On the rank of $\mathbb{Z}_2$-matrices with free entries
on the diagonal

E. Kogan *

Abstract

For an $n \times n$ matrix $M$ with entries in $\mathbb{Z}_2$ denote by $R(M)$ the minimal rank of all the matrices obtained by changing some numbers on the main diagonal of $M$. We prove that for each non-negative integer $k$ there is a polynomial in $n$ algorithm deciding whether $R(M) \leq k$ (whose complexity may depend on $k$). We also give a polynomial in $n$ algorithm computing a number $m$ such that $m/2 \leq R(M) \leq m$. These results have applications to graph drawings on non-orientable surfaces.

For an $n \times n$ matrix $M$ with entries in $\mathbb{Z}_2$ denote by $R(M)$ the minimal rank of all the matrices obtained by changing some numbers on the main diagonal of $M$. We present two algorithms estimating $R(M)$. These results (Theorems 1 and 2) have applications to graph drawings on non-orientable surfaces, see Appendix A. For a brief overview of the history of the problem, see Remark 4.

Denote by $M_n(\mathbb{Z}_2)$ the set of all $n \times n$ matrices with entries in $\mathbb{Z}_2$.

**Theorem 1.** Let $k$ be a fixed non-negative integer. There is an algorithm with the complexity of $O(n^{k+4})$ deciding for an arbitrary matrix $M \in M_n(\mathbb{Z}_2)$ whether $R(M) \leq k$.

The proof of Theorem 1 uses Lemma 5 and well known Lemma 6.

**Theorem 2.** There is an algorithm with the complexity of $O(n^4)$ calculating for an arbitrary matrix $M \in M_n(\mathbb{Z}_2)$ a number $k$ such that

$$k/2 \leq R(M) \leq k.$$

*We would like to thank R. Fulek, B. Mohar, V. Podolskii, D. Pozdeev, A. Skopenkov, I. Soloveychik, M. Tancer, and D. Trushin for helpful discussions and suggestions.
The proof of Theorem 2 uses Lemma 7 which in turn uses well known Lemma 6.

**Remark 3.** Let $M$ be an $n \times n$ matrix with entries in $\mathbb{Z}_2$. A matrix $M'$ obtained from $M$ by changing some numbers on the main diagonal of $M$ is called $k$-good if $rk M' \leq k$ and $k$-realising if $rk M' = k$. The algorithm presented in the proof of Theorem 1 can be easily modified to calculate the numbers on the main diagonal of a $k$-good matrix if such numbers exist. The algorithm presented in the proof of Theorem 2 can be easily modified to calculate the numbers on the main diagonal of a $k$-realising matrix.

**Remark 4.** There is a related Low-Rank Matrix Completion Problem (LRMC) which has been extensively studied, see survey [NKS].

Consider the following problem which we call $F$-LRMC. Let $F$ be a field. Let $M$ be an $n \times m$ matrix with entries in $F$. Let $\Omega$ be a set of cells of $M$. What is the minimal rank of all the $n \times m$ matrices obtained from $M$ by changing some entries having indices not in $\Omega$?

LRMC is a special case of $F$-LRMC with $F = \mathbb{R}$.

In our paper, a special case of $\mathbb{Z}_2$-LRMC is considered with $m = n$ and $\Omega = \overline{\Omega_{\text{diag}}} := \{(i, j) \mid i, j \in \{1, \ldots, n\}, i \neq j\}$.

In [GKO], $\mathbb{Z}_p$-LRMC is considered. In the case of the main diagonal having entries to change (i.e. in the case of $\Omega = \overline{\Omega_{\text{diag}}}$) the results obtained in [GKO] give algorithms with larger complexity than the algorithm given by Theorem 1.

$\mathbb{Z}_2$-LRMC is NP-hard [Pe96].

In [NKS], LRMC is considered. The methods used in [NKS] give results based on minimizing different approximations of the rank of a matrix, for example, the sum of the singular values of a matrix [SVD]. Thus, none of the results of [NKS] can be applied to the problem of finding $R(M)$.

