Thin circulant matrices and lower bounds on the complexity of some Boolean operators*

M. I. Grinchuk, I. S. Sergeev

Abstract

We prove a lower bound $\Omega\left(\frac{k+l}{k^2l^2}N^{2-\frac{k+l+2}{kl}}\right)$ on the maximal possible weight of a (k,l)-free (that is, free of all-ones $k\times l$ submatrices) Boolean circulant $N\times N$ matrix. The bound is close to the known bound for the class of all (k,l)-free matrices. As a consequence, we obtain new bounds for several complexity measures of Boolean sums' systems and a lower bound $\Omega(N^2\log^{-6}N)$ on the monotone complexity of the Boolean convolution of order N.

Keywords: complexity, circulant matrix, thin matrix, Zarankiewicz problem, monotone circuit, rectifier circuit, Boolean sum, Boolean convolution.

1 Introduction

Hereafter, a Boolean matrix is called (k,l)-*free* (or thin) if it does not contain an all-ones $k\times l$ submatrix. In the case $k=l$ we write simply k-*free*. Further, assume $2\le k\le l$.

An $N\times N$ matrix $(c_{i,j})$ is *circulant* (or cyclic), if either $c_{i,j}=c_{0,(i+j)\bmod N}$ for all i,j, or $c_{i,j}=c_{0,(i-j)\bmod N}$ for all i,j.

In [2] the first author proved the existence of k-free Boolean circulant $N\times N$ matrices of weight[1] $\Omega\left(k^{-4}N^{2-\sqrt{3/k}}\right)$ and obtained corollaries for the complexity[2] of Boolean sums' systems[3] with circulant matrices, with respect to implementation

*Original text published in Russian in *Diskretnyi Analiz i Issledovanie Operatsii* (*Discrete analysis and operations research*). 2011. 18(5), 38–53.

[1] *Weight* of a (Boolean) matrix is the number of non-zero entries in it.

[2] The reader can find the notions of *complexity*, *depth*, *rectifier circuit*, *circuit of functional elements* e.g. in [4, 5].

[3] *Boolean sum* is a function of the form $x_1\vee\ldots\vee x_n$. A system of Boolean sums with an $N\times N$ matrix $(c_{i,j})$ is a mapping with components $\bigvee_{j=1}^{N}c_{i,j}x_j$, $1\le i\le N$.

via rectifier circuits of depth 2 or unbounded depth. Precisely, the bound for the first measure is $\Omega(N^2 \log^{-10} N)$, and for the second it is $\Omega(N^2 \log^{-12} N)$.

In fact, the method has a potential for improvement of the above bounds, which is of interest due to connection to the Zarankiewicz problem (the problem is discussed in details e.g. in [6]). This potential is in application of a more accurate bound on the cardinality of the sum of two sets in a Euclidean space following from [8, 10].

Below, we show the existence of (k,l)-free circulant $N \times N$ matrices of weight $\Omega\left(\frac{k+l}{k^2 l^2} N^{2-\frac{k+l+2}{kl}}\right)$. For comparison, the classic Erdös—Spencer result [6] states just a slightly better bound $\Omega_{k,l}\left(N^{2-\frac{k+l-2}{kl-1}}\right)$ in the class of all (k,l)-free matrices.

Hence, for a system of Boolean sums with an appropriate circulant matrix the following complexity bounds hold:

— $\Omega(N^2 \log^{-6} N)$ with respect to implementation via circuits of functional elements[4] over the basis $\{\vee, \wedge\}$;

— $\Omega(N^2 \log^{-5} N)$ with respect to implementation via circuits over the basis $\{\vee\}$, or via rectifier circuits;

— $\Omega(N^2 \log^{-4} N)$ with respect to implementation via depth-2 rectifier circuits.

The paper [1] considers the ratio $\lambda(N) = \max_A \frac{L_\vee(A)}{L_\oplus(A)}$, where $L_\vee(A)$ is the circuit complexity of the Boolean sums' system with matrix A over the basis $\{\vee\}$, $L_\oplus(A)$ is the circuit complexity of the linear operator with matrix A over the basis $\{\oplus\}$, and the maximum is taken over all Boolean $N \times N$ matrices. The result of the present paper leads to a bound $\lambda(N) = \Omega\left(\frac{N}{(\log N)^6 \log\log N}\right)$, which in a sense close to an upper bound $\lambda(N) = O\left(\frac{N}{\log N}\right)$.

As another corollary, we obtain that the circuit complexity of the Boolean convolution of order N over the basis $\{\vee, \wedge\}$ is $\Omega(N^2 \log^{-6} N)$. Specifically, this bound holds for the number of disjunctors (that is, $\vee$-gates) in any monotone circuit computing the convolution. Some recent papers (e.g. [3, 7]) mention the bound $\Omega(N^{3/2})$ as a record, though a stronger bound follows from [2] directly[5]. The obtained lower bound is close to the trivial upper bound $O(N^2)$.

[4]Further, we simply call them circuits.

