Quantum–Classical Equivalence for AND-Functions

Sreejata Kishor Bhattacharya
TIFR, Mumbai, India

Farzan Byramji
UC San Diego, La Jolla, CA, USA

Arkadev Chattopadhyay
TIFR, Mumbai, India

Yogesh Dahiya
UC San Diego, La Jolla, CA, USA

Shachar Lovett
UC San Diego, La Jolla, CA, USA

Abstract

A major open problem in quantum communication complexity is whether quantum protocols can be exponentially more efficient than classical protocols for computing total Boolean functions; the prevailing conjecture is that they cannot be so. In a seminal work, Razborov (2002) resolved this question for AND-functions of the form

$$F(x, y) = f(x_1 \wedge y_1, \ldots, x_n \wedge y_n),$$

when the outer function f is symmetric, by proving that their bounded-error quantum and classical communication complexities are polynomially related. Since then, extending this result to *all* AND-functions has remained open and has been posed by several authors.

In this work, we settle this problem in a strong way. We show that for every Boolean function f, the bounded-error quantum and classical deterministic communication complexities of the function $f \circ \text{AND}_2$ are polynomially related, up to polylogarithmic factors in n. We prove this by showing that both are characterized – up to polynomial loss – by the logarithm of the De Morgan sparsity of f.

Our results build on the recent work of Chattopadhyay, Dahiya, and Lovett [8] on structural characterizations of non-sparse Boolean functions, which we extend to resolve the conjecture for general AND-functions.

2012 ACM Subject Classification Theory of computation → Communication complexity; Theory of computation → Quantum communication complexity

Keywords and phrases Communication complexity, quantum communication complexity, De Morgan sparsity, approximate gamma two norm, And-functions

Digital Object Identifier 10.4230/LIPIcs.CCC.2026.23

Related Version *Full Version*: https://eccc.weizmann.ac.il/report/2026/013/

Funding *Sreejata Kishor Bhattacharya*: Supported by the Department of Atomic Energy, Govt. of India, under project #RTI4014 and by a Google PhD Fellowship.
Farzan Byramji: Supported by Simons Investigator Award #929894, and NSF awards CCF-2425349 and AF: Medium 2212136.
Arkadev Chattopadhyay: Supported by the Department of Atomic Energy, Govt. of India, under project #RTI4014 and by a Google India Faculty Award.
Yogesh Dahiya: Supported by Simons Investigator Award #929894 and NSF award CCF-2425349.
Shachar Lovett: Supported by Simons Investigator Award #929894 and NSF award CCF-2425349.

41st Computational Complexity Conference (CCC 2026).
Editor: Dana Moshkovitz; Article No. 23; pp. 23:1–23:24

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

Understanding when/if the laws of quantum mechanics can be exploited to provide significant computational advantage is a major research theme. Two classical results show that indeed quantum algorithms can have surprising power. Shor's polynomial time algorithm on quantum computers for integer factoring beats all *known* classical factoring algorithms by an *exponential* speed-up. Grover search in the quantum query model beats all *possible* classical algorithms by a *quadratic* factor. The first result on exponential speed-up is *conditional* on the hardness of factoring by classical computers. The second result is provable advantage but only quadratic and we know no better is possible in that query model. In this work, we're concerned with provable exponential advantage for quantum models when the task is to compute a total Boolean function. It is well known that both the factoring problem and the search problem for a marked item that Grover's algorithm solves, can readily be re-phrased in terms of total Boolean functions.

Communication complexity, introduced by Yao [41], studies the amount of communication required to compute a function whose input is distributed among multiple parties. Since its inception, communication complexity has become a central tool in theoretical computer science, with applications ranging from streaming algorithms and time–space tradeoffs to data structure lower bounds and circuit complexity. See the textbooks [22, 28] for excellent introductions to the area and its applications.

In the most standard setting – the two-party model – two players, Alice and Bob, wish to compute a Boolean function $F : X \times Y \to \{0, 1\}$, where Alice receives $x \in X$ and Bob receives $y \in Y$. They exchange messages according to a pre-agreed protocol in order to compute $F(x, y)$, with the goal of minimizing the total number of bits communicated in the worst case.

Several variants of communication complexity arise depending on the type of interaction allowed and whether the protocol is permitted to err with small probability. In this work, we focus on three such models: two classical models and one quantum model.

In the classical deterministic model, the *deterministic communication complexity* of F, denoted $\mathrm{D}^{cc}(F)$, is the minimum number of bits that must be exchanged by a protocol that computes $F(x, y)$ correctly on all inputs. In the public-coin randomized model, Alice and Bob have access to shared public randomness and are required to compute $F(x, y)$ with error probability at most $1/3$ on every input; the corresponding complexity measure is the *randomized communication complexity*, denoted $\mathrm{R}^{cc}(F)$.

The quantum variant of communication complexity, also introduced by Yao [42], allows the messages exchanged between the parties to be quantum states (qubits). At the end of the protocol, one of the parties performs a measurement on its quantum state to produce the output. In addition, the parties may share an arbitrary entangled state prior to the start of the protocol, at no communication cost [10]. The number of qubits exchanged by the best such protocol that computes $F(x, y)$ with error probability at most $1/3$, for the worst input, is called the *bounded-error quantum communication complexity* of F, and is denoted $\mathrm{Q}^{cc}(F)$.

A central question in quantum computation is to understand when quantum models enable efficient solutions to problems that are believed to be hard for classical models. In the context of communication complexity, this question asks when quantum protocols can be super-polynomially more efficient than classical (randomized) protocols, and has been a major driving force behind research in the area.

Exponential quantum advantages were discovered early on for structured problems arising from partial functions, sampling problems and relations, see [29, 4, 2, 14, 32]. This understanding has been refined over a number of papers, see [13, 12, 15], with latest work

exhibiting exponential quantum advantage over such promised problems even when the quantum protocol is extremely restricted and no restrictions are imposed on competing classical protocols. In a very recent work, exponential quantum advantage was shown for a total search problem in TFNP by Goos et. al. [16]. However, as re-iterated by [16], no such separation is known for *total* Boolean functions which remains the holy grail of the search for quantum advantage. The largest gap currently known between bounded-error quantum and randomized communication complexity for a total function is only polynomial: a cubic separation (up to polylogarithmic factors), obtained by lifting cubic separations between quantum and randomized query complexity [3, 39].

These results have led to the prevailing belief that, in the absence of special structure – most notably for total Boolean functions – quantum and randomized communication complexities are always polynomially related. This belief was explicitly formulated by Shi and Zhu [40] as the *Log-Equivalence Conjecture* (LEC).

▶ **Conjecture 1.1** (Log-Equivalence Conjecture (LEC) [40]). *For every total Boolean function, the bounded-error quantum and randomized communication complexities are polynomially related in the two-party communication model.*

Despite intensive research efforts, the problem of proving or disproving the LEC remains wide open. In light of the lack of progress on general total functions, several authors [5, 31, 18, 40] have proposed studying *composed* communication problems of the form $F = f \circ g$, where $g : \{0,1\}^b \times \{0,1\}^b \to \{0,1\}$ is a small, preferably constant-size, gadget. Among such restricted classes, AND-functions of the form

$$F(x,y) = f(x_1 \wedge y_1, \ldots, x_n \wedge y_n)$$

have received particular attention. This is partly motivated by the fact that some of the most studied problems in communication complexity, such as *Set Disjointness* and *Inner Product*, are AND-functions.

To the best of our knowledge, AND-functions, especially in the context of quantum communication complexity, were first systematically studied by Buhrman and de Wolf [5]. They showed that for this class, deterministic communication complexity and zero-error quantum communication complexity are polynomially related when the outer function f is symmetric or monotone. However, the relationship between bounded-error quantum and randomized communication complexity for AND-functions remained poorly understood at the time. Indeed, very few lower bounds were known against bounded-error quantum protocols, essentially limited to those – such as Inner Product – obtained via the discrepancy method [21]. In particular, no polynomial lower bound on the quantum communication complexity of Set Disjointness was known.

In a major breakthrough, Razborov [31] established the optimal $\Omega(\sqrt{n})$ lower bound for the bounded-error quantum communication complexity of Set Disjointness over a universe of size n. More generally, his method yielded that for AND-functions $f \circ \text{AND}_2$ with symmetric outer function f, the bounded-error quantum communication complexity is polynomially equivalent even to the deterministic communication complexity. Extending this result to *all* AND-functions remained open since then and this has been investigated by several authors [5, 31, 18, 35, 40, 37]. Our main result resolves this problem.

▶ **Theorem 1.2.** *Let $f : \{0,1\}^n \to \{0,1\}$ be any Boolean function. Then:*

1. $D^{cc}(f \circ \text{AND}_2) = O\big(Q^{cc}(f \circ \text{AND}_2)^7 \cdot (\log n)^2\big)$.
2. $D^{cc}(f \circ \text{AND}_2) = O\big(R^{cc}(f \circ \text{AND}_2)^5 \cdot (\log n)^2\big)$.

Ignoring polylogarithmic factors in n, this shows that for all AND-functions, deterministic communication complexity is polynomially equivalent to bounded-error quantum communication complexity. Notably, prior to our work it was not even known whether randomized communication complexity is polynomially equivalent to deterministic complexity for this class of functions. Before describing our results, we briefly review prior developments related to this question.

While not stated explicitly in Razborov's work, his quantum lower bound for Set Disjointness can be used to obtain lower bounds for a broader class of composed functions. In particular, consider any constant-size gadget $g : \{0,1\}^b \times \{0,1\}^b \to \{0,1\}$ whose communication matrix contains both AND_2 and OR_2 as submatrices. For such gadgets, the bounded-error quantum communication complexity of $f \circ g$ is $\Omega(\sqrt{\text{bs}(f)})$ [43], where $\text{bs}(f)$ denotes the block sensitivity of f. This follows from the fact that a promised Set Disjointness instance of size $\text{bs}(f)$ can be embedded into $f \circ g$.

Combining this lower bound with Nisan's classical result [27], which upper bounds the deterministic query complexity of f by a polynomial in $\text{bs}(f)$, yields a quantum–classical equivalence for composed functions $f \circ g$ whenever g embeds both AND_2 and OR_2. We refer to this class as AND-OR-*functions*.

Subsequently, using different techniques, Sherstov [35] gave an independent proof of quantum–classical equivalence for AND-OR-functions via his pattern matrix method. In related and independent work, Shi and Zhu [40] proved quantum–classical equivalence for composed functions with gadgets satisfying certain pseudorandomness properties, although their gadgets were required to have size $\Omega(\log n)$.

Another important line of work concerns XOR-functions, i.e., functions composed with the XOR_2 gadget. Shi and Zhang [44] showed that, up to polylogarithmic factors, the Log-Equivalence Conjecture holds for XOR-functions when the outer function f is symmetric. Subsequently, Montanaro and Osborne [26] proved a polynomial equivalence between deterministic and zero-error quantum communication complexity for XOR-functions with monotone outer functions. With the AND-function case resolved in this work, extending the Log-Equivalence Conjecture to XOR-functions in full generality emerges as a natural next milestone toward a complete understanding of the conjecture.

