Electronic Colloquium on Computational Complexity, Report No. 306 (2026)

Exponential Separation of Sign-Rank and Margin Complexity

Gabriel Hart* Kaave Hosseini*

Abstract

We construct an explicit family of total sign matrices whose sign-rank is exponentially larger than their margin complexity. Previously, no superpolynomial separation was known for total matrices. In particular, we give a quantitative answer to a question of Linial, Mendelson, Schechtman, and Shraibman (Combinatorica, 2007), who asked whether the logarithmic dependence on matrix size in the Johnson–Lindenstrauss upper bound for sign-rank is necessary. Our result shows that this dependence on matrix size is nearly tight if the dependence on margin complexity is required to remain polynomial. The construction is the OR of Equality tests on a large enough alphabet. This also exponentially improves sign-rank lower bounds for sparse disjointness and Hamming-distance thresholds.

The proof uses a new one-sided version of Forster's spectral bound for sparse matrices that gives exponentially stronger lower bounds than previous lower bounds. Using the same method, we also obtain direct-sum theorems for unbounded-error communication complexity for OR compositions of expanders and set-membership tests.

1 Introduction

The *sign-rank* of a 0/1 matrix F, denoted by $\mathrm{rank}^{\pm}(F)$, is the minimum rank of a real matrix whose entries are positive where F is one and negative where F is zero. Geometrically, sign-rank is the least dimension in which the rows can be represented by points and the columns by homogeneous halfspaces with the prescribed incidence pattern. Paturi and Simon introduced sign-rank in communication complexity and showed that the logarithm of sign-rank characterizes unbounded-error communication complexity [PS86], denoted by $\mathrm{UPP}(F)$; namely, $\log(\mathrm{rank}^{\pm}(F)) = \mathrm{UPP}(F) \pm O(1)$. In this model of communication, Alice and Bob have private randomness and must output the correct answer with probability strictly greater than 1/2.

One of the simplest functions with respect to sign-rank is *Equality* (i.e., the identity matrix). The equality function $\mathsf{EQ}_n(x, y) = \mathbf{1}[x = y]$ on n-bit strings, satisfies $\mathrm{rank}^{\pm}(\mathsf{EQ}_n) = 3$ for every $n \geq 2$ (see Section 2). Moreover, Equality is closed under conjunction, namely, the AND of k equality tests on k pairs of n-bit blocks is simply equality on the concatenated kn-bit strings. Hence, its sign-rank remains three, independent of k. Motivated by several applications discussed later, we ask: *what happens to sign-rank if AND is replaced by OR?*

Let $\mathsf{OR}_k \circ \mathsf{EQ}_n$ denote the OR of k equality tests, so that $\mathsf{OR}_k \circ \mathsf{EQ}_n(x, y) = 1$ if and only if $x_i = y_i$ for at least one $i \in [k]$. Unlike AND, the case of $\mathsf{OR}_k \circ \mathsf{EQ}_n$ does not reduce to a single equality test on the concatenated inputs. The $\mathsf{OR}_k \circ \mathsf{EQ}_n$ function has a long history in communication complexity, beginning with the work of Mehlhorn and Schmidt [MS82]. For example the same function, as Block-Equality, was used by Impagliazzo and Williams [IW10] in communication complexity separations, and by Chattopadhyay, Mahajan, Mande, and Saurabh [CMMS20] to prove lower bounds for linear decision lists. Its bounded-error public-coin randomized communication complexity is $\Theta(k)$ for

*Department of Computer Science, University of Rochester.
Emails: ghart3@ur.rochester.edu, kaave.hosseini@rochester.edu.

ISSN 1433-8092

$n \geq 2$, independently of the block length n [HW07, KS92]. For unbounded-error communication, however, the known sign-rank bounds left an exponential gap.

Known lower-bound methods for sign-rank give a lower bound of $\mathrm{rank}^{\pm}(\mathsf{OR}_k \circ \mathsf{EQ}_n) \geq \Omega(k)$ when $n \geq 2$. On the other hand, an elementary polynomial construction gives an upper bound of 3^k on its sign-rank, independently of the block length n, leaving an exponential gap between the upper and lower bounds. See Section 2 for these lower and upper bounds. Our first result shows that the sign-rank is indeed exponential in k when n is sufficiently large.

Theorem 1. *For every integer $k \geq 1$ and $n \geq 10k2^k$,*

$$2^k \leq \mathrm{rank}^{\pm}(\mathsf{OR}_k \circ \mathsf{EQ}_n) < 2^{k+1}.$$

The upper bound holds for every $n \geq 1$.

Remark 2. It is impossible to retain a similar lower bound if n is subexponential in k. Indeed, an application of the Johnson–Lindenstrauss lemma gives an upper bound of $\mathrm{rank}^{\pm}(\mathsf{OR}_k \circ \mathsf{EQ}_n) \leq O(k^3 n)$ (see Equation (3)). Hence, if $n = 2^{o(k)}$, one obtains a $\mathrm{rank}^{\pm}(\mathsf{OR}_k \circ \mathsf{EQ}_n) \leq 2^{o(k)}$.

Note that for every integer $0 < t < k$ and $n \geq 10t2^t$, we also have $\mathrm{rank}^{\pm}(\mathsf{OR}_k \circ \mathsf{EQ}_n) \geq 2^t$. Indeed, fix the last $k - t$ pairs of blocks to unequal values and apply Theorem 1 to the first t blocks. Consequently, for $2 \leq n < 10k2^k$, these bounds imply $\Omega(n/\log n) \leq \mathrm{rank}^{\pm}(\mathsf{OR}_k \circ \mathsf{EQ}_n) \leq O(k^3 n)$.

Theorem 1 can be viewed as a direct-sum lemma for this function: for $n \geq 10k2^k$,

$$\mathrm{UPP}(\mathsf{OR}_k \circ \mathsf{EQ}_n) = k \pm O(1). \tag{1}$$

To prove Theorem 1, we develop a new lower-bound on sign-rank because previous lower-bound methods provably cannot produce a lower bound larger than $O(k^2)$ for $\mathrm{rank}^{\pm}(\mathsf{OR}_k \circ \mathsf{EQ}_n)$; see Section 1.2. Here, it is convenient to work with $\{-1, 1\}$-valued matrices. We write $\|A\|_{\mathrm{op}} = \sup_{\|v\|_2=1} \|Av\|_2$ for the operator norm. In other words, $\|A\|_{\mathrm{op}}$ is the largest singular value of A. Forster's lower bound on sign-rank [For02] states that for every sign matrix $F \in \{-1, 1\}^{M \times N}$,

$$\mathrm{rank}^{\pm}(F) \geq \frac{\sqrt{MN}}{\|F\|_{\mathrm{op}}}. \tag{2}$$

However, this bound can not give a lower bound larger than $\Theta(k)$ for sign-rank of $\mathsf{OR}_k \circ \mathsf{EQ}_n$. See Section 1.2. To circumvent this issue, based on a simple observation, we show that for sparse matrices it is enough to have a "one-sided spectral witness" to get large sign-rank. Namely, find a matrix A that satisfies $A_{ij} \leq -1$ when $F_{ij} = -1$, while allowing arbitrary signs where $F_{ij} = 1$. This relaxed one-sided condition allows us to construct a witness with smaller operator norm which can exponentially improve on Forster's bound.

Lemma 3 (One-sided spectral witness bound (informal))**.** *Let $F \in \{-1, 1\}^{M \times N}$ be a sign matrix, and let $A \neq 0$ be a real matrix of the same size satisfying $A_{ij} \leq -1$ whenever $F_{ij} = -1$. If F is sparse enough (i.e., has few ones), then*

$$\mathrm{rank}^{\pm}(F) \geq \frac{\sqrt{MN}}{\|A\|_{\mathrm{op}}}.$$

For the precise statement and the sparsity requirement on F, see Lemma 13.

1.1 Applications

Here we discuss the main motivations and applications to study Theorem 1 and Lemma 3.

Communication complexity. An important analytic parameter of a Boolean matrix F is its *margin*, the largest $\gamma > 0$ for which there are vectors of norms at most one, in an unrestricted dimension, satisfying the following inequalities:

$$\begin{cases} \langle u_x, v_y \rangle \geq \gamma, & F(x,y) = 1, \\ \langle u_x, v_y \rangle \leq -\gamma, & F(x,y) = 0. \end{cases}$$

Here $\langle u_x, v_y \rangle$ denotes the Euclidean inner product which lies in $[-1, 1]$. The reciprocal of γ, namely $1/\gamma$, is called *margin complexity* and is denoted by $\mathrm{mc}(F)$. Linial and Shraibman [LS09] proved that margin complexity is equivalent, up to a constant factor, to inverse discrepancy which is a main lower bound technique in randomized communication. If $R^{\mathrm{pub}}_{1/3}(F)$ denotes public-coin randomized communication complexity with error at most $1/3$, then

$$\Omega(\log \mathrm{mc}(F)) \leq \mathrm{R}^{\mathrm{pub}}_{1/3}(F) \leq O(\mathrm{mc}(F)^2).$$

An important question is how large sign-rank can be when margin complexity is small. Ben-David, Eiron, and Simon [BDES02, Corollary 20] and also Linial, Mendelson, Schechtman, and Shraibman [LMSS07] showed that Johnson–Lindenstrauss dimension reduction [JL84] gives a general upper bound: every $N \times N$ 0/1 matrix with $N \geq 2$ satisfies

$$\mathrm{rank}^{\pm}(F) \leq O(\mathrm{mc}(F)^2 \log N). \tag{3}$$

They further asked the following question:

Question 4 ([LMSS07, p. 460])**.** Is the $\log N$ factor in Equation (3) necessary?