In [SFH], a problem similar to $\mathbb{Z}_2$-LRMC is considered with the difference being the assumption that $\Omega$ is formed by choosing cells at random independently from each other with the same probability. The provided algorithm is probabilistic.

In [FK19], the following variation of $\mathbb{Z}_2$-LRMC is considered. Take a symmetric matrix $M$ with entries in $\mathbb{Z}_2$ whose rows and columns are indexed by the edges of a connected graph. Let $\Omega$ be equal to the set of pairs of independent edges of the graph. What is the minimal rank of all symmetric matrices obtained from $M$ by changing some entries having indices not in $\Omega$?
Lemma 5. There is an algorithm with the complexity of $O(n^4)$ which for an arbitrary matrix $M \in M_n(\mathbb{Z}_2)$ finds some numbers from $\mathbb{Z}_2$ to replace the entries on the main diagonal of $M$ with such that the resulting matrix $M'$ is non-degenerate.

Proof. The proof consists of two parts: a description and an estimation of the complexity of the algorithm and a proof of that the algorithm gives the required numbers.

Part 1. Description and estimation of the complexity of the algorithm.

Denote by $a_i$ the desired numbers to put on the main diagonal of $M'$. Let us calculate these numbers one by one from top-left to bottom-right. Suppose we have calculated $a_1, \ldots, a_{i-1}$. Then calculate the corner minor $\Delta_i$ with the size $i \times i$ assuming that a zero is on the $i$-th place on the main diagonal of the matrix $M'$. Put $a_i = 1 + \Delta_i$ on the $i$-th place on the diagonal.

There is an algorithm computing the determinant of a $m \times m$ matrix with the complexity of $O(m^3)$ [Ga]. Then, since the presented algorithm is essentially a computation of the determinants of $n$ submatrices with sizes $1 \times 1, 2 \times 2, \ldots, n \times n$, its complexity is $O(1^3 + 2^3 + \ldots + n^3) = O(n^4)$.

Part 2. Proof that the algorithm gives the required numbers.

Let us prove by induction by $i$ that after step $i$ of the first part of the algorithm the corner minor $\Delta'_i$ of $M'$ with the side $i$ is equal to 1.

Base. $i = 1$. $\Delta_1 = 0 \Rightarrow$ we put 1 in the first diagonal element. Hence $\Delta'_1 = 1$.

Step. $i \to i + 1$. By the induction hypothesis $\Delta'_i = 1$. By the decomposition formula of the determinant $\Delta'_{i+1}$ by the last row of the corresponding submatrix of $M'$

$$\Delta'_{i+1} = \Delta_{i+1} + (1 + \Delta_{i+1})\Delta'_i = \Delta_{i+1} + (1 + \Delta_{i+1}) = 1$$

Thus, $\det M' = \Delta'_n = 1$. \qed

A matrix is called diagonal if all its entries outside of the main diagonal are equal to 0.

Lemma 6. Let $M, D$ be matrices of the same size with entries in $\mathbb{Z}_2$ such that $D$ is diagonal. Then $\text{rk}(M + D) \geq \text{rk} M - \text{rk} D$.

Proof. Since $D$ has $\text{rk} D$ nonzero rows, upon addition to $M$ it changes $\text{rk} D$ rows. Since there are $\text{rk} M$ linearly independent rows in the matrix $M$ and at least $\text{rk} M - \text{rk} D$ of them remain the same after the addition of $D$, it
follows that there are $\text{rk} M - \text{rk} D$ linearly independent rows in the matrix $M + D$. Hence $\text{rk}(M + D) \geq \text{rk} M - \text{rk} D$. \qed

Proof of Theorem 4. The proof consists of two parts: a description and an estimation of the complexity of the algorithm and a proof that the algorithm gives the right answer.

Part 1. Description and estimation of the complexity of the algorithm.

Since $k$ is fixed, it is sufficient to estimate the complexity of the algorithm for $n \geq 2k$.

Apply the algorithm given by Lemma 5. Denote by $M'$ the non-degenerate matrix given by the applied algorithm.