More recently, Chattopadhyay, Dahiya, and Lovett [8] revisited the Log-Equivalence Conjecture. In all previously known classes of functions satisfying the conjecture, an even stronger statement holds: deterministic communication complexity is polynomially related to bounded-error quantum communication complexity. In [8], the authors studied composed functions of the form $f \circ \text{EQ}_4$, where EQ_4 denotes equality on four bits. Since $\text{AND} \circ \text{EQ}_4$ (which is just the equality function) belongs to this class, deterministic communication complexity is exponentially separated from randomized communication complexity; nevertheless, they showed that the Log-Equivalence Conjecture continues to hold (up to polylogarithmic factors). Our work builds on and extends the ideas developed in [8], which we elaborate on in subsequent sections.

1.1 Our Results

For any Boolean function $f : \{0,1\}^n \to \{0,1\}$, our main result shows that the bounded-error quantum and deterministic communication complexities of the AND-function $f \circ \text{AND}_2$ are polynomially related, up to polylogarithmic factors in n. That is,

$$\text{D}^{cc}(f \circ \text{AND}_2) = \text{Q}^{cc}(f \circ \text{AND}_2)^{O(1)} \cdot (\log n)^{O(1)}.$$

To establish such a relationship, it is helpful to characterize $\mathrm{D}^{cc}(f \circ \textsc{And}_2)$ in terms of structural properties of the outer function f. This was achieved by Knop, Lovett, McGuire, and Yuan [19], who showed that the deterministic communication complexity of $f \circ \textsc{And}_2$ is characterized – up to polynomial loss and ignoring poly-$\log(n)$ factors – by the logarithm of the *De Morgan sparsity* of f.

Recall that every Boolean function f admits a unique multilinear polynomial representation over the reals,

$$f(x) = \sum_{S \subseteq [n]} a_S \prod_{i \in S} x_i.$$

The (De Morgan) *sparsity* of f, denoted $\mathrm{spar}(f)$, is the number of nonzero coefficients a_S.

Knop et al. [19] showed that the deterministic communication complexity of an $\textsc{And}$-function satisfies

$$\mathrm{D}^{cc}(f \circ \textsc{And}_2) = O\big((\log \mathrm{spar}(f))^5 \cdot \log n\big).$$

This characterization has an immediate and important consequence. Since the sparsity of f coincides with the rank of the communication matrix of $f \circ \textsc{And}_2$, this bound yields a resolution of the log-rank conjecture for $\textsc{And}$-functions, up to a $\log n$ factor.

Given this characterization of $\mathrm{D}^{cc}(f \circ \textsc{And}_2)$, our task reduces to showing that large sparsity of f forces large bounded-error quantum communication complexity. Concretely, we show that

$$\mathrm{Q}^{cc}(f \circ \textsc{And}_2) = (\log \mathrm{spar}(f))^{\Omega(1)},$$

ignoring polylogarithmic factors in n. In fact, we prove a stronger statement: large sparsity of f implies a large *approximate* γ_2 norm of the communication matrix $M_{f \circ \textsc{And}_2}$. This is a strengthening of a direct lower bound on quantum communication complexity, since the approximate γ_2 norm is a known lower bound on bounded-error quantum communication complexity.

The approximate γ_2 norm of a two-party Boolean function $F : X \times Y \to \{0,1\}$, denoted $\widetilde{\gamma}_2(F)$, is defined as the minimum total weight of a rectangle decomposition that approximates the communication matrix M_F within constant error:

$$\widetilde{\gamma}_2(F) = \min\Big\{ \sum_i |\alpha_i| \;\Big|\; \big\| M_F - \sum_i \alpha_i R_i \big\|_\infty \le 1/3 \Big\},$$

where each $R_i(x,y) = g_i(x)h_i(y)$ is a combinatorial rectangle.

Any bounded-error quantum protocol for F using c qubits of communication induces a pointwise approximation of M_F of the form

$$M_F \approx \sum_i \alpha_i R_i,$$

with total weight $\sum_i |\alpha_i| \le 2^{O(c)}$. While such decompositions are immediate for deterministic and randomized protocols, they also hold in the quantum setting by unpacking the definition of quantum communication protocols (see [25]).

Our main technical contribution is to show that large sparsity of f lifts to a lower bound on the approximate γ_2 norm of $f \circ \textsc{And}_2$.

▶ **Theorem 1.3.** *For every total Boolean function $f : \{0,1\}^n \to \{0,1\}$,*

$$\log \widetilde{\gamma}_2(f \circ \textsc{And}_2) = \Omega\left(\left(\frac{\log \mathit{spar}(f)}{\log n} \right)^{1/3} \right).$$

We now record several immediate consequences of Theorem 1.3. Recall that for a communication problem $F : X \times Y \to \{0,1\}$, the *rank* of F is the rank (over the reals) of its communication matrix M_F, and the *approximate rank*, denoted $\widetilde{\mathrm{rank}}(F)$, is the minimum rank of a real matrix that approximates M_F entrywise within a small error $1/3$.

The approximate γ_2 norm can be viewed as a convex relaxation of the approximate rank and is known to be essentially equivalent to it. It is shown in [23] that, on the logarithmic scale, the two measures coincide up to an additive $O(\log\log|X||Y|)$ term: for every communication problem $F : X \times Y \to \{0,1\}$,

$$\Omega\big(\log\widetilde{\gamma}_2(F)\big) \;\le\; \log\widetilde{\mathrm{rank}}(F) \;\le\; O\big(\log\widetilde{\gamma}_2(F) + \log\log|X||Y|\big).$$

Combining this observation with our result and the result of Knop et al. [19], we obtain

$$\begin{aligned}\mathrm{R}^{cc}(f\circ\textsc{And}_2) \le \mathrm{D}^{cc}(f\circ\textsc{And}_2) &= O\big((\log\mathrm{spar}(f))^5\cdot\log n\big)\\ &= O\Big((\log\widetilde{\mathrm{rank}}(f\circ\textsc{And}_2))^{15}\cdot(\log n)^6\Big).\end{aligned}$$

While the exponents above are not optimized and can be improved (see Section 4.2), this implies that the *log-approximate-rank conjecture* holds for AND-functions, up to poly-log(n) factors.

The Log-Approximate-Rank Conjecture (LARC), a term first coined by Lee and Shraibman [24], is a natural approximate analogue of the classical Log-Rank conjecture. It asserts that, just as the logarithm of the rank of the communication matrix is believed to characterize deterministic communication complexity up to polynomial loss, the logarithm of the approximate rank should similarly characterize randomized communication complexity up to polynomial loss.

Somewhat surprisingly, the LARC was recently shown to be false by Chattopadhyay, Mande, and Sherif [9], by exhibiting an XOR-function as a counterexample. In contrast, our results show that LARC *does* hold for AND-functions, up to polylogarithmic factors in n.

In fact, our results give the following chain of inequalities:

$$\begin{aligned}&\Omega\big(\log\widetilde{\gamma}_2(f\circ\textsc{And}_2)\big) \;\le\; \mathrm{Q}^{cc}(f\circ\textsc{And}_2) \;\le\; \mathrm{R}^{cc}(f\circ\textsc{And}_2) \;\le\; \mathrm{D}^{cc}(f\circ\textsc{And}_2),\\ &\qquad \mathrm{D}^{cc}(f\circ\textsc{And}_2) \;\le\; (\log\mathrm{spar}(f))^{O(1)}(\log n)^{O(1)} \;\le\; \log\widetilde{\gamma}_2(f\circ\textsc{And}_2)^{O(1)}(\log n)^{O(1)}.\end{aligned}$$

As a result, for an AND-function $F := f\circ\textsc{And}_2$, up to polylogarithmic factors in n, the quantities

$$\mathrm{D}^{cc}(F),\ \mathrm{R}^{cc}(F),\ \mathrm{Q}^{cc}(F),\ \log\widetilde{\gamma}_2(F),\ \log\widetilde{\mathrm{rank}}(F),\ \log\mathrm{rank}(F),\ \text{and}\ \log\mathrm{spar}(f)$$

are all polynomially equivalent. In other words, for AND-functions, diverse notions of complexity – from communication-theoretic to algebraic/analytic – coincide. Interestingly, this striking equivalence of various measures was suspected in the early work of Buhrman and de Wolf [5] on AND-functions more than two decades ago. Our results finally confirm their intuition.

1.2 Proof overview

As discussed above, our main technical contribution is Theorem 1.3, which lifts the sparsity of a Boolean function f to a lower bound on the approximate γ_2-norm of the lifted function $f\circ\textsc{And}_2$. We now give a high-level overview of the proof.

We begin with a concrete example that will later help illustrate the general argument. The canonical Boolean function with large sparsity is the n-bit $\textsc{Or}_n$ function, whose sparsity equals 2^n-1. When composed with the $\textsc{And}_2$ gadget, the function $\textsc{Or}_n\circ\textsc{And}_2$ corresponds to the *Set Intersection* problem, the negation of the well-known Set Disjointness function.

Lower bound for Set Intersection. We illustrate our approach using the Set Intersection function

$$(\mathrm{OR}_n \circ \mathrm{AND}_2)(x, y) \;=\; \mathrm{OR}_n(x_1 \wedge y_1, \ldots, x_n \wedge y_n),$$

where we write $z_i := x_i \wedge y_i$.

The key structural property of OR_n is that it remains an OR under every restriction $\rho_z \in \{0, *\}^n$: the restricted function retains full degree on the free variables. This hardness under $\{0, *\}$-restrictions is precisely what gives OR_n its large sparsity. More importantly, as we show later, the existence of many such *max-degree* restrictions is a general consequence of large sparsity, allowing the argument to extend beyond Set Intersection to arbitrary AND-functions.

To lift this hardness to the communication setting, we map restrictions on the z-variables to restrictions on the input pairs (x_i, y_i). Given $\rho_z \in \{0, *\}^n$, we define a lifted restriction ρ by

$$(\rho(x_i), \rho(y_i)) = \begin{cases} (\Delta, 0), & \text{if } \rho_z(z_i) = 0, \\ (*, 1), & \text{if } \rho_z(z_i) = *, \end{cases}$$

so that $(x_i \wedge y_i)|_\rho = \rho_z(z_i)$. Here Δ denotes a free but *masked* variable: although syntactically free, the restricted function does not depend on it.

These lifted restrictions have three crucial properties: (i) all y-variables are fixed, (ii) the restricted function $(\mathrm{OR}_n \circ \mathrm{AND}_2)|_\rho$ computes an OR on exactly the $*$-variables, and (iii) it is independent of the masked variables. As a result, if ρ_z has exactly d stars, then

$$\deg\big(\mathbb{E}_{x_{M(\rho)}}[(\mathrm{OR}_n \circ \mathrm{AND}_2)|_\rho]\big) = d,$$

where $M(\rho)$ denotes the masked variables and $\mathbb{E}_{x_{M(\rho)}}[\cdot]$ denotes expectation over independent uniform assignments to them.

