ROBUST LOGARITHMIC LOWER BOUND ON SHARED-RESOURCE COST FOR f-ROUTING

KEVIN BOGNER
COSIC, KU Leuven, Belgium
kevin.bogner@kuleuven.be

Abstract. In one-round f-routing, Alice receives an n-bit string and an unknown qubit, while Bob receives another n-bit string. They exchange one simultaneous message each, and the value of f determines which party must recover the qubit. Message lengths, local systems, and local operations are unrestricted. We charge only $E_{\dim}(\rho_{LR}) = \log_2 \min\{\operatorname{rank} \rho_L, \operatorname{rank} \rho_R\}$, the logarithm of the smaller marginal support dimension of the shared state prepared before the inputs arrive. The state may be arbitrary and mixed. For the inner product modulo 2, we prove that every protocol with worst-case error at most 0.09 in both routing cases, measured in the full, unhalved diamond norm, satisfies $d \log_2(2d) = \Omega(n)$ for $d = \min\{\operatorname{rank} \rho_L, \operatorname{rank} \rho_R\}$. Thus $d = \Omega(n/\log n)$ and $E_{\dim}(\rho_{LR}) \geq \log_2 n - \log_2 \log_2 n - O(1)$. The closest earlier growing Schmidt-rank lower bound for an explicit routing function assumes zero error in one routing case. Our proof converts correctness into a matrix whose entries have a constant gap between the two routing cases. It approximates this matrix by one whose rank depends on d, but not on the dimensions of messages or local systems. A sign-rank lower bound for the matrix of the inner product modulo 2 completes the argument. A lower bound polynomial in n on $E_{\dim}$ remains open.

Disclosure of AI use. The mathematical content of this paper and its Lean 4 formalization were developed with substantial use of GPT-5.6 Sol (OpenAI). The author reviewed and verified the resulting material and takes full responsibility for the content of the paper and its formalization.

1. Introduction

In f-routing, Alice receives an n-bit string x and an unknown qubit Q. Bob receives an n-bit string y. They share a state prepared before the inputs arrive and exchange one simultaneous message each. No further communication is allowed. Alice must recover Q if $f(x, y) = 0$, and Bob must recover it if $f(x, y) = 1$. Correctness is measured in the full, unhalved diamond norm. Thus recovery must also preserve entanglement between Q and an arbitrary reference system. This is the standard one-round f-routing model [May26, BHM+26].

We place no bound on message lengths, local systems, or local operations. We charge only the shared state ρ_{LR} prepared before the inputs arrive. Its cost is

$$E_{\dim}(\rho_{LR}) = \log_2 \min\{\operatorname{rank} \rho_L, \operatorname{rank} \rho_R\}.$$

If $d = \min\{\operatorname{rank} \rho_L, \operatorname{rank} \rho_R\}$, then d is the smaller marginal support dimension. The state may be mixed, and $E_{\dim}$ is not an entanglement measure. It also charges separable correlations and shared classical randomness. Other work considers costs based on EPR pairs, Schmidt rank, or local dimension [May26, ACM25]. Lemma 2 shows that free local processing can generate the original mixed state from a pure shared state of Schmidt rank d. This preserves the induced routing channels and does not increase the cost.

The main result concerns the inner product modulo 2. It allows worst-case error in both routing cases.

Theorem 1. *There is an absolute constant $c > 0$ with the following property. Let $f_n(x, y) = \bigoplus_{i=1}^{n} x_i y_i$ be the inner product modulo 2 of $x, y \in \{0, 1\}^n$. For all sufficiently large n and every protocol for f_n in the standard model of Section 2 with worst-case error at most* 0.09*, if $d = \min\{\operatorname{rank} \rho_L, \operatorname{rank} \rho_R\}$, then $d \log_2(2d) \geq cn$. Consequently, the cost satisfies $E_{\dim}(\rho_{LR}) \geq \log_2 n - \log_2 \log_2 n - O(1)$.*

Lean certification. A Lean 4 formalization machine-checks the complete internal proof of Theorem 1. The information-disturbance tradeoff, the comparison for measured quantum triangular discrimination, and the sign-rank bound [KSW08, Liu25, For02] are treated as hypotheses. The formalization is available at `https://github.com/kevinbogner/f-routing-lean`.

