, which is impossible, since a_2b_2 is neither a constant nor a parity function. Thus $f'_{11} \neq 0$. Similarly, $f'_{12}, g'_{12}, g'_{22} \neq 0$.

If $f'_{11}, f'_{12} \neq 1$, then, by lemma 9, $f'_{12} = f'_{11} + 1$. Similarly, if $g'_{12}, g'_{22} \neq 1$, then $g'_{22} = g'_{12} + 1$. If both these equations are true, then

$$\begin{aligned} a_2b_2 &= f'_{11} \cdot g'_{12} + f'_{12} \cdot g'_{22} \\ &= f'_{11} \cdot g'_{12} + (f'_{11} + 1) \cdot (g'_{12} + 1) \\ &= 1 + f'_{11} + g'_{12}. \end{aligned}$$

This is impossible, since $1 + f'_{11} + g'_{12}$ is a constant or parity function. Therefore, at least one of f'_{11}, f'_{12}, g'_{12}, and g'_{22} is 1. Without loss of generality, say $f'_{11} = 1$. Then $a_2b_2 = g'_{12} + f'_{12} \cdot g'_{22}$.

Now either $g'_{12} = 1$ or $g'_{22} = 1$. Otherwise, by lemma 9, $a_2b_2 = g'_{12} + f'_{12} \cdot (g'_{12} + 1)$ or, equivalently, $a_2b_2 + (f'_{12} + 1) \cdot (g'_{12} + 1) = 1$. But this would contradict lemma 9, since $a_2, b_2, g'_{12} + 1 \neq 0, 1$ and $b_2 \neq a_2$.

Furthermore, $g'_{22} \neq 1$ or else $a_2 b_2 = g'_{12} + f'_{12}$. This is impossible because $g'_{12} + f'_{12}$ is a constant or parity function. Therefore, $g'_{12} = 1$.

Then $a_2 b_2 = 1 + f'_{12} \cdot g'_{22}$ or, equivalently, $a_2 b_2 + f'_{12} \cdot g'_{22} = 1$. Since $a_2 \neq 0, 1$ and $b_2 \neq 1, a_2$, it follows from lemma 9 and the commutativity of f'_{12} and g'_{22}, that neither f'_{12} nor g'_{22} can be parity functions; thus, they are constant. But this implies that $a_2 b_2$ is also constant, which it is not. Hence $M(2) \leq 2$.

□

4 Conclusion

The following theorem summarizes the results of sections 2 and 3.

Theorem 2 *Let* $u = \frac{1}{2}\left(\lceil n/3 \rceil - 1\right)(n-1)(n-2)$ *and* $v = \sqrt{u^2 - 1/27}$. *Then for* $n \geq 4$,

$$M(n) \leq n\left(1 + \sqrt[3]{u+v} + \sqrt[3]{u-v}\right) = n^2/\sqrt[3]{3} + O(n).$$

In addition, $M(2) = 2$, $M(3) = 6$, *and* $M(4) = 9$.

The above theorem states the best asymptotic results, but for some small values of n, lemmas 4 and 6 give better results, as the following table shows. The theorem gives the best results for larger values of n than those shown in the table.

n	1	2	3	4	5	6	7	8	9	10
Lemma 4	–	2	6	9	16	22	33	40	55	65
Lemma 5	–	–	–	12	17	23	35	43	53	70
Lemma 6	–	4	7	11	18	22	32	43	57	64

n	11	12	13	14	15	16	17	18	19	20
Lemma 4	83	95	117	130	156	172	201	219	251	271
Lemma 5	82	95	118	134	150	179	198	217	252	274
Lemma 6	81	99	120	130	154	179	207	220	251	283

n	21	22	23	24	25	26	27	28	29	30
Lemma 4	307	330	369	393	436	463	510	538	588	619
Lemma 5	297	337	363	390	435	465	495	546	579	612
Lemma 6	318	334	372	411	453	472	517	563	612	634

n	31	32	33	34	35	36	37	38	39	40
Lemma 4	673	706	763	798	859	896	960	999	1067	1109
Lemma 5	669	705	742	804	844	884	952	995	1039	1112
Lemma 6	686	739	795	820	879	939	1002	1030	1096	1163

n	41	42	43	44	45	46	47	48	49	50
Lemma 4	1180	1223	1298	1344	1422	1470	1552	1602	1687	1739
Lemma 5	1159	1207	1285	1335	1386	1471	1524	1579	1668	1726
Lemma 6	1233	1264	1337	1411	1488	1522	1602	1683	1767	1804

Acknowledgements

We are grateful to Gudmund Frandsen for helpful discussions. We would also like to thank Richard Anstee, Jørgen Brandt, Cary Huffman, and Richard Nowakowski.

References

[1] J. Boyar and G. Brassard and R. Peralta, Subquadratic zero-knowledge, in: Proceedings of the 32nd IEEE Symposium on Foundations of Computer Science (IEEE Computer Society Press, Los Amitos, 1991) 69-78.

[2] R. Salem and D. C. Spencer, On sets of integers which contain no three terms in arithmetical progression, in: Proceedings of the National Academy of Sciences of the United States of America 28 (1942) 561-563.

[3] E. Szemerédi, personal communication through G. Brassard, S. Kannan, S. Rudich and G. Tardos, 1991.

[4] A.V. Aho and J.E. Hopcroft and J.D. Ullman, The Design and Analysis of Computer Algorithms (Addison-Wesley Publishing Company, Reading, 1974).

[5] S. Winograd, On the number of multiplications necessary to compute certain functions, Comm. on Pure and Applied Mathematics 23 (1970) 165-179.

[6] R.K. Guy and S. Znám, A problem of Zarankiewicz, in: W.T. Tutte, ed., Recent Progress in Combinatorics: Proceedings of the Third Waterloo Conference on Combinatorics, May 1968 (Academic Press, New York, 1969) 237-243.

[7] S. Roman, A problem of Zarankiewicz, Journal of Combinatorial Theory (A) 18 (1975) 187-198.

[8] T. Kövari and V.T. Sós and P. Turán, On a problem of K. Zarankiewicz, Colloq. Math. 3 (1954) 50-57.

[9] C. Hyltén-Cavallius, On a combinatorical problem, Colloq. Math. 6 (1958) 59-65.

[10] K. Zarankiewicz, Problem P 101, Colloq. Math. 2 (1951) 301.