Electronic Colloquium on Computational Complexity, Report No. 8 (2009)

Min-Rank Conjecture for Log-Depth Circuits *

Stasys Jukna†‡ Georg Schnitger§

January 15, 2009

Abstract

A completion of an m-by-n matrix A with entries in $\{0, 1, *\}$ is obtained by setting all $*$-entries to constants 0 or 1. A system of semi-linear equations over GF_2 has the form $M\boldsymbol{x} = f(\boldsymbol{x})$, where M is a completion of A and $f : \{0, 1\}^n \to \{0, 1\}^m$ is an operator, the ith coordinate of which can only depend on variables corresponding to $*$-entries in the ith row of A. We conjecture that no such system can have more than $2^{n-\epsilon\cdot \mathrm{mr}(A)}$ solutions, where $\epsilon > 0$ is an absolute constant and $\mathrm{mr}(A)$ is the smallest rank over GF_2 of a completion of A. The conjecture is related to an old problem of proving super-linear lower bounds on the size of log-depth boolean circuits computing linear operators $\boldsymbol{y} = M\boldsymbol{x}$. The conjecture is also a generalization of a classical question about how much larger can non-linear codes be than linear ones. We prove some special cases of the conjecture and establish some structural properties of solution sets.

Keywords: Boolean circuits; Partial matrix; Matrix completion; Min-rank; Matrix rigidity; Sum-Sets; Cayley graphs; Error-Correcting Codes

1 Introduction

One of the challenges in circuit complexity is to prove a super-linear lower bound for log-depth circuits over $\{\&, \vee, \neg\}$ computing an explicitly given boolean operator $f : \{0, 1\}^n \to \{0, 1\}^n$. Attempts to solve it have led to several weaker problems which are often of independent interest. The problem is open even if we impose an additional restriction that the depth of the circuit is $O(\log n)$. It is even open for *linear* log-depth circuits, that is for log-depth circuits over the basis $\{\oplus, 1\}$, in spite of the apparent simplicity of such circuits. It is clear that the operators computed by linear circuits must also be linear, that is, be matrix-vector products $\boldsymbol{x} \to M\boldsymbol{x}$ over the field $GF_2 = (\{0, 1\}, \oplus, \cdot)$,

An important result of Valiant [27] reduces the lower bounds problem for log-depth circuits over $\{\&, \vee, \neg\}$ to proving lower bounds for certain depth-2 circuits, where we allow *arbitrary* boolean functions as gates.

*Research of both authors supported by a DFG grant SCHN 503/4-1.

†Institute of Mathematics and Computer Science, Akademijos 4, LT-80663 Vilnius, Lithuania.

‡Corresponding author. *Current address* is that of the second author. E-mail: jukna@thi.informatik.uni-frankfurt.de. Fax: +49 69 798-28814

§University of Frankfurt, Institut of Computer Science, Robert-Mayer-Str. 11-5, D-60054 Frankfurt, Germany. E-mail: georg@thi.informatik.uni-frankfurt.de

ISSN 1433-8092

1.1 Reduction to depth-2 circuits

A depth-2 circuit of *width* w has n boolean variables $x_1, \ldots, x_n$ as input nodes, w arbitrary boolean functions $h_1, \ldots, h_w$ as gates on the middle layer, and m arbitrary boolean functions $g_1, \ldots, g_m$ as gates on the output layer. Direct input-output wires, connecting input variables with output gates, are allowed. Such a circuit computes an operator $f = (f_1, \ldots, f_m) : \{0,1\}^n \to \{0,1\}^m$ if, for every $i = 1, \ldots, m$,

$$f_i(\boldsymbol{x}) = g_i(\boldsymbol{x}, h_1(\boldsymbol{x}), \ldots, h_w(\boldsymbol{x})).$$

The *degree* of such a circuit is the maximum, over all output gates g_i, of the number of wires going directly from input variables $x_1, \ldots, x_n$ to the gate g_i. That is, we ignore the wires incident with the gates on the middle layer. Let $\deg_w(f)$ denote the smallest degree of a depth-2 circuit of width w computing f.

It is clear that $\deg_n(f) = 0$: just put the functions $f_1, \ldots, f_n$ on the middle layer. Hence, this parameter is only nontrivial for $w < n$. Especially interesting is the case when $w = O(n/\ln\ln n)$:

Lemma 1 (Valiant [27])**.** *If* $\deg_w(f) = n^{\Omega(1)}$ *for* $w = O(n/\ln\ln n)$*, then the operator* f *cannot be computed by a circuit of depth* $O(\ln n)$ *using* $O(n)$ *constant fanin gates.*

Recently, there was a substantial progress in proving lower bounds on the *size* of (that is, on the total number of wires in) depth-2 circuits. Superlinear lower bounds of the form n times poly-log(n) where proved using graph-theoretic arguments by analyzing some superconcentration properties of the circuit as a graph [6, 14, 15, 18, 16, 2, 20, 21, 22]. Higher lower bounds of the form $\Omega(n^{3/2})$ were proved using information theoretical arguments [4, 9]. But the highest known lower bound on the *degree* of width w circuits has the form $\Omega((n/w)\ln(n/w))$ [20], and is too weak to have a consequence for log-depth circuits.

A natural question therefore was to improve the lower bound on the degree at least for *linear* circuits, that is, for depth-2 circuits whose middle gates as well as output gates are linear boolean functions (parities of their inputs). Such circuits compute linear operators $f_M(\boldsymbol{x}) = M\boldsymbol{x}$ for some $(0,1)$-matrix; we work over GF_2. By Valiant's reduction, this would give a super-linear lower bound for log-depth circuits over $\{\oplus, 1\}$.

This last question attracted attention of many researchers because of its relation to a purely algebraic characteristic of the underlying matrix M—its rigidity. The *rigidity* $\mathcal{R}_M(r)$ of a $(0,1)$-matrix M is the smallest number of entries of M that must be changed in order to reduce its rank over GF_2 until r. It is not difficult to show (see [27]) that any linear depth-2 circuit of width w computing $M\boldsymbol{x}$ must have degree at least $\mathcal{R}_M(w)/n$: If we set all direct input-output wires to 0, then the resulting degree-0 circuit will compute some linear transformation $M'\boldsymbol{x}$ where the rank of M' does not exceed the width w. On the other hand, M' differs from M in at most dn entries, where d is the degree of the original circuit. Hence, $\mathcal{R}_M(w) \le dn$ from which $d \ge \mathcal{R}_M(w)/n$ follows.

Motivated by its connection to proving lower bounds for log-depth circuits, matrix rigidity (over different fields) was considered by many authors, [23, 1, 17, 7, 16, 20, 25, 24, 10, 11, 19, 26] among others. It is therefore somewhat surprising that the highest known lower bounds on $\mathcal{R}_M(r)$ (over the field GF_2), proved in [7, 25] also have the form $\Omega((n^2/r)\ln(n/r))$, resulting to the same lower bound $\Omega((n/w)\ln(n/w))$ on the degree of linear circuits as that for general depth-2 circuits proved in [20]. This phenomenon is particularly surprising, because general circuits may use *arbitrary* (not just linear) boolean functions as gates. We suspect that the absence of higher lower bounds for linear circuits than those for non-linear ones could be not just a coincidence.

Conjecture 1 (Linearization conjecture for depth-2 circuits)**.** *Depth-*2 *circuits can be linearized. That is, every depth-*2 *circuit computing a linear operator can be transformed into an equivalent* linear *depth-*2 *circuit without substantial increase of its width or its degree.*

If true, the conjecture would have important consequences for log-depth circuits. Assuming this conjecture, any proof that every depth-2 circuit of width $w = O(n/\ln\ln n)$ with unbounded fanin parity gates for a given linear operator $M\boldsymbol{x}$ requires degree $n^{\Omega(1)}$ would imply that $M\boldsymbol{x}$ requires a super-linear number of gates in any log-depth circuit over $\{\&, \vee, \neg\}$. In particular, this would mean that proving high lower bounds on matrix rigidity is a much more difficult task than assumed before: such bounds would yield super-linear lower bounds for log-depth circuits over a general basis $\{\&, \vee, \neg\}$, not just for circuits over $\{\oplus, 1\}$.

As the first step towards Conjecture 1, in this paper we relate it to a purely combinatorial conjecture about partially defined matrices—the *min-rank conjecture*, and prove some results supporting this last conjecture. This turns the problem about the linearization of depth-2 circuits into a problem of Combinatorial Matrix Theory concerned with properties of completions of partially defined matrices (see, e.g., the survey [8]). Hence, the conjecture may also be of independent interest.

Unfortunately, we were not able to prove the conjecture in its full generality. So far, we are only able to prove that some its special cases are true. This is not very surprising because the conjecture touches a basic problem in circuit complexity: Can non-linear gates help to compute linear operators? This paper is just the first step towards this question.

1.2 The Min-Rank Conjecture

A *completion* of a $(0, 1, *)$-matrix A is a $(0, 1)$-matrix M obtained from A by setting all $*$'s to constants 0 and 1. A *canonical completion* of A is obtained by setting all $*$'s in A to 0.

If A is an m-by-n matrix, then each its completion M defines a linear operator mapping each vector $\boldsymbol{x} \in \{0, 1\}^n$ to a vector $M\boldsymbol{x} \in \{0, 1\}^m$. Besides such (linear) operators we also consider general ones. Each operator $G : \{0, 1\}^n \to \{0, 1\}^m$ can be looked at as a sequence $G = (g_1, \ldots, g_m)$ of m boolean functions $g_i : \{0, 1\}^n \to \{0, 1\}$.

We say that an operator G is *consistent* with an m-by-n $(0, 1, *)$-matrix $A = (a_{ij})$ if the ith boolean function g_i can only depend on those variables x_j for which $a_{ij} = *$. That is, the ith component g_i of G can only depend on variables on which the ith row of A has stars.

Definition 1. With some abuse in notation, we call a set $L \subseteq \{0, 1\}^n$ a *solution* for a partial matrix A if there is a completion M of A and an operator G such that G is consistent with A and $M\boldsymbol{x} = G(\boldsymbol{x})$ holds for all $\boldsymbol{x} \in L$. A solution L is *linear* if it forms a linear subspace of $\{0, 1\}^n$ over GF_2.

We are interested in how much the maximum $\mathrm{opt}(A) = \max_L |L|$ over all solutions L for A can exceed the maximum $\mathrm{lin}(A) = \max_L |L|$ over all linear solutions L for A. It can be shown (Corollary 4 below) that

$$\mathrm{lin}(A) = 2^{n-\mathrm{mr}(A)},$$

where $\mathrm{mr}(A)$ is the *min-rank* of A defined as the smallest possible rank of its completion:

$$\mathrm{mr}(A) = \min\{\mathrm{rank}(M) \colon M \text{ is a completion of } A\}.$$

If we only consider *constant* operators G, that is, operators with $G(\boldsymbol{x}) = \boldsymbol{b}$ for some $\boldsymbol{b} \in \{0,1\}^m$ and all $\boldsymbol{x} \in \{0,1\}^n$, then Linear Algebra tells us that no solution for A can have more than 2^{n-r} vectors, where $r = \mathrm{rank}(M)$ is the rank (over GF_2) of the canonical completion M of A, obtained by setting all stars to 0.

If we only consider *affine* operators G, that is, operators of the form $G(\boldsymbol{x}) = H\boldsymbol{x} \oplus \boldsymbol{b}$ where H is an m-by-n $(0,1)$-matrix, then no solution for A can have more than $2^{n-\mathrm{mr}(A)}$ vectors, because then the consistency of $G(\boldsymbol{x})$ with A ensures that, for every completion M of A, the matrix $M \oplus H$ is a completion of A.

Remark 1. This last observation implies, in particular, that $\mathrm{opt}(A) \leq 2^{n-\mathrm{mr}(A)}$ for all $(0,1,*)$-matrices A with at most one $*$ in each row: In this case each g_i can depend on at most one variable, and hence, must be a linear boolean function.