This problem has received a lot of attention [HHL22, HHP$^+$22, HHM23, HH24, HZ24, GHIS25, FHV26]. Hatami, Hosseini, and Meng [HHM23] answered Question 4 for partial matrices: using the Borsuk–Ulam theorem, they constructed partial $N \times N$ matrices from the Gap Hamming family with margin complexity $O(1)$ and sign-rank $\Theta(\log N)$. However, the case of total matrices has remained a challenge. As an immediate application of Theorem 1, we give a quantitative answer to Question 4 for total matrices. It is easy to see that $\mathrm{mc}(\mathsf{OR}_k \circ \mathsf{EQ}_n) = O(k)$. (See Section 2.) Also, the matrix $\mathsf{OR}_k \circ \mathsf{EQ}_n$ has order $N = 2^{kn}$, giving the following corollary.

Corollary 5. *For every sufficiently large integer N, there is an explicit $N \times N$ total 0/1 matrix F satisfying*

$$\mathrm{mc}(F) = O(\log \log N), \qquad \mathrm{rank}^{\pm}(F) = \Omega\left(\frac{\log N}{(\log \log N)^2}\right).$$

This means that in Equation (3) reducing the exponent of $\log N$ by any fixed $\varepsilon > 0$ forces at least exponential dependence on $\mathrm{mc}(F)$. Namely, for every fixed $\varepsilon > 0$, no bound of the form

$$\mathrm{rank}^{\pm}(F) \leq 2^{o(\mathrm{mc}(F))} (\log N)^{1-\varepsilon}$$

holds for all total matrices.

A qualitative version of this question, formulated explicitly in [HHP$^+$22, Question 1.4] and repeated in several recent works, asks whether there is any function g such that, $\mathrm{rank}^{\pm}(F) \leq g(\mathrm{mc}(F))$ independently of matrix size. Corollary 5 shows that such a function, if it exists, must be at least exponential. It is also worth mentioning a complementary line of work, initiated by Babai, Frankl, and Simon [BFS86] and developed over the years [BVdW07, She08, She11, She13, Tha16, She21], to construct matrices with small sign-rank but large margin complexity. This

direction led to construction of a $2^n \times 2^n$ matrix F with sign-rank three and margin complexity $2^{\Omega(n)}$ [HHL22, ACHS24]. Hence giving a separation of $\mathrm{UPP}(F) = O(1)$ and $\mathrm{R}^{\mathrm{pub}}_{1/3}(F) \geq \Omega(n)$.

We end this discussion with a question. The exponential separation of Corollary 5 does not transfer to the communication models, since $\mathrm{R}^{\mathrm{pub}}_{1/3}(\mathsf{OR}_k \circ \mathsf{EQ}_n) = \Theta(k)$ [HW07, KS92].[HW07, KS92].

Question 6. Is there a family of total Boolean matrices $(F_k)_{k\geq 2}$ such that

$$\mathrm{R}^{\mathrm{pub}}_{1/3}(F_k) = O(\log k) \qquad \text{and} \qquad \mathrm{UPP}(F_k) = \Omega(k)?$$

A positive answer would show that shared randomness can yield an exponential saving in communication even when private-coin protocols are allowed an arbitrarily small advantage over random guessing. Note that, Newman's theorem [New91] implies that every $N \times N$ matrix F satisfies

$$\mathrm{UPP}(F) \leq \mathrm{R}^{\mathrm{pub}}_{1/3}(F) + O(\log\log N). \tag{4}$$

If such a family can be chosen with F_k of size $2^{2^{O(k)}}$, it would show that the additive $O(\log\log N)$ overhead in Newman's theorem is optimal even when the private-coin protocol is allowed unbounded error.

Learning theory. In learning theory, sign-rank is known as *dimension complexity*. Margin-dependent mistake bounds date back to the perceptron analyses of the 1960s [Blo62, Nov62], while dimension complexity and its relation to VC dimension were studied since the 1980s [AHW87, AMY16]. The roles of embedding dimension and margin in learning via halfspaces were studied systematically by Ben-David, Eiron, and Simon [BDES02].

Let $\mathcal{C} \subseteq \{0,1\}^{\mathcal{X}}$ be a binary concept class, and consider the evaluation function $\Phi : \mathcal{X} \times \mathcal{C} \to \{0,1\}$ given by $\Phi(x,c) = c(x)$. The dimension complexity of $\mathcal{C}$ is the least dimension of a common feature map realizing all concepts by homogeneous linear thresholds, hence precisely the sign-rank of Φ. An explicit feature map realizing a concept class by linear thresholds gives a PAC learning algorithm via linear programming, with sample complexity linear in the dimension for fixed error and confidence. On the other hand, a feature representation with margin at least $1/k$ in any dimension allows the hard-margin support vector machine (SVM) algorithm to PAC learn the class using $O(k^2)$ samples (for fixed error and confidence), independently of the feature dimension; see [KMS20, Section 3.2]. However, it was not known whether dimension complexity can also be bounded by a polynomial of k.

Let $\mathcal{X} = \mathbb{Z}^k$ and $\mathcal{C}_k = \{c_y : y \in \mathcal{X}\}$, where

$$c_y(x) = \mathbf{1}[\exists i \in [k] : x_i = y_i].$$

Using similar construction for $\mathsf{OR}_k \circ \mathsf{EQ}_n$ it is easy to show that margin complexity of this class is at most $2k+1$, see Section 2. Nevertheless, from Theorem 1 it follows that the dimension complexity of this concept class is at least 2^k. As far as we know, this is the first concept class showing that sample complexity of hard margin SVM can be exponentially smaller than dimension complexity.

Direct-sum for OR compositions of expanders. Let $B \in \{0,1\}^{n\times n}$ have exactly D ones in every row and column, where $1 \leq D < n$, and put $\lambda = \|B - (D/n)J\|_{\mathrm{op}}$. If B is the adjacency matrix of an undirected graph, λ is the largest absolute value of an eigenvalue on the subspace orthogonal to the constant vector. For $D \leq n/2$, Alon, Moran, and Yehudayoff [AMY16] showed that

$\mathrm{rank}^{\pm}(B) \geq D/\lambda$. As another application of Lemma 3, we extend this bound to OR compositions of such sparse matrices. Define

$$(\mathsf{OR}_k \circ B)(x, y) = \mathbf{1}[\exists i \in [k] : B(x_i, y_i) = 1], \qquad x, y \in [n]^k.$$

For sufficiently large n, Theorem 14 gives

$$\left(\frac{D}{\lambda}\right)^k \leq \mathrm{rank}^{\pm}(\mathsf{OR}_k \circ B) \leq (3D)^k.$$

Consequently, if $\lambda = O(\sqrt{D})$, then

$$\frac{k}{2}\,\mathrm{UPP}(B) - O(k) \leq \mathrm{UPP}(\mathsf{OR}_k \circ B) \leq 2k\ \mathrm{UPP}(B) + O(k).$$

The same approach applies to OR compositions of the set-membership problem [MNSW98]. For $x \in [n]^k$ and $\mathbf{S} = (S_1, \ldots, S_k) \in \binom{[n]}{D}^k$, define

$$L_{n,k,D}(x, \mathbf{S}) = \mathbf{1}[\exists i \in [k] : x_i \in S_i].$$

Corollary 15 shows that, for fixed $k, D \geq 1$ and all sufficiently large n,

$$\mathrm{rank}^{\pm}(L_{n,k,D}) = (D+1)^{\Theta(k)}.$$

Sparse disjointness and Hamming-distance thresholds. Sparse set disjointness $\mathsf{DISJ}_{k,n}$ asks whether two subsets $A, B \subseteq [n]$, each of size at most k, are disjoint. This function has been studied extensively in randomized communication complexity [HW07, DKS12, ST13, BCK+16, HPZZ21], where $\mathrm{R}^{\mathrm{pub}}_{1/3}(\mathsf{DISJ}_{k,n}) = \Theta(k)$ for $n \geq k$ [HW07, KS92, Raz92]. For unbounded-error communication, however, there was an exponential gap between lower and upper bounds of $\Omega(\log k)$ and $O(k)$. An immediate corollary of Theorem 1 gives tight bounds.

Corollary 7. *For every integer $k \geq 1$ and $n \geq k2^{10k2^k}$,*

$$k - O(1) \leq \mathrm{UPP}(\mathsf{DISJ}_{k,n}) \leq 3k + O(1).$$

The upper bound holds for every $n \geq 1$.

The same corollary applies to Kneser graphs which is the special case of $\mathsf{DISJ}_{k,n}$ where sets have size exactly k and have been significantly studied in topological combinatorics since the seminal work of Lovasz [Lov78].

