Electronic Colloquium on Computational Complexity, Report No. 193 (2026)

ECCC

Algebraic Complexity Approach to Sign-Rank

Mika Göös
EPFL

Kaave Hosseini
University of Rochester

Valentin Imbach
EPFL

Anastasia Sofronova
EPFL

September 18, 2026

Abstract

An outstanding open problem asks if there exists a boolean matrix with unbounded sign-rank, but bounded randomised communication complexity. We make progress towards this question by proving lower bounds against *real linear sketches* (a model weaker than sign-rank): Alice and Bob send few linear measurements to a referee, who makes a decision based on the evaluation of a low-degree polynomial. Our proof uses the *Combinatorial Nullstellensatz* and the *rank method* from algebraic circuit complexity.

Contents

ISSN 1433-8092

1 Introduction

The *sign-rank* $\mathrm{rk}_{\pm}(M)$ of a boolean matrix $M \in \{0,1\}^{N\times N}$ is the least rank of a real matrix $S \in \mathbb{R}^{N\times N}$ such that $\mathrm{sgn}(S_{ij}) = (-1)^{M_{ij}}$ for all i, j. This notion was introduced by Paturi and Simon [PS86] and it has since been extensively studied in communication complexity and learning theory (see Section 1.4 for references). It corresponds to one of the most powerful models of communication, unbounded-error protocols. In fact, it is so powerful that several non-trivial matrices can have *constant* sign-rank, independent of matrix size N:

- The Equality (or *identity*) matrix (ij-th entry is 1 iff $i = j$) has sign-rank 3.
- The Greater-Than matrix (ij-th entry is 1 iff $j \geqslant i$) has sign-rank 2.
- The k-Hamming Distance (k-HD) matrices (rows/columns indexed by binary strings; entry is 1 iff Hamming distance between the strings equals k) have sign-rank $2^{O(k)}$ [GHIS25].

The final example, in particular, is surprising: before the constant upper bound was discovered, several papers [HHP+22, HHM23] had put forth the conjecture that k-HD should require super-constant sign-rank.

Our problem. Despite decades of study, several fundamental problems about sign-rank remain poorly understood. In this paper, we address the following outstanding problem, which compares sign-rank with (public-coin) randomised communication complexity (R). The problem was first asked by Linial, Mendelson, Schechtman, and Shraibman [LMSS07] (although our formulation is from [HHP+22]). It asks "merely" to prove a super-constant sign-rank lower bound for a specific structured matrix. Nevertheless, the example of k-HD demonstrates how tricky matrices of constant sign-rank can be.

Problem 1. Exhibit a matrix M with $\mathrm{R}(M) \leqslant O(1)$ but $\mathrm{rk}_{\pm}(M) \geqslant \omega(1)$.

This problem has several equivalent interpretations in communication complexity and learning theory, arising out of the many ways in which the class of matrices with $\mathrm{R}(M) \leqslant O(1)$ can be characterised (constant discrepancy, point–halfspace arrangements with constant margin, PAC learnable under pure differential privacy); see Section 1.4 for these connections.

What makes Problem 1 particularly interesting is that most existing techniques yield lower bounds not only on $\mathrm{rk}_{\pm}(M)$ but also on $\mathrm{R}(M)$. Indeed, Hatami, Hatami, Pires, Tao, and Zhao [HHP+22] discuss how existing methods (e.g., [PS86, For02, APP+05, AMY16]) all fail to resolve this problem. So far, the problem has been solved for *partial* matrices by Hatami, Hosseini, and Meng [HHM23] who developed a novel lower-bound technique based on the Borsuk–Ulam theorem. For total matrices, Problem 1 has emerged as one of the foremost open problems in the study of *constant-cost communication* (introduced in [HHH23, HWZ22] and surveyed in [HH24]) that treats two complexity measures as "equivalent," if they can be bound as functions of one another, independent of N; see Figure 1.

The k-HD matrices stood as promising candidates (proposed in [HHP+22, HHM23]) to witness the separation in Problem 1, until the constant upper bound was discovered [GHIS25]. Currently, there exists only one known candidate matrix in the literature that is due to Sherstov and Storozhenko [SS24].

Our contributions. The goal of this paper is to make progress on Problem 1:

- We introduce a new communication model called *real linear sketching* (see Section 1.1). This model is weaker than sign-rank, but *it captures the sign-rank upper bound for* k-HD *from* [GHIS25].
- We resolve an analogue of Problem 1 by ruling out a class of constant-cost sketching protocols for the Sherstov–Storozhenko matrix (see Section 1.2). In particular, this means that *the sign-rank upper bound technique from* [GHIS25] *does not extend to the Sherstov–Storozhenko matrix.* Our lower bound will crucially use the *Combinatorial Nullstellensatz* and the central *rank method* from algebraic circuit complexity.

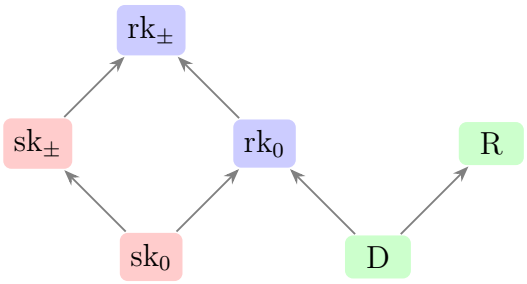

Figure 1: Relationships between complexity measures. An arrow $\mathrm{C}_1 \longrightarrow \mathrm{C}_2$ means that there is some function $\alpha\colon \mathbb{N} \to \mathbb{N}$ such that $\mathrm{C}_2(M) \leqslant \alpha(\mathrm{C}_1(M))$ for all matrices M. For example, the standard matrix rank (rk) is equivalent to deterministic communication complexity (D) since $\log \mathrm{rk}(M) \leqslant \mathrm{D}(M) \leqslant 2^{\mathrm{rk}(M)}$. In this paper, we introduce the sign-sketch ($\mathrm{sk}_\pm$) and support-sketch (sk_0) complexities, which are naturally related to sign-rank ($\mathrm{rk}_\pm$) and support-rank (rk_0). Problem 1 asks to show there is no arrow $\mathrm{R} \longrightarrow \mathrm{rk}_\pm$.

1.1 Our model: Real linear sketches

From now on, we restrict our attention to 2^n-by-2^n boolean matrices corresponding to two-party functions of the form $f\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$. In the following definition, the way a two-party function f has chosen to represent its inputs (i.e., rows/columns) as n-bit strings is important: relabelling the inputs might change the sketching complexity of the function (which is not true for usual communication protocols).

Definition 1 (Sketch protocols). In a *sketch protocol*, Alice receives input $x \in \{0,1\}^n$ and Bob receives input $y \in \{0,1\}^n$. The players each send k linear measurements $Ax \in \mathbb{R}^k$ and $By \in \mathbb{R}^k$ (where $A, B\colon \mathbb{R}^n \to \mathbb{R}^k$ are linear maps) to a third-party referee, who then makes a decision based on evaluating a fixed, degree-d, real polynomial $P\colon \mathbb{R}^{2k} \to \mathbb{R}$ on the $2k$ measurements (Ax, By). We consider two possible acceptance criteria:

- *Sign-sketch:* Referee accepts iff $P(Ax, By) > 0$.
- *Support-sketch:* Referee accepts iff $P(Ax, By) \neq 0$.

The *cost* of such a protocol is $k + d$. For a function $f\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, we define its *sign-sketch complexity* $\mathrm{sk}_\pm(f)$ as the least cost of a sign-sketch protocol that computes f correctly. The *support-sketch complexity* $\mathrm{sk}_0(f)$ is defined similarly.

Figure 1 shows how these new sketch complexities are related to previously studied measures. It is easy to see that $\mathrm{sk}_\pm(f) \leqslant 2\,\mathrm{sk}_0(f)$, since we can simply square the polynomial P. Moreover, it is not hard to see that sign-sketches imply upper bounds on sign-rank (see Section 5 for the proof). Analogously, support-sketches imply upper bounds on *support-rank* $\mathrm{rk}_0(M)$, defined as the least rank of an $S \in \mathbb{R}^{N \times N}$ such that $M_{ij} = 0$ iff $S_{ij} = 0$, that is, M and S have the same support. Studying support-rank has proven particularly useful in past work, as it enjoys closure properties ($\wedge$ and $\vee$) incomparable to those for sign-rank ($\neg$ and $\oplus$).

To better grasp our new model, consider the following examples.

(E1) The *negation* of Equality (Eq) admits a constant-cost support-sketch, that is, $\mathrm{sk}_0(\neg \mathrm{Eq}) \leqslant 2$. Each player just sends a single measurement, $Ax \coloneqq \sum_{i=1}^n 2^i x_i$, and $B \coloneqq A$. The referee then tests whether $Ax \neq By$ via the 2-variate degree-1 polynomial $P(a, b) \coloneqq a - b$. By contrast, we claim that Eq itself has unbounded support-sketch complexity, $\mathrm{sk}_0(\mathrm{Eq}) \geqslant \omega(1)$. This follows from the stronger fact that its support-rank is unbounded; in fact, $\mathrm{rk}_0(\mathrm{Eq}) = 2^n$.

(E2) Our most important example is k-Hamming Distance. It is our chief motivation for the sketching model: the constant sign-rank upper bound from [GHIS25] can, in fact, be witnessed by a support-

sketch. In particular, the technique of the paper yields $\mathrm{sk}_0(\mathrm{HD}_{\geqslant k}) \leqslant O(k^2)$ where $\mathrm{HD}_{\geqslant k}(x,y) \coloneqq 1$ iff the Hamming distance between $x, y \in \{0,1\}^n$ is at least k.

(E3) Every function admits a support-sketch with $k = 1$. Indeed, the players choose A and B to be injective on their input domain $\{0,1\}^n$ (e.g., as in (E1)). The referee can then use interpolation to design a polynomial of huge degree (exponential in n) that vanishes precisely on $\{(Ax, By) : f(x,y) = 0\}$.

(E4) The *parity* function has $\mathrm{sk}_{\pm}(\textsc{Xor}_{2n}) \geqslant 2n$. This is because a degree-d sign-sketch (regardless of k) implies a degree-d *sign representation* (aka polynomial threshold representation) for $f(x,y)$, that is, a polynomial Q such that $\mathrm{sgn}(Q(x,y)) = (-1)^{f(x,y)}$. It is a classic fact going back to Minsky and Papert [MP69] that the parity function requires full *sign-degree* ($\deg_{\pm}$), namely $\deg_{\pm}(\textsc{Xor}_{2n}) = 2n$.

The examples (E1) and (E4) demonstrate that our sketching models are incomparable to deterministic communication complexity, that is, there are no arrows $\mathrm{sk}_0 \longrightarrow \mathrm{D}$ and $\mathrm{D} \longrightarrow \mathrm{sk}_{\pm}$ in Figure 1. While example (E4) identifies sign-degree as a lower bound on sketching complexity, $\mathrm{sk}_{\pm}(f) \geqslant \deg_{\pm}(f)$, our goal is to obtain sketching lower bounds even for functions of constant sign-degree, $\deg_{\pm}(f) \leqslant O(1)$. This includes, in particular, the Sherstov–Storozhenko function to be introduced next.