We conjecture that a similar upper bound also holds for *any* operator G, as long as it is consistent with A. That is, we conjecture that linear operators are almost optimal.

Conjecture 2 (Min-Rank Conjecture). *There exists a constant $\epsilon > 0$ such that for every m-by-n $(0,1,*)$-matrix A we have that $\mathrm{opt}(A) \leq 2^{n-\epsilon\cdot\mathrm{mr}(A)}$ or, equivalently,*

$$\mathrm{opt}(A) \leq 2^n \left(\frac{\mathrm{lin}(A)}{2^n}\right)^{\epsilon}. \tag{1}$$

Remark 2. Valiant [27] reduces log-depth circuits of linear size to depth-2 circuits of width $O(n/\log\log n)$. Hence, to have consequences for log-depth circuits, it would be enough that the conjecture holds at least for $\epsilon = o(1/\log\log n)$.

To illustrate the introduced concepts, let us consider the following system of 3 equations in 6 variables:

$$\begin{aligned} x_1 \oplus x_6 &= x_3 \cdot x_5 \\ x_2 \oplus x_3 \oplus x_4 &= x_1 \cdot (x_5 \oplus x_6) \\ x_4 &= (x_2 \oplus x_5) \cdot (x_3 \oplus x_6)\,. \end{aligned} \tag{2}$$

The corresponding $(0,1,*)$-matrix for this system is

$$A = \begin{pmatrix} 1 & 0 & * & 0 & * & 1 \\ * & 1 & 1 & 1 & * & * \\ 0 & * & * & 1 & * & * \end{pmatrix}, \tag{3}$$

and the system itself has the form $M\boldsymbol{x} = G(\boldsymbol{x})$, where M is the canonical completion of A:

$$M = \begin{pmatrix} 1 & 0 & \underline{0} & 0 & \underline{0} & 1 \\ \underline{0} & 1 & 1 & 1 & \underline{0} & \underline{0} \\ 0 & \underline{0} & \underline{0} & 1 & \underline{0} & \underline{0} \end{pmatrix},$$

and $G = (g_1, g_2, g_3) : \{0,1\}^6 \to \{0,1\}^3$ is an operator with

$$\begin{aligned} g_1(\boldsymbol{x}) &= x_3 \cdot x_5\,; \\ g_2(\boldsymbol{x}) &= x_1 \cdot (x_5 \oplus x_6)\,; \\ g_3(\boldsymbol{x}) &= (x_2 \oplus x_5) \cdot (x_3 \oplus x_6)\,. \end{aligned}$$

The min-rank of A is equal 2, and is achieved by the following its completion:

$$M' = \begin{pmatrix} 1 & 0 & \underline{0} & 0 & \underline{0} & 1 \\ \underline{0} & 1 & 1 & 1 & \underline{0} & \underline{0} \\ 0 & \underline{1} & \underline{1} & 1 & \underline{0} & \underline{0} \end{pmatrix} .$$

Remark 3 (Canonical completions)**.** Recall that solutions for a given $(0,1,*)$-matrix A are subsets $L \subseteq \{0,1\}^n$ of vectors such that $M\boldsymbol{x} = G(\boldsymbol{x})$ for some completion M of A and some operator G consistent with A. Since, besides the consistency, there are no other restrictions on G, we can always assume that M is the canonical completion of A (with all stars set to 0). Indeed, if $L' = \{\boldsymbol{x} \colon M'\boldsymbol{x} = G'(\boldsymbol{x})\}$ is some solution for A, then the ith row $\boldsymbol{m}_i'$ of M' must have the form $\boldsymbol{m}_i' = \boldsymbol{m}_i \oplus \boldsymbol{p}_i$, where $\boldsymbol{m}_i \in \{0,1\}^n$ is the ith row of the canonical completion M of A, and $\boldsymbol{p}_i \in \{0,1\}^n$ is a vector with no 1's in positions where the ith row of A has no stars. We can then define an operator $G = (g_1, \ldots, g_m)$ by [1] $g_i(\boldsymbol{x}) := g_i'(\boldsymbol{x}) \oplus \langle \boldsymbol{p}_i, \boldsymbol{x} \rangle$. Since G' was consistent with A, the new operator G is also consistent with A. Moreover, for every vector $\boldsymbol{x} \in \{0,1\}^n$, we have that $\langle \boldsymbol{m}_i', \boldsymbol{x} \rangle = g_i'(\boldsymbol{x})$ iff $\langle \boldsymbol{m}_i, \boldsymbol{x} \rangle = g_i(\boldsymbol{x})$.

1.3 Our results

In Section 2 we prove the main consequence of the min-rank conjecture for boolean circuits: If true, it would imply that non-linear gates are powerless when computing linear operators $M\boldsymbol{x}$ by depth-2 circuits (Lemmas 2 and 3).

In Sections 3 and 4 we prove some partial results supporting Conjectures 1 and 2. We first show (Corollary 2) that every depth-2 circuit of width w computing a linear operator can be transformed into an equivalent *linear* depth-2 circuit of the same degree and width at most w plus the maximum number of wires in a matching formed by the input-output wires of the original circuit.

We then prove two special cases of Min-Rank Conjecture. A set of $(0,1,*)$-vectors is *independent* if they cannot be made linearly dependent over GF_2 by setting stars to constants 0 and 1. If A is a $(0,1,*)$-matrix, then the upper bound $\mathrm{opt}(A) \leq 2^{n-r}$ holds if the matrix A contains r independent columns (Theorem 1). The same upper bound also holds if A contains r independent rows, and the sets of star positions in these rows form a chain with respect to set-inclusion (Theorem 2).

After that we concentrate on the *structure* of solutions. In Section 5 we show that solutions for a $(0,1,*)$-matrix A are precisely independent sets in a Cayley graph over the Abelian group $(\{0,1\}^n, \oplus)$ generated by a special set $K_A \subseteq \{0,1\}^n$ of vectors defined by the matrix A (Theorem 3).

In Section 6 we first show that every linear solution for A lies in the kernel of some completion of A (Theorem 4). This, in particular, implies that $\mathrm{lin}(A) = 2^{n-\mathrm{mr}(A)}$ (Corollary 4), and gives an alternative definition of the min-rank $\mathrm{mr}(A)$ as the smallest rank of a boolean matrix H such that $H\boldsymbol{x} \neq \boldsymbol{0}$ for all $\boldsymbol{x} \in K_A$ (Corollary 5). In Section 7 we show that non-linear solutions must be "very non-linear": they cannot contain linear subspaces of dimension exceeding the maximum number of $*$'s in a row of A (Theorem 5).

In Section 8 we consider the relation of the min-rank conjecture with error-correcting codes. We define $(0,1,*)$-matrices A, the solutions for which are error-correcting codes, and show that the min-rank conjecture for these matrices is true: In this case the conjecture

[1] As customary, the scalar product of two vectors $\boldsymbol{x}, \boldsymbol{y} \in \{0,1\}^n$ over GF_2 is $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^n x_i y_i \pmod 2$.

is implied by well known lower and upper bounds on the size of linear and nonlinear error correcting codes (Lemma 9).

2 Min-rank conjecture and depth-2 circuits

Let F be a depth-2 circuit computing a linear operator $\boldsymbol{x} \to M\boldsymbol{x}$, where M is an m-by-n $(0,1)$-matrix. Say that the (i,j)th entry of M is *seen* by the circuit, if there is a direct wire from x_j to the ith output gate. Replace all entries of M seen by the circuit with $*$'s, and let A_F be the resulting $(0,1,*)$-matrix. Note that the original matrix M is one of the completions of A_F; hence, $\mathrm{rank}(M) \geq \mathrm{mr}(A_F)$.

Proposition 1. *Every linear depth-2 circuit F has* $\mathrm{width}(F) \geq \mathrm{mr}(A_F)$.

Proof. Let $M\boldsymbol{x}$ be a linear operator computed by F. Every assignment of constants to direct input-output wires leads to a depth-2 circuit of degree $d = 0$ computing a linear operator $B\boldsymbol{x}$, where B is a completion of A_F. This operator takes $2^{\mathrm{rank}(B)}$ different values. Hence, the operator $H : \{0,1\}^n \to \{0,1\}^w$ computed by $w = \mathrm{width}(F)$ boolean functions on the middle layer of F must take at least so many different values, as well. This implies that the width w must be large enough to fulfill $2^w \geq 2^{\mathrm{rank}(B)}$, from which $w \geq \mathrm{rank}(B) \geq \mathrm{mr}(A_F)$ follows. □

Lemma 2. *Every depth-2 circuit F computing a linear operator can be transformed into an equivalent linear depth-2 circuit of the same degree and width at most* $\mathrm{mr}(A_F)$.

Together with Proposition 1, this implies that $\mathrm{width}(F) = \mathrm{mr}(A_F)$ for every optimal linear depth-2 circuit F.

Proof. Let $\boldsymbol{x} \to M\boldsymbol{x}$ be the operator computed by F, and let $A = A_F$ be the $(0,1,*)$-matrix of F. We can construct the desired *linear* depth-2 circuit computing $M\boldsymbol{x}$ as follows. Take a completion B of A with $\mathrm{rank}(B) = \mathrm{mr}(A)$. By the definition of completions, the ith row $\boldsymbol{b}_i$ of B has the form $\boldsymbol{b}_i = \boldsymbol{a}_i \oplus \boldsymbol{p}_i$, where $\boldsymbol{a}_i$ is the ith row of A with all stars set to 0, and $\boldsymbol{p}_i$ is a $(0,1)$-vector having no 1's in positions, where this row of A has non-stars. The ith row $\boldsymbol{m}_i$ of the original $(0,1)$-matrix M is of the form $\boldsymbol{m}_i = \boldsymbol{a}_i \oplus \boldsymbol{m}_i'$, where $\boldsymbol{m}_i'$ is a $(0,1)$-vector which coincides with $\boldsymbol{m}_i$ in all positions, where the ith row of A has stars, and has 0's elsewhere.

Take any $r = \mathrm{mr}(A)$ linearly independent rows $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_r$ of B, and add r linear gates computing the scalar products $\langle \boldsymbol{b}_1, \boldsymbol{x}\rangle, \ldots, \langle \boldsymbol{b}_r, \boldsymbol{x}\rangle$ over GF_2 on the middle layer. Join these linear gates with all input and all output gates. Note that the ith output gate can compute both scalar products $\langle \boldsymbol{p}_i, \boldsymbol{x}\rangle$ and $\langle \boldsymbol{m}_i', \boldsymbol{x}\rangle$ by only using existing direct wires from inputs $x_1, \ldots, x_n$ to this gate. Hence, using the r linear gates $\langle \boldsymbol{b}_1, \boldsymbol{x}\rangle, \ldots, \langle \boldsymbol{b}_r, \boldsymbol{x}\rangle$ on the middle layer, the ith output gate can also compute the whole scalar product

$$\langle \boldsymbol{m}_i, \boldsymbol{x}\rangle = \langle \boldsymbol{a}_i, \boldsymbol{x}\rangle \oplus \langle \boldsymbol{m}_i', \boldsymbol{x}\rangle = \langle \boldsymbol{b}_i, \boldsymbol{x}\rangle \oplus \langle \boldsymbol{p}_i, \boldsymbol{x}\rangle \oplus \langle \boldsymbol{m}_i', \boldsymbol{x}\rangle .$$

We have thus constructed an equivalent linear depth-2 of the same degree and of width $r = \mathrm{mr}(A_F)$. □

By Lemma 2, the main question is: How much the width of a circuit F can be smaller than the min-rank of its matrix A_F? Ideally, we would like to have that $\mathrm{width}(F) \geq \epsilon \cdot \mathrm{mr}(A_F)$: then the width of the resulting *linear* circuit would be at most $1/\epsilon$ times larger than that of the original circuit F.