The theorem implies $d = \Omega(n/\log n)$. The charged cost is therefore at least $\log_2 n - \log_2 \log_2 n - O(1)$. The closest earlier result gives growing Schmidt-rank lower bounds for explicit routing functions but assumes zero error in one routing case [ACM25]. For the stated function and cost, the theorem instead allows error in both routing cases. It also permits arbitrary mixed shared states, unrestricted message lengths, and local systems of arbitrary finite dimension.

The proof converts routing correctness into a rank bound for a matrix. The cost-preserving purification first reduces the shared resource to a pure state of Schmidt rank d. For every input pair, Bob's two states are then sums of at most d^2 tensor products. In each product, one factor depends only on x and the other only on y. Correctness and the information-disturbance tradeoff give a constant trace-norm gap between the two routing cases. The symmetric logarithmic derivative quadratic form preserves this gap and is compatible with the tensor-product expansions. Mixing each state with a small amount of a product state gives uniform bounds for the inverse that defines the quadratic form. A polynomial approximation and a Fourier expansion then show that the resulting matrix can be approximated entrywise by a real matrix of rank at most

$$\exp\big(O(d\log(2d))\big)\log(2N), \qquad N = 2^n.$$

This bound has no dependence on message or workspace dimensions. For the inner product modulo 2, any real matrix with the required entrywise signs has rank at least $\sqrt{N}$. Comparing the two rank bounds gives $d\log(2d) = \Omega(n)$.

Robust lower bounds are known in other settings. Some charge all attack registers in a bounded quantum-storage model [BCS22]. Others charge gates and measurements [ACCM26], communication alone or together with the shared resource [AKL+25, GMOW26], or entanglement of formation for selected unitary tasks in non-local quantum computation [CM26]. Those results use different tasks or cost measures. Further work relates routing to other non-local tasks [ABM+24, BHM+26]. It remains open to prove, for an explicit function in the model considered here, a two-sided-error lower bound on $E_{\dim}$ that grows asymptotically faster than logarithmically. In particular, a lower bound polynomial in n on $E_{\dim}$ remains open [May26].

2. Routing model

All Hilbert spaces are finite-dimensional. Logarithms without a base subscript are natural. Fix a total Boolean function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$. Alice receives $x \in \{0,1\}^n$ and a qubit Q. Bob receives $y \in \{0,1\}^n$. Before the inputs arrive they share a state ρ_{LR}, with Alice holding L and Bob holding R. The state can depend on n and on the complete function f, but not on the inputs or on the state of Q. Figure 1 shows the protocol.

In the first round, Alice and Bob apply arbitrary channels $\mathcal{A}_x : QL \to A_0C_{A\to B}$ and $\mathcal{B}_y : R \to B_0C_{B\to A}$. Alice keeps A_0 and sends the message $C_{A\to B}$ to Bob. Bob keeps B_0 and sends $C_{B\to A}$ to Alice. The messages cross simultaneously, so afterwards Alice holds $M_A = A_0C_{B\to A}$ and Bob holds $M_B = B_0C_{A\to B}$. For fixed inputs, everything before recovery is therefore one channel from the input qubit to these two systems,

$$\mathcal{N}^{x,y}_{Q\to M_AM_B}(\,\cdot\,) = (\mathcal{A}_x \otimes \mathcal{B}_y)(\,\cdot\, \otimes \rho_{LR}),$$