1.2 Main results

The Sherstov–Storozhenko [SS24] function is defined as follows. For input length $n = m^2$ we interpret the inputs $x, y \in \{0,1\}^n$ as m-by-m boolean matrices, denote by $x \oplus y$ their modulo-2 sum, and set

$$\mathrm{SS}(x,y) \coloneqq \begin{cases} 1 & \text{if } \mathrm{rk}_{\mathbb{F}_2}(x \oplus y) > 1, \\ 0 & \text{else.} \end{cases}$$

In other words, $\mathrm{SS}(x,y) = 0$ iff the 1-entries of $x \oplus y$ form a combinatorial rectangle (set of rows $\times$ set of columns). Note that if the inputs to SS are restricted to be diagonal matrices, we recover the function $\mathrm{HD}_{\geqslant 2}$. We conjecture that SS is strictly harder than $\mathrm{HD}_{\geqslant 2}$ for sign-rank, that is, $\mathrm{rk}_{\pm}(\mathrm{SS}) \geqslant \omega(1)$, and thus SS would witness the separation in Problem 1 (it is known [SS24, GHR25] that SS admits a constant-cost randomised protocol). However, at present, we are not able to show $\mathrm{rk}_0(\mathrm{SS}) \geqslant \omega(1)$ nor $\mathrm{sk}_{\pm}(\mathrm{SS}) \geqslant \omega(1)$. Instead, our main result shows a support-sketch lower bound. This implies, in particular, that the sign-rank upper bound technique for $\mathrm{HD}_{\geqslant k}$ (example (E2)) does not extend to SS.

Theorem 1. $\mathrm{sk}_0(\mathrm{SS}) \geqslant \Omega(n^{1/6})$.

To prove Theorem 1, we rely on several tools (Combinatorial Nullstellensatz and the rank method) that have not been previously used in the study of sign-rank; see Section 1.6 for a detailed proof overview.

Extension: Boolean closure. Support-rank lower bounds can sometimes feel brittle as the measure is not closed under complement. For instance, recall from (E1) that $\mathrm{rk}_0(\neg\,\textsc{Eq}) \leqslant O(1)$ while $\mathrm{rk}_0(\textsc{Eq}) \geqslant \omega(1)$. In fact, one can show that $\mathrm{rk}_0(\neg\,\mathrm{SS}) \geqslant \omega(1)$ simply because $\neg\,\mathrm{SS}$ contains an unbounded-size Eq sub-matrix. To make sure that a support-rank/sketch result such as Theorem 1 is truly non-trivial—and hence provides some evidence that the sign-rank, too, is unbounded—it is desirable to extend the result to hold even against *boolean closures*. Formally, an n-bit two-party function f is a *boolean combination* of some n-bit functions $f_1, \ldots, f_t$ if there exists some combining function $\Gamma\colon \{0,1\}^t \to \{0,1\}$ such that $f(x,y) = \Gamma(f_1(x,y), \ldots, f_t(x,y))$ for all x, y. We define the *boolean closure* $\overline{\mathrm{rk}}_0(f)$ of support-rank as the minimum $\sum_i \mathrm{rk}_0(f_i)$ over all ways of expressing f as a boolean combination of $\{f_i\}$. The analogous closure

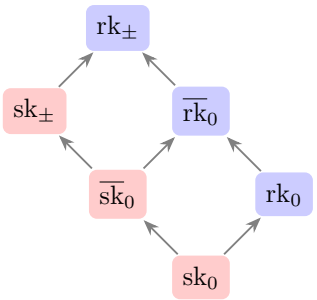

Figure 2: Boolean closures of support-rank/sketch.

for support-sketches, $\overline{\mathrm{sk}}_0(f)$, is defined similarly. Figure 2 illustrates how these measures relate to existing ones (see Section 5 for proofs). In particular, the boolean closures are still efficiently simulated by the corresponding sign measures.

We indeed show our Theorem 1 can be strengthened to hold against boolean closures. The paper [GHIS25] asked to prove lower bound against $\overline{\mathrm{rk}}_0$ as a milestone towards Problem 1; the following theorem carries this out for the support-sketch model.

Theorem 2. $\overline{\mathrm{sk}}_0(\mathrm{SS}) \geqslant \omega(1)$.

1.3 Connection to algebraic circuit complexity

A support-sketch can also be viewed as a constrained form of a nondeterministic polynomial representation. A nondeterministic polynomial for a boolean function f is a polynomial $Q \in \mathbb{R}[x]$ such that

$$Q(x) \neq 0 \iff f(x) = 1.$$

This notion was introduced by de Wolf in the context of quantum query complexity [DW03]. In our setting, the polynomial is required to have the more structured form $Q(x, y) = P(Ax, By)$, where A and B are rank-k linear maps and P has degree d. This can also be viewed as a form of depth-3 algebraic circuit complexity. If $Ax = (\ell_1(x), \ldots, \ell_k(x))$ and $By = (\ell_{k+1}(y), \ldots, \ell_{2k}(y))$, then

$$P(\ell_1, \ldots, \ell_{2k}) = \sum_{|\alpha| \leqslant d} c_\alpha \prod_{i=1}^{2k} \ell_i^{\alpha_i},$$

which gives a $\Sigma\Pi\Sigma$ circuit with product fan-in (middle layer) at most d and top fan-in at most $\binom{2k+d}{d}$. Moreover, all of its bottom linear forms lie in a space of dimension at most $2k$. The dimension of span of bottom linear forms is known as *rank* of the depth-3 circuit [SS11, SY10], which is the minimum number of independent variables needed to express the circuit. In general, consider a real $\Sigma\Pi\Sigma$ circuit

$$C(x, y) = \sum_{i=1}^{t} \prod_{j=1}^{d_i} \ell_{i,j}(x, y), \quad d \coloneqq \max_i d_i,$$

and let r be its rank. In particular, $r \leqslant td$. It is easy to see that

$$d + r/2 \leqslant \mathrm{sk}_0(C) \leqslant d + r.$$

This gives a rephrasing of our main result (Theorem 1) as a real algebraic circuit lower bound.

Theorem 3 ($\Sigma\Pi\Sigma$ lower bound)**.** *Let $C(x, y)$ be a real $\Sigma\Pi\Sigma$ circuit with rank r and maximum product fan-in d such that, on boolean inputs $x, y \in \{0, 1\}^n$,*

$$C(x, y) \neq 0 \iff \mathrm{SS}(x, y) = 1.$$

Then, $r + d \geqslant \Omega(n^{1/6})$.

This lower bound is particularly curious in the context of Polynomial Identity Testing (PIT); see surveys [Sax14, SY10]. Indeed, testing whether $\mathrm{SS}(x, y) = 1$ is a simple PIT problem over $\mathbb{F}_2$. Let $M = x \oplus y$ be viewed as an m-by-m matrix over $\mathbb{F}_2$. Define a polynomial $p_M \colon (\mathbb{F}_2^m)^4 \to \mathbb{F}_2$ of degree four as

$$p_M(a_1, a_2, b_1, b_2) \coloneqq \det \begin{pmatrix} a_1^\top M b_1 & a_1^\top M b_2 \\ a_2^\top M b_1 & a_2^\top M b_2 \end{pmatrix} = (a_1^\top M b_1)(a_2^\top M b_2) + (a_1^\top M b_2)(a_2^\top M b_1).$$

It is easy to see that

$$p_M \equiv 0 \iff \mathrm{rk}_{\mathbb{F}_2}(M) \leqslant 1.$$

Thus, evaluating p_M at random points $a_1, a_2, b_1, b_2 \in \mathbb{F}_2^m$ gives a constant-query, one-sided-error test for whether $\mathrm{rk}_{\mathbb{F}_2}(M) > 1$. This is indeed the idea behind the constant-cost protocol for SS [SS24, GHR25]. Moreover, for every fixed choice of $a_1, a_2, b_1, b_2 \in \mathbb{F}_2^m$, the polynomial

$$C(x, y) \coloneqq p_M(a_1, a_2, b_1, b_2) = (a_1^\top(x \oplus y)b_1)(a_2^\top(x \oplus y)b_2) + (a_1^\top(x \oplus y)b_2)(a_2^\top(x \oplus y)b_1)$$

is a $\mathbb{F}_2$-$\Sigma\Pi\Sigma$ circuit with top fan-in and product fan-in 2, and rank at most 4. Moreover, it is easy to see that for a random choice of such C, if $\mathrm{SS}(x, y) = 1$, then $C(x, y) = 1$ with probability at least 2^{-4}; and also if $\mathrm{SS}(x, y) = 0$, then always $C(x, y) = 0$. Taking constantly many samples results in the following statement.

Theorem 4. *There is a distribution over $\mathbb{F}_2$-$\Sigma\Pi\Sigma$ circuits C of constant rank and constant product fan-in such that*

$$\Pr_C[C(x, y) = \mathrm{SS}(x, y)] \geqslant 2/3.$$

Moreover, the error is one-sided; if $\mathrm{SS}(x, y) = 0$ then $C(x, y) = 0$.

1.4 Motivation and related work

Unbounded-error communication. Paturi and Simon [PS86] showed that sign-rank characterises the communication complexity in the unbounded-error communication model (U), where randomness is private and the success probability is required to be strictly larger than $\frac{1}{2}$. Specifically, they showed that

$$\mathrm{U}(M) = \log \mathrm{rk}_{\pm}(M) \pm O(1).$$

In particular, this means U $\longleftrightarrow$ $\mathrm{rk}_{\pm}$ in the language of Figure 1. Newman's theorem [New91] states that public-coin randomness can be converted into private-coin randomness with just $O(\log n)$ communication overhead. Thus, in particular, $\mathrm{U}(M) \leqslant \mathrm{R}(M) + O(\log n)$. From this perspective, Problem 1 asks whether one can get rid of the $O(\log n)$ dependence in this application of Newman's theorem.

Margin and learning theory. There is also a geometric formulation via the *margin complexity* $\mathrm{mar}(M)$ of a sign matrix $M \in \{\pm 1\}^{N \times N}$. This is defined by

$$\mathrm{mar}(M) \coloneqq \sup \min_{i,j} |\langle u_i, v_j \rangle|,$$

where the supremum ranges over all dimensions and all unit vectors u_i, v_j satisfying $M_{ij} = \mathrm{sgn}\langle u_i, v_j \rangle$. Sign-rank asks for a realisation in low dimension, whereas margin asks for a realisation with a large separation from the defining hyperplanes, with no restriction on the dimension. Constant-cost randomised protocols give constant-margin representations and vice versa [LMSS07]; that is, we have "R $\longleftrightarrow$ mar" in the language of Figure 1. Thus Problem 1 can be viewed as asking whether constant margin implies constant-dimensional sign representations. Margin is a fundamental parameter in learning theory that controls the sample complexity of the support vector machine algorithm. The paper [LMSS07] observed that one can use the Johnson–Lindenstrauss lemma (in place of Newman's theorem) to upper bound sign-rank using margin with logarithmic dependence on the size of the matrix. From this perspective, Problem 1 asks whether one can get rid of the size dependence in this application of Johnson–Lindenstrauss. Finally, yet another characterisation of matrices with $\mathrm{R}(M) \leqslant O(1)$ is that they correspond to PAC-learnability under pure differential privacy [FX14, BNS19].

We also mention that, very recently, several new frameworks and invariants for lower bounding sign-rank have been developed from a topological and learning theoretic perspective. Among them are lower bounding sign-rank by list replicability [CMW25, BHH+26a, BHH+26b], or by $\mathbb{Z}_2$-index of certain simplicial complexes [FHV26, BHH+26b], and approximate sign-rank lower bounds via a strengthening of Forster's method [BHKR26].

Converse problem. The problem opposite of Problem 1 asks to construct matrices with small sign-rank but large randomised complexity. This direction was initiated by [BFS86] and since then it has been studied exhaustively [BVdW07, She08, She11, She13, Tha16, She21, HHL20, ACHS24]. One culmination of this line of work [HHL20, ACHS24] exhibits a matrix of sign-rank 3 but randomised complexity $\Omega(n)$.