There is an algorithm searching through all diagonal $n \times n$ matrices with $l$ zeroes and $n - l$ identities on the main diagonal with the complexity of $O\left(n \binom{n}{l}\right)$. Hence there is an algorithm searching through all diagonal $n \times n$ matrices with $\leq k$ zeroes on the main diagonal with the complexity of

$$O\left(n \binom{n}{0} + n \binom{n}{1} + \ldots + n \binom{n}{k}\right) = O\left(n(k + 1)\binom{n}{\min(n/2, k)}\right) = O\left(n(k + 1)\binom{n}{k}\right) = O\left(n \cdot n^k\right) = O\left(n^{k+1}\right).$$

Apply the algorithm searching through all diagonal $n \times n$ matrices with $\leq k$ zeroes on the main diagonal. For every such matrix $D$ calculate $\text{rk}(M' + D)$. If for at least one such matrix $D$ the rank of $M' + D$ is less than or equal to $k$ then return that it is possible to achieve a rank $\leq k$. Otherwise return that it is not possible to achieve a rank $\leq k$.

For each matrix $M' + D$ where $D$ is a diagonal matrix with $\leq k$ zeroes on the main diagonal the algorithm described above calculates $\text{rk}(M' + D)$. Since the rank of a matrix can be calculated by an algorithm with the complexity of $O(n^3)$, the total complexity of the algorithm described in the previous paragraph is $O(n^{k+1}n^3) = O(n^{k+4})$. Thus, the complexity of the whole algorithm is $O(n^4) + O(n^{k+4}) = O(n^{k+4})$.

Part 2. Proof that the algorithm gives the right answer.

Note that any matrix $M''$ formed as a result of putting numbers from $\mathbb{Z}_2$ in the elements on the main diagonal of $M$ can be uniquely represented as the sum $M' + D$ where $D$ is a diagonal matrix. By Lemma 5 every diagonal
matrix $D$ with $l < n - k$ identities on the main diagonal satisfies

$$\text{rk}(M' + D) \geq \text{rk} M' - \text{rk} D = n - l > n - (n - k) = k.$$ 

Thus, it is sufficient to search through matrices $D$ with at least $n - k$ identities on the main diagonal which is exactly what the algorithm does.

**Lemma 7.** Let $M$, $D$ be matrices of the same size with entries in $\mathbb{Z}_2$ such that $M$ is non-degenerate and $D$ is diagonal. Then $\text{rk}(M + D) \geq \text{rk}(M + I)/2$ where $I$ is the identity matrix.

**Proof.** Denote by $n$ the amount of columns of $M$ and $D$. By Lemma 6 we have

$$2 \text{rk}(M + D) = \text{rk}(M + D) + \text{rk}((M + I) + (I + D))$$
$$\geq (\text{rk} M - \text{rk} D) + (\text{rk}(M + I) - \text{rk}(I + D))$$
$$= n - \text{rk} D + \text{rk}(M + I) - (n - \text{rk} D) = \text{rk}(M + I).$$

**Proof of Theorem 2.** Let us describe the algorithm.

Apply the algorithm given by Lemma 5 to the matrix $M$. Denote by $M'$ the matrix obtained by this algorithm. Denote $k := \text{rk}(M' + I)$. The number $k$ can be computed in time $O(n^3)$.

By Lemma 7 and the fact that $M' + I$ can be obtained from $M$ by changing some numbers on the main diagonal,

$$k/2 \leq R(M) \leq k$$

as required.

The total complexity of the algorithm is $O(n^4) + O(n^3) = O(n^4)$.

**Remark 8.** Let $M$ be an $n \times n$ matrix with entries in $\mathbb{Z}_2$. Then $R(M) \leq n - 1$.

**Proof.** If we change the numbers on the main diagonal of $M$ so that the sum of the entries in each row is even, the resulting matrix $M'$ will obviously be degenerate. Hence $\text{rk} M' \leq n - 1$ and $R(M) \leq n - 1$. 

5
A Appendix

This appendix describes the applications of the main results to graph drawings on non-orientable surfaces (Corollaries A1, A2). The problem of graph drawings on non-orientable surfaces has been extensively studied (for example, see [MT01]). See the definitions of an hieroglyph, weak realizability, the disk with \( k \) Möbius strips below. Our results are different from [Mo89] because there, realizability is studied which comes down to calculating \( r_k M \) (see Theorem A3), and we study weak realizability which comes down to calculating \( R(M) \). The paper [Bi20] gives an algorithmic criterion for the weak realizability of an hieroglyph on the Möbius strip (the disk with 1 Möbius strip).

