

ON BLOCKY RANKS OF MATRICES

DANIEL AVRAHAM AND AMIR YEHUDAYOFF

Abstract. A matrix is blocky if it is a “blowup” of a permutation matrix. The *blocky rank* of a matrix M is the minimum number of blocky matrices that linearly span M . Hambardzumyan, Hatami and Hatami defined blocky rank and showed that it is connected to communication complexity and operator theory. We describe additional connections to circuit complexity and combinatorics, and we prove upper and lower bounds on blocky rank in various contexts.

Keywords. Communication complexity, matrix rank, threshold circuits, fat matchings.

Subject classification. 68Q11, 68R10.

1. Introduction

Matrices serve as a model for many objects; linear operators in algebra, communication problems in computational complexity, concept classes in machine learning, and more. There are many ways to measure the complexity of matrices; there are various notions of rank (the “usual” rank, approximate rank, non-negative rank, sign rank, etc.), of communication complexity (deterministic, randomized, quantum, etc.), in learning theory (VC dimension, Littlestone dimension, margin complexity, etc.), and more. We focus on the notion of *blocky rank* recently defined by [Hambardzumyan, Hatami & Hatami \(2023\)](#).

A standard mechanism for defining a complexity measure has two stages. In the first stage, we define the *building blocks* of the model (in our case, matrices of blocky rank one). In the second stage, complexity is defined as the *minimum number of operations*

that are needed to generate the target (in our case, sum operations).

DEFINITION 1.1. *All identity matrices have blocky rank one. The set of matrices of blocky rank one is also closed under three operations: duplicating a row or a column, permuting the rows or columns, and adding a zero row or a zero column. In other words, a matrix has blocky rank one if up to a permutation of the rows and columns it has blocks of ones of different sizes on the “diagonal” followed by some amount of zeros.*

DEFINITION 1.2. *The blocky rank $\text{blocky}(M)$ of a matrix M is the minimum integer R so that M can be written as a linear combination of R matrices B_1, \dots, B_R , each of blocky rank one. In this work, we always work over the field \mathbb{R} .*

Motivation to study blocky rank, and its relatives, comes from various areas. In communication complexity, it is related to understanding randomized communication problems, deterministic communication complexity with equality oracle queries, and other communication complexity models (Hambardzumyan *et al.* (2023) and Pitassi, Shirley & Shraibman (2023)). In operator theory, it is related to idempotents in Schur algebras (see Hambardzumyan *et al.* (2023) and references within). In circuit complexity, it is related to depth-two threshold circuits. In combinatorics, it is related to covering problems in graphs. In machine learning, it is related to closure properties of Littlestone classes (Alon, Beimel, Moran & Stemmer (2020)).

1.1. Generic matrices. A typical first question about complexity is “what is the complexity of a random object?” The “obvious” upper bound on the blocky rank of an $n \times n$ Boolean matrix is n , because a Boolean matrix with one nonzero row has blocky rank one. The following theorem provides a lower bound for random matrices.

THEOREM 1.3. *If M is a uniformly random $n \times n$ Boolean matrix, then*

$$\Pr[\text{blocky}(M) \geq \frac{n}{4 \log(2n)}] \geq 1 - 2^{-\frac{n^2}{2}}.$$

[Theorem 1.3](#) is proved in [Section 2.1](#). The lower bound has a factor of $\log n$ compared to the obvious upper bound. This factor turns out to be needed. The blocky rank of *all* Boolean matrices is much smaller than n . This immediately follows from a result of [Pudlák & Rödl \(1994\)](#).¹ In fact, this holds for a large complexity measure. The *blocky partition number* $\text{blocky-par}(M)$ of M is the minimum number of blocky matrices that sum to M . It trivially holds that $\text{blocky}(M) \leq \text{blocky-par}(M)$.

THEOREM 1.4. *For every Boolean $n \times n$ matrix M ,*

$$\text{blocky-par}(M) \leq O\left(\frac{n}{\log n}\right).$$

The two theorems are reminiscent of Shannon's lower bounds and Lupanov's upper bound in the context of Boolean circuit complexity. Shannon proved that the circuit complexity of a random n -variate Boolean function is at least $\Omega\left(\frac{2^n}{n}\right)$, while the obvious upper bound is larger. Lupanov proved that in fact the lower bound is sharp; every n -variate Boolean function has a Boolean circuit of size at most $O\left(\frac{2^n}{n}\right)$.