This motivates the following *restriction-and-averaging procedure*. We sample an unlifted restriction $\rho_z \in \{0, *\}^n$ uniformly at random subject to having $d = \Theta(n^{2/3})$ stars (we will explain shortly why this specific parameter is chosen), lift it to obtain ρ, and then take expectation over the masked variables.

Under this restriction-and-averaging procedure, $\mathrm{OR}_n \circ \mathrm{AND}_2$ retains its hardness – in terms of degree – with exact degree $\Theta(n^{2/3})$. On the other hand, we show that the same procedure simplifies any small-weight rectangle decomposition arising from a small approximate γ_2 norm. Concretely, it converts such a decomposition into a low-degree polynomial (more precisely, one whose total Fourier mass on high-degree monomials is small).

This leads to a contradiction. Indeed, from a small approximate γ_2 norm of size $2^{O(n^{1/3})}$ for $\mathrm{OR}_n \circ \mathrm{AND}_2$, we would obtain a polynomial of degree $O(n^{1/3})$ that approximates $\mathbb{E}_{x_{M(\rho)}}[(\mathrm{OR}_n \circ \mathrm{AND}_2)|_\rho]$, even though the latter has exact degree $\Theta(n^{2/3})$. This contradicts the known quadratic relationship between degree and approximate degree, forcing the approximate γ_2 norm of $\mathrm{OR}_n \circ \mathrm{AND}_2$ to be $2^{\Omega(n^{1/3})}$. Extracting low-degree approximating polynomials from a decomposition exhibiting low approximate γ_2 norm of a matrix, in the context of communication complexity, was initiated in the work of Razborov [31], followed by other works, including that of [6, 38]. Our particular method is inspired by the work of [8] and the work of Krause and Pudlák [20] who used it in the context of proving lower bounds on the Fourier sparsity of lifted functions.

We now briefly explain why our restriction-and-averaging procedure simplifies a small-weight rectangle decomposition into a low-degree polynomial. Let

$$\Pi(x,y) \;=\; \sum_{i=1}^{m} b_i\, g_i(x)\, h_i(y)$$

be an approximator for Set Intersection with small total weight $\sum_i |b_i|$. Under any lifted restriction, all y-variables are fixed, so each rectangle collapses to a function of the x-variables alone. Moreover, for any Boolean function $g : \{0,1\}^n \to \{0,1\}$ on x-variables, we show that the expected Fourier mass of g above level $k \geq n^{1/3}$ after restriction and averaging decays as $2^{-\Omega(k)}$ (this is where our choice of $d = \Theta(n^{2/3})$ plays a part). Intuitively, this happens because any Fourier character involving a masked variable vanishes upon averaging. An averaging argument then yields a restriction ρ such that the Fourier mass of $\mathbb{E}_{x_{M(\rho)}}[\Pi|_\rho]$ above level $k = O(n^{1/3})$ is negligible, provided $\sum_i |b_i| \leq 2^{O(n^{1/3})}$. Discarding higher-degree monomials produces a low-degree $O(n^{1/3})$ approximator for $\mathbb{E}_{x_{M(\rho)}}[(\mathrm{OR}_n \circ \mathrm{AND}_2)|_\rho]$, contradicting its degree $\Theta(n^{2/3})$.

This shows a $2^{\Omega(n^{1/3})}$ lower bound on the approximate γ_2 norm of Set Intersection. Moreover, if you observe closely, the same argument also yields a $2^{\Omega(n^{1/3})}$ lower bound for the approximate γ_2 norm of Inner Product function $\mathrm{XOR}_n \circ \mathrm{AND}_2$, another central problem in communication complexity.

Generalization to arbitrary functions with large sparsity. In the discussion above for Set Intersection and Inner Product, we relied crucially on a structural property of the outer functions OR_n and XOR_n: under every restriction in $\{0,*\}^n$, the restricted function retains full degree on the surviving free variables. As hinted earlier, the existence of many such *max-degree* restrictions is not specific to these functions. Rather, it is a general consequence of large sparsity – and this is the key insight of our work that allows us to extend the argument to arbitrary AND-functions.

More concretely, we show that any Boolean function $f : \{0,1\}^n \to \{0,1\}$ of large sparsity admits a fixed set of variables $V \subseteq [n]$, which we call the *core variables*, with the following property. For every assignment $\alpha \in \{0,*\}^V$, the remaining variables in $[n] \setminus V$ can be fixed – possibly depending on α – so that the resulting restriction of f has full degree on the free variables. We refer to the resulting collection of restrictions as a *semi-adaptive max-degree restriction tree*. For OR_n and XOR_n, the core set is simply $V = [n]$, but in general this need not be the case. For a function of sparsity s, we show that one can always find V of size $\Omega(\log s / \log n)$, which we call the *depth* of the restriction tree.

Having obtained such a collection of restrictions $\mathcal{D}$ for f, we lift them to restrictions for the two-party function $f \circ \mathrm{AND}_2$, extending the construction used for Set Intersection. Given a restriction $\rho_z \in \mathcal{D}$, we define a lifted restriction ρ on the variables (x_i, y_i) by

$$(\rho(x_i), \rho(y_i)) = \begin{cases} (\Delta, 0), & \text{if } \rho_z(z_i) = 0 \text{ and } i \in V, \\ (*, 1), & \text{if } \rho_z(z_i) = * \text{ and } i \in V, \\ (1, 1), & \text{if } \rho_z(z_i) = 1 \text{ and } i \notin V, \\ (1, 0), & \text{if } \rho_z(z_i) = 0 \text{ and } i \notin V, \end{cases}$$

ensuring that $(x_i \wedge y_i)|_\rho = \rho_z(z_i)$. As in the Set Intersection case, this construction guarantees that $(f \circ \mathrm{AND}_2)|_\rho$ coincides with $f|_{\rho_z}$ (up to renaming of variables) and is independent of the masked variables $M(\rho)$. As a result, for every max-degree restriction ρ_z,

$$\deg\big(\mathbb{E}_{x_{M(\rho)}}\big[(f \circ \mathrm{AND}_2)|_\rho\big]\big) = \deg(f|_{\rho_z}) = |\rho_z^{-1}(*)| = |\rho^{-1}(*)|.$$

Thus, the lifted function $f \circ \textsc{And}_2$ retains its hardness under the same restriction-and-averaging procedure used for Set Intersection. Moreover, the lifting has an additional convenient feature: all non-core x-variables are fixed to 1. This allows the same Fourier-analytic argument to go through, showing that the restriction-and-averaging procedure simplifies any small-weight rectangle decomposition into a low-degree polynomial.

The only quantitative change from the Set Intersection analysis is that the effective parameter n is replaced by the number of core variables $|V| = \Omega(\log s / \log n)$. Carrying out the argument yields

$$\log \widetilde{\gamma}_2(f \circ \textsc{And}_2) = \Omega\left(\left(\frac{\log s}{\log n}\right)^{1/3}\right),$$

which is exactly the bound stated in Theorem 1.3.

Finally, we place our work in the context of the results of Chattopadhyay, Dahiya, and Lovett [8]. Among other results, that work studies the sparsity of Boolean functions and provides a structural characterization in terms of *max-degree restriction trees*. Their restriction trees are fully adaptive: the choice of which variable to restrict next depends on the outcomes of previous restrictions, and there is no fixed set of core variables, as in our setting. Moreover, their framework applies to a broader class of complexity measures, which they call one-sided and two-sided nice measures.

While this adaptive viewpoint is powerful, we do not know how to use such fully adaptive restriction trees to control analytic quantities such as the approximate γ_2 norm. In contrast, our work constructs a more structured, *semi-adaptive* form of max-degree restriction trees directly from sparsity, with a fixed core set of variables on which all $\{0, *\}$-restrictions occur. This additional structure is crucial for our analysis and enables us to lift sparsity to lower bounds on the approximate γ_2 norm.

Organization

In Section 2, we introduce the necessary preliminaries and notation. In Section 3, we formally define semi-adaptive max-degree restriction trees and show how to construct them for Boolean functions with large sparsity. Finally, in Section 4, we prove our main technical contribution: a lifting theorem that lifts the sparsity of a Boolean function f into a lower bound on the approximate γ_2 norm of its lifted function $f \circ \textsc{And}_2$, and we discuss the resulting consequences for And-functions.

2 Preliminaries

All functions considered are defined on the Boolean hypercube $\{0,1\}^n$ and unless stated otherwise, all polynomials are real and multilinear.

Multilinear representations

Over the Boolean domain, Boolean functions admit a canonical polynomial representation.

▶ **Definition 2.1** (Multilinear polynomial representation). *A polynomial $Q \in \mathbb{R}[x_1, \ldots, x_n]$ is* multilinear *if each variable appears with degree at most one in every monomial. Every function $f : \{0,1\}^n \to \mathbb{R}$ admits a unique multilinear polynomial representation; that is, there exists a unique multilinear polynomial $Q \in \mathbb{R}[x_1, \ldots, x_n]$ such that $Q(x) = f(x)$ for all $x \in \{0,1\}^n$.*

We freely identify Boolean functions with their unique multilinear polynomial representations, and use the two interchangeably.

Polynomial complexity measures

Let $Q \in \mathbb{R}[x_1, \ldots, x_n]$ be a multilinear polynomial written as

$$Q(x) = \sum_{S \subseteq [n]} a_S \prod_{i \in S} x_i.$$

The *degree* of Q is $\deg(Q) := \max\{|S| : a_S \neq 0\}$, and the *sparsity* of Q, denoted $\mathrm{spar}(Q)$, is the number of nonzero coefficients a_S.

For a function $f : \{0,1\}^n \to \mathbb{R}$, let $\mathcal{P}(f)$ denote its unique multilinear polynomial representation. We define

$$\deg(f) := \deg(\mathcal{P}(f)), \qquad \mathrm{spar}(f) := \mathrm{spar}(\mathcal{P}(f)).$$

Approximate degree

Let $f : \{0,1\}^n \to \mathbb{R}$ and $\varepsilon > 0$. The *ε-approximate degree* of f is defined as

$$\widetilde{\deg}_\varepsilon(f) := \min\big\{\deg(Q) : |Q(x) - f(x)| \leq \varepsilon \text{ for all } x \in \{0,1\}^n\big\}.$$

When $\varepsilon = 1/3$, we write $\widetilde{\deg}(f) := \widetilde{\deg}_{1/3}(f)$.

▶ **Theorem 2.2** ([7, Theorem 10, Section 3.4]). *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function. For any $0 < \varepsilon < 1/2$,*

$$\widetilde{deg}_\varepsilon(f) = O\big(\widetilde{deg}(f) \cdot \log(1/\varepsilon)\big).$$

▶ **Theorem 2.3** ([1, Theorem 4]). *For every Boolean function $f : \{0,1\}^n \to \{0,1\}$,*

$$deg(f) = O\big(\widetilde{deg}(f)^2\big).$$

Fourier basis

Another fundamental representation of Boolean functions is given by the *Fourier basis*, where each monomial $\chi_S(x) = (-1)^{\sum_{i \in S} x_i}$ represents the ± 1-valued parity function on the subset $S \subseteq [n]$ of variables.

