ELSEVIER
Theoretical Computer Science 235 (2000) 89–107
Theoretical Computer Science
www.elsevier.com/locate/tcs

Some structural properties of low-rank matrices related to computational complexity

Bruno Codenotti [a], Pavel Pudlák [b,1], Giovanni Resta [a,*]

[a] *Istituto Matematica Computazionale, Consiglio Naz. delle Ricerche, Via S. Maria 46, 56126-Pisa, Italy*
[b] *Mathematical Institute, AVČR, Žitná 25, 115 67 Praha 1, Czech Republic*

Abstract

We consider the problem of the presence of short cycles in the graphs of nonzero elements of matrices which have sublinear rank and nonzero entries on the main diagonal, and analyze the connection between these properties and the rigidity of matrices. In particular, we exhibit a family of matrices which shows that sublinear rank does not imply the existence of triangles. This family can also be used to give a constructive bound of the order of $k^{3/2}$ on the Ramsey number $R(3,k)$, which matches the best-known bound. On the other hand, we show that sublinear rank implies the existence of 4-cycles. Finally, we prove some partial results towards establishing lower bounds on matrix rigidity and consequently on the size of logarithmic depth arithmetic circuits for computing certain explicit linear transformations. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Matrix rigidity; Low-rank matrices; Cycles in graphs; Ramsey number; Linear circuits

1. Introduction

The problem of relating the rank of a matrix to its structural properties given by the pattern of its nonzero entries is a classical problem in mathematics. In complexity theory the most famous instance of this problem is the relation between the communication complexity of a $\{0,1\}$ matrix and its rank over the field of reals [16, 12]. In this paper we consider general matrices over arbitrary fields and we study cycles in the graphs of their nonzero elements. Our goal is to prove lower bounds on the rigidity of matrices which would imply nonlinear lower bounds on some algebraic circuits. This

* Corresponding author.
E-mail Address: resta@imc.pi.cnr.it (G. Resta)

[1] Partially supported by grant A1019602 of the Academy of Sciences of the Czech Republic and by a cooperative research grant INT-9600919/ME-103 from the NSF (USA) and the MŠMT (Czech Republic), part of this research was done while visiting Istituto di Matematica Computazionale del CNR, Pisa.

PII: S0304-3975(99)00185-1

research goes in the direction proposed by Valiant [19], who defined the concept of rigidity of matrices and proved that lower bounds on the size of logarithmic depth circuits can be proved by constructing matrices with high rigidity. The *rigidity of a matrix M* is defined as the function $R_M(r)$, which for a given r gives the minimum number of entries of M which one has to change in order to reduce its rank to r or less. Valiant proved the following result.

Theorem 1 (Valiant [19]). *If for some $\varepsilon>0$ there exists $\delta>0$ such that the $n\times n$ matrix M_n has rigidity $R_{M_n}(\varepsilon n)\geqslant n^{1+\delta}$ over a field F, then the transformation $x\to M_n x$ cannot be computed by linear size and logarithmic depth circuits with gates computing linear functions over F.*

Another relation of this type was found by Razborov [15]. He proved that some weaker bounds on rigidity would imply that a $\{0,1\}$ matrix defines a function which is not in the communication complexity version of the polynomial hierarchy. However, the existing lower bounds on the rigidity are not sufficient even for that. Although both a random matrix and a matrix whose entries are different indeterminates have rigidity even larger than required by Theorem 1 (close to n^2), very little is known about *explicit* matrices. The best-known lower bounds on the rigidity of explicit matrices are of the form $\Omega((n^2/r)\log(n/r))$ [6], which gives only linear lower bounds on $R_M(\varepsilon n)$. It seems that Hadamard matrices have large rigidity over the real field, but the best bound is so far only $\Omega(n^2/r)$ (see [9]).

Let us call an *alternating cycle* an oriented graph which is a cycle, when the orientation is forgotten, and such that the orientation of the arcs on the cycle alternates with one exception (put otherwise, there is a vertex v on the cycle such that if we go around the cycle from v to v, the orientation of the edges alternates).

Let $A=(a_{ij})$ be an $n\times n$ matrix. The graph of nonzero entries of A is the directed graph with vertex set $\{1,\dots,n\}$, where (i,j) is an arc iff $a_{ij}\neq 0$.

Given a matrix, we call [2,2] *configuration* a 2×2 submatrix consisting of nonzero elements. In graph theoretical terms a [2,2] configuration corresponds to either an alternating 3-cycle, which is usually called a transitive triangle (and this occurs when the 2×2 submatrix has one entry on the main diagonal), or an alternating 4-cycle (when none of the entries of the 2×2 submatrix is on the main diagonal), or a 2-cycle (when the 2×2 submatrix has two entries on the main diagonal).

We are especially interested in odd alternating cycles as subgraphs of the graph of nonzero entries of low-rank matrices, because of a connection to matrix rigidity. The reason for such special attention to *odd* lengths is that an odd alternating cycle corresponds to a configuration in the matrix where one element is on the main diagonal, while an even alternating cycle (of length greater than or equal to 4) is not connected to it. This connection with the main diagonal will allow us to argue about the large rigidity of certain explicitly defined families of matrices. In particular, the truth of the following conjecture would imply the nonlinear rigidity of certain matrices, and thus nonlinear lower bounds on the size of some circuits.

Note that we have to make some nontriviality assumption, such as having nonzero elements on the main diagonal or having a *support* (see Section 4), in order to get any interesting implication from low-rank.

Conjecture 1 (The Odd Alternating Cycle Conjecture). *For every field F, there exist an odd k and $\varepsilon > 0$ such that every $n \times n$ matrix M with nonzero entries on the main diagonal, and such that $\mathrm{rank}(M) \leqslant \varepsilon n$, contains an alternating cycle of length k.*

It is worthwhile to notice that the Odd Alternating Cycle Conjecture, in addition to motivating some of the problems analyzed in this paper, also implies a nonlinear lower bound on the computation of cyclic shifts on semilinear circuits, a model introduced and studied in [13].

In Section 2 we provide the main motivation for studying short cycles in low-rank matrices, by giving an explicit construction of a family of rigid matrices, assuming the Odd Alternating Cycle Conjecture be true. (We call a family of matrices rigid, if for some fixed $\varepsilon > 0$ the rigidity $R_M(\varepsilon n)$ of any $n \times n$ matrix M from the family is superlinear.) Even if the conjecture fails, these matrices may be good candidates for large rigidity.

In Section 3 we describe a construction which shows that, for $k = 3$, Conjecture 1 does not hold, over any field. We also show that this construction can be modified to give a very simple constructive lower bound on the Ramsey number $R(3,k) \geqslant k^{3/2}$, which matches the best known one, due to Alon [1]. If the field has characteristic different from 2, the counterexample to Conjecture 1 can be provided by a symmetric matrix. For $GF[2]$ we instead describe (see Section 5) a family of symmetric matrices of rank $n/4 + 2$ with 1's on the main diagonal and without a triangle, which is so far the best bound in this case.

