

Exponential Quantum Advantage in Number-on-Forehead Communication

Haoyu Wang*

Pei Wu[†]Guangxu Yang[‡]

Abstract

We give the first exponential quantum advantage in the general interactive three-party Number-on-Forehead (NOF) model for a decision problem. Previous separations hold only for restricted protocols like one-way communication for a relation. We construct an explicit partial Boolean function, the Interleaved Unitary Product problem, that requires only $O(\log n)$ quantum communication but $\tilde{\Omega}(n^{1/32})$ randomized communication in the NOF model. This function builds on the two-party unitary product problem of Arunachalam, Girish, and Lifshitz (TQC 2024).

The main technical obstacle is that discrepancy, the standard lower-bound method for NOF, also lower-bounds quantum communication. We instead develop a regularity-based argument for randomized NOF lower bounds, building on the approach of Kelley, Lovett, and Meka (STOC 2024) and adapting the regularity decomposition of Abboud, Fischer, Kelley, Lovett, and Meka (STOC 2024) to cylinder intersections. Combined with matrix-product estimates of Arunachalam, Girish, and Lifshitz (TQC 2024), this yields our randomized lower bound.

1 Introduction

When can quantum messages convey information more efficiently than classical messages? Communication complexity, introduced by Yao [Yao79], provides a natural setting in which this question can be studied through unconditional lower bounds. A long line of work has established exponential separations between quantum and classical communication complexity [BCW98, Raz99, BYJK04, GKK⁺07, RK11, Gav16, GRT22, Gav19, Gav20, GGJL24, Gav26]. These separations identify tasks for which quantum messages convey useful information far more efficiently than classical messages, without relying on unproved computational hardness assumptions.

Communication separations have also motivated experimental demonstrations of quantum advantage [TSB⁺05, XAW⁺15, GXY⁺16, KKD19, SLW⁺25]. One example is *quantum information supremacy*, proposed by Aaronson, Buhrman, and Kretschmer [ABK24], which uses one-way communication separations to demonstrate an advantage in the information resources required for a task. Recently, a trapped-ion experiment has demonstrated such an advantage, supported by unconditional classical lower bounds [KGD⁺25]. These developments motivate understanding the scope of quantum communication advantage along two dimensions: the information available to the players and the quantum resources available to their protocols.

*Penn State University. Email: hjw5492@psu.edu.

[†]Penn State University. Email: pei.wu@psu.edu.

[‡]University of Southern California. Email: guangxuy@usc.edu.

Quantum advantage with overlapping inputs. The separations discussed above establish quantum advantages in two-party models, where the players have access to disjoint parts of the input. Despite this progress, much less is understood about quantum communication advantage when the players’ input access overlaps. The importance of input-access assumptions is also highlighted by the trapped-ion experiment [KGD⁺25]: the authors identify a potential loophole arising from the fact that Bob’s measurement choice was generated before Alice’s state was prepared, so the preparation device could in principle have advance knowledge of that choice. This motivates a natural research question: *How does overlap in the players’ input access affect quantum communication advantage?*

Quantum advantage with one clean qubit. A complementary question is: *What are the minimal quantum resources needed for an exponential communication advantage?* We focus on initial-state purity. Motivated by nuclear magnetic resonance implementations, where the initial state can be highly mixed, Knill and Laflamme introduced the one-clean-qubit model of quantum computation, also known as DQC1 [KL98]. Only one qubit is initially pure, while the remaining workspace is maximally mixed. The model thus restricts initial-state purity while allowing ideal unitary control. Klauck and Lim introduced a communication version of this model [KL18], making it possible to investigate the power of limited purity through unconditional communication lower bounds. Arunachalam, Girish, and Lifshitz subsequently established an exponential separation between one-clean-qubit quantum communication and interactive randomized communication in the two-party setting [AGL23].

The Number-on-Forehead (NOF) model, introduced by Chandra, Furst, and Lipton [CFL83], is one of the best-known models of multiparty communication and captures a maximal form of input overlap. Lower bounds in the NOF model have applications in various areas including circuit complexity [HG90, BT94, Cha07, VW07], additive combinatorics [LS⁺17, JLL⁺25], distributed computing [DKO14], pseudorandom generators [BNS92, GR14], streaming algorithms [KMPV19], and cryptography [CGK19, MV13, LMQW20]. It therefore provides a natural setting that brings these two questions together. The input consists of k parts, $x_1, \dots, x_k$, and player i sees every part except x_i . Thus, each input part is visible to all but one player, while no player sees the entire input. For $k = 2$, this recovers the standard two-party model, up to relabeling the inputs, while for $k \geq 3$, the players’ views overlap.

The above considerations lead to the following concrete question in multiparty communication:

Does one clean qubit suffice for an exponential quantum advantage in the multi-party NOF communication model?

In this paper, we answer the above question affirmatively in the three-party NOF model. We establish the first unconditional exponential separation between quantum and randomized communication for a partial Boolean function in the three-party NOF model.

1.1 Our contributions

To obtain this separation, we introduce the Interleaved Unitary Product problem, which is inspired by the two-party communication problem in [AGL23].

Definition 1.1 (Interleaved Unitary Product (IUP) problem). Let $m \geq 1$ be an integer. The input consists of complex matrices $A_i, B_i, C_i, D_i, E_i \in \mathbb{C}^{m \times m}$, $i \in \{1, 2\}$. Each input matrix U is promised to satisfy $\|U^\dagger U - I_m\|_F \leq 10^{-4}$. Group the inputs as

$$X = (A_1, C_1, A_2, C_2), \quad Y = (B_1, D_1, B_2, D_2), \quad Z = (E_1, E_2).$$

Define $W = A_1 B_1 C_1 D_1 E_1 A_2 B_2 C_2 D_2 E_2$. In the three-party NOF model, Alice, Bob, and Charlie see (Y, Z) , (X, Z) , and (X, Y) , respectively. Their task is to output

$$b = \begin{cases} 1, & \text{if } \operatorname{Re} \operatorname{Tr}(W)/m \geq 0.9, \\ 0, & \text{if } \operatorname{Re} \operatorname{Tr}(W)/m \leq 0.1 \end{cases}$$

under the promise that $\operatorname{Re} \operatorname{Tr}(W)/m$ doesn't fall into $(0.1, 0.9)$.

To obtain a Boolean function, encode X, Y , and Z as bit strings $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \{0, 1\}^n$, using $O(\log n)$ bits for each matrix entry and padding $\mathbf{z}$ with zeros. This allows matrices of dimension $m = \Theta(\sqrt{n/\log n})$; see Definition 2.1 for the formal description.

Our first result shows that the interleaved unitary product problem can be tested with logarithmic quantum communication under the one-clean-qubit guarantee. The protocol is a natural adaptation of that in [AGL23] and is presented with its analysis in Section 3.

Theorem 1.2 (Quantum upper bound). *There is a bounded-error one-clean-qubit quantum NOF protocol for the IUP problem using $O(\log n)$ qubits of communication.*¹

In contrast, every classical randomized protocol requires polynomial communication, even with unrestricted interaction among all three players. This lower bound is our main technical contribution.

Theorem 1.3 (Randomized lower bound). *Every bounded-error randomized NOF protocol for the IUP problem requires $\widetilde{\Omega}(n^{1/32})$ bits of communication.*

We prove this lower bound in Section 4. Together, Theorems 1.2 and 1.3 give the claimed exponential separation for a partial Boolean function, while requiring only one clean qubit on the quantum side.

1.2 Related work

Gavinsky and Pudlák [GP08] gave relational problems with $O(\log n)$ -qubit simultaneous quantum protocols and $n^{\Omega(1/k^2)}$ randomized lower bounds in a non-interactive NOF model. Their separation is exponential for constant k and superpolynomial for $k = o(\sqrt{\log n / \log \log n})$. In their classical model, the first $k - 1$ players each send a message to the final player without seeing one another's messages. Yang and Zhang [YZ25b] gave a gadgeted Hidden Matching construction for three-party simultaneous NOF, obtaining an $O(\log n)$ quantum upper bound and an $\Omega(n^{1/16})$ randomized lower bound. Their subsequent work [YZ26] extended this direction to one-way NOF communication, giving an $O(\log n)$ quantum upper bound and an $\Omega(n^{1/3}/2^{k/3})$ randomized lower bound. These results consider relational problems and restrictions on classical interaction. Our result allows *unrestricted* interaction among all three classical players and establishes the separation for a partial *Boolean* function.

1.3 Technical Overview

The bulk of technical work in this paper is the classical lower bound in the NOF model. The overlap of information among the players is exactly what makes NOF lower bounds notoriously difficult

¹Here, “one clean qubit” means one per run with a constant acceptance-probability gap. We use $O(1)$ independent repetitions to achieve any fixed constant error, preserving $O(\log n)$ total communication. Such repetitions are allowed in the original DQC1 model of Knill and Laflamme [KL98]. Arunachalam, Girish, and Lifshitz [AGL23] likewise use $O(1)$ repetitions for amplification, using $O(1)$ clean qubits in total. Klauck and Lim [KL18] instead incorporate the amplification overhead into the cost.

to prove. A quantum-classical separation faces a second obstacle: the discrepancy and generalized discrepancy methods, which underlie many known lower bounds for interactive NOF protocols, also lower bound quantum communication, and therefore cannot separate the two models.

Our classical lower bound builds on a strong corruption bound of cylinder intersections, which shows that distinguishing two well-separated distributions requires the claimed communication cost. To prove this corruption bound, we adapt the regularity decomposition for Boolean matrices developed by Abboud, Fischer, Kelley, Lovett, and Meka [AFK⁺24] to cylinder intersections, which allows us to treat the structure (sparse here) and pseudorandom cases separately. Besides, the matrix-product estimates of Arunachalam, Girish, and Lifshitz [AGL23] and the regularity estimates of Kelley, Lovett, and Meka [KLM24] provide the key technical tools for establishing the corruption bound.

Hard distributions and corruption bound. We first work with exact matrices in $SU(m)$. Under μ_0 , all inputs are independent Haar matrices; under μ_1 , all but the last matrix are sampled independently from the Haar measure. The last matrix is then set to the inverse of their product, ensuring that $W = I_m$. Thus μ_0 produces 0-inputs with high probability while μ_1 produces 1-inputs. On the other hand, Haar invariance makes every player's view identically distributed under both measures: the distinguishing information lies in the global correlation. In the NOF model, each fixed transcript corresponds to a cylinder intersection with indicator $c(X, Y, Z) = f(Y, Z)g(X, Z)h(X, Y)$. We prove $\mu_1(c) \leq \alpha\mu_0(c) + e^{-\beta m^{1/16}}$ for constants $\alpha > 1$ and $\beta > 0$.

Step 1. Decompose cylinder intersections. Fix $d = \Theta(m^{1/16})$. By our cylinder intersection decomposition adapted from [AFK⁺24, Theorem 3.1], we decompose c into k disjoint pieces: $c_i = f_i g_i h_i$ on $\mathcal{B}_i := \mathcal{X}_i \times \mathcal{Y}_i \times \mathcal{Z}_i$, $1 \leq i \leq k$, where $c = \sum_{i=1}^k c_i$, $k \leq e^{O(d^7)}$ and $\sum_i \mu_0(\mathcal{B}_i) = O(d^2)$. For $u_i \in \{f_i, g_i, h_i\}$, let $\delta_{u,i} := \mathbb{E}_{\mathcal{B}_i}[u_i]$. There are two kinds of pieces: one kind of pieces is *sparse*, satisfying $\delta_{f,i} \leq 2^{-d}$, $\delta_{g,i} \leq 2^{-d}$, or $\delta_{h,i} \leq 2^{-\Omega(d)}$; the remaining pieces are *pseudorandom*: both f_i and g_i satisfy the regularity and min-degree properties of [KLM24], which allow their product average to be compared with the product of their averages.

Step 2. Derive the grid-norm bound. The regularity notion of [KLM24] uses grid norms $\|\cdot\|_{U(\ell,d)}$ to measure uniformity relative to density. From the matrix-product estimate of [AGL23], we derive the following grid norm bound: for every integer $p \geq 2$,

$$\mu_1(c_i) \leq O\left(\mu_0(\mathcal{B}_i) \cdot \delta_{h,i}^{1-1/p} \cdot \|f_i\|_{U(2,p),i} \cdot \|g_i\|_{U(2,p),i}\right) + e^{-\Omega(\sqrt{m}/p)}, \quad (1.1)$$

where the norms $\|\cdot\|_{U(2,p),i}$ are local versions on $\mathcal{B}_i$.

Step 3. Establish the corruption bound. For regular pieces, the regularity property bounds the two norms by the local expectations, i.e., $\|f_i\|_{U(2,d),i} \|g_i\|_{U(2,d),i} = O(\delta_{f,i} \delta_{g,i})$; besides, the estimate of [KLM24] also gives $\mu_0(c_i) \geq \Omega(\mu_0(\mathcal{B}_i) \delta_{f,i} \delta_{g,i} \delta_{h,i})$. Together with (1.1), this yields the corruption bound on regular cylinder intersections: $\mu_1(c_i) \leq O(\mu_0(c_i)) + e^{-\Omega(\sqrt{m}/d)}$.

For sparse pieces, by standard calculation $\|f_i\|_{U(2,2),i} \|g_i\|_{U(2,2),i} = O(\sqrt{\delta_{f,i} \delta_{g,i}})$; applying (1.1) with $p = 2$ and $\sum_i \mu_0(\mathcal{B}_i) = O(d^2)$, we have $\sum_{i: c_i \text{ is sparse}} \mu_1(c_i) \leq O(d^2) 2^{-\Omega(d)} + k e^{-\Omega(\sqrt{m})}$.

Combining the corruption bounds of regular and sparse pieces, for any cylinder intersection c ,

$$\mu_1(c) \leq \alpha \cdot \mu_0(c) + e^{-\beta m^{1/16}},$$

where $\alpha > 1, \beta > 0$ are constants.

Step 4. Deduce the communication lower bound. Fix a randomized protocol with error $\leq 1/3$ and cost t . Summing over all possible 2^t cylinder intersections, we have

$$2/3 \leq \sum \mu_1(\mathbf{c}) \leq \alpha \cdot \sum \mu_0(\mathbf{c}) + 2^t \cdot e^{-\beta m^{1/16}} \leq 1/3 + 2^t \cdot e^{-\beta m^{1/16}}.$$

Hence $t = \Omega(m^{1/16}) = \Omega((n/\log n)^{1/32})$.