▶ **Definition 2.4** (Fourier complexity measures). *Let $f : \{0,1\}^n \to \mathbb{R}$ have the Fourier expansion*

$$f(x) = \sum_{S \subseteq [n]} \widehat{f}(S)\, \chi_S(x).$$

The Fourier degree *of f is the maximum size of a set $S \subseteq [n]$ with $\widehat{f}(S) \neq 0$. Since $\chi_S(x) = \prod_{i \in S}(1 - 2x_i)$, the Fourier degree coincides with the ordinary degree.*

The Fourier sparsity *of f, denoted $\left\|\widehat{f}\right\|_0$, is the number of nonzero Fourier coefficients $\widehat{f}(S)$. The* Fourier ℓ_1-norm *of f is*

$$\left\|\widehat{f}\right\|_1 := \sum_{S \subseteq [n]} |\widehat{f}(S)|.$$

For $t \geq 0$, the Fourier ℓ_1-mass above level t *is*

$$\left\|\widehat{f}\right\|_1^{\geq t} := \sum_{\substack{S \subseteq [n] \\ |S| \geq t}} |\widehat{f}(S)|.$$

Communication complexity

We assume familiarity with the standard model of communication complexity and refer the reader to [22] for background. In this model, two parties – Alice and Bob – aim to compute a Boolean function $F : X \times Y \to \{0,1\}$, where Alice receives $x \in X$ and Bob receives $y \in Y$. The *deterministic communication complexity* of F, denoted $\mathrm{D}^{cc}(F)$, is the minimum number of bits exchanged by any deterministic protocol that always outputs $F(x,y)$. In the public-coin randomized model, Alice and Bob have access to shared randomness and must compute $F(x,y)$ with error at most $1/3$; the corresponding measure is the *randomized communication complexity* $\mathrm{R}^{cc}(F)$.

We also assume familiarity with quantum communication complexity [11]. We use $\mathrm{Q}^{cc}(F)$ to denote the bounded-error (error at most $1/3$) quantum communication complexity of F in the model with unlimited shared entanglement.

The γ_2 and approximate γ_2 norms

▶ **Definition 2.5** (γ_2 and approximate γ_2 norms). *Let $F : X \times Y \to \{0,1\}$ be a Boolean function, and let $M_F \in \{0,1\}^{X \times Y}$ denote its communication matrix, defined by $M_F(x,y) = F(x,y)$.*

The γ_2 norm *of F is*

$$\gamma_2(F) = \min\Big\{\sum_i |\alpha_i| : M_F = \sum_i \alpha_i R_i\Big\},$$

where each R_i is a combinatorial rectangle*, i.e., $R_i(x,y) = g_i(x)h_i(y)$ for Boolean functions $g_i : X \to \{0,1\}$ and $h_i : Y \to \{0,1\}$.*

For $0 < \varepsilon < 1/2$, the ε-approximate γ_2 norm *is*

$$\gamma_2^{\varepsilon}(F) = \min\Big\{\gamma_2(A) : A \in \mathbb{R}^{X \times Y},\ \|A - M_F\|_\infty \le \varepsilon\Big\}.$$

We write $\widetilde{\gamma}_2(F) := \gamma_2^{1/3}(F)$.

▶ Remark 2.6. The γ_2 norm is often defined via a factorization-based formulation. For a real matrix M, one may equivalently define

$$\gamma_2(M) = \min_{X,Y : XY^\top = M} r(X)\, r(Y),$$

where $r(X)$ denotes the maximum ℓ_2-norm of a row of X. This formulation, sometimes called the *μ-norm*, is equivalent to the rectangle-based definition up to constant factors; see, e.g., Chapter 2 of [24]. We use the rectangle-based definition throughout.

▶ **Theorem 2.7** ([25, Theorem 1]). *For every Boolean function $F : X \times Y \to \{0,1\}$,*

$$R^{cc}(F) = \Omega(\log \widetilde{\gamma}_2(F)) \qquad \text{and} \qquad Q^{cc}(F) = \Omega(\log \widetilde{\gamma}_2(F)).$$

3 Semi-Adaptive Max-Degree Restriction Trees from Sparsity

▶ **Definition 3.1** (Restrictions). *A* restriction *ρ on a set of variables $V \subseteq \{x_1, \dots, x_n\}$ is a partial assignment $\rho : V \to \{0,1,*\}$, where for $x_i \in V$, $\rho(x_i) \in \{0,1\}$ indicates that x_i is fixed, and $\rho(x_i) = *$ means x_i is left free. For a polynomial $Q \in \mathbb{R}[x_1, \dots, x_n]$, we write $Q|_\rho$ for the polynomial obtained by substituting $x_i = \rho(x_i)$ for all fixed variables x_i.*

A central tool in our work is to extract structural consequences of a function having large polynomial sparsity. In particular, we seek restrictions under which a function remains maximally hard, in the sense of retaining full degree.

At a high level, large sparsity guarantees the existence of many restrictions under which the function retains full degree. More concretely, if a multilinear polynomial has sparsity s, then there exists a set of variables V of size $\Omega(\log s/\log n)$ such that, for every assignment of the variables in V to values in $\{0,*\}$, one can fix the remaining variables so that the restricted polynomial has full degree in the variables left free.

We capture this collection of restrictions using what we call *semi-adaptive max-degree restriction trees*. The term *semi-adaptive* reflects the following structure. There is a fixed set of variables V such that every assignment in $\{0,*\}^V$ appears as a restriction, rather than variables being chosen adaptively based on previous assignments, as in fully adaptive restriction trees (e.g., in the work of [8]). However, the restrictions are not fully non-adaptive: although the set V is fixed, for each assignment $\rho \in \{0,*\}^V$, the fixing of the remaining variables in $[n] \setminus V$ that ensures full degree may depend on ρ. This intermediate structure motivates the term *semi-adaptive*.

The qualifier *max-degree* indicates that under every restriction in the tree, the polynomial retains full degree on the variables that remain free. This notion compactly encodes the key structural consequence of large sparsity that we exploit later.

We now formalize this notion.

▶ **Definition 3.2** (Max-degree restriction). *Let $Q \in \mathbb{R}[x_1,\dots,x_n]$ be a nonzero multilinear polynomial and let $\rho : \{x_1,\dots,x_n\} \to \{0,1,*\}$ be a restriction. We say that ρ is a* max-degree restriction *of Q if $deg(Q|_\rho) = |\rho^{-1}(*)|$, that is, the restricted polynomial has full degree in its free variables.*

▶ **Definition 3.3** (Semi-adaptive restriction tree). *A* semi-adaptive restriction tree of depth *d on n variables is a collection $\mathcal{D}$ of 2^d restrictions $\rho : \{x_1,\dots,x_n\} \to \{0,1,*\}$ for which there exists a fixed set of variables $V \subseteq \{x_1,\dots,x_n\}$ with $|V| = d$, called the* core variables*, such that:*

- *For every $\rho \in \mathcal{D}$ and every $x_i \notin V$, $\rho(x_i) \in \{0,1\}$.*
- *For every assignment $\alpha \in \{0,*\}^V$, there exists a unique $\rho \in \mathcal{D}$ such that $\rho(x_i) = \alpha(x_i)$ for all $x_i \in V$.*

Equivalently, $\mathcal{D}$ consists of all restrictions obtained by assigning each variable in V either 0 or $*$, while fixing all variables outside V as a function of this assignment.

▶ **Definition 3.4** (Semi-adaptive max-degree restriction tree). *Let $Q \in \mathbb{R}[x_1,\dots,x_n]$ be a multilinear polynomial and let $\mathcal{D}$ be a semi-adaptive restriction tree. We say that $\mathcal{D}$ is a* max-degree restriction tree *for Q if every $\rho \in \mathcal{D}$ is a max-degree restriction of Q, i.e., $deg(Q|_\rho) = |\rho^{-1}(*)|$.*

Examples

To illustrate the definitions, we give two informative examples.

▶ **Example 3.5.** OR_n. The function $\mathrm{OR}_n(x_1,\dots,x_n)$ has sparsity $2^n - 1$. It admits a semi-adaptive max-degree restriction tree of depth n with core variables $V = \{x_1,\dots,x_n\}$. Let $\mathcal{D} = \{\, \rho : \{x_1,\dots,x_n\} \to \{0,*\} \,\}$. For any $\rho \in \mathcal{D}$, the restricted function $\mathrm{OR}_n|_\rho$ is an OR over the free variables $\rho^{-1}(*)$, and hence $\deg(\mathrm{OR}_n|_\rho) = |\rho^{-1}(*)|$. Thus $\mathcal{D}$ is a semi-adaptive max-degree restriction tree for OR_n.

▶ **Example 3.6.** $\text{AND}_n \circ \text{OR}_2$. Consider

$$(\text{AND}_n \circ \text{OR}_2)(x_1, \ldots, x_n, y_1, \ldots, y_n) = \text{AND}_n(\text{OR}_2(x_1, y_1), \ldots, \text{OR}_2(x_n, y_n)),$$

which has sparsity 3^n. We construct a semi-adaptive max-degree restriction tree of depth n with core variables $V = \{x_1, \ldots, x_n\}$. For a restriction $\rho_x \in \{0, *\}^{\{x_1, \ldots, x_n\}}$, define a restriction $\rho_y^{\rho_x}$ on $\{y_1, \ldots, y_n\}$ by

$$\rho_y^{\rho_x}(y_i) = \begin{cases} 0, & \text{if } \rho_x(x_i) = *, \\ 1, & \text{if } \rho_x(x_i) = 0. \end{cases}$$

Set $\mathcal{D} = \left\{ \rho_x \cup \rho_y^{\rho_x} \;\middle|\; \rho_x \in \{0, *\}^{\{x_1, \ldots, x_n\}} \right\}$. For any $\rho \in \mathcal{D}$, each gate $\text{OR}_2(x_i, y_i)$ evaluates to a free variable when $\rho(x_i) = *$, and to the constant 1 when $\rho(x_i) = 0$. As a result, $(\text{AND}_n \circ \text{OR}_2)|_\rho$ computes an AND over exactly the free core variables $\rho^{-1}(*)$, and therefore $\deg\big((\text{AND}_n \circ \text{OR}_2)|_\rho\big) = |\rho^{-1}(*)|$. Thus, $\mathcal{D}$ is a semi-adaptive max-degree restriction tree of depth n for $\text{AND}_n \circ \text{OR}_2$.

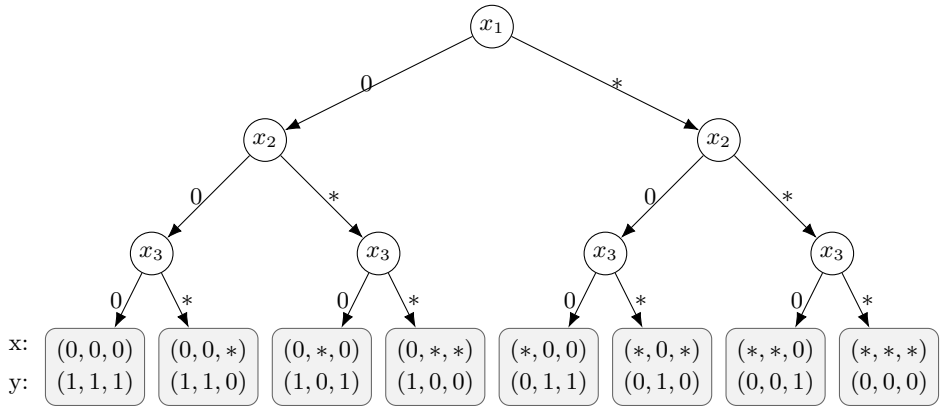

Figure 1 A semi-adaptive max-degree restriction tree for $\text{AND}_3 \circ \text{OR}_2$ with core variables $V = \{x_1, x_2, x_3\}$. Leaves correspond to the $2^{|V|}$ restrictions in the tree. For each leaf restriction, the restricted function $(\text{AND}_3 \circ \text{OR}_2)|_\rho$ computes an AND over exactly the variables left free.