The lower bound also immediately transfers to a harder problem of Hamming-distance thresholds. Let $\mathrm{HD}^d_{<k}$ be the 0/1 matrix on $\{0,1\}^d \times \{0,1\}^d$ whose entry is one exactly when $d_H(x, y) < k$. It was conjectured that even for $k = 2$, the sign-rank has to grow with the dimension d [HHP+22, HHM23]. However, Göös, Harms, Imbach, and Sokolov [GHIS25] proved the dimension-independent upper bound

$$\sup_{d \geq 1} \mathrm{rank}^{\pm}(\mathrm{HD}^d_{<k}) \leq 2^{4k} + 1$$

using *support-rank*. Proving an exponential lower bound on sign-rank remained open. The following corollary is also an immediate consequence of Theorem 1.

Corollary 8 (Hamming-distance thresholds)**.** *For every integer $k \geq 2$ and all sufficiently large d,*

$$\mathrm{rank}^{\pm}(\mathrm{HD}^d_{<k}) \geq 2^{\lfloor k/2 \rfloor}.$$

Arbitrary predicates. Chattopadhyay and Mande [CM22, Theorems 5.1 and 7.1] studied the sign-rank of a natural decision list that outputs the parity of the index of the last successful equality test, or 0 if none succeeds. For k tests on blocks of length $k^{1/3} + \log k$, they determined its sign-rank to be $2^{\Theta(k^{1/3})}$. Our result complements theirs in the regime of larger block lengths: for sufficiently large n, Theorem 1 gives a lower bound of $2^{\lceil k/2 \rceil}$ for their function. Indeed, fixing every even-indexed pair of blocks to unequal values restricts this decision list to $\mathsf{OR}_{\lceil k/2 \rceil} \circ \mathsf{EQ}_n$.

More generally, Hatami et al. [HHP+22, Theorem 3.8] proved that $\mathrm{rank}^{\pm}(f \circ \mathsf{EQ}_n) \leq 4^k$ for every predicate $f : \{0,1\}^k \to \{0,1\}$ and every $n \geq 1$. It is natural to ask whether a complexity measure of f roughly characterizes $\mathrm{rank}^{\pm}(f \circ \mathsf{EQ}_n)$ for sufficiently large n. We show that such a characterization would necessarily have to capture arbitrary matrix sign-rank: every Boolean matrix B can be encoded by a predicate f_B such that $\mathrm{rank}^{\pm}(f_B \circ \mathsf{EQ}_n) = \Theta(\mathrm{rank}^{\pm}(B))$ for every $n \geq 1$. We give the construction and proof in Section 4.3.

1.2 Technical motivations

We briefly discuss limitations of prior lower bounds on sign-rank. For more details, see [HHP+22] *VC dimension* is a lower bound on sign-rank but for every sign matrix F, the VC-dimension bound is at most $\mathrm{mc}(F)^2$. Forster's lower bound on sign-rank is at most $\mathrm{mc}(F)$, since average margin is at least margin. Forster's spectral refinements, including the Razborov–Sherstov bound [RS10] and its applications via the pattern-matrix method of Sherstov [She11], face the same $\mathrm{mc}(F)$ barrier. Since $\mathrm{mc}(\mathsf{OR}_k \circ \mathsf{EQ}_n) \leq O(k)$ by the margin bound given in the background Section 2, these methods give at most $O(k^2)$.

Another approach lower-bounds sign-rank using the size of the largest monochromatic rectangle [APP+05], but its quantitative limitations compared to margin complexity is not understood. However, it gives only $\mathrm{rank}^{\pm}(\mathsf{OR}_k \circ \mathsf{EQ}_n) \geq \Omega(k)$. It is easy to see that under every product distribution on the rows and columns, there is a monochromatic rectangle of measure at least 2^{-2k-1}. There is a recent topological framework to lower bound sign-rank of arbitrary sign-matrices [FHV26] that can beat the previous bounds in some settings. However, this approach also is not able to give a lower bound better than $O(k^2)$ for $\mathrm{rank}^{\pm}(\mathsf{OR}_k \circ \mathsf{EQ}_n)$ [FH26]

Our one-sided witness bound of Lemma 3 avoids the sign-agreement requirement on the one entries of the matrix which can allow for exponentially stronger lower bounds in sparse settings.

Idea of proof of Lemma 3. Forster's istoropic normalization gives a real matrix B' with the same sign pattern as F with total squared mass $MN/\mathrm{rank}^{\pm}(F)$. Using a covering argument and sparsity of F we zero out a small fraction of rows and columns of B', leaving a matrix B whose positive entries are small and whose squared mass is almost unchanged. The one-sided condition on A then shows that $\sum_{i,j} A_{ij}B_{ij}$ is at least $MN/\mathrm{rank}^{\pm}(F)$ minus a small error. On the other hand, it is easy to see that by Cauchy-Schwarz, $\sum_{i,j} A_{ij}B_{ij} \leq \|A\|_{\mathrm{op}}\sqrt{MN}$ which implies the final bound of $\mathrm{rank}^{\pm}(F) \geq \frac{\sqrt{MN}}{\|A\|_{\mathrm{op}}}$.

Organization. Section 2 reviews preliminaries and basic upper and lower bounds on sign-rank of $\mathsf{OR}_k \circ \mathsf{EQ}_n$. Section 3 proves the improved upper bounds. Section 4 proves one-sided spectral lower bound for sign-rank and its applications.

2 Preliminaries

Let $F \in \{0,1\}^{M\times N}$. The sign-rank of F, denoted by $\mathrm{rank}^{\pm}(F)$, is the minimum rank of a real matrix $H \in \mathbb{R}^{M\times N}$ satisfying

$$H_{ij} > 0 \quad \text{if } F_{ij} = 1, \qquad H_{ij} < 0 \quad \text{if } F_{ij} = 0.$$

Such a matrix H is called a *realization* of F. Multiplying H by -1 shows that $\mathrm{rank}^{\pm}(F) = \mathrm{rank}^{\pm}(\overline{F})$. Sign-rank cannot increase upon restricting the row or column index sets.

Sign-rank has fundamental applications in learning theory [Fel17, KS01, FGV21, KS07, SSBD14, SS05, She08], circuit complexity [BT16, RS10, SW19], discrete and real algebraic geometry [AFR85, EMRPS14, FGL$^+$12, FPS$^+$17, Suk16], communication complexity [CM18, HHL22, She08, She12], and metric dimensionality reduction and Banach space theory [Mat96, Nao18].

We discuss the elementary upper and lower bounds on the sign-rank of our main matrices.

Equality and conjunction. Identify $\{0,1\}^n$ with $[2^n]$. Let $\mathsf{EQ}_n(x,y) = \mathbf{1}[x = y]$. The polynomial

$$H(x,y) = \frac{1}{2} - (x-y)^2 = -x^2 + 2xy + (\tfrac{1}{2} - y^2)$$

is positive when $x = y$ and negative otherwise. It is the sum of three rank-one matrices, so $\mathrm{rank}^{\pm}(\mathsf{EQ}_n) \le 3$. The standard matching lower bound for equality on an alphabet of size at least four gives $\mathrm{rank}^{\pm}(\mathsf{EQ}_n) = 3$ for $n \ge 2$.

An elementary upper bound for sign-rank of $\mathsf{OR}_k \circ \mathsf{EQ}_n$. The above idea can be generalized to OR of equalities. Identify the n-bit strings with the integers $0, \dots, 2^n - 1$, so the inputs are $x, y \in \{0, \dots, 2^n - 1\}^k$. Consider

$$H(x,y) = \frac{1}{2} - \prod_{i=1}^{k} (x_i - y_i)^2.$$

If some $x_i = y_i$, the product vanishes and $H(x,y) = 1/2$. Otherwise every squared difference is a positive integer, so $H(x,y) \le -1/2$. Thus H realizes $\mathsf{OR}_k \circ \mathsf{EQ}_n$. On expanding the product, H is a sum of at most 3^k rank-one matrices, giving

$$\mathrm{rank}^{\pm}(\mathsf{OR}_k \circ \mathsf{EQ}_n) \le 3^k. \tag{5}$$

Theorem 1 improves it to $2^{k+1} - 1$.

An elementary lower bound for sign-rank of $\mathsf{OR}_k \circ \mathsf{EQ}_n$. Identify the n-bit strings with the integers $0, \dots, 2^n - 1$. For $n \ge 2$, choose the k columns $y^{(1)}, \dots, y^{(k)}$, where $y^{(i)}$ has a 1 in coordinate i and 0 elsewhere. Restrict the rows to $x \in \{1,2\}^k$. Since x and $y^{(i)}$ can agree only in coordinate i,

$$\mathsf{OR}_k \circ \mathsf{EQ}_n(x, y^{(i)}) = 1 \quad \iff \quad x_i = 1.$$

Choosing each x_i independently realizes every binary pattern on these k columns. Thus these columns are shattered by the rows, giving VC dimension at least k. Since VC dimension is a lower bound on sign-rank,

$$\mathrm{rank}^{\pm}(\mathsf{OR}_k \circ \mathsf{EQ}_n) \ge k.$$

Margin complexity. For $F \in \{0,1\}^{M\times N}$, its margin complexity $\mathrm{mc}(F)$ is the reciprocal of the largest $\gamma > 0$ for which there exist vectors $u_1, \ldots, u_M$ and $v_1, \ldots, v_N$ in $\mathbb{R}^d$ for arbitrarily large d, of Euclidean norm at most one satisfying

$$\langle u_i, v_j\rangle \geq \gamma \quad (F_{ij} = 1), \qquad \langle u_i, v_j\rangle \leq -\gamma \quad (F_{ij} = 0).$$

As with sign-rank, margin complexity is invariant under complementation and cannot increase under restriction to a submatrix. It follows from John's theorem from convex geometry that one always has $1 \leq \mathrm{mc}(F) \leq \sqrt{\mathrm{rank}(F)+1}$ [LMSS07] where $\mathrm{rank}(F)$ is rank over the reals. Moreover, it is known [LS09] that margin complexity is equivalent to inverse of *discrepancy* in communication complexity up to a constant multiplicative factor.