1.5 Open problems

Our current method relies heavily on applying Combinatorial Nullstellensatz to the explicit polynomial P in a support-sketch. Finding a support-rank analogue of this step is open, as is finding any technique to replace Nullstellensatz that can reason about the signs of a polynomial. This yields the following two incomparable challenge problems that serve as milestones towards Problem 1.

Problem 2. Show $\mathrm{rk}_0(\mathrm{SS}) \geqslant \omega(1)$.

Problem 3. Show $\mathrm{sk}_{\pm}(\mathrm{SS}) \geqslant \omega(1)$.

Another direction is to search for other candidate functions besides the function SS. It is widely conjectured [GKPW18, CLV19, HHH23, HH24, GHR+26, BHT25] that $\mathrm{R}(f) \leqslant O(1)$ implies that f contains a monochromatic rectangle of density $\Omega(1)$. (This is indeed satisfied by SS.) Such constant-density monochromatic rectangles are known to exist for functions with $\mathrm{rk}_{\pm}(f) \leqslant O(1)$ [APP+05]. Thus, if this conjecture turns out to be false, the counterexample function would then witness the separation in Problem 1, too.

1.6 Proof overview of Theorem 1

We develop a general lower bound technique for any Xor-*function*. We say that a function $f\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ is an Xor-*function* if for some $g\colon \{0,1\}^n \to \{0,1\}$ we have $f = g \circ \textsc{Xor}$, namely,

$$\forall\, x, y \in \{0,1\}^n \;:\quad f(x,y) = g(x \oplus y).$$

Note that $\mathrm{SS} = \textsc{Rank}_{>1} \circ \textsc{Xor}$ where for $x \in \mathbb{F}_2^{m\times m}$, $n = m^2$,

$$\textsc{Rank}_{>1}(x) \coloneqq \begin{cases} 1 & \text{if } \mathrm{rk}_{\mathbb{F}_2}(x) > 1, \\ 0 & \text{else.} \end{cases}$$

At a high level, our proof of Theorem 1 is an iterative degree-reduction argument. We start with the polynomial P from a given support-sketch computing a function f. We then find a small subset of variables of P that intersects all its top-degree monomials. By restricting these variables we obtain a polynomial P' of smaller degree, while at the same time we make sure, using a "self-reducibility" property of f, that the restricted function f' still computes a copy of f, but on slightly fewer variables. Iterating this process eventually reduces the polynomial (and hence the function) to a constant. Since it takes several steps to reduce f to a constant, it follows that P must have had high degree to start with. This iterative procedure has two main ingredients, using *Combinatorial Nullstellensatz* [Alo99] and the *rank method* [Nis91].

Nullstellensatz and no spread monomials. Let $g\colon \{0,1\}^n \to \{0,1\}$ and suppose $g(x \oplus y) = 0$ if and only if $P(Ax, By) = 0$ where P has degree d and A, B are linear maps to $\mathbb{R}^k$. Consider the change of variables $z_i^+ = x_i + y_i$ and $z_i^- = x_i - y_i$. Hence $Q(z) \coloneqq P(Cz) = P(Ax, By)$ where $z = (z^+, z^-) \in \mathbb{R}^{2n}$ and

$$C = \frac{1}{2}\begin{pmatrix} A & A \\ B & -B \end{pmatrix}.$$

x_i	y_i	$x_i \oplus y_i$	$z_i^+ = x_i + y_i$	$z_i^- = x_i - y_i$
0	0	0	0	**0**
1	1	0	2	**0**
0	1	1	**1**	-1
1	0	1	**1**	1

Figure 3: Fixing $x_i \oplus y_i$ causes either z_i^+ or z_i^- to become fixed while the other ranges over two values.

If we fix the parity $x_i \oplus y_i$ (to either 0 or 1), exactly one of z_i^+, z_i^- becomes fixed and the other ranges over two values; see Figure 3. We use this property to find for a given 0-input $a \in g^{-1}(0)$ a large product grid of corresponding inputs z where $Q(z)$ vanishes. Our goal is to show that this is impossible if $k + d$ is small.

More specifically, for $g = \textsc{Rank}_{>1}$, recall that the 0-inputs are precisely matrices whose 1-entries form a combinatorial rectangle. In this case the coordinates of z^+ and z^- are indexed by $[m] \times [m]$. Suppose a top-degree monomial M of Q is "spread out" meaning that the matrix entries corresponding to its (plus or minus type) variables are lying in distinct rows and columns. Based on the (plus or minus) type of each variable of M we can prescribe entries of $x \oplus y$ in a way that $x \oplus y$ has rank at most 1 and all variables of M independently vary over two values. The Combinatorial Nullstellensatz (Theorem 5) guarantees that Q is non-zero somewhere on this grid, contradicting that the entire grid consists of 0-inputs of Q. In summary, this part implies that Q does not have such "spread-out" top-degree monomials.

In the more general case, we capture the notion of "spread-out" based on some dependency graph defined for the function g of interest. The corresponding statement is given in Lemma 2.

Rank method and no spread implies hitting sets. The second ingredient is to show that an absence of top-degree spread-out monomials of Q forces global structure: all top-degree monomials of Q can be hit by a few variables. In the case of $g = \textsc{Rank}_{>1}$ the bound is $O(mkd)$. To prove this, the key object of study here is a type of coefficient matrix D_Q obtained from first derivatives of the top homogeneous part of Q. On the one hand, we show that the rank of D_Q is upper-bounded by k, and, on the other hand, we show that if the top-degree monomials have no small hitting set, we can build a large triangular sub-matrix of D_Q hence lower-bounding its rank. The final result is a second technical ingredient Lemma 3 which may be of independent interest.

Iteration and putting the lemmas together. The first ingredient rules out spread-out top-degree monomials, and combined with the second, we get a small hitting set for all top degree monomials. We restrict all the variables in the hitting set which reduces the degree by one. Repeating for $d, d-1, \cdots, 1$ gives a bound on the total number of variables. In the case of $g = \textsc{Rank}_{>1}$ this is $O(mkd \times d) = O(mkd^2) = O(\sqrt{n}kd^2)$ which implies the bound of $\Omega(m^{1/3})$ on $k + d$, namely, the support-sketch complexity of the SS function.

2 Setup: Resilience and self-reducibility

Our lower bound proof for $\text{SS} = \textsc{Rank}_{>1} \circ \textsc{Xor}$ works more generally for any function $f \circ \textsc{Xor}$ as long as f satisfies the following two properties, which we start formalising below.

- *Resilience*, stating that certain partial inputs to f can be completed to inputs that evaluate to 0. This is used to construct the grid of inputs on which we apply the Combinatorial Nullstellensatz.
- *Self-Reducibility*, stating that certain inputs of f can be restricted in such a way that we recover f on smaller inputs. We will use this to concentrate the complexity of f in the top-degree of the polynomial.

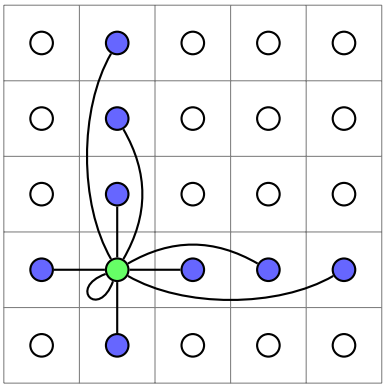

Figure 4: The dependency graph G_{RANK} associated with the boolean function $\text{RANK}_{>1}$ on $\{0,1\}^{5\times 5}$. The neighbourhood of the green vertex $x_{4,1}$ is highlighted in blue.

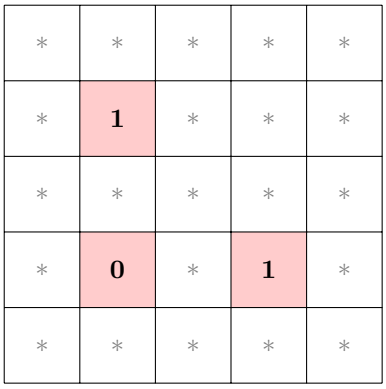

Figure 5: The three red entries form a "corner" with 1s at the tips and a 0 at the joint. This is a local obstruction to the 1-entries forming a combinatorial rectangle.

2.1 Dependency graphs

Both of the properties are defined relative to a *dependency graph* G on the input variables of f, which we identify with $[n]$. Intuitively, an edge $i \sim_G j$ in this graph signifies that the two variables x_i and x_j are "locally dependent" on each other, in the context of evaluating f. We follow the convention that $i \sim_G i$ for all $i \in [n]$, i.e. that each variable is dependent on itself. For $H \subseteq [n]$, define

$$N_G(H) = \{i \in [n] : i \sim_G j \text{ for some } j \in H\}.$$

We sometimes drop the subscript G. Further, define $\Delta(G)$ as the maximal value of $|N_G(\{i\})|$ over all $i \in [n]$. We denote by $\varnothing$ the empty graph given by $i \sim_\varnothing j \iff i = j$. Note that $N_\varnothing(H) = H$ for all $H \subseteq [n]$ and $\Delta(\varnothing) = 1$. A set $S \subseteq [n]$ is *independent* in G if for any distinct $x, y \in S$ we have $x \not\sim y$.

When dealing with $\text{RANK}_{>1} \circ \text{XOR}$, we will work relative to the dependency graph G_{RANK}, given by

$$x_{ij} \sim_{G_{\text{RANK}}} x_{i'j'} \iff i = i' \text{ or } j = j'.$$

In other words, two variables are connected if and only if their corresponding entries in the matrix lie in the same row or column, see Figure 4. Intuitively, this is a good notion of "local dependence" when deciding if the 1-entries form a combinatorial rectangle: The local obstruction of this fact consists of 3 variables spanning just two rows and two columns, as illustrated in Figure 5.

2.2 Resilience

Recall that for a boolean function $f\colon \{0,1\}^n \to \{0,1\}$ and some $I \subseteq [n]$, a *restriction* of f on I is a function $f'\colon \{0,1\}^{n-|I|} \to \{0,1\}$ obtained by fixing the variables indexed by I to some boolean values.

Definition 2 (Resilience). Let $f\colon \{0,1\}^n \to \{0,1\}$ and let G be a graph on $[n]$. The *resilience* of f with respect to G, written $\text{res}_G(f)$, is the largest integer t with the following property: For every independent set S in G with $|S| \leqslant t$ and every restriction of f on S, the resulting function is non-constant. If there is no largest such integer, we set $\text{res}_G(f) = \infty$.

Example 1. We have $\text{res}_\varnothing(\text{HW}_{\geqslant k}) = k - 1$ (assuming $n > k$), where $\text{HW}_{\geqslant k}$ is the function that accepts all n-bit strings of Hamming weight $\geqslant k$. Also, $\text{res}_\varnothing(\text{PARITY}) = n - 1$, where PARITY is the function that accepts all n-bit strings with odd Hamming weight. Finally, $\text{res}_{G_{\text{RANK}}}(\text{RANK}_{>1}) = \infty$, since no restriction on any independent set is constant, see Figure 6.

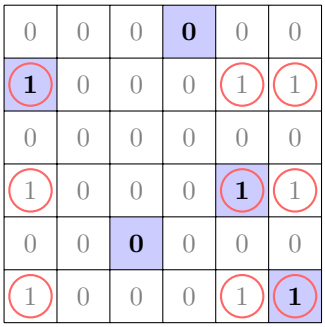

Figure 6: After fixing some entries on a G_{RANK}-independent set (in blue), we can still choose to either create a 1-rectangle (in red), or to avoid it. Thus, $\text{res}_G(\text{RANK}_{>1}) = \infty$.