The following lemma shows that large sparsity guarantees the existence of deep semi-adaptive max-degree restriction trees.

▶ **Lemma 3.7.** *Let $Q : \{0,1\}^n \to \mathbb{R}$ be a nonzero multilinear polynomial of sparsity s. Then Q admits a semi-adaptive max-degree restriction tree of depth $\Omega(\log s / \log n)$.*

Proof. Let $\mathcal{M}_Q \subseteq 2^{\{x_1, \ldots, x_n\}}$ denote the family of supports of monomials appearing in Q. Let $V \subseteq \{x_1, \ldots, x_n\}$ be a largest set shattered by $\mathcal{M}_Q$. Recall that V being shattered by $\mathcal{M}_Q$ means that for every $S \subseteq V$, there exists $T \in \mathcal{M}_Q$ such that $T \cap V = S$. The Sauer–Shelah–Perles lemma [33, 34] states that that for a set system $\mathcal{F}$ containing subsets of $[n]$, if d is the maximum size of a set shattered by $\mathcal{F}$, then $|\mathcal{F}| \leq O(n^d)$. Using this lemma, we obtain $|V| = \Omega(\log s / \log n)$; write $|V| = d$.

For each subset $S \subseteq V$, we construct a restriction ρ_S. The resulting family $\{\rho_S : S \subseteq V\}$ will form a semi-adaptive max-degree restriction tree for Q.

Fix an arbitrary $S \subseteq V$. Define the *free restriction* $\rho_S^{\text{free}} : V \to \{0, *\}$ by setting $\rho_S^{\text{free}}(x) = *$ for all $x \in S$ and $\rho_S^{\text{free}}(x) = 0$ for all $x \in V \setminus S$. Under this restriction, every monomial of Q containing a variable from $V \setminus S$ vanishes. As a result, the restricted polynomial can be written as $Q|_{\rho_S^{\text{free}}} = \sum_{T \subseteq S} \left(\prod_{x \in T} x \right) \cdot R_T$, where each R_T is a multilinear polynomial over the variables $\{x_i : x_i \notin V\}$.

Since V is shattered by $\mathcal{M}_Q$, there exists a monomial of Q whose support intersects V exactly in S. Equivalently, the coefficient polynomial R_S is nonzero. We now fix the remaining variables to witness this nonzeroness. Choose a *fixing restriction* $\rho_S^{\text{fix}} : \{x_1, \ldots, x_n\} \setminus V \to \{0,1\}$ such that $R_S|_{\rho_S^{\text{fix}}} \neq 0$.

Let $\rho_S := \rho_S^{\text{free}} \cup \rho_S^{\text{fix}}$. Then $Q|_{\rho_S}$ has degree exactly $|S|$, and its set of free variables is $\rho_S^{-1}(*) = S$. Hence, ρ_S is a max-degree restriction of Q.

Therefore, the family $\{\rho_S : S \subseteq V\}$ forms a semi-adaptive max-degree restriction tree of depth $|V| = \Omega(\log s / \log n)$. ◀

4 Lifting with the AND_2 Gadget

In this section, we present our main technical contribution: a lifting theorem that lifts the sparsity of a Boolean function $f : \{0,1\}^n \to \{0,1\}$ into a lower bound on the approximate γ_2 norm of the lifted function $F := f \circ \text{AND}_2$. We begin with a high-level overview of the proof.

Proof overview

Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function of sparsity s. Our goal is to lower bound the approximate γ_2 norm of the lifted function $f \circ \text{AND}_2$ in terms of $\log s$. Suppose, toward a contradiction, that $f \circ \text{AND}_2$ admits a small approximate γ_2 norm. Then there exists an approximator

$$\Pi(x, y) = \sum_{i=1}^{m} b_i \, g_i(x) h_i(y)$$

with $\sum_i |b_i|$ small. Our goal is to show that such an approximator cannot exist. The proof proceeds via a carefully designed random restriction argument that preserves the hardness of $f \circ \text{AND}_2$, in terms of degree, while simplifying any such approximator.

Step 1: Structure from sparsity. A key consequence of large sparsity is that f admits a *semi-adaptive max-degree restriction tree* $\mathcal{D}$ of depth $d = \Omega(\log s / \log n)$. Equivalently, there exists a fixed set of core variables $V \subseteq \{z_1, \ldots, z_n\}$, with $|V| = d$, such that for every assignment in $\{0, *\}^V$, the remaining variables can be fixed so that the restricted function has full degree in the surviving free variables. This provides a large family of restrictions under which f remains maximally hard in terms of degree.

Step 2: Lifting restrictions through the AND_2 gadget. We lift the restriction tree $\mathcal{D}$ for f to a collection of restrictions $\mathcal{D} \circ \text{AND}_2$ for the lifted function

$$(f \circ \text{AND}_2)(x_1, \ldots, x_n, y_1, \ldots, y_n) = f(\text{AND}_2(x_1, y_1), \ldots, \text{AND}_2(x_n, y_n)).$$

Each restriction $\rho_f \in \mathcal{D}$ is mapped to a restriction ρ on the variables (x_i, y_i) so that $\text{AND}_2(x_i, y_i)|_\rho = \rho_f(z_i)$ for every i. Concretely, the lifted restriction is given by

$$(\rho(x_i), \rho(y_i)) = \begin{cases} (\Delta, 0), & \text{if } \rho_f(z_i) = 0 \text{ and } z_i \in V, \\ (*, 1), & \text{if } \rho_f(z_i) = * \text{ and } z_i \in V, \\ (1, 1), & \text{if } \rho_f(z_i) = 1 \text{ and } z_i \notin V, \\ (1, 0), & \text{if } \rho_f(z_i) = 0 \text{ and } z_i \notin V. \end{cases}$$

Here both $*$ and Δ denote free variables. Algorithm 1 formalizes this construction.

The lifted restrictions satisfy two immediate properties. First, all y-variables are fixed under every lifted restriction. Second, the lifting introduces a special type of free variable, called a *masked variable*: these are free x_i-variables assigned Δ, whose corresponding y_i is fixed to 0; consequently, the restricted function is independent of them.

As a result, for every lifted restriction ρ, averaging over the masked variables yields a function that coincides (up to renaming variables) with $f|_{\rho_f}$. For a restriction ρ, we denote the set of masked variables by

$$M(\rho) := \{\, x_i \mid \rho(x_i) = \Delta \,\},$$

and write $\mathbb{E}_{x_{M(\rho)}}[\cdot]$ for expectation over independent, uniform assignments to the variables in $M(\rho)$. We will use this notation throughout.

Since $\mathcal{D}$ is a max-degree restriction tree, it follows that

$$\deg\big(\mathbb{E}_{x_{M(\rho)}}[(f \circ \textsc{And}_2)|_\rho]\big) = |\rho^{-1}(*)|.$$

Step 3: Random restrictions and Fourier decay. We place a uniform distribution over lifted restrictions $\rho \in \mathcal{D} \circ \textsc{And}_2$ having exactly pd many $*$-variables. We analyze the effect of sampling such a restriction and then taking expectation over the masked variables on the following:

1. *The target function $f \circ \textsc{And}_2$.* This operation preserves hardness: by construction, the resulting function has exact degree pd.
2. *An arbitrary Boolean function g on the x-variables.* We show that the same operation causes the Fourier ℓ_1-mass above level k to decay exponentially, provided $k \gtrsim p^2 d$.

To derive a contradiction later, we need the Fourier mass above level $\sqrt{pd}$ to be exponentially small, since this would imply an approximating polynomial of degree $O(\sqrt{pd})$ for the restricted target function. Since exponential decay is guaranteed only above level $p^2 d$, we choose parameters so that $p^2 d \approx \sqrt{pd}$. Solving this relation yields $p = \Theta(d^{-1/3})$, for which $pd = \Theta(d^{2/3})$ and $\sqrt{pd} = \Theta(d^{1/3})$. Thus the Fourier tail decays exponentially above the precise level needed for the contradiction.

Step 4: Deriving a contradiction. Applying this random restriction-and-averaging procedure to the approximator Π, each function $h_i(y)$ collapses to a constant, while the Fourier tails of the corresponding $g_i(x)$ terms decay rapidly. Since $\sum_i |b_i|$ is small, we conclude that the expected Fourier mass of $\mathbb{E}_{x_{M(\rho)}}[\Pi|_\rho]$ above level $k = \Theta(\sqrt{pd}) = \Theta(d^{1/3})$ is negligible.

Discarding this high-degree mass yields a polynomial of degree $O(\sqrt{pd})$ that still approximates $\mathbb{E}_{x_{M(\rho)}}[(f \circ \textsc{And}_2)|_\rho]$. However, by the construction above, this target function has exact degree pd, contradicting the general fact that the exact degree of a Boolean function is at most quadratic in its approximate degree. Choosing parameters appropriately, this contradiction implies

$$\log \widetilde{\gamma}_2(f \circ \textsc{And}_2) = \Omega\left(\left(\frac{\log s}{\log n}\right)^{1/3}\right),$$

completing the proof.

4.1 Lifting De Morgan Sparsity to the Approximate γ_2-Norm via the $\textsc{And}_2$ Gadget

We now formalize the proof strategy outlined in the proof overview. The first step is to lift max-degree restrictions for f to a collection of restrictions for the lifted function $f \circ \textsc{And}_2$.Algorithm 1 describes this lifting procedure.

Algorithm 1 LIFTEDRESTRICTION.

Input: A semi-adaptive max-degree restriction tree $\mathcal{D}$ for $f : \{0,1\}^n \to \{0,1\}$
Output: A collection of lifted restrictions $\mathcal{D} \circ \text{AND}_2$, where each $\rho : \{x_i, y_i\}_{i=1}^n \to \{0, 1, *, \Delta\}$ is a restriction for $f \circ \text{AND}_2$
Let V be the set of *core variables* of the semi-adaptive restriction tree $\mathcal{D}$.
for each restriction $\rho_f \in \mathcal{D}$ **do**
    Define the lifted restriction $\rho = \text{Lift}_{\mathcal{D}}(\rho_f)$ as follows:
    for each $i \in [n]$ **do**
        Set $(\rho(x_i), \rho(y_i)) = \begin{cases} (\Delta, 0), & \text{if } \rho_f(z_i) = 0 \text{ and } z_i \in V, \\ (*, 1), & \text{if } \rho_f(z_i) = * \text{ and } z_i \in V, \\ (1, 1), & \text{if } \rho_f(z_i) = 1 \text{ and } z_i \notin V, \\ (1, 0), & \text{if } \rho_f(z_i) = 0 \text{ and } z_i \notin V. \end{cases}$
    end for
end for
return $\mathcal{D} \circ \text{AND}_2 := \{\text{Lift}(\rho_f) : \rho_f \in \mathcal{D}\}$

A lifted restriction assigns each variable a value in $\{0, 1, *, \Delta\}$. When applying such a restriction to $f \circ \text{AND}_2$, both symbols $*$ and Δ are treated as free variables. Variables marked by Δ are called *masked variables*; although syntactically free, the restricted function does not depend on them. Tracking masked variables explicitly will be convenient for the subsequent analysis.