Our results should be contrasted with those by Rosenfeld [17], and Alon and Szegedy [2]. (Their results are stated in terms of vectors in E^r, but can be easily translated to statements on real-valued matrices.) Rosenfeld proved that a symmetric positive definite $n \times n$ matrix of rank $\leqslant n/2$ with ones on the main diagonal contains a triangle. The combination of Rosenfeld's and our result shows that the assumption of positive semidefiniteness is essential. On the other hand, Alon and Szegedy proved that, for every $\delta > 0$ there exists k such that there are symmetric positive semidefinite matrices with ones on the main diagonal and rank $\leqslant n^{\delta}$ with no $k \times k$ principal submatrices of nonzero elements. The minimal value of k for which they get sublinear rank can be computed from their proof, but for small values, in particular for $k = 4$, it is open whether such matrices must contain K_4.

In Section 4 we prove that matrices with sublinear rank over the real or complex field must contain a [2,2] configuration, in fact they must contain an alternating 4-cycle. For fields of nonzero characteristic the corresponding statement is an open problem.

In Section 5 we analyze some special cases, obtained by making further assumptions, under which Conjecture 1 is true for $k = 3$; we also consider the Conjecture over $GF[2]$ in the symmetric case.

Although our results do not improve on any current lower bound on circuit complexity, we nevertheless think that we made a visible progress in the area. Fundamental problems in circuit complexity cannot be solved by gradually increasing lower bounds. There is need of progress in associated combinatorial and algebraic problems, and this paper is a step in this direction.

2. New candidates for high rigidity

We now describe an explicit construction of circulant $\{0,1\}$ matrices which have rigidity of the order of $n(\log n)^{1/(k-1)}$, provided that the the Odd Alternating Cycle Conjecture is true. Recall that a circulant matrix is a matrix fully determined by its first row, each other row being a cyclic shift of the previous one. For technical reasons, we number rows and columns of the matrices starting from 0, rather than 1. We construct a circulant $\{0,1\}$ matrix C_n' whose first row has nonzero entries in columns $1, b, b^2, \ldots, b^{m-1}$, where the choices of b and m are described below.

Lemma 2. *Let $n = 2^{2m}-1$, and define $a = 2^{2m-1}+2^{m-1}$, and $b = a+1$. The following relations hold over $\mathbf{Z}_n$ for $1 \leqslant h \leqslant m$:*

$$a^2 \equiv_n a, \tag{1}$$

$$b^h \equiv_n 2^{2m-1} - 2^{m-1} + 2^{h-1} + 2^{h+m-1}. \tag{2}$$

Proof. From $2^{2m} \equiv_n 1$, we easily obtain (1), since

$$a^2 = 2^{4m-2} + 2 \cdot 2^{2m-1}2^{m-1} + 2^{2m-2} \equiv_n 2 \cdot 2^{2m-2} + 2^{m-1} \equiv_n a.$$

Hence we also have that $a^h \equiv_n a$, for $h>0$. Relation (2) is obtained as follows:

$$\begin{aligned} b^h &= (a+1)^h = 1 + \sum_{i=1}^{h} \binom{h}{i} a^i = (2^h-1)a + 1 \\ &\equiv_n 2^{h-1} + 2^{h+m-1} - 2^{2m-1} - 2^{m-1} + 1 \equiv_n 2^{2m-1} - 2^{m-1} + 2^{h-1} + 2^{h+m-1}, \end{aligned}$$

where we used (1) and $2^{2m} \equiv_n 1$ to simplify the expressions. $\square$

Corollary 3. *The set $\{1, b, b^2, \ldots, b^{m-1}\}$, with the elements taken modulo n, has size m and it is a subgroup of the multiplicative group $\mathbf{Z}_n^*$.*

Proof. The size is immediate from (2). To see that it is a subgroup, just check that $b^m \equiv_n 1$. $\square$

Let us consider, for an integer α invertible over $\mathbf{Z}_n$, a matrix C_n'' defined by

$$c''_{i,j} = c'_{\alpha i, \alpha j}, \tag{3}$$

where indices run from zero and are computed over $\mathbf{Z}_n$.

It is easy to see that the effect of (3) is to permute the diagonals in such a way that C''_n is still circulant. In particular, if $\alpha = b^{-j}$, with $1 \leqslant j \leqslant m-1$, the elements of the diagonal corresponding to b^j are moved to diagonal 1, and, since $\{1, b, b^2, \dots, b^{m-1}\}$ and $\{b^{-j}, b^{-j+1}, \dots, b^{m-1-j}\}$ coincide (by Corollary 3), we have $C''_n = C'_n$.

We summarize relevant properties of C'_n in the following observations.

Observation 1. *Let $n = 2^{2m-1}$. There are $m-1$ permutation matrices Q_k such that the automorphism $Q_k C'_n Q_k^{\mathrm{T}} = C''_n = C'_n$ corresponds to transformation* (3). *In particular the permutation matrix Q_h, defined as $q_{ij} = 1$ iff $j = b^{-h} i$ and* 0 *elsewhere, takes the elements of diagonal b^h onto diagonal* 1.

Observation 2. *Let M be the matrix obtained from C'_n by deleting its first column and last row. M has a principal submatrix of order $n/4$ which is an identity matrix, since it is easy to verify, from* (2), *that $n/2 < b^j \pmod n < 3n/4$, for $1 \leqslant j \leqslant m-1$.*

The above two observations can be used to prove the following Theorem.

Theorem 4. *Assuming the Odd Alternating Cycle Conjecture for a field F and an odd k there exists an $\varepsilon > 0$ such that*

$$R_{C'_n}(\varepsilon n) = \Omega(n(\log n)^{1/(k-1)}).$$

Proof. By Observation 2, we have that the submatrix M (associated to the first diagonal of C'_n) contains an $n/4 \times n/4$ identity matrix.

Let us assume that the Odd Alternating Cycle Conjecture be true, for an odd k and a constant $\varepsilon > 0$. Then it is easy to see that matrices of rank at most $\frac{1}{2}\varepsilon n$ must contain a linear number of alternating k-cycles. In order to decrease the rank of M below $\frac{1}{2}\varepsilon n$, we must thus either introduce a linear number of alternating k-cycles or change a linear number of the diagonal entries to 0. By Observation 2, we actually have a linear number of alternating k-cycles which do not contain entries from other diagonals of C'_n.

By Observation 1, we can rearrange C'_n by means of permutations so that the elements of each diagonal can in turn be moved to the first diagonal. This implies that we can repeat the previous argument for all the m *diagonals* of C'_n. Thus either more than half of the elements on more than half of the diagonals are changed to 0, in which case we are done, as this gives $\Omega(nm) = \Omega(n \log n)$ changes, or there are $\Omega(nm)$ alternating k-cycles. To get a lower bound on the number of changes in the latter case, we let d be the average number of changes in a row. We may assume that each row and each column contains at most $4d$ changes. The number of alternating k-cycles can be easily upper bounded by a function of the order of nd^{k-1}. Thus we get $nd^{k-1} = \Omega(nm)$, whence the number of changes must be $\Omega(n(\log n)^{1/(k-1)})$. $\square$

3. Sublinear rank matrices without triangles

We show here the construction of an $n \times n$ matrix with ones on the main diagonal, rank of the order of $n^{2/3}$, and such that the graph obtained by associating edges to nonzero entries of the matrix does not have transitive triangles. As a byproduct of our construction, we find another constructive bound on the Ramsey number $R(3,n)$ which is simpler than the construction obtained by Alon [1].