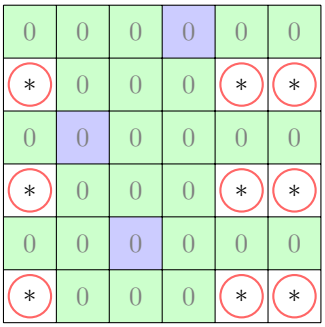

Figure 7: For any set of variables I (in blue) of $\text{RANK}_{>1}$, there is a restriction (to zero) on a superset of $N(I)$ (in green) which recovers $\text{RANK}_{>1}$ (in red). Thus, $\text{RANK}_{>1}$ is G_{RANK}-self-reducible.

2.3 Self-reducibility

Definition 3 (Self-reducibility). Let $f = (f_n)$ be a family of functions and let $G = (G_n)$ be a corresponding family of dependency graphs on $[n]$, where $\Delta(G_n)$ is non-decreasing in n. We say f is *G-self-reducible* if for every set of variables indexed by $I \subseteq [n]$ with $|I| < n/\Delta(G_n)$, there is a restriction of f_n on a superset of $N(I)$ which recovers some $f_{n'}$ with input size $n' \geqslant n - \Delta(G_n)|I|$.

Example 2. Both $\text{HW}_{\geqslant k}$ and PARITY are $\varnothing$-self-reducible, since after restricting some variables to zero, we recover the same function, just on a smaller domain.

Moreover, $\text{RANK}_{>1}$ is G_{RANK}-self-reducible: Note that $\Delta(G_{\text{RANK}}) = 2\sqrt{n}-1$ and let I be a set of variables of $\text{RANK}_{>1}$ with $|I| \leqslant \sqrt{n}$. When $\text{RANK}_{>1}$ is restricted on $N_{G_{\text{RANK}}}(I)$, the unrestricted inputs still contain a combinatorial square R of side length at least $\sqrt{n} - |I|$. Restricting all entries outside of R to 0, we recover the function $\text{RANK}_{>1}$ on input length $(\sqrt{n} - |I|)^2 \geqslant n - \Delta(G_{\text{RANK}}) \cdot |I|$. This is illustrated in Figure 7.

3 Proof of main result (Theorem 1)

The proof of Theorem 1 immediately follows from the following more general lemma. Indeed, we have seen that $\text{RANK}_{>1}$ is G_{RANK}-self-reducible in Example 2. We also know that $\text{res}_{G_{\text{RANK}}}(\text{RANK}_{>1}) = \infty$ from Example 1 and $\Delta(G_{\text{RANK}}) = 2\sqrt{n} - 1$. Plugging these facts in the lemma below proves Theorem 1.

Lemma 1. *Let $f = (f_n)$ be a G_n-self-reducible family. If $\text{res}_{G_n}(f_n) \geqslant \Omega(n/\Delta(G_n))^{1/3}$, then*

$$\text{sk}_0(f_n \circ \text{XOR}) \geqslant \Omega(n/\Delta(G_n))^{1/3}.$$

The proof of Lemma 1 follows from the next two Lemmas 2 and 3. We postpone their proofs to Sections 3.1 and 3.2 respectively, since it requires the introduction of new concepts. To state the lemmas, we first need to define notions of *independence* and *hitting sets*.

Independent and hitting sets. For a monomial M appearing in $P \in \mathbb{R}[x_1, x_2, \ldots, x_n]$, write $\text{supp}(M)$ for the multiset of indices of the variables appearing in M. Given a graph G on $[n]$, we call M *independent* in G if M is multilinear and $\text{supp}(M)$ is an independent set in G. Further, we denote by $G^{(2)}$ the graph on $[2n]$ given by

$$i \sim_{G^{(2)}} j \iff (i \mod n) \sim_G (j \mod n).$$

Furthermore, let $\mathcal{M}$ be the family of all top-degree monomials of P. A set $H \subseteq [n]$ is a *G-hitting set* of P if

$$\forall M \in \mathcal{M}: \quad \operatorname{supp}(M) \cap N(H) \neq \varnothing.$$

We denote by $h_G(P)$ the size of the smallest G-hitting set in P.

The main technical lemmas. The next lemma is where we apply Combinatorial Nullstellensatz.

Lemma 2. *Let $f\colon \{0,1\}^n \to \{0,1\}$ and let G be a graph on $[n]$. Let P be a real, degree-d polynomial over $2n$ variables such that*

$$\forall x, y \in \{0,1\}^n: \quad f(x \oplus y) = 0 \iff P(x+y,\ x-y) = 0.$$

If $\operatorname{res}_G(f) \geqslant d$, then P has no degree-d monomial which is independent in $G^{(2)}$.

The dual notions of independent set and hitting set are usually related as follows: Independent set size is a lower bound on hitting set size. The following lemma provides a converse to this paradigm in our setting.

Lemma 3. *Let $P \in \mathbb{R}[x_1, \ldots, x_k]$ and, $A : \mathbb{R}^n \to \mathbb{R}^k$ linear, and G a graph on $[n]$. Let $Q(x) \coloneqq P(Ax)$ and suppose that Q is non-zero and $d = \deg(Q)$. If Q has no top-degree monomial which is independent in G, then $h_G(Q) \leqslant dk$.*

Proof of Lemma 1. Let $\Delta_n = \Delta(G_n)$. Suppose that $f_n \circ \textsc{Xor}$ has a support-sketch protocol with k sketches and a degree-d polynomial P. For a contradiction, assume that there is a constant $0 < c < 2^{-1/3}$ with

$$k + d \leqslant c \left(\frac{n}{\Delta_n}\right)^{1/3} \leqslant \frac{\operatorname{res}_{G_n}(f_n)}{2}.$$

We repeatedly apply the following reduction argument, which decreases both the input size m and the degree d_m. We will ensure that $m \geqslant n/2$, where by monotonicity of Δ_m we have

$$\operatorname{res}_{G_m}(f_m) \geqslant 2c \left(\frac{m}{\Delta_m}\right)^{1/3} \geqslant c \left(\frac{n}{\Delta_n}\right)^{1/3} \geqslant d. \tag{1}$$

Given a sketch-protocol for $f_m \circ \textsc{Xor}$ with polynomial P_m of degree d_m, we perform the change of variables $z^+ = x + y$ and $z^- = x - y$. Then, by Lemmas 2 and 3, we get a $G_m^{(2)}$-hitting set H of size at most $2d_m k$. Let I denote the set of variables of f_m by collapsing z_i^+ with z_i^- in H. Note that $|I| \leqslant |H| \leqslant 2d_m k$. By self-reducibility, we can restrict a superset of $N(I)$ to recover $f_{m'}$, where

$$m' \geqslant m - \Delta_m |I| \geqslant m - 2k\Delta_m d_m.$$

Performing this restriction to both the z^+ and z^- variables, we recover $f_{m'} \circ \textsc{Xor}$ and a new support-sketch protocol with polynomial $P_{m'}$ of degree $d_{m'}$. But since H was a $G_m^{(2)}$-hitting set, we must have $d_{m'} < d_m$. Repeating this, we reach degree 0 while the overall loss in input size is at most

$$2k\Delta_n \sum_{i=1}^{d} i \leqslant k\Delta_n d(d+1) \leqslant \Delta_n (k+d)^3 \leqslant c^3 n \leqslant n/2,$$

ensuring that Equation (1) holds at every step. Hence, we have arrived at a non-constant function f_m with a degree-zero support-sketch protocol, which is a contradiction. □

3.1 Combinatorial Nullstellensatz and proof of Lemma 2

Our goal in this subsection is to prove Lemma 2. The key observation is that 0-inputs of an XOR-function naturally give rise to product grids in the variables $x+y$ and $x-y$. If a top-degree monomial of the sketching polynomial is sufficiently spread out, then one can choose such a grid on which the polynomial must vanish identically, while the Combinatorial Nullstellensatz guarantees that it is non-zero at some point of the grid. This contradiction rules out independent top-degree monomials.

Theorem 5 (Combinatorial Nullstellensatz [Alo99])**.** *Let $P \in \mathbb{R}[x_1, \dots, x_n]$ be a real polynomial. Let $M = \prod_{i=1}^{n} x_i^{t_i}$ be a monomial of P with maximal degree and non-zero coefficient. Then, for any $S_1, \dots, S_n \subseteq \mathbb{R}$ with $|S_i| \geqslant t_i + 1$ there is at least one $s \in S_1 \times \cdots \times S_n$ such that $P(s) \neq 0$.*

We will use a special case of the Combinatorial Nullstellensatz where $|S_i| = t_i + 1 \in \{1, 2\}$ for all i, and hence the monomial M in question is multilinear. See Figure 8 for an example.

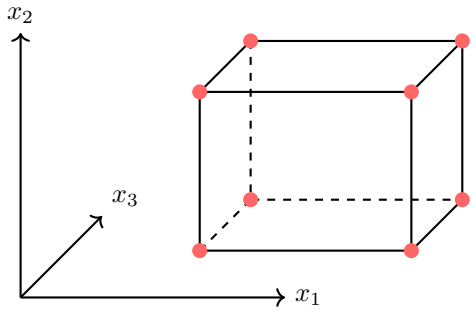

Figure 8: The special case of the Nullstellensatz we use, with $n = 3$: If $P \in \mathbb{R}[x_1, x_2, x_3]$ has degree 3 and vanishes on all 8 vertices of some axis-aligned cuboid (in red), then the coefficient of $x_1x_2x_3$ in P must be 0.

Proof of Lemma 2. For contradiction, assume we have such a monomial M. Note that M cannot contain both $z_i^+ = x_i + y_i$ and $z_i^- = x_i - y_i$ for any i because of independence in $G^{(2)}$, as z_i^+ and z_i^- are substituted in P as variables with numbers i and $n + i$. For every z_i^+ in M, we restrict to $x_i = y_i$. For every z_i^- in M, we restrict to $x_i \neq y_i$. This restricts the corresponding inputs of f to 0 and 1, respectively. Moreover, since the monomial is independent in $G^{(2)}$, the set of restricted input bits of f forms an independent set in G of size $\deg(P)$. Because $\mathrm{res}_G(f) \geqslant \deg(P)$, we can extend this restriction to an input $a \in f^{-1}(0)$. We now fix all x_i and y_i that are not part of the above relations in such a way that $x_i \oplus y_i$ agrees with a_i, thus also fixing both z_i^+ and z_i^-. Now, the only z_i^+ that are not fixed are exactly the ones appearing in M, each of which can independently take a value $z_i^+ \in \{0, 2\}$. Similarly, the only z_i^- not yet fixed are the ones in M, with $z_i^- \in \{-1, 1\}$. Therefore, together with the fixed $z^{\pm}$-variables, these values form a grid that satisfies the conditions of Combinatorial Nullstellensatz (Theorem 5) for the monomial M. Thus, there must exist a point on this grid where the polynomial does not vanish, which contradicts the fact that $a \in f^{-1}(0)$. □

3.2 Rank method and proof of Lemma 3

Here we prove Lemma 3 by developing a rank argument. For a polynomial Q of degree d, we will denote by $Q^{(d)}$ the *degree-d homogeneous part* of Q.

Definition 4. Let $Q \in \mathbb{R}[x_1, \ldots, x_n]$ have degree $d \geqslant 1$. Define the *first-derivative coefficient matrix* D_Q with rows indexed by $i \in [n]$ and columns indexed by degree-$(d-1)$ monomials in the variables of Q. The entry for row i and monomial M is

$$D_Q(i, M) \coloneqq \text{ coefficient of } M \text{ in } \frac{\partial Q^{(d)}}{\partial x_i}.$$

Note that by definition, $\mathrm{rk}(D_Q) = \dim \mathrm{span}\left\{\frac{\partial Q^{(d)}}{\partial x_1}, \cdots, \frac{\partial Q^{(d)}}{\partial x_n}\right\}$. The crucial property is that $D_Q(i, M) \neq 0$ if and only if $x_i M$ is a degree-d monomial of Q. In fact, we have $D_Q(i, M) = (r+1) \cdot c$ where r is the exponent of x_i in M and c is the coefficient of $x_i M$ in Q.