The theorems above have the following additional combinatorial interpretation. The *clique cover* number of a graph is the least number of (induced) cliques that are required to cover it. The *intersection* number of a graph is the least k so that the graph can be represented as the intersection graphs over a universe of size k . An *intersection graph* consists of a set $S_v \subseteq [k]$ for each vertex v , so that every two vertices $u \neq v$ are connected by an edge iff $S_v \cap S_u \neq \emptyset$.

[Erdős, Goodman & Pósa \(1966\)](#) showed that the clique cover number is equal to the intersection number. [Bollobás, Erdős, Spencer & West \(1993\)](#) proved that the clique cover number of a uniformly random graph on n vertices is at least $\Omega\left(\frac{n^2}{\log^2 n}\right)$ and at most $O\left(\frac{n^2 \log \log n}{\log^2 n}\right)$. Part of their motivation was to understand the interval number of random graphs. [Frieze & Reed \(1995\)](#) improved the upper bound to a sharp $O\left(\frac{n^2}{\log^2 n}\right)$. Roughly speaking,

¹We had a weaker $O\left(\frac{n \log \log n}{\log n}\right)$ upper bound in a previous version of our paper. We thank an anonymous referee for informing us of the better bound.

the connection between their $\frac{n^2}{\log^2 n}$ and our $\frac{n}{\log n}$ is that a typical clique is of size $\log n$, and $\frac{n}{\log n}$ cliques can (sometimes) be glued to a single blocky matrix, so the total number of blocky matrices becomes $\approx \frac{n^2}{\log^2 n} / \frac{n}{\log n}$. It is worth noting that the upper bounds from Bollobás *et al.* (1993) and Frieze & Reed (1995) hold for random graphs and are false for some graphs, whereas the upper bound above holds for all matrices.

Let us now make the connection more formal. We work with bipartite graphs, because they correspond to (general) Boolean matrices.² A *blocky* graph is a bipartite graph that consists of a disjoint union of full bipartite graphs. Equivalently, the adjacency matrix of a blocky graph has blocky rank one. The *blocky cover number* of a bipartite graph is the minimum number of (induced) blocky graphs that are required to cover it.

The blocky cover number can be thought of as a variant of the intersection number. An *intersection representation* of a graph is a map s that assigns to each vertex v a vector $s(v) \in \{0, 1\}^k$, and two vertices $v \neq u$ are connected by an edge if there is $i \in [k]$ so that $s(v)_i = s(u)_i > 0$. The *intersection number* of a graph is the minimum k for which there is such a representation. We can extend this definition to larger alphabets. An *agreement representation* consists of a map s that assigns to each vertex v a vector $s(v) \in \{0, 1, \dots, L\}^k$, so that every two vertices $v \neq u$ are connected by an edge if there is $i \in [k]$ so that $s(v)_i = s(u)_i > 0$. The integer k is called the *universe size* of the graph (the integer L is not assumed to be bounded).

If G is a bipartite graph, then G has an agreement representation with universe size one iff G is blocky. More generally, the cover number of G using blocky graphs is equal to the least universe size of an agreement representation of G . This is analogous to the fact that the clique cover number is equal to the intersection number (Erdős *et al.* (1966)).

Instead of covering the edges of a graph, we can ask to partition them to structured parts. The *blocky partition number* of a

²The two color classes of the graph correspond to rows and columns in the matrix, and there is an edge between vertices i and j if the corresponding entry in the matrix is one.

bipartite graph is the minimum number of pairwise edge-disjoint (induced) blocky graphs that are required to cover it. [Theorem 1.4](#) says that the blocky partition number of *every* bipartite graph with n vertices is $O(\frac{n}{\log n})$.

1.2. The greater-than matrix. A more interesting but often more difficult question is understanding the complexity of specific objects (and not of random objects). We move to investigating the blocky rank of specific matrices.

The first matrix we consider is the $n \times n$ greater-than matrix GT_n defined by $\text{GT}_n(x, y) = 1_{x \leq y}$, where we think of the rows and columns of GT_n as integers in $[n]$. The greater-than matrix is the adjacency matrix of the *half graph*.

THEOREM 1.5. $\text{blocky}(\text{GT}_n) = \Theta(\log n)$.

The upper bound is relatively straightforward and was proved a long time ago in the context of Schur algebras ([Kwapień & Pełczyński \(1970\)](#)). It actually states that the blocky partition number of the half graph is at most $\lceil \log n \rceil + 1$; see [Claim 4.4](#) below. In particular, even the *monotone* blocky rank of GT_n is at most order $\log n$ (in monotone ranks, we only allow to use positive coefficients).