Theorem 5. *For every m, there is an explicitly definable square matrix M of size $n \approx m^{3/2}$ which has* 1*'s on the main diagonal, rank $\leqslant m$, and such that the associated graph of nonzero elements does not contain a transitive triangle.*

Proof. We shall use an auxiliary undirected graph G with no cycles of length less than 6 or 8, and with a nonlinear number of edges. The following simple construction [20] provides us with an example of an infinite family of such graphs.

For every prime number p and for $k=2$ or $k=3$, we construct a bipartite graph $H_k(p)$ as follows. The vertex set of $H_k(p)$ is the union of two sets V_1 and V_2 with $|V_1| = |V_2| = p^k$ each. Each vertex $a \in V_1$ has a unique label $(a_0, a_1, \dots, a_{k-1})$ with $0 \leqslant a_i \leqslant p-1$. Similarly each $b \in V_2$ has a unique label $(b_0, b_1, \dots, b_{k-1})$. (The labels are simply the numbers from 0 to $p^k - 1$ expressed in base p notation.) The edge set of $H_k(p)$ is $\{(a,b)\colon\ a \in V_1,\ b \in V_2,\ b_j \equiv_p a_j + a_{j+1} b_{k-1},\ \text{for } j = 0,\dots,k-2\}$. The bipartite graphs $H_k(p)$ have $n = 2p^k$ vertices and $p^{k+1} = (n/2)^{((k+1)/k)}$ edges by construction. It is possible to prove [20] that $H_2(p)$, which has $\mathrm{O}(n^{3/2})$ edges, contains no cycles of length less than 6, and that $H_3(p)$, which has $\mathrm{O}(n^{4/3})$ edges, contains no cycles of length less than 8.

Other known constructions rely on finite projective geometry. A projective plane of order n can be defined as a set P of n^2+n+1 points and a set L of n^2+n+1 lines, such that any two points determine a line, any two lines determine a point, every point has $n+1$ lines on it, and every line contains $n+1$ points. Projective planes exists for every n equal to a prime power. The incidence graph of a projective plane is a bipartite graph whose vertices correspond to points and lines, and edges link each line with its incident points. Such a graph, by the properties of the projective plane, contains no cycles of length less than 6, and has $\mathrm{O}(n^3)$ edges. The asymptotic edge density is thus the same of the $H_2(p)$ graphs.

In this proof, we will be needing a graph without 6 cycles (either $H_2(p)$ or the graph of the projective plane), while for its extension to the symmetric case we will take $H_3(p)$.

Let now G be a graph without cycles of length less than 6, with m vertices and $\mathrm{O}(m^{3/2})$ edges. Starting from G, we first describe an oriented graph H, and then a matrix which have nonzero entries corresponding to edges of H. For the sake of a simpler description, in the following we assume that G is the bipartite graph associated with a projective plane.

The vertices of the graph H are pairs (P,L), where P corresponds to a point on a line L. The edges of H are given by pairs $((P,L),(P',L'))$, where P, P', L, L' are all different and $P \in L'$, i.e., the point P is incident to the line L'. It is straightforward to verify that there is no transitive triangle in H (using only the fact that there are no cycles of length less than 6 in G).

Now we associate to H a matrix as follows. We index both rows and columns by pairs (P,L). To a row (P,L), we assign the vector whose coordinates are the vertices of G (that is, points and lines) and which has an entry equal to -1 on P and an entry equal to 1 on L, all the other entries being zero. To a column (P,L), we assign the characteristic vector of the set $(L\backslash P)\cup\{L\}$, (i.e., the set consisting of the vertex L and all neighbours of L except P). Thus the matrix obtained as the product of the vectors associated to its rows and those associated with its columns has 1's on the main diagonal, -1's on the entries corresponding to the edges of H, and 0 elsewhere. The rank of this matrix is at most m, the number of vertices of G, while its size is equal to the number of edges of G, which is of the order of $m^{3/2}$. □

Notice that the above construction does not produce $\{0,1\}$ matrices, except for the field $GF[2]$.

It is an easy observation that the graph H has oriented 3-cycles if G contains 6-cycles. This prevents from applying the construction to the symmetric case, where oriented 3-cycles become triangles. However, to get a graph without oriented 3-cycles, we can start from a bipartite graph without cycles of length less than 8, like the previously described $H_3(p)$, and with $O(n^{4/3})$ edges. Then it is possible to proceed as in the proof, obtaining an $n\times n$ matrix of rank of the order of $n^{3/4}$, and then symmetrize the construction by adding the matrix and its transpose, thus getting an $n\times n$ symmetric matrix still of rank $O(n^{3/4})$, and without triangles.

However, note that the above approach fails over $GF[2]$, because symmetrization produces (over $GF[2]$) zeros on the main diagonal.

Using the fact that the rank (over any field) of a matrix is an upper bound on the size of the maximal independent set of a graph associated with the zero-nonzero pattern of the matrix, the above construction — after symmetrization — provides an explicit Ramsey graph. More precisely, it gives an n-vertex graph without triangles and with independent sets of size $O(n^{3/4})$. As it is, the bound is worse than the already mentioned Alon's bound. However, it is possible to work on the original construction, do another kind of symmetrization, and get the same asymptotic bound as Alon's, while significantly gaining in simplicity. The idea is to consider only the upper triangle of the matrix obtained from the graph H, and copy it in the lower part of the matrix. In this way, we obtain a symmetric matrix without triangles, without having to start with a sparser graph G. On the other hand, we loose the "low-rank" property, whereas the rank of the original matrix still bounds the size of the maximal independent set. We summarize these considerations in the following theorem.

Theorem 6. *Let G be the graph of a projective geometry with $2m$ vertices. Take a linear ordering of pairs (P,L), $P\in L$ and construct a* symmetric *graph J on this ordering by connecting (P,L) with (P',L'), where (P,L) is less than (P',L') in the ordering, iff P, P', L, L' are all different and $P\in L'$. Then the graph J neither contains a triangle nor an independent set of size $2m+1$, while the number of its vertices is of the order of $m^{3/2}$.*

On the other hand, one can prove an $\Omega(\sqrt{n \log n})$ lower bound on the rank matrices with ones on the main diagonal and without triangles using the well-known bound on the Ramsey number $R(3,k) = \Theta(k^2/\log k)$. Namely, let an $n \times n$ matrix M be given, and $n \geqslant R(3,k)$. Color the edges of the complete graph on n vertices blue, if the corresponding entry in the right upper half of M is nonzero, and red otherwise. If M does not contain a transitive triangle in the right upper part, the complete graph does not contain a blue triangle. Hence there must be a red complete subgraph on k elements. This corresponds to a $k \times k$ principal submatrix with zeros above the main diagonal, which has rank k. Thus the least rank which implies the existence of triangles is between $c_1\sqrt{n \log n}$ and $c_2 n^{2/3}$, for some constants c_1, c_2.

4. 4-cycles in low-rank matrices

Triangles are special cases of a [2,2] configuration. We have seen so far that sublinear rank matrices need not contain them, except in special cases, which will be dealt with in the next section.