2 Preliminaries

Notation. We use standard matrix notation. The norms $\|\cdot\|_2$, $\|\cdot\|_F$, and $\|\cdot\|_{\text{op}}$ denote the Euclidean, Frobenius, and operator norms, respectively. Write $\text{U}(m)$ and $\text{SU}(m)$ for the unitary and special unitary groups.

For $\delta \geq 0$, denote the set of matrices close to unitary by

$$\mathcal{M}_m(\delta) = \{M \in \mathbb{C}^{m \times m} : \|M^\dagger M - I_m\|_F \leq \delta\}.$$

Let ν denote Haar probability measure on $\text{SU}(m)$, and $\nu^{\otimes q}$ its q -fold product. Haar measure is invariant under left and right multiplication and inversion. For a probability measure μ , write

$$\mu(\mathbf{u}) = \mathbb{E}_\mu[\mathbf{u}], \quad \mu(\mathcal{S}) = \mu(\mathbf{1}_\mathcal{S}), \quad \|\mathbf{u}\|_p = (\mathbb{E}_\mu[|\mathbf{u}|^p])^{1/p} \quad (1 \leq p < \infty),$$

whenever these expectations are defined. For a nonempty finite set $\mathcal{S}$, $x \sim \mathcal{S}$ means uniform sampling. Samples are independent unless a joint distribution or conditioning is specified.

Communication complexity. We use the three-player number-on-forehead (NOF) model [CFL83, KN97]. For inputs $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \{0, 1\}^n$, Alice, Bob, and Charlie see $(\mathbf{y}, \mathbf{z})$, $(\mathbf{x}, \mathbf{z})$, and $(\mathbf{x}, \mathbf{y})$, respectively. Classical protocols allow public randomness and unrestricted blackboard interaction. Local computation is free, and communication cost is the worst-case total number of bits or qubits exchanged. Unless stated otherwise, protocols have error at most $1/3$ on every promised input.

Fixing the random coins of a classical NOF protocol of cost at most t on all inputs partitions its input space into at most 2^t *cylinder intersections*, one for each transcript. Their indicators have the form

$$\mathbf{c}(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \mathbf{f}(\mathbf{y}, \mathbf{z})\mathbf{g}(\mathbf{x}, \mathbf{z})\mathbf{h}(\mathbf{x}, \mathbf{y}),$$

where the factors are measurable Boolean functions of the indicated arguments.

Input representations. When bit strings are given as input, players interpret them as complex matrices. Let $b = \lceil \log n \rceil$, use $b + 2$ bits to represent a number in $[-2, 2)$ with precision 2^{-b} . For example, given a bit string $\mathbf{x} \in \{0, 1\}^n$, divide it into blocks of length $b + 2$. Each block $\mathbf{a} = (\mathbf{a}_1, \dots, \mathbf{a}_{b+2}) \in \{0, 1\}^{b+2}$ is decoded as a rational number in two's-complement representation, i.e.,

$$\text{dec}(\mathbf{a}) = -2 \cdot \mathbf{a}_{b+2} + \sum_{r=1}^{b+1} 2^{r-1-b} \cdot \mathbf{a}_r \in [-2, 2) \cap 2^{-b} \cdot \mathbb{Z}.$$

Each consecutive pair of blocks forms a complex number, with the first as the real part and the second as the imaginary part. Decode the bit string from left to right into complex numbers, then arrange them row by row into $m \times m$ matrices, where $m = \Theta(\sqrt{n/\log n})$. Ignoring redundant bits, decode $\mathbf{x}$ into $X = (A_1, C_1, A_2, C_2)$, and similarly decode $\mathbf{y}$ and $\mathbf{z}$ into $Y = (B_1, D_1, B_2, D_2)$ and $Z = (E_1, E_2)$.

Definition 2.1 (IUP function). Given $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \{0, 1\}^n$, in the NOF view, Alice sees $(\mathbf{y}, \mathbf{z})$, Bob sees $(\mathbf{x}, \mathbf{z})$ and Charlie sees $(\mathbf{x}, \mathbf{y})$. They decode the matrices in their respective views, obtaining (Y, Z) , (X, Z) , and (X, Y) , respectively. Define

$$F_n(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \begin{cases} 0, & \text{if } \text{Re Tr}(W)/m \leq 0.1 \text{ \& } X, Y, Z \subseteq \mathcal{M}_m(10^{-4}); \\ 1, & \text{if } \text{Re Tr}(W)/m \geq 0.9 \text{ \& } X, Y, Z \subseteq \mathcal{M}_m(10^{-4}); \\ *, & \text{Otherwise.} \end{cases}$$

Under the promise that $F_n \neq *$, the players communicate to determine whether F_n is 0 or 1.

Quantum communication. We follow the standard quantum formalism [NC10]. Our protocol uses no prior entanglement and passes a register consisting of a control qubit and an m -dimensional target with basis $\{|j\rangle : j \in [m]\}$. We use $|+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$ and $|-\rangle = \frac{|1\rangle-|0\rangle}{\sqrt{2}}$. For $U \in \text{U}(m)$, define the controlled operation $\text{ctrl}(U) = |0\rangle\langle 0| \otimes I_m + |1\rangle\langle 1| \otimes U$. In the one-clean-qubit communication model [KL18], the initial state is $\varrho_0 = |0\rangle\langle 0| \otimes I_{2^q}/2^q$ for an integer $q \geq 0$. Players apply input-dependent unitaries and exchange these qubits, without additional quantum memory, before a fixed projective measurement. We allow a constant number of independent runs, each using one clean qubit, as specified in the introduction.

3 Quantum protocol

In this section, we prove Theorem 1.2 by giving an $O(\log n)$ -qubit protocol for the IUP problem. To be concise, let P_1, P_2, P_3 denote Alice, Bob, and Charlie, respectively. After decoding, their views are (Y, Z) , (X, Z) , and (X, Y) , where

$$X = (A_1, C_1, A_2, C_2), \quad Y = (B_1, D_1, B_2, D_2), \quad Z = (E_1, E_2).$$

And Recall that

$$W = A_1 B_1 C_1 D_1 E_1 A_2 B_2 C_2 D_2 E_2.$$

We first replace each input matrix U by its unitary polar factor $\tilde{U} = U(U^\dagger U)^{-1/2}$, and let $\tilde{W}$ be the product of these factors in the same order as in W . The promise $U \in \mathcal{M}_m(\sigma)$, with $\sigma = 10^{-4}$, ensures that this replacement preserves the trace gap.

Claim 3.1. *For every input matrix U , $\tilde{U}^\dagger \tilde{U} = I_m$, $\|\tilde{U} - U\|_{\text{op}} \leq \sigma$. Moreover,*

$$\frac{|\text{Re Tr } \tilde{W} - \text{Re Tr } W|}{m} \leq 10\sigma(1 + \sigma)^{9/2}.$$

Claim 3.1 is proved in Appendix C. Using the register and controlled operations defined in Section 2, the players execute the following protocol.

Protocol 1 Quantum protocol for gap trace estimation

- 1: Each player computes $\tilde{U} = U(U^\dagger U)^{-1/2}$ for every input matrix U in their view.
 - 2: $L \leftarrow 100$.
 - 3: **for** $\ell = 1, \dots, L$ **do**
 - 4: P_2 samples j_ℓ uniformly from $[m]$, independently of earlier trials, and prepares $|\Phi\rangle \leftarrow |+\rangle |j_\ell\rangle$.
 - 5: **for** $r = 1, \dots, 10$ **do**
 - 6: P_{i_r} applies $\text{ctrl}(\tilde{U}_r)$ to $|\Phi\rangle$, according to the table:

r	1	2	3	4	5	6	7	8	9	10
U_r	E_2	D_2	C_2	B_2	A_2	E_1	D_1	C_1	B_1	A_1
i_r	2	1	3	1	3	2	1	3	1	3
 - 7: **if** $r < 10$ **then**
 - 8: P_{i_r} sends the register to $P_{i_{r+1}}$.
 - 9: **end if**
 - 10: **end for**
 - 11: P_3 measures $|\Phi\rangle$ with $\{|+\rangle\langle+| \otimes I_m, |-\rangle\langle-| \otimes I_m\}$, and let $\omega_\ell = \begin{cases} 1, & |+\rangle\langle+| \otimes I_m, \\ -1, & |-\rangle\langle-| \otimes I_m; \end{cases}$
 - 12: **end for**
 - 13: P_3 outputs $\mathbf{1}\left[L^{-1} \sum_{\ell=1}^L \omega_\ell > 1/2\right]$.
-

Each controlled operation is performed by a player who sees its matrix. Averaging over j_ℓ , the initial message state is $|+\rangle\langle+| \otimes I_m/m$. The operations apply the factors from right to left in $\tilde{W}$, so, conditioned on the sampled $j_\ell \in [m]$, the final state is

$$|\Phi\rangle = \frac{|0\rangle |j_\ell\rangle + |1\rangle \tilde{W} |j_\ell\rangle}{\sqrt{2}}.$$

Thus,

$$\Pr[\omega_\ell = 1 \mid j_\ell] = \frac{1}{4} \left\| |j_\ell\rangle + \tilde{W} |j_\ell\rangle \right\|_2^2 = \frac{1 + \text{Re } \tilde{W}_{j_\ell j_\ell}}{2}, \text{ and } \Pr[\omega_\ell = -1 \mid j_\ell] = \frac{1 - \text{Re } \tilde{W}_{j_\ell j_\ell}}{2}.$$

Taking the expectation over j_ℓ gives

$$\mathbb{E}[\omega_\ell] = \frac{1}{m} \sum_{j=1}^m \text{Re } \tilde{W}_{jj} = \frac{\text{Re Tr } \tilde{W}}{m}.$$

By Claim 3.1 and taking $\sigma = 10^{-4}$, replacing the input matrices by their polar factors changes the normalized trace by less than 0.002. Thus,

$$\mathbf{F}_n = 1 \implies \mathbb{E}[\omega_\ell] > 0.898, \quad \mathbf{F}_n = 0 \implies \mathbb{E}[\omega_\ell] < 0.102.$$

Notice that $j_1, \dots, j_L$ are independent and $\text{Var}(\omega_\ell) \leq 1$. For $L = 100$, Chebyshev's inequality therefore yields

$$\Pr\left[\mathbf{1}\left[\frac{1}{L} \sum_{\ell=1}^L \omega_\ell > \frac{1}{2}\right] \neq \mathbf{F}_n\right] \leq \frac{1}{3}.$$

Finally, each iteration uses nine message transmissions, each consisting of $1 + \lceil \log_2 m \rceil$ qubits. Recall that $m = \Theta(\sqrt{n/\log n})$; hence the total communication cost is $O(\log n)$.

Remark: The $O(\log n)$ quantum upper bound also holds in the more restricted number-in-hand (NIH) model, where P_1 , P_2 , and P_3 hold $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$, respectively. Run Protocol 1 in the same matrix order, assigning each controlled operation to the player holding its matrix; P_3 initializes the register and P_1 measures and outputs the answer. The communication cost and error probability are unchanged. This quantum protocol for IUP does not rely on the overlap of inputs. Together with Theorem 1.3, this establishes an exponential separation even between quantum protocols in the NIH model and randomized protocols in the NOF model.

4 Randomized lower bound

In this section, we prove Theorem 1.3, establishing the claimed randomized communication lower bound. The high level strategy is to define two input distributions under which the IUP function is 0 and 1 with high probability, respectively; then establish a corruption bound showing that distinguishing these distributions with significant constant advantage requires $\Omega((n/\log n)^{1/32})$ communication cost. Moreover, we adapt the regularity decomposition of Abboud, Fischer, Kelley, Lovett, and Meka [AFK⁺24] to split cylinder intersections into sparse parts and regular (pseudorandom-like) parts. Then we combine the regularity guarantees with the matrix-product estimate of Arunachalam, Girish, and Lifshitz [AGL23] to bound these parts separately and obtain the corruption bound.

At first, we define two distributions μ_1 and μ_0 of (X, Y, Z) on $\text{SU}(m)^4 \times \text{SU}(m)^4 \times \text{SU}(m)^2$, where $\text{SU}(m)$ denotes the special unitary group. Let ν be the Haar measure of $\text{SU}(m)$.

1. μ_0 : For each $U \in \{A_1, B_1, C_1, D_1, E_1, A_2, B_2, C_2, D_2, E_2\}$, sample them independently from ν .
2. μ_1 : For all the matrices except E_2 , sample them i.i.d by ν . Then let

$$P = A_1 B_1 C_1 D_1, \quad Q = A_2 B_2 C_2 D_2, \quad \text{and} \quad E_2 = (P E_1 Q)^{-1}.$$

Hence under μ_1 , $W = I_m$; under μ_0 , $\text{Re Tr}(W)/m$ is close to zero with high probability. Haar invariance gives the following marginal identities.

Claim 4.1. μ_0 and μ_1 have identical marginals on each NOF view, i.e.,

$$(X, Y) \sim \nu^{\otimes 4} \times \nu^{\otimes 4}, \quad (X, Z) \sim \nu^{\otimes 4} \times \nu^{\otimes 2}, \quad (Y, Z) \sim \nu^{\otimes 4} \times \nu^{\otimes 2}.$$

The proof of Claim 4.1 is in Appendix C.

To apply a protocol on bit-string inputs to Haar-distributed matrices, we introduce the following round process, which maps the coordinates of a continuous complex matrix to the decoding grid $[-2, 2) \cap 2^{-b} \cdot \mathbb{Z}$. For any measurable function u and set $\mathcal{S}$, write $\mu_j(u) := \mathbb{E}_{\mu_j}[u]$ and $\mu_j(\mathcal{S}) := \mu_j(\mathbf{1}_{\mathcal{S}})$ for $j \in \{0, 1\}$.