A variant of the blocky rank of the greater-than matrix was studied in the context of closure properties of “threshold classes” in machine learning ([Alon et al. \(2020\)](#); [Ghazi et al. \(2021\)](#)). There are many variants of blocky rank we can study: a monotone version where the linear combination just uses positive numbers, an approximate version where we just need to approximate the target matrix, a signed version where we just need to get the sign pattern correctly, and so forth. Here is a variant that is related to closure properties in machine learning. For a tuple $B = (B_1, \dots, B_R)$ of $n \times n$ Boolean matrices, and a function $F : \{0, 1\}^R \rightarrow \{0, 1\}$, let $F(B)$ be the $n \times n$ matrix obtained by applying F entry-wise: for all i, j ,

$$(F(B))_{i,j} = F((B_1)_{i,j}, \dots, (B_R)_{i,j}).$$

DEFINITION 1.6. *The functional blocky rank $\text{fun-blocky}(M)$ of a Boolean matrix M is the minimum number R so that there is a tuple $B = (B_1, \dots, B_R)$ of blocky matrices and $F : \{0, 1\}^R \rightarrow \{0, 1\}$ so that $M = F(B)$.*

The lower bound $\text{fun-blocky}(\text{GT}_n) \geq \Omega(\log \log n)$ is implicit in the work of Alon, Beimel, Moran & Stemmer (2020). The better lower bound $\text{fun-blocky}(\text{GT}_n) \geq \Omega(\frac{\log n}{\log \log n})$ is implicit in the work of Ghazi, Golowich, Kumar & Manurangsi (2021). These two works consider a more general framework, and their arguments are based on Ramsey theory. And even the stronger lower bound is off by a $\log n$ factor. We remove this factor, and get a sharp bound.

THEOREM 1.7. $\text{fun-blocky}(\text{GT}_n) \geq \frac{1}{8} \log n$.

The lower bound is proved in Section 3.2. This argument too is related to covering graphs. The Graham-Pollak theorem states that the edges of the full graph on n vertices cannot be *partitioned* to less than $n - 1$ complete bipartite graphs (Graham & Pollak (1971)). Orlin (1977) suggested to study the problem of *covering* the *cocktail party* graph (a full graph minus a perfect matching). Part of his motivation came from computational complexity theory (see Remark 3.7 in his paper). Gregory & Pullman (1982) proved that the clique cover number of the cocktail party graph is $\Theta(\log n)$.

We consider the following bipartite strengthening of their result. The *bipartite cocktail party graph* is the full bipartite graph minus a perfect matching.

THEOREM 1.8. *The blocky cover number of the bipartite cocktail party graph with n vertices on each side is at least $\frac{1}{4} \log n$.*

This is quantitatively weaker but more general than the lower bound of Gregory and Pullman. The cocktail party graph contains a copy of the bipartite cocktail party graph. And a clique in the cocktail party graph corresponds to a *connected* blocky graph in that copy. Our lower bound holds also when we are allowed to cover the bipartite cocktail party graph by any blocky graphs (not necessarily connected).

Avishay Tal shared with us the following observation. For every $n \times n$ Boolean matrix M , the functional blocky rank of M is at most

$O(\log n)$. The reason is that the matrix B defined by $B_{x,y} = x_i$ for some fixed i is blocky, and with $2\lceil \log n \rceil$ such matrices we can encode both x, y . This shows that, somewhat unusually, GT_n is an explicit matrix of essentially maximum functional blocky rank.

1.3. The inner-product matrix. The second matrix we consider is the inner-product matrix; let IP_n be the $\{0, 1\}^n \times \{0, 1\}^n$ matrix defined by $\text{IP}_n(x, y) = \sum_i x_i y_i \bmod 2$. This matrix has been studied in various contexts, like circuit complexity, communication complexity, margin complexity and more. We focus here on its connection to circuit complexity; in particular, to depth-two threshold circuits. There is a long line of research on this topic; see [Amano \(2020\)](#); [Hajnal *et al.* \(1993\)](#); [Kane & Williams \(2016\)](#); [Mahajan \(2019\)](#); [Paturi & Saks \(1994\)](#); [Roychowdhury *et al.* \(1994\)](#); [Williams \(2018\)](#) and references within.