Here we consider both general [2,2] configurations and 4-cycles, and show that they must appear in any matrix with sublinear rank. We first analyze the special case of $\{0,1\}$ matrices with constant row sums, and then generalize the result obtained in this case to arbitrary real matrices. We analyze real matrices, although all the results can be easily extended to complex matrices.

We will take advantage of the following lemma [3].

Lemma 7. *Le A be an $n \times n$ real symmetric matrix. Let* $\mathrm{tr}(A)$ *denote the trace of A, i.e., the sum of the diagonal entries of A, and* $\mathrm{rk}(A)$ *the rank of A, over the real field. We have*

$$\mathrm{rk}(A) \geqslant \frac{(\mathrm{tr}(A))^2}{\mathrm{tr}(A^2)}. \tag{4}$$

Proof. Let k be the rank of A. Since A is symmetric, k is equal to the number of nonzero eigenvalues of A. Let λ_i, $i = 1,\dots,n$, be the eigenvalues of A, and assume that the first k of them are different from zero. By elementary properties of the trace, we obtain

$$\mathrm{tr}(A^2) = \sum_i^n \lambda_i^2 = \sum_i^k \lambda_i^2 \quad \text{and} \quad \mathrm{tr}(A))^2 = \left(\sum_i^n \lambda_i\right)^2 = \left(\sum_i^k \lambda_i\right)^2$$

and the thesis follows from Cauchy–Schwarz inequality, since

$$\left(\sum_i^k \lambda_i\right)^2 \leqslant k \sum_i^k \lambda_i^2.$$

Lemma 8. *Let $B = I + A$ be a $\{0,1\}$ $n \times n$ symmetric matrix with* 1'*s on the main diagonal and row sums all equal to $d \geqslant 2$.*

1. *If B does not contain 2×2 full submatrices, except for those with two entries on the main diagonal, then the following equalities hold*:

$$\begin{aligned} &\mathrm{tr}(A) = 0, && \mathrm{tr}(B) = n,\\ &\mathrm{tr}(A^2) = n(d-1), && \mathrm{tr}(B^2) = nd,\\ &\mathrm{tr}(A^3) = 0, && \mathrm{tr}(B^3) = n(3d-2),\\ &\mathrm{tr}(A^4) = n(2d^2-5d+3), && \mathrm{tr}(B^4) = n(2d^2+d-2). \end{aligned} \tag{5}$$

2. *If B does not contain 2×2 full submatrices, except for those with one or two entries on the main diagonal, then the following equalities and inequalities hold*:

$$\begin{aligned} &\mathrm{tr}(A) = 0, && \mathrm{tr}(B) = n,\\ &\mathrm{tr}(A^2) = n(d-1), && \mathrm{tr}(B^2) = nd,\\ &\mathrm{tr}(A^3) \leqslant n(d^2-3d+2), && \mathrm{tr}(B^3) \leqslant nd^2,\\ &\mathrm{tr}(A^4) = n(2d^2-5d+3), && \mathrm{tr}(B^4) \leqslant n(6d^2-11d+6). \end{aligned} \tag{6}$$

Proof. 1. Let us consider the undirected graph G associated with A. Since A does not contain 2×2 full submatrices, G cannot have cycles of length 1, 3 or 4. It is well known that the entry $a_{ii}^{(k)}$ on the main diagonal of A^k is equal to the number of closed walks of length k in G, which originate and terminate at node i. Since the nodes of G have degree $d-1$, then $a_{ii}^{(2)} = d-1$. The lack of self-loops and 3-cycles implies that $a_{ii}^{(3)} = 0$. The closed walks of length 4 from i to i consist of the $(d-1)^2$ walks $i \to j \to i \to h \to i$, with $j, h \neq i$, and of the $(d-1)(d-2)$ walks $i \to j \to h \to j \to i$, for $j \neq i$ and $h \neq i, j$.

2. We evaluate $\mathrm{tr}(A^3)$ and $\mathrm{tr}(A^4)$ under the assumption that A can contain 3-cycles. The number of closed walks of length 3 from i to i is equal to the number of 3-cycles, since A does not contain self-loops, and thus we obtain $\mathrm{tr}(A^3) \leqslant n(d-1)(d-2) = n(d^2-3d+2)$. It is easy to see that the number of closed walks of length 4 from i to i does not increase w.r.t. the previous case, hence we still have $\mathrm{tr}(A^4) = n(2d^2-5d+3)$.

In both cases, the values of $\mathrm{tr}(B^k)$ are easily obtained by expanding $B^k = (I + A)^k$. □

Theorem 9. *Let B be a $\{0,1\}$ $n \times n$ symmetric matrix with* 1'*s on the main diagonal constant row sums.*

1. *If* $\mathrm{rk}(B) < \frac{1}{2}n$, *then B contains a 2×2 full submatrix without two entries on the main diagonal*;
2. *if* $\mathrm{rk}(B) < \frac{1}{6}n$, *then B contains a 2×2 full submatrix with no entries on the main diagonal.*

Proof. Assuming by contradiction that B does not contain a 2×2 full submatrix, then equalities (5) hold.

Basic properties of rank imply that $\mathrm{rk}(B) \geqslant \mathrm{rk}[B(B-I)] = \mathrm{rk}(B^2 - B)$. Since $B^2 - B$ is symmetric we can apply inequality (4) to $B^2 - B$, obtaining

$$\mathrm{rk}(B) \geqslant \mathrm{rk}(B^2 - B) \geqslant \frac{[n(d-1)]^2}{n(2d^2 - 4d + 2)} = \frac{1}{2}n \tag{7}$$

for $d \geqslant 2$, which is a contradiction with the assumption $\mathrm{rk}(B) < \frac{1}{2}n$, and the thesis follows.

The proof of the second case, $\mathrm{rk}(B) < \frac{1}{6}n$, is obtained similarly, applying the inequalities (6) and (4) to B^2. $\square$

We now proceed to generalizing the above results. The intermediate step is provided by Lemma 10 below, which gives a lower bound on the rank of doubly stochastic matrices without [2, 2] configurations. This Lemma is then instrumental to obtain a more general result.

In the following we denote by $A \cdot B$ the element-wise product of two matrices.

Lemma 10. *Let B be an $n \times n$ matrix such that the matrix $B \cdot B$ is doubly stochastic, i.e., $\sum_j b_{ij}^2 = \sum_i b_{ij}^2 = 1$ for $i, j = 1, \ldots, n$.*
- *If B does not contain a full 2×2 submatrix, then $\mathrm{rk}(B) > n/2$;*
- *if B does not contain full 2×2 submatrices, except for those with two entries on the main diagonal, then $\mathrm{rk}(B) > n/4$;*
- *if B does not contain full 2×2 submatrices, except for those with one or two entries on the main diagonal, then $\mathrm{rk}(B) > n/6$.*

Proof. We have $\mathrm{rk}(B) \geqslant \mathrm{rk}(BB^\top)$, and we can apply the inequality

$$\mathrm{rk}(C) \geqslant \frac{(\mathrm{tr}(C))^2}{\mathrm{tr}(C^2)}$$

to $C = BB^\top$.