In F has no direct input-output wires, then $A_F = M$, and we have that

$$\mathrm{width}(F) \geq \mathrm{rank}(M)\,. \tag{4}$$

The argument is the same as in the proof of Proposition 1. Since the operator $M\boldsymbol{x}$ takes $2^{\mathrm{rank}(M)}$ different values, the operator $H : \{0,1\}^n \to \{0,1\}^w$ computed by $w = \mathrm{width}(F)$ boolean functions on the middle layer of F must take at least so many different values, as well; hence, the width w must be large enough to fulfill $2^w \geq 2^{\mathrm{rank}(M)}$.

Lemma 3. *For every depth-2 circuit F computing a linear operator in n variables, we have that*

$$\mathrm{width}(F) \geq n - \log_2 \mathrm{opt}(A_F)\,.$$

Hence, the Min-Rank Conjecture implies that $\mathrm{width}(F) \geq \epsilon \cdot \mathrm{mr}(A_F)$.

Proof. Let M be an m-by-n $(0,1)$-matrix. Take a depth-2 circuit F of width w computing $M\boldsymbol{x}$, and let $A = A_F$ be the corresponding $(0,1,*)$-matrix. Let $H = (h_1, \ldots, h_w)$ be an operator computed at the gates on the middle layer, and $G = (g_1, \ldots, g_m)$ an operator computed at the gates on the output layer. Hence,

$$M\boldsymbol{x} = G(\boldsymbol{x}, H(\boldsymbol{x})) \qquad \text{for all } \boldsymbol{x} \in \{0,1\}^n.$$

Fix a vector $\boldsymbol{b} \in \{0,1\}^w$ for which the set

$$L = \{\boldsymbol{x} \in \{0,1\}^n : M\boldsymbol{x} = G(\boldsymbol{x}, \boldsymbol{b})\}$$

is the largest one; hence, $|L| \geq 2^{n-w}$. Note that the operator $G'(\boldsymbol{x}) := G(\boldsymbol{x}, \boldsymbol{b})$ must be consistent with A: its ith component $g'_i(\boldsymbol{x})$ can only depend on input variables x_j to which the ith output gate g_i is connected. Hence, L is a solution for A, implying that $\mathrm{opt}(A) \geq |L| \geq 2^{n-w}$ from which the desired lower bound $w \geq n - \log_2 \mathrm{opt}(A)$ on the width of F follows. □

Corollary 1. *Conjecture 2 implies Conjecture 1.*

Proof. Let F be a depth-2 circuit computing a linear operator in n variables. Assuming Conjecture 2, Lemma 3 implies that $\epsilon \cdot \mathrm{mr}(A_F) \leq n - \log_2 \mathrm{opt}(A_F) \leq \mathrm{width}(F)$. By Lemma 2, the circuit F can be transformed into an equivalent linear depth-2 circuit of the same degree and width at most $\mathrm{mr}(A_F) \leq \mathrm{width}(F)/\epsilon$. □

Hence, together with Valiant's result, the min-rank conjecture implies that a linear operator $M\boldsymbol{x}$ requires a super-linear number of gates in any log-depth circuit over $\{\&, \vee, \neg\}$, if every depth-2 circuit for $M\boldsymbol{x}$ over $\{\oplus, 1\}$ of width $w = O(n/\ln\ln n)$ requires degree $n^{\Omega(1)}$.

Finally, let us show that the only "sorrow", when trying to linearize a depth-2 circuit, is the possible non-linearity of *output* gates—non-linearity of gates on the middle layer is no problem.

Lemma 4. *Let F be a depth-2 circuit computing a linear operator. If all gates on the output layer are linear boolean functions, then F can be transformed into an equivalent linear depth-2 circuit of the same degree and width.*

Proof. Let M be an m-by-n $(0,1)$-matrix, and let F be a depth-2 circuit of width w computing $M\boldsymbol{x}$. Let $H = (h_1, \ldots, h_w)$ be the operator $H : \{0,1\}^n \to \{0,1\}^w$ computed by the gates on the middle layer. Assume that all output gates of F are linear boolean functions. Let B be the m-by-n adjacency $(0,1)$-matrix of the bipartite graph formed by the direct input-output wires, and C be the m-by-w adjacency $(0,1)$-matrix of the bipartite graph formed by the wires joining the gates on the middle layer with those on the output layer. Then

$$M\boldsymbol{x} = B\boldsymbol{x} \oplus C \cdot H(\boldsymbol{x}) \qquad \text{for all } \boldsymbol{x} \in \{0,1\}^n,$$

where $C \cdot H(\boldsymbol{x})$ is the product of the matrix C with the vector $\boldsymbol{y} = H(\boldsymbol{x})$. Hence,

$$C \cdot H(\boldsymbol{x}) = D\boldsymbol{x} \tag{5}$$

is a linear operator with $D = M \oplus B$. Write each vector $\boldsymbol{x} = (x_1, \ldots, x_n)$ as the linear combination

$$\boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{e}_i \tag{6}$$

of unit vectors $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n \in \{0,1\}^n$, and replace the operator H computed on the middle layer by a *linear* operator

$$H'(\boldsymbol{x}) := \sum_{i=1}^{n} x_i H(\boldsymbol{e}_i) \pmod 2\,. \tag{7}$$

Then, using the linearity of the matrix-vector product, we obtain that (with all sums mod 2):

$$\begin{aligned}
C \cdot H(\boldsymbol{x}) &= D \cdot \Big(\sum x_i \boldsymbol{e}_i\Big) && \text{by (5) and (6)} \\
&= \sum x_i D\boldsymbol{e}_i && \text{linearity} \\
&= \sum x_i C \cdot H(\boldsymbol{e}_i) && \text{by (5)} \\
&= C \cdot \Big(\sum x_i H(\boldsymbol{e}_i)\Big) && \text{linearity} \\
&= C \cdot H'\Big(\sum x_i \boldsymbol{e}_i\Big) && \text{by (7)} \\
&= C \cdot H'(\boldsymbol{x}) && \text{by (6)}\,.
\end{aligned}$$

Hence, we again have that $M\boldsymbol{x} = B\boldsymbol{x} \oplus C \cdot H'(\boldsymbol{x})$, meaning that the obtained *linear* circuit computes the same linear operator $M\boldsymbol{x}$. □

3 Bounds on opt(A)

Recall that $\mathrm{opt}(A)$ is the largest possible number of vectors in a solution for a given $(0,1,*)$-matrix A. The simplest properties of this parameter are summarized in the following proposition.

Proposition 2. *Let A be an m-by-n $(0,1,*)$-matrix. If A' is obtained by removing some rows of A, then $\mathrm{opt}(A') \geq \mathrm{opt}(A)$. If $A = [B, C]$ where B is an m-by-p submatrix of A for some $1 \leq p \leq n$, then*

$$\mathrm{opt}(B) \cdot \mathrm{opt}(C) \leq \mathrm{opt}(A) \leq \mathrm{opt}(B) \cdot 2^{n-p}\,.$$

Proof. The first claim $\mathrm{opt}(A') \leq \mathrm{opt}(A)$ is obvious, since addition of new equations can only decrease the number of solutions in any system of equations.

To prove $\mathrm{opt}(A) \leq \mathrm{opt}(B) \cdot 2^{n-q}$, take an optimal solution $L_A = \{\boldsymbol{x} : M\boldsymbol{x} = G(\boldsymbol{x})\}$ for A; hence, $|L_A| = \mathrm{opt}(A)$. Fix a vector $\boldsymbol{b} \in \{0,1\}^{n-p}$ for which the set

$$L_B = \{\boldsymbol{y} \in \{0,1\}^p : (\boldsymbol{y}, \boldsymbol{b}) \in L_A\}$$

is the largest one; hence, $|L_B| \geq \mathrm{opt}(A)/2^{n-p}$. The completion M of A has the form $M = [M', M'']$, where M' is a completion of B and M'' is a completion of C. If we define an operator $G' : \{0,1\}^p \to \{0,1\}^m$ by

$$G'(\boldsymbol{y}) := G(\boldsymbol{y}, \boldsymbol{b}) \oplus M''\boldsymbol{b},$$

then $M'\boldsymbol{y} = G'(\boldsymbol{y})$ for all $\boldsymbol{y} \in L_B$. Hence, L_B is a solution for B, implying that $\mathrm{opt}(A) \leq |L_B| \cdot 2^{n-p} \leq \mathrm{opt}(B) \cdot 2^{n-p}$.

To prove $\mathrm{opt}(A) \geq \mathrm{opt}(B) \cdot \mathrm{opt}(C)$, let $L_B = \{\boldsymbol{y} \in \{0,1\}^p : M'\boldsymbol{y} = G'(\boldsymbol{y})\}$ be an optimal solution for B, and let $L_C = \{\boldsymbol{z} \in \{0,1\}^{n-p} : M''\boldsymbol{z} = G''(\boldsymbol{z})\}$ be an optimal solution for C. For any pair $\boldsymbol{x} = (\boldsymbol{y}, \boldsymbol{z}) \in L_B \times L_C$, we have that $M\boldsymbol{x} = G(\boldsymbol{x})$, where $M = [M', M'']$ and $G(\boldsymbol{y}, \boldsymbol{z}) := G'(\boldsymbol{y}) \oplus G''(\boldsymbol{z})$. Hence, the set $L_B \times L_C \subseteq \{0,1\}^n$ is a solution for A, implying that $\mathrm{opt}(B) \cdot \mathrm{opt}(C) = |L_B \times L_C| \leq \mathrm{opt}(A)$, as claimed. □

Let A be an m-by-n $(0,1,*)$-matrix. The min-rank conjecture claims that the largest number $\mathrm{opt}(A)$ of vectors in a solution for A can be upper bounded in terms of the min-rank of A as $\mathrm{opt}(A) \leq 2^{n-\epsilon\cdot\mathrm{mr}(A)}$. The claim is obviously true if the min-rank of A is "witnessed" by some its $(0,1)$-submatrix.

Proposition 3. *If A is an m-by-n $(0,1,*)$-matrix, then $\mathrm{opt}(A) \leq 2^{n-\mathrm{rank}(B)}$ for every $(0,1)$-submatrix B of A.*

Proof. Let B be a p-by-q $(0,1)$-submatrix of A. Since B has no stars, only constant operators can be consistent with B. Hence, if $L \subseteq \{0,1\}^q$ is a solution for B, then there must be a vector $\boldsymbol{b} \in \{0,1\}^p$ such that $B\boldsymbol{x} = \boldsymbol{b}$ for all $\boldsymbol{x} \in L$. This implies $|L| \leq 2^{q-\mathrm{rank}(B)}$. Together with Proposition 2, this yields $\mathrm{opt}(A) \leq 2^{q-\mathrm{rank}(B)} \cdot 2^{n-q} = 2^{n-\mathrm{rank}(B)}$. □

The *max-rank* $\mathrm{Mr}(A)$ of a $(0,1,*)$-matrix A is a maximal possible rank of its completion. A cover of A is a set X of its lines (rows and columns) covering all stars. Let $\mathrm{cov}(A)$ denotes the smallest possible number of lines in a cover of A.

Lemma 5. *For every m-by-n $(0,1,*)$-matrix A, we have that*

$$\mathrm{opt}(A) \leq 2^{n-\mathrm{Mr}(A)+\mathrm{cov}(A)}.$$

Proof. Given a cover X of the stars in A by lines, remove all these lines, and let A_X be the resulting $(0,1)$-submatrix of A. It is known (see [5]) that

$$\mathrm{Mr}(A) = \min_X \mathrm{rank}(A_X) + |X|,$$

where the minimum is over all covers X of A. (For $(0,*)$-matrices A this equality is the well-known Frobenius–König theorem stating that $\mathrm{cov}(A)$ is the largest number of $*$-entries of A, no two on the same line). Fix a set X of lines achieving this minimum. For such a choice of X, Proposition 3 yields $\mathrm{opt}(A) \leq 2^{n-\mathrm{rank}(A_X)}$, where $\mathrm{rank}(A_X) = \mathrm{Mr}(A) - |X| \geq \mathrm{Mr}(A) - \mathrm{cov}(A)$. □

Given a depth-2 circuit F, let $\mathrm{m}(F)$ denote the largest number of wires in a matching formed by direct input-output wires.