A *linear threshold function* (LTF) function is of the form $T(z) = \text{sign}(b + \sum_i a_i z_i)$ for $a_1, \dots, a_n, b \in \mathbb{Z}$ where sign is 1 on $[0, \infty)$ and 0 on $(-\infty, 0)$. A *majority gate* is a special kind of LTFs in which all constants a_1, \dots, a_n are in $\{-1, 0, 1\}$. A $\text{MAJ} \circ \text{LTF}$ circuit computes a function of the form $D(z) = m(T_1(z), \dots, T_s(z))$ where each T_i is an LTF and m is a majority gate. The *size* of the circuit is $|D| = s$.

[Hajnal, Maass, Pudlák, Szegedy & Turán \(1993\)](#) proved a lower bound of roughly $2^{n/3}$ for the size of $\text{MAJ} \circ \text{LTF}$ circuits computing the inner product function. [Amano \(2020\)](#) constructed a $\text{MAJ} \circ \text{LTF}$ circuit of size $(1.899)^n$ computing the inner product function. The blocky rank perspective allows to improve the lower bound.

THEOREM 1.9. *Any $\text{MAJ} \circ \text{LTF}$ circuit computing IP_n has size at least $\Omega\left(\frac{2^{n/2}}{n}\right)$.*

The theorem is proved in [Section 4.1](#). The proof proceeds by bounding the correlation between IP_n and a threshold gate. An upper bound of $\approx 2^{-n/3}$ on the correlation was proved in [Hajnal *et al.* \(1993\)](#). We improve the bound to $\approx 2^{-n/2}$ which is basically sharp.

1.4. Depth-two threshold circuits. Finally, we describe a general connection between blocky rank and circuit complexity, specif-

ically, depth-two threshold circuits. A similar connection was already discovered by [Jukna \(2006\)](#).

Proving strong lower bounds for $\text{LTF} \circ \text{LTF}$ circuits is a long-standing open problem. [Kane & Williams \(2016\)](#) proved the best known lower bound for this model. They proved that the size of every $\text{LTF} \circ \text{LTF}$ circuit computing the n -variate Andreev function must be of size $\Omega(n^{3/2})$.

Roychowdhury, Orlitsky, and Siu observed that we do not even know how to prove lower bounds for $\Sigma \circ \text{LTF}$ circuits, where the upper gate just computes a linear function (with arbitrary coefficients); see [Roychowdhury et al. \(1994\)](#); [Williams \(2018\)](#). We observe that lower bounds on blocky rank yield circuit lower bounds in this model.

THEOREM 1.10. *Let M be a $\{0, 1\}^n \times \{0, 1\}^n$ matrix. If $M = \sum_{i=1}^s w_i T_i$ where each $w_i \in \mathbb{R}$ and each T_i is an LTF, then*

$$s \geq \frac{\text{blocky}(M)}{2(n+1)}.$$

The theorem is proved in [Section 4.2](#). It shows that proving strong lower bound on the blocky rank of explicit Boolean matrices might be difficult but rewarding.

REMARK 1.11. *A similar theorem holds for the signed version of blocky rank and general $\text{LTF} \circ \text{LTF}$ circuits.*

The theorem suggests that even proving relatively weak lower bounds (say, polynomial in n) on the blocky rank of an explicit $2^n \times 2^n$ matrix is interesting. The lower bound from [Kane & Williams \(2016\)](#) relies on the anti-concentration phenomenon, which does not seem directly relevant to blocky rank. So, even obtaining an $\Omega(n^{5/2})$ lower bound on the blocky rank (which would yield the same circuit lower bound) seems interesting to us.

2. General matrices

2.1. A lower bound for random matrices. In this sub-section we prove [Theorem 1.3](#); if M is a uniformly random $n \times n$ Boolean

matrix then

$$\Pr[\text{blocky}(M) \geq \frac{n}{4 \log(2n)}] \geq 1 - 2^{-\frac{n^2}{2}}.$$

The lower bound follows from a counting argument showing that there are few Boolean matrices with low blocky rank.

LEMMA 2.1. *If $V \subset \mathbb{R}^n$ is a linear subspace of dimension k , then*

$$|V \cap \{0, 1\}^n| \leq 2^k.$$

PROOF. We can choose a basis for V in echelon form. That is, there are $v_1, \dots, v_k \in V$ and $i_1 < \dots < i_k$ in $[n]$ so that $(v_j)_{i_j} = 1$ and $(v_j)_{i_\ell} = 0$ for all $\ell < j$. If $\sum_i a_i v_i \in \{0, 1\}^n$, it follows that given a_1, \dots, a_i , there are at most two possible options for a_{i+1} . The total number of possibilities for a_1, \dots, a_k is at most 2^k . \square