Basic structure of lifted restrictions

Let $V \subseteq \{z_1, \ldots, z_n\}$ denote the set of core variables of the semi-adaptive restriction tree $\mathcal{D}$, and let $V_x := \{x_i : z_i \in V\}$ be the corresponding set of x-variables, which we refer to as the *core x-variables*. We will use this notation throughout.

By construction, every lifted restriction $\rho \in \mathcal{D} \circ \text{AND}_2$ fixes all y-variables and fixes all x-variables outside V_x to 1. Moreover, the assignment on V_x uniquely determines the entire restriction: for every $\alpha \in \{\Delta, *\}^{V_x}$, there exists a unique $\rho \in \mathcal{D} \circ \text{AND}_2$ such that $\rho(x_i) = \alpha(x_i)$ for all $x_i \in V_x$. This follows directly from the defining property of $\mathcal{D}$, which guarantees that for every $\beta \in \{0, *\}^V$ there is a unique $\rho_f \in \mathcal{D}$ satisfying $\rho_f(z_i) = \beta(z_i)$ for all $z_i \in V$. As a result, the family $\mathcal{D} \circ \text{AND}_2$ is parametrized by assignments to the core x-variables.

The construction of the lifted restrictions, together with the fact that $\mathcal{D}$ is a max-degree restriction tree, implies the following structural properties. In particular, these properties show that $f \circ \text{AND}_2$ retains its hardness (in terms of degree) under the lifted restrictions.

▷ Claim 4.1. Let $\mathcal{D} \circ \text{AND}_2$ be the collection of lifted restrictions for $f \circ \text{AND}_2$ obtained from a semi-adaptive max-degree restriction tree $\mathcal{D}$ for f via Algorithm 1. Then the following properties hold:

1. For every $\rho \in \mathcal{D} \circ \text{AND}_2$, the restricted function $(f \circ \text{AND}_2)|_\rho$ does not depend on the masked variables $M(\rho)$.
2. For every $\rho \in \mathcal{D} \circ \text{AND}_2$, $\deg\big(\mathbb{E}_{x_{M(\rho)}}[(f \circ \text{AND}_2)|_\rho]\big) = \deg\big((f \circ \text{AND}_2)|_\rho\big) = |\rho^{-1}(*)|$.

Proof. For (1), if x_i is masked in ρ, then by construction $\rho(y_i) = 0$. Hence $\text{AND}_2(x_i, y_i)|_\rho = 0$ regardless of the value of x_i, and therefore $(f \circ \text{AND}_2)|_\rho$ is independent of all masked variables.

For (2), fix $\rho \in \mathcal{D} \circ \text{AND}_2$, and let $\rho_f \in \mathcal{D}$ be the restriction used to generate ρ, i.e., $\rho = \text{Lift}(\rho_f)$. By construction, for every $i \in [n]$ we have $\text{AND}_2(x_i, y_i)|_\rho = \rho_f(z_i)$. Thus, $(f \circ \text{AND}_2)|_\rho$ depends on the variables $\{x_i, y_i\}$ only through the tuple $(\rho_f(z_1), \ldots, \rho_f(z_n))$.

After ignoring the masked variables (which $(f \circ \text{AND}_2)|_\rho$ does not depend on by part (1)), the resulting function coincides with $f|_{\rho_f}$. Since ρ_f is a max-degree restriction of f, we obtain $\deg((f \circ \text{AND}_2)|_\rho) = \deg(f|_{\rho_f}) = |\rho_f^{-1}(*)| = |\rho^{-1}(*)|$.

Finally, since $(f \circ \text{AND}_2)|_\rho$ does not depend on the variables in $M(\rho)$, the function $\mathbb{E}_{x_{M(\rho)}}[(f \circ \text{AND}_2)|_\rho]$ is obtained by simply viewing $(f \circ \text{AND}_2)|_\rho$ as a function on the remaining free variables. In particular, taking expectation over the masked variables leaves the function unchanged as a polynomial in the remaining free variables. Hence, $\deg\big(\mathbb{E}_{x_{M(\rho)}}[(f \circ \text{AND}_2)|_\rho]\big) = \deg\big((f \circ \text{AND}_2)|_\rho\big)$. ◁

Next, we introduce a probability distribution on the restrictions in $\mathcal{D} \circ \text{AND}_2$ for use in a random restriction argument. Fix a parameter $p \in (0,1)$, and recall that every restriction $\rho \in \mathcal{D} \circ \text{AND}_2$ satisfies $|\rho^{-1}(*)| + |\rho^{-1}(\Delta)| = d$, where d is the depth of the semi-adaptive restriction tree $\mathcal{D}$.

We consider the uniform distribution over all restrictions in $\mathcal{D} \circ \text{AND}_2$ that leave exactly pd variables marked by $*$. Formally, let $U := \{\, \rho \in \mathcal{D} \circ \text{AND}_2 \mid |\rho^{-1}(*)| = pd \,\}$. By construction, $|U| = \binom{d}{pd}$. We sample a restriction uniformly at random from U, i.e., $\Pr[\rho] = 1/\binom{d}{pd}$ for all $\rho \in U$. We denote this distribution by $\mathcal{U}_p(\mathcal{D} \circ \text{AND}_2)$.

Next, we analyze the effect of sampling a restriction $\rho \sim \mathcal{U}_p(\mathcal{D} \circ \text{AND}_2)$, applying it to $f \circ \text{AND}_2$, and then taking expectation over the masked variables $M(\rho)$. By Claim 4.1, the resulting function retains degree pd. On the other hand, we show that applying the same process – sampling $\rho \sim \mathcal{U}_p(\mathcal{D} \circ \text{AND}_2)$ and taking expectation over $M(\rho)$ – to an arbitrary Boolean function g over the x-variables causes its high-degree Fourier mass to decay exponentially above level k, provided $k \gtrsim p^2 d$. Optimizing this tradeoff leads to the choice $p = \Theta(d^{-1/3})$, for which $pd = \Theta(d^{2/3})$ and the Fourier tail decays exponentially above level $k = \Theta(\sqrt{pd}) = \Theta(d^{1/3})$. Combining these two observations yields our main result: a lower bound on the approximate γ_2 norm of $f \circ \text{AND}_2$ in terms of the sparsity of f.

▷ Claim 4.2. Let $\mathcal{D} \circ \text{AND}_2$ be the collection of lifted restrictions for $f \circ \text{AND}_2$ obtained from a semi-adaptive max-degree restriction tree $\mathcal{D}$ of depth d via Algorithm 1. Let $g : \{0,1\}^n \to \{0,1\}$ be an arbitrary Boolean function on the x-variables $\{x_1, \ldots, x_n\}$. Then for every integer $k \geq 0$,

$$\mathbb{E}_{\rho \sim \mathcal{U}_p(\mathcal{D} \circ \text{AND}_2)}\left[\left\|\widehat{\mathbb{E}_{x_{M(\rho)}}[g|_\rho]}\right\|_1^{\geq k}\right] \leq \sum_{t=k}^{pd} \left(p\sqrt{\frac{ed}{t}}\right)^t.$$

In particular, if $k \geq 4ep^2 d$, then

$$\mathbb{E}_{\rho \sim \mathcal{U}_p(\mathcal{D} \circ \text{AND}_2)}\left[\left\|\widehat{\mathbb{E}_{x_{M(\rho)}}[g|_\rho]}\right\|_1^{\geq k}\right] \leq 2^{-k+1}.$$

Proof. By construction, every restriction $\rho \in \mathcal{D} \circ \text{AND}_2$ fixes all x-variables outside the set V_x to 1. Let $h : \{0,1\}^{|V_x|} \to \{0,1\}$ be the Boolean function obtained from g by fixing all variables outside V_x to 1. Then for every $\rho \in \mathcal{D} \circ \text{AND}_2$, we have $g|_\rho = h$. Without loss of generality, assume $V_x = \{x_1, \ldots, x_d\}$.

Taking expectation over any variable kills all Fourier monomials containing it. Therefore,

$$\begin{aligned}\mathbb{E}_{\rho \sim \mathcal{U}_p(\mathcal{D} \circ \text{AND}_2)}\left[\left\|\widehat{\mathbb{E}_{x_{M(\rho)}}[g|_\rho]}\right\|_1^{\geq k}\right] &= \mathbb{E}_{\rho \sim \mathcal{U}_p(\mathcal{D} \circ \text{AND}_2)}\left[\left\|\widehat{\mathbb{E}_{x_{M(\rho)}}[h]}\right\|_1^{\geq k}\right] \\ &= \sum_{\substack{S \subseteq [d] \\ |S| \geq k}} |\widehat{h}(S)| \cdot \Pr_\rho\big[S \subseteq \rho^{-1}(*)\big].\end{aligned}$$

Under the uniform distribution over restrictions with exactly pd variables marked by $*$, the probability $\Pr_\rho[S \subseteq \rho^{-1}(*)]$ is zero when $|S| > pd$, and for $|S| \le pd$ we have

$$\Pr_\rho\big[S \subseteq \rho^{-1}(*)\big] = \frac{\binom{d-|S|}{pd-|S|}}{\binom{d}{pd}} \;\le\; p^{|S|}.$$

Therefore,

$$\sum_{\substack{S\subseteq[d]\\ k\le|S|\le pd}} |\widehat{h}(S)| \cdot \Pr_\rho\big[S \subseteq \rho^{-1}(*)\big] \;\le\; \sum_{t=k}^{pd}\sum_{|S|=t} |\widehat{h}(S)|\, p^t.$$

Applying the Cauchy–Schwarz inequality, together with Parseval's identity (which implies $\sum_S \widehat{h}(S)^2 \le 1$ for the Boolean function h), we obtain $\sum_{|S|=t} |\widehat{h}(S)| \le \sqrt{\binom{d}{t}}$. Hence,

$$\sum_{t=k}^{pd}\sum_{|S|=t} |\widehat{h}(S)|\, p^t \;\le\; \sum_{t=k}^{pd}\sqrt{\binom{d}{t}}\, p^t \;\le\; \sum_{t=k}^{pd}\left(p\sqrt{\frac{ed}{t}}\right)^t.$$

Finally, suppose $k \ge 4ep^2 d$. Then for every $t \ge k$,

$$p\sqrt{\frac{ed}{t}} \le p\sqrt{\frac{ed}{k}} \le \frac{1}{2}.$$

Therefore,

$$\sum_{t=k}^{pd}\left(p\sqrt{\frac{ed}{t}}\right)^t \le \sum_{t=k}^{\infty} 2^{-t} \le 2^{-k+1}.$$

This completes the proof. ◁

▶ **Theorem 1.3** (Restated). *For every total Boolean function $f : \{0,1\}^n \to \{0,1\}$,*

$$\log \widetilde{\gamma_2}(f \circ \mathrm{AND}_2) = \Omega\left(\left(\frac{\log spar(f)}{\log n}\right)^{1/3}\right).$$

Proof. Let f have sparsity s. By Lemma 3.7, there exists a semi-adaptive max-degree restriction tree $\mathcal{D}$ for f of depth $d = c_1 \log s/\log n$, for a suitable absolute constant $c_1 > 0$.