Lemma 4. *Let $P \in \mathbb{R}[u_1, \cdots, u_k]$ and $A : \mathbb{R}^n \to \mathbb{R}^k$ be a linear map. For $Q(x) \coloneqq P(Ax)$, we have*

$$\mathrm{rk}(D_Q) \leqslant \mathrm{rk}(A) \leqslant k.$$

Proof. Let d be the degree of Q and let $A(x) = (L_1(x), \cdots, L_k(x))$, where $L_j(x) = \sum A_{ji} x_i$. Then, we have $Q^{(d)} = P^{(d)}(L_1(x), \cdots, L_k(x))$. By the chain rule,

$$\frac{\partial Q^{(d)}}{\partial x_i} = \sum_{j=1}^{k} \frac{\partial P^{(d)}}{\partial u_j}(Ax) \cdot \frac{\partial L_j}{\partial x_i} = \sum_{j=1}^{k} A_{j,i} \cdot \frac{\partial P^{(d)}}{\partial u_j}(Ax).$$

Hence the partial derivatives of $Q^{(d)}$ are linear combinations of k fixed polynomials, with coefficient vectors given by the columns of A. We claim that $\mathrm{rk}(D_Q) \leqslant \mathrm{rk}(A)$. Let $\frac{\partial P^{(d)}}{\partial u_j}(Ax) = \sum_M B_{j,M} M$ where M ranges over all degree-$(d-1)$ monomials in $x_1, \cdots, x_n$. Then

$$\frac{\partial Q^{(d)}}{\partial x_i} = \sum_M \left(\sum_{j=1}^{k} A_{j,i} B_{j,M} \right) M,$$

hence, $D_Q = A^\top B$ and $\mathrm{rk}(D_Q) \leqslant \mathrm{rk}(A^\top) = \mathrm{rk}(A) \leqslant k$. □

The next lemma converts a bound on derivative coefficient rank into the existence of either a hitting set or a monomial whose support is an independent set.

Lemma 5. *Let $Q \in \mathbb{R}[x_1, \ldots, x_n]$ be a polynomial of degree $d \geqslant 1$, and let G be a graph on $[n]$. If no degree-d monomial of Q is independent in G, then*

$$h_G(Q) \leqslant d \cdot \mathrm{rk}(D_Q).$$

Proof. We show that D_Q has a triangular sub-matrix of size at least $h_G(Q)/d$ (with no zeros on the diagonal). View each monomial M as a multiset of its d variables. Define $E(M)$ as the number of edges of G among the variables appearing in M:

$$E(M) \coloneqq \#\big\{\{x_i, x_j\} \subseteq M : i \sim_G j\big\}.$$

Note that a variable with multiplicity $k \geqslant 2$ in M contributes $\binom{k}{2}$ edges to $E(M)$. Thus, M is independent in G if and only if $E(M) = 0$. Hence, by our hypothesis, every degree-d monomial of Q has $E(M) \geqslant 1$.

The pivoting step. Suppose $I \subseteq [n]$ is a set of variables with $|I| < h_G(Q)$. Then, some degree-d monomial avoids $N_G(I)$. Choose one such monomial M which minimises $E(M)$, and pick a variable x_i of M with

$$E(M/x_i) < E(M).$$

Such an x_i exists because $E(M) \geqslant 1$. Let $M' \coloneqq M/x_i$. Then:

(a) We have $D_Q(i, M') \neq 0$, because $M'x_i = M$ is a top-degree monomial of Q.

(b) Moreover, $i \notin N_G(I)$, since we picked $M = M'x_i$ to avoid $N_G(I)$.

(c) Finally, for every j with $D_Q(j, M') \neq 0$, we have $j \in N_G(\mathrm{supp}(M') \cup I)$. If not, then $x_j M'$ would be a top-degree monomial in Q avoiding $N_G(I)$. But since $j \notin N_G(\mathrm{supp}(M'))$, we would have $E(x_j M') < E(M)$, contradicting minimality of $E(M)$.

Building a triangular sub-matrix. We will construct a triangular sub-matrix (with no zero on the diagonal) in D_Q, by repeatedly using the pivoting step to pick rows $i_1, i_2, \ldots$ and columns corresponding to degree-$(d-1)$ monomials $M'_1, M'_2, \ldots$ in Q.

For $t \geqslant 1$, suppose we have already found pivot rows $i_1, \ldots, i_{t-1}$ and monomials $M'_1, \ldots, M'_{t-1}$. Set

$$I_{t-1} \coloneqq \mathrm{supp}(M'_1 x_{i_1}) \cup \mathrm{supp}(M'_2 x_{i_2}) \cup \cdots \cup \mathrm{supp}(M'_{t-1} x_{i_{t-1}}).$$

As long as $|I_{t-1}| < h_G(Q)$, the pivoting step applied to I_{t-1} produces a degree-d monomial M_t, a pivot row $i_t \in \mathrm{supp}(M_t)$ and a column $M'_t \coloneqq M_t/x_{i_t}$ which satisfy the three properties listed above. We now show that the new sub-matrix of D_Q induced by rows $\{i_1, \ldots, i_t\}$ and columns corresponding to $\{M'_1, \ldots, M'_t\}$ is again triangular (after reordering), with non-zero diagonal:

Fix $s < t$ and consider entry $D_Q(i_t, M'_s)$. By property (c) of the s-th pivoting step, this entry can be non-zero only if $i_t \in N_G(\mathrm{supp}(M'_s) \cup I_{s-1}) \subseteq N_G(I_{t-1})$. But property (b) of the t-th pivoting step guarantees $i_t \notin N_G(I_{t-1})$. Hence $D_Q(i_t, M'_s) = 0$ for all $s < t$, while $D_Q(i_t, M'_t) \neq 0$ (property (a) of the t-th pivoting step).

The process terminates only when $h_G(Q) \leqslant |I_t|$. Since $|I_t| \leqslant \sum_{i=1}^{t} |\mathrm{supp}(M'_i x_i)| = td$, we get a triangular sub-matrix of size at least $h_G(Q)/d$, as desired. □

Proof of Lemma 3. By Lemma 4, we have $\mathrm{rk}(D_Q) \leqslant \mathrm{rk}(A) \leqslant k$. By Lemma 5, $h_G(Q) \leqslant d \cdot \mathrm{rk}(D_Q)$. The result follows from combining the two bounds. □

4 Boolean closures (Theorem 2)

In this section we will extend our lower bounds to the boolean closure of support-sketches. The proof mirrors the main argument in [FGHH25], and we will only make explicit the parts that deviate from it.

Let $A \in \{0,1\}^{n \times n}$ be a matrix. For an index $i \in [n]$, we denote by A_i the i-th row of A. For a permutation π acting on $[n]$, we denote by A_π the matrix with rows $A_{\pi(1)}, A_{\pi(2)}, \ldots, A_{\pi(n)}$.

Definition 5 (Minor-Closure). We say that a family $\mathcal{F}$ of boolean functions on $\{0,1\}^*$ is *minor-closed*, if for any $f\colon \{0,1\}^n \to \{0,1\}$ in $\mathcal{F}$ and any affine function $A\colon \{0,1\}^m \to \{0,1\}^n$ with $m \geqslant n$, the composition $f \circ A$ is also in $\mathcal{F}$.

Since $\{0,1\}^n \times \{0,1\}^n \cong \{0,1\}^{2n}$, this definition naturally extends to two-party functions.

Lemma 6. *The family of functions for which* $\mathrm{sk}_0 \leqslant k + d$ *is minor-closed.*

Proof. Note that affine change of variables turns linear sketches into affine sketches. However, the additive constant can be absorbed in the polynomial, whose degree is unchanged. □

Definition 6 (Stackability). We say that a function f on boolean square matrices is *stackable* if

$$\forall X \in \{0,1\}^{n\times n}\colon \quad f(X) = f\begin{pmatrix} X & 0 \\ X & 0 \end{pmatrix}.$$

Example 3. Any property of boolean matrices that only depends on the rank of the matrix is stackable. The same is true for support-rank and sign-rank.

Definition 7 (Stability). We say that a function $f\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ is *t-stable* if there do not exist inputs $x_1, y_1 \ldots, x_t, y_t \in \{0,1\}^n$ such that for all $1 \leqslant i, j \leqslant t$ we have

$$f(x_i, y_j) = 1 \iff i \geqslant j.$$

We call a family of functions $\mathcal{F}$ *stable* if there exists a $t \geqslant 2$ such that every $f \in \mathcal{F}$ is t-stable.

If a function is t-stable, then clearly it is also $(t+1)$-stable.

Lemma 7. *Any function $f\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ is t-stable for both*

$$t = \mathrm{rk}_0(f) + 1 \quad \textit{and} \quad t = \mathrm{rk}_0(\neg f) + 2.$$

Proof. Assume that f is not t-stable, witnessed by x_i and y_i with $i \in [t]$. Let $M \in \mathbb{R}^{2^n \times 2^n}$ with $\mathrm{rk}(M) = \mathrm{rk}_0(f)$ and the same support-pattern as f. The sub-matrix of M induced by the x_i and y_i is a triangular matrix, and thus has rank t. We conclude that $t \leqslant \mathrm{rk}(M) \leqslant \mathrm{rk}_0(f)$. Similarly, $t \leqslant \mathrm{rk}_0(\neg f) + 1$, since the rank of the induced sub-matrix in $\neg M$ is $t - 1$. □

Lemma 8. *For a number of sketches k and degree d, a function with $\mathrm{sk}_0 \leqslant k + d$ is $O(k+d)^d$-stable.*

Proof. We know that $\mathrm{rk}_0(f) \leqslant O(k+d)^d$ by Proposition 4. The statement thus follows directly from Lemma 7. □

Definition 8 (Row-Swap). Given matrices $A, B \in \{0,1\}^{n\times n}$, a *row-swap* consists of picking a transposition $\pi = (i, i+1) \in S_n$ of consecutive indices, and replacing A, B with A_π, B_π respectively. The *type* of this row-swap is the tuple $(A_i, A_{i+1}, B_i, B_{i+1}) \in \{0,1\}^{4n}$.

Definition 9 (Row-Shuffle-Invariance). Let $f\colon \{0,1\}^{n\times n} \times \{0,1\}^{n\times n} \to \{0,1\}$ and $\Delta \subseteq \{0,1\}^{4n}$. We say f is *Δ-row-shuffle-invariant* if it is invariant under applying, to its two inputs, any row-swap with type in Δ.
For $f\colon \{0,1\}^{m\times m} \times \{0,1\}^{m\times m} \to \{0,1\}$ and $\Delta \subseteq \{0,1\}^{4n}, m \geqslant n$, we say that f is Δ-row-shuffle-invariant if it is invariant under all row swaps of types $(a0^{m-n}, b0^{m-n}, c0^{m-n}, d0^{m-n})$ for $(a,b,c,d) \in \Delta$. We say f is *row-shuffle-invariant* if it is invariant under row-swaps of all possible types.

Note that a function f is row-shuffle-invariant exactly if it is invariant under shuffling the rows of the two input matrices with the same, arbitrary permutation. We similarly define *column-shuffle-invariance*. If such a function is both row-shuffle-invariant and column-shuffle-invariant, we call it *shuffle-invariant*.