[5]The bound $\Omega(N^{3/2})$ corresponds to the number of disjunctors in a monotone circuit (the survey [3] is inaccurate at this point). However, the recent paper [7] declares the same bound for the number of conjunctors ($\wedge$-gates; proof is omitted there).

2 Some properties of "rectangles"

Now, we present the main result following the proof strategy from [2]. Let $k, l \in \mathbb{N}$, $2 \le k \le l$. Denote $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$. We define *rectangle* as an element of the set

$$R_{k,l} = \{(x_1, \ldots, x_k, y_1, \ldots, y_l) \in \mathbb{Z}_+^{k+l} \mid \forall_{i \ne j}(x_i \ne x_j), \forall_{i \ne j}(y_i \ne y_j)\}.$$

Let $E = (a_1, \ldots, a_k, b_1, \ldots, b_l)$ be a rectangle. Let $m(E) = |\{a_i + b_j \mid 1 \le i \le k,\ 1 \le j \le l\}|$ denote the number of points in the rectangle E.

Consider the system $S(E)$ of linear equations

$$\{x_r + y_s = x_u + y_v \mid a_r + b_s = a_u + b_v,\ 1 \le r, u \le k,\ 1 \le s, v \le l\}$$

over the field $\mathbb{R}$. The set of solutions constitutes a linear subspace T_E in $\mathbb{R}^{k+l}$. Let $n(E)$ be its dimension. Let $C(E)$ denote the set of rectangles $\{(x_1, \ldots, x_k, y_1, \ldots, y_l)\}$ satisfying $S(E)$ and failing to satisfy any other equation $x_r + y_s = x_u + y_v$ (in [2], $C(E)$ is called equivalence class).

We have to estimate the number of rectangles with bounded (by a number N) coordinates and fixed number of points. An implicit relation between the number of rectangles and the number of points will be further established with the help of intermediate parameter $n(E)$. First, we will count the number of rectangles E with a given value of $n(E)$. Next, we will derive relations between $n(E)$ and $m(E)$.

To roughly estimate the number of rectangles with bounded coordinates $0 \le x_1, \ldots, y_l < N$ in $C(E)$ we use the following lemma.

Lemma 1. *Let $N \in \mathbb{N}$. Then $|C(E) \cap \{0, \ldots, N-1\}^{k+l}| \le N^{n(E)}$.*

Proof. The coordinates $x_1, \ldots, x_k, y_1, \ldots, y_l$ of a vector from T_E are defined by values of $n(E)$ free variables. There are at most $N^{n(E)}$ ways to arrange such values, given that the vector is from $C(E) \cap \{0, \ldots, N-1\}^{k+l}$. □

The second lemma estimates the number of classes with a given value of $n(E)$. (We use notation C_n^k for binomial coefficients.)

Lemma 2. *Let $n \in \mathbb{N}$. Then $|\{C(E) \mid n(E) = n\}| \le C_{k^2l^2}^{k+l-n}$.*

Proof. The class $C(E)$ is uniquely defined by the system $S(E)$, which in its turn is uniquely defined by a linearly independent subsystem of $k + l - n$ equations. The number of such subsystems is bounded from above by the number of ways to choose $k + l - n$ equations from k^2l^2 ones. □

Now, we manage to obtain relations between $n(E)$ and $m(E)$. This piece of proof differs from [2].

Let ξ_i denote the unit vector in the space $\mathbb{R}^{k+l}$ with i-th coordinate being 1 and other coordinates being 0.

Let $n = n(E)$. For unification, let us introduce notation $x_{i+k} = y_i$, $1 \le i \le l$. Set

$$\bar{x} = (\underbrace{1, \dots, 1}_{k}, \underbrace{0, \dots, 0}_{l}), \qquad \bar{y} = (\underbrace{0, \dots, 0}_{k}, \underbrace{1, \dots, 1}_{l}).$$

Notice that $\bar{x}, \bar{y} \in T_E$ (regardless of E). Let T'_E be the space of solutions of the system

$$S'(E) = S(E) \cup \{x_1 = y_1 = 0\}.$$

Then $\dim T'_E = n - 2$ and $T_E = T'_E + \{\alpha\bar{x} + \beta\bar{y} \mid \alpha, \beta \in \mathbb{R}\}$ ($A + B$ hereafter denotes the element-wise sum (Minkowski sum) of sets A and B). Write

$$T'_E = \left\{ (x_1, \dots, x_{k+l}) \;\middle|\; x_i = \sum_{j=1}^{n-2} \alpha_{i,j} x_{i_j},\ 1 \le i \le k + l \right\},$$

where $x_{i_1}, \dots, x_{i_{n-2}}$ is a set of free variables of the system $S'(E)$, and $\alpha_{i,j}$ are real constants. Then setting $i_{n-1} = 1$ and $i_n = k + 1$ we conclude that $x_{i_1}, \dots, x_{i_n}$ is a set of free variables of the system $S(E)$, and

$$T_E = \left\{ (x_1, \dots, x_{k+l}) \;\middle|\; x_i = \sum_{j=1}^{n} \alpha_{i,j} x_{i_j},\ 1 \le i \le k + l \right\},$$

where $\alpha_{i,n-1} = \alpha_{k+j,n} = 1$, and $\alpha_{i,n} = \alpha_{k+j,n-1} = 0$ for $1 \le i \le k$, $1 \le j \le l$.