For EQ_n, let e_x be the standard basis vector indexed by $x \in [2^n]$ and set $u_x = \frac{(\sqrt{2}e_x, 1)}{\sqrt{3}}, v_y = \frac{(\sqrt{2}e_y, -1)}{\sqrt{3}}$. Then

$$\langle u_x, v_y\rangle = \frac{2\mathbf{1}[x=y]-1}{3},$$

so $\mathrm{mc}(\mathsf{EQ}_n) \leq 3$.

Similar construction can be generalized to OR of equality tests. The vectors

$$u_x = (\sqrt{2}e_{x_1}, \ldots, \sqrt{2}e_{x_k}, 1), \qquad v_y = (\sqrt{2}e_{y_1}, \ldots, \sqrt{2}e_{y_k}, -1)$$

have squared norm $2k+1$ and inner product $2|\{i : x_i = y_i\}|-1$. Normalizing gives margin $1/(2k+1)$, hence margin complexity at most $2k+1$. A lower bound of $\Omega(k)$ is also easy to obtain but we skip it here since we don't need it. For a proof using discrepancy see [SZ09, Section 4.2].

3 Sign-rank Upper Bounds

In this section we establish the improved sign-rank upper bounds. We first consider OR compositions of Boolean matrices with few ones in each row, which includes $\mathsf{OR}_k \circ \mathsf{EQ}_n$ as a special case, and then give a separate bound for sparse disjointness.

Let $B \in \{0,1\}^{u\times v}$ have at most D ones in each row. The matrix $\mathsf{OR}_k \circ B$ has rows indexed by $x \in [u]^k$, columns indexed by $y \in [v]^k$, and entries

$$(\mathsf{OR}_k \circ B)(x,y) = \bigvee_{i=1}^{k} B(x_i, y_i).$$

An easy upper bound generalizing the polynomial construction given in Section 2 for $\mathsf{OR}_k \circ \mathsf{EQ}_n$ is

$$\mathrm{rank}^{\pm}(\mathsf{OR}_k \circ B) \leq (2D+1)^k.$$

which we leave as an exercise. The following lemma improves this upper bound to $2(D+1)^k - 1$. Our proof refines the classical sign-change construction of Alon, Frankl, and Rödl [AFR85]; see also [AMY16, Lemma 18]. They observed that if in a sign-matrix, each row has at most s sign-changes, then each row can be realized by a univariate polynomial of degree at most s. Indeed, choose a root between integers i and $i+1$ if the sign changes from column i to $i+1$. For B, since each row has D many ones, there are at most $2D+1$ sign changes, hence $\mathrm{rank}^{\pm}(B) \leq 2D+1$. To handle the more general case of $\mathsf{OR}_k \circ B$, we need to be more careful and space the column locations and the roots geometrically, that is associate column i to the value T^i for some large enough $T > 1$ and choose roots of the polynomial between these values. This will allow us to start with one polynomial construction for one B and add more B's one at a time until we arrive at $\mathsf{OR}_k \circ B$.

Lemma 9 (OR compositions of sparse rows)**.** *Let $k, D, u, v \geq 1$ be integers, and let $B \in \{0,1\}^{u \times v}$ have at most D ones in every row. Then*

$$\operatorname{rank}^{\pm}(\mathsf{OR}_k \circ B) \leq 2(D+1)^k - 1.$$

In particular, since EQ_n has exactly one 1 in each row, for every $n \geq 1$,

$$\operatorname{rank}^{\pm}(\mathsf{OR}_k \circ \mathsf{EQ}_n) \leq 2^{k+1} - 1.$$

Proof. We first construct a rational function for each row that is small at the zero entries of B and at least $1/2$ at its one entries. We then clear denominators to obtain polynomials.

A construction for one copy of B. For $T > 1$ and integer column labels $b, c \in [v]$, observe that

$$\frac{T^{b+c}}{T^{2b} + T^{2c}} = \frac{1}{T^{b-c} + T^{c-b}}.$$

This quantity equals $1/2$ when $b = c$ and is at most $1/T$ when $b \neq c$, since then $|b - c| \geq 1$.

For each row a, sum this expression over the columns c where $B(a, c) = 1$, defining

$$R_T(a, b) = \sum_{\substack{c \in [v] \\ B(a,c)=1}} \frac{T^{b+c}}{T^{2b} + T^{2c}}.$$

If $B(a, b) = 0$, all terms are at most $1/T$, and there are at most D terms. If $B(a, b) = 1$, the term $c = b$ contributes $1/2$. Thus

$$B(a,b) = 0 \implies R_T(a,b) \leq \frac{D}{T}, \qquad B(a,b) = 1 \implies R_T(a,b) \geq \frac{1}{2}.$$

We clear the denominators by defining $E_T, F_T \in \mathbb{R}^{u \times v}$ as

$$E_T(a,b) = \prod_{\substack{c \in [v] \\ B(a,c)=1}} (T^{2b} + T^{2c}), \qquad F_T(a,b) = E_T(a,b) R_T(a,b).$$

For $T > 4D$, the matrix $F_T - \frac{1}{4} E_T$ is a sign realization of B.

Adding one copy of B at a time. We construct a sign realization H_j of $\mathsf{OR}_j \circ B$, with rows indexed by $[u]^j$ and columns indexed by $[v]^j$. Start with the one-entry matrix $H_0 = [-1]$.

Suppose H_{j-1} has already been constructed. Multiplying it by a positive scalar, we may assume that

$$\max_{x,y} |H_{j-1}(x,y)| \leq \frac{1}{4}.$$

This changes neither its signs nor its rank. Fix this rescaled matrix, let

$$m = \min_{x,y} |H_{j-1}(x,y)| > 0,$$

and choose a new parameter $T = T_j > 1$ sufficiently large that $D/T < m$.

For $x \in [u]^{j-1}$, $y \in [v]^{j-1}$, $a \in [u]$, and $b \in [v]$, define

$$H_j((x,a),(y,b)) = E_T(a,b) H_{j-1}(x,y) + F_T(a,b).$$

Since $E_T(a,b) > 0$, the sign of this entry is the sign of

$$H_{j-1}(x,y) + R_T(a,b).$$

If $B(a,b) = 0$, the added quantity lies in $[0,m)$, so it preserves the previous sign. If $B(a,b) = 1$, then

$$H_{j-1}(x,y) + R_T(a,b) \geq -\frac{1}{4} + \frac{1}{2} > 0.$$

Hence H_j realizes $\mathsf{OR}_j \circ B$.

Bounding the rank. Fix the parameter T chosen at the current stage, and regard $z = T^b$ as the column variable. For each fixed row a, the expression

$$E_T(a,b) = \prod_{\substack{c\in[v]\\ B(a,c)=1}} (z^2 + T^{2c})$$

is a polynomial in z^2 of degree at most D.

Expanding $F_T = E_T R_T$ gives

$$F_T(a,b) = z \sum_{\substack{c\in[v]\\ B(a,c)=1}} T^c \prod_{\substack{d\in[v]\\ B(a,d)=1\\ d\neq c}} (z^2 + T^{2d}).$$

Thus $F_T(a,b)$ is z times a polynomial in z^2 of degree at most $D-1$.

Consequently, every row of E_T is a linear combination of the same $D+1$ column functions

$$1,\ T^{2b}, \ldots, T^{2Db},$$

and every row of F_T is a linear combination of the same D column functions

$$T^b,\ T^{3b}, \ldots, T^{(2D-1)b}.$$

Therefore

$$\operatorname{rank} E_T \leq D+1, \qquad \operatorname{rank} F_T \leq D.$$

Let $J_{p,q}$ denote the $p \times q$ all-ones matrix. The definition of H_j can be written as

$$H_j = H_{j-1} \otimes E_T + J_{u^{j-1},v^{j-1}} \otimes F_T.$$

Using subadditivity of rank and multiplicativity under tensor products,

$$\operatorname{rank} H_j \leq (\operatorname{rank} H_{j-1})(\operatorname{rank} E_T) + \operatorname{rank} F_T \leq (D+1)\operatorname{rank} H_{j-1} + D.$$

Since $\operatorname{rank} H_0 = 1$, induction gives

$$\operatorname{rank}^{\pm}(\mathsf{OR}_k \circ B) \leq \operatorname{rank} H_k \leq 2(D+1)^k - 1.$$

Finally, EQ_n has exactly one one in each row, so taking $D = 1$ gives the claimed bound $2^{k+1} - 1$. □

For sparse disjointness, we can use further symmetry to get direct and better bounds.