Corollary 2. *Every depth-2 circuit F computing a linear operator can be transformed into an equivalent linear depth-2 circuit F' of the same degree and*

$$\mathrm{width}(F') \leq \mathrm{width}(F) + \mathrm{m}(F)\,.$$

Proof. Let A_F be the $(0,1,*)$-matrix of F. By Lemmas 3 and 5, we have that $\mathrm{width}(F) \geq \mathrm{Mrk}(A_F) - \mathrm{cov}(A_F)$ where, by Frobenius–König theorem, $\mathrm{cov}(A_F) = \mathrm{m}(F)$. By Lemma 2, the circuit F can be transformed into an equivalent linear depth-2 circuit of the same degree and width at most $\mathrm{mr}(A_F) \leq \mathrm{Mrk}(A_F) \leq \mathrm{width}(F) + \mathrm{m}(F)$. ☐

4 Row and column min-rank

We are now going to show that the min-rank conjecture holds for a stronger version of min-rank—row min-rank and column min-rank.

If A is a $(0,1,*)$-matrix of min-rank r then, for every assignment of constants to stars, the resulting $(0,1)$-matrix will have r linearly independent columns as well as r linearly independent rows. However, for different assignments these columns/rows may be different. It is natural to ask whether the min-rank conjecture is true if the matrix A has r columns (or r rows) that remain linearly independent under any assignment of constants to stars?

Namely, say that $(0,1,*)$-vectors are *dependent* if they can be made linearly dependent over GF_2 by setting each $*$-entry to a constant 0 or 1; otherwise, the vectors are *independent*.

Remark 4. The dependence of $(0,1,*)$-vectors can be defined by adding to $\{0,1\}$ a new element $*$ satisfying $\alpha \oplus * = * \oplus \alpha = *$ for $\alpha \in \{0,1,*\}$. Then a set of $(0,1,*)$-vectors is dependent iff some its subset sums up to a $\{0,*\}$-vector. In particular, if in each coordinate at least one of the vectors contains at least one $*$, then the vectors cannot be independent, just because then all vectors sum up to an all-$*$ vector.

Remark 5. A basic fact of Linear Algebra, leading to Gauss-Algorithm, is that linear independence of vectors $\boldsymbol{x}, \boldsymbol{y} \in \{0,1\}^n$ implies that the vectors $\boldsymbol{x}+\boldsymbol{y}$ and $\boldsymbol{y}$ are linear independent as well. For $(0,1,*)$-vectors this does not hold anymore. Take, for example, $\boldsymbol{x} = (0,1)$ and $\boldsymbol{y} = (1,*)$. Then $\boldsymbol{x} \oplus \boldsymbol{y} = (1,*) = \boldsymbol{y}$.

For a $(0,1,*)$-matrix A, define its *column min-rank* $\mathrm{mr}_{\mathrm{col}}(A)$ as the maximum number of independent columns, and its *row min-rank* $\mathrm{mr}_{\mathrm{row}}(A)$ as the maximum number of independent rows. In particular, both $\mathrm{mr}_{\mathrm{row}}(A)$ and $\mathrm{mr}_{\mathrm{col}}(A)$ are at least r if A contains an $r \times r$ "triangular" submatrix, that is, a submatrix with zeroes below (or above) the diagonal and ones on the diagonal:

$$\Delta = \begin{pmatrix} 1 & \circledast & \circledast & \circledast \\ 0 & 1 & \circledast & \circledast \\ 0 & 0 & 1 & \circledast \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where $\circledast \in \{0,1,*\}$. It is clear that neither $\mathrm{mr}_{\mathrm{col}}(A)$ nor $\mathrm{mr}_{\mathrm{row}}(A)$ can exceed the min-rank of A. Later (Lemma 10 below) we will give an example of a matrix A where both $\mathrm{mr}_{\mathrm{col}}(A)$

and $\mathrm{mr}_{\mathrm{row}}(A)$ are by a logarithmic factor smaller than $\mathrm{mr}(A)$. The question about a more precise relation between these parameters remains open (see Problem 1).

Albeit for $(0,1)$-matrices we always have that their row-rank coincides with column-rank, for $(0,1,*)$-matrices this is no more true. In particular, for some $(0,1,*)$-matrices A, we have that $\mathrm{mr}_{\mathrm{row}}(A) \neq \mathrm{mr}_{\mathrm{col}}(A)$.

Example 1. Consider the following $(0,1,*)$-matrix:

$$A = \begin{pmatrix} 1 & 1 & * & 1 \\ 1 & 0 & 1 & * \\ 1 & * & 0 & 0 \end{pmatrix} .$$

Then $\mathrm{mr}_{\mathrm{row}}(A) = \mathrm{mr}(A) = 3$ but $\mathrm{mr}_{\mathrm{col}}(A) = 2$. To see that $\mathrm{mr}_{\mathrm{row}}(A) = 3$, just observe that the rows cannot be made linearly dependent by setting the stars to 0 or 1: the sum of all three vectors is not a $\{0,*\}$-vector because of the 1st column, and the pairwise sums are not $\{0,*\}$-vectors because, for each pair of rows there is a column containing 0 and 1. To see that $\mathrm{mr}_{\mathrm{col}}(A) = 2$, observe that the last three columns are dependent (each row has a star). Moreover, for every pair of these columns, there is an assignment of constants to stars such that either the resulting $(0,1)$-columns are equal or their sum equals the first column.

We first show that the min-rank conjecture holds with "min-rank" replaced by "column min-rank".

Theorem 1 (Column min-rank)**.** *Let A be a $(0,1,*)$-matrix with n columns and of column min-rank r. Then* $\mathrm{opt}(A) \leq 2^{n-r}$.

Proof. Any m-by-n $(0,1,*)$-matrix B of column min-rank r must contain an $m \times r$ submatrix A of min-rank r. Since $\mathrm{opt}(B) \leq \mathrm{opt}(A) \cdot 2^{n-r}$ (Proposition 2), it is enough to show that $\mathrm{opt}(A) \leq 1$ for all m-by-r $(0,1,*)$-matrices A of min-rank r.

To do this, let L be a solution for A. Then there is an operator $G = (g_1, \ldots, g_m) : \{0,1\}^r \to \{0,1\}^m$ such that G is consistent with A and $\langle \boldsymbol{a}_i, \boldsymbol{x} \rangle = g_i(\boldsymbol{x})$ holds for all $\boldsymbol{x} \in L$ and all $i = 1, \ldots, m$. Here $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ are the rows of A with all stars set to 0.

For the sake of contradiction, assume that $|L| \geq 2$ and fix any two vectors $\boldsymbol{x} \neq \boldsymbol{y} \in L$. Our goal is to construct a vector $\boldsymbol{c} \in \{0,1\}^m$ and a completion M of A such that $M\boldsymbol{x} = M\boldsymbol{y} = \boldsymbol{c}$. Since M must have rank r, this will give the desired contradiction, because at most $2^{r-\mathrm{rank}(M)} = 2^0 = 1$ vectors $\boldsymbol{z}$ can satisfy $M\boldsymbol{z} = \boldsymbol{c}$.

If M is a completion of $A = (a_{ij})$, then its ith row must have the form $\boldsymbol{m}_i = \boldsymbol{a}_i \oplus \boldsymbol{p}_i$ where $\boldsymbol{p}_i \in \{0,1\}^n$ is some vector with no 1's in positions where the ith row of A has no stars. To construct the desired vector $\boldsymbol{p}_i$ for each $i \in [m]$, we consider two possible cases. (Recall that the vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are fixed.)

Case 1: $\langle \boldsymbol{a}_i, \boldsymbol{x} \rangle = \langle \boldsymbol{a}_i, \boldsymbol{y} \rangle$. In this case we can take $\boldsymbol{p}_i = \boldsymbol{0}$ and $c_i = \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle$. Then $\langle \boldsymbol{m}_i, \boldsymbol{x} \rangle = \langle \boldsymbol{m}_i, \boldsymbol{y} \rangle = \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle = c_i$, as desired.

Case 2: $\langle \boldsymbol{a}_i, \boldsymbol{x} \rangle \neq \langle \boldsymbol{a}_i, \boldsymbol{y} \rangle$. In this case we have that $g_i(\boldsymbol{x}) \neq g_i(\boldsymbol{y})$, that is, the vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ must differ in some position j where the ith row of A has a star. Then we can take $\boldsymbol{p}_i := \boldsymbol{e}_j$ (the jth unit vector) and $c_i := \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle \oplus x_j$. With this choice of $\boldsymbol{p}_i$, we again have

$$\langle \boldsymbol{m}_i, \boldsymbol{x} \rangle = \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle \oplus \langle \boldsymbol{p}_i, \boldsymbol{x} \rangle = \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle \oplus \langle \boldsymbol{e}_j, \boldsymbol{x} \rangle = \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle \oplus x_j = c_i$$

and, since $\langle \boldsymbol{a}_i, \boldsymbol{x} \rangle \neq \langle \boldsymbol{a}_i, \boldsymbol{y} \rangle$ and $x_j \neq y_j$,

$$\langle \boldsymbol{m}_i, \boldsymbol{y} \rangle = \langle \boldsymbol{a}_i, \boldsymbol{y} \rangle \oplus \langle \boldsymbol{p}_i, \boldsymbol{y} \rangle = \langle \boldsymbol{a}_i, \boldsymbol{y} \rangle \oplus \langle \boldsymbol{e}_j, \boldsymbol{y} \rangle = \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle \oplus x_j = c_i \, .$$

□

Example 2. It is not difficult to verify that, for the $(0, 1, *)$-matrix A given by (3), we have that $\mathrm{mr}_{\mathrm{col}}(A) = \mathrm{mr}(A) = 2$. Hence, no linear solution of the system of semi-linear equations (2) can have more than $\mathrm{lin}(A) = 2^{6-2} = 32$ vectors. Theorem 1 implies that, in fact, *no* solution can have more than this number of vectors.

The situation with *row* min-rank is more complicated. In this case we are only able to prove an upper bound $\mathrm{opt}(A) \le 2^{n-r}$ under an additional restriction that the star-positions in the rows of A form a chain under set-inclusion.

Recall that $(0, 1, *)$-vectors are *independent* if they cannot be made linearly dependent over GF_2 by setting stars to constants. The row min-rank of a $(0, 1, *)$-matrix is the largest number r of its independent rows. Since adding new rows can only decrease $\mathrm{opt}(A)$, it is enough to consider r-by-n $(0, 1, *)$-matrices A with $\mathrm{mr}(A) = r$.

If $r = 1$, that is, if A consists of just one row, then $\mathrm{opt}(A) \le 2^{n-1} = 2^{n-r}$ holds. Indeed, since $\mathrm{mr}(A) = 1$, this row cannot be a $(0, *)$-row. So, there must be at least one 1 in, say, the 1st position. Let $L_A = \{\boldsymbol{x} \colon \langle \boldsymbol{a}_1, \boldsymbol{x} \rangle = g_1(\boldsymbol{x})\}$ be a solution for A, where $\boldsymbol{a}_1$ the row of A with all stars set to 0. Take the unit vector $\boldsymbol{e}_1 = (1, 0, \ldots, 0)$ and split the vectors in $\{0, 1\}^n$ into 2^{n-1} pairs $\{\boldsymbol{x}, \boldsymbol{x} \oplus \boldsymbol{e}_1\}$. Since the boolean function g_1 cannot depend on the first variable x_1, we have that $g_i(\boldsymbol{x} \oplus \boldsymbol{e}_1) = g_i(\boldsymbol{x})$. But $\langle \boldsymbol{a}_i, \boldsymbol{x} \oplus \boldsymbol{e}_1 \rangle = \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle \oplus 1 \neq \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle$. Hence, at most one of the two vectors $\boldsymbol{x}$ and $\boldsymbol{x} \oplus \boldsymbol{e}_1$ from each pair $\{\boldsymbol{x}, \boldsymbol{x} \oplus \boldsymbol{e}_1\}$ can lie in L_A, implying that $|L_A| \le 2^{n-1}$.