Consider a linear mapping ψ_E from $\mathbb{R}^{k+l}$ to the space $\mathbb{R}^{n-2}$ with Euclidean metrics and orthonormal basis $\{e_1, \dots, e_{n-2}\}$ defined by $\psi_E : \xi_i \to \sum_{j=1}^{n-2} \alpha_{i,j} e_j$ for any i. In particular, $\psi_E(\xi_1) = \psi_E(\xi_{k+1}) = 0$ (0 hereafter stands for the zero vector of a space if it does not lead to a misunderstanding).

Set $A_E = \psi_E(\{\xi_1, \dots, \xi_k\})$, $B_E = \psi_E(\{\xi_{k+1}, \dots, \xi_{k+l}\})$.

Recall that the dimension $\dim A$ of a set A in a Euclidean space is the minimum of dimensions of affine subspaces containing A.

Lemma 3. *$|A_E + B_E| = m(E)$, $\dim(A_E + B_E) = n - 2$.*

Proof. The first equality holds due to the following chain of equivalent transformations:

$$a_r + b_s = a_u + b_v \iff$$

$$(\ (x_1, \dots, x_k, y_1, \dots, y_l) \in T_E \Longrightarrow x_r + y_s = x_u + y_v\ ) \iff$$

$$\left(\ (x_1, \dots, x_k, y_1, \dots, y_l) \in T_E \Longrightarrow \sum_{j=1}^{n} (\alpha_{r,j} + \alpha_{k+s,j}) x_{i_j} = \sum_{j=1}^{n} (\alpha_{u,j} + \alpha_{k+v,j}) x_{i_j}\ \right)$$

$$\iff \forall_{j,\ 1 \le j \le n} (\alpha_{r,j} + \alpha_{k+s,j} = \alpha_{u,j} + \alpha_{k+v,j}) \iff$$

$$\forall_{j,\, 1\le j\le n-2}(\alpha_{r,j}+\alpha_{k+s,j}=\alpha_{u,j}+\alpha_{k+v,j}) \iff$$

$$\psi_E(\xi_r+\xi_{k+s})=\sum_{j=1}^{n-2}(\alpha_{r,j}+\alpha_{k+s,j})e_j=\sum_{j=1}^{n-2}(\alpha_{u,j}+\alpha_{k+v,j})e_j=\psi_E(\xi_u+\xi_{k+v}).$$

The second equality is straightforward, since $0\in A_E\cap B_E$ and $\{e_1,\dots,e_{n-2}\}\subset A_E\cup B_E$. □

In the next section, we will estimate $m(E)$.

3 The cardinality of the sum of two sets in a Euclidean space

The following result is due to I. Ruzsa [10].

Theorem 1 (Ruzsa [10])**.** *Let A and B be finite sets in the Euclidean space $\mathbb{R}^n$ satisfying $|A|\le|B|$ and $\dim(A+B)=n$. Then*

$$|A+B|\ge n|A|+|B|-\frac{n(n+1)}{2}.$$

Ruzsa also provided a more accurate bound

$$|A+B|\ge|B|+\sum_{i=1}^{|A|-1}\min\{n,\,|B|-i\}.$$

W.l.o.g. we can assume $0\in A\cap B$ throughout this section.

Already, these bounds are sufficient to principally achieve results announced in the introduction. However, the bounds are not asymptotically tight for large n. On the contrary, the bound $|A+B|\ge\lfloor n^2/4\rfloor$ established in [2] is rough for small n (though its advantage is the simplicity of the proof). The method [8] allows to exhibit tight bounds.

Let $\{e_1,\dots,e_n\}$ be an orthonormal basis of a Euclidean space $\mathbb{E}^n$. Following [8], we define *long simplex* as a set F of the form

$$\{me_1|m=0,\dots,|F|-k\}\cup\{e_{i_1},\dots,e_{i_{k-1}}\},\tag{1}$$

with numbers $1,i_1,\dots,i_{k-1}$ being pairwise different, $k\ge 0$.

The next lemma is a reformulation of the Corollary 3.8 [8].

Lemma 4. *Under conditions of Theorem* 1 *the minimum of $|A+B|$ is either $|A||B|$ (in this case $\dim A+\dim B=n$), or it is witnessed by a pair of long simplices.*

The proof can be found in [8]. It is crucial to observe that the sets A and B delivering the minimum in the lemma satisfy the definition (1) with the same basis and, in particular, with the same vector e_1.

Tight bounds (for any values of parameters) were not determined in [8]. Though, they can be easily derived from the lemma above.