Rounding Process. Recall that $b = \lceil \log n \rceil$. Replace every real or imaginary coordinate $t \in [-1, 1]$ by

$$2^{-b} \left\lfloor 2^b t - \frac{1}{2} \right\rfloor \in [-2, 2) \cap 2^{-b} \mathbb{Z}.$$

This is nearest-grid rounding, with ties rounded down. Encode these coordinates using the inverse of dec , in the decoding order specified in Section 2, and set unused bits to zero. Denote the resulting encoding maps by $R_x : \text{SU}(m)^4 \rightarrow \{0, 1\}^n$, $R_y : \text{SU}(m)^4 \rightarrow \{0, 1\}^n$, and $R_z : \text{SU}(m)^2 \rightarrow \{0, 1\}^n$; write $R(X, Y, Z) = (R_x(X), R_y(Y), R_z(Z))$. For $j \in \{0, 1\}$, let $\tilde{\mu}_j$ be the distribution of $R(X, Y, Z)$ when $(X, Y, Z) \sim \mu_j$. Then, for every Boolean input set $\mathcal{S}$,

$$\tilde{\mu}_j(\mathcal{S}) = \mu_j(R^{-1}(\mathcal{S})) \quad (j = 0, 1). \quad (4.1)$$

Claim 4.2. *For all sufficiently large n , the inputs under μ_0 and μ_1 satisfy*

$$\Pr_{\mu_1}[\text{Re Tr}(W)/m \geq 0.9] = 1, \quad \Pr_{\mu_0}[\text{Re Tr}(W)/m \leq 0.1] \geq 1 - \frac{1}{m}.$$

After the rounding process, every matrix belongs to $\mathcal{M}_m(10^{-4})$, and

$$\Pr_{\tilde{\mu}_1}[F_n = 1] = 1, \quad \Pr_{\tilde{\mu}_0}[F_n = 0] \geq 1 - \frac{1}{m}.$$

The proof of Claim 4.2 is in Appendix C. We can therefore work with Haar matrices and apply the Boolean protocol after rounding. Equation (4.1) preserves every acceptance probability exactly.

Cylinder intersection. Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be measurable spaces. Given measurable Boolean functions

$$f : \mathcal{Y} \times \mathcal{Z} \rightarrow \{0, 1\}, \quad g : \mathcal{X} \times \mathcal{Z} \rightarrow \{0, 1\}, \quad h : \mathcal{X} \times \mathcal{Y} \rightarrow \{0, 1\},$$

define the cylinder intersection indicator by

$$c : \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \rightarrow \{0, 1\}, \quad c(X, Y, Z) := f(Y, Z) \cdot g(X, Z) \cdot h(X, Y).$$

The randomized lower bound follows directly from the following corruption bound, which shows any short classic protocol can not distinguish μ_0 and μ_1 significantly on cylinder intersections.

Lemma 4.3 (corruption bound). *Let $\mathcal{X} = \mathcal{Y} = \text{SU}(m)^4$ and $\mathcal{Z} = \text{SU}(m)^2$. There exist universal constants $\alpha > 1$ and $\beta > 0$ such that for each cylinder intersection c , we have*

$$\mu_1(c) \leq \alpha \cdot \mu_0(c) + \exp(-\beta m^{1/16}).$$

The proof of Lemma 4.3 is deferred to Section 4.1. We are now ready to prove Theorem 1.3.

Proof of Theorem 1.3. By constant amplification and truncation, take a protocol Π with cost at most t on all inputs, and error at most $\eta = 1/[4(\alpha + 1)]$, where α is the constant in the corruption bound. The amplification changes the cost by only a constant factor. Let ρ denote the random coins. Then

$$\Pr_{\rho}[\Pi_{\rho}(\mathbf{x}, \mathbf{y}, \mathbf{z}) \neq F_n(\mathbf{x}, \mathbf{y}, \mathbf{z})] \leq \eta \quad \text{whenever } F_n(\mathbf{x}, \mathbf{y}, \mathbf{z}) \neq *.$$

Fix ρ ; let $s \leq 2^t$ be the number of accepting transcript leaves, and let $\tilde{c}_{\ell}$ be the indicator of the ℓ -th leaf, for $\ell \in [s]$. Distinct leaves correspond to disjoint sets of inputs. Each $\tilde{c}_{\ell}$ is a Boolean cylinder intersection. Since the rounding process R encodes X, Y, Z separately, $c_{\ell} := \tilde{c}_{\ell} \circ R$ is also a cylinder intersection.

The probability identity (4.1) gives $\tilde{\mu}_j(\tilde{c}_{\ell}) = \mu_j(c_{\ell})$ for $j = 0, 1$. Apply the corruption bound of Lemma 4.3:

$$\tilde{\mu}_1(\tilde{c}_{\ell}) = \mu_1(c_{\ell}) \leq \alpha \mu_0(c_{\ell}) + e^{-\beta m^{1/16}} = \alpha \tilde{\mu}_0(\tilde{c}_{\ell}) + e^{-\beta m^{1/16}}.$$

Summing over the accepting leaves and averaging over ρ yields

$$\Pr_{\tilde{\mu}_1, \rho}[\Pi_{\rho} = 1] \leq \alpha \Pr_{\tilde{\mu}_0, \rho}[\Pi_{\rho} = 1] + 2^t e^{-\beta m^{1/16}}.$$

Claim 4.2 says that $\tilde{\mu}_1$ is supported on 1-inputs and $\tilde{\mu}_0$ assigns probability at most $1/m$ to inputs with $F_n \neq 0$. Together with the error bound above, this gives

$$\Pr_{\tilde{\mu}_1, \rho}[\Pi_{\rho} = 1] \geq 1 - \eta, \quad \Pr_{\tilde{\mu}_0, \rho}[\Pi_{\rho} = 1] \leq \eta \Pr_{\tilde{\mu}_0}[F_n = 0] + \Pr_{\tilde{\mu}_0}[F_n \neq 0] \leq \eta + \frac{1}{m}.$$

Hence, for sufficiently large m ,

$$2^t e^{-\beta m^{1/16}} \geq 1 - (\alpha + 1)\eta - \frac{\alpha}{m} \geq \frac{1}{2}, \quad t \geq \frac{\beta}{\ln 2} m^{1/16} - 1.$$

Therefore the protocol must cost $t = \Omega(m^{1/16}) = \Omega((n/\log n)^{1/32})$ bits. $\square$

4.1 Corruption bound via regularity decomposition

In this section, we prove Lemma 4.3. We first prove the bound for finite-type cylinder intersections, whose factors are constant on finite equal-measure partitions, so that we can apply the matrix decomposition of Abboud, Fischer, Kelley, Lovett, and Meka [AFK⁺24]. We then extend the bound to all measurable cylinder intersections by approximation.

At first, we define the *finite-type* cylinder intersection.

Definition 4.4 (finite type). A cylinder intersection $\mathbf{c} = \mathbf{fgh}$ is *finite-type* if there are finite measurable partitions $\mathcal{P}_X, \mathcal{P}_Y, \mathcal{P}_Z$ of $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$, respectively, such that for every $\mathcal{X}' \in \mathcal{P}_X, \mathcal{Y}' \in \mathcal{P}_Y, \mathcal{Z}' \in \mathcal{P}_Z$,

$$\nu^{\otimes 4}(\mathcal{X}') = \frac{1}{|\mathcal{P}_X|}, \quad \nu^{\otimes 4}(\mathcal{Y}') = \frac{1}{|\mathcal{P}_Y|}, \quad \nu^{\otimes 2}(\mathcal{Z}') = \frac{1}{|\mathcal{P}_Z|},$$

and the restrictions

$$\mathbf{f}|_{\mathcal{Y}' \times \mathcal{Z}'}, \quad \mathbf{g}|_{\mathcal{X}' \times \mathcal{Z}'}, \quad \mathbf{h}|_{\mathcal{X}' \times \mathcal{Y}'} \quad \text{are constant.}$$

All averages of functions on the cells $\mathcal{X}', \mathcal{Y}', \mathcal{Z}'$ or a union of cells use *normalized* Haar measure.

The following claim shows that every measurable cylinder intersection admits a finite-type approximation.

Claim 4.5. *For any measurable cylinder intersection indicator $\mathbf{c} : \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \rightarrow \{0, 1\}$ and any integer $r \geq 1$, there exists a finite-type cylinder intersection $\mathbf{c}_r$ such that $|\mu_i(\mathbf{c} - \mathbf{c}_r)| \leq 1/r$ for $i = 0, 1$.*

The proof of Claim 4.5 is in Appendix C.

To prove Lemma 4.3, we follow the structure-versus-pseudorandomness approach: we decompose cylinder intersections into sparse parts and pseudorandom parts, and bound them separately. We use grid regularity, introduced in the context of communication complexity by Kelley, Lovett, and Meka [KLM24], building on the work of Kelley and Meka [KM23] on 3-term arithmetic progressions. We obtain the regularity decomposition on cylinder intersections by adapting the results of Abboud, Fischer, Kelley, Lovett, and Meka [AFK⁺24].

Definition 4.6 (Grid norm, grid regularity, and min-degree). Let $\mathcal{S}, \mathcal{T}$ be probability spaces. For a non-negative function $\mathbf{u} : \mathcal{S} \times \mathcal{T} \rightarrow \mathbb{R}_+$ and integers $k, \ell \geq 1$, define the *grid norm*² by

$$\|\mathbf{u}\|_{U(k, \ell)} := \left(\mathbb{E}_{\substack{s_1, \dots, s_k \sim \mathcal{S} \\ t_1, \dots, t_\ell \sim \mathcal{T}}} \left[\prod_{a=1}^k \prod_{b=1}^\ell \mathbf{u}(s_a, t_b) \right] \right)^{1/(k\ell)},$$

where all samples are independent. We use $k = 2$ throughout; for an integer $d \geq 1$,

$$\|\mathbf{u}\|_{U(2, d)} = \left(\mathbb{E}_{s, s' \sim \mathcal{S}} \left[\left(\mathbb{E}_{t \sim \mathcal{T}} [\mathbf{u}(s, t) \mathbf{u}(s', t)] \right)^d \right] \right)^{1/(2d)}.$$

²We follow the notation of Kelley, Lovett, and Meka [KLM24] and Abboud, Fischer, Kelley, Lovett, and Meka [AFK⁺24]. The letter U stands for uniformity, as in the Gowers uniformity norms of additive combinatorics.

Let $\delta := \mathbb{E}_{s \sim \mathcal{S}, t \sim \mathcal{T}}[u(s, t)]$. Given $\epsilon > 0$, u is (ϵ, d) -regular if ³

$$\|u\|_{U(2,d)} \leq (1 + \epsilon) \cdot \delta.$$

The function u satisfies ϵ -min-degree if

$$\text{for any } s \in \mathcal{S}, \quad \mathbb{E}_{t \in \mathcal{T}}[u(s, t)] \geq (1 - \epsilon) \cdot \delta.$$

In the following argument, we refer to grid regularity simply as regularity.

Theorem 4.7 (Regularity decomposition, adapted from [AFK⁺24]). *Fix $\epsilon > 0$ and an integer $d \geq 1$. For any finite-type cylinder intersection $c = fgh$ on $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, there exist k sets $\mathcal{B}_i = \mathcal{X}_i \times \mathcal{Y}_i \times \mathcal{Z}_i \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, indexed by $i \in [k]$, where $\mathcal{X}_i, \mathcal{Y}_i, \mathcal{Z}_i$ are nonempty unions of cells from their respective finite-type partitions, satisfying*

1. *For each $i \in [k]$, there exist Boolean functions f_i and g_i , which are constant on products of the partition cells and supported on $\mathcal{Y}_i \times \mathcal{Z}_i$ and $\mathcal{X}_i \times \mathcal{Z}_i$, s.t. $c = \sum_{i=1}^k f_i g_i h$; ⁴*
2. *$\sum_{i=1}^k \mu_0(\mathcal{B}_i) \leq O(d^2)$ and $k \leq \exp(c_\epsilon d^7)$, where c_ϵ is a constant only depending on ϵ ;*
3. *For each $i \in [k]$, let $\delta_{f,i} := \mathbb{E}_{Y \sim \mathcal{Y}_i, Z \sim \mathcal{Z}_i}[f_i(Y, Z)]$ and $\delta_{g,i} := \mathbb{E}_{X \sim \mathcal{X}_i, Z \sim \mathcal{Z}_i}[g_i(X, Z)]$. ⁵ Either*
 - (a) *$\delta_{f,i} \leq 2^{-d}$ or $\delta_{g,i} \leq 2^{-d}$;*
 - (b) *or both f_i and g_i are (ϵ, d) -regular and satisfy ϵ -min-degree on $\mathcal{Y}_i \times \mathcal{Z}_i$ and $\mathcal{X}_i \times \mathcal{Z}_i$, respectively. ⁶*

The proof of Theorem 4.7 is in Appendix A. Since f_i and g_i are 0 outside the support, let $h_i := h \cdot \mathbf{1}_{\mathcal{X}_i} \cdot \mathbf{1}_{\mathcal{Y}_i}$ and $c_i := f_i g_i h_i$; then $c = \sum_{i=1}^k c_i$. We further define $\delta_{h,i} := \mathbb{E}_{x \sim \mathcal{X}_i, y \sim \mathcal{Y}_i}[h_i(x, y)]$.

With Theorem 4.7 in hand, we treat $\{c_i\}_{i=1}^k$ in two cases: *regular* and *sparse*. Define the regular set and sparse set as

$$\mathcal{I}_r := \{i \in [k] : \delta_{f,i} > 2^{-d}, \delta_{g,i} > 2^{-d}, \delta_{h,i} > 2^{-\epsilon d}\}, \quad \text{and} \quad \mathcal{I}_s := [k] \setminus \mathcal{I}_r.$$

For the next two lemmas, fix a sufficiently small constant $\epsilon > 0$ and a sufficiently large integer d .

Lemma 4.8 (Regular case). *There exist universal constants $\kappa > 0$ and $\alpha > 1$ such that for any $i \in \mathcal{I}_r$, we have*

$$\mu_1(c_i) \leq \alpha \cdot \mu_0(c_i) + \exp(-\kappa \sqrt{m}/d).$$

Lemma 4.9 (Sparse case). *There exist universal constants $C, \kappa > 0$ such that*

$$\sum_{i \in \mathcal{I}_s} \mu_1(c_i) \leq C d^2 2^{-\epsilon d/2} + |\mathcal{I}_s| e^{-\kappa \sqrt{m}/2}.$$

We prove Lemma 4.8 and Lemma 4.9 in Section 4.1.1 and Section 4.1.2 respectively. Now combining these two lemmas we are ready to prove Lemma 4.3.

³We write (ϵ, d) -regular for $(\epsilon, 2, d)$ -regular in Abboud, Fischer, Kelley, Lovett, and Meka [AFK⁺24], since only the case $k = 2$ is used here.

⁴Both functions are defined on the full input spaces and vanish outside the indicated subsets. They need not be restrictions of f and g .