with the output regrouped into M_A and M_B. The designated party must then recover the qubit with a map that may depend on both input strings, $\mathcal{D}^{x,y}_A : M_A \to Q$ for Alice and $\mathcal{D}^{x,y}_B : M_B \to Q$ for Bob. For a linear map Φ, $\|\Phi\|_\diamond$ is the full, unhalved diamond norm. The protocol is ϵ-correct if the following bounds hold for every input pair:

$$f(x,y) = 0 \Longrightarrow \|\mathcal{D}^{x,y}_A \circ \operatorname{Tr}_{M_B} \circ \mathcal{N}^{x,y} - \mathrm{id}_Q\|_\diamond \le \epsilon,$$
$$f(x,y) = 1 \Longrightarrow \|\mathcal{D}^{x,y}_B \circ \operatorname{Tr}_{M_A} \circ \mathcal{N}^{x,y} - \mathrm{id}_Q\|_\diamond \le \epsilon.$$

The first condition says that when $f(x,y) = 0$, Alice must recover the qubit from her final system alone. The second says the same for Bob when $f(x,y) = 1$. The tolerance ϵ bounds

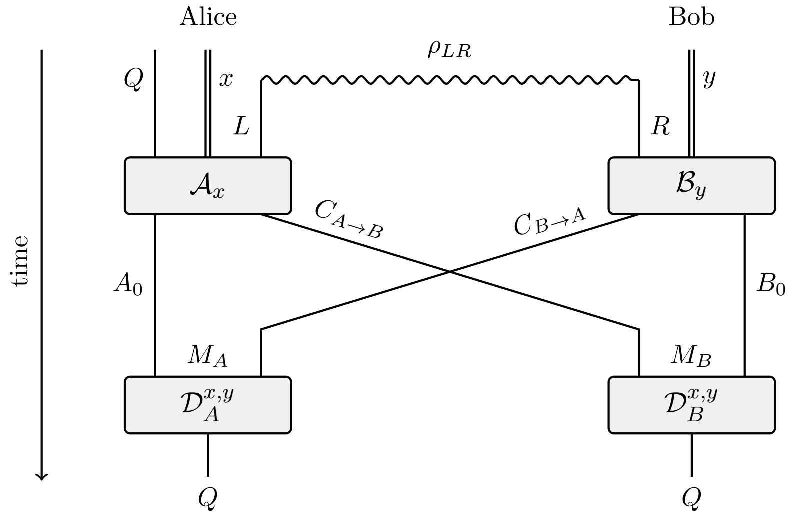

FIGURE 1. One round of f-routing. Time runs downward. Alice receives x and the qubit Q, Bob receives y, and they share the state ρ_{LR} prepared before the inputs arrive. Double lines carry the classical inputs. Alice applies $\mathcal{A}_x$ to QL, keeps A_0, and sends the message $C_{A\to B}$. Bob applies $\mathcal{B}_y$ to R, keeps B_0, and sends $C_{B\to A}$. The messages cross simultaneously. If $f(x,y)=0$, Alice applies $\mathcal{D}_A^{x,y}$ to $M_A = A_0 C_{B\to A}$ and outputs the qubit. If $f(x,y)=1$, Bob applies $\mathcal{D}_B^{x,y}$ to $M_B = B_0 C_{A\to B}$.

how far the whole process may deviate from returning the qubit unchanged. Messages, local outputs, local ancillas, environments, Kraus ranks, local randomness, and local computation are not charged. The model above is standard [May26, BHM$^+$26]. The main nonstandard choice is the cost measure $E_{\dim}$ from the introduction.

The first part of the proof of Theorem 1 reduces every protocol to a normal form that the rest of the paper assumes. The shared state ρ_{LR} is replaced by a pure state whose Schmidt rank is the smaller marginal rank d, so each half of the new state has dimension d. The first-round channels are replaced by isometries, channels of the form $V(\,\cdot\,)V^\dagger$ with $V^\dagger V = I$, which discard nothing. The lemma proves that the cost and all channels $\mathcal{N}^{x,y}$ are unchanged.