Theorem 2. *Let $A, B \subset \mathbb{R}^n$, $K = |A| \le |B| = L$, and $\dim(A+B) = n$. We have:*

(i) if $n = K + L - 2$, then $|A+B| = KL$;

(ii) if $n \le L - K$, then $|A+B| \ge L + n(K-1)$;

(iii) if $L - K \le n \le L$, then

$$|A+B| \ge (n+1)K - \frac{(n-L+K)(n-L+K+1)}{2};$$

(iv) if $L \le n \le K + L - 3$, then

$$|A+B| \ge KL - \frac{(K+L-n)(K+L-n-1)}{2}.$$

Proof. In the case $\dim A + \dim B = n$, the set $A + B$ has the maximal possible cardinality KL, thus, (i) follows. Therefore, in the case $n < K + L - 2$, we may assume that $\dim A + \dim B > n$.

So, by Lemma 4, it suffices to consider sets A, B being long simplices (1). Assume w.l.o.g.

$$A = C_A \cup D \cup D_A, \qquad B = C_B \cup D \cup D_B,$$

where

$$C_A = \{me_1 | m = 0, \dots, K - s - s_A - 1\},$$

$$C_B = \{me_1 | m = 0, \dots, L - s - s_B - 1\},$$

$$D = \{e_2, \dots, e_{s+1}\}, \quad D_A = \{e_{s+2}, \dots, e_{s+s_A+1}\}, \quad D_B = \{e_{s+s_A+2}, \dots, e_n\},$$

$s = |D|$, $s_A = |D_A|$, $s_B = |D_B|$, $s + s_A + s_B = n - 1$. Hence,

$$|A+B| = |C_A + C_B| + |(C_A \cup C_B) + D| + |C_A + D_B| + \\ |C_B + D_A| + |D + D| + |D + (D_A \cup D_B)| + |D_A + D_B|.$$

It can be verified directly that

$$|C_A + C_B| = K + L - s - n, \quad |(C_A \cup C_B) + D| = s(\max\{L - s_B, K - s_A\} - s),$$

$$|C_A + D_B| = (K - s - s_A)s_B, \quad |C_B + D_A| = (L - s - s_B)s_A,$$

$$|D+D| = \frac{s(s+1)}{2}, \quad |D+(D_A \cup D_B)| = s(s_A+s_B), \quad |D_A+D_B| = s_A s_B.$$

Summing all, we obtain

$$|A+B| = (s_A+1)K+(s_B+1)L+s\cdot\max\{L-s_B,\, K-s_A\}-n-\frac{s(s+1)}{2}-s_A s_B. \quad (2)$$

Thus, the problem reduced to finding the minimum of the expression (2). Let s^*, s_A^*, s_B^* denote the values of parameters s, s_A, s_B delivering this minimum. Let us list restrictions on the parameters:

$$s+s_A+s_B = n-1, \qquad s+s_A \le K-1, \qquad s+s_B \le L-1. \qquad (*)$$

Consider (ii). Suppose $n \le L-K$. Then

$$L-s_B \ge K+(n-s_B) \ge K \ge K-s_A.$$

Thus, minimization of (2) (with eliminated constant terms) is equivalent to maximization of the expression

$$s_B(n+L-K-1-s_B)+\frac{s(s+1)}{2}. \quad (3)$$

For a fixed s the value of (3) grows when s_A decreases (and s_B increases accordingly), since $2s_B < n+L-K-1$ and due to the fact that the function $x(a-x)$ monotonically grows in the interval $[0,\, a/2]$. Yet, the conditions $(*)$ are not violated. Hence, $s_A^* = 0$.

Set $s_B = n-1-s$. Then, after elimination of constant terms the expression (3) reduces to

$$-\frac{s(s+1)}{2} - ((L-K)-n)s.$$

Consequently, $s^* = 0$. By the assignment $s_B = n-1$ and $s = s_A = 0$ in (2), we derive the inequality (ii).

Let us prove (iii). Assume $L-K \le n \le L$. Consider two cases.

Case A. Suppose $L-s_B \ge K-s_A$. As above, the problem reduces to maximization of (3). Note that for a fixed s_B the value of (3) grows with decreasing of s_A (and corresponding increasing of s), and the conditions $(*)$ are not violated. Therefore, either $s_B^* \le L-K$ and $s_A^* = 0$, or $s_B^* = L-K+s_A^*$.

In the former subcase, assign $s = n-1-s_B$. Then, after elimination of constant terms the expression (3) reduces to

$$s_B(2(L-K)-1-s_B),$$

hence, $s_B^* \in \{L-K-1,\, L-K\}$.

In the latter subcase, assign $s_B = L - K + s_A$ and $s = n - 1 - L + K - 2s_A$. Then, the expression (3) has the form

$$s_A(s_A + L - K - n).$$

The second factor is $s_B - n$, and so it is negative. Consequently, $s_A^* = 0$, and $s_B^* = L - K$ follows as well, as in the previous subcase.

Case B. Suppose $L - s_B \leq K - s_A$. Then,

$$s + s_A = n - 1 - s_B \leq L - 1 - s_B \leq K - 1 - s_A \leq K - 1.$$

So, only the first of conditions $(*)$ is essential. Here, minimization of (2) is equivalent to maximization of the expression

$$s_A(n - L + K - 1 - s_A) + \frac{s(s+1)}{2}. \tag{4}$$

For a fixed s_A the value of (4) grows, when s increases and s_B accordingly decreases, thus, $s_B^* = L - K + s_A^*$. That is the very situation already discussed in the second subcase of the case A.