The value of $\mathrm{tr}(BB^\top) = \mathrm{tr}(C)$ can be easily calculated, since $\sum_j b_{ij}^2 = 1$ for $i = 1, \ldots, n$. Indeed we have $c_{ij} = \sum_h b_{ih} b_{jh}$, hence $c_{ii} = \sum_h b_{ih}^2 = 1$, and $\mathrm{tr}(C) = \sum_i c_{ii} = n$. The trace of C^2 can be computed similarly:

$$\mathrm{tr}(C^2) = \sum_{i,j} c_{ij} c_{ji} = \sum_{i,j} \left(\sum_h b_{ih} b_{jh} \right) \left(\sum_k b_{ik} b_{jk} \right) = \sum_{i,j,h,k} b_{ih} b_{jh} b_{ik} b_{jk}.$$

Let δ_{ij} be equal to 1 if $i = j$, and 0 otherwise. Then we have

$$\mathrm{tr}(C^2) = [\delta_{ij} + \delta_{hk} - \delta_{ij}\delta_{hk} + (1 - \delta_{ij})(1 - \delta_{hk})] \sum_{i,j,h,k} b_{ih} b_{jh} b_{ik} b_{jk}$$

and thus

$$\operatorname{tr}(C^2) = \sum_{i,h,k} b_{ih}^2 b_{ik}^2 + \sum_{i,j,h} b_{ih}^2 b_{jh}^2 - \sum_{i,h} b_{ih}^4 + \sum_{i\neq j, h\neq k} b_{ih} b_{jh} b_{ik} b_{jk}.$$

Exploiting the fact that $B \cdot B$ is doubly stochastic, we get

$$\sum_{i,h,k} b_{ih}^2 b_{ik}^2 = \sum_i \left[\sum_j b_{ij}^2\right]^2 = n \quad \text{and} \quad \sum_{i,j,h} b_{ih}^2 b_{jh}^2 = \sum_h \left[\sum_k b_{kh}^2\right]^2 = n$$

and, since $\sum_{ij} b_{ij}^4 > 0$, we have $\operatorname{tr}(C^2) < 2n + T$, with

$$T = \sum_{i\neq j, h\neq k} b_{ih} b_{jh} b_{ik} b_{jk}\,.$$

The products in T correspond exactly to all 2×2 nonzero submatrices of B (each counted 4 times due to symmetries). Hence, if B does not contain such submatrices, we have $T = 0$ and $\operatorname{tr}(C^2) < 2n$, from which

$$\operatorname{rk}(B) = \operatorname{rk}(C) \geqslant \frac{\operatorname{tr}(C)^2}{\operatorname{tr}(C^2)} > \frac{n^2}{2n} = \frac{n}{2} \tag{8}$$

as claimed.

If there are 2×2 nonzero submatrices with two entries on the main diagonal of B, then the nonzero terms in T are those obtained setting either $i = h$, $j = k$ or $i = k$, $j = h$. The two cases are disjoint (since $i \neq j$ and $h \neq k$) and thus we get

$$\sum_{i\neq j, h\neq k} b_{ih} b_{jh} b_{ik} b_{jk} = 2 \sum_{i\neq j} b_{ii} b_{ij} b_{ji} b_{jj} \leqslant 2 \sum_{i,j} b_{ij} b_{ji} \leqslant 2 \sum_{i,j} \frac{b_{ij}^2 + b_{ji}^2}{2} = 2n$$

by stochasticity and using the inequality $xy \leqslant (x^2 + y^2)/2$. Summarizing we obtain $\operatorname{tr}(C^2) < 4n$, from which, proceeding as in (8), we have $\operatorname{rk}(B) > n/4$.

If there are also the 2×2 submatrices with one entry on the main diagonal, then at least one of $i = h$, $i = k$, $j = h$, $j = k$ holds, so that we can write

$$\begin{aligned}\sum_{i\neq j, h\neq k} b_{ih} b_{ik} b_{jh} b_{jk} &\leqslant \sum_{ijh} b_{ih} b_{ii} b_{jh} b_{ji} + \sum_{ijk} b_{ij} b_{ik} b_{jj} b_{jk} \\ &\quad + \sum_{ijh} b_{ih} b_{ij} b_{jh} b_{jj} + \sum_{ijk} b_{ii} b_{ik} b_{jh} b_{jk}.\end{aligned}$$

We have

$$\sum_{ijh} b_{ih} b_{ii} b_{jh} b_{ji} = \sum_{ij} b_{ii} b_{ji} \sum_h b_{ih} b_{jh} \leqslant \sum_{ij} b_{ii} b_{ji} \sum_h \frac{b_{ih}^2 + b_{jh}^2}{2} = \sum_{ij} b_{ii} b_{ji}$$

and

$$\sum_{ij} b_{ii} b_{ji} \leqslant \sum_i \sum_j \frac{b_{ii}^2 + b_{ji}^2}{2} \leqslant n.$$

The other three terms can be bounded analogously, thus giving $\mathrm{tr}(C^2)<2n+4n=6n$. Proceeding as in (8), we have $\mathrm{rk}(B)>n/6$.

Definition 11. If A is an $n\times n$ matrix and σ is a permutation of $\{1,\dots,n\}$, then the sequence of elements $a_{1,\sigma(1)},\dots,a_{n,\sigma(n)}$ is called the *diagonal* corresponding to σ.

Definition 12. Let A be a square matrix. A is said to have *support* if it contains a diagonal of nonzero elements. A is said to have *total support* if every nonzero element of A lies on a diagonal of nonzero elements.

Lemma 11. *Let A be a square matrix with total support, then there exists a matrix B with the same rank and nonzero pattern of A such that $B\cdot B$ is doubly stochastic.*

Proof. The matrix $A'=A\cdot A$, whose nonzero pattern is the same of A, is nonnegative and has total support. By [18] there exist two positive diagonal matrices D_1 and D_2 such that $D_1A'D_2$ is doubly stochastic. It is easy to verify that the matrix $B=D_1^{1/2}AD_2^{1/2}$ meets our conditions, since it has the same pattern and rank of A and

$$B\cdot B=(D_1^{1/2}AD_2^{1/2})\cdot(D_1^{1/2}AD_1^{1/2})=(D_1^{1/2}\cdot D_1^{1/2})(A\cdot A)(D_2^{1/2}\cdot D_2^{1/2})=D_1A'D_2. \qquad \square$$

Theorem 12. *Let A be an $n\times n$ matrix with total support.*
- *If* $\mathrm{rk}(A)\leqslant n/2$, *then A contains at least one full 2×2 submatrix*;
- *if* $\mathrm{rk}(A)\leqslant n/4$, *then A contains at least one full 2×2 submatrix with at most one entry on the main diagonal*;
- *if* $\mathrm{rk}(A)\leqslant n/6$, *then A contains at least one full 2×2 submatrix with no entry on the main diagonal.*

Proof. Let us suppose, by contradiction, that A does not contain a full 2×2 submatrix. Then by Lemmas 10 and 11 there exists a matrix B such that $\mathrm{rk}(A)=\mathrm{rk}(B)>n/2$. This contradicts the assumption $\mathrm{rk}(A)\leqslant n/2$, and hence proves the thesis. The other two cases, $\mathrm{rk}(A)\leqslant n/4$ and $\mathrm{rk}(A)\leqslant n/6$, are similar. $\square$