Definition 10 (Xor-Invariance). Let $f\colon \{0,1\}^{n\times n} \times \{0,1\}^{n\times n} \to \{0,1\}$. We say f is Xor-*invariant*, if for any $x, y, z \in \{0,1\}^{n\times n}$ it holds that $f(x,y) = f(x \oplus z, y \oplus z)$.

Definition 11. Let $\mathcal{Q}$ be a family of two-party functions. We define its *boolean closure* $\overline{\mathcal{Q}}$ as the family of all functions $f\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ for which the following holds: There exists a constant $s \in \mathbb{N}$, such that for all $n \in \mathbb{N}$ we have

$$\forall\, x, y \in \{0,1\}^n : \quad f(x,y) = \Gamma\big(q_1(x,y),\ q_2(x,y),\ \ldots\ ,\ q_s(x,y)\big),$$

for some $\Gamma : \{0,1\}^s \to \{0,1\}$ and some $q_1, q_2, \ldots, q_s \in \mathcal{Q}$ (which can both depend on n).

Informally, the following theorem states that if an $f \in \overline{\mathcal{Q}}$ exhibits certain symmetries, then we can express f as a uniform boolean combination of some functions in $\mathcal{Q}$ that share the same symmetries as f.

Theorem 6. *Let $\mathcal{Q}$ be a stable, minor-closed family of boolean two-party functions on bit-strings. Let the function $f\colon \{0,1\}^{n\times n} \to \{0,1\}$ be shuffle-invariant and stackable.*
If $f \circ \textsc{Xor} \in \overline{\mathcal{Q}}$, then there exists a fixed function $\Gamma\colon \{0,1\}^s \to \{0,1\}$ such that for all $n \in \mathbb{N}$, there are shuffle-invariant Xor*-functions $q_1, q_2, \ldots, q_s \in \mathcal{Q}$ satisfying*

$$\forall\, x, y \in \{0,1\}^{n\times n} \ : \quad (f \circ \textsc{Xor})(x,y) = \Gamma\big(q_1(x,y), \ldots, q_s(x,y)\big).$$

We prove Theorem 6 by establishing the following incremental results:

Part 1 The result holds if the $q_i \in \mathcal{Q}$ are allowed to be arbitrary.

Part 2 Using that f is row-shuffle-invariant, the q_i can be made Δ_n-row-shuffle-invariant with

$$\Delta_n = \{(a,b,c,d) \in \{0,1\}^{4n} \mid a = c \text{ or } b = d\}.$$

Part 3 Using that $\mathcal{Q}$ is stable, the q_i can be made fully row-shuffle-invariant.

Part 4 Using that f is column-shuffle-invariant, the q_i can also be made column-shuffle invariant.

Part 5 Using that f is stackable, the q_i can further be made Xor-invariant.

Part 1 holds by the definition of $\overline{\mathcal{Q}}$. For Part 2, we refer the reader to [FGHH25], Proposition 41. The only difference, apart from notational convention, is that we assume $\mathcal{Q}$ to be minor-closed, which is more restrictive than their notion of "query family". However, this does not affect the proofs they give.

Below we refer to a dimension n being (Δ-)shuffle-invariant meaning that q_i for this dimension are (Δ-)shuffle-invariant. We present the reformulation of the statement from [FGHH25] and explain the differences in notation below.

Lemma 9 (Proposition 41 in [FGHH25])**.** *Let $a \in \{0,1\}^n$, $\Delta_a^i = \{(a^i, *, a^i, *)\} \cup \{(*, a^i, *, a^i)\}$, $B \subseteq \{0,1\}^n$ and $\Delta^i = \bigcup_{b\in B} \Delta_b^i$. If for every n, Δ^i-row-shuffle invariance holds for dimension $n \cdot i$, then $\Delta^i \cup \Delta_a^i$-row-shuffle invariance also holds for every $n \cdot i$.*

Notational differences. The proof in [FGHH25] deals with rectangular matrices; one can keep the matrices square by padding them with copies of a, which does not affect the proof. This is because by construction, the padded rows always coincide outside of the embedded smaller matrices, so swapping shorter rows (*dominoes*) equal to $a \in \{0,1\}^n$ in [FGHH25] is the same as our swapping rows from $\{0,1\}^{n\cdot i}$ with swap type in Δ_a^i. The rest of the proof carries over verbatim. □

Applying Lemma 9 repeatedly for every a proves part 2.

The proof of part 3 in [FGHH25] uses the previously established "inner-permutation-invariance", which is particular to the $(4,4)$-HD problem. However, this assumption is not necessary and we give an alternative proof below, which also significantly simplifies this part of the argument in [FGHH25].

Proof of Part 3. We can assume that all dimensions n are already Δ_n-shuffle-invariant.

For a given n, we use Algorithm 1 to make it fully shuffle-invariant. During its runtime, the algorithm extends the sets of shuffle invariance for n as well as for some bounded number of dimensions greater than n. We specify later what the extension procedure itself looks like; the algorithm describes the order of applying it to different dimensions. Let $t \geqslant 2$ be such that $\mathcal{Q}$ is t-stable.

Algorithm 1 Symmetrisation procedure

```
Δ = ∅
for 0 ⩽ i < 2^{4n} do
    Let p_i be the 4n-bit representation of i
    for j ← n + 2t · (2^{4n} − i − 1); j ⩾ n; j −= 2t do
        Make j invariant wrt Δ ∪ Δ_j ∪ p_i
    end for
    Δ ← Δ ∪ p_i
end for
```

Note that when the algorithm tries to ensure the shuffle invariance for some dimension m with respect to $\Delta \cup \Delta_m \cup p$, we have already ensured that $m+2t$ is invariant with respect to $\Delta \cup \Delta_{m+2t}$. We now need to show that we can indeed extend the shuffle invariance for m in such a case, and this will prove the correctness of the procedure.

Let $p = (a, b, c, d)$ and let us pad each row with 0s according to the dimension m. Consider the embedding $\phi\colon \{0,1\}^{m\times m} \to \{0,1\}^{(m+2t)\times(m+2t)}$ that adds to the input matrix t many rows equal to a and t many rows equal to d, padding each row with $2t$ trailing zeros. Concretely, for all $i \in [m+2t]$, we have

$$\phi(x)_i = \begin{cases} x_i\ 0^{2t} & \text{if } i \leqslant m, \\ a\ 0^{2t} & \text{if } m < i \leqslant m+t, \\ d\ 0^{2t} & \text{if } m+t < i \leqslant m+2t. \end{cases}$$

Consider Γ and q'_i satisfying Equation $(*)$ for input matrices of size $(m+2t)\times(m+2t)$ such that each q'_i is $\Delta \cup \Delta_{m+2t}$-row-shuffle-invariant. Now define q_i by

$$\forall x, y \in \{0,1\}^{m\times m}: \quad q_i(x,y) = q'_i(\phi(x), \phi(y)).$$

Since q_i is an affine restriction of q'_i and $\mathcal{Q}$ is minor-closed, we have $q_i \in \mathcal{Q}$. Moreover, q_i inherits $\Delta \cup \Delta_m$-shuffle-invariance from q'_i. We now prove that each q_i is p-row-shuffle-invariant as well. If not, then for some i there exist $x, y \in \{0,1\}^{m\times m}$ and a transposition $\pi = (j, j+1) \in S_m$ with $q_i(x,y) \neq q_i(x_\pi, y_\pi)$ and

$$\begin{aligned} x_j &= a, \quad & y_j &= c, \\ x_{j+1} &= b, \quad & y_{j+1} &= d. \end{aligned}$$

For all $k \in [t+1]$, we now define the matrix $x(k) \in \{0,1\}^{(m+2t)\times(m+2t)}$ as follows: The rows indexed j and $j+1$ are replaced by $2t+2$ alternating copies of a and d. Then, the k-th copy of d we just added is replaced by b. Finally, all rows are padded with $2t$ trailing zeros. Similarly define $y(k)$, where instead the k-th copy of a is replaced by c. Now note that, by Δ-row-shuffle-invariance, for $u, v \in [t+1]$ we have

$$q'_i(x'(u), y'(v)) = \begin{cases} q'_i(\phi(x), \phi(y)) & = q_i(x,y), & \text{if } u \geqslant v \\ q'_i(\phi(x_\pi), \phi(y_\pi)) & = q_i(x_\pi, y_\pi), & \text{if } u < v. \end{cases}$$

But this contradicts t-stability of $\mathcal{Q}$. The matrices in question are illustrated in Figure 9. □

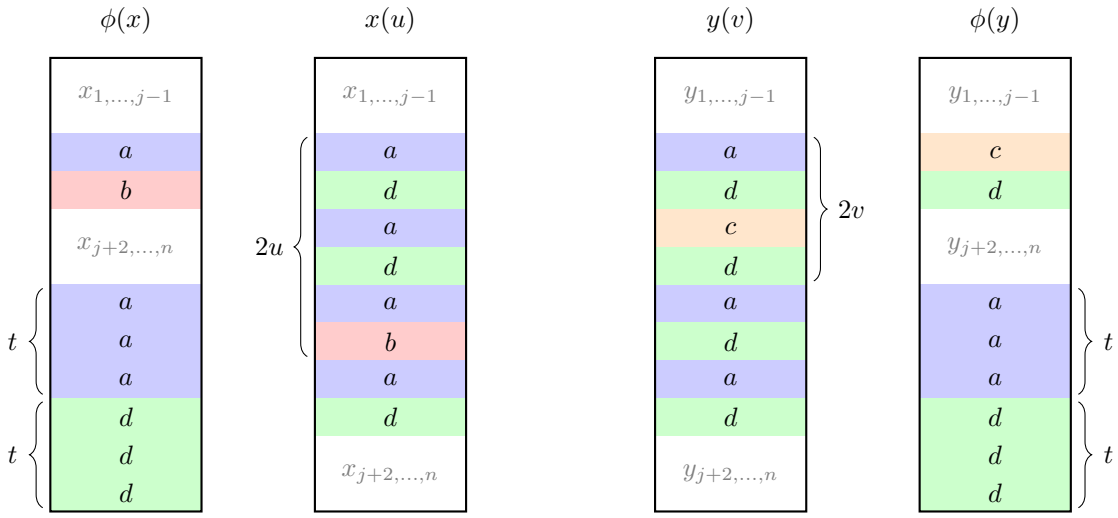

Figure 9: If $u \geqslant v$, the matrices $\phi(u), \phi(u)$ can be obtained from matrices $x(u), y(v)$ with row-swaps of type $(a0^{2t}, *, a0^{2t}, *)$ and $(d0^{2t}, *, d0^{2t}, *)$, which are both in Δ_{m+2t}. If $u < v$, the same is true with $\phi(x_\pi), \phi(y_\pi)$ instead.