⁵ $\delta_{f,i}$ and $\delta_{g,i}$ use normalized Haar measure on $\mathcal{Y}_i \times \mathcal{Z}_i$ and $\mathcal{X}_i \times \mathcal{Z}_i$; this expectation does not necessarily equal the expectation on the whole input space $\mathcal{Y} \times \mathcal{Z}$ or $\mathcal{X} \times \mathcal{Z}$.

⁶Regularity and min-degree use normalized Haar measure on the indicated subsets $\mathcal{Y}_i \times \mathcal{Z}_i$ and $\mathcal{X}_i \times \mathcal{Z}_i$.

Proof of Lemma 4.3. Fix $\epsilon = 10^{-4}$ and first suppose $\mathbf{c}$ is finite-type. Theorem 4.7 gives Boolean functions $\mathbf{c}_i$ with $\mathbf{c} = \sum_{i=1}^k \mathbf{c}_i$ and $k \leq e^{c_\epsilon d^7}$. Their supports are disjoint because $\mathbf{c}$ is Boolean. Lemma 4.8 compares each regular term with its μ_0 -mass, while Lemma 4.9 bounds the total sparse mass. Taking a common smaller $\kappa > 0$ and $C \geq 1$, we obtain

$$\begin{aligned} \mu_1(\mathbf{c}) &= \sum_{i \in \mathcal{I}_r} \mu_1(\mathbf{c}_i) + \sum_{i \in \mathcal{I}_s} \mu_1(\mathbf{c}_i) \\ &\leq \alpha \sum_{i \in \mathcal{I}_r} \mu_0(\mathbf{c}_i) + C d^2 2^{-\epsilon d/2} + |\mathcal{I}_r| e^{-\kappa \sqrt{m}/d} + |\mathcal{I}_s| e^{-\kappa \sqrt{m}/2} \\ &\leq \alpha \mu_0(\mathbf{c}) + C d^2 2^{-\epsilon d/2} + \exp(c_\epsilon d^7 - \kappa \sqrt{m}/d). \end{aligned}$$

Choose a fixed $\gamma > 0$ with $c_\epsilon \gamma^8 \leq \kappa/2$, and put $d = \lfloor \gamma m^{1/16} \rfloor$. For sufficiently large m , d satisfies the hypotheses of both lemmas, and

$$c_\epsilon d^7 \leq \frac{\kappa \sqrt{m}}{2d}, \quad C d^2 2^{-\epsilon d/2} + \exp\left(-\frac{\kappa \sqrt{m}}{2d}\right) \leq \exp(-\beta m^{1/16})$$

for a universal $\beta > 0$.

For any measurable $\mathbf{c}$, Claim 4.5 supplies finite-type $\mathbf{c}_r$ with $|\mu_j(\mathbf{c} - \mathbf{c}_r)| \leq 1/r$ for $j = 0, 1$. Applying the bound just proved to $\mathbf{c}_r$ gives

$$\mu_1(\mathbf{c}) \leq \mu_1(\mathbf{c}_r) + \frac{1}{r} \leq \alpha \mu_0(\mathbf{c}) + \frac{\alpha + 1}{r} + e^{-\beta m^{1/16}}.$$

Putting $r \rightarrow \infty$ completes the proof. $\square$

4.1.1 Corruption bound on regular cylinder intersections.

In this section, we prove Lemma 4.8, i.e., the corruption bound of the regular part. Let $\mathcal{X} = \mathcal{Y} = \text{SU}(m)^4$ and $\mathcal{Z} = \text{SU}(m)^2$. Recall that $\mathbf{f} : \mathcal{Y} \times \mathcal{Z} \rightarrow \{0, 1\}$, $\mathbf{g} : \mathcal{X} \times \mathcal{Z} \rightarrow \{0, 1\}$ and $\mathbf{h} : \mathcal{X} \times \mathcal{Y} \rightarrow \{0, 1\}$; ν is the Haar measure of $\text{SU}(m)$; by Claim 4.1, for both μ_0 and μ_1 we have

$$(X, Y) \sim \nu^{\otimes 4} \times \nu^{\otimes 4}, \quad (X, Z) \sim \nu^{\otimes 4} \times \nu^{\otimes 2}, \quad (Y, Z) \sim \nu^{\otimes 4} \times \nu^{\otimes 2}.$$

We first bound $\mu_1(\mathbf{c}_i)$ by the grid norms of $\mathbf{f}_i, \mathbf{g}_i$; this bound will also handle the sparse terms. For regular terms, a moment bound then gives a matching lower bound on $\mu_0(\mathbf{c}_i)$.

At first, we derive the following grid-norm bound from Arunachalam, Girish, and Lifshitz [AGL23, Lemma 5.2] using a moment argument and Cauchy–Schwarz, which might be of independent interest.

Theorem 4.10 (Grid norm bound). *There exist universal constants $\lambda > 1$ and $\kappa > 0$ such that, for any $\mathbf{f} : \mathcal{Y} \times \mathcal{Z} \rightarrow \{0, 1\}$, $\mathbf{g} : \mathcal{X} \times \mathcal{Z} \rightarrow \{0, 1\}$ and integer $d \geq 2$, we have*

$$\|\mathbb{E}_{\mu_1}[\mathbf{f}(Y, Z) \cdot \mathbf{g}(X, Z) \mid X, Y]\|_d \leq \lambda^{1+2/d} \|\mathbf{f}\|_{U(2,d)} \|\mathbf{g}\|_{U(2,d)} + \exp(-\kappa \sqrt{m}/d).$$

The proof of Theorem 4.10 is in Appendix B.1. We next apply this bound to each $\mathbf{c}_i = \mathbf{f}_i \mathbf{g}_i \mathbf{h}_i$. For an integer $p \geq 2$, define the local grid norms using normalized Haar measure:

$$\|\mathbf{f}_i\|_{U(2,p),i} := \left(\mathbb{E}_{Y, Y' \sim \mathcal{Y}_i} \left[\left(\mathbb{E}_{Z \sim \mathcal{Z}_i} [\mathbf{f}_i(Y, Z) \mathbf{f}_i(Y', Z)] \right)^p \right] \right)^{1/(2p)},$$

$$\|\mathbf{g}_i\|_{U(2,p),i} := \left(\mathbb{E}_{X, X' \sim \mathcal{X}_i} \left[\left(\mathbb{E}_{Z \sim \mathcal{Z}_i} [\mathbf{g}_i(X, Z) \mathbf{g}_i(X', Z)] \right)^p \right] \right)^{1/(2p)}.$$

The following corollary follows from Theorem 4.10 by Hölder’s inequality and normalization on $\mathcal{B}_i$.

Corollary 4.11 (Grid norm bound on $\mathcal{B}_i$). *With the constants $\lambda > 1$ and $\kappa > 0$ from Theorem 4.10, for every $i \in [k]$ and integer $p \geq 2$,*

$$\mu_1(\mathbf{c}_i) \leq \lambda^{1+2/p} \mu_0(\mathcal{B}_i) \delta_{\mathbf{h},i}^{1-1/p} \|\mathbf{f}_i\|_{U(2,p),i} \|\mathbf{g}_i\|_{U(2,p),i} + e^{-\kappa\sqrt{m}/p}.$$

Proof of Corollary 4.11. By Claim 4.1, (X, Y) has product Haar measure under μ_1 . Conditioning on (X, Y) and applying Hölder's inequality gives

$$\mu_1(\mathbf{c}_i) = \mathbb{E}_{X,Y}[\mathbf{h}_i \mathbb{E}_{\mu_1}[\mathbf{f}_i \mathbf{g}_i \mid X, Y]] \leq \|\mathbf{h}_i\|_{p/(p-1)} \|\mathbb{E}_{\mu_1}[\mathbf{f}_i \mathbf{g}_i \mid X, Y]\|_p.$$

Theorem 4.10 applies to the Boolean functions $\mathbf{f}_i, \mathbf{g}_i$ with moment p ; using $\|\mathbf{h}_i\|_{p/(p-1)} \leq 1$ gives

$$\mu_1(\mathbf{c}_i) \leq \lambda^{1+2/p} \|\mathbf{h}_i\|_{p/(p-1)} \|\mathbf{f}_i\|_{U(2,p)} \|\mathbf{g}_i\|_{U(2,p)} + e^{-\kappa\sqrt{m}/p}.$$

All norms here use Haar measure on the full input spaces. The support restrictions give

$$\begin{aligned} \|\mathbf{f}_i\|_{U(2,p)} &= \nu^{\otimes 4}(\mathcal{Y}_i)^{1/p} \nu^{\otimes 2}(\mathcal{Z}_i)^{1/2} \|\mathbf{f}_i\|_{U(2,p),i}, \\ \|\mathbf{g}_i\|_{U(2,p)} &= \nu^{\otimes 4}(\mathcal{X}_i)^{1/p} \nu^{\otimes 2}(\mathcal{Z}_i)^{1/2} \|\mathbf{g}_i\|_{U(2,p),i}, \\ \|\mathbf{h}_i\|_{p/(p-1)} &= [\nu^{\otimes 4}(\mathcal{X}_i) \nu^{\otimes 4}(\mathcal{Y}_i) \delta_{\mathbf{h},i}]^{1-1/p}. \end{aligned}$$

Multiplying these identities yields

$$\|\mathbf{h}_i\|_{p/(p-1)} \|\mathbf{f}_i\|_{U(2,p)} \|\mathbf{g}_i\|_{U(2,p)} = \mu_0(\mathcal{B}_i) \delta_{\mathbf{h},i}^{1-1/p} \|\mathbf{f}_i\|_{U(2,p),i} \|\mathbf{g}_i\|_{U(2,p),i},$$

which proves the bound. $\square$

The next theorem bounds the relative deviation of $\mathbb{E}_{Z \sim \mathcal{Z}_i}[\mathbf{f}_i(Y, Z) \mathbf{g}_i(X, Z)]$ from $\delta_{\mathbf{f},i} \delta_{\mathbf{g},i}$, using normalized Haar measure on $\mathcal{X}_i \times \mathcal{Y}_i$. The functions supplied by Theorem 4.7 are constant on finite equal-measure cells, so their normalized Haar averages equal the arithmetic means of their constant values on the equal-measure cells.

Theorem 4.12 ([KLM24, Lemma 4.8]). *Fix $i \in [k]$, $0 < \epsilon < 1/20$, and an even integer $\ell \geq 2$ with $\lceil \ell/\epsilon \rceil \leq d$. If the nonzero functions $\mathbf{f}_i, \mathbf{g}_i$ are (ϵ, d) -regular and satisfy ϵ -min-degree on $\mathcal{Y}_i \times \mathcal{Z}_i$ and $\mathcal{X}_i \times \mathcal{Z}_i$, respectively, then*

$$\left\| \frac{\mathbb{E}_{Z \sim \mathcal{Z}_i}[\mathbf{f}_i(Y, Z) \mathbf{g}_i(X, Z)]}{\delta_{\mathbf{f},i} \delta_{\mathbf{g},i}} - 1 \right\|_{\ell} \leq 20\epsilon.$$

The norm uses normalized Haar measure on $\mathcal{X}_i \times \mathcal{Y}_i$.

Remark 4.13. In our notation, Lemma 4.8 of Kelley, Lovett, and Meka [KLM24] requires (ϵ, p) -regularity, where $p = \lceil \ell/\epsilon \rceil$. Since $p \leq d$, Hölder's inequality and the assumed (ϵ, d) -regularity give

$$\|\mathbf{f}_i\|_{U(2,p),i} \leq \|\mathbf{f}_i\|_{U(2,d),i} \leq (1 + \epsilon) \delta_{\mathbf{f},i}, \quad \|\mathbf{g}_i\|_{U(2,p),i} \leq \|\mathbf{g}_i\|_{U(2,d),i} \leq (1 + \epsilon) \delta_{\mathbf{g},i},$$

with normalized Haar measure on $\mathcal{Y}_i \times \mathcal{Z}_i$ and $\mathcal{X}_i \times \mathcal{Z}_i$, respectively. Thus both functions satisfy the required (ϵ, p) -regularity.

Now we are ready to prove Lemma 4.8.

Proof of Lemma 4.8. Fix $0 < \epsilon \leq 10^{-4}$, $d \geq 8/\epsilon$, and $i \in \mathcal{I}_r$. We build a lower bound on $\mu_0(\mathbf{c}_i)$ and an upper bound on $\mu_1(\mathbf{c}_i)$, then compare them.