Lemma 10 (Sparse disjointness)**.** *For all integers $k, n \geq 1$,*

$$\operatorname{rank}^{\pm}(\mathsf{DISJ}_{k,n}) \leq \binom{3k}{k} \leq 2^{3k}.$$

Proof. Pad both sets to size k using disjoint pools of fresh elements, and identify the enlarged ground set with distinct integers. Define

$$H(A,B) = \prod_{a \in A} \prod_{b \in B} (a-b)^2 - \frac{1}{2}.$$

The product is zero when A and B intersect, and is a positive integer otherwise. Thus H realizes disjointness.

Fix the row set A. As a polynomial in the k column variables, $H(A,B)$ is symmetric and has degree at most $2k$ in each variable. Group monomials that differ only by a permutation of these variables. Each group is a function of B multiplied by a coefficient depending only on A, and hence gives a rank-one matrix. The groups are indexed by

$$0 \leq e_1 \leq \cdots \leq e_k \leq 2k,$$

so there are $\binom{3k}{k}$ of them. Therefore $\operatorname{rank}^{\pm}(\mathsf{DISJ}_{k,n}) \leq \operatorname{rank} H \leq \binom{3k}{k} \leq 2^{3k}$. □

4 Sign-rank Lower Bounds

We first discuss the lower bound method and then apply it to specific constructions mentioned in the introduction. We give some notation first. We use $\langle \cdot, \cdot \rangle$ for the usual inner product of vectors and, for matrices of the same size, write $\langle A, B \rangle = \sum_{i,j} A_{ij} B_{ij}$. We write $\| \cdot \|$ for the Euclidean norm of vectors, $\| \cdot \|_{\mathrm{op}}$ for the operator norm of matrices, and $\|A\|_F = (\sum_{i,j} A_{ij}^2)^{1/2}$ for the Frobenius norm.

4.1 A one-sided spectral lemma

The main ingredient in our lower bound is the following isotropic realization lemma.

Lemma 11 (Isotropic realization [Bar98, For02])**.** *If $\operatorname{rank}^{\pm}(F) = r$, there are unit vectors $u_x, v_y \in \mathbb{R}^r$ realizing F such that*

$$\mathbb{E}_{x,y} |\langle u_x, v_y \rangle|^2 = \frac{1}{r}.$$

We next show that, for a sparse matrix, deleting a small fraction of rows and columns makes all positive inner products in a low-dimensional realization small.

Lemma 12. *Let $F \in \{0,1\}^{X \times Y}$ have at most $\eta |X||Y|$ entries equal to 1, where $0 < \eta < 1$, and let $H(x,y) = \langle u_x, v_y \rangle$ be a unit-vector realization in $\mathbb{R}^r$. For $0 < \tau < 1$, there are sets $E \subseteq X$ and $G \subseteq Y$ such that*

$$\frac{|E|}{|X|} + \frac{|G|}{|Y|} \leq 2\sqrt{\eta (1 + 4/\tau)^r},$$

and

$$0 < H(x,y) \leq \tau \quad \text{whenever } F(x,y) = 1,\ x \notin E,\ y \notin G.$$

Proof. Using a $\tau/2$-net of the unit sphere, partition the row vectors into at most $N_\tau = (1+4/\tau)^r$ cells of diameter at most τ, see [Ver18, Corollary 4.2.13]. If $H(x,y) > \tau$, then every row x' in the same cell as x satisfies $H(x',y) \geq H(x,y) - \|u_{x'} - u_x\| > 0$. Thus column y of F has a one at every row in that cell.

We discard small cells of rows and columns containing many ones. For a parameter $\alpha > 0$ to be chosen below, let E consist of all rows in cells containing at most $\alpha|X|$ rows, and let G consist of all columns containing more than $\alpha|X|$ ones. If $x \notin E$ and $H(x,y) > \tau$, the observation above forces column y to contain more than $\alpha|X|$ ones, so $y \in G$. This proves $0 < H(x,y) \leq \tau$ whenever $F(x,y) = 1,\ x \notin E,\ y \notin G$.

There are at most N_τ cells, and each cell contributing to E contains at most $\alpha|X|$ rows. Thus $\frac{|E|}{|X|} \leq \alpha N_\tau$. Every column in G contains more than $\alpha|X|$ ones, so

$$\alpha|X|\,|G| \leq \sum_{x,y} F(x,y) \leq \eta|X||Y|, \qquad \text{hence} \qquad \frac{|G|}{|Y|} \leq \frac{\eta}{\alpha}.$$

Combining these bounds gives $\frac{|E|}{|X|} + \frac{|G|}{|Y|} \leq \alpha N_\tau + \frac{\eta}{\alpha}$. Choosing $\alpha = \sqrt{\eta/N_\tau}$ makes the right-hand side equal to $2\sqrt{\eta N_\tau}$, as required. □

We now combine these two lemmas to obtain a sign-rank lower bound from a one-sided witness.

Lemma 13 (One-sided spectral witness)**.** *Let $d \geq 2$ be an integer and $L \geq 1$. Suppose $F \in \{0,1\}^{M\times N}$ has at most $(dL)^{-6d}MN$ entries equal to 1. If there exists a real matrix $A \in \mathbb{R}^{M\times N}$ such that*

$$\begin{aligned}
&(1) \quad A_{ij} \leq -1 && \text{whenever } F_{ij} = 0,\\
&(2) \quad \frac{1}{MN}\sum_{i,j} |A_{ij}| \leq L,\\
&(3) \quad \|A\|_{\mathrm{op}} \leq \sqrt{MN}/d,
\end{aligned}$$

then $\mathrm{rank}^{\pm}(F) \geq d$.

Proof of Lemma 13. Suppose, for a contradiction, that $r = \mathrm{rank}^{\pm}(F) \leq d-1$. By Lemma 11, choose a unit-vector realization $H(i,j) = \langle u_i, v_j\rangle$ with $\sum_{i,j} H(i,j)^2 = \frac{MN}{r}$.

We first make the positive entries of this realization small by zeroing a small fraction of rows and columns. Set $\eta = (dL)^{-6d}$ and $\tau = (dL)^{-3}$. Lemma 12 gives exceptional row and column sets E, G whose total fraction is at most $\delta = 2\sqrt{\eta(1+4/\tau)^r} \leq 2(dL)^{-3}$, where the inequality follows from $1 + 4/\tau \leq (dL)^6$ and $r \leq d-1$.

Let B be obtained from H by setting the rows in E and the columns in G to zero. Then $-1 \leq B_{ij} \leq 0$ when $F_{ij} = 0$, while $0 \leq B_{ij} \leq \tau$ when $F_{ij} = 1$. Since $|H(i,j)| \leq 1$,

$$\sum_{i,j} B_{ij}^2 \geq \sum_{i,j} H(i,j)^2 - \delta MN = MN\left(\frac{1}{r} - \delta\right).$$

The remaining entries at positions where $F_{ij} = 1$ contribute at most $\eta\tau^2 MN$. Consequently, $\sum_{i,j:F_{ij}=0} B_{ij}^2 \geq MN\left(\frac{1}{r} - \delta - \eta\tau^2\right)$.

We now use this to lower-bound $\langle A, B\rangle$. When $F_{ij} = 0$, assumption (1) gives $A_{ij}B_{ij} \geq |B_{ij}| \geq B_{ij}^2$. When $F_{ij} = 1$, we have $A_{ij}B_{ij} \geq -\tau|A_{ij}|$. Thus, by assumption (2),

$$\langle A, B\rangle \geq \sum_{i,j:F_{ij}=0} B_{ij}^2 - \tau\sum_{i,j}|A_{ij}| \geq MN\left(\frac{1}{r} - \delta - \eta\tau^2 - L\tau\right).$$

On the other hand, zeroing rows and columns of H amounts to replacing the corresponding vectors u_i and v_j by zero. We therefore still have $B = U'V'^{\mathsf{T}}$, with every row of U' and V' of norm at most one. Their Frobenius norms thus satisfy $\|U'\|_F \le \sqrt{M},, \|V'\|_F \le \sqrt{N}$. By Cauchy-Schwarz, we get

$$|\langle A, B\rangle| = |\langle AV', U'\rangle| \le \|AV'\|_F \|U'\|_F \le \|A\|_{\mathrm{op}} \|V'\|_F \|U'\|_F \le \frac{MN}{d}.$$

Comparing the two bounds, we obtain $\frac{1}{r} - \frac{1}{d} \le \delta + \eta\tau^2 + L\tau$. But $r \le d-1$ makes the left-hand side at least $1/(d(d-1))$, whereas

$$\delta + \eta\tau^2 + L\tau \le 2(dL)^{-3} + (dL)^{-6d-6} + L(dL)^{-3} < \frac{4}{d^3} \le \frac{1}{d(d-1)}.$$

This is a contradiction. □

4.2 Applications

We use the Lemma 13 to obtain our lower bound results. The witnesses are built by tensor products, so we briefly discuss the notation here. For matrices A and B, the tensor product $A \otimes B$ has one row for each pair of rows of A and B, and one column for each pair of columns. Its entries are $(A \otimes B)_{(i,i'),(j,j')} = A_{ij}B_{i'j'}$. Similarly, the k-fold tensor power $A^{\otimes k}$ has rows and columns indexed by k-tuples of row and column indices of A, with $(A^{\otimes k})_{(i_1,\ldots,i_k),(j_1,\ldots,j_k)} = \prod_{t=1}^{k} A_{i_t j_t}$. The operator norm is multiplicative under tensor products (see [HJ91, Theorem 4.2.15]):

$$\|A \otimes B\|_{\mathrm{op}} = \|A\|_{\mathrm{op}} \|B\|_{\mathrm{op}}, \qquad \|A^{\otimes k}\|_{\mathrm{op}} = \|A\|_{\mathrm{op}}^{k}.$$

OR of Equalities

Theorem 1. *For every integer $k \ge 1$ and $n \ge 10k2^k$,*

$$2^k \le \mathrm{rank}^{\pm}(\mathsf{OR}_k \circ \mathsf{EQ}_n) < 2^{k+1}.$$

The upper bound holds for every $n \ge 1$.