Lemma 2 (Cost-preserving purification)**.** *Every protocol with shared state ρ_{LR} and $d = \min\{\operatorname{rank}\rho_L, \operatorname{rank}\rho_R\}$ has an equivalent protocol with a pure shared state of Schmidt rank d. Both marginals of the new state have rank d, and every channel $\mathcal{N}^{x,y}$ is reproduced exactly after local systems are discarded. The first-round channels can also be taken to be isometries, with the extra output systems kept on their own sides.*

Proof. Assume that $\operatorname{rank}\rho_L = d \le \operatorname{rank}\rho_R$. The other case follows by exchanging the roles of L and R. Choose a purification $|\Psi\rangle_{LRE}$ and regard E as a system on Bob's side. Viewed as a bipartite pure state between L and the pair RE, $|\Psi\rangle$ has Schmidt rank d, since its marginal on L is the original ρ_L, which has rank d. Let $K \subseteq R\otimes E$ be the d-dimensional subspace spanned by the Schmidt vectors on Bob's side. There is an isometry $V : K \to R\otimes E$ and a pure state $|\psi\rangle_{LK}$ such that $(I_L \otimes V)|\psi\rangle = |\Psi\rangle$. Use $|\psi\rangle_{LK}$ as the new shared resource. As the first part of Bob's first-round channel, apply V, discard E, and then apply the old first-round channel to R. The state on LR is then exactly ρ_{LR}, so every later channel is unchanged. Both marginals of $|\psi\rangle$ have rank d.

Finally make the first-round channels isometries. By the Stinespring dilation, every channel can be implemented as an isometry into a larger output space followed by discarding an extra register, called its environment. Implement both first-round channels this way and keep the environments instead of discarding them, Alice hers and Bob his. The recovery channels can discard these registers first, so the original channels are reproduced exactly. Now nothing before recovery discards information, and for every input pair the map from Q to the enlarged systems

$M_A M_B$ is an isometry. Every system is held inside the two laboratories, and the shared state is untouched, so the cost E_{dim} is unchanged. $\square$

Applying Lemma 2, replace the shared state ρ_{LR} by an equivalent pure state, written in Schmidt form as

$$|\psi\rangle_{LR} = \sum_{r=1}^{d} \sqrt{\lambda_r}\, |r\rangle_L |r\rangle_R, \qquad \lambda_r > 0, \qquad \sum_r \lambda_r = 1.$$

The vectors $|r\rangle_L$ and $|r\rangle_R$ form orthonormal bases of the two d-dimensional halves, and the weights λ_r form a probability distribution. The rest of the proof analyzes Bob's system just before recovery. Studying Bob's side is enough. When Bob must recover the qubit, his states for the two computational-basis inputs must be far apart. When Alice must recover it, those two states on Bob's side must be close.

Keep the local Stinespring environments. Bob's system then has a fixed decomposition $M_B = X \otimes Y$. Here $X = C_{A\to B}$ is the message from Alice, and Y contains Bob's register B_0 and his Stinespring environment. For $q \in \{0,1\}$, define a channel $\mathcal{E}_x^q : L \to X$ as follows. Fix the input qubit to $|q\rangle$, apply Alice's first-round channel, and trace out everything Alice keeps. Also let $\mathcal{F}_y : R \to Y$ be the channel that applies Bob's first-round isometry and traces out the message $C_{B\to A}$ sent to Alice. Bob's state before recovery is

$$\rho_{x,y}^q = (\mathcal{E}_x^q \otimes \mathcal{F}_y)(|\psi\rangle\langle\psi|) = \sum_{r,s=1}^{d} \sqrt{\lambda_r \lambda_s}\, \mathcal{E}_x^q(|r\rangle\langle s|) \otimes \mathcal{F}_y(|r\rangle\langle s|). \tag{1}$$

For each fixed q, every state is a sum of at most d^2 tensor products. In each product, the first factor depends only on x and the second only on y. This is a product expansion with separated row and column factors. An expansion of this form appears in [ACM25]; the version here displays the channels and the products of Schmidt coefficients.