Via assignment $s_A = 0$, $s_B = L - K$, $s = n - 1 - L + K$ in (2), we obtain the inequality (iii) (the assignment is in a sense correct also in the case $L - K = n$).

Now, turn to (iv). Assume $L \leq n \leq K + L - 3$. Again, consider two cases.

Case A. Suppose $L - s_B \geq K - s_A$. In this case, the latter of conditions $(*)$ follows from the second:

$$s + s_B \leq s + s_A + L - K \leq L - 1.$$

Again, the problem is to maximize the expression (3). Observe that for a fixed s_B the value of (3) grows when s_A decreases (and s correspondingly increases), and conditions $(*)$ are not violated. Hence, $s_A^* = s_B^* - L + K$ (it is the minimal possible value of s_A for a fixed s_B).

Under the assignment $s = n + L - K - 1 - 2s_B$ and elimination of constant terms, the expression (3) reduces to

$$s_B(s_B - n - L + K).$$

Since $2s_B < (L - K + s_A) + (n - s_A) = n + L - K$, the second factor is negative and greater than the first factor by absolute value. Consequently, the maximum is achieved on the minimal possible value of s_B under the conditions $(*)$. Hence, we deduce that $s_B^* = n - K$.

Case B. Suppose $L - s_B \leq K - s_A$. In this case, the second condition in $(*)$ is inessential:

$$s + s_A \leq s + s_B - L + K \leq K - 1.$$

We have to maximize (4). Observe that it grows when s_A is fixed, s increases and s_B decreases, and conditions $(*)$ are fulfilled. Thus, $s_B^* = L - K + s_A^*$. So, we are under the conditions of the already investigated case A.

Under assignment $s_A = n - L$, $s_B = n - K$, $s = L + K - 1 - n$ in (2), we exhibit the inequality (iv). □

As follows from the proof, the bounds of the theorem are achievable.

Under the conditions of Theorem 2, define the function

$$\rho(K, L) = \max_{1 \le n \le K+L-2} \frac{n+2}{|A+B|}. \tag{5}$$

Lemma 5. $\rho(2, L) = \frac{L+2}{2L}$. *If* $K \ge 3$, *then*

$$\rho(K, L) = \max\left\{ \frac{K+L}{KL},\ \frac{K+L-1}{KL-3},\ \frac{2(L+2)}{K(2L-K+1)} \right\} < \frac{K+L+2}{KL}.$$

In particular, $\rho(K, K) = \frac{2(K+2)}{K(K+1)}$.

Proof. Define additionally $\rho(K, L, n) = \frac{n+2}{\min_{A,B} |A+B|}$. By the definition, $\rho(K, L) = \max_n \rho(K, L, n)$.

First, we need to verify that the function $\rho(K, L, n)$ achieves its maximum at the endpoints of intervals defined in pp. (ii)–(iv) of Theorem 2.

In the case $1 \le n \le L - K$, the function

$$\rho^{-1}(K, L, n) = \frac{L + n(K-1)}{n+2} = K - 1 + \frac{L - 2K + 2}{n+2}$$

is evidently monotone (hereafter, we consider $\rho(K, L, n)$ as a function of variable n).

In the case $L - K \le n \le L$, denote $n' = n - (L - K)$. Then

$$\rho^{-1}(K, L, n) = K - \frac{n'(n'+1) + 2K}{2(n' + L - K + 2)}.$$

The subtrahend function is convex downward for $n' \ge 0$, since it has the form $c\frac{n'(n'+1)+a}{n'+1+b}$ with $a, b, c \ge 0$. Therefore, with respect to the interval $[0,\ K]$ it takes its maximal value in the endpoints (it holds for $K \ge 3$; for $K = 2$ the argument of the maximum lies in the interval $[0,\ 1]$). Consequently, there takes its maximum the function $\rho(K, L, n)$.

In the case $L \le n \le K + L - 3$, denote $n' = n - L$. Then

$$\rho^{-1}(K, L, n) = K - \frac{2(n'+2)K + (K - n')(K - n' - 1)}{2(n' + L + 2)} =$$

$$= K - \frac{n'(n'+1) + K(K+3)}{2(n' + L + 2)}.$$

We treat this case the same way as the previous one.

Thus, for $K \geq 3$ we have

$$\arg \max_{1 \leq n \leq K+L-2} \rho(K, L, n) \in \{1,\ L-K,\ L,\ K+L-3,\ K+L-2\},$$

$$\arg \max_{1 \leq n \leq K+L-2} \rho(2, L, n) \in \{1,\ L-2,\ L-1,\ L\}.$$

Let us check that $\rho(K, L, 1) \leq \rho(K, L, K+L-2)$. Indeed,

$$\rho(K, L, 1) = \frac{3}{K+L-1} \leq \frac{4}{K+L} \leq \frac{1}{K} + \frac{1}{L} = \frac{K+L}{KL} = \rho(K, L, K+L-2),$$

due to the well-known inequality $\frac{a^2}{b} + \frac{c^2}{d} \geq \frac{(a+c)^2}{b+d}$, where $b, d > 0$.