LEMMA 2.2. *For $n > 2$, the number of blocky matrices of size $n \times n$ is at most $\frac{1}{2}(2n)^{2n}$.*

PROOF. By permuting the rows and columns, every blocky matrix can be brought into a block diagonal form. A matrix in block diagonal form can be represented by two sets $\{i_1 < i_2 < \dots < i_r\}$ and $\{j_1 < j_2 < \dots < j_r\}$ in $[n]$ so that the first block is of size $i_1 \cdot j_1$, the second $(i_2 - i_1) \cdot (j_2 - j_1)$ and so on. The case when there are zero rows or columns is encoded by $i_r < n$ or $j_r < n$. There is at most $2^n \cdot 2^n = 2^{2n}$ ways to choose this representation and $n! \cdot n! \leq \frac{1}{2}n^{2n}$ ways to order the rows and columns. \square

PROOF (Theorem 1.3). For fixed blocky matrices B_1, \dots, B_R ,

$$|\text{span}\{B_1, \dots, B_R\} \cap \{0, 1\}^{n \times n}| \leq 2^R.$$

The number of Boolean matrices of blocky rank at most R is therefore at most $2^R (\frac{1}{2}(2n)^{2n})^R = (2n)^{2nR}$. If $R \leq \frac{n}{4 \log(2n)}$, then $(2n)^{2nR} \leq 2^{n^2/2}$. \square

2.2. An upper bound for all matrices. In this sub-section, we explain the proof of [Theorem 1.4](#); for every Boolean $n \times n$ matrix M ,

$$\text{blocky}(M) \leq O\left(\frac{n}{\log n}\right).$$

The proof is based on the work of Pudlák and Rödl [Pudlák & Rödl \(1994\)](#). In that paper, blocky matrices are called “fat matchings”. Theorem 11.1 in [Pudlák & Rödl \(1994\)](#) says that if M is a Boolean $n \times n$ matrix, then M can be covered by $R \leq O\left(\frac{n}{\log n}\right)$ blocky matrices B_1, \dots, B_R ; that is, $M \leq \sum_r B_r$ and for all $r \in [R]$ we have $B_r \leq M$. The proof of Theorem 11.1, however, shows that $M = \sum_r B_r$ so that $\text{blocky}(M) \leq R$.

3. Blocky ranks of greater-than

3.1. An upper bound on blocky rank. It is known that the blocky rank of GT_n is at most logarithmic; see, e.g. [Kwapień & Pełczyński \(1970\)](#). We include a proof for completeness.

CLAIM 3.1. *For all n ,*

$$\text{blocky}(\text{GT}_n) \leq \lceil \log n \rceil + 1.$$

PROOF. We prove the claim for GT_n for $n = 2^k$. The proof is by induction on k . For the base case $k = 0$, we have $\text{blocky}(G_1) \leq 1$. The matrix $\text{GT}_{2^{k+1}}$ can be written as

$$\text{GT}_{2^{k+1}} = \begin{pmatrix} \text{GT}_{2^k} & J \\ 0 & \text{GT}_{2^k} \end{pmatrix},$$

where J is the all-ones matrix. Let B_1, \dots, B_{k+1} be the matrices so that $\text{GT}_{2^k} = B_1 + \dots + B_{k+1}$. We can write $\text{GT}_{2^{k+1}}$ as

$$\text{GT}_{2^{k+1}} = \begin{pmatrix} B_1 & 0 \\ 0 & B_1 \end{pmatrix} + \dots + \begin{pmatrix} B_{k+1} & 0 \\ 0 & B_{k+1} \end{pmatrix} + \begin{pmatrix} 0 & J \\ 0 & 0 \end{pmatrix}.$$

□

3.2. A lower bound on functional blocky rank. In this section we prove [Theorem 1.7](#) stating that $\text{fun-blocky}(\text{GT}_n) \geq \frac{1}{8} \log n$. Let C be the $n \times n$ matrix with zeros on the diagonal and ones elsewhere; it corresponds to the bipartite cocktail party graph. The following lemma provides a lower bound on the blocky cover number of the bipartite cocktail party graph.

LEMMA 3.2. *If $C = \vee(B)$ where \vee denotes the OR function and $B = (B_1, \dots, B_R)$ is a tuple of $n \times n$ blocky matrices, then $R \geq \frac{1}{2} \log n$.*

PROOF ([Lemma 3.2](#)). The proof is by induction on n . For $n = 1$, the claim is trivial. For the inductive step, the ones of the matrix B_1 correspond to pairwise disjoint sets $S_1, \dots, S_A \subseteq [n]$ and $T_1, \dots, T_A \subseteq [n]$. That is, $(B_1)_{i,j} = 1$ iff $i \in S_a$ and $j \in T_a$ for some $a \in [A]$.