Step 1. Lower bound $\mu_0(\mathbf{c}_i)$. Since $\delta_{\mathbf{f},i}, \delta_{\mathbf{g},i} > 2^{-d}$, Theorem 4.7 gives regularity and min-degree on the respective subsets:

$$\begin{aligned} \|\mathbf{f}_i\|_{U(2,d),i} &\leq (1+\epsilon)\delta_{\mathbf{f},i}, & \|\mathbf{g}_i\|_{U(2,d),i} &\leq (1+\epsilon)\delta_{\mathbf{g},i}, \\ \mathbb{E}_{Z \sim \mathcal{Z}_i}[\mathbf{f}_i(Y, Z)] &\geq (1-\epsilon)\delta_{\mathbf{f},i} \quad (\forall Y \in \mathcal{Y}_i), & \mathbb{E}_{Z \sim \mathcal{Z}_i}[\mathbf{g}_i(X, Z)] &\geq (1-\epsilon)\delta_{\mathbf{g},i} \quad (\forall X \in \mathcal{X}_i). \end{aligned}$$

Take $\ell = 2\lfloor \epsilon d/2 \rfloor$. Then $\ell \geq \epsilon d/2 \geq 2$ and $\lceil \ell/\epsilon \rceil \leq d$, so Theorem 4.12 gives

$$\left\| \frac{\mathbb{E}_{Z \sim \mathcal{Z}_i}[\mathbf{f}_i(Y, Z)\mathbf{g}_i(X, Z)]}{\delta_{\mathbf{f},i}\delta_{\mathbf{g},i}} - 1 \right\|_{\ell} \leq 20\epsilon.$$

Hölder's inequality and $\|\mathbf{h}_i\|_{\ell/(\ell-1)} = \delta_{\mathbf{h},i}^{1-1/\ell}$, with normalized Haar measure on $\mathcal{X}_i \times \mathcal{Y}_i$, therefore imply

$$|\mathbb{E}_{X \sim \mathcal{X}_i, Y \sim \mathcal{Y}_i}[\mathbf{h}_i(X, Y)(\mathbb{E}_{Z \sim \mathcal{Z}_i}[\mathbf{f}_i(Y, Z)\mathbf{g}_i(X, Z)] - \delta_{\mathbf{f},i}\delta_{\mathbf{g},i})]| \leq 20\epsilon\delta_{\mathbf{f},i}\delta_{\mathbf{g},i}\delta_{\mathbf{h},i}^{1-1/\ell}.$$

Since $\delta_{\mathbf{h},i} > 2^{-\epsilon d}$, we have $\delta_{\mathbf{h},i}^{-1/\ell} \leq 4$. Hence

$$\begin{aligned} \mu_0(\mathbf{c}_i) &= \mu_0(\mathcal{B}_i)\mathbb{E}_{X \sim \mathcal{X}_i, Y \sim \mathcal{Y}_i}[\mathbf{h}_i(X, Y)\mathbb{E}_{Z \sim \mathcal{Z}_i}[\mathbf{f}_i(Y, Z)\mathbf{g}_i(X, Z)]] \\ &\geq \mu_0(\mathcal{B}_i)\delta_{\mathbf{f},i}\delta_{\mathbf{g},i}\delta_{\mathbf{h},i}(1 - 20\epsilon\delta_{\mathbf{h},i}^{-1/\ell}) \\ &\geq (1 - 80\epsilon)\mu_0(\mathcal{B}_i)\delta_{\mathbf{f},i}\delta_{\mathbf{g},i}\delta_{\mathbf{h},i}. \end{aligned} \tag{4.2}$$

Step 2. Upper bound $\mu_1(\mathbf{c}_i)$. Corollary 4.11 bounds $\mu_1(\mathbf{c}_i)$ by the local grid norms. Taking $p = d$ and substituting the regularity bounds in Step 1 gives

$$\mu_1(\mathbf{c}_i) \leq \lambda^{1+2/d}(1+\epsilon)^2\mu_0(\mathcal{B}_i)\delta_{\mathbf{f},i}\delta_{\mathbf{g},i}\delta_{\mathbf{h},i}^{1-1/d} + e^{-\kappa\sqrt{m}/d}. \tag{4.3}$$

Step 3. Compare the bounds. Since $\delta_{\mathbf{h},i}^{-1/d} < 2^\epsilon$, the lower bound (4.2) and upper bound (4.3) imply

$$\mu_1(\mathbf{c}_i) \leq \frac{\lambda^{1+2/d}(1+\epsilon)^2 2^\epsilon}{1-80\epsilon}\mu_0(\mathbf{c}_i) + e^{-\kappa\sqrt{m}/d} \leq 4\lambda^2\mu_0(\mathbf{c}_i) + e^{-\kappa\sqrt{m}/d}.$$

Thus $\alpha = 4\lambda^2$ suffices. $\square$

4.1.2 Corruption bound on sparse cylinder intersections.

In this section, we prove Lemma 4.9. We use Corollary 4.11 with $p = 2$: for Boolean functions, the grid norm $U(2, 2)$ is at most the square root of the density.

Proof of Lemma 4.9. For any Boolean function $\mathbf{u} : \mathcal{S} \times \mathcal{T} \rightarrow \{0, 1\}$ on probability spaces $\mathcal{S}, \mathcal{T}$, take independent samples $s, s' \sim \mathcal{S}$ and $t, t' \sim \mathcal{T}$. Since $0 \leq \mathbf{u} \leq 1$,

$$\|\mathbf{u}\|_{U(2,2)}^4 = \mathbb{E}[\mathbf{u}(s, t)\mathbf{u}(s', t)\mathbf{u}(s, t')\mathbf{u}(s', t')] \leq \mathbb{E}[\mathbf{u}(s, t)\mathbf{u}(s', t')] = (\mathbb{E}\mathbf{u})^2.$$

Applying this to $\mathbf{f}_i, \mathbf{g}_i$ on their respective subsets gives

$$\|\mathbf{f}_i\|_{U(2,2),i} \leq \delta_{\mathbf{f},i}^{1/2}, \quad \|\mathbf{g}_i\|_{U(2,2),i} \leq \delta_{\mathbf{g},i}^{1/2}.$$

For $i \in \mathcal{I}_s$, at least one of $\delta_{\mathbf{f},i}, \delta_{\mathbf{g},i}, \delta_{\mathbf{h},i}$ is at most $2^{-\epsilon d}$, since $0 < \epsilon \leq 1$. Corollary 4.11 with $p = 2$ therefore yields

$$\mu_1(\mathbf{c}_i) \leq \lambda^2\mu_0(\mathcal{B}_i)\sqrt{\delta_{\mathbf{f},i}\delta_{\mathbf{g},i}\delta_{\mathbf{h},i}} + e^{-\kappa\sqrt{m}/2} \leq \lambda^2\mu_0(\mathcal{B}_i)2^{-\epsilon d/2} + e^{-\kappa\sqrt{m}/2}.$$

Summing and using $\sum_i \mu_0(\mathcal{B}_i) = O(d^2)$ from Theorem 4.7 gives a universal constant $C > 0$ such that

$$\sum_{i \in \mathcal{I}_s} \mu_1(c_i) \leq \lambda^2 2^{-\epsilon d/2} \sum_{i \in \mathcal{I}_s} \mu_0(\mathcal{B}_i) + |\mathcal{I}_s| e^{-\kappa \sqrt{m}/2} \leq C d^2 2^{-\epsilon d/2} + |\mathcal{I}_s| e^{-\kappa \sqrt{m}/2}. \quad \square$$

AI Declaration

AI are used extensively for this work. Below, we list our AI usages following the AI Methodology template recommended by the [Simons Institute's AI + TCS Working Group](#).

Activity	AI used
Asking the research question	No
Coming up with the approach	Yes
Proof development	Yes
Writing and exposition	Yes
Checking for bugs	Yes
Other supporting tasks	Yes

Details.

1. **Asking the research question.** The authors formulated the research question and the goal of establishing a $\text{poly}(\log n)$ versus $\text{poly}(n)$ quantum-classical separation in the standard three-party NOF model.
2. **Coming up with the approach.** During discussions with the authors on how to use their prior work [YZ25a, YZ26, WW26] to construct a three-party function along the lifting route, GPT-5.6 Pro suggested using the ABCD problem of Arunachalam, Girish, and Lifshitz [AGL23] as the building block; following subsequent trials and manual screening, we formalized the Interleaved Unitary Product problem (IUP) (Definition 1.1).
3. **Proof development.** Focusing on the IUP problem, the authors proposed applying the structure-versus-pseudorandomness framework to the analysis of cylinder intersections. Following this suggestion, GPT-5.6 Pro identified the regularity decomposition of Abboud, Fischer, Kelley, Lovett, and Meka [AFK⁺24] and adapted it to the key regularity decomposition theorem (Theorem 4.7). GPT-5.6 Pro developed the proof of the grid-norm bound (Theorem 4.10). GPT-5.6 Pro also developed the regular and sparse estimates (Lemmas 4.8 and 4.9) and the resulting corruption bound (Lemma 4.3).
4. **Writing and exposition.** The authors structured the manuscript and wrote a first draft. AI tools (Codex GPT-6, Claude Opus 5.5) were subsequently used to polish the language and presentation.
5. **Checking for bugs.** AI helped check the proof and did not identify consequential errors in the proofs developed along the successful route; it later helped simplify these proofs.
6. **Other supporting tasks.** GPT-5.6 Pro also assisted with literature searches and language polishing.

The authors reviewed and finalized all mathematical claims and proofs, and take full responsibility for the correctness, exposition, and attribution in the final manuscript.

References

- [ABK24] Scott Aaronson, Harry Buhrman, and William Kretschmer. A qubit, a coin, and an advice string walk into a relational problem. In Venkatesan Guruswami, editor, *Proceedings of the 15th Innovations in Theoretical Computer Science Conference*, volume 287 of *LIPICs*, pages 1:1–1:24. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2024.
- [AFK⁺24] Amir Abboud, Nick Fischer, Zander Kelley, Shachar Lovett, and Raghu Meka. New graph decompositions and combinatorial Boolean matrix multiplication algorithms. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 935–943, 2024. Full version: arXiv:2308.12451.
- [AGL23] Srinivasan Arunachalam, Uma Girish, and Noam Lifshitz. One clean qubit suffices for quantum communication advantage. *arXiv preprint arXiv:2310.02406*, 2023.
- [BCW98] Harry Buhrman, Richard Cleve, and Avi Wigderson. Quantum vs. classical communication and computation. In *Proceedings of the thirtieth annual ACM symposium on Theory of computing*, pages 63–68, 1998.
- [BNS92] László Babai, Noam Nisan, and Mária Szegedy. Multiparty protocols, pseudorandom generators for logspace, and time-space trade-offs. *Journal of Computer and System Sciences*, 45(2):204–232, 1992.
- [BT94] Richard Beigel and Jun Tarui. On acc. *computational complexity*, 4(4):350–366, 1994.
- [BYJK04] Ziv Bar-Yossef, Thathachar S Jayram, and Iordanis Kerenidis. Exponential separation of quantum and classical one-way communication complexity. In *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*, pages 128–137, 2004.
- [Car10] Eric A. Carlen. Notes on divisibility of non atomic measures for Math 501, Fall 2010. Lecture notes, Rutgers University, October 2010.
- [CFL83] Ashok K Chandra, Merrick L Furst, and Richard J Lipton. Multi-party protocols. In *Proceedings of the fifteenth annual ACM symposium on Theory of computing*, pages 94–99, 1983.
- [CGK19] Henry Corrigan-Gibbs and Dmitry Kogan. The function-inversion problem: Barriers and opportunities. In *Theory of Cryptography Conference*, pages 393–421. Springer, 2019.
- [Cha07] Amit Chakrabarti. Lower bounds for multi-player pointer jumping. In *Twenty-Second Annual IEEE Conference on Computational Complexity (CCC’07)*, pages 33–45. IEEE, 2007.
- [DKO14] Andrew Drucker, Fabian Kuhn, and Rotem Oshman. On the power of the congested clique model. In *Proceedings of the 2014 ACM symposium on Principles of distributed computing*, pages 367–376, 2014.
- [Gav16] Dmitry Gavinsky. Entangled simultaneity versus classical interactivity in communication complexity. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 877–884, 2016.

- [Gav19] Dmitry Gavinsky. Quantum versus classical simultaneity in communication complexity. *IEEE Transactions on Information Theory*, 65(10):6466–6483, 2019.
- [Gav20] Dmitry Gavinsky. Bare quantum simultaneity versus classical interactivity in communication complexity. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 401–411, 2020.
- [Gav26] Dmytro Gavinsky. On the quantum communication complexity of total functions. *arXiv preprint arXiv:2608.18784*, 2026.
- [GGJL24] Mika Göös, Tom Gur, Siddhartha Jain, and Jiawei Li. Quantum communication advantage in TFNP. *arXiv preprint arXiv:2411.03296*, 2024.
- [GKK⁺07] Dmitry Gavinsky, Julia Kempe, Iordanis Kerenidis, Ran Raz, and Ronald De Wolf. Exponential separations for one-way quantum communication complexity, with applications to cryptography. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 516–525, 2007.
- [GP08] Dmitry Gavinsky and Pavel Pudlák. Exponential separation of quantum and classical non-interactive multi-party communication complexity. In *2008 23rd Annual IEEE Conference on Computational Complexity*, pages 332–339, 2008.
- [GR14] Anat Ganor and Ran Raz. Space pseudorandom generators by communication complexity lower bounds. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2014)*, pages 692–703. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2014.
- [GRT22] Uma Girish, Ran Raz, and Avishay Tal. Quantum versus randomized communication complexity, with efficient players. *computational complexity*, 31(2):17, 2022.
- [GXY⁺16] Jian-Yu Guan, Feihu Xu, Hua-Lei Yin, Yuan Li, Wei-Jun Zhang, Si-Jing Chen, Xiao-Yan Yang, Li Li, Li-Xing You, Teng-Yun Chen, Zhen Wang, Qiang Zhang, and Jian-Wei Pan. Observation of quantum fingerprinting beating the classical limit. *Physical Review Letters*, 116(24):240502, 2016.
- [HG90] J. Hastad and M. Goldmann. On the power of small-depth threshold circuits. In *Proceedings [1990] 31st Annual Symposium on Foundations of Computer Science*, pages 610–618 vol.2, 1990.
- [JLL⁺25] Michael Jaber, Yang P. Liu, Shachar Lovett, Anthony Ostuni, and Mehtaab Sawhney. Quasipolynomial bounds for the corners theorem, 2025.
- [KGD⁺25] William Kretschmer, Sabee Grewal, Matthew DeCross, Justin A. Gerber, Kevin Gilmore, Dan Gresh, Nicholas Hunter-Jones, Karl Mayer, Brian Neyenhuis, David Hayes, and Scott Aaronson. Demonstrating an unconditional separation between quantum and classical information resources. *arXiv:2509.07255*, 2025.
- [KKD19] Niraj Kumar, Iordanis Kerenidis, and Eleni Diamanti. Experimental demonstration of quantum advantage for one-way communication complexity surpassing best-known classical protocol. *Nature Communications*, 10:4152, 2019.
- [KL98] Emanuel Knill and Raymond Laflamme. Power of one bit of quantum information. *Physical Review Letters*, 81(25):5672–5675, 1998.