Proof. The upper bound follows from Lemma 9. For the lower bound, identify the alphabet with $[2^n]$ and take

$$A = -\left(J_{2^n} - 2^{n-1}\mathrm{Id}_{2^n}\right)^{\otimes k}, \qquad d = 2^k, \qquad L = (3/2)^k.$$

Every zero entry of $\mathsf{OR}_k \circ \mathsf{EQ}_n$ has all coordinates different, so the corresponding entry of A is -1. The eigenvalues of $J_{2^n} - 2^{n-1}\mathrm{Id}_{2^n}$ are $\pm 2^{n-1}$, giving

$$\|A\|_{\mathrm{op}} = 2^{(n-1)k} = \frac{2^{nk}}{d}, \qquad \frac{1}{2^{2nk}} \sum_{x,y} |A(x,y)| = \left(\frac{3}{2} - 2^{1-n}\right)^k \le L.$$

The one-density is at most $k/2^n$. Since

$$\log_2 k + 6 \cdot 2^k \log_2(3^k) = \log_2 k + 6k2^k \log_2 3 \le 10k2^k,$$

the assumed lower bound on n ensures $k/2^n \le (dL)^{-6d}$. Lemma 13 now gives $\mathrm{rank}^{\pm}(\mathsf{OR}_k \circ \mathsf{EQ}_n) \ge 2^k$. □

Corollary 5. *For every sufficiently large integer N, there is an explicit $N \times N$ total $0/1$ matrix F satisfying*

$$\mathrm{mc}(F) = O(\log\log N), \qquad \mathrm{rank}^{\pm}(F) = \Omega\Big(\frac{\log N}{(\log\log N)^2}\Big).$$

Proof of Corollary 5. For $k \geq 1$, put $n_k = 10k2^k$ and $N_k = 2^{kn_k}$. By Theorem 1 and the margin bound in Section 2, the matrix $\mathsf{OR}_k \circ \mathsf{EQ}_{n_k}$ has order N_k, margin complexity $O(k)$, and sign-rank at least 2^k. Given a sufficiently large integer N, choose the largest k such that $N_k \leq N$, and duplicate rows and columns to obtain an $N \times N$ matrix F. Duplication preserves both sign-rank and margin complexity. Since

$$10k^2 2^k \leq \log_2 N < 10(k+1)^2 2^{k+1} \leq 80k^2 2^k,$$

we have $k = \Theta(\log\log N)$ and $2^k = \Omega(\log N/(\log\log N)^2)$, proving the claim. □

Sparse disjointness

Corollary 7. *For every integer $k \geq 1$ and $n \geq k2^{10k2^k}$,*

$$k - O(1) \leq \mathrm{UPP}(\mathsf{DISJ}_{k,n}) \leq 3k + O(1).$$

The upper bound holds for every $n \geq 1$.

Proof of Corollary 7. The lower bound follows by associating each word $x \in (\{0,1\}^b)^k$ with the set $A_x = \{(1,x_1),\ldots,(k,x_k)\} \subseteq [k] \times \{0,1\}^b$. Then $\mathsf{OR}_k \circ \mathsf{EQ}_b(x,y) = 1$ exactly when $A_x \cap A_y \neq \varnothing$. Thus the complement of $\mathsf{OR}_k \circ \mathsf{EQ}_b$ is a submatrix of sparse disjointness on a universe of size $k2^b$. Taking $b = 10k2^k$ and applying Theorem 1 gives the lower bound whenever $n \geq k2^{10k2^k}$.

The upper bound follows from Lemma 10. □

Hamming-distance thresholds

Corollary 8 (Hamming-distance thresholds)**.** *For every integer $k \geq 2$ and all sufficiently large d,*

$$\mathrm{rank}^{\pm}(\mathrm{HD}^d_{<k}) \geq 2^{\lfloor k/2 \rfloor}.$$

Proof of Corollary 8. Put $t = \lfloor k/2 \rfloor$ and $b = 10t2^t$. As in the preceding proof, associate each word $x \in (\{0,1\}^b)^t$ with the set $A_x = \{(i,x_i) : i \in [t]\} \subseteq [t] \times \{0,1\}^b$. For the characteristic vectors of two such t-sets, $d_H(\mathbf{1}_A, \mathbf{1}_B) = 2t - 2|A \cap B|$. When k is odd, append one coordinate on which the row and column inputs differ. The resulting distance is less than k exactly when $A \cap B \neq \varnothing$, so $\mathsf{OR}_t \circ \mathsf{EQ}_b$ is a submatrix of $\mathrm{HD}^{d_0}_{<k}$, where $d_0 = t2^{10t2^t} + k - 2t$. Since $b = 10t2^t$, Theorem 1 gives sign-rank at least 2^t. Appending zero coordinates to both inputs proves the same bound for every $d \geq d_0$. □

Direct-sum bounds on expanders. Another application of Lemma 13 is a generalization of the lower bound on sign-rank of $\mathsf{OR}_k \circ \mathsf{EQ}_n$ that can be viewed as a direct-sum result. Let $B \in \{0,1\}^{n \times n}$ have exactly D ones in every row and column, where $1 \leq D < n$. Equivalently, B is the biadjacency matrix of a D-regular bipartite graph with n vertices on each side. For $k \geq 1$, define

$$(\mathsf{OR}_k \circ B)(x,y) = \mathbf{1}[\exists i \in [k] : B(x_i,y_i) = 1], \qquad x,y \in [n]^k,$$

and put $\lambda = \|B - (D/n)J\|_{\mathrm{op}}$.

Theorem 14 (OR composition of regular relations)**.** *Let B, D, n, λ be as above, and let $k \geq 1$. If $n \geq 2^{(6D)^{k+1}}$, then*

$$\left(\frac{D}{\lambda}\right)^k \leq \mathrm{rank}^{\pm}(\mathsf{OR}_k \circ B) \leq (3D)^k.$$

The upper bound holds for any $n \geq 1$.

Proof. The upper bound follows from Lemma 9. Since $B - (D/n)J$ has rank at most $n-1$ and squared Frobenius norm $D(n-D)$, $\lambda^2 \geq \frac{D(n-D)}{n-1} \geq 1$. Put $\alpha = 1 + \frac{D}{\lambda}, \qquad \beta = 1 + \frac{2D}{\lambda}, \qquad d = \lfloor \alpha^k \rfloor$. Regularity gives $\lambda \leq D$, so $\alpha \geq 2$ and $d \geq 2$. Set $Z = B - \frac{D}{n}J, \qquad Q = \frac{1}{n}J - \frac{1}{\lambda}Z$. Every row and column of Z sums to zero. With respect to the constant and mean-zero subspaces on both sides, Q is the direct sum of the identity on constants and $-Z/\lambda$ on their orthogonal complements. Hence $\|Q\|_{\mathrm{op}} = 1$. At every nonedge (a, b), $Q(a, b) = \frac{1 + D/\lambda}{n} = \frac{\alpha}{n}$. Moreover,

$$\begin{aligned}\frac{1}{n}\sum_{a,b} |Q(a,b)| &\leq 1 + \frac{1}{n\lambda}\sum_{a,b}\left|B(a,b) - \frac{D}{n}\right| \\ &= 1 + \frac{2D(n-D)}{n\lambda} \leq \beta.\end{aligned}$$

Take $A = -\frac{n^k}{d}Q^{\otimes k}, \qquad L = \frac{\beta^k}{d}$. Since $d \leq \alpha^k \leq \beta^k$, we have $L \geq 1$. At every zero entry of $\mathsf{OR}_k \circ B$ all coordinate pairs are nonedges, and therefore $A(x, y) = -\frac{\alpha^k}{d} \leq -1$. The operator norm and the sum of absolute entries satisfy

$$\|A\|_{\mathrm{op}} = \frac{n^k}{d}, \qquad \frac{1}{n^{2k}}\sum_{x,y} |A(x,y)| \leq \frac{\beta^k}{d} = L.$$

The one-density of $\mathsf{OR}_k \circ B$ is

$$1 - \left(1 - \frac{D}{n}\right)^k \leq \frac{kD}{n} \leq \beta^{-6kd} = (dL)^{-6d},$$

where the second inequality follows from the assumption on n, since

$$kD\beta^{6kd} \leq kD(1+2D)^{6k(D+1)^k} \leq 2^{(6D)^{k+1}}.$$

Applying Lemma 13 proves $\mathrm{rank}^{\pm}(\mathsf{OR}_k \circ B) \geq d \geq \alpha^k - 1 \geq (D/\lambda)^k$. □

List membership as a special case. For integers $n \geq D \geq 1$, let $x \in [n]^k$ and $\mathbf{S} = (S_1, \ldots, S_k) \in \binom{[n]}{D}^k$, and define $L_{n,k,D}(x, \mathbf{S}) = \mathbf{1}[\exists i \in [k] : x_i \in S_i]$.