To extend this argument for matrices with more rows, we need the following definition. Let $A = (a_{ij})$ be an r-by-n $(0, 1, *)$-matrix, and $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r$ be the rows of A with all stars set to 0. Let $S_i = \{j \colon a_{ij} = *\}$ be the set of star-positions in the ith row of A. It will be convenient to describe the star-positions by diagonal matrices. Namely, let D_i be the incidence matrix of stars in the ith row of A. That is, D_i is a diagonal n-by-n $(0, 1)$-matrix whose jth diagonal entry is 1 iff $j \in S_i$. In particular, $D_i \boldsymbol{x} = \boldsymbol{0}$ means that $x_j = 0$ for all $j \in S_i$.

Definition 2. A matrix A is *isolated* if there exist vectors $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_r \in \{0, 1\}^n$ such that, for all $1 \le i \le r$, we have $D_i \boldsymbol{z}_i = \boldsymbol{0}$ and

$$\langle \boldsymbol{a}_j, \boldsymbol{z}_i \rangle = \begin{cases} 1 & \text{if } j = i; \\ 0 & \text{if } j < i. \end{cases}$$

If $D_1 \boldsymbol{z}_i = \ldots = D_i \boldsymbol{z}_i = \boldsymbol{0}$, then the matrix is *strongly isolated.*

Lemma 6. *If A is a strongly isolated r-by-n $(0, 1, *)$-matrix, then* $\mathrm{opt}(A) \le 2^{n-r}$.

Proof. Let $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r$ be the rows of A with all stars set to 0. We prove the lemma by induction r. The basis case $r = 1$ is already proved above. For the induction step $r - 1 \mapsto r$, let

$$L_A = \{\boldsymbol{x} \in \{0, 1\}^n \colon \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle = g_i(\boldsymbol{x}) \text{ for all } i = 1, \ldots, r\}$$

be an optimal solution for A, and let B be a submatrix of A consisting of its first $r - 1$ rows. Then

$$L_B = \{\boldsymbol{x} \in \{0, 1\}^n \colon \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle = g_i(\boldsymbol{x}) \text{ for all } i = 1, \ldots, r - 1\}$$

is a solution for B. Since A is strongly isolated, the matrix B is strongly isolated as well. The induction hypothesis implies that $|L_B| \le 2^{n-(r-1)}$.

Let $\boldsymbol{z} = \boldsymbol{z}_r$ be the r-th isolating vector. For each row $i = 1, \ldots, r-1$, the conditions $\langle \boldsymbol{z}, \boldsymbol{a}_i \rangle = 0$ and $D_i \boldsymbol{z} = \boldsymbol{0}$ imply that $\langle (\boldsymbol{x} \oplus \boldsymbol{z}), \boldsymbol{a}_i \rangle = \langle \boldsymbol{x}, \boldsymbol{a}_i \rangle$ and $g_i(\boldsymbol{x} \oplus \boldsymbol{z}) = g_i(\boldsymbol{x})$. That is,

$$\boldsymbol{x} \in L_B \text{ iff } \boldsymbol{x} \oplus \boldsymbol{z} \in L_B.$$

For the rth row, the conditions $\langle \boldsymbol{z}, \boldsymbol{a}_r \rangle = 1$ and $D_r \boldsymbol{z} = \boldsymbol{0}$ imply that $\langle (\boldsymbol{x} \oplus \boldsymbol{z}), \boldsymbol{a}_r \rangle \neq \langle \boldsymbol{x}, \boldsymbol{a}_r \rangle$ whereas $g_r(\boldsymbol{x} \oplus \boldsymbol{z}) = g_r(\boldsymbol{x})$. That is,

$$\boldsymbol{x} \in L_A \text{ iff } \boldsymbol{x} \oplus \boldsymbol{z} \notin L_A.$$

Hence, for every vector $\boldsymbol{x} \in L_B$, only one of the vectors $\boldsymbol{x}$ and $\boldsymbol{x} \oplus \boldsymbol{z}$ can belong to L_A, implying that $\mathrm{opt}(A) = |L_A| \leq |L_B|/2 \leq 2^{n-r}$. □

We are now going to show that $(0,1,*)$-matrices with some conditions on the distribution of stars in them are strongly isolated. For this, we need the following two facts. A *projection* of a vector $\boldsymbol{x} = (x_1, \ldots, x_n)$ onto a set of positions $I = \{i_1, \ldots, i_k\}$ is the vector

$$\boldsymbol{x}\restriction_I = (x_{i_1}, \ldots, x_{i_k})\,.$$

A $(0,1,*)$-vector $\boldsymbol{x}$ is *independent of* $(0,1,*)$-vectors $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k$ if no completion of $\boldsymbol{x}$ can be written as a linear combination of some completions of these vectors.

Proposition 4. *Let $\boldsymbol{x}, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_k$ be $(0,1,*)$-vectors, and $I = \{i \colon x_i \neq *\}$. If $\boldsymbol{x}$ is independent of $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k$, then $\boldsymbol{x}\restriction_I$ is also independent of $\boldsymbol{y}_1\restriction_I, \ldots, \boldsymbol{y}_k\restriction_I$.*

Proof. Assume that $\boldsymbol{x}\restriction_I$ is dependent on the projections $\boldsymbol{y}_1\restriction_I, \ldots, \boldsymbol{y}_k\restriction_I$. Then there is an assignment of stars to constants in the vectors $\boldsymbol{y}_i$ such that $\boldsymbol{x}\restriction_I$ can be written as a linear combination of the projections $\boldsymbol{y}'_1\restriction_I, \ldots, \boldsymbol{y}'_k\restriction_I$ on I of the resulting $(0,1)$-vectors $\boldsymbol{y}'_1, \ldots, \boldsymbol{y}'_k$. But since $\boldsymbol{x}$ has stars in all positions outside I, these stars can be set to appropriate constants so that the resulting $(0,1)$-vector $\boldsymbol{x}'$ will be a linear combination of $\boldsymbol{y}'_1, \ldots, \boldsymbol{y}'_k$, a contradiction. □

Proposition 5. *Let $\boldsymbol{a} \in \{0,1\}^n$ be a vector and M be an m-by-n $(0,1)$-matrix of rank $r \leq n-1$. If $\boldsymbol{a}$ is linearly independent of the rows of M, then there exists a set $Z \subseteq \{0,1\}^n$ of $|Z| \geq 2^{n-r-1}$ vectors such that, for all $\boldsymbol{z} \in Z$, we have $\langle \boldsymbol{z}, \boldsymbol{a} \rangle = 1$ and $M\boldsymbol{z} = \boldsymbol{0}$.*

Proof. Let $Z = \{\boldsymbol{z} \colon M\boldsymbol{z} = \boldsymbol{0}, \langle \boldsymbol{a}, \boldsymbol{z} \rangle = 1\}$, and let M' be the matrix M with an additional row $\boldsymbol{a}$. Note that $Z = \ker(M) \setminus \ker(M')$. Since $\mathrm{rank}(M') = \mathrm{rank}(M) + 1 \leq n$, we have that $|\ker(M')| = |\ker(M)|/2$, implying that

$$|Z| = |\ker(M) \setminus \ker(M')| = |\ker(M)|/2 \geq 2^{n-r-1}\,.$$ □

Lemma 7. *If A is an r-by-n $(0,1,*)$-matrix with $\mathrm{mr}(A) = r$, then A is isolated.*

Proof. Let $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r$ be the rows of A with all stars set to 0. Let $I \subseteq \{1, \ldots, n\}$ be the set of all star-free positions in the ith row of A, and consider an $(r-1)$-by-$|I|$ $(0,1)$-matrix M_i whose rows are the projections $\boldsymbol{a}'_j = \boldsymbol{a}_j\restriction_I$ of vectors $\boldsymbol{a}_j$ with $j \neq i$ onto the set I. By Proposition 4, the projection $\boldsymbol{a}'_i = \boldsymbol{a}_i\restriction_I$ of the ith vector $\boldsymbol{a}_i$ onto I cannot be written as a linear combination of the rows of M_i; hence, $\mathrm{rank}(M_i) \leq |I| - 1$. Since $2^{|I| - \mathrm{rank}(M_i) - 1} \geq 2^0 = 1$, Proposition 5 gives us a vector $\boldsymbol{z}'_i \in \{0,1\}^{|I|}$ such that $\langle \boldsymbol{z}'_i, \boldsymbol{a}'_i \rangle = 1$ and $\langle \boldsymbol{z}'_i, \boldsymbol{a}'_j \rangle = 0$ for all $j \neq i$. But then $\boldsymbol{z}_i := (\boldsymbol{z}'_i, \boldsymbol{0})$ is the desired $(0,1)$-vector: $D_i \boldsymbol{z}_i = D_i \cdot \boldsymbol{0} = \boldsymbol{0}$, $\langle \boldsymbol{z}_i, \boldsymbol{a}_i \rangle = \langle \boldsymbol{z}'_i, \boldsymbol{a}'_i \rangle = 1$, and $\langle \boldsymbol{z}_i, \boldsymbol{a}_j \rangle = \langle \boldsymbol{z}'_i, \boldsymbol{a}'_j \rangle = 0$ for all rows $j \neq i$. □

Say that an r-by-n $(0,1,*)$-matrix A is *star-monotone* if the sets $S_1,\ldots,S_r$ of star-positions in its rows form a chain, that is, if $S_1 \subseteq S_2 \subseteq \ldots \subseteq S_r$.

Theorem 2 (Star-monotone matrices)**.** *Let A be a $(0,1,*)$-matrix with n columns. If A contains an r-by-n star-monotone submatrix of min-rank r, then* $\mathrm{opt}(A) \le 2^{n-r}$.

Proof. Since addition of new rows can only decrease the size of a solution, we can assume that A itself is an r-by-n star-monotone matrix of min-rank r. Let $\boldsymbol{a}_1,\ldots,\boldsymbol{a}_r$ be the rows of A with all stars set to 0. By Lemma 7, the matrix A is isolated. That is, there exist vectors $\boldsymbol{z}_1,\ldots,\boldsymbol{z}_r \in \{0,1\}^n$ such that: $\langle \boldsymbol{a}_i, \boldsymbol{z}_j\rangle = 1$ iff $i=j$, and $D_i\boldsymbol{z}_i = \boldsymbol{0}$ for all $1 \le i \le r$. Since $S_j \subseteq S_i$ for all $j<i$, this last condition implies that $D_j\boldsymbol{z}_i = \boldsymbol{0}$ for all $1 \le j < i \le r$, that is, A is strongly isolated. Hence, we can apply Lemma 6. □

5 Solutions as independent sets in Cayley graphs

Let $A=(a_{ij})$ be an m-by-n $(0,1,*)$-matrix. In the definition of solutions L for A we take a completion M of A and an operator $G(\boldsymbol{x})$, and require that $M\boldsymbol{x} = G(\boldsymbol{x})$ for all $\boldsymbol{x} \in L$. The operator $G=(g_1,\ldots,g_m)$ can be arbitrary—the only restriction is that its ith component g_i can only depend on variables corresponding to stars in the ith row of A. In this section we show that the actual *form* of operators G can be ignored—only star-positions are important. To do this, we associate with A the following set of "forbidden" vectors:

$$K_A = \{\boldsymbol{x} \in \{0,1\}^n : \exists i \in [m] \;\; D_i\boldsymbol{x} = \boldsymbol{0} \text{ and } \langle \boldsymbol{a}_i, \boldsymbol{x}\rangle = 1\},$$

where D_i is the incidence n-by-n $(0,1)$-matrix of stars in the ith row of A, and $\boldsymbol{a}_i$ is the ith row of A with all stars set to 0. Hence, K_A is a union $K_A = \bigcup_{i=1}^m K_i$ of m affine spaces

$$K_i = \left\{ \boldsymbol{x} \colon \begin{pmatrix} D_i \\ \boldsymbol{a}_i \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} \boldsymbol{0} \\ 1 \end{pmatrix} \right\}.$$

For sets of vectors $S,T \subseteq \{0,1\}^n$, let

$$S+T = \{\boldsymbol{x} \oplus \boldsymbol{y} \colon \boldsymbol{x} \in S, \boldsymbol{y} \in T\}.$$

Theorem 3. *A set $L \subseteq \{0,1\}^n$ is a solution for A if and only if $(L+L)\cap K_A = \emptyset$.*

Proof. Observe that the sum $\boldsymbol{x} \oplus \boldsymbol{y}$ of two vectors belongs to K_A iff these vectors coincide on all stars of at least one row of A such that $\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle \ne \langle \boldsymbol{a}_i, \boldsymbol{y}\rangle$. By this observation, we see that the condition $(L+L)\cap K_A = \emptyset$ is equivalent to:

$$\forall \boldsymbol{x},\boldsymbol{y} \in L \;\; \forall i \in [m]: \; D_i\boldsymbol{x} = D_i\boldsymbol{y} \;\text{ implies }\; \langle \boldsymbol{a}_i, \boldsymbol{x}\rangle = \langle \boldsymbol{a}_i, \boldsymbol{y}\rangle. \tag{8}$$

We now turn to the actual proof of Theorem 3.