Corollary 13. *Let A be an $n\times n$ matrix with support.*
- *If* $\mathrm{rk}(A)\leqslant n/2$, *then A contains at least one full 2×2 submatrix*;
- *if* $\mathrm{rk}(A)\leqslant n/4$, *then A contains at least one full 2×2 submatrix with at most one entry on the main diagonal*;
- *if* $\mathrm{rk}(A)\leqslant n/6$, *then A contains at least one full 2×2 submatrix with no entry on the main diagonal.*

Proof. If A does not have total support, then there exist two permutation matrices P and Q such that $B=PAQ$ is a block triangular matrix, whose diagonal blocks $B_1,B_2,\dots,B_t$ have total support [4]. The matrices B_i are uniquely determined within arbitrary permutation of their rows, but their ordering in B is not necessarily unique. As a consequence we cannot guarantee that the multiplication by P and Q maps all the diagonal elements of A inside the union of the blocks B_i. However, we can choose P and Q in such a way that every diagonal element of A is either outside the union of the diagonal matrices B_i or on the diagonal of one of them.

The sum of the ranks of the diagonal blocks of B does not exceed the rank of A, and thus, if $\mathrm{rk}(A)\leqslant n/k$, there is at least a block, say B_j, of size m and rank less than m/k. Depending on k we consider three cases.

If $\mathrm{rk}(B_j)\leqslant m/2$, by Theorem 12, we have that B_j, and thus $A=P^{-1}BQ^{-1}$ contains a 2×2 submatrix of nonzeros, as claimed. In fact it is easy to see that the multiplication by permutation matrices, like P and Q and their inverses, preserves such a structure.

If $\mathrm{rk}(B_j)\leqslant m/4$, by Theorem 12, we have that B_j contains a 2×2 submatrix of nonzeros, with at most one entry on the main diagonal of B_j. Since all the diagonal elements of A which appear inside B_j are located on the main diagonal of B_j, we have that the 2×2 submatrix of B_j, once mapped again in A, can contain at most one diagonal entry, as required.

The case $\mathrm{rk}(B_j)\leqslant m/6$ is similar to the case $\mathrm{rk}(B_j)\leqslant m/4$. $\square$

5. Further results

5.1. Symmetric case over $GF[2]$

We consider here symmetric $\{0,1\}$ matrices with low-rank over $GF[2]$, a case for which we cannot apply the result of Section 3. In the following, we describe a family of matrices of rank $n/4+2$ and without triangles.

Let I_k, J_k, and P_k denote the identity matrix, the matrix with all the entries equal to 1, and the matrix with the $(i,k-i)$th entries equal to 1, respectively, all of size k. Let us consider the following $n\times n$ matrix, for $n=4k$, written in block form:

$$A_n=I_n+B_n=\begin{pmatrix} I_k & I_k & I_k & J_k-P_k \\ I_k & I_k & J_k-I_k & P_k \\ I_k & J_k-I_k & I_k & P_k \\ J_k-P_k & P_k & P_k & I_k \end{pmatrix}.$$

This is a family of symmetric matrices with the following properties:

- A_n is triangle-free. This property can be easily verified computing the trace of B_n^3. We have

$$B_n^2=\begin{pmatrix} (k-2)J_k+3I_k & 2(J_k-I_k) & 2(J_k-I_k) & 2P_k \\ 2(J_k-I_k) & (k-2)J_k+3I_k & J_k-I_k & 2(J_k-P_k) \\ 2(J_k-I_k) & J_k-I_k & (k-2)J_k+3I_k & 2(J_k-P_k) \\ 2P_k & 2(J_k-P_k) & 2(J_k-P_k) & (k-2)J_k+3I_k \end{pmatrix},$$

from which we readily see that $\mathrm{Tr}(B_n^3)=4\mathrm{Tr}(6J_k-6I_k)=0$.

- $\text{Rank}_2(A) = r = n/4+2$. Indeed the matrices A_n can be obtained as $A_n = UU^\top$, where $U^\top$ is the following $r \times n$ matrix:

$$\left(\begin{array}{c|ccc|c|ccc|c|ccc|ccc|c} 1 & 0 & \cdots & 0 & 0 & 1 & \cdots & 1 & 0 & 1 & \cdots & 1 & 1 & \cdots & 1 & 0 \\ 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & 1 & \cdots & 1 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 & 1 & \cdots & 1 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 \\ \hline 0 & & & & 0 & & & & 0 & & & & & & & 1 \\ \vdots & & I_{r-3} & & \vdots & & I_{r-3} & & \vdots & & I_{r-3} & & & J_{r-3} - P_{r-3} & & \vdots \\ 0 & & & & 0 & & & & 0 & & & & & & & 1 \end{array}\right). \tag{9}$$

- A_n is regular of degree r.
- B_n has independent sets of size $k = r - 2$.
- $\text{Rank}_{\mathbf{R}}(A_n) = n$, and A_n has exactly five distinct integer eigenvalues. More precisely A_n has the following eigensystem:
  - ○ $\lambda_1 = r$ with multiplicity 1 and eigenvector $(1, \ldots, 1)^\top$.
  - ○ $\lambda_2 = r - 4$ with multiplicity 1 and eigenvector

$$(\overbrace{-1, \ldots, -1}^{k}, \overbrace{1, \ldots, 1}^{k}, \overbrace{1, \ldots, 1}^{k}, \overbrace{-1, \ldots, -1}^{k})^\top.$$

  - ○ $\lambda_3 = 4 - r$ with multiplicity 2 and eigenvectors

$$(\overbrace{-1, \ldots, -1}^{k}, \overbrace{0, \ldots, 0}^{k}, \overbrace{0, \ldots, 0}^{k}, \overbrace{1, \ldots, 1}^{k})^\top$$

and

$$(\overbrace{-1, \ldots, -1}^{k}, \overbrace{1, \ldots, 1}^{k}, \overbrace{-1, \ldots, -1}^{k}, \overbrace{1, \ldots, 1}^{k})^\top.$$

  - ○ $\lambda_4 = -2$ with multiplicity $k - 1$ and eigenvectors

$$\left(\begin{array}{c|c|c|c|c|c|c|c} 1 & & -1 & & -1 & & & 1 \\ \vdots & -P_{k-1} & \vdots & P_{k-1} & \vdots & P_{k-1} & -I_{k-1} & \vdots \\ 1 & & -1 & & -1 & & & 1 \end{array}\right)^\top.$$

∘ $\lambda_5 = 2$ with multiplicity $n - r - 1 = 3(k-1)$ and eigenvectors

$$\left(\begin{array}{c|c|c|c}
\begin{array}{c|c} -I_{k-1} & \begin{matrix}1\\ \vdots\\ 1\end{matrix}\end{array} & 0 & 0 & \begin{array}{c|c} \begin{matrix}-1\\ \vdots\\ -1\end{matrix} & P_{k-1}\end{array} \\ \hline
\begin{array}{c|c} \begin{matrix}-1\\ \vdots\\ -1\end{matrix} & P_{k-1}\end{array} & 0 & \begin{array}{c|c} \begin{matrix}-1\\ \vdots\\ -1\end{matrix} & P_{k-1}\end{array} & 0 \\ \hline
\begin{array}{c|c} \begin{matrix}-1\\ \vdots\\ -1\end{matrix} & P_{k-1}\end{array} & \begin{array}{c|c} \begin{matrix}-1\\ \vdots\\ -1\end{matrix} & P_{k-1}\end{array} & 0 & 0
\end{array}\right)^{\top}.$$

The verification of the above properties can now be done by direct inspection.