Corollary 15 (List membership)**.** *For every $k, D \geq 1$ and all sufficiently large n,*

$$\mathrm{rank}^{\pm}(L_{n,k,D}) = (D+1)^{\Theta(k)}. \tag{6}$$

The sign-rank upper bound holds for every $n \geq D$.

Proof. Assume $n > D$, put $b = \binom{n}{D}$, and define $W(x, S) = 1 - \frac{n}{D+\sqrt{D}}\mathbf{1}[x \in S]$. Counting sets containing one or two prescribed elements gives

$$WW^{\mathsf{T}} = \frac{nb}{(1+\sqrt{D})^2}\left(\frac{n-D}{n-1}\mathrm{Id}_n + \frac{D-1}{n-1}\frac{J_n}{n}\right).$$

The matrix in parentheses is a convex combination of matrices of operator norm one. Hence $\|W\|_{\mathrm{op}} \le \frac{\sqrt{nb}}{1+\sqrt{D}}$. Also, $W(x,S) = 1$ when $x \notin S$, and $\frac{1}{nb}\sum_{x,S}|W(x,S)| \le 1 + \frac{D}{D+\sqrt{D}} \le 2$.

Take $A = -W^{\otimes k}$ and $d = \lfloor (1+\sqrt{D})^k \rfloor$. Then $A = -1$ on every zero entry of $L_{n,k,D}$, and

$$\|A\|_{\mathrm{op}} \le \frac{(nb)^{k/2}}{d}, \qquad \frac{1}{(nb)^k}\sum_{x,\mathbf{S}}|A(x,\mathbf{S})| \le 2^k.$$

The one-density is at most kD/n, so Lemma 13, with $L = 2^k$, applies for all sufficiently large n and gives $\mathrm{rank}^{\pm}(L_{n,k,D}) \ge \lfloor (1+\sqrt{D})^k \rfloor$. For the upper bound, every column of the membership matrix contains D ones, so apply Lemma 9 to its transpose. □

4.3 Arbitrary outer predicates

Remark 16 (Encoding arbitrary sign-rank)**.** Let $B \in \{0,1\}^{m\times m}$, and define $f_B : \{0,1\}^m \times \{0,1\}^m \to \{0,1\}$ as follows. If the first zero positions in its two arguments are i and j, set $f_B(s,t) = B_{ij}$; if either argument has no zero, set $f_B(s,t) = 0$. For every $n \ge 1$,

$$\mathrm{rank}^{\pm}(B) \le \mathrm{rank}^{\pm}(f_B \circ \mathsf{EQ}_n) \le 16\,\mathrm{rank}^{\pm}(B) + 1.$$

In particular, $\mathrm{rank}^{\pm}(f_B \circ \mathsf{EQ}_n) = \Theta(\mathrm{rank}^{\pm}(B))$ with absolute implicit constants.

Proof. Identify the block alphabet with $\{0,\dots,q-1\}$, where $q = 2^n$. On row input (a,a') and column input (b,b'), the composed function outputs B_{ij}, where i and j are the first differing coordinates of (a,b) and (a',b'); if either pair of words is equal, it outputs zero.

Lower bound. Let $e_i \in \{0,1\}^m$ have its only nonzero coordinate at i. Restrict the row inputs to $(e_i, 0^m)$ and the column inputs to $(0^m, e_j)$. The first differing coordinates of these two pairs are i and j, so the corresponding entry is B_{ij}. Thus B is a submatrix, proving the lower bound.

Upper bound. Regard the block alphabet as $\{0,\dots,2^n-1\}$, and let A be a realization of B of rank $r = \mathrm{rank}^{\pm}(B)$. For row input (a,a') and column input (b,b'), define $z_i = T^{m-i}(a_i - b_i), w_i = T^{m-i}(a'_i - b'_i)$, and set

$$H_T((a,a'),(b,b')) = \left(\sum_i z_i\right)\left(\sum_j w_j\right) z^{\mathsf{T}} A w.$$

For large T, the first disagreements dominate these sums. Specifically, if i and j are the first differing coordinates of the two pairs, the leading term of H_T is $A_{ij} z_i^2 w_j^2$, which has the sign of A_{ij}. If either pair is equal, then $H_T = 0$. Since there are finitely many inputs, choosing T sufficiently large and then $\varepsilon > 0$ sufficiently small makes $H_T - \varepsilon J$ a realization of $f_B \circ \mathsf{EQ}_n$.

Finally, factoring A into r rank-one terms expresses H_T as a sum of r products of four linear forms in z and w. Each linear form is a difference between a row function and a column function, so each product expands into at most 16 rank-one matrices. Hence $\mathrm{rank}^{\pm}(f_B \circ \mathsf{EQ}_n) \le \mathrm{rank}(H_T - \varepsilon J) \le 16r + 1$. □

Acknowledgment. KH thanks Florian Frick, Daniel Štefankovič, and Aliaksei Vasileuski for useful conversations. KH also thanks the Simons Institute for theory of computing where this work was partially conducted while visiting the program Pseudorandomness and HDX.

AI use statement. The problem of giving an exponential lower bound on sign-rank of $\mathsf{OR}_k \circ \mathsf{EQ}_n$ and applications in separating sign-rank and margin complexity, and other applications were given by the authors without AI use. A ad hoc proof of Theorem 1 with a weaker bound was obtained during interactions with LLMs. The authors subsequently simplified and streamlined the proof and formulated the one-sided witness Lemma 13. LLMs were used during writing the manuscript.

References

[ACHS24] Manasseh Ahmed, Tsun-Ming Cheung, Hamed Hatami, and Kusha Sareen. Communication complexity and discrepancy of halfplanes. In *40th International Symposium on Computational Geometry (SoCG 2024)*, volume 293 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 5:1–5:17. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2024.

[AFR85] Noga Alon, Péter Frankl, and Vojtěch Rödl. Geometrical realization of set systems and probabilistic communication complexity. In *26th Annual Symposium on Foundations of Computer Science (FOCS 1985)*, pages 277–280. IEEE, 1985.

[AHW87] Noga Alon, David Haussler, and Emo Welzl. Partitioning and geometric embedding of range spaces of finite Vapnik–Chervonenkis dimension. In *Proceedings of the Third Annual Symposium on Computational Geometry*, pages 331–340. Association for Computing Machinery, 1987.

[AMY16] Noga Alon, Shay Moran, and Amir Yehudayoff. Sign rank versus VC dimension. In *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 47–80. PMLR, 2016.

[APP+05] Noga Alon, János Pach, Rom Pinchasi, Radoš Radoičić, and Micha Sharir. Crossing patterns of semi-algebraic sets. *Journal of Combinatorial Theory, Series A*, 111(2):310–326, 2005.

[Bar98] Franck Barthe. On a reverse form of the Brascamp–Lieb inequality. *Inventiones Mathematicae*, 134(2):335–361, 1998.

[BCK+16] Joshua Brody, Amit Chakrabarti, Ranganath Kondapally, David P. Woodruff, and Grigory Yaroslavtsev. Certifying equality with limited interaction. *Algorithmica*, 76(3):796–845, 2016.

[BDES02] Shai Ben-David, Nadav Eiron, and Hans Ulrich Simon. Limitations of learning via embeddings in euclidean half spaces. *Journal of Machine Learning Research*, 3(Nov):441–461, 2002.

[BFS86] László Babai, Péter Frankl, and Janos Simon. Complexity classes in communication complexity theory (preliminary version). In *27th Annual Symposium on Foundations of Computer Science (FOCS 1986)*, pages 337–347. IEEE, 1986.

[Blo62] H. D. Block. The perceptron: A model for brain functioning. I. *Reviews of Modern Physics*, 34(1):123–135, 1962.

[BT16] Mark Bun and Justin Thaler. Improved bounds on the sign-rank of AC^0. In *43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016)*,

volume 55 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 37:1–37:14. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2016.

[BVdW07] Harry Buhrman, Nikolay Vereshchagin, and Ronald de Wolf. On computation and communication with small bias. In *Twenty-Second Annual IEEE Conference on Computational Complexity (CCC 2007)*, pages 24–32. IEEE, 2007.

[CM18] Arkadev Chattopadhyay and Nikhil S. Mande. Separation of unbounded-error models in multi-party communication complexity. *Theory of Computing*, 14(21):1–23, 2018.

[CM22] Arkadev Chattopadhyay and Nikhil S. Mande. A short list of equalities induces large sign-rank. *SIAM Journal on Computing*, 51(3):820–848, 2022. Preliminary version in FOCS 2018.

[CMMS20] Arkadev Chattopadhyay, Meena Mahajan, Nikhil Mande, and Nitin Saurabh. Lower bounds for linear decision lists. *Chicago Journal of Theoretical Computer Science*, 2020:1–15, 2020. Article 1.

[DKS12] Anirban Dasgupta, Ravi Kumar, and D. Sivakumar. Sparse and lopsided set disjointness via information theory. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2012)*, volume 7408 of *Lecture Notes in Computer Science*, pages 517–528. Springer, 2012.