($\Rightarrow$) Let L be a solution for A. Hence, there is a consistent with A operator $G=(g_1,\ldots,g_m)$ such that $\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle = g_i(\boldsymbol{x})$ for all $\boldsymbol{x} \in L$ and all rows $i \in [m]$. To show that then L must satisfy (8), take any two vectors $\boldsymbol{x},\boldsymbol{y} \in L$ and assume that $D_i\boldsymbol{x} = D_i\boldsymbol{y}$. This means that vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ must coincide in all positions where the ith row of A has stars. Since g_i can only depend on these positions, this implies $g_i(\boldsymbol{x}) = g_i(\boldsymbol{y})$, and hence, $\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle = \langle \boldsymbol{a}_i, \boldsymbol{y}\rangle$.

($\Leftarrow$) Assume that $L \subseteq \{0,1\}^n$ satisfies (8). We have to show that then there exists a consistent with A operator $G=(g_1,\ldots,g_m)$ such that $\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle = g_i(\boldsymbol{x})$ for all $\boldsymbol{x} \in L$ and

$i \in [m]$; here, as before, $\boldsymbol{a}_i$ is the ith row of A with all stars set to 0. The ith row of A splits the set L into two subsets

$$L_i^0 = \{\boldsymbol{x} \in L \colon \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle = 0\} \text{ and } L_i^1 = \{\boldsymbol{x} \in L \colon \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle = 1\}.$$

Condition (8) implies that $D_i\boldsymbol{x} \neq D_i\boldsymbol{y}$ for all $(\boldsymbol{x}, \boldsymbol{y}) \in L_i^0 \times L_i^1$. That is, if S_i is the set of star-positions in the ith row of A, then the projections $\boldsymbol{x}\lceil_{S_i}$ of vectors $\boldsymbol{x}$ in L_i^0 onto these positions must be different from all the projections $\boldsymbol{y}\lceil_{S_i}$ of vectors $\boldsymbol{y}$ in L_i^1. Hence, we can find a boolean function $g_i : \{0,1\}^{S_i} \to \{0,1\}$ taking different values on these two sets of projections. This function will then satisfy $g_i(\boldsymbol{x}) = \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle$ for all $\boldsymbol{x} \in L$. ☐

A *coset* of a set of vectors $L \subseteq \{0,1\}^n$ is a set $\boldsymbol{v} + L = \{\boldsymbol{v} \oplus \boldsymbol{x} \colon \boldsymbol{x} \in L\}$ with $\boldsymbol{v} \in \{0,1\}^n$. Since $(\boldsymbol{v} + L) + (\boldsymbol{v} + L) = L + L$, Theorem 3 implies:

Corollary 3. *Every coset of a solution for a $(0,1,*)$-matrix A is also a solution for A.*

Remark 6. A Cayley graph over the Abelian group $(\{0,1\}^n, \oplus)$ generated by a set $K \subseteq \{0,1\}^n$ of vectors has all vectors in $\{0,1\}^n$ as vertices, and two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are joined by an edge iff $\boldsymbol{x} \oplus \boldsymbol{y} \in K$. Theorem 3 shows that solutions for a $(0,1,*)$-matrix A are precisely the independent sets in a Cayley graph generated by a special set K_A.

Remark 7. If A is an m-by-n $(0,1)$-matrix, that is, has no stars at all, then $K_A = \{\boldsymbol{x} \colon A\boldsymbol{x} \neq \boldsymbol{0}\}$. Hence, in this case, a set $L \subseteq \{0,1\}^n$ is a solution for A iff there is a vector $\boldsymbol{b} \in \{0,1\}^m$ such that $A\boldsymbol{x} = \boldsymbol{b}$ for all $\boldsymbol{x} \in L$. That is, in this case, $\ker(A)$ is an optimal solution.

6 Structure of linear solutions

By Theorem 3, a set of vectors $L \subseteq \{0,1\}^n$ is a solution for an m-by-n $(0,1,*)$-matrix A if and only if $(L + L) \cap K_A = \emptyset$, where $K_A \subseteq \{0,1\}^n$ is the set of "forbidden" vectors defined by

$$K_A = \{\boldsymbol{x} \in \{0,1\}^n \colon \exists i \in [m] \;\; D_i\boldsymbol{x} = \boldsymbol{0} \text{ and } \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle = 1\};$$

here $D_1, \ldots, D_m$ are diagonal $n \times n$ $(0,1)$-matrices corresponding the stars in the matrix A, and $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ are the rows of A with all stars set to 0. Thus, *linear* solutions are precisely vector subspaces of $\{0,1\}^n$ avoiding the set K_A. Which subspaces these are?

Each subspace of $\{0,1\}^n$ is a kernel $\ker(H) = \{\boldsymbol{x} \colon H\boldsymbol{x} = \boldsymbol{0}\}$ of some $(0,1)$-matrix H. Hence, linear solution for A are given by matrices H such that $H\boldsymbol{x} \neq \boldsymbol{0}$ for all $\boldsymbol{x} \in K_A$; in this case we say that the matrix H *separates* K_A from zero. By the *span-matrix* of a $(0,1)$-matrix H we will mean the matrix $\widehat{H}$ whose rows are all linear combinations of the rows of H.

Lemma 8. *Let A be a $(0,1,*)$-matrix. Then*

(i) *Every completion M of A separates K_A from zero.*

(ii) *A $(0,1)$-matrix H separates K_A from zero iff $\widehat{H}$ contains a completion of A.*

Proof. (i) For the sake of contradiction, assume that some vector $\boldsymbol{x} \in K_A$ lies in the kernel of some completion M of A. Then $M\boldsymbol{x} = \boldsymbol{0}$, $D_i\boldsymbol{x} = \boldsymbol{0}$ and $\langle \boldsymbol{a}_i, \boldsymbol{x} \rangle = 1$ for some i. The ith row of M has the form $\boldsymbol{m}_i = \boldsymbol{a}_i \oplus \boldsymbol{p}_i$ where $D_i\boldsymbol{p}_i = \boldsymbol{p}_i$. Since $D_i\boldsymbol{x} = \boldsymbol{0}$, we have that $\langle \boldsymbol{p}_i, \boldsymbol{x} \rangle = 0$. Hence, $\langle \boldsymbol{m}_i, \boldsymbol{x} \rangle = \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle \oplus \langle \boldsymbol{p}_i, \boldsymbol{x} \rangle = \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle = 1$, a contradiction with $M\boldsymbol{x} = \boldsymbol{0}$.

(ii) To prove ($\Leftarrow$), suppose that some completion M of A is a submatrix of $\widehat{H}$. Let $\boldsymbol{x} \in K_A$. By (i), we know that then $M\boldsymbol{x} \neq \mathbf{0}$, and hence, also $\widehat{H}\boldsymbol{x} \neq \mathbf{0}$. Since $H\boldsymbol{x} = \mathbf{0}$ would imply $\widehat{H}\boldsymbol{x} = \mathbf{0}$, we also have that $H\boldsymbol{x} \neq \mathbf{0}$.

To prove ($\Rightarrow$), suppose that H separates K_A. Then, for every row $i \in [m]$ and every vector $\boldsymbol{x} \in \{0,1\}^n$, $H\boldsymbol{x} = \mathbf{0}$ and $D_i\boldsymbol{x} = \mathbf{0}$ imply that $\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle = 0$.

Claim 1. *Let $\boldsymbol{a} \in \{0,1\}^n$ and M be a $(0,1)$-matrix with n columns. If for every $\boldsymbol{x} \in \{0,1\}^n$, $M\boldsymbol{x} = \mathbf{0}$ implies $\langle \boldsymbol{a}, \boldsymbol{x}\rangle = 0$, then $\boldsymbol{a}$ is a linear combination of rows of M.*

Proof. Extend M to a matrix M' by adding a new row $\boldsymbol{a}$. The condition implies that $\ker(M') = \ker(M)$. Hence, $\mathrm{rank}(M') = \mathrm{rank}(M)$, implying that $\boldsymbol{a}$ must be linearly dependent on the rows of M. ☐

Hence, for each i, the vector $\boldsymbol{a}_i$ must lie in the vector space spanned by the rows of H and D_i, that is, $\boldsymbol{a}_i = \boldsymbol{\alpha}_i^\top H \oplus \boldsymbol{\beta}_i^\top D_i$ for some vectors $\boldsymbol{\alpha}_i$ and $\boldsymbol{\beta}_i$. In other words, the ith linear combination $\boldsymbol{\alpha}_i^\top H$ of the rows of H is the ith row $\boldsymbol{a}_i \oplus \boldsymbol{\beta}_i^\top D_i$ of a particular completion M of A, implying that M is a submatrix of $\widehat{H}$, as desired. ☐

Theorem 4. *Let A be $(0,1,*)$-matrix. A linear subspace is a solution for A if and only if it is contained in a kernel of some completion of A.*

Proof. ($\Leftarrow$): If a linear subspace $L \subseteq \{0,1\}^n$ lies in a kernel of some completion of A, then L is a solution for A, by Lemma 8(i).

($\Rightarrow$): Let $L \subseteq \{0,1\}^n$ be an arbitrary linear solution for A. Then $L + L = L$ and $L \cap K_A = \emptyset$. Take a $(0,1)$-matrix H with $L = \ker(H)$. Since $\ker(H) \cap K_A = \emptyset$, the matrix H separates K_A from zero. Lemma 8(ii) implies that then $\widehat{H}$ must contain some completion M of A. But then $L = \ker(H) = \ker(\widehat{H}) \subseteq \ker(M)$, as claimed. ☐

Corollary 4. *For any $(0,1,*)$-matrix A we have that $\mathrm{lin}(A) = 2^{n-\mathrm{mr}(A)}$.*

Proof. By Theorem 4, $\mathrm{lin}(A)$ is the maximum of $|\ker(M)| = 2^{n-\mathrm{rank}(M)}$ over all completions M of A. Since $\mathrm{mr}(A)$ is the minimum of $\mathrm{rank}(M)$ over all completions M of A, we are done. ☐

Corollary 5 (Alternative definition of min-rank)**.** *For every $(0,1,*)$-matrix A we have*

$$\mathrm{mr}(A) = \min\{\mathrm{rank}(H) \colon H \text{ separates } K_A \text{ from zero}\}\,.$$

Proof. Let R be the smallest possible rank of a $(0,1)$-matrix separating K_A from zero. To prove $\mathrm{mr}(A) \geq R$, let M be a completion of A with $\mathrm{rank}(M) = \mathrm{mr}(A)$. By Lemma 8(i), the matrix M separates K_A form zero. Hence, $R \leq \mathrm{rank}(M) = \mathrm{mr}(A)$.