Let us introduce the notion of weak realizability.

The **disk with \( k \) Möbius strips** is the figure shown on the left of fig. More precisely, the disk with \( k \) Möbius strips is the union of a disk and \( k \) pairwise disjoint ribbons having their ends glued to \( 2k \) pairwise disjoint arcs on the boundary circle of the disk (the ribbons do not have to lie in the plane of the disk) so that

(a) the orientations of the ends of each ribbon given by an orientation of the boundary circle of the disk have “the same direction along the ribbon”,

and

(b) the ribbons are “separated”, i.e. there are \( k \) pairwise disjoint arcs \( A_i \) of the boundary circle of the disk such that the ends of the \( i \)-th ribbon are glued to two disjoint arcs contained in \( A_i \) (\( i = 1, 2, \ldots, k \)).

An **hieroglyph** on \( n \) letters is an unoriented cyclic letter sequence of length \( 2n \) such that each letter from the sequence appears in the sequence twice.

Take a hieroglyph \( H \) on \( n \) letters. Take a convex polygon with \( 2n \) sides. Put the letters in the hieroglyph on the sides of the convex polygon in the nonoriented cyclic order. For each letter glue the ends of a ribbon to the pair of sides corresponding to the letter so that the glued ribbons are pairwise disjoint. Call the resulting surface a **disk with ribbons** corresponding to

---

1Since any nonorientable 2-surface of nonorientable genus \( k \) is homeomorphic to the disk with \( k \) Möbius strips, in this definition the term “disk with \( k \) Möbius strips” can be replaced with “non-orientable 2-surface of nonorientable genus \( k \)".
the hieroglyph $H$ (see fig. 1). Since each ribbon can be either twisted or not twisted, several surfaces may correspond to a single hieroglyph.

A hieroglyph $H$ is **weakly realizable** on the disk with $k$ Möbius strips if some disk with ribbons corresponding to $H$ can be cut out of the disk with $k$ Möbius strips.

For a hieroglyph $H$ denote by $R(H)$ the minimal number $k$ such that the hieroglyph $H$ is weakly realizable on the disk with $k$ Möbius strips. Such a number exists by the classification theorem for compact surfaces with boundary.

**Corollary A1.** For each non-negative integer $k$ there is an algorithm with the complexity of $O(n^{k+4})$ which for an arbitrary hieroglyph $H$ on $n$ letters decides whether $R(H) \leq k$.

**Corollary A2.** There is an algorithm with the complexity of $O(n^4)$ which for an arbitrary hieroglyph $H$ on $n$ letters computes a number $k$ such that

$$k/2 \leq R(H) \leq k.$$

Two letters $a, b$ in a hieroglyph $H$ **overlap in $H$** if they interlace in the cyclic sequence of the hieroglyph (i.e. they appear in the sequence in the order $abab$ but not $aabb$). Take an $n \times n$ matrix with entries in $\mathbb{Z}_2$ with zeroes on the main diagonal. Put 1 in the cell $(i, j)$ for $i \neq j$ if the letters $i, j$
overlap in \( H \) and 0 otherwise. Call the resulting matrix the **overlap matrix** of the hieroglyph \( H \).

**Theorem A3** (Mohar). Let \( M \) be the overlap matrix of a hieroglyph \( H \). Then \( R(H) = R(M) \).

Theorem **A3** is a corollary of [Mo89, Theorem 3.1] (see also [Sk20, §2.8, statement 2.8.8(c)]).

The following Lemma **A4** is well known.

**Lemma A4.** There is an algorithm with the complexity of \( O(n^2) \) which for an arbitrary hieroglyph on \( n \) letters constructs its overlap matrix.

*Proof.* The following is a rough description of the algorithm.