[EMRPS14] Marek Eliáš, Jiří Matoušek, Edgardo Roldán-Pensado, and Zuzana Safernová. Lower bounds on geometric Ramsey functions. *SIAM Journal on Discrete Mathematics*, 28(4):1960–1970, 2014.

[Fel17] Vitaly Feldman. A general characterization of the statistical query complexity. In *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 785–830. PMLR, 2017.

[FGL+12] Jacob Fox, Mikhail Gromov, Vincent Lafforgue, Assaf Naor, and János Pach. Overlap properties of geometric expanders. *Journal für die reine und angewandte Mathematik*, 671:49–83, 2012.

[FGV21] Vitaly Feldman, Cristóbal Guzmán, and Santosh Vempala. Statistical query algorithms for mean vector estimation and stochastic convex optimization. *Mathematics of Operations Research*, 46(3):912–945, 2021.

[FH26] Florian Frick and Kaave Hosseini. An upper bound on $\mathbb{Z}_2$-index of sign-complex of Kneser graphs. Private communication, 2026.

[FHV26] Florian Frick, Kaave Hosseini, and Aliaksei Vasileuski. A $\mathbb{Z}_2$-topological framework for sign-rank lower bounds. arXiv preprint arXiv:2604.01510, 2026.

[For02] Jürgen Forster. A linear lower bound on the unbounded error probabilistic communication complexity. *Journal of Computer and System Sciences*, 65(4):612–625, 2002.

[FPS+17] Jacob Fox, János Pach, Adam Sheffer, Andrew Suk, and Joshua Zahl. A semi-algebraic version of Zarankiewicz's problem. *Journal of the European Mathematical Society*, 19(6):1785–1810, 2017.

[GHIS25] Mika Göös, Nathaniel Harms, Valentin Imbach, and Dmitry Sokolov. Sign-rank of k-Hamming distance is constant. In *2025 IEEE 66th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 2353–2368. IEEE, 2025.

[HH24] Hamed Hatami and Pooya Hatami. Guest column: Structure in communication complexity and constant-cost complexity classes. *ACM SIGACT News*, 55(1):67–93, 2024.

[HHL22] Hamed Hatami, Kaave Hosseini, and Shachar Lovett. Sign-rank vs. discrepancy. *Theory of Computing*, 18(19):1–22, 2022. Preliminary version in CCC 2020.

[HHM23] Hamed Hatami, Kaave Hosseini, and Xiang Meng. A Borsuk–Ulam lower bound for sign-rank and its applications. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing (STOC 2023)*, pages 463–471. Association for Computing Machinery, 2023.

[HHP+22] Hamed Hatami, Pooya Hatami, William Pires, Ran Tao, and Rosie Zhao. Lower bound methods for sign-rank and their limitations. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2022)*, volume 245 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 22:1–22:24. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2022.

[HJ91] Roger A. Horn and Charles R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.

[HPZZ21] Dawei Huang, Seth Pettie, Yixiang Zhang, and Zhijun Zhang. The communication complexity of set intersection and multiple equality testing. *SIAM Journal on Computing*, 50(2):674–717, 2021.

[HW07] Johan Håstad and Avi Wigderson. The randomized communication complexity of set disjointness. *Theory of Computing*, 3(11):211–219, 2007.

[HZ24] Nathaniel Harms and Viktor Zamaraev. Randomized communication and implicit representations for matrices and graphs of small sign-rank. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1810–1833. SIAM, 2024.

[IW10] Russell Impagliazzo and Ryan Williams. Communication complexity with synchronized clocks. In *2010 IEEE 25th Annual Conference on Computational Complexity (CCC)*, pages 259–269. IEEE, 2010.

[JL84] William B. Johnson and Joram Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. In Richard Beals, Anatole Beck, Alexandra Bellow, and Arshag Hajian, editors, *Conference on Modern Analysis and Probability*, volume 26 of *Contemporary Mathematics*, pages 189–206. American Mathematical Society, 1984.

[KMS20] Pritish Kamath, Omar Montasser, and Nathan Srebro. Approximate is good enough: Probabilistic variants of dimensional and margin complexity. In *Proceedings of the Thirty Third Conference on Learning Theory*, volume 125 of *Proceedings of Machine Learning Research*, pages 2236–2262. PMLR, 2020.

[KS92] Bala Kalyanasundaram and Georg Schnitger. The probabilistic communication complexity of set intersection. *SIAM Journal on Discrete Mathematics*, 5(4):545–557, 1992. Preliminary version in the Second Annual Structure in Complexity Theory Conference, 1987.

[KS01] Adam R. Klivans and Rocco A. Servedio. Learning DNF in time $2^{\widetilde{O}(n^{1/3})}$. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing (STOC 2001)*, pages 258–265. ACM, 2001.

[KS07] Adam R. Klivans and Alexander A. Sherstov. Unconditional lower bounds for learning intersections of halfspaces. *Machine Learning*, 69(2–3):97–114, 2007.

[LMSS07] Nati Linial, Shahar Mendelson, Gideon Schechtman, and Adi Shraibman. Complexity measures of sign matrices. *Combinatorica*, 27(4):439–463, 2007.

[Lov78] László Lovász. Kneser's conjecture, chromatic number, and homotopy. *Journal of Combinatorial Theory, Series A*, 25(3):319–324, 1978.

[LS09] Nati Linial and Adi Shraibman. Learning complexity vs communication complexity. *Combinatorics, Probability and Computing*, 18(1–2):227–245, 2009.

[Mat96] Jiří Matoušek. On the distortion required for embedding finite metric spaces into normed spaces. *Israel Journal of Mathematics*, 93:333–344, 1996.

[MNSW98] Peter Bro Miltersen, Noam Nisan, Shmuel Safra, and Avi Wigderson. On data structures and asymmetric communication complexity. *Journal of Computer and System Sciences*, 57(1):37–49, 1998.

[MS82] Kurt Mehlhorn and Erik M. Schmidt. Las Vegas is better than determinism in VLSI and distributed computing (extended abstract). In *Proceedings of the Fourteenth Annual ACM Symposium on Theory of Computing (STOC 1982)*, pages 330–337. Association for Computing Machinery, 1982.

[Nao18] Assaf Naor. Metric dimension reduction: A snapshot of the Ribe program. In *Proceedings of the International Congress of Mathematicians – Rio de Janeiro 2018*, volume 1, pages 759–838. World Scientific, 2018.

[New91] Ilan Newman. Private vs. common random bits in communication complexity. *Information Processing Letters*, 39(2):67–71, 1991.

[Nov62] A. B. J. Novikoff. On convergence proofs for perceptrons. Technical report, Stanford Research Institute, Menlo Park, California, November 1962. Presented at the Symposium on Mathematical Theory of Automata, Polytechnic Institute of Brooklyn, April 24–26, 1962.

[PS86] Ramamohan Paturi and Janos Simon. Probabilistic communication complexity. *Journal of Computer and System Sciences*, 33(1):106–123, 1986.

[Raz92] Alexander A. Razborov. On the distributional complexity of disjointness. *Theoretical Computer Science*, 106(2):385–390, 1992. Preliminary version in ICALP 1990.

[RS10] Alexander A. Razborov and Alexander A. Sherstov. The sign-rank of AC^0. *SIAM Journal on Computing*, 39(5):1833–1855, 2010.

[She08] Alexander A. Sherstov. Halfspace matrices. *Computational Complexity*, 17(2):149–178, 2008.

[She11] Alexander A. Sherstov. The pattern matrix method. *SIAM Journal on Computing*, 40(6):1969–2000, 2011.

[She12] Alexander A. Sherstov. The communication complexity of gap Hamming distance. *Theory of Computing*, 8(8):197–208, 2012.

[She13] Alexander A. Sherstov. Optimal bounds for sign-representing the intersection of two halfspaces by polynomials. *Combinatorica*, 33(1):73–96, 2013.

[She21] Alexander A. Sherstov. The hardest halfspace. *Computational Complexity*, 30(2):1–85, 2021. Article 11.

[SS05] Nathan Srebro and Adi Shraibman. Rank, trace-norm and max-norm. In *Learning Theory (COLT 2005)*, volume 3559 of *Lecture Notes in Computer Science*, pages 545–560. Springer, 2005.

[SSBD14] Shai Shalev-Shwartz and Shai Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 2014.

[ST13] Mert Sağlam and Gábor Tardos. On the communication complexity of sparse set disjointness and exists-equal problems. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 678–687. IEEE, 2013.

[Suk16] Andrew Suk. Semi-algebraic Ramsey numbers. *Journal of Combinatorial Theory, Series B*, 116:465–483, 2016.

[SW19] Alexander A. Sherstov and Pei Wu. Near-optimal lower bounds on the threshold degree and sign-rank of AC^0. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing (STOC 2019)*, pages 401–412. ACM, 2019.

[SZ09] Yaoyun Shi and Yufan Zhu. Quantum communication complexity of block-composed functions. *Quantum Information and Computation*, 9(5–6):444–460, 2009.

[Tha16] Justin Thaler. Lower bounds for the approximate degree of block-composed functions. In *43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016)*, volume 55 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 17:1–17:15. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2016.

[Ver18] Roman Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*, volume 47 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, 2018.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il