Proof of Part 4. Assume that f is also column-shuffle-invariant. For any function $g\colon \{0,1\}^{n\times n} \times \{0,1\}^{n\times n} \to \{0,1\}$, define the function g^T by

$$\forall x, y \in \{0,1\}^{n\times n}: \quad g^T(x,y) = g(x^T, y^T).$$

Note that since f is column-shuffle-invariant, f^T is row-shuffle-invariant. Thus, by the previous part, there exist Γ and row-shuffle-invariant q_i such that

$$(f \circ \textsc{Xor})(x,y) = (f^T \circ \textsc{Xor})(x^T, y^T) = \Gamma\big(q_1(x^T, y^T),\ \ldots,\ q_s(x^T, y^T)\big) = \Gamma\big(q_1^T(x,y),\ \ldots,\ q_s^T(x,y)\big).$$

But note that for each i, the function q_i^T is column-shuffle-invariant. Since transposition of the input is just a reordering of the variables, we also have $q_i^T \in \mathcal{Q}$ by minor-closure. We conclude that part 3 also holds if "row-shuffle-invariant" is replaced by "column-shuffle-invariant". To see why we can guarantee both properties simultaneously, it suffices to notice that the construction in parts 2 and 3 preserves column-shuffle-invariance. □

Proof of Part 5. Assume that f is stackable. For each $i \in [2^n]$, let $\Psi(i) \in \{0,1\}^{n\times n}$ be the matrix with all rows equal to the binary expansion of $i-1$. For $x \in \{0,1\}^{n\times n}$, denote by $\Phi(x) \in \{0,1\}^{n2^n \times n2^n}$ the matrix obtained by stacking the matrices

$$x \oplus \Psi(1),\ x \oplus \Psi(2),\ \ldots,\ x \oplus \Psi(2^n)$$

on top of each other, padding each row with $n(2^n - 1)$ many trailing zeros, as illustrated in Figure 10. We have

$$\forall x, y \in \{0,1\}^{n\times n}: \quad (f \circ \textsc{Xor})(\Phi(x), \Phi(y)) = (f \circ \textsc{Xor})(x,y).$$

Using the result from the previous part, for matrices of size $n2^n \times n2^n$, there are Γ and q_i' for which

$$\forall x, y \in \{0,1\}^{n\times n}: \quad (f \circ \textsc{Xor})(x,y) = \Gamma\big(q_1'\big(\Phi(x), \Phi(y)\big),\ \ldots,\ q_s'\big(\Phi(x), \Phi(y)\big)\big), \tag{$*$}$$

where each q_i' is row-shuffle invariant and column-shuffle-invariant. Define q_i by

$$\forall x, y \in \{0,1\}^{n\times n}: \quad q_i(x,y) = q_i'\big(\Phi(x), \Phi(y)\big).$$

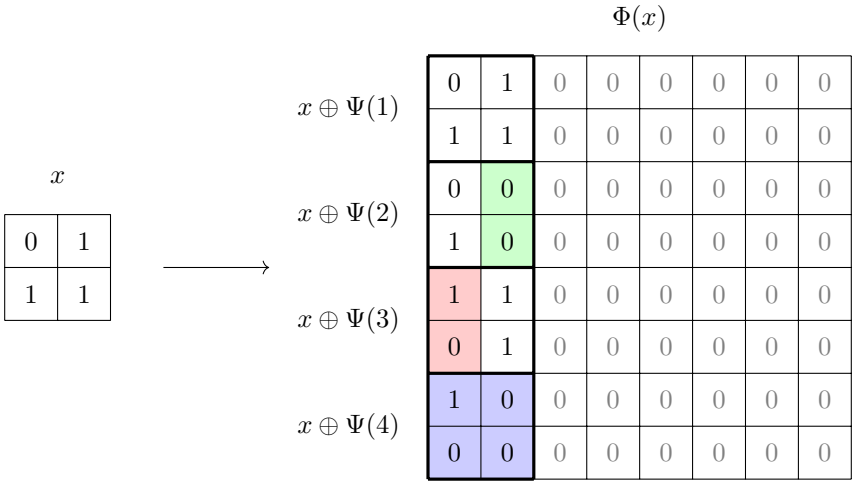

Figure 10: For every subset C of columns of x, the stack in $\Phi(x)$ contains a copy of x with all bits in C flipped (highlighted). The rows of $\Phi(x)$ are padded with zeros to make the matrix square.

Since q_i is an affine restriction of q'_i, we also have $Q_i \in \mathcal{Q}$ by minor-closure. Note that q_i inherits shuffle-invariance from q'_i. Moreover, for any $z \in \{0,1\}^{n \times n}$, the matrix $\Phi(x \oplus z)$ can be obtained from $\Phi(x)$ by permuting its rows. Thus, each q_i is also XOR-invariant, as desired. □

Definition 12 (Promise-Rank). For some integer $2 \leqslant t \leqslant n$, we define the partial function $\text{PRANK}^t_{>1} \colon \{0,1\}^{n \times n} \to \{0,1,*\}$ as follows: For any $x \in \{0,1\}^{n \times n}$,

- if the 1s in x form a $t \times t$ combinatorial rectangle, then $\text{PRANK}^t_{>1}(x) = 0$, else
- if the 1s in x form a $t \times t$ combinatorial rectangle with a 1×1 hole, then $\text{PRANK}^t_{>1}(x) = 1$,
- else $\text{PRANK}^t_{>1}(x) = *$ is unspecified.

This function is essentially the same as $\text{RANK}_{>1}$, but with the promise that the 1s of the input form a $t \times t$ rectangle, possibly with a single entry missing. Note that for this function, all yes-inputs and all no-inputs are in the same orbit of the symmetry group generated by shuffle-invariance. See Figure 11.

Lemma 10. *For any $n/2 \geqslant t \geqslant \Omega(n^{1/6})$, we have*

$$\text{sk}_0(\text{PRANK}^t_{>1} \circ \text{XOR}) \geqslant \Omega(n^{1/6}) \quad \textit{and} \quad \text{sk}_0(\text{PRANK}^t_{\leqslant 1} \circ \text{XOR}) \geqslant \Omega(n^{1/6}).$$

Proof. We first note that the definitions of resilience and self-reducibility can naturally be extended to partial functions. In particular, we have $\text{res}_{G_{\text{RANK}}}(\text{PRANK}^t_{>1}) = t$ (using $n/2 \geqslant t \geqslant 2$) and that $\text{PRANK}^t_{>1}$ is G_{RANK}-self-reducible. Both properties also hold for the negated function $\text{PRANK}^t_{\leqslant 1}$. The statement then follows directly from Lemma 1, whose proof is also valid for for partial functions. □

We can now extend of our main result to the closure of support-sketches.

$x \in \{0,1\}^{6\times 6}$

1	0	1	0	1	1
1	0	1	0	1	1
1	0	1	0	1	1
0	0	0	0	0	0
1	0	1	0	1	1
0	0	0	0	0	0

$y \in \{0,1\}^{6\times 6}$

1	1	0	1	0	1
1	1	0	1	0	1
0	0	0	0	0	0
1	0	0	1	0	1
1	1	0	1	0	1
0	0	0	0	0	0

Figure 11: We have $\mathrm{PRANK}^4_{>1}(x) = 0$ and $\mathrm{PRANK}^4_{>1}(y) = 1$. By permuting rows and columns, all yes-instances and all no-instances can be obtained from x and y, respectively.

Proof of Theorem 2. Suppose $\overline{\mathrm{sk}}_0(\mathrm{RANK}_{>1} \circ \mathrm{XOR}) = k+d$, for some constants k and d. Recall that $\mathrm{RANK}_{>1}$ is shuffle-invariant and stackable. We also saw in Lemma 6 that the family of functions f with $\overline{\mathrm{sk}}_0(f) \leqslant k+d$ is minor-closed and stable. Thus, by Theorem 6, there is a fixed Γ such that

$$\forall\, n \in \mathbb{N} \quad \exists\, q_i, \overline{\mathrm{sk}}_0(q_i) \leqslant k+d : \quad (\mathrm{RANK}_{>1} \circ \mathrm{XOR})(x,y) = \Gamma\big(q_1(x,y),\ \ldots,\ q_s(x,y)\big), \tag{$\dagger$}$$

where each q_i is a shuffle-invariant XOR-function.

Let $n/2 \geqslant t \geqslant \Omega(n^{1/6})$, and consider a matrix $A \in \{0,1\}^{n\times n}$ whose 1-entries form a $t \times t$ rectangle R. Let B be the matrix that differs from A only in the bottom-right corner of R. For $t \geqslant 2$, we clearly have $\mathrm{rk}(A) = 1$ and $\mathrm{rk}(B) = 2$, hence

$$0 = (\mathrm{RANK}_{>1} \circ \mathrm{XOR})(A,0) \neq (\mathrm{RANK}_{>1} \circ \mathrm{XOR})(B,0) = 1.$$

It follows that there is some i in Equation ($\dagger$) for which $q_i(A,0) \neq q_i(B,0)$. But by the symmetries of the q_i, it follows that q_i is a completion of either $\mathrm{PRANK}^t_{>1}$ or its negation. Thus, by Lemma 10,

$$k + d \geqslant \Omega(n^{1/6}),$$

a contradiction for large enough n. □

5 Basic relationships between models

In this section we give simple proofs for the relationships illustrated in Figures 1 and 2.

Proposition 1. *Let f and g be n-bit two-party functions. Then,*

$$\mathrm{sk}_0(f \wedge g) \leqslant \mathrm{sk}_0(f) + \mathrm{sk}_0(g) \quad \textit{and} \quad \mathrm{sk}_0(f \vee g) \leqslant \mathrm{sk}_0(f) + \mathrm{sk}_0(g).$$

Moreover, $\mathrm{sk}_\pm(\neg f) = \mathrm{sk}_\pm(f)$.

Proof. Let P_f and P_g denote the polynomials used in the least-cost support-sketch protocols of f and g, respectively. To bound $\mathrm{sk}_0(f \wedge g)$, note that $P_f P_g$ is non-zero if and only if both P_f and P_g are non-zero.

For $\mathrm{sk}_0(f \vee g)$, note instead that for a suitable $\alpha \in \mathbb{R}$ the polynomial $P_f + \alpha P_g$ vanished on all inputs if and only if both P_f and P_g vanish on all inputs.

For the last property, note that by slightly shifting the polynomial P in the sign-sketch protocol of f, we can assume that it does not vanish on any input. Then, $-P$ yields a sign-sketch protocol for $\neg f$. □

Proposition 2. *For every f we have $\mathrm{sk}_{\pm}(f) \leqslant 2\,\mathrm{sk}_0(f)$.*

Proof. Let P denote the polynomial used in the least-cost support-sketch protocol of f. It suffices to observe that the polynomial P^2 is positive if and only if P is non-zero. □

The following is the sketching analogue of the fact that sign-rank can be bound as a function of support-rank [GHIS25, Lemma II.3].

Proposition 3. *For every f we have $sk_{\pm}(f) \leqslant 2\,\overline{\mathrm{sk}}_0(f)$.*

Proof. Let f be a boolean combination of the two-party functions $f_1, f_2, \ldots f_m$. By induction on m, we show

$$\mathrm{sk}_{\pm}(f) \leqslant 2\sum_{i=1}^{m} \mathrm{sk}_0(f_i),$$

where, crucially, every sketch in the sign-sketch protocol for f also appears as a sketch in the support-protocol for one of the f_i. For $m = 1$, this follows from Proposition 2. For $m > 1$, we write

$$f = (g_1 \wedge f_m) \vee (g_0 \wedge \neg f_m),$$

where g_1 and g_0 are boolean combinations of $f_1, f_2, \ldots, f_{m-1}$. By the inductive hypothesis, there is a polynomial P_1 that achieves a sign-sketch protocol for g_1 of cost at most

$$2\sum_{i=1}^{m-1} \mathrm{sk}_0(f_i),$$

and define P_0 similarly for g_0. Let H be a polynomial with least support-sketch cost for f_m. Now note that for sufficiently large real α, the polynomial $P_0 + \alpha H^2 P_1$ is positive on some input (x, y) if and only if $f(x, y) = 1$. Thus, it gives a sign-sketch protocol for f. By assumption, every sketch in this protocol is a sketch used in the optimal support-sketch protocols of some f_i. Thus, bounding the degree, the cost of the sign-sketch protocol is at most

$$2\,\mathrm{sk}_0(f_m) + 2\sum_{i=1}^{m-1} \mathrm{sk}_0(f_i) = \sum_{i=1}^{m} \mathrm{sk}_0(f_i).$$

□

Proposition 4. *Let f be a two-party function with $\mathrm{sk}_0(f) = k + d$, where k is the number of sketches, and d the degree in the support-sketch protocol. Then,*

$$\mathrm{rk}_0(f) \leqslant O(k + d)^d.$$

The analogous statement involving $\mathrm{sk}_{\pm}$ and $\mathrm{rk}_{\pm}$ holds also.