Define two random subsets S and T of $[n]$ as follows. Let $\epsilon_1, \dots, \epsilon_A$ be i.i.d. uniformly distributed in $\{0, 1\}$. Let S be the complement of $\bigcup_{a:\epsilon_a=1} S_a$ and T be the complement of $\bigcup_{a:\epsilon_a=0} T_a$. Let $I = \{i \in [n] : (i, i) \in S \times T\}$.

The projection of B_1 to $S \times T$ is the zero matrix (with probability one). For each $i \in [n]$, the probability that $(i, i) \in S \times T$ is one quarter, because $(B_1)_{i,i} = 0$. There is a choice for $S \times T$ so that $|I| \geq \frac{n}{4}$.

Let C' be the matrix C after deleting all rows and columns not in I . The matrix C' is a cocktail party matrix of dimension $|I|$, and the matrix B_1 does not contribute to its representation. The inductive hypothesis completes the proof. \square

PROOF ([Theorem 1.7](#)). Assume that $\text{GT}_n = F(B)$ for $B = (B_1, \dots, B_R)$ where each B_r is blocky. Assume towards a contradiction that $R < \frac{1}{8} \log n$. For $i, j \in [n]$, denote by $B_{i,j} \in \{0, 1\}^R$ the vector

$$B_{i,j} = ((B_1)_{i,j}, (B_2)_{i,j}, \dots, (B_R)_{i,j}) \in \{0, 1\}^R.$$

There is a set $I \subseteq [n]$ of size

$$m := |I| \geq \frac{n}{2^R} \geq \sqrt{n}$$

so that all the m vectors $B_{i,i}$ for $i \in I$ are the same. Delete the rest $n - m$ rows and columns from GT_n and from B , and focus on the remaining $m \times m$ part. Denote by G the obtained copy of GT_m , and denote by B' the obtained tuple of matrices so that $G = F(B')$. It follows that there is $f \in \{0, 1\}^R$ so that for each $i \in [m]$, we have $B'_{i,i} = f$. Order B' so that the first k entries in f are ones, and the last $R - k$ are zeros.

CLAIM 3.3. *For every $i \neq j$ in $[m]$, there is $r > k$ so that (exactly) one of $(B'_r)_{i,j}$ and $(B'_r)_{j,i}$ is one.*

PROOF. For each $r \leq k$, because B'_r is blocky and $(B'_r)_{i,i} = (B'_r)_{j,j} = 1$, we know that $(B'_r)_{i,j} = (B'_r)_{j,i}$. Because $G_{i,j} \neq G_{j,i}$, the two lists $((B'_{k+1})_{i,j}, \dots, (B'_R)_{i,j})$ and $((B'_{k+1})_{j,i}, \dots, (B'_R)_{j,i})$ must be different. \square

The matrix $\vee(B'_{k+1}, \dots, B'_R, B'^T_{k+1}, \dots, B'^T_R)$ is therefore zero on the diagonal and one elsewhere (where T denotes transposition). [Lemma 3.2](#) implies that $2(R - k) \geq \frac{1}{2} \log m$. \square

4. Circuit complexity

4.1. MAJ \circ LTF circuits. In this sub-section, we prove [Theorem 1.9](#); if $D \in \text{MAJ} \circ \text{LTF}$ and $D = \text{IP}_n$ then $|D| \geq \Omega(\frac{2^{n/2}}{n})$. We use the blocky rank perspective to prove circuit lower bounds for inner-product.

DEFINITION 4.1. *The nuclear norm of the matrix M is*

$$\|M\|_\nu = \inf \left\{ \sum_{i=1}^t p_i : M = \sum_{i=1}^t p_i A_i \quad \forall i \ p_i > 0, \text{rank}(A_i) = 1, \|A_i\|_\infty \leq 1 \right\}.$$

CLAIM 4.2. *If B is a blocky matrix, then $\|B\|_\nu \leq 1$.*

PROOF. The claim follows from the well-known fact that the nuclear norm of the unit matrix is at most one (see, e.g. [Hambarzumyan et al. \(2023\)](#)). We include a proof for completeness.

It is sufficient to consider $n \times n$ identity matrices for the case that $n + 1$ is prime. Then, for all $x, y \in [n]$,

$$1_{x=y} = \sum_{z \in \{0, 1, \dots, n\}} \frac{1}{n+1} \cdot e^{2\pi i(x-y)z/(n+1)}.$$

□

Consider the following generalization of LTFs.