To prove $\mathrm{mr}(A) \leq R$, let H be a $(0,1)$-matrix such that H separates K_A form zero and $\mathrm{rank}(H) = R$. By Lemma 8(ii), the matrix $\widehat{H}$ must contain a completion M of A. Hence, $\mathrm{mr}(A) \leq \mathrm{rank}(M) \leq \mathrm{rank}(\widehat{H}) = \mathrm{rank}(H) = R$. ☐

7 Structure of general solutions

The following theorem says that non-linear solutions must be "very non-linear": they cannot contain large linear subspaces. Recall that in Valiant's setting (cf. Lemma 1) we may assume that each row of a $(0,1,*)$-matrix contains at most $s = n^\delta$ stars, where $\delta > 0$ is an arbitrary small constant.

Theorem 5. *Let $L \subseteq \{0,1\}^n$ be a solution for an m-by-n $(0,1,*)$-matrix A, and let s be the maximum number of stars in a row of A. If L contains an affine subspace of dimension $s+1$, then some coset of L lies in a linear solution for A.*

Proof. Take an arbitrary solution L' for A, and suppose that L' contains an affine subspace $\boldsymbol{v} + W$ of dimension $s+1$. By Corollary 3, the coset $L = \boldsymbol{v} + L'$ of L' is also a solution for A and contains the vector space W.

Since L is a solution for A, W is a linear solution for A as well. Hence, by Theorem 4, W is contained in a kernel of some completion M of A. Our goal is to show that then the entire solution L must be contained in $\ker(M)$. To show this, we will use the following simple fact.

Claim 2. *Let $W \subseteq \{0,1\}^n$ be a linear subspace of dimension $k+1$. Then, for every k-element subset $S \subseteq [n]$ and for every vector $\boldsymbol{y} \in \{0,1\}^n$, there is a vector $\boldsymbol{x} \in W$ such that $\boldsymbol{x} \neq \boldsymbol{0}$ and $\boldsymbol{y}\restriction_S = \boldsymbol{x}\restriction_S$.*

Proof of Claim. For a vector $\boldsymbol{y} \in \{0,1\}^n$ and a k-element subset $S \subseteq \{1, \ldots, n\}$, consider the $n-k$ dimensional subspace $V = \{\boldsymbol{x} \colon \boldsymbol{y}\restriction_S = \boldsymbol{x}\restriction_S\}$. Then

$$\dim(V \cap W) = \dim V + \dim W - \dim(V + W) \geq (n-k) + k + 1 - n = 1 > 0,$$

and hence, $|V \cap W| \geq 2$. □

Assume now that $L \not\subseteq \ker(M)$, and take a vector $\boldsymbol{y} \in L \setminus \ker(M)$. Since $\boldsymbol{y} \notin \ker(M)$, we have that $\langle \boldsymbol{m}_i, \boldsymbol{y}\rangle = 1$ for at least one row $\boldsymbol{m}_i$ of M. Let S be the set of star-positions in the ith row of A (hence, $|S| \leq s$), and let $\boldsymbol{a}_i$ be this row of A with all stars set to 0. By Claim 2, there must be a vector $\boldsymbol{x} \in W \subseteq L \cap \ker(M)$ with $\boldsymbol{y}\restriction_S = \boldsymbol{x}\restriction_S$, that is, $D_i(\boldsymbol{x} \oplus \boldsymbol{y}) = \boldsymbol{0}$. But $\boldsymbol{x} \in \ker(M)$ implies that $\langle \boldsymbol{m}_i, \boldsymbol{x}\rangle = 0$. Hence, $\langle \boldsymbol{m}_i, \boldsymbol{x} \oplus \boldsymbol{y}\rangle = \langle \boldsymbol{m}_i, \boldsymbol{x}\rangle \oplus \langle \boldsymbol{m}_i, \boldsymbol{y}\rangle = \langle \boldsymbol{m}_i, \boldsymbol{y}\rangle = 1$. Since the vector $\boldsymbol{m}_i$ can only differ from $\boldsymbol{a}_i$ is star-positions of the ith row of A and, due to $D_i(\boldsymbol{x} \oplus \boldsymbol{y}) = \boldsymbol{0}$, the vector $\boldsymbol{x} \oplus \boldsymbol{y}$ has no 1's in these positions, we obtain that $\langle \boldsymbol{a}_i, \boldsymbol{x} \oplus \boldsymbol{y}\rangle = 1$. Hence, the vector $\boldsymbol{x} \oplus \boldsymbol{y}$ belongs to K_A, a contradiction with $\boldsymbol{x}, \boldsymbol{y} \in L$.

This completes the proof of Theorem 5. □

8 Relation to codes

Let $1 \leq r < n$ be integers. A (binary) error-correcting code of minimal distance $r+1$ is a set $C \subseteq \{0,1\}^n$ of vectors, any two of which differ in at least $r+1$ coordinates. A code is *linear* if it forms a linear subspace over GF_2. The question on how good linear codes are, when compared to non-linear ones, is a classical problem in Coding Theory. We now will show that this is just a special case of a more general "opt(A) versus lin(A)" problem for $(0,1,*)$-matrices, and that Min-Rank Conjecture in this special case holds true.

An (n,r)-*code matrix*, or just an r-*code matrix* if the number n of columns is not important, is a $(0,1,*)$-matrix with n columns and $m = (r+1)\binom{n}{r}$ rows, each of which consists of $n-r$

stars and at most one 0. The matrix is constructed as follows. For every r-element subset S of $[n] = \{1, \ldots, n\}$ include in A a block of $r+1$ rows $\boldsymbol{a}$ with $a_i = *$ for all $i \notin S$, $a_i \in \{0,1\}$ for all $i \in S$, and $|\{i \in S\colon a_i = 0\}| \leq 1$. That is, each of these rows has stars outside S and has at most one 0 within S. For $r = 3$ and $S = \{1,2,3\}$ such a block looks like

$$A_S = \begin{pmatrix} 1 & 1 & 1 & * & \cdots & * \\ 0 & 1 & 1 & * & \cdots & * \\ 1 & 0 & 1 & * & \cdots & * \\ 1 & 1 & 0 & * & \cdots & * \end{pmatrix}.$$

A Hamming ball around the all-0 vector $\mathbf{0}$ is defined by

$$\mathrm{Ball}(r) = \{\boldsymbol{x} \in \{0,1\}^n \colon 0 \leq |\boldsymbol{x}| \leq r\},$$

where $|\boldsymbol{x}| = x_1 + \cdots + x_n$ is the number of 1's in $\boldsymbol{x}$.

Observation 1. *If A is an r-code matrix, then $K_A = \mathrm{Ball}(r) \setminus \{\mathbf{0}\}$.*

Proof. It is easy to see that no vector $\boldsymbol{x} \in \{0,1\}^r$, $\boldsymbol{x} \neq \mathbf{0}$ can be orthogonal to all $r+1$ vectors in $\{0,1\}^r$ with at most one 0. By this observation, a vector $\boldsymbol{x}$ belongs to K_A iff there is an r-element set $S \subseteq [n]$ of positions such that $\boldsymbol{x}\restriction_S \neq \mathbf{0}$ and $\boldsymbol{x}\restriction_{\overline{S}} = \mathbf{0}$, that is, iff $\boldsymbol{x} \neq \mathbf{0}$ and $\boldsymbol{x} \in \mathrm{Ball}(r)$. □

Observation 2. *If A is an (n,r)-code matrix, then the solutions for A are error-correcting codes of minimal distance $r+1$, and linear solutions for A are linear codes.*

Proof. We have $(L+L) \cap (\mathrm{Ball}(r) \setminus \{\mathbf{0}\}) = \emptyset$ iff $|\boldsymbol{x} \oplus \boldsymbol{y}| \geq r+1$ for all $\boldsymbol{x} \neq \boldsymbol{y} \in L$, that is, iff every two vectors $\boldsymbol{x} \neq \boldsymbol{y} \in L$ differ in at least $r+1$ positions. Hence, every solution for an r-code matrix A is a code of minimal distance at least $r+1$, and linear solutions are linear codes. □

Lemma 9. *For code matrices, the min-rank conjecture holds with a constant $\epsilon > 0$.*

Proof. Let A be an (n,r)-code matrix; hence, $K_A = \mathrm{Ball}(r) \setminus \{\mathbf{0}\}$. Set $t := \lfloor (r-1)/2 \rfloor$. Since $|\boldsymbol{x} \oplus \boldsymbol{y}| \leq 2t < r$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathrm{Ball}(t)$, the sum of any two vectors $\boldsymbol{x} \neq \boldsymbol{y}$ from $\mathrm{Ball}(t)$ lies in K_A, implying that $\mathrm{Ball}(t)$ is a clique in the Cayley graph generated by K_A. Since, by Remark 6, solutions for A are independent sets in this graph, and since in any graph the number of its vertices divided by the clique number is an upper bound on the size of any independent set, we obtain:

$$\mathrm{opt}(A) \leq 2^n / |\mathrm{Ball}(t)| = 2^n \Big/ \sum_{i=0}^{t} \binom{n}{i}, \tag{9}$$

which is the well-known Hamming bound for codes. On the other hand, Gilbert-Varshamov bound says that linear codes in $\{0,1\}^n$ of dimension k and minimum distance d exist, if

$$\sum_{i=0}^{d-2} \binom{n-1}{i} < 2^{n-k}.$$

Hence,

$$\mathrm{lin}(A) \geq 2^n \Big/ \sum_{i=0}^{r} \binom{n}{i}. \tag{10}$$

Together with (9), this implies that the inequality (1) holds with ϵ about 1/2. □

The example of code matrices also shows that the gap between min-rank and row/column min-rank may be at least logarithmic in n.

Lemma 10. *If A is an (n,r)-code matrix, then* $\mathrm{mr}(A) = \Omega(r \ln(n/r))$ *but both row and column min-ranks of A do not exceed $r+1$.*

Proof. To prove $\mathrm{mr}(A) = \Omega(r \ln(n/r))$, recall that $K_A = \mathrm{Ball}(r) \setminus \{\mathbf{0}\}$. Hence, Corollary 5 implies that $\mathrm{mr}(A)$ is the smallest possible rank of a $(0,1)$-matrix H such that $\ker(H) \cap \mathrm{Ball}(r) \subseteq \{\mathbf{0}\}$. On the other hand, for any such matrix H, its kernel $L = \ker(H)$ is a (linear) code of minimal distance at least $r+1$ containing $|L| = 2^{n-\mathrm{rank}(H)}$ vectors. Since, by Hamming bound (9), no code L of distance at least $r+1$ can have more than $N = 2^n/(n/r)^{O(r)}$ vectors, we have that

$$\mathrm{rank}(H) = n - \log_2 |L| \geq n - \log_2 N = \Omega(r \ln(n/r)) .$$

To prove that $\mathrm{mr}_{\mathrm{col}}(A) \leq r+1$, suppose that A contains some $m \times k$ submatrix B of min-rank k. Since all k columns must be independent, at least one row $\boldsymbol{b}$ of B must be $*$-free and contain an odd number $|\boldsymbol{b}|$ of 1's. But every row of A (and hence, also $\boldsymbol{b}$) can contain at most one 0, implying that $|\boldsymbol{b}| \geq k-1$. Together with $|\boldsymbol{b}| \leq r$, this implies that $k \leq r+1$.