Let $c_2, c_3 > 0$ be absolute constants such that Theorem 2.2 converts a 0.44-approximator of degree r into a 1/3-approximator of degree at most $c_2 r$, and such that every Boolean function g satisfies $\deg(g) \le c_3 \cdot (\widetilde{\deg}(g))^2$. Choose an absolute constant $c > 0$ sufficiently small so that $c^{3/2} \le \frac{1}{4ec_2\sqrt{2c_3}}$. Set

$$p := cd^{-1/3} \qquad \text{and} \qquad k := \frac{c^{1/2}d^{1/3}}{c_2\sqrt{2c_3}}.$$

Observe that $pd = cd^{2/3}$, while $k = \Theta(d^{1/3}) = \Theta(\sqrt{pd})$.

Now suppose, for the sake of contradiction, that there exists a decomposition of the communication matrix $M_{f\circ \mathrm{AND}_2}$ of the form

$$\Pi(x,y) = \sum_{i=1}^{m} b_i\, g_i(x) h_i(y),$$

where each g_i, h_i is Boolean, such that $\|\Pi - M_{f\circ\mathrm{AND}_2}\|_\infty \le 1/3$ and $\sum_{i=1}^m |b_i| \le \frac{1}{20}\, 2^k$. We derive a contradiction, thereby proving the theorem.

Let $\mathcal{D} \circ \text{AND}_2$ be the collection of lifted restrictions for $f \circ \text{AND}_2$ obtained from $\mathcal{D}$ via Algorithm 1. Sample a restriction $\rho \sim \mathcal{U}_p(\mathcal{D} \circ \text{AND}_2)$. We study the effect of applying ρ and then taking expectation over the masked variables $M(\rho)$. For notational convenience, define

$$F_\rho := \mathbb{E}_{x_{M(\rho)}}\left[(f \circ \text{AND}_2)|_\rho\right], \qquad G_{i,\rho} := \mathbb{E}_{x_{M(\rho)}}\left[g_i|_\rho\right].$$

Both are functions of the starred x-variables under ρ.

1. **Hardness of the restricted function.** By Claim 4.1, $\deg(F_\rho) = |\rho^{-1}(*)|$. Since every restriction in the support of $\mathcal{U}_p(\mathcal{D} \circ \text{AND}_2)$ satisfies $|\rho^{-1}(*)| = pd$, we have $\deg(F_\rho) = cd^{2/3}$.
2. **Simplification of the approximator Π.** For every restriction $\rho \in \mathcal{D} \circ \text{AND}_2$, all y-variables are fixed. As a result, for each term in the decomposition $\Pi(x, y) = \sum_{i=1}^m b_i\, g_i(x) h_i(y)$, the restricted function $h_i|_\rho$ becomes a constant, which we denote by $a_{\rho,i} \in \{0, 1\}$. Thus, $\Pi|_\rho = \sum_{i=1}^m b_i\, a_{\rho,i}\, g_i|_\rho$, which is a function only of the core x-variables, since all non-core x-variables are fixed to 1 in every $\rho \in \mathcal{D} \circ \text{AND}_2$.
   Fix a restriction ρ. The ℓ_1-mass of the Fourier spectrum of $\mathbb{E}_{x_{M(\rho)}}\left[\Pi|_\rho\right]$ above level k can be bounded as follows:

$$\begin{aligned}
\left\|\widehat{\mathbb{E}_{x_{M(\rho)}}[\Pi|_\rho]}\right\|_1^{\geq k} &= \sum_{\substack{S \subseteq [n] \\ |S| \geq k}} \left|\sum_{i=1}^m b_i\, a_{\rho,i}\, \widehat{G_{i,\rho}}(S)\right| \\
&\leq \sum_{\substack{S \subseteq [n] \\ |S| \geq k}} \sum_{i=1}^m |b_i| |\widehat{G_{i,\rho}}(S)| \\
&= \sum_{i=1}^m |b_i| \sum_{\substack{S \subseteq [n] \\ |S| \geq k}} |\widehat{G_{i,\rho}}(S)| \\
&= \sum_{i=1}^m |b_i| \left\|\widehat{G_{i,\rho}}\right\|_1^{\geq k}.
\end{aligned}$$

   Moreover,

$$4ep^2 d = 4ec^2 d^{1/3} \leq \frac{c^{1/2} d^{1/3}}{c_2 \sqrt{2c_3}} = k,$$

   where the inequality follows from the choice of c. Taking expectation over ρ and applying Claim 4.2, we obtain

$$\begin{aligned}
\mathbb{E}_{\rho \sim \mathcal{U}_p(\mathcal{D} \circ \text{AND}_2)}\left[\left\|\widehat{\mathbb{E}_{x_{M(\rho)}}[\Pi|_\rho]}\right\|_1^{\geq k}\right] &\leq \sum_{i=1}^m |b_i|\, \mathbb{E}_{\rho \sim \mathcal{U}_p(\mathcal{D} \circ \text{AND}_2)}\left[\left\|\widehat{G_{i,\rho}}\right\|_1^{\geq k}\right] \\
&\leq \sum_{i=1}^m |b_i| \cdot 2^{-k+1} \;\leq\; \frac{1}{10},
\end{aligned}$$

   where the final inequality follows from the assumed bound on $\sum_i |b_i|$.
   Thus, there exists a restriction $\rho \sim \mathcal{U}_p(\mathcal{D} \circ \text{AND}_2)$ such that the ℓ_1-mass of the Fourier spectrum of $\Pi|_\rho$ above level k, after averaging over the masked variables $M(\rho)$, is at most 0.1.

Combining the two items, there exists a restriction ρ such that $\deg(F_\rho) = pd = cd^{2/3}$, while the ℓ_1-mass of the Fourier spectrum of $\mathbb{E}_{x_{M(\rho)}}[\Pi|_\rho]$ above level k is at most 0.1.

Let $\tilde{\Pi}$ be the polynomial obtained from $\mathbb{E}_{x_{M(\rho)}}[\Pi|_\rho]$ by deleting all Fourier monomials of degree at least k. Since the discarded Fourier mass is at most 0.1 and Π is a 1/3-approximator of $f \circ \text{AND}_2$, the polynomial $\tilde{\Pi}$ has degree $< k$ and 0.44-approximates F_ρ. By standard error reduction (Theorem 2.2), $\tilde{\Pi}$ can be converted into a 1/3-approximator of degree at most $c_2 k$. As a result,

$$\widetilde{\deg}(F_\rho) < c_2 k = \frac{c^{1/2} d^{1/3}}{\sqrt{2c_3}}.$$

On the other hand, $\deg(F_\rho) = cd^{2/3}$ which contradicts the general inequality $\deg(g) \leq c_3 \cdot (\widetilde{\deg}(g))^2$ for Boolean functions g (Theorem 2.3). This contradiction completes the proof. ◀

4.2 Consequences

Knop et al. [19] showed that $\log \text{spar}(f)$ characterizes the deterministic communication complexity of AND-functions ($f \circ \text{AND}_2$), up to polynomial loss and polylogarithmic factors in n. In particular, they proved that for every Boolean function f,

$$\text{D}^{cc}(f \circ \text{AND}_2) = O\big((\log \text{spar}(f))^5 \cdot \log n\big).$$

Combining this bound with Theorem 1.3, together with the fact that the logarithm of the approximate γ_2 norm lower bounds bounded-error quantum communication complexity, we immediately obtain that for every Boolean function f,

$$\text{D}^{cc}(f \circ \text{AND}_2) = O\big(\text{Q}^{cc}(f \circ \text{AND}_2)^{15} \cdot (\log n)^6\big).$$

A tighter relationship can be obtained using a more refined structural result of Knop et al., which relates deterministic AND-query complexity to sparsity and a combinatorial measure known as *monotone block sensitivity*.

▶ **Definition 4.3** (Monotone Block Sensitivity)**.** *The* monotone block sensitivity *of a Boolean function $f : \{0,1\}^n \to \{0,1\}$, denoted MBS($f$), is a variant of block sensitivity that only considers flipping 0's to 1's. A subset $B \subseteq [n]$ is called a* sensitive 0-block of f at input x *if $x_i = 0$ for all $i \in B$, and $f(x) \neq f(x \oplus 1_B)$, where $x \oplus 1_B$ denotes the input obtained by flipping all bits in B from 0 to 1. For an input $x \in \{0,1\}^n$, let MBS(f, x) denote the maximum number of pairwise disjoint sensitive 0-blocks of f at x. Then, MBS(f) $= \max_{x \in \{0,1\}^n}$ MBS(f, x).*

Knop et al. [19] showed that deterministic communication complexity can be bounded in terms of both sparsity and monotone block sensitivity.

▷ Claim 4.4 ([19, Lemma 3.2, Claim 4.4, Lemma 4.6, Theorem 1.2]). For every Boolean function $f : \{0,1\}^n \to \{0,1\}$,

$$\text{D}^{cc}(f \circ \text{AND}_2) = O\big(\text{MBS}(f)^2 \cdot \log \text{spar}(f) \cdot \log n\big).$$

Intuitively, a large value of MBS(f) indicates that a large-arity PROMISE-OR function can be embedded into f via suitable restrictions and identifications of variables. When such a function f is lifted via composition with AND_2, this structure gives rise to an embedded instance of the *unique set disjointness* problem.

The *unique set disjointness* function UDISJ_k is a partial Boolean function on inputs $x, y \in \{0,1\}^k$, defined as

$$\mathrm{UDISJ}_k(x,y) = \begin{cases} 0, & \text{if } |x \wedge y| = 0, \\ 1, & \text{if } |x \wedge y| = 1, \\ \text{undefined}, & \text{otherwise}, \end{cases}$$

where $x \wedge y$ denotes the bitwise AND and $|\cdot|$ the Hamming weight. That is, under the promise that the inputs are either bitwise disjoint or intersect in exactly one coordinate, the function distinguishes between these two cases.

The following result of Knop et al. shows that large monotone block sensitivity forces large embedded instances of unique set disjointness.

▷ Claim 4.5 ([19, Claim 4.7]). Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function with $\mathrm{MBS}(f) = k$. Then the communication matrix of $f \circ \textsc{And}_2$ contains, as a submatrix (up to flipping output bits), the communication matrix of UDISJ_k.

Using known lower bounds for unique set disjointness, we obtain the following.