It is an open problem whether or not the above family is extremal, i.e., whether all symmetric $\{0,1\}$ matrices with rank over $GF[2]$ at most $n/4+1$ must contain triangles.

5.2. Special cases for which Conjecture 1 holds

We now consider sublinear rank matrices satisfying additional constraints which guarantee that they must contain triangles. The main result of this section is Theorem 15, where we show that a low-rank matrix which can be factored in terms of sparse matrices must contain triangles.

5.2.1. Low rank and sparse factors

It is known that a symmetric $\{0,1\}$ matrix M can be factored over $GF[2]$ as $M = AA^{\top}$, where the number of columns of A is equal to either r or $r+1$, and r is the rank of M over $GF[2]$, see [11]. If at least one diagonal element is equal to 1, then the number of columns of A is r. The above decomposition of a symmetric matrix M over $GF[2]$ can be interpreted as representing M as an intersection matrix as follows. The rows and columns are indexed by sets of some family of subsets of $\{1,\dots,r\}$, where r is the rank of M. The (i,j)th entry of M is 1 iff the intersection of the index sets corresponding to row i and column j is odd. The rows of A are the characteristic vectors of the sets. We shall call set systems also *hypergraphs*. If all sets have size k, then we speak of *k-hypergraphs*. This representation allows us to investigate the presence of triangles in M in a purely combinatorial way, by treating the rows of the matrix A as a set system. Since we assume that all the diagonal entries are nonzero, we have that all sets have odd cardinalities. A triangle corresponds to three sets, every two of which intersecting in odd sets.

Definition 3. A sunflower (also called a *delta system*, or a star) with l petals and core Y is a family of sets $X_1,\dots,X_l$ such that $X_i \cap X_j = Y$ for every $i \neq j$.

Note that the assumption that $X_1,\dots,X_l$ form a sunflower with an odd core is stronger than the assumption that every two sets intersect in an odd set. A classical result of Erdős and Radó [5] states that for a given k and l, every sufficiently large k-hypergraph contain a sunflower with l petals. It has been observed in [8] that a k-hypergraph, k odd, on an n element set of vertices with at least $n^{(k-1)/2}l^{(k+1)/2}(1\cdot 3\cdots k/2\cdot 4\cdots(k-1))$, edges contains a sunflower with l petals and with an *even* core. A k-hypergraph can be easily constructed which shows that this bound cannot be essentially improved. On the other hand, the next theorem shows that the case of odd cores is essentially different. The theorem is an easy consequence of a theorem of Füredi [7]. We are indebted to V. Rödl for suggesting this proof, which replaces our original proof, which was self contained, but more complicated.

Theorem 14. *For every positive integers k,l, with k odd and $l\geqslant 3$, there exists an integer K such that for every $n\geqslant 1$ and every k-hypergraph H on n vertices with at least Kn edges, there exists a sunflower in H with l petals and an odd core.*

Proof. Let a k-hypergraph on n vertices be given. By Füredi's theorem there exists a subhypergraph H' with $|H'|\geqslant\varepsilon|H|$ such that for every two different sets $X,Y\in H'$, the intersection $X\cap Y$ is a core of a sunflower with l petals contained in H'. The constant $\varepsilon>0$ depends only on k and l. Thus we only need to show that there is at least one odd intersection. Suppose $K\varepsilon>1$, then $|H'|>1$. Hence, by the odd town theorem (an easy algebraic argument) there is at least one odd intersection and we are done. □

This theorem implies that if $M=AA^\top$, where the $n\times n/K$ matrix A is "sparse" and M has 1's on the main diagonal, then the associated matrix must contain a complete graph on l vertices. We shall show that a similar statement holds also with $M=AB^\top$, where A and B may be different sparse matrices.

Theorem 15. $\forall k,l$, $\exists\varepsilon>0$ *such that* $\forall n$, *if* $M=AB$ (*over* $GF[2]$), *where M is an $n\times n$ matrix with ones on the main diagonal, A is an $n\times r$ matrix, B is an $r\times n$ matrix, $r\leqslant\varepsilon n$, A* (*resp.* B) *has at most k ones in each column* (*resp. row*), *then there exists an $l\times l$ principal submatrix of ones in M.* (*Note that M does not need to be symmetric.*)

Theorem 15 can be reformulated in an equivalent way, in set intersection terms:

Theorem 16. $\forall k,l$, $\exists\varepsilon>0$, $\forall n$, *and for all sets* $(A_1,B_1),\dots,(A_n,B_n)$, $|A_i|,|B_i|\leqslant k$, $A_i,B_i\subseteq X$, $|X|=r\leqslant\varepsilon n$, *where for all i, if $|A_i\cap B_i|$ is odd, then there exist $i_1,\dots,i_l$, such that for all $1\leqslant\alpha,\beta\leqslant l$, with $\alpha\neq\beta$, we have that $|A_{i_\alpha}\cap B_{i_\beta}|$ is odd.*

Theorem 16 follows from a stronger result, which we prove below.

Theorem 17. $\forall k,l$, $\exists\varepsilon>0$, $\forall n$, *and for all sets* $(A_1,B_1),\dots,(A_n,B_n)$, $|A_i|,|B_i|\leqslant k$, $A_i,B_i\subseteq X$, $|X|=r\leqslant\varepsilon n$, *if for all i, $|A_i\cap B_i|$ is odd, then there exists a set D of* *odd cardinality and* $i_1,\ldots,i_l$ *such that for all* $1 \leqslant \alpha, \beta \leqslant l$, *with* $\alpha \neq \beta$, *we have that* $A_{i_\alpha} \cap B_{i_\beta} = D$.

Proof. Let $C_i = A_i \cap B_i$, i.e., $|C_i|$ is odd. By Theorem 14, we have that there exist $j_1,\ldots,j_m$ and D such that $\forall \alpha \neq \beta$, $C_{j_\alpha} \cap C_{j_\beta} = D$.

Now, for every i, choose the mappings

$f_i : P(A_i) \to \{0,\ldots,2^k - 1\}$,

$g_i : P(B_i) \to \{0,\ldots,2^k - 1\}$

and assign the colour $(f_i(A_i \cap B_j), g_j(A_i \cap B_j))$ to the pair (i,j), for $i<j$. By Ramsey Theorem, there exists $\{i_1,\ldots,i_l\} \subseteq \{j_1,\ldots,j_m\}$ such that all pairs have the same colour.

Claim. $\forall i,i',j,j' \in \{i_1,\ldots,i_l\}, i<j, i'<j'$, *we have* $A_i \cap B_j = A_{i'} \cap B_{j'}$.

Proof. $A_i \cap B_j = A_i \cap B_{j'} = A_{i'} \cap B_{j'}$.