Notice further that

$$\rho(K, L, L-K) = \frac{1}{K}\left(1 + \frac{1}{L-(K-1)}\right) \leq \frac{1}{K}\left(1 + \frac{K}{L}\right) = \rho(K, L, K+L-2).$$

Yet,

$$\rho(2, L, L-1) = \frac{L+1}{2L-1} \leq \frac{L+2}{2L} = \rho(2, L, L).$$

Therefore, it is proved that $\rho(2, L) = \rho(2, L, L) = \frac{L+2}{2L}$ and

$$\rho(K, L) = \max\{\rho(K, L, K+L-2),\ \rho(K, L, K+L-3),\ \rho(K, L, L)\} = \max\left\{\frac{K+L}{KL},\ \frac{K+L-1}{KL-3},\ \frac{2(L+2)}{K(2L-K+1)}\right\}.$$

Applying the simple estimation

$$\frac{2(L+2)}{K(2L-K+1)} = \frac{L + (K+3)\frac{L}{2L-K+1}}{KL} \leq \frac{L + (K+3)\frac{K}{K+1}}{KL} < \frac{K+L+2}{KL},$$

the inequality $\rho(K, L) < \frac{K+L+2}{KL}$ can be easily checked. The last statement of the lemma concerning $\rho(K, K)$ is easy to verify. □

4 Weight of thin circulant matrices

A circulant matrix is entirely defined by its one row, say, the first row. Let $c_j = c_{0,j}$, $0 \leq j \leq N-1$, denote the entries of the row, where N is the size of the matrix. For convenience, assume that the other entries satisfy $c_{i,j} = c_{(i+j) \bmod N}$ (that is, 1-uniform diagonals of the matrix are parallel to the secondary diagonal).

Then, the condition that a matrix $(c_{i,j})$ contains an all-ones submatrix constituted by rows with numbers $a_1, \ldots, a_k$ and by columns with numbers $b_1, \ldots, b_l$ can be written as

$$c_{(a_i+b_j) \bmod N} = 1, \qquad 1 \le i \le k,\ 1 \le j \le l.$$

Let $\gamma_0, \ldots, \gamma_{N-1}$ be independent random variables taking value 1 with probability p and value 0 with probability $1-p$. Denote $\gamma = \sum \gamma_i$.

Hereafter, we denote by $\mathbf{P}(Q)$ the probability of the event Q. Let $\mathbf{M}\xi$ and $\mathbf{D}\xi$ denote the expectation and the variance of a random variable ξ, respectively.

Lemma 6. $\mathbf{P}\left(\gamma \ge pN - 2\sqrt{pN}\right) \ge 3/4$.

Proof. The required inequality follows from the Chebyshev's inequality

$$\mathbf{P}\left(|\gamma - \mathbf{M}\gamma| > \varepsilon\right) < \frac{\mathbf{D}\gamma}{\varepsilon^2}$$

by setting $\mathbf{M}\gamma = pN$, $\mathbf{D}\gamma = p(1-p)N$ $\varepsilon = 2\sqrt{pN}$. □

Set formally $\gamma_i = 0$, when $i \ge N$. Let $Q(E, \gamma_0, \ldots, \gamma_{N-1})$ with $E = (a_1, \ldots, a_k, b_1, \ldots, b_l) \in R_{k,l} \cap \{0, \ldots, N-1\}^{k+l}$ denote the event

$$\forall_{i,j}(\gamma_{a_i+b_j} = 1).$$

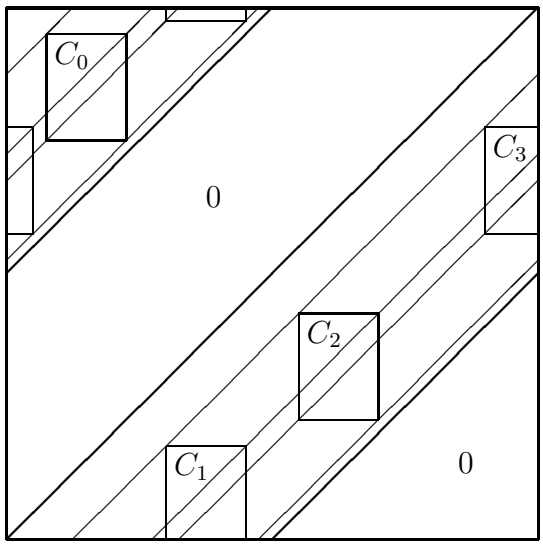
Substantially, it implies that a random circulant $2N \times 2N$ matrix Γ with the first row $(\gamma_0, \ldots, \gamma_{N-1}, 0, \ldots, 0)$ contains an all-ones $k \times l$ submatrix in the intersection of rows $a_1, \ldots, a_k$ and columns $b_1, \ldots, b_l$.