- [KL18] Hartmut Klauck and Debbie Lim. The power of one clean qubit in communication complexity. *arXiv preprint arXiv:1807.07762*, 2018.
- [KLM24] Zander Kelley, Shachar Lovett, and Raghu Meka. Explicit separations between randomized and deterministic number-on-forehead communication. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 1299–1310, 2024.
- [KM23] Zander Kelley and Raghu Meka. Strong bounds for 3-progressions. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 933–973, 2023.
- [KMPV19] John Kallaugher, Andrew McGregor, Eric Price, and Sofya Vorotnikova. The complexity of counting cycles in the adjacency list streaming model. In *Proceedings of the 38th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems*, pages 119–133, 2019.
- [KN97] Eyal Kushilevitz and Noam Nisan. *Communication Complexity*. Cambridge University Press, 1997.
- [Kna16] Anthony W. Knaapp. *Advanced Real Analysis*. Published by the author, digital second edition, 2016.
- [LMQW20] Xin Li, Fermi Ma, Willy Quach, and Daniel Wichs. Leakage-resilient key exchange and two-seed extractors. In *Annual International Cryptology Conference*, pages 401–429. Springer, 2020.
- [LS⁺17] Nati Linial, Adi Shraibman, et al. On the communication complexity of high-dimensional permutations. *arXiv preprint arXiv:1706.02207*, 2017.
- [MV13] Eric Miles and Emanuele Viola. Shielding circuits with groups. In *Proceedings of the forty-fifth annual ACM symposium on Theory of Computing*, pages 251–260, 2013.
- [NC10] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 10th anniversary edition, 2010.
- [Raz99] 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.
- [RK11] Oded Regev and Bo’az Klartag. Quantum one-way communication can be exponentially stronger than classical communication. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 31–40, 2011.
- [SLW⁺25] Ao Shen, Yu-Shuo Lu, Xiping Wu, Jinping Lin, Xiao-Yu Cao, Chengfang Ge, Shan-Feng Shao, Hua-Lei Yin, Lai Zhou, and Zhiliang Yuan. Experimental quantum fingerprinting without the shared randomness loophole. *Physical Review Letters*, 135(1):010801, 2025.
- [TSB⁺05] Pavel Trojek, Christian Schmid, Mohamed Bourennane, Časlav Brukner, Marek Żukowski, and Harald Weinfurter. Experimental quantum communication complexity. *Physical Review A*, 72(5):050305, 2005.
- [VW07] E. Viola and A. Wigderson. One-way multi-party communication lower bound for pointer jumping with applications. In *2007 48th Annual IEEE Symposium on Foundations of Computer Science*, pages 427–437, Los Alamitos, CA, USA, oct 2007. IEEE Computer Society.

- [WW26] Haoyu Wang and Pei Wu. Randomized and quantum lifting for one-way conservative NOF model. In *51st International Symposium on Mathematical Foundations of Computer Science (MFCS 2026)*, pages 78:1–78:17. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2026.
- [XAW⁺15] Feihu Xu, Juan Miguel Arrazola, Kejin Wei, Wenyan Wang, Pablo Palacios-Avila, Chen Feng, Shihan Sajeed, Norbert Lütkenhaus, and Hoi-Kwong Lo. Experimental quantum fingerprinting with weak coherent pulses. *Nature Communications*, 6:8735, 2015.
- [Yao79] 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. ACM, 1979.
- [YZ25a] Guangxu Yang and Jiapeng Zhang. Deterministic lifting theorems for one-way number-on-forehead communication. *arXiv preprint arXiv:2506.12420*, 2025.
- [YZ25b] Guangxu Yang and Jiapeng Zhang. Quantum versus classical separation in simultaneous number-on-forehead communication. *CoRR*, abs/2506.16804, 2025.
- [YZ26] Guangxu Yang and Jiapeng Zhang. Exponential separation of quantum and classical one-way Numbers-on-Forehead communication. In *Proceedings of the IEEE Symposium on Foundations of Computer Science (FOCS)*, 2026. To appear.

A Regularity decomposition for cylinder intersections

Abboud, Fischer, Kelley, Lovett, and Meka [AFK⁺24] give a regularity decomposition of two Boolean matrices.⁷ We check that their recursion preserves the product at every (X, Y, Z) ; multiplying by $\mathbf{h}(X, Y)$ then gives the required cylinder decomposition. We will restate all the cited theorems and algorithms below for completeness.

Throughout the argument, $\mathcal{I}, \mathcal{J}, \mathcal{K}$ are nonempty finite sets, $F : \mathcal{J} \times \mathcal{K} \rightarrow \{0, 1\}$ and $G : \mathcal{I} \times \mathcal{K} \rightarrow \{0, 1\}$. A matrix on a subdomain is extended by zero outside that subdomain; its density, grid norm, and min-degree use the uniform probability measure on its stated domain. We write $G[\mathcal{I}', \mathcal{K}']$ for a restriction. The notation (X, Y, Z, A, B) of Abboud, Fischer, Kelley, Lovett, and Meka [AFK⁺24] corresponds to $(\mathcal{J}, \mathcal{K}, \mathcal{I}, F, G^T)$ here.

The following GoodCube procedure selects $\mathcal{I}^* \times \mathcal{K}^*$ and splits F on $\mathcal{J} \times \mathcal{K}^*$.

Lemma A.1 (GoodCube, [AFK⁺24, Lemma 5.3]). *Let $0 < \epsilon < 1$, $0 < \eta < 1/2$, and let $d \geq 1$ be an integer. If $\mathbb{E}G \geq 2^{-d}$, there are nonempty $\mathcal{I}^* \subseteq \mathcal{I}$, $\mathcal{K}^* \subseteq \mathcal{K}$, and a list $(\mathcal{I}_i, \mathcal{J}_i, \mathcal{K}_i, F_i)_{i=1}^\ell$, where $\mathcal{I}_i \subseteq \mathcal{I}^*$, $\mathcal{J}_i \subseteq \mathcal{J}$, $\mathcal{K}_i \subseteq \mathcal{K}^*$, and $F_i : \mathcal{J}_i \times \mathcal{K}_i \rightarrow \{0, 1\}$, with the following properties. Let $G_i = G[\mathcal{I}_i, \mathcal{K}_i]$.*

1. $F(y, z)\mathbf{1}_{\mathcal{K}^*}(z) = \sum_{i=1}^\ell F_i(y, z)$ for every $(y, z) \in \mathcal{J} \times \mathcal{K}$.
2. For each i , either $\mathbb{E}F_i \leq 2^{-d}$, or both F_i, G_i are (ϵ, d) -regular and satisfy ϵ -min-degree.
3. $\mathbb{E}G[\mathcal{I}^*, \mathcal{K}^*] \geq \mathbb{E}G$, and

$$\sum_i |\mathcal{I}_i| |\mathcal{J}_i| |\mathcal{K}_i| \leq (d+2) |\mathcal{I}^*| |\mathcal{J}| |\mathcal{K}^*|,$$

⁷All theorem and algorithm numbers for Abboud et al. [AFK⁺24] refer to the full version.

$$\sum_i |\mathcal{I}^* \setminus \mathcal{I}_i| |\mathcal{J}_i| |\mathcal{K}_i| \leq \eta(d+2) |\mathcal{I}^*| |\mathcal{J}| |\mathcal{K}^*|.$$

4. For a constant $c_\epsilon > 0$ depending only on ϵ ,

$$|\mathcal{I}^*| |\mathcal{K}^*| \geq e^{-c_\epsilon d^4 / \eta} |\mathcal{I}| |\mathcal{K}|, \quad \ell \leq e^{c_\epsilon d^3}.$$

The output is computed by a deterministic algorithm, denoted by **GoodCube**.

The recursion below returns the pairs from **GoodCube**, then recurses on $\mathcal{I}^* \setminus \mathcal{I}_i$ and on $G - G[\mathcal{I}^*, \mathcal{K}^*]$. It follows Algorithm 5.5 of Abboud et al. [AFK⁺24], with the parameters and terminal step specified below.

AB-decomposition. The AB-decomposition procedure is as follows.

Input: $\mathcal{I}, \mathcal{J}, \mathcal{K}, F, G$, parameters $0 < \epsilon < 1$, integer $d \geq 1$, and a counter $r \in \{0, \dots, d\}$. The first call uses $r = 0$; a call with an empty domain returns the empty list.

Output: A list of tuples $(\mathcal{I}_i, \mathcal{J}_i, \mathcal{K}_i, F_i, G_i)$.

1. If $\mathbb{E}G \leq 2^{-d}$, return $\{(\mathcal{I}, \mathcal{J}, \mathcal{K}, F, G)\}$.
2. If $r = d$, return the following list and stop.
 - (a) If $|\mathcal{I}| |\mathcal{K}| \geq 2^d$, partition the 1 entries of G into groups of size $\lfloor 2^{-d} |\mathcal{I}| |\mathcal{K}| \rfloor$, except possibly the last group. For each group's indicator $G^{(j)}$, return $(\mathcal{I}, \mathcal{J}, \mathcal{K}, F, G^{(j)})$.
 - (b) Otherwise, for each $G(x, z) = 1$, put $\mathcal{J}_{x,z} = \{y : F(y, z) = 1\}$. If this set is nonempty, return

$$(\{x\}, \mathcal{J}_{x,z}, \{z\}, F[\mathcal{J}_{x,z}, \{z\}], G[\{x\}, \{z\}]).$$

3. Set $\eta = 1/[4(d+2)^2]$. Since $\mathbb{E}G > 2^{-d}$, apply the GoodCube procedure of Lemma A.1 to $(\mathcal{I}, \mathcal{J}, \mathcal{K}, F, G; \epsilon, \eta, d)$. For its output, put

$$G_i = G[\mathcal{I}_i, \mathcal{K}_i], \quad G'_i = G[\mathcal{I}^* \setminus \mathcal{I}_i, \mathcal{K}_i], \quad G^* = G[\mathcal{I}^*, \mathcal{K}^*].$$

4. Make the recursive calls

$$\mathcal{O}_i = \text{AB-decomposition}(\mathcal{I}^* \setminus \mathcal{I}_i, \mathcal{J}_i, \mathcal{K}_i, F_i, G'_i; \epsilon, d, r+1) \quad (i = 1, \dots, \ell),$$

$$\mathcal{O}^* = \text{AB-decomposition}(\mathcal{I}, \mathcal{J}, \mathcal{K}, F, G - G^*; \epsilon, d, r).$$

5. Return the list $(\mathcal{I}_i, \mathcal{J}_i, \mathcal{K}_i, F_i, G_i)_{i=1}^\ell$ followed by $\mathcal{O}_1, \dots, \mathcal{O}_\ell, \mathcal{O}^*$.

Step 2(a) returns at most 2^{d+1} tuples, each with $\mathbb{E}G^{(j)} \leq 2^{-d}$; both output matrices in step 2(b) are identically 1. Fixing orders on the finite sets makes the choices deterministic.

The small omitted sets bound the total size of the returned domains, and the density removed at each call bounds the number of calls. The following is the bound of Abboud et al. [AFK⁺24, Theorem 3.1] with the stated terminal step.

Lemma A.2 (AB-decomposition bounds). *The first call of AB-decomposition returns k tuples satisfying*

$$\sum_{i=1}^k |\mathcal{I}_i| |\mathcal{J}_i| |\mathcal{K}_i| \leq 4(d+2)^2 |\mathcal{I}| |\mathcal{J}| |\mathcal{K}|, \quad k \leq \exp(c_\epsilon d^7).$$

Every output is sparse in at least one matrix, with density at most 2^{-d} , or both matrices are (ϵ, d) -regular and satisfy ϵ -min-degree.

Proof of Lemma A.2.

Step 1. Bound the total size of the returned domains. Write $V = |\mathcal{I}||\mathcal{J}||\mathcal{K}|$ and $\delta = \mathbb{E}G$ for the current call. The sum $\sum_i |\mathcal{I}_i||\mathcal{J}_i||\mathcal{K}_i|$ over its returned tuples is at most V if $\delta \leq 2^{-d}$, and otherwise at most

$$2^{r+2}(d+2)(d+2+\log_2 \delta)V. \quad (\text{A.1})$$

Induct lexicographically on $(d-r, |\{(x, z) : G(x, z) = 1\}|)$. The lists returned in steps 1 and 2 of AB-decomposition satisfy $\sum_i |\mathcal{I}_i||\mathcal{J}_i||\mathcal{K}_i| \leq 2^{d+1}V$. In a recursive call put $V^* = |\mathcal{I}^*||\mathcal{J}||\mathcal{K}^*|$. The GoodCube size bounds in Lemma A.1(3) are

$$\sum_i |\mathcal{I}_i||\mathcal{J}_i||\mathcal{K}_i| \leq (d+2)V^*, \quad \sum_i |\mathcal{I}^* \setminus \mathcal{I}_i||\mathcal{J}_i||\mathcal{K}_i| \leq \eta(d+2)V^*.$$

For each recursive call at $r+1$, the induction hypothesis bounds the total size of its returned domains by $2^{r+3}(d+2)^2$ times the product of the sizes of its three input sets. Hence the pairs supplied by GoodCube and the outputs in $\mathcal{O}_1, \dots, \mathcal{O}_\ell$ contribute at most

$$(d+2)V^* + 2^{r+3}(d+2)^2\eta(d+2)V^* \leq 2^{r+2}(d+2)V^*.$$

GoodCube also guarantees $\mathbb{E}G[\mathcal{I}^*, \mathcal{K}^*] \geq \delta$. Thus the remaining density satisfies

$$\delta^* := \mathbb{E}(G - G^*) = \delta - \frac{V^*}{V} \mathbb{E}G[\mathcal{I}^*, \mathcal{K}^*] \leq \delta(1 - V^*/V).$$

If $\delta^* > 2^{-d}$, including the outputs in $\mathcal{O}^*$ gives (A.1), since

$$\log_2 \delta^* \leq \log_2 \delta - V^*/V \implies V^* + (d+2+\log_2 \delta^*)V \leq (d+2+\log_2 \delta)V.$$

If $\delta^* \leq 2^{-d}$, the sum over all returned tuples is at most $2^{r+2}(d+2)V^* + V$, which is also at most (A.1) since $d+2+\log_2 \delta > 2$. Taking $r=0$ proves the claimed bound on this sum.

Step 2. Bound the number of outputs. The bounds on $|\mathcal{I}^*||\mathcal{K}^*|$ and ℓ in Lemma A.1(4), with $\eta = 1/[4(d+2)^2]$, give numbers $M = \exp(O_\epsilon(d^6)) \geq 1$ and $L = \exp(O_\epsilon(d^3)) \geq 1$ such that

$$\frac{V^*}{V} \geq \frac{1}{M}, \quad \ell \leq L, \quad \delta - \delta^* \geq \delta \frac{V^*}{V} > \frac{2^{-d}}{M}.$$

Let k_r be the maximum number of tuples returned by a call with counter r . Each chain of calls on $G - G^*$ has at most $2^d M$ nonterminal calls, so

$$k_d \leq 2^{d+1}, \quad k_r \leq 2^d M L (1 + k_{r+1}) + 2^{d+1} \quad (0 \leq r < d).$$

Induction on $d-r$ gives

$$k \leq k_0 \leq (2^{d+2} M (L+1))^{d+1} = \exp(O_\epsilon(d^7)).$$

In AB-decomposition, steps 1 and 2(a) return pairs with a matrix of density at most 2^{-d} , while step 2(b) returns two all-1 matrices. The pairs supplied by GoodCube are sparse or regular by Lemma A.1(2). The bounds on recursive calls also show that AB-decomposition terminates. $\square$

We now prove the regularity decomposition on cylinder intersections.