Thus there exists a set D' of odd cardinality such that $A_i \cap B_j = D'$, for all $i,j \in \{i_1,\ldots, i_l\}$, $i<j$. Symmetrically, there exists D'' such that $A_i \cap B_j = D''$, for all $i,j \in \{i_1,\ldots,i_l\}$, $i>j$. W.l.o.g. we can assume that $l \geqslant 4$. Then $D' \subseteq A_1 \cap B_2 \cap A_2 \cap B_3 \cap A_3 \cap A_4 \subseteq A_2 \cap B_2 \cap A_3 \cap B_3 = D$. But also $D \subseteq A_1 \cap B_1 \cap A_2 \cap B_2 \subseteq A_1 \cap B_2 = D'$. Therefore $D = D'$. By symmetry $D = D''$.

Note that we have applied Theorem 14 although the sets C_i do not have all the same cardinality because there is only a constant number of different cardinalities, so that we can take those i's for which the size of intersection occurs more frequently. □

Note that sunflower theorems had been used at least twice to prove lower bounds on the size of circuits [14, 8].

5.2.2. *Rank equal to the maximal independent set*

We now consider the problem of rank vs. triangles, under the additional assumption that the rank is equal to the size of the largest independent set in the associated graph. Our main motivation is that, due to the connection with Ramsey theory, we would like to know how big the gap can be between rank and size of an independent set, in the case of triangle free graphs. Note that in our Ramsey construction of Section 3, the rank of the matrix is full!

Here we show that an $2n \times 2n$ 0–1 matrix M, with rank n, maximal independent set of size n, and without "triangles" must be bipartite.

Let I denote the identity matrix. We represent M (or $P^\top MP$, where P is a suitable permutation matrix) as a 2×2 block matrix, with $n \times n$ blocks:

$$M = \begin{pmatrix} I & B \\ C & E \end{pmatrix}$$

such that $E \geqslant I$, (element-wise). Note that the case $E = I$ would imply that the graph is bipartite.

We will show that M (in order not to have triangles) must be bipartite.

We can write

$$M = \begin{pmatrix} I & B \\ C & E \end{pmatrix} \quad \text{and} \quad M' = \begin{pmatrix} I & B \\ 0 & E - CB \end{pmatrix},$$

where M' is obtained from M by Gaussian elimination, and thus has the same rank as M. It is easy to see that if $E - CB \neq 0$ then the rank of M' would exceed n, hence we can assume that $E = CB$. If $E = I$ then M is clearly bipartite. Assuming by contradiction that $E \neq I$ we obtain that M must have triangles. In fact, since $E \geqslant I$, $E \neq I$ implies that E must have at least an off-diagonal nonzero entry, say e_{pq}, with $p \neq q$. Since $E = CB$ we have

$$1 = E_{pq} = \sum_{k=1}^{n} c_{pk} b_{kq}$$

which implies that there is at least one index k such that both c_{pk} and b_{kq} are equal to 1. The thesis follows by observing that e_{pq}, c_{pk} and b_{kq} form a triangle in M.

The above-described result is just a first step towards understanding whether or not there must be a significant gap between rank and size of a maximal independent set in a triangle free graph of low-rank.

6. Conclusions

The Odd Alternating Cycle Conjecture fails for $k = 3$, and a similar statement is false for complete graphs on k vertices, for k sufficiently large (even for positive semidefinite matrices). However, in spite of some effort, we were not able to disprove the conjecture even for $k = 5$. Furthermore, it is true for some special cases (e.g., products of sparse matrices) and the statement is true for $k = 4$ over the real field. These results do not give better bounds on the size of linear circuits, but we are thinking of some generalizations which may have such consequences. Thus we believe that this is a promising research area, which will eventually lead to results on the complexity of linear circuits and solve problems posed more than 20 years ago. As the construction of a Ramsey graph shows, this research also belongs to a mainstream area in combinatorics.

To conclude, we shall insist on three open problems, which we believe are the most important among those mentioned in this paper:

1. Does the Odd Alternating Cycle Conjecture hold for $k = 5$?
2. Does the Odd Alternating Cycle Conjecture hold for $k = 3$ for symmetric matrices over $GF[2]$?
3. Does sublinear rank and nonzero elements on the main diagonal imply the existence of a 2×2 submatrix of nonzero elements also for fields of nonzero characteristic?

Acknowledgements

Pavel Pudlák thanks Vojta Rödl for several fruitful discussions.

References

[1] N. Alon, Explicit Ramsey graphs and orthonormal labelings, Electron. J. Combin. 1 (1994) R12.
[2] N. Alon, M. Szegedy, Large sets of nearly orthogonal vectors, preprint.
[3] R. Bellman, Introduction to Matrix Analysis, McGraw-Hill, New York, 1960.
[4] R.A. Brualdi, H.J. Ryser, Combinatorial Matrix Theory, 1991, Cambridge University Press, Cambridge.
[5] P. Erdős, R. Rado, Intersection theorems for systems of sets, J. London Math. Soc. 35 (1960) 85–90.
[6] J. Friedman, A note on matrix rigidity, Combinatorica 13(2) (1993) 235–239.
[7] Z. Füredi, On finite set-systems whose every intersection is a Kernel of a star, Discr. Math. 47 (1983) 129–132.
[8] J. Håstad, S. Jukna, P. Pudlák, Top-down lower bounds for depth-three circuits, Comput. Complexity 5 (1995) 99–112.
[9] B.S. Kashin, A.A. Razborov, Improved lower bounds on the rigidity of Hadamard matrices, preprint, 1997.
[10] A. Kotlov, L. Lovász, The rank and size of graphs, J. Graph Theory 23(1) (1996) 185–189.
[11] A. Lempel, Matrix Factorization over $GF[2]$ and Trace-Orthogonal Bases of $GF[2^n]$, SIAM J. Comput. 4(2) (1975) 175–186.
[12] N. Nisan, A. Wigderson, On Rank vs Communication Complexity, Proc. 35th IEEE FOCS, 1994, pp. 831–836.
[13] P. Pudlák, V. Rödl, J. Sgall, Boolean circuits, tensor ranks and communication complexity, SIAM J. Comput. 26/3 (1997) 605–633.
[14] A.A. Razborov, Lower bounds on the monotone complexity of some boolean functions, Sov. Math. Dokl. 31 (1985) 354–357.
[15] A.A. Razborov, On rigid matrices, typeset manuscript in Russian, 1989.
[16] R. Raz, B. Spiker, On the log-rank conjecture in communication complexity, Proc. 34th IEEE FOCS, 1993, pp. 168–176.
[17] M. Rosenfeld, Almost orthogonal lines in E^d, DIMACS Ser. Discrete Math. 4 (1991) 489–492.
[18] R. Sinkhorn, P. Knopp, Concerning nonnegative matrices and doubly stochastic matrices, Pacific J. Math. 21 (2) (1967) 343–348.
[19] L.G. Valiant, Graph-theoretic arguments in low level complexity, Proc. 6th MFCS, Lecture Notes in Mathematics, Vol. 53, Springer, Berlin, 1977, pp. 162–176.
[20] R. Wenger, Extremal Graphs with no C_4's, C_6's or C_{10}'s, J. Combin. Theory Ser. B 52 (1991) 113–116.