To prove that $\mathrm{mr}_{\mathrm{row}}(A) \leq r+1$, recall that each row of A consists of $n-r$ stars and at most one 0; the remaining r (or $r-1$) entries are 1's. Suppose now that A contains some set $X = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ of independent rows. That is, no subset of these rows can be made linearly dependent by setting $*$'s to 0 or 1. The rows in X must be, in particular, *pairwise* independent. This means that for every pair of rows in X we must have the following configuration:

$$\begin{pmatrix} \cdots & 0 & \cdots \\ \cdots & 1 & \cdots \end{pmatrix} .$$

Moreover, since none of the rows can have more than one 0, the following two configurations

$$\begin{pmatrix} \cdots & 0 & \cdots \\ \cdots & 0 & \cdots \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} \cdots & 0 & \cdots & * & \cdots \\ \cdots & * & \cdots & 0 & \cdots \end{pmatrix}$$

are forbidden. Hence, there must be $k = |X|$ different columns $i_1 < i_2 < \ldots < i_k$ such that, for every row $\boldsymbol{x}_j \in X$, we have that $\boldsymbol{x}_j(i_j) = 0$ and $\boldsymbol{x}_{j+1}(i_j) = \ldots = \boldsymbol{x}_k(i_j) = 1$. If, say, $k = 4$ then A must contain the following submatrix

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

where the entries above the diagonal belong to $\{1, *\}$. But then the last row $\boldsymbol{x}_k$ must have at least $k-1$ ones. Since no row can have more than r ones, this implies $|X| = k \leq r+1$. ☐

9 Conclusion and open problems

In this paper we rise a conjecture about systems of semi-linear equations and show its relation to proving super-linear lower bounds for log-depth circuits. We then give a support for the conjecture by proving that some its weaker versions are true. We also show that solutions are independent sets in particular Cayley graphs, thus turning the conjecture in a more general (combinatorial) setting. Using this, we prove several structural properties of sets of solutions that might be useful when tackling the original conjecture.

We defined solutions for a given m-by-n $(0,1,*)$-matrix A as sets $L \subseteq \{0,1\}^n$ of vectors $\boldsymbol{x}$ satisfying a system of equations

$$\langle \boldsymbol{a}_i, \boldsymbol{x} \rangle = g_i(D_i \boldsymbol{x}) \qquad i = 1, \ldots, m\,, \tag{11}$$

where $\boldsymbol{a}_i$ is the ith row of A with all stars replaced by 0, g_i is an arbitrary boolean function, and D_i is a diagonal n-by-n $(0,1)$-matrix corresponding to stars in the ith row of A. We have also shown (see Remark 6) that solution for A are precisely the independent sets in a Cayley graph over the Abelian group $(\{0,1\}^n, \oplus)$ generated by a special set of vectors

$$K_A = \{\boldsymbol{x} \colon \exists i\ D_i \boldsymbol{x} = \boldsymbol{0} \text{ and } \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle = 1\}\,. \tag{12}$$

The following two questions about possible generalizations of the min-rank conjecture naturally arise:

1. What if instead of diagonal matrices D_i in (11) we would allow other $(0,1)$-matrices?
2. What if instead of special generating sets K_A, defined by (12), we would allow other generating sets?

The following two examples show that the min-rank conjecture cannot be fetched too far: its generalized versions are false.

Example 3 (Bad generating sets K)**.** Let G be a Cayley graph generated by the set $K \subseteq \{0,1\}^n$ of all vectors with more than $n - 2\sqrt{n}$ ones. If $L \subseteq \{0,1\}^n$ consist of all vectors with at most $n/2 - \sqrt{n}$ ones, then $(L + L) \cap K = \emptyset$, that is, L is an independent set in G of size $|L| \geq 2^{n - O(\log n)}$. But any *linear* independent set L' in G is a vector space of dimension at most $n - 2\sqrt{n}$. Hence, $|L'| \leq 2^{n-2\sqrt{n}}$, and the gap $|L|/|L'|$ can be as large as $2^{\Omega(\sqrt{n})}$.

Note, however, that there is a big difference between the set K we constructed and the sets K_A arising form $(0,1,*)$-matrices A: generating sets K_A must be almost "closed downwards". In particular, if $\boldsymbol{x} \in K_A$ then *all* nonzero vectors, obtained from $\boldsymbol{x}$ by flipping some even number of its 1's to 0's, must also belong to K_A. Hence, this example does not refute the min-rank conjecture as such.

Example 4 (Bad matrices D_i)**.** Let us now look what happens if we allow the matrices $D_1, \ldots, D_m$ in the definition of a system of semi-linear equations (11) be *arbitrary* $n \times n$ $(0,1)$-matrices. A completion M of A can then be defined as a $(0,1)$-matrix with rows $\boldsymbol{m}_i = \boldsymbol{a}_i + \boldsymbol{\alpha}_i^\top D_i$. Now define $\mathrm{mr}(A|D_1, \ldots, D_r)$ as the minimal rank of such a completion of A. Observe that this definition coincides with the "old" min-rank, if we take the D_i's to be the diagonal matrices corresponding the stars in the ith row of A.

However, Example 3 shows that the min-rank conjecture is false in this generalized setting. To see why, we can define appropriate matrices $A, D_1, \ldots, D_m$ such that the corresponding set K_A defined by (12) consists of vectors with more than $n - 2\sqrt{n}$ ones: for an arbitrary vector $\boldsymbol{v}$ with more than $n - 2\sqrt{n}$ ones just define $\boldsymbol{a}_i$ and D_i such that the system $D_i \boldsymbol{x} = \boldsymbol{0}$, $\langle \boldsymbol{a}_i, \boldsymbol{x} \rangle = 1$ has $\boldsymbol{v}$ as its only solution.

Except of the obvious open problem to prove or disprove the linearization conjecture (Conjecture 1) or the min-rank conjecture (Conjecture 2), there are several more concrete problems.

We have shown (Lemma 10) that the gap between min-rank and row/column min-ranks may be as large as $\ln n$. It would be interesting to find $(0, 1, *)$-matrices A with larger gap.

Problem 1. *How large the gap* $\mathrm{mr}(A)/\max\{\mathrm{mr}_{\mathrm{col}}(A), \mathrm{mr}_{\mathrm{row}}(A)\}$ *can be?*

The next question concerns the clique number $\omega(G_A)$ of (that is, the largest number of vertices in) Cayley graphs G_A generated by the sets of the sets $K_A \subseteq \{0, 1\}^n$ of the form (12). By Remark 6, solutions for A are independent sets in this graph. Hence, $\mathrm{opt}(A)$ is just the independence number $\alpha(G_A)$ of this graph. Since in any N-vertex graph G we have that $\omega(G) \cdot \alpha(G) \leq N$, this yields $\mathrm{opt}(A) \leq 2^n/\omega(G_A)$. On the other hand, it is easy to see that $\omega(G_A) \leq 2^{\mathrm{rank}(M)}$, where M is a canonical completion of A obtained by setting all $*$'s to 0: If $C \subseteq \{0, 1\}^n$ is a clique in G_A, then we must have $M\boldsymbol{x} \neq M\boldsymbol{y}$ for all $\boldsymbol{x} \neq \boldsymbol{y} \in C$, because otherwise the vector $\boldsymbol{x} \oplus \boldsymbol{y}$ would not belong to K_A.

Problem 2. *Give an upper bound on* $\omega(G_A)$ *in terms of min-rank* $\mathrm{mr}(A)$ *of* A.

Finally, it would be interesting to eliminate an annoying requirement in Theorem 2 that the matrix A must be star-monotone.

Problem 3. *If* A *is an* r*-by-*n $(0, 1, *)$*-matrix of min-rank* r*, is then* $\mathrm{opt}(A) \leq 2^{n-r}$*?*

References

[1] N. Alon, On the rigidity of an Hadamard matrix, manuscript, 1990.

[2] N. Alon, P. Pudlák. Superconcentrators of depth 2 and 3; odd levels help (rarely), J. Comp. Sys. Sci. 48 (1994) 194–202.

[3] N. Alon, M. Karchmer, A. Wigderson, Linear circuits over GF(2), SIAM. J. Comput. 19(6) (1990) 1064–1067.

[4] D. Y. Cherukhin, The lower estimate of complexity in the class of schemes of depth 2 without restrictions on a basis, Moscow Univ. Math. Bull. 60(4) (2005) 42–44.

[5] N. Cohen, C. R. Johnson, L. Rodman, H. J. Woederman, Ranks of completions of partial matrices, in: H. Dym et al., editors, The Gohberg Anniversary Collection, Volume I, volume 40 of Operator Theory: Advances and Applications, Birkhäuser Verlag, Basel, 1989, pp. 165–185.

[6] D. Dolev, C. Dwork, N. Pippenger, A. Wigderson, Superconcentrators, generalizer and generalized connectors with limited depth, in: Proc. 15th STOC 1983 pp. 42–51.

[7] J. Friedman, A note on matrix rigidity, Combinatorica 13 (1993) 235–239.

[8] C. R. Johnson. Matrix completion problems: a survey, in C. R. Johnson, ed., Matrix Theory and Applications, Proc. of AMS Symp. in Applied Math., vol. 40 (1990) pp. 171–198.

[9] S. Jukna, Entropy of operators or why matrix multiplication is hard for depth-two circuits, Theory of Comput. Syst. (2008), doi 10.1007/s00224-008-9133-y.

[10] S. Lokam, On the rigidity of Vandermonde matrices, Theoretical Comput. Sci. 237(1-2) (2000) 477-483.

[11] S. Lokam, Spectral methods for matrix rigidity with applications to size-depth trade-offs and communication complexity, J. Comput. Syst. Sci. 63(3) (2001) 449-473.

[12] J. Morgenstern, Note on a lower bound on the linear complexity of fast Fourier transform, J. ACM 20(2) (1973) 305–306.

[13] J. Morgenstern, The linear complexity of computation, J. ACM 22(2) (1975) 184–194.

[14] N. Pippenger, Superconcentrators, SIAM J. Comput. 6 (1977) 298–304.

[15] N. Pippenger, Superconcentrators of depth 2, J. Comput. Syst. Sci. 24 (1982) 82–90.

[16] P. Pudlák, Communication in bounded depth circuits, Combinatorica 14(2) (1994) 203–216.

[17] P. Pudlák, Z. Vavrín, Computation of rigidity of order n^2/r for one simple matrix, Comment. Math. Univ. Carolinae 32 (1991) 213–218.

[18] P. Pudlák and P. Savický. On shifting networks. *Theoretical Comput. Sci.* 116: 415–419, 1993.

[19] R. Paturi, P. Pudlák, Circuit lower bounds and linear codes, J. Math. Sciences, 134(5) (2006) 2425–2434.

[20] P. Pudlák, V. Rödl, J. Sgall, Boolean circuits, tensor ranks, and communication complexity, SIAM J. Comput. 26(3) (1997) 605–633.

[21] J. Radhakrishnan, A. Ta-Shma, Bounds for dispersers, extractors, and depth-two superconcentrators, SIAM J. Discrete Math. 13(1) (2000) 2–24.

[22] R. Raz, A. Shpilka, Lower bounds for matrix product in bounded depth circuits with arbitrary gates, SIAM J. Comput. 32(2) (2003) 488–513.

[23] A. A. Razborov, On rigid matrices, manuscript, 1989 (in Russian).

[24] A. A. Razborov, B. Khasin, Improved lower bounds on the rigidity of Hadamard matrices, Mat. Zametki 63(4) (1998) 534–540 (in Russian).

[25] M. A. Shokrollahi, D. A. Spielman, V. Stetmann, A remark on matrix rigidity, Inform. Process. Letters 64(6): 283–285, 1997.

[26] R. de Wolf, Lower bounds on matrix rigidity via a quantum argument, in: 33rd Int. Colloq. on Automata, Languages and Programming (ICALP'06), in: Springer Lect. Notes in Comput. Sci., vol. 4051 (2006) pp. 62–71.

[27] L. Valiant, Graph-theoretic methods in low-level complexity, in: Proceedings of MFCS 1977, in: Springer Lect. Notes in Comput. Sci., vol. 53 (1977) pp. 162–176.

[28] L. Valiant, Why is boolean complexity theory difficult?, in: M. S. Paterson (Ed.), Boolean Function Complexity, Cambridge Univ. Press (1992) pp. 84-94.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de/