Observe that any all-ones submatrix of a matrix Γ can be translated to an all-ones submatrix entirely contained in the upper left $N \times N$ submatrix (that is, constituted by rows and columns numbered from 0 to $N-1$) of Γ by a cyclic shift (of numbers of rows and columns). Generation of all-ones submatrices by cyclic shifts is illustrated on the picture below; submatrices C_i are shown as rectangles, the submatrix C_0 is a desired one.

Therefore, the matrix Γ is (k,l)-free iff its left upper $N \times N$ submatrix is.

Theorem 3. *There exists a (k,l)-free circulant $N \times N$ matrix of weight $\Omega\left(\frac{k+l}{k^2l^2} N^{2-\rho(k,l)}\right)$.*

Proof. It follows directly from the definition that the probability of the event



$Q(E, \gamma_0, \ldots, \gamma_{N-1})$ is at most $p^{m(E)}$. Then

$$\mathbf{P}\left(\exists_{E\in R_{k,l}}(Q(E,\gamma_0,\ldots,\gamma_{N-1}))\right) \le \sum_{E\in R_{k,l}\cap\{0,\ldots,N-1\}^{k+l}} \mathbf{P}\left(Q(E,\gamma_0,\ldots,\gamma_{N-1})\right) =$$

$$\sum_{n=3}^{k+l} \sum_{\substack{E\in R_{k,l}\cap\{0,\ldots,N-1\}^{k+l},\\ n(E)=n}} \mathbf{P}\left(Q(E,\gamma_0,\ldots,\gamma_{N-1})\right) \le$$

$$\sum_{n=3}^{k+l} \sum_{\substack{E\in R_{k,l}\cap\{0,\ldots,N-1\}^{k+l},\\ n(E)=n}} p^{m(E)} \le \sum_{n=3}^{k+l} \sum_{\substack{C(E)\subset R_{k,l},\\ n(E)=n}} N^n p^{m(E)} \le$$

$$\sum_{n=3}^{k+l} \sum_{\substack{C(E)\subset R_{k,l},\\ n(E)=n}} \left(pN^{\rho(k,l)}\right)^{m(E)}.$$

Here, the second from the last inequality follows from Lemma 1, and the last one is justified by Lemma 3 and the definition (5).

Set $p = \left(\frac{k+l}{ek^2l^2}\right) N^{-\rho(k,l)}$, and continue exploiting the inequality of Lemma 2:

$$\sum_{n=3}^{k+l} \sum_{\substack{C(E)\subset R_{k,l},\\ n(E)=n}} \left(pN^{\rho(k,l)}\right)^{m(E)} \le \sum_{n=3}^{k+l} \sum_{\substack{C(E)\subset R_{k,l},\\ n(E)=n}} \left(\frac{k+l}{ek^2l^2}\right)^{m(E)} \le$$

$$\sum_{n=3}^{k+l} C_{k^2l^2}^{k+l-n} \left(\frac{k+l}{ek^2l^2}\right)^{k+l-1} \le \sum_{n=3}^{k+l} \left(\frac{ek^2l^2}{k+l-n}\right)^{k+l-n} \left(\frac{k+l}{ek^2l^2}\right)^{k+l-1} =$$

$$\sum_{n=3}^{k+l} \left(\frac{k+l}{ek^2l^2}\right)^{n-1} \left(1+\frac{n}{k+l-n}\right)^{k+l-n} \le \sum_{n=3}^{k+l} \left(\frac{k+l}{ek^2l^2}\right)^{n-1} e^n =$$

$$e\sum_{n=3}^{k+l} \left(\frac{k+l}{k^2l^2}\right)^{n-1} \le \frac{e(k+l)}{k^2l^2} \le e/4.$$

Here, we use well-known inequalities $C_n^m \le \left(\frac{en}{m}\right)^m$ and $(1+1/x)^x < e$ for $x > 0$, and assume $x^x \mid_{x=0} = 1$ (this quantity appears in the form $(k+l-n)^{k+l-n} \mid_{n=k+l}$).

Hence, as follows form the note before the theorem, a random circulant $(2N \times 2N)$ matrix Γ is (k,l)-free with probability at least $(4-e)/4$. In the sight of Lemma 6, we can conclude that this random matrix is (k,l)-free and also has weight $2N\gamma \ge 2N(pN - 2\sqrt{pN}) = \Omega(pN^2)$ with positive probability. □

5 Corollaries

Theorem 3 and Lemma 5 lead to

Corollary 1. *There exists a (k,l)-free $N \times N$ circulant matrix of weight $\Omega\left(\frac{k+l}{k^2l^2} N^{2-\frac{k+l+2}{kl}}\right)$.*

In the case $k = l = \Theta(\log N)$, the weight of a circulant matrix provided by the corollary is $\Omega\left(N^2 \log^{-3} N\right)$. This fact together with complexity bounds for Boolean sums' systems with (k,l)-free matrices [9] (see also [2, 11]) yields

Corollary 2. *There exists a circulant $N \times N$ matrix such that for the complexity of the corresponding system of Boolean sums the following bounds hold: $\Omega\left(N^2 \log^{-4} N\right)$ with respect to implementation via depth-2 rectifier circuits, $\Omega\left(N^2 \log^{-5} N\right)$ — for circuits over the basis $\{\vee\}$ or unbounded-depth rectifier circuits, $\Omega\left(N^2 \log^{-6} N\right)$ — for the number of disjunctors in a circuit over the basis $\{\vee, \wedge\}$.*

For the same choice of the parameters, the function $\lambda(N)$ defined in the introduction can be bounded as follows (taking [1] into account).