DEFINITION 4.3. *An $[n] \times [n]$ matrix M is monotone if for every x and $y < y'$ in $[n]$,*

$$M_{x,y} \leq M_{x,y'}.$$

CLAIM 4.4. *If M is a Boolean $n \times n$ monotone matrix, then M is a sum of at most $\lceil \log n \rceil + 1$ blocky matrices.*

PROOF. Because duplicating rows and columns do not increase the blocky rank, the blocky rank of M is at most that of \mathbf{GT}_n . Now, use [Claim 3.1](#). □

COROLLARY 4.5. *The nuclear norm of an $n \times n$ Boolean monotone matrix is at most $\lceil \log(n) \rceil + 1$.*

The property of inner-product we rely on is Lindsey's lemma. This was done in many works, including [Hajnal *et al.* \(1993\)](#). The proof of the lemma uses the fact that the rows of \mathbf{IP}_n are orthogonal using the Cauchy–Schwarz inequality.

LEMMA 4.6 (Lindsey's lemma). *If M is a $\{0, 1\}^n \times \{0, 1\}^n$ matrix of rank one so that $\|M\|_\infty \leq 1$, then*

$$\left| \sum_{x,y \in \{0,1\}^n} (-1)^{\mathbf{IP}_n(x,y)} M(x,y) \right| \leq 2^{\frac{3n}{2}}.$$

We can conclude the following strengthening of the correlation bound from [Hajnal *et al.* \(1993\)](#).

LEMMA 4.7. *If T is an LTF in the $2n$ variables $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ then*

$$\left| \sum_{x,y \in \{0,1\}^n} (-1)^{\mathbb{IP}_n(x,y)} T(x,y) \right| \leq (n+1)2^{\frac{3n}{2}}.$$

PROOF. The matrix T is monotone (up to a permutation of the rows and columns). [Corollary 4.5](#) bounds the nuclear norm of T from above; we can write

$$T = \sum_i p_i B_i$$

where each $p_i > 0$, where each B_i is of rank one and $\|B_i\|_\infty \leq 1$, and where $\sum_i p_i \leq (n+1)$. [Lemma 4.6](#) implies

$$\begin{aligned} \left| \sum_{x,y \in \{0,1\}^n} (-1)^{\mathbb{IP}_n(x,y)} T(x,y) \right| &= \left| \sum_{x,y \in \{0,1\}^n} (-1)^{\mathbb{IP}_n(x,y)} \sum_i p_i B_i(x,y) \right| \\ &\leq \sum_i p_i \left| \sum_{x,y \in \{0,1\}^n} (-1)^{\mathbb{IP}_n(x,y)} B_i(x,y) \right| \\ &\leq (n+1)2^{\frac{3n}{2}}. \end{aligned}$$

□

PROOF ([Theorem 1.9](#)). Assume that

$$\mathbb{IP}_n(x,y) = \text{sign}\left(-b + \sum_{i=1}^s w_i T_i(x,y)\right)$$

where each T_i is an LTF and $w_i \in \{-1, 0, 1\}$. It follows that $|b| \leq s$ because otherwise the right-hand side is constant. For all x, y ,

$$(-1)^{1+\mathbb{IP}_n(x,y)} \left(1 + 2\left(-b + \sum_i w_i T_i(x,y)\right)\right) \geq 1.$$

Summing over all x, y ,

$$\left| \sum_{x,y} (-1)^{\mathbb{IP}_n(x,y)} \left(1 + 2\left(-b + \sum_i w_i T_i(x,y)\right)\right) \right| \geq 2^{2n}.$$

Lemma 4.7 implies that

$$\begin{aligned} & \left| \sum_{x,y} (-1)^{\text{IP}_n(x,y)} \left(1 + 2 \left(-b + \sum_i w_i T_i(x,y) \right) \right) \right| \\ & \leq 2(s+1)2^n + 2 \cdot \sum_i w_i \left| \sum_{x,y} (-1)^{\text{IP}_n(x,y)} T_i(x,y) \right| \\ & \leq s \cdot 2(n+2)2^{\frac{3n}{2}}. \end{aligned}$$

□

4.2. Σ oLTF circuits. In this sub-section, we prove Theorem 1.10; if M is a $\{0,1\}^n \times \{0,1\}^n$ matrix so that $M = \sum_{i=1}^s w_i T_i$ where each $w_i \in \mathbb{R}$ and each T_i is an LTF then

$$s \geq \frac{\text{blocky}(M)}{2(n+1)}.$$

In other words, blocky rank lower bounds imply circuit lower bounds.