The rest of the proof uses the average $\tau_{x,y} = \frac{1}{2}(\rho_{x,y}^0 + \rho_{x,y}^1)$ and the difference $\Delta_{x,y} = \frac{1}{2}(\rho_{x,y}^0 - \rho_{x,y}^1)$. Equivalently, $\rho_{x,y}^0 = \tau_{x,y} + \Delta_{x,y}$ and $\rho_{x,y}^1 = \tau_{x,y} - \Delta_{x,y}$. The pair $(\tau_{x,y}, \Delta_{x,y})$ therefore contains the same information as the two states. The size of $\Delta_{x,y}$ measures how well Bob can tell whether the input qubit was $|0\rangle$ or $|1\rangle$. Correctness makes it large when Bob must recover the qubit and small when Alice must. The average $\tau_{x,y}$ is Bob's state for an equally weighted mixture of the two computational-basis inputs. The next section measures $\Delta_{x,y}$ relative to this average. Both operators are linear combinations of $\rho_{x,y}^0$ and $\rho_{x,y}^1$, so both have the product expansion above.

3. Routing gap

This section is the second part of the proof of Theorem 1. It proves that the difference $\Delta_{x,y}$ separates the two routing outcomes by a constant gap. The difference is large for every input pair with $f(x,y) = 1$ and small for every input pair with $f(x,y) = 0$. We first bound the trace-norm difference $\|\rho_{x,y}^0 - \rho_{x,y}^1\|_1$. We then use this bound to derive bounds on the statistic used in the rest of the proof. For $f(x,y) = 1$, the lower bound follows directly from correctness because Bob must recover the qubit. Input pairs with $f(x,y) = 0$ require the next lemma.

The lemma gives explicit error bounds for the fact that Alice and Bob cannot both recover the qubit. It applies to any isometry that splits a qubit between two systems A and B. In the application, the isometry is the channel $\mathcal{N}^{x,y}$ in the normal form of Lemma 2. The system A becomes M_A, which Alice holds before recovery, and B becomes M_B, which Bob holds. The channel $\mathcal{D}$ becomes Alice's recovery map. When $f(x,y) = 0$, correctness says that Alice can recover the qubit. The lemma then shows that Bob's two states $\rho_{x,y}^0$ and $\rho_{x,y}^1$ are nearly equal, so he cannot also recover it. This standard result is called the information-disturbance tradeoff [KSW08]. A diamond-norm statement for f-routing appears in [ABM+24].

Lemma 3 (Information-disturbance tradeoff)**.** *Let $V : Q \to A \otimes B$ be an isometry, and let $\mathcal{N}_A(\cdot) = \mathrm{Tr}_B\, V(\cdot)V^*$ and $\mathcal{N}_B(\cdot) = \mathrm{Tr}_A\, V(\cdot)V^*$. If a channel $\mathcal{D} : A \to Q$ satisfies $\|\mathcal{D} \circ \mathcal{N}_A -$*

$\mathrm{id}_Q\|_\diamond \le \epsilon$, *then there is a state* ζ_B *such that* $\|\mathcal{N}_B - \mathcal{R}_{\zeta_B}\|_\diamond \le 2\sqrt{\epsilon}$, *where* $\mathcal{R}_{\zeta_B}(X) = \mathrm{Tr}(X)\zeta_B$ *is the replacer channel that discards its input and outputs* ζ_B*.*

Proof. See [ABM+24]. The proof applies the continuity theorem for Stinespring representations [KSW08] to $\mathcal{D} \circ \mathcal{N}_A$ and the identity channel. A Stinespring isometry for $\mathcal{D} \circ \mathcal{N}_A$ can retain B. The theorem shows that this isometry is $\sqrt{\epsilon}$-close to an implementation of the identity in which B is in a fixed state. Taking the reduced channel on B gives $\mathcal{R}_{\zeta_B}$. Converting the isometry bound to diamond norm gives the factor 2. □