Proof of Theorem 4.7.

Step 1. Preserve $F(y, z)G(x, z)$ at every triple. GoodCube preserves F on $\mathcal{J} \times \mathcal{K}^*$ by Lemma A.1(1):

$$\sum_i F_i(y, z) = F(y, z)\mathbf{1}_{\mathcal{K}^*}(z), \quad F_i(y, z)(G_i(x, z) + G'_i(x, z)) = F_i(y, z)G(x, z)\mathbf{1}_{\mathcal{I}^*}(x).$$

The second identity follows from $\mathcal{I}^* = \mathcal{I}_i \sqcup (\mathcal{I}^* \setminus \mathcal{I}_i)$. Summing it and adding $F(G - G^*)$ gives

$$F(y, z)G(x, z) = \sum_i F_i(y, z)G_i(x, z) + \sum_i F_i(y, z)G'_i(x, z) + F(y, z)(G - G^*)(x, z). \quad (\text{A.2})$$

The last two terms are the recursive products in step 4 of AB-decomposition. In step 2(a) of AB-decomposition, the indicators $G^{(j)}$ sum to G ; in step 2(b), each triple (x, y, z) with $F(y, z)G(x, z) = 1$ appears only in the output indexed by (x, z) . Induction therefore gives

$$F(y, z)G(x, z) = \sum_{i=1}^k F_i(y, z)G_i(x, z). \quad (\text{A.3})$$

Step 2. Return functions on $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$. Let $\mathcal{I}, \mathcal{J}, \mathcal{K}$ index the equal-measure partitions defining the finite-type functions $\mathbf{f}, \mathbf{g}, \mathbf{h}$. Write x, y, z for the labels of the sets containing X, Y, Z , and define the Boolean matrices

$$F(y, z) = \mathbf{f}(Y, Z), \quad G(x, z) = \mathbf{g}(X, Z), \quad H(x, y) = \mathbf{h}(X, Y).$$

Apply AB-decomposition with the integer $q = d + 1$ in place of d and $\min\{\epsilon, 1/2\} \in (0, 1)$ in place of ϵ . For each returned tuple, let $\mathcal{X}_i, \mathcal{Y}_i, \mathcal{Z}_i$ be the unions of the sets indexed by $\mathcal{I}_i, \mathcal{J}_i, \mathcal{K}_i$, respectively, and set

$$\mathbf{f}_i(Y, Z) = F_i(y, z), \quad \mathbf{g}_i(X, Z) = G_i(x, z), \quad \mathcal{B}_i = \mathcal{X}_i \times \mathcal{Y}_i \times \mathcal{Z}_i.$$

Their supports have the required form. Multiplying (A.3) by $H(x, y)$ gives

$$\sum_i \mathbf{f}_i(Y, Z)\mathbf{g}_i(X, Z)\mathbf{h}(X, Y) = F(y, z)G(x, z)H(x, y) = \mathbf{c}(X, Y, Z).$$

All terms are Boolean, so they are also pairwise disjoint wherever they equal 1.

Equal cell measures identify the normalized integrals with uniform finite averages. Hence the densities, grid norms, and averages defining min-degree agree with those of F_i, G_i . Lemma A.2 classifies each output: either a density is at most $2^{-q} < 2^{-d}$, or both functions are regular and satisfy min-degree. In the latter case, monotonicity of probability-space L^p norms gives

$$\|\mathbf{f}_i\|_{U(2,d),i} \leq \|\mathbf{f}_i\|_{U(2,q),i} \leq (1 + \epsilon)\delta_{\mathbf{f},i}, \quad \|\mathbf{g}_i\|_{U(2,d),i} \leq \|\mathbf{g}_i\|_{U(2,q),i} \leq (1 + \epsilon)\delta_{\mathbf{g},i},$$

and increasing the min-degree parameter to ϵ only weakens its lower bound.

The same lemma bounds the finite total size and number of outputs, which translate to

$$\sum_i \mu_0(\mathcal{B}_i) = \frac{\sum_i |\mathcal{I}_i||\mathcal{J}_i||\mathcal{K}_i|}{|\mathcal{I}||\mathcal{J}||\mathcal{K}|} \leq 4(d+3)^2 = O(d^2), \quad k \leq \exp(c_\epsilon d^7),$$

after increasing c_ϵ . This proves all three claims. $\square$

B The grid norm bound

Arunachalam, Girish, and Lifshitz [AGL23] bound how much the constraint $ABCD = I_m$ can increase the average of $\mathbf{u}(A, C)\mathbf{w}(B, D)$. Applying this estimate to the two matrix products in $W = PE_1QE_2$ will give the grid norm bound.

Theorem B.1 ([AGL23, Lemma 5.2]). *There is a universal constant $\kappa > 0$ such that the following holds. Let A, B, C, D be independent Haar matrices in $\text{SU}(m)$. For measurable Boolean functions $\mathbf{u}, \mathbf{w} : \text{SU}(m)^2 \rightarrow \{0, 1\}$ with $\mathbb{E}[\mathbf{u}], \mathbb{E}[\mathbf{w}] \geq e^{-\kappa\sqrt{m}}$, we have*

$$\left| \mathbb{E}[\mathbf{u}(A, C)\mathbf{w}(B, (ABC)^{-1})] - \mathbb{E}[\mathbf{u}(A, C)\mathbf{w}(B, D)] \right| \leq \frac{\mathbb{E}[\mathbf{u}]\mathbb{E}[\mathbf{w}]}{30}.$$

Write $\mathcal{G} = \text{SU}(m)$ and $\mathcal{H} = \mathcal{G}^2$, with coordinatewise multiplication. Let $\kappa_0 > 0$ be the constant in Theorem B.1, and put $\lambda_0 = 31/30$ and $\lambda = 2\lambda_0$. In either group, conditioning on $ABCD = T$ means sampling A, B, C independently by Haar measure and setting $D = (ABC)^{-1}T$. Unconditioned averages below use independent Haar pairs.

The next lemma removes the density condition by bounding small averages directly, then applies the resulting estimate twice to handle both matrix products.

Lemma B.2 (Weighted product estimates). *1. For $T \in \mathcal{G}$ and measurable $\mathbf{a}, \mathbf{b} : \mathcal{G}^2 \rightarrow [0, 1]$,*

$$\mathbb{E}[\mathbf{a}(A, C)\mathbf{b}(B, D) \mid ABCD = T] \leq \lambda_0 \mathbb{E}[\mathbf{a}]\mathbb{E}[\mathbf{b}] + e^{-\kappa_0\sqrt{m}}. \quad (\text{B.1})$$

2. For $L \in \mathcal{H}$ and measurable $\mathbf{a}, \mathbf{b} : \mathcal{H}^2 \rightarrow [0, 1]$,

$$\mathbb{E}[\mathbf{a}(A, C)\mathbf{b}(B, D) \mid ABCD = L] \leq \lambda_0^2 \mathbb{E}[\mathbf{a}]\mathbb{E}[\mathbf{b}] + (1 + \lambda_0)e^{-\kappa_0\sqrt{m}}. \quad (\text{B.2})$$

Proof. For (1), first let $\mathbf{a}, \mathbf{b}$ be Boolean. If both averages are at least $e^{-\kappa_0\sqrt{m}}$, apply the relative-error bound in Theorem B.1 to $\mathbf{a}$ and $(B, D) \mapsto \mathbf{b}(B, DT)$. Haar invariance preserves $\mathbb{E}[\mathbf{b}]$, giving

$$\mathbb{E}[\mathbf{a}(A, C)\mathbf{b}(B, D) \mid ABCD = T] \leq \lambda_0 \mathbb{E}[\mathbf{a}]\mathbb{E}[\mathbf{b}].$$

Otherwise, both (A, C) and (B, D) still have product Haar marginals, so

$$\mathbb{E}[\mathbf{a}\mathbf{b} \mid ABCD = T] \leq \min\{\mathbb{E}[\mathbf{a}], \mathbb{E}[\mathbf{b}]\} < e^{-\kappa_0\sqrt{m}}.$$

For functions in $[0, 1]$, integrate the Boolean inequality using $\mathbf{a} = \int_0^1 \mathbf{1}[\mathbf{a} \geq s] ds$ and $\mathbf{b} = \int_0^1 \mathbf{1}[\mathbf{b} \geq t] dt$.

For (2), write $A = (A_1, A_2)$, and similarly for B, C, D, L . Fix the second coordinates and apply (1) to the first coordinates. The resulting functions $\mathbf{a}_2(A_2, C_2) = \mathbb{E}_{A_1, C_1}[\mathbf{a}(A, C)]$ and $\mathbf{b}_2(B_2, D_2) = \mathbb{E}_{B_1, D_1}[\mathbf{b}(B, D)]$ lie in $[0, 1]$. Averaging over the second coordinates and applying (1) again gives

$$\mathbb{E}[\mathbf{a}\mathbf{b} \mid ABCD = L] \leq \lambda_0 \mathbb{E}[\mathbf{a}_2\mathbf{b}_2 \mid A_2B_2C_2D_2 = L_2] + e^{-\kappa_0\sqrt{m}} \leq \lambda_0^2 \mathbb{E}[\mathbf{a}]\mathbb{E}[\mathbf{b}] + (1 + \lambda_0)e^{-\kappa_0\sqrt{m}}. \quad \square$$

We now prove Theorem 4.10 using Lemma B.2 in the following section.

B.1 Proof of Theorem 4.10

Sampling μ_1 through P, Q . For the proof below, write $A = (A_1, A_2)$ and similarly for B, C, D . Equivalently, sample μ_1 by first sampling independent Haar matrices P, Q and putting $L = (P, Q)$. Given L , sample E_1 by Haar measure and set $E_2 = (PE_1Q)^{-1}$; independently, sample A, B, C by

product Haar measure and set $D = (ABC)^{-1}L$. Thus (X, Y) and Z are independent conditional on L , with

$$ABCD = L, \quad PE_1QE_2 = I_m.$$

Unconditionally, $L \sim \nu^{\otimes 2}$ and $Z \sim \nu^{\otimes 2}$. Lemma B.2(1) therefore controls functions of (P, Q) and (E_1, E_2) , while part (2) controls functions of $X = (A_1, C_1, A_2, C_2)$ and $Y = (B_1, D_1, B_2, D_2)$ conditional on L .

We expand the d -th moment using conditional samples of Z . The product bounds above replace the dependent averages by independent ones, whose second moments give the grid norms. Fix $\mathbf{u} : \mathcal{G}^2 \rightarrow [0, 1]$.

Step 1. Bound $\mathbb{E}_L[(\mathbb{E}[\mathbf{u}(Z) \mid L])^d]$. Define $\mathbf{r}(L) = \mathbb{E}[\mathbf{u}(Z) \mid L]$ and $\mathcal{S} = \{L : \mathbf{r}(L) > 2\lambda_0\mathbb{E}[\mathbf{u}]\}$. Lemma B.2(1) controls functions of the alternating factors in $PE_1QE_2 = I_m$. Apply it with $(A, B, C, D) = (P, E_1, Q, E_2)$, $T = I_m$, and $(\mathbf{a}, \mathbf{b}) = (\mathbf{1}_{\mathcal{S}}, \mathbf{u})$. Both functions lie in $[0, 1]$, and the displayed product condition holds; hence

$$\mathbb{E}[\mathbf{r}\mathbf{1}_{\mathcal{S}}] \leq \lambda_0\mathbb{E}[\mathbf{u}] \Pr[L \in \mathcal{S}] + e^{-\kappa_0\sqrt{m}} \leq \frac{1}{2}\mathbb{E}[\mathbf{r}\mathbf{1}_{\mathcal{S}}] + e^{-\kappa_0\sqrt{m}}.$$

Since $0 \leq \mathbf{r} \leq 1$, for every integer $d \geq 1$ we obtain

$$\mathbb{E}_L[(\mathbb{E}[\mathbf{u}(Z) \mid L])^d] \leq \min\{1, (2\lambda_0\mathbb{E}[\mathbf{u}])^d\} + 2e^{-\kappa_0\sqrt{m}}. \quad (\text{B.3})$$

Step 2. Bound the moments of the products of $\mathbf{f}$ and $\mathbf{g}$. Given L , let $Z^{(1)}, \dots, Z^{(d)}$ be independent samples from the conditional distribution of Z ; write their tuple as $Z^{[d]}$. For fixed $L, Z^{[d]}$, Lemma B.2(2) replaces the average of a function of X times a function of Y by their independent averages. Apply it to $\prod_{j=1}^d \mathbf{g}(X, Z^{(j)})$ and $\prod_{j=1}^d \mathbf{f}(Y, Z^{(j)})$. These are Boolean functions of (A, C) and (B, D) , respectively. Given L , the samples $Z^{(1)}, \dots, Z^{(d)}$ are mutually independent and independent of (X, Y) . Thus

$$\begin{aligned} & \|\mathbb{E}_{\mu_1}[\mathbf{f}\mathbf{g} \mid X, Y]\|_d^d \\ &= \mathbb{E}_{L, Z^{[d]}} \left[\mathbb{E} \left[\prod_{j=1}^d \mathbf{f}(Y, Z^{(j)}) \mathbf{g}(X, Z^{(j)}) \mid L, Z^{[d]} \right] \right] \\ &\leq \lambda_0^2 \mathbb{E}_{L, Z^{[d]}} \left[\left(\mathbb{E}_Y \left[\prod_{j=1}^d \mathbf{f}(Y, Z^{(j)}) \right] \right) \left(\mathbb{E}_X \left[\prod_{j=1}^d \mathbf{g}(X, Z^{(j)}) \right] \right) \right] + (1 + \lambda_0)e^{-\kappa_0\sqrt{m}}. \end{aligned} \quad (\text{B.4})$$

Here the inner averages on the last line use $X \sim \nu^{\otimes 4}$ and $Y \sim \nu^{\otimes 4}$ independently. For independent $Y, Y' \sim \nu^{\otimes 4}$, apply the conditional-moment bound (B.3) to $\mathbf{u}(Z) = \mathbf{f}(Y, Z)\mathbf{f}(Y', Z)$ and average over Y, Y' :

$$\begin{aligned} & \mathbb{E}_{L, Z^{[d]}} \left[\left(\mathbb{E}_Y \left[\prod_{j=1}^d \mathbf{f}(Y, Z^{(j)}) \right] \right)^2 \right] = \mathbb{E}_{Y, Y', L} \left[(\mathbb{E}[\mathbf{f}(Y, Z)\mathbf{f}(Y', Z) \mid L, Y, Y'])^d \right] \\ & \leq \min\{1, (2\lambda_0)^d \|\mathbf{f}\|_{U(2, d)}^{2d}\} + 2e^{-\kappa_0\sqrt{m}}. \end{aligned} \quad (\text{B.5})$$

The last line uses $\mathbb{E}[\min\{1, V\}] \leq \min\{1, \mathbb{E}[V]\}$ for a nonnegative random variable V , and the definition of $\|\cdot\|_{U(2, d)}$; the same bound holds for $\mathbf{g}$.