PROOF (Theorem 1.10). If $M = \sum_{i=1}^s w_i T_i$ then by Claim 4.4 we have

$$\text{blocky}(M) \leq \sum_i \text{blocky}(T_i) \leq s \cdot (n+1).$$

□

Funding Open access funding provided by Technion - Israel Institute of Technology.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- NOGA ALON, AMOS BEIMEL, SHAY MORAN & URI STEMMER (2020). Closure properties for private classification and online prediction. In *COLT*, 119–152.
- KAZUYUKI AMANO (2020). On the size of depth-two threshold circuits for the inner product mod 2 function. In *LATA*, 235–247.
- BÉLA BOLLOBÁS, PAUL ERDŐS, JOEL SPENCER & DOUGLAS B WEST (1993). Clique coverings of the edges of a random graph. *Combinatorica* **13**(1), 1–5.
- PAUL ERDŐS, ADOLPH W GOODMAN & LOUIS PÓSA (1966). The representation of a graph by set intersections. *Canadian Journal of Mathematics* **18**, 106–112.
- ALAN FRIEZE & BRUCE REED (1995). Covering the edges of a random graph by cliques. *Combinatorica* **15**(4), 489–497.
- BADIH GHAZI, NOAH GOLOWICH, RAVI KUMAR & PASIN MANURANGSI (2021). Near-tight closure bounds for the Littlestone and threshold dimensions. In *ALT*, 686–696.
- RONALD L GRAHAM & HENRY O POLLAK (1971). On the addressing problem for loop switching. *The Bell system technical journal* **50**(8), 2495–2519.
- DAVID A GREGORY & NORMAN J PULLMAN (1982). On a clique covering problem of Orlin. *Discrete Mathematics* **41**(1), 97–99.
- ANDRÁS HAJNAL, WOLFGANG MAASS, PAVEL PUDLÁK, MARIO SZEGEDY & GYÖRGY TURÁN (1993). Threshold circuits of bounded depth. *Journal of Computer and System Sciences* **46**(2), 129–154.
- LIANNA HAMBARDZUMYAN, HAMED HATAMI & POOYA HATAMI (2023). Dimension-free bounds and structural results in communication complexity. *Israel Journal of Mathematics* **253**(2), 555–616.

STASYS JUKNA (2006). On graph complexity. *Combinatorics, Probability and Computing* **15**(6), 855–876.

DANIEL M KANE & RYAN WILLIAMS (2016). Super-linear gate and super-quadratic wire lower bounds for depth-two and depth-three threshold circuits. In *STOC*, 633–643.

STANISŁAW KWAPIEŃ & ALEKSANDER PEŁCZYŃSKI (1970). The main triangle projection in matrix spaces and its applications. *Studia Mathematica* **34**(1), 43–67.

MEENA MAHAJAN (2019). Depth-2 threshold circuits. *Resonance* **24**(3), 371–380.

JAMES ORLIN (1977). Contentment in graph theory: covering graphs with cliques. In *Indagationes Mathematicae*, volume 80, 406–424.

RAMAMOHAN PATURI & MICHAEL E SAKS (1994). Approximating threshold circuits by rational functions. *Information and Computation* **112**(2), 257–272.

TONIANN PITASSI, MORGAN SHIRLEY & ADI SHRAIBMAN (2023). The strength of equality oracles in communication. In *ITCS*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik.

PAVEL PUDLÁK & VOJTECH RÖDL (1994). Some combinatorial-algebraic problems from complexity theory. *Discrete Mathematics* **1**(136), 253–279.

VWANI ROYCHOWDHURY, KAI-YEUNG SIU & ALON ORLITSKY (1994). Neural models and spectral methods. In *Theoretical Advances in Neural Computation and Learning*, 3–36. Springer.

RICHARD RYAN WILLIAMS (2018). Limits on representing boolean functions by linear combinations of simple functions: thresholds, reLUs, and low-degree polynomials. In *CCC*, 1–24.

Manuscript received 8 April 2023

DANIEL AVRAHAM

Department of Mathematics

Technion-IIT

Haifa, Israel

`kspy-d@campus.technion.ac.il`

AMIR YEHUDAYOFF

Department of Mathematics

Technion-IIT

Haifa, Israel

Department of Computer Science

Copenhagen University

Copenhagen, Denmark

`amir.yehudayoff@gmail.com`