Take an \( n \times n \) matrix \( M \) with zeroes on the main diagonal. For each letter \( i \) do the following. The two occurrences of \( i \) split the hieroglyph into two sequences of letters \( A_i \) and \( B_i \). Go through all the letters in any one sequence \( A_i \) or \( B_i \). For each occurrence of such a letter \( j \) add 1 to the cell \((i, j)\) of \( M \). Return the resulting matrix.

The complexity of the presented algorithm is obviously \( O(n^2) \).

The algorithm gives the required matrix since the letters \( i, j \) overlap if and only if \( j \) appears in the sequence \( A_i \) or \( B_i \) an odd number of times (once).

\[ \square \]

*Proof of Corollary A1.* The following is a description of the algorithm.

Denote by \( M \) the overlap matrix of the hieroglyph \( H \). It can be calculated by the algorithm given by Lemma **A4**. Apply the algorithm given by Theorem **A4** to \( M \). Return the result of the last applied algorithm.

The total complexity of the presented algorithm is \( O(n^2) + O(n^{k+4}) = O(n^{k+4}) \).

The presented algorithm gives the right answer by Theorem **A3**.

\[ \square \]

The *proof of Corollary** A2 can be obtained from the proof of Corollary **A1 by replacing Theorem **A1 and \( O(n^{k+4}) \) by Theorem **A2 and \( O(n^4) \), respectively.

**Remark A5.** For a multigraph \( G \), a **half-edge** of \( G \) is an orientation of one of its edges. For a vertex \( v \) of a multigraph an (oriented or unoriented) **local rotation at** \( v \) is an (oriented or unoriented) cyclic ordering of the half-edges incident to \( v \). A multigraph \( G \) with a family of unoriented local rotations at its vertices is called **weakly realizable** on the disk with \( k \) Möbius strips.
if $G$ can be drawn on the disk with $k$ Möbius strips so that for each vertex $v$ of $G$ the local orientation at $v$ matches the cyclic ordering at which the half-edges incident to $v$ intersect the boundary of a small circle around $v$. A natural generalization of finding $R(H)$ for a hieroglyph $H$ is finding the least integer $k$ such that a given multigraph with a family of unoriented local rotations at its vertices is weakly realizable on the disk with $k$ Möbius strips. However, this problem cannot be reduced to the problem of finding $R(H)$ by contracting spanning trees of the components of the multigraph because upon contracting an edge $(v, u)$, it is unclear how to combine the unoriented local orientations at $v$ and $u$.

References

[Bi20] A. Bikeev. The realizability of discs with ribbons on a Möbius strip (in Russian). https://arxiv.org/abs/2012.12070

[FK19] R. Fulek, J. Kynčl. $\mathbb{Z}_2$-genus of graphs and minimum rank of partial symmetric matrices. https://arxiv.org/abs/1903.08637

[Ga] https://en.wikipedia.org/wiki/Gaussian_elimination

[GKO+] R. Ganian, I. Kanj, S. Ordyniak, S. Szeider. Parameterized algorithms for the matrix completion problem. https://arxiv.org/abs/1804.03423

[Mo89] B. Mohar. An obstruction to embedding graphs in surfaces. Discrete Math. 1989. V. 78. Pp. 135-142.

[MT01] B. Mohar, C. Thomassen. Graphs on Surfaces. The Johns Hopkins University Press. 2001.

[NKS] L. T. Nguyen, J. Kim, B. Shim. Low-rank matrix completion: a contemporary survey. https://arxiv.org/abs/1907.11705

[Pe96] R. Peeters. Orthogonal representations over finite fields and the chromatic number of graphs. Combinatorica. 1996. V. 16(3). Pp. 417-431.

[SFH] J. Saunderson, M. Fazel, B. Hassibi. Simple algorithms and guarantees for low rank matrix completion over $F_2$. In Proceedings of
The IEEE International Symposium on Information Theory, pp. 86-90. IEEE, 2016.

[Sk20] A. Skopenkov. Algebraic Topology from Geometric Viewpoint (in Russian). MCCME. Moscow. 2020 (2nd edition). http://www.mccme.ru/circles/oim/obstruct.pdf

[SVD] https://en.wikipedia.org/wiki/Singular_value_decomposition