Step 3. Bound $\|\mathbb{E}_{\mu_1}[\mathbf{fg} \mid X, Y]\|_d$. For $0 \leq a, b \leq 1$ and $q \geq 0$, $\sqrt{(a+q)(b+q)} \leq \sqrt{ab} + 2\sqrt{q} + q$. The expectation of the product in (B.4) is at most the geometric mean of the two second moments, by Cauchy–Schwarz. Substitute their bounds (B.5) and the preceding inequality with $q = 2e^{-\kappa_0\sqrt{m}}$. Since $2\sqrt{2}\lambda_0^2 + 2\lambda_0^2 + 1 + \lambda_0 < 8$, we get

$$\|\mathbb{E}_{\mu_1}[\mathbf{fg} \mid X, Y]\|_d^d \leq \lambda_0^2(2\lambda_0)^d (\|\mathbf{f}\|_{U(2,d)}\|\mathbf{g}\|_{U(2,d)})^d + 8e^{-\kappa_0\sqrt{m}/2}.$$

Recall $\lambda = 2\lambda_0$. Taking the d th root gives

$$\|\mathbb{E}_{\mu_1}[\mathbf{fg} \mid X, Y]\|_d \leq \lambda^{1+2/d} \|\mathbf{f}\|_{U(2,d)}\|\mathbf{g}\|_{U(2,d)} + (8e^{-\kappa_0\sqrt{m}/2})^{1/d}. \quad (\text{B.6})$$

If $\sqrt{m} \geq 4\log(8)/\kappa_0$, then

$$(8e^{-\kappa_0\sqrt{m}/2})^{1/d} \leq e^{-\kappa_0\sqrt{m}/(4d)}.$$

For the remaining dimensions, put $p = \|\mathbf{f}\|_{U(2,d)}\|\mathbf{g}\|_{U(2,d)}$. By Claim 4.1, $\mu_1(\mathbf{f}) = \mu_0(\mathbf{f})$ and $\mu_1(\mathbf{g}) = \mu_0(\mathbf{g})$. Since $\mathbf{fg} \in \{0, 1\}$ and each grid norm dominates the corresponding average, Jensen's inequality gives

$$\|\mathbb{E}_{\mu_1}[\mathbf{fg} \mid X, Y]\|_d^d \leq \min\{\mu_0(\mathbf{f}), \mu_0(\mathbf{g})\} \leq \sqrt{p}.$$

If $p \geq 1/\lambda$, the main term in the claimed bound is at least 1. Otherwise, the last display gives $\|\mathbb{E}_{\mu_1}[\mathbf{fg} \mid X, Y]\|_d \leq p^{1/(2d)} < \lambda^{-1/(2d)}$. Choose $0 < \kappa \leq \min\{\kappa_0/4, \kappa_0 \log(\lambda)/(8\log(8))\}$; then $\lambda^{-1/(2d)} \leq e^{-\kappa\sqrt{m}/d}$ in these remaining dimensions. This proves the stated bound for every $d \geq 2$.

C Additional Proofs

Proof of Claim 3.1. The singular values of U lie in $[\sqrt{1-\sigma}, \sqrt{1+\sigma}]$, since $\|U^\dagger U - I_m\|_{\text{op}} \leq \|U^\dagger U - I_m\|_{\text{F}} \leq \sigma < 1$. Thus U is invertible. Let $\mathcal{S}_U$ be the set of singular values of U ; then its polar factor satisfies

$$\tilde{U}^\dagger \tilde{U} = I_m, \quad \|\tilde{U} - U\|_{\text{op}} = \max_{s \in \mathcal{S}_U} \frac{|s^2 - 1|}{s + 1} \leq \sigma, \quad \|U\|_{\text{op}} \leq \sqrt{1 + \sigma}.$$

Write $U_1, \dots, U_{10}$ for the factors in W , in their product order. Replacing one factor at a time gives

$$\tilde{W} - W = \sum_{r=1}^{10} \tilde{U}_1 \cdots \tilde{U}_{r-1} (\tilde{U}_r - U_r) U_{r+1} \cdots U_{10}.$$

The preceding bounds and $|\text{Tr } M| \leq m\|M\|_{\text{op}}$ now give

$$\frac{|\text{Re Tr } \tilde{W} - \text{Re Tr } W|}{m} \leq \|\tilde{W} - W\|_{\text{op}} \leq \sum_{r=1}^{10} \sigma(1 + \sigma)^{(10-r)/2} \leq 10\sigma(1 + \sigma)^{9/2}. \quad \square$$

Proof of Claim 4.1. Let $U_1, \dots, U_{10}$ be the factors of W in order. Under μ_1 ,

$$(U_1, \dots, U_9) \sim \nu^{\otimes 9}, \quad U_{10} = (U_1 \cdots U_9)^{-1}.$$

Fix $j \in [9]$ and condition on $(U_r)_{r \in [9] \setminus \{j\}}$. Then U_j remains Haar, and

$$U_{10} = (U_{j+1} \cdots U_9)^{-1} U_j^{-1} (U_1 \cdots U_{j-1})^{-1}.$$

Both surrounding products are fixed. Thus Haar invariance gives, for every measurable $S \subseteq \text{SU}(m)$,

$$\Pr_{\mu_1}[U_{10} \in S \mid (U_r)_{r \in [9] \setminus \{j\}}] = \nu(S).$$

Together with the independent Haar distribution of the other eight matrices, this implies

$$(U_r)_{r \in [10] \setminus \{j\}} \sim \nu^{\otimes 9} \quad (j \in [10]).$$

The views (X, Y) , (X, Z) , and (Y, Z) contain 8, 6, and 6 distinct matrices, respectively, so each has the same product Haar distribution as under μ_0 . $\square$

Proof of Claim 4.2. Let $\tilde{U}_r$ denote the rounded version of a Haar matrix U ; the subscript r marks rounding, as opposed to the polar factor $\tilde{U}$. Put $\Delta = \tilde{U}_r - U$. Each real coordinate changes by at most 2^{-b-1} , so

$$\|\Delta\|_F \leq \frac{m2^{-b}}{\sqrt{2}} \leq \frac{m}{n}, \quad \|\tilde{U}_r^\dagger \tilde{U}_r - I_m\|_F \leq 2\|\Delta\|_F + \|\Delta\|_F^2 \leq \frac{2m}{n} + \frac{m^2}{n^2} < 10^{-4}$$

for all sufficiently large n , because $m = \Theta(\sqrt{n/\log n})$. Let $\tilde{W}_r$ be the product of the ten rounded factors. Expanding the product difference one factor at a time, as in Claim 3.1, gives

$$\frac{|\operatorname{Re} \operatorname{Tr} \tilde{W}_r - \operatorname{Re} \operatorname{Tr} W|}{m} \leq \|\tilde{W}_r - W\|_{\text{op}} \leq \frac{10m}{n} \left(1 + \frac{m}{n}\right)^9 < 10^{-3}.$$

Under μ_1 , $W = I_m$, so

$$\frac{\operatorname{Re} \operatorname{Tr} W}{m} = 1, \quad \frac{\operatorname{Re} \operatorname{Tr} \tilde{W}_r}{m} > 1 - 10^{-3} > 0.9, \quad \mathbf{F}_n(\mathbf{R}(X, Y, Z)) = 1.$$

Under μ_0 , W is Haar. Haar invariance gives

$$\mathbb{E}_{\mu_0}[W_{ii} \overline{W_{jj}}] = \frac{\mathbf{1}[i=j]}{m}, \quad \mathbb{E}_{\mu_0}[|\operatorname{Tr} W|^2] = 1.$$

Markov's inequality gives the continuous-input estimate

$$\Pr_{\mu_0}[\operatorname{Re} \operatorname{Tr} W > 0.1m] \leq \frac{\mathbb{E}_{\mu_0}[|\operatorname{Tr} W|^2]}{0.1^2 m^2} = \frac{100}{m^2} \leq \frac{1}{m}.$$

The rounded matrices satisfy $\|\tilde{U}_r^\dagger \tilde{U}_r - I_m\|_F \leq 10^{-4}$, and rounding changes the normalized trace by less than 10^{-3} . Hence

$$\Pr_{\mu_0}[\mathbf{F}_n \neq 0] \leq \Pr_{\mu_0}[\operatorname{Re} \operatorname{Tr} W > 0.099m] \leq \frac{\mathbb{E}_{\mu_0}[|\operatorname{Tr} W|^2]}{0.099^2 m^2} \leq \frac{1}{m}$$

for all sufficiently large m . The event $\mathbf{F}_n \neq 0$ includes inputs outside the promise, so this proves the stated probability of $\mathbf{F}_n = 0$. $\square$

Proof of Claim 4.5. All pairwise expectations below use independent Haar inputs.

Step 1. Construct Boolean approximations on finite partitions. Let $\mathcal{E} = \{\mathbf{f} = 1\}$. By compact inner and open outer approximation for Haar measure (see Knapp [Kna16, Section VI.2]), there are a compact set $\mathcal{K}$ and an open set $\mathcal{O}$ with

$$\mathcal{K} \subseteq \mathcal{E} \subseteq \mathcal{O}, \quad (\nu^{\otimes 4} \otimes \nu^{\otimes 2})(\mathcal{O} \setminus \mathcal{K}) < \frac{1}{6r}.$$

Every point of $\mathcal{K}$ has an open rectangle neighborhood contained in $\mathcal{O}$. A finite subcover gives a union of rectangles $\mathcal{D}$ with $\mathcal{K} \subseteq \mathcal{D} \subseteq \mathcal{O}$. Thus $\mathbf{f}' := \mathbf{1}_{\mathcal{D}}$ satisfies

$$\mathbb{E}_{Y,Z}[\mathbf{f} - \mathbf{f}'] = (\nu^{\otimes 4} \otimes \nu^{\otimes 2})(\mathcal{E} \Delta \mathcal{D}) \leq (\nu^{\otimes 4} \otimes \nu^{\otimes 2})(\mathcal{O} \setminus \mathcal{K}) < \frac{1}{6r}.$$

Construct $\mathbf{g}', \mathbf{h}'$ similarly. Partition each of $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ according to membership in all rectangle sides used in these three constructions. The resulting finite partitions $\mathcal{P}_X, \mathcal{P}_Y, \mathcal{P}_Z$ make $\mathbf{f}', \mathbf{g}', \mathbf{h}'$ constant on the corresponding cell products, with each approximation error at most $1/(6r)$.

Step 2. Make the partitions equal in measure. Let $s = \max\{|\mathcal{P}_X|, |\mathcal{P}_Y|, |\mathcal{P}_Z|\}$ and choose an integer $q \geq 12rs$. Since Haar measure is nonatomic for $m \geq 2$, every measurable set contains subsets of any prescribed measure between zero and its own measure (see Carlen [Car10, p. 2]). Thus each $S \in \mathcal{P}_X$ splits into $\lfloor q\nu^{\otimes 4}(S) \rfloor$ sets of measure $1/q$ and a remainder R_S of measure less than $1/q$. Their union $R_X := \bigcup_{S \in \mathcal{P}_X} R_S$ satisfies

$$\nu^{\otimes 4}(R_X) < \frac{s}{q}, \quad q\nu^{\otimes 4}(R_X) = q - \sum_{S \in \mathcal{P}_X} \lfloor q\nu^{\otimes 4}(S) \rfloor \in \mathbb{Z}_{\geq 0}.$$

Therefore R_X also splits into sets of measure $1/q$, completing an equal-measure partition up to null sets. Repeat this on $\mathcal{Y}, \mathcal{Z}$, obtaining remainders with $\nu^{\otimes 4}(R_Y), \nu^{\otimes 2}(R_Z) < s/q$. Define

$$\mathbf{f}_r = \mathbf{f}' \mathbf{1}_{\mathcal{Y} \setminus R_Y} \mathbf{1}_{\mathcal{Z} \setminus R_Z}, \quad \mathbf{g}_r = \mathbf{g}' \mathbf{1}_{\mathcal{X} \setminus R_X} \mathbf{1}_{\mathcal{Z} \setminus R_Z}, \quad \mathbf{h}_r = \mathbf{h}' \mathbf{1}_{\mathcal{X} \setminus R_X} \mathbf{1}_{\mathcal{Y} \setminus R_Y}.$$

Outside the remainders, the new cells refine the old ones; on the remainder cells, these functions are zero. Assign null sets to cells and extend values constantly there. Thus $\mathbf{c}_r := \mathbf{f}_r \mathbf{g}_r \mathbf{h}_r$ is finite-type.

Step 3. Bound the approximation error. For $\mathbf{f}_r$,

$$\mathbb{E}_{Y,Z} |\mathbf{f} - \mathbf{f}_r| \leq \frac{1}{6r} + \nu^{\otimes 4}(R_Y) + \nu^{\otimes 2}(R_Z) \leq \frac{1}{6r} + \frac{2s}{q} \leq \frac{1}{3r},$$

and the same bound holds for $\mathbf{g}_r, \mathbf{h}_r$. By Claim 4.1, both μ_0 and μ_1 have the product Haar marginals used above. Hence, for $i \in \{0, 1\}$,

$$\begin{aligned} |\mu_i(\mathbf{c} - \mathbf{c}_r)| &\leq \mathbb{E}_{\mu_i} |\mathbf{f} \mathbf{g} \mathbf{h} - \mathbf{f}_r \mathbf{g}_r \mathbf{h}_r| \\ &\leq \mathbb{E}_{Y,Z} |\mathbf{f} - \mathbf{f}_r| + \mathbb{E}_{X,Z} |\mathbf{g} - \mathbf{g}_r| + \mathbb{E}_{X,Y} |\mathbf{h} - \mathbf{h}_r| \\ &\leq 1/r. \end{aligned}$$

□