Proof. Let P denote the $2k$-variate, degree-d polynomial used in the support-sketch protocol for f, and let $A, B\colon \mathbb{R}^n \to \mathbb{R}^k$ be the linear sketches of Alice and Bob. After expanding, P is a sum of $O(k+d)^d$ monomials, which we can each split into a product of two factors, one only depending on Ax, the other only depending on By. Thus, P can be written as an inner product of two real vectors $u(x)$ and $v(y)$, of dimension $O(k+d)^d$ each, yielding the desired support-rank decomposition. The same argument works for the sign-measures. □

Acknowledgements

All ideas in this paper are due to humans; AI was only used for literature search and grammar. We thank Robert Andrews, Nathan Harms, William Lauga, and Rafael Oliveira for discussions. M.G. and A.S. are supported by the Swiss State Secretariat for Education, Research, and Innovation (SERI) under contract number MB22.00026.

References

[ACHS24] Manasseh Ahmed, Tsun-Ming Cheung, Hamed Hatami, and Kusha Sareen. Communication complexity and discrepancy of halfplanes. In *Proceedings of the Symposium on Computational Geometry (SoCG)*, pages 5:1–5:17. Schloss Dagstuhl, 2024. doi:10.4230/LIPIcs.SoCG.2024.5.

[Alo99] Noga Alon. Combinatorial Nullstellensatz. *Combinatorics, Probability and Computing*, 8(1-2):7–29, 1999. doi:10.1017/S0963548398003411.

[AMY16] Noga Alon, Shay Moran, and Amir Yehudayoff. Sign rank versus VC dimension. In *Proceedings of the Conference on Learning Theory (COLT)*, volume 49, pages 47–80. PMLR, 2016. doi:10.48550/arXiv.1503.07648.

[APP+05] Noga Alon, János Pach, Rom Pinchasi, Radoš Radoičić, and Micha Sharir. Crossing patterns of semi-algebraic sets. *Journal of Combinatorial Theory*, 111(2):310–326, 2005. doi:10.1016/j.jcta.2004.12.008.

[BFS86] László Babai, Peter Frankl, and Janos Simon. Complexity classes in communication complexity theory. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS)*, pages 337–347. IEEE, 1986. doi:10.1109/SFCS.1986.15.

[BHH+26a] Ari Blondal, Hamed Hatami, Pooya Hatami, Chavdar Lalov, and Sivan Tretiak. Borsuk-ulam and replicable learning of large-margin halfspaces. In Aditya Bhaskara and Artur Czumaj, editors, *Proceedings of the Symposium on Theory of Computing (STOC)*, pages 529–540. ACM, 2026. doi:10.1145/3798129.3800771.

[BHH+26b] Ari Blondal, Hamed Hatami, Pooya Hatami, Chavdar Lalov, and Sivan Tretiak. Sign-rank, index, and list replicability: Connections and separations. *CoRR*, abs/2606.18236, 2026. doi:10.48550/arXiv.2606.18236.

[BHKR26] Riju Bindua, Hamed Hatami, Hasti Karimi, and Robert Robere. Lower bounds for approximate sign rank. *CoRR*, abs/2605.01038, 2026. doi:10.48550/arXiv.2605.01038.

[BHT25] Igor Balla, Lianna Hambardzumyan, and István Tomon. Factorization norms and an inverse theorem for MaxCut. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS)*, pages 947–963. IEEE, 2025. doi:10.1109/focs63196.2025.00049.

[BNS19] Amos Beimel, Kobbi Nissim, and Uri Stemmer. Characterizing the sample complexity of pure private learners. *Journal of Machine Learning Research*, 20:146:1–146:33, 2019. URL: http://jmlr.org/papers/v20/18-269.html.

[BVdW07] Harry Buhrman, Nikolay Vereshchagin, and Ronald de Wolf. On computation and communication with small bias. In *Proceedings of the Conference on Computational Complexity (CCC)*, pages 24–32. IEEE, 2007. doi:10.1109/CCC.2007.18.

[CLV19] Arkadev Chattopadhyay, Shachar Lovett, and Marc Vinyals. Equality alone does not simulate randomness. In *Proceedings of the Conference on Computational Complexity (CCC)*, pages 14:1–14:11. Schloss Dagstuhl, 2019. doi:10.4230/LIPIcs.CCC.2019.14.

[CMW25] Bogdan Chornomaz, Shay Moran, and Tom Waknine. Spherical dimension. Technical report, arXiv, 2025. doi:10.48550/arXiv.2503.10240.

[DW03] Ronald De Wolf. Nondeterministic quantum query and communication complexities. *SIAM Journal on Computing*, 32(3):681–699, 2003. doi:10.1137/S0097539702407345.

[FGHH25] Yuting Fang, Mika Göös, Nathaniel Harms, and Pooya Hatami. Constant-cost communication does not reduce to k-Hamming distance. In *Proceedings of the Symposium on Theory of Computing (STOC)*, 2025. doi:10.1145/3717823.3718129.

[FHV26] Florian Frick, Kaave Hosseini, and Aliaksei Vasileuski. A $\mathbb{Z}_2$-topological framework for sign-rank lower bounds. Technical report, arXiv, 2026. doi:10.48550/arXiv.2604.01510.

[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. doi:10.1016/S0022-0000(02)00019-3.

[FX14] Vitaly Feldman and David Xiao. Sample complexity bounds on differentially private learning via communication complexity. In *Proceedings of the Conference on Learning Theory (COLT)*, pages 1000–1019. PMLR, 2014. doi:10.48550/arXiv.1402.6278.

[GHIS25] Mika Göös, Nathaniel Harms, Valentin Imbach, and Dmitry Sokolov. Sign-rank of k-Hamming distance is constant. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS)*, pages 2353–2368. IEEE, December 2025. doi:10.1109/focs63196.2025.00123.

[GHR25] Mika Göös, Nathaniel Harms, and Artur Riazanov. Equality is far weaker than constant-cost communication. In *International Conference on Randomization and Computation (RANDOM)*, volume 353 of *LIPIcs*, pages 58:1–58:14. Schloss Dagstuhl, 2025. doi:10.4230/LIPIcs.APPROX/RANDOM.2025.58.

[GHR+26] Mika Göös, Nathaniel Harms, Artur Riazanov, Anastasia Sofronova, Dmitry Sokolov, and Weiqiang Yuan. Pseudodeterministic communication complexity. In *Proceedings of the Symposium on Theory of Computing (STOC)*, pages 2030–2039, 2026. doi:10.1145/3798129.3800907.

[GKPW18] Mika Göös, Pritish Kamath, Toniann Pitassi, and Thomas Watson. Query-to-communication lifting for P^{NP}. *Computational Complexity*, 28(1):113–144, 2018. doi:10.1007/s00037-018-0175-5.

[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. doi:10.1145/3654780.3654788.

[HHH23] Lianna Hambardzumyan, Hamed Hatami, and Pooya Hatami. Dimension-free bounds and structural results in communication complexity. *Israel Journal of Mathematics*, 253(2):555–616, 2023. doi:10.1007/s11856-022-2365-8.

[HHL20] Hamed Hatami, Kaave Hosseini, and Shachar Lovett. Sign rank vs discrepancy. In *Proceedings of the Conference on Computational Complexity (CCC)*, pages 18–1. Schloss Dagstuhl, 2020. doi:10.4230/LIPIcs.CCC.2020.18.

[HHM23] Hamed Hatami, Kaave Hosseini, and Xiang Meng. A Borsuk-Ulam lower bound for sign-rank and its applications. In *Proceedings of the Symposium on Theory of Computing (STOC)*, pages 463–471, 2023. doi:10.1145/3564246.3585210.

[HHP+22] Hamed Hatami, Pooya Hatami, William Pires, Ran Tao, and Rosie Zhao. Lower bound methods for sign-rank and their limitations. In *International Conference on Randomization and Computation (RANDOM)*, pages 22–1. Schloss Dagstuhl, 2022. doi:10.4230/LIPIcs.APPROX/RANDOM.2022.22.

[HWZ22] Nathaniel Harms, Sebastian Wild, and Viktor Zamaraev. Randomized communication and implicit graph representations. In *Proceedings of the Symposium on Theory of Computing (STOC)*, 2022. doi:10.1145/3519935.3519978.

[LMSS07] Nati Linial, Shahar Mendelson, Gideon Schechtman, and Adi Shraibman. Complexity measures of sign matrices. *Combinatorica*, 27:439–463, 2007. doi:10.1007/s00493-007-2160-5.

[MP69] Marvin Minsky and Seymour Papert. *Perceptrons: An Introduction to Computational Geometry*. The MIT Press, 1969. doi:10.7551/mitpress/11301.001.0001.

[New91] Ilan Newman. Private vs. common random bits in communication complexity. *Information Processing Letters*, 39(2):67–71, 1991. doi:10.1016/0020-0190(91)90157-D.

[Nis91] Noam Nisan. Lower bounds for non-commutative computation. In *Proceedings of the Symposium on Theory of Computing (STOC)*, pages 410–418. ACM, 1991. doi:10.1145/103418.103462.

[PS86] Ramamohan Paturi and Janos Simon. Probabilistic communication complexity. *Journal of Computer and System Sciences*, 33(1):106–123, 1986. doi:10.1016/0022-0000(86)90046-2.

[Sax14] Nitin Saxena. Progress on polynomial identity testing-ii. In *Perspectives in Computational Complexity: The Somenath Biswas Anniversary Volume*, pages 131–146. Springer, 2014. doi:10.1007/978-3-319-05446-9_7.

[She08] Alexander A Sherstov. Halfspace matrices. *Computational Complexity*, 17(2):149–178, 2008. doi:10.1007/s00037-008-0242-4.

[She11] Alexander A Sherstov. The pattern matrix method. *SIAM Journal on Computing*, 40(6):1969–2000, 2011. doi:10.1137/080733644.

[She13] Alexander A. Sherstov. Optimal bounds for sign-representing the intersection of two halfspaces by polynomials. *Combinatorica*, 33(1):73–96, 2013. doi:10.1007/s00493-013-2759-7.

[She21] Alexander A Sherstov. The hardest halfspace. *computational complexity*, 30(2):11, 2021. doi:10.1007/s00037-021-00211-4.

[SS11] Nitin Saxena and Comandur Seshadhri. An almost optimal rank bound for depth-3 identities. *SIAM journal on computing*, 40(1):200–224, 2011. doi:10.1137/090770679.

[SS24] Alexander Sherstov and Andrey Storozhenko. The communication complexity of approximating matrix rank. In *Proceedings of the Symposium on Foundations of Computer Science (FOCS)*, pages 433–462. IEEE, 2024. doi:10.1109/FOCS61266.2024.00035.

[SY10] Amir Shpilka and Amir Yehudayoff. Arithmetic circuits: A survey of recent results and open questions. *Foundations and Trends in Theoretical Computer Science*, 5(3-4):207–388, 2010. doi:10.1561/0400000039.

[Tha16] Justin Thaler. Lower bounds for the approximate degree of block-composed functions. In *Proceedings of the International Colloquium on Automata, Languages, and Programming (ICALP)*. Schloss Dagstuhl, 2016. doi:10.4230/LIPIcs.ICALP.2016.17.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il