▶ **Theorem 4.6** ([17, 30, 31, 36]). *$R^{cc}(\mathrm{UDISJ}_k) = \Omega(k)$ and $Q^{cc}(\mathrm{UDISJ}_k) = \Omega(\sqrt{k})$.*

▷ Claim 4.7. If $f : \{0,1\}^n \to \{0,1\}$ satisfies $\mathrm{MBS}(f) = k$, then

$$\mathrm{R}^{cc}(f \circ \textsc{And}_2) = \Omega(k) \qquad \text{and} \qquad \mathrm{Q}^{cc}(f \circ \textsc{And}_2) = \Omega(\sqrt{k}).$$

Proof. This follows immediately from Claim 4.5 and Theorem 4.6. ◁

Combining Claim 4.7 with Claim 4.4, we obtain the following relationships between deterministic, randomized, and quantum communication complexity for And-functions.

▶ **Theorem 1.2** (Restated). *Let $f : \{0,1\}^n \to \{0,1\}$ be any Boolean function. Then:*

1. $D^{cc}(f \circ \textsc{And}_2) = O\big(Q^{cc}(f \circ \textsc{And}_2)^7 \cdot (\log n)^2\big)$.
2. $D^{cc}(f \circ \textsc{And}_2) = O\big(R^{cc}(f \circ \textsc{And}_2)^5 \cdot (\log n)^2\big)$.

Proof. Combine Theorem 1.3, Theorem 2.7, Claim 4.4 and Claim 4.7. ◀

References

1 Scott Aaronson, Shalev Ben-David, Robin Kothari, Shravas Rao, and Avishay Tal. Degree vs. approximate degree and quantum implications of huang's sensitivity theorem. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1330–1342, 2021. doi:10.1145/3406325.3451047.

2 Andris Ambainis, Leonard J Schulman, Amnon Ta-Shma, Umesh Vazirani, and Avi Wigderson. The quantum communication complexity of sampling. *SIAM Journal on Computing*, 32(6):1570–1585, 2003. doi:10.1137/S009753979935476.

3 Nikhil Bansal and Makrand Sinha. K-forrelation optimally separates quantum and classical query complexity. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1303–1316, 2021. doi:10.1145/3406325.3451040.

4 Harry Buhrman, Richard Cleve, John Watrous, and Ronald de Wolf. Quantum fingerprinting. *Physical review letters*, 87(16):167902, 2001.

5 Harry Buhrman and Ronald de Wolf. Communication complexity lower bounds by polynomials. In *Proceedings of the 16th Annual IEEE Conference on Computational Complexity*, pages 120–130, 2001. doi:10.1109/CCC.2001.933879.

6 Harry Buhrman, Nikolai K. Vereshchagin, and Ronald de Wolf. On computation and communication with small bias. In *22nd Annual IEEE Conference on Computational Complexity (CCC 2007), 13-16 June 2007, San Diego, California, USA*, pages 24–32. IEEE Computer Society, 2007. doi:10.1109/CCC.2007.18.

7 Mark Bun and Justin Thaler. Approximate degree in classical and quantum computing. *Found. Trends Theor. Comput. Sci.*, 15(3-4):229–423, 2022. doi:10.1561/0400000107.

8 Arkadev Chattopadhyay, Yogesh Dahiya, and Shachar Lovett. Restriction trees for sparsity and applications. In *Proceedings of the 58th Annual ACM Symposium on Theory of Computing (STOC 2026)*, 2026. URL: https://eccc.weizmann.ac.il/report/2025/172/.

9 Arkadev Chattopadhyay, Nikhil S. Mande, and Suhail Sherif. The log-approximate-rank conjecture is false. *J.ACM*, 67(4):1–28, 2020. doi:10.1145/3396695.

10 Richard Cleve and Harry Buhrman. Substituting quantum entanglement for communication. *Physical Review A*, 56(2):1201, 1997.

11 Ronald de Wolf. Quantum communication and complexity. *Theoretical computer science*, 287(1):337–353, 2002. doi:10.1016/S0304-3975(02)00377-8.

12 Dmitry Gavinsky. Bare quantum simultaneity versus classical interactivity in communication complexity. In Konstantin Makarychev, Yury Makarychev, Madhur Tulsiani, Gautam Kamath, and Julia Chuzhoy, editors, *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, STOC 2020, Chicago, IL, USA, June 22-26, 2020*, pages 401–411. ACM, 2020. doi:10.1145/3357713.3384243.

13 Dmitry Gavinsky. Entangled simultaneity versus classical interactivity in communication complexity. *IEEE Trans. Inf. Theory*, 66(7):4641–4651, 2020. doi:10.1109/TIT.2020.2976074.

14 Dmitry Gavinsky, Julia Kempe, Oded Regev, and Ronald de Wolf. Bounded-error quantum state identification and exponential separations in communication complexity. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of Computing*, pages 594–603, 2006. doi:10.1145/1132516.1132602.

15 Uma Girish, Ran Raz, and Avishay Tal. Quantum versus randomized communication complexity, with efficient players. In James R. Lee, editor, *12th Innovations in Theoretical Computer Science Conference, ITCS 2021, Virtual Conference, January 6-8, 2021*, LIPIcs, pages 54:1–54:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021. doi:10.4230/LIPIcs.ITCS.2021.54.

16 Mika Göös, Tom Gur, Siddhartha Jain, and Jiawei Li. Quantum communication advantage in TFNP. In Michal Koucký and Nikhil Bansal, editors, *Proceedings of the 57th Annual ACM Symposium on Theory of Computing, STOC 2025, Prague, Czechia, June 23-27, 2025*, pages 1465–1475. ACM, 2025. doi:10.1145/3717823.3718155.

17 Bala Kalyanasundaram and Georg Schintger. The probabilistic communication complexity of set intersection. *SIAM Journal on Discrete Mathematics*, 5(4):545–557, 1992. doi:10.1137/0405044.

18 Hartmut Klauck. Lower bounds for quantum communication complexity. *SIAM J. Comput.*, 37(1):20–46, 2007. doi:10.1137/S0097539702405620.

19 Alexander Knop, Shachar Lovett, Sam McGuire, and Weiqiang Yuan. Log-rank and lifting for and-functions. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 197–208, 2021. doi:10.1145/3406325.3450999.

20 Matthias Krause and Pavel Pudlák. On computing boolean functions by sparse real polynomials. In *36th Annual Symposium on Foundations of Computer Science, FOCS 1995, Milwaukee, Wisconsin, USA, 23-25 October 1995*, pages 682–691. IEEE Computer Society, 1995. doi:10.1109/SFCS.1995.492670.

21 Ilan Kremer. Quantum communication. *Master's Thesis, Hebrew University, Computer Science Department*, 1995.

22 Eyal Kushilevitz and Noam Nisan. *Communication Complexity*. Cambridge University Press, 1997.

23 Troy Lee and Adi Shraibman. An approximation algorithm for approximation rank. In *2009 24th Annual IEEE Conference on Computational Complexity*, pages 351–357. IEEE, 2009. doi:10.1109/CCC.2009.25.

24 Troy Lee and Adi Shraibman. Lower bounds in communication complexity. *Foundations and Trends® in Theoretical Computer Science*, 3(4):263–399, 2009. doi:10.1561/0400000040.

25 Nati Linial and Adi Shraibman. Lower bounds in communication complexity based on factorization norms. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 699–708, 2007. doi:10.1145/1250790.1250892.

26 Ashley Montanaro and Tobias Osborne. On the communication complexity of XOR functions. *arXiv preprint arXiv:0909.3392*, 2009. arXiv:0909.3392.

27 Noam Nisan. CREW prams and decision trees. *SIAM J. Comput.*, 20(6):999–1007, 1991. doi:10.1137/0220062.

28 Anup Rao and Amir Yehudayoff. *Communication complexity: and applications*. Cambridge University Press, 2020.

29 Ran Raz. Exponential separation of quantum and classical communication complexity. In *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pages 358–367, 1999. doi:10.1145/301250.301343.

30 Alexander A Razborov. On the distributional complexity of disjointness. In *International Colloquium on Automata, Languages, and Programming*, pages 249–253. Springer, 1990. doi:10.1007/BFB0032036.

31 Alexander A Razborov. Quantum communication complexity of symmetric predicates. *Izvestiya: Mathematics*, 67(1):145, 2003.

32 Oded Regev and Bo'az Klartag. Quantum one-way communication can be exponentially stronger than classical communication. In Lance Fortnow and Salil P. Vadhan, editors, *Proceedings of the 43rd ACM Symposium on Theory of Computing, STOC 2011, San Jose, CA, USA, 6-8 June 2011*, pages 31–40. ACM, 2011. doi:10.1145/1993636.1993642.

33 Norbert Sauer. On the density of families of sets. *Journal of Combinatorial Theory, Series A*, 13(1):145–147, 1972. doi:10.1016/0097-3165(72)90019-2.

34 Saharon Shelah. A combinatorial problem; stability and order for models and theories in infinitary languages. *Pacific Journal of Mathematics*, 41(1):247–261, 1972.

35 Alexander A Sherstov. The pattern matrix method for lower bounds on quantum communication. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 85–94, 2008. doi:10.1145/1374376.1374392.

36 Alexander A Sherstov. The pattern matrix method (journal version). *arXiv preprint arXiv:0906.4291*, 2009. arXiv:0906.4291.

37 Alexander A. Sherstov. On quantum-classical equivalence for composed communication problems. *Quantum Inf. Comput.*, 10(5&6):435–455, 2010. doi:10.26421/QIC10.5-6-5.

38 Alexander A. Sherstov. The multiparty communication complexity of set disjointness. *SIAM J. Comput.*, 45(4):1450–1489, 2016. doi:10.1137/120891587.

39 Alexander A Sherstov, Andrey A Storozhenko, and Pei Wu. An optimal separation of randomized and quantum query complexity. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1289–1302, 2021. doi:10.1145/3406325.3451019.

40 Yaoyun Shi and Yufan Zhu. Quantum communication complexity of block-composed functions. *Quantum Inf. Comput.*, 9(5&6):444–460, 2009. doi:10.26421/QIC9.5-6-7.

41 Andrew Chi-Chih Yao. Some complexity questions related to distributive computing (preliminary report). In *Proceedings of the eleventh annual ACM symposium on Theory of computing*, pages 209–213, 1979. doi:10.1145/800135.804414.

42 Andrew Chi-Chih Yao. Quantum circuit complexity. In *Proceedings of 1993 IEEE 34th Annual Foundations of Computer Science*, pages 352–361. IEEE, 1993. doi:10.1109/SFCS.1993.366852.

43 Shengyu Zhang. On the tightness of the buhrman-cleve-wigderson simulation. In Yingfei Dong, Ding-Zhu Du, and Oscar Ibarra, editors, *Algorithms and Computation*, pages 434–440, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg. doi:10.1007/978-3-642-10631-6_45.

44 Zhiqiang Zhang and Yaoyun Shi. Communication complexities of symmetric XOR functions. *Quantum information and computation*, 9(3&4):255–263, 2009. doi:10.26421/QIC9.3-4-5.