Corollary 3. $\lambda(N) = \Omega\left(\frac{N}{(\log N)^6 \log\log N}\right)$.

Boolean convolution of order N is the function

$$U_N(x_0, \ldots, x_{N-1}, y_0, \ldots, y_{N-1}) = (u_0, \ldots, u_{2N-2}), \quad u_k = \bigvee_{i+j=k} x_i y_j.$$

Cyclic Boolean convolution of order N is defined as

$$Z_N(x_0, \ldots, x_{N-1}, y_0, \ldots, y_{N-1}) = (z_0, \ldots, z_{N-1}), \quad z_k = \bigvee_{i+j\equiv k \bmod N} x_i y_j.$$

Let $V(f)$ be the minimal number of disjunctors in a circuit over the basis $\{\vee, \wedge\}$ that implements a function f. Then, the following relations are straight from the definition of convolutions:

$$V(Z_N) \le V(U_N) + N - 1, \qquad V(U_N) \le V(Z_{2N-1}).$$

A cyclic Boolean convolution (up to a permutation of its components) can be viewed as a system of Boolean sums of arguments $x_0, \ldots, x_{N-1}$ with a variable circulant matrix defined by the row $y_{N-1}, \ldots, y_0$. Since the complexity of a circuit (here, in the sense of the complexity measure $V(f)$) does not increases after a replacement of some inputs by constants, we can conclude that the complexity of the cyclic convolution of order N is at least the complexity of a system of Boolean sums with an arbitrary circulant $N \times N$ matrix. So, by Corollary 2, we obtain

Corollary 4. $V(U_N)$, $V(Z_N) = \Omega\left(N^2 \log^{-6} N\right)$.

References

[1] Gashkov S. B., Sergeev I. S. On the complexity of linear Boolean operators with thin matrices // J. Applied and Industrial Math. — 2011. — V. 5(2). — P. 202–211.

[2] Grinchuk M. I. Complexity of implementing cyclic Boolean matrices by means of gate circuits // Soviet Math. (Izvestiya VUZ. Matematika). — 1988. — V. 32(7). — 65–72.

[3] Korshunov A. D. Monotone Boolean functions // Russian Math. Surveys. — 2003. — V. 58(5). — P. 929–1001.

[4] Lupanov O. B. On rectifier and switching-and-rectifier circuits // Dokl. Akad. Nauk SSSR. — 1956. — V. 111(6). — P. 1171–1174. (in Russian) [translation is available at http://www.thi.informatik.uni-frankfurt.de/~jukna/boolean/lupanov56.pdf]

[5] Lupanov O. B. Asymptotic estimates for the complexity of control systems. — Moscow: Moscow State University Publishing House, 1984. (in Russian)

[6] Erdös P., Spencer J. Probabilistic methods in combinatorics. — New York—London: Wiley-Intersci. Ser. Discrete Math. Optim., Academic Press, 1974.

[7] Blum N. On negations in Boolean networks // LNCS. — V. 5760. — Berlin—Heidelberg: Springer-Verlag, 2009. — P. 18–29.

[8] Gardner R. J., Gronchi P. A Brunn—Minkowski inequality for the integer lattice // Trans. AMS. — 2001. — V. 353(10). — P. 3995–4024.

[9] Mehlhorn K. Some remarks on boolean sums // Acta Inf. — 1979. — V. 12. — P. 371–375.

[10] Ruzsa I. Z. Sum of sets in several dimensions // Combinatorica. — 1994. — V. 14. — P. 485–490.

[11] Wegener I. The complexity of boolean functions. — Stuttgart: Wiley, 1987.

Grinchuk Mikhail Ivanovich, e-mail: grinchuk@nw.math.msu.su
Sergeev Igor Sergeevich, e-mail: isserg@gmail.com

Notes (2017)

By now, $\lambda(N)$ is proven to be $\Omega(N/\log^2 N)$. There are several ways to show it, see e.g. [Jukna S., Sergeev I. Complexity of linear boolean operators. Foundations and Trends in Theoretical Computer Science. 2013. V. 9(1). 1–123] and references there.

An explicit circulant matrix A achieving $L_\vee(A)/L_\oplus(A) = N^{1-o(1)}$ was constructed in [Gashkov S. B., Sergeev I. S. A method for deriving lower bounds for the complexity of monotone arithmetic circuits computing real polynomials. Sbornik: Mathematics. 2012. V. 203(10), 1411–1447] with the use of a combinatorial result by J. Kóllar, L. Rónyai and T. Szabó.