The next proposition combines the two cases into bounds on the trace-norm difference between Bob's states $\rho^0_{x,y}$ and $\rho^1_{x,y}$ from Equation (1). Correctness gives the lower bound when $f(x,y) = 1$, since Bob must recover the qubit. Applying Lemma 3 to Alice's recovery gives the upper bound when $f(x,y) = 0$. This is the last time correctness enters the proof. Later steps use only these two bounds and the structure of the states. A form of the $f(x,y) = 0$ estimate appears in [ABM+24].

Proposition 4 (Trace-norm routing gap)**.** *For a protocol with diamond-norm error* ϵ *under the unhalved convention, the states in Equation* (1) *satisfy*

$$f(x,y) = 1 \Longrightarrow \|\rho^0_{x,y} - \rho^1_{x,y}\|_1 \ge 2(1-\epsilon), \qquad f(x,y) = 0 \Longrightarrow \|\rho^0_{x,y} - \rho^1_{x,y}\|_1 \le 4\sqrt{\epsilon}.$$

For $\epsilon = 0.09$*, the two bounds are* 1.82 *and* 1.2*.*

Proof. Suppose first that $f(x,y) = 1$. Let $\mathcal{D}_B$ be Bob's recovery map. For $q = 0,1$, $\|\mathcal{D}_B(\rho^q_{x,y}) - |q\rangle\langle q|\|_1 \le \epsilon$. The two basis states $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$ are orthogonal, so their trace-norm difference is 2. The triangle inequality and the fact that a channel cannot increase the trace-norm difference give $2 \le 2\epsilon + \|\rho^0_{x,y} - \rho^1_{x,y}\|_1$. Rearranging gives $\|\rho^0_{x,y} - \rho^1_{x,y}\|_1 \ge 2(1-\epsilon)$.

Now suppose that $f(x,y) = 0$. In the normal form, the global channel $\mathcal{N}^{x,y}$ with all retained systems is an isometry from Q to $M_A \otimes M_B$. Apply Lemma 3 to Alice's recovery. There is a fixed state ζ_{M_B} such that Bob's channel is within $2\sqrt{\epsilon}$ in diamond norm of the replacer channel with output ζ_{M_B}. Therefore $\|\rho^0_{x,y} - \rho^1_{x,y}\|_1 \le \|\rho^0_{x,y} - \zeta_{M_B}\|_1 + \|\rho^1_{x,y} - \zeta_{M_B}\|_1 \le 4\sqrt{\epsilon}$. □

The routing gap is now a bound on $\|\rho^0_{x,y} - \rho^1_{x,y}\|_1$. The third part of the proof needs a quantity that can be analyzed using the product form in Equation (1). The trace norm is the trace of an operator absolute value, which does not preserve that form. The quantity below instead uses addition, subtraction, the inverse of a linear map, and a trace. These operations can be analyzed using the same product form. Lemma 5 gives corresponding bounds for this quantity. The resulting constants appear in Equation (2).

We use the following conventions from here on. For a positive operator T, let P_T be its support projection. Inverses and complex powers of T are taken on $\operatorname{supp} T$ and extended by zero on the kernel. The same convention applies to Moore–Penrose inverses of positive semidefinite linear operators on Hilbert–Schmidt space.

For density operators ρ and σ, put $\Sigma = \rho + \sigma$, $Z = \rho - \sigma$, and $\mathcal{L}_\Sigma(X) = (\Sigma X + X\Sigma)/2$, and define the symmetric logarithmic derivative (SLD) quadratic form

$$\chi_{\mathrm{SLD}}(\rho,\sigma) = \mathrm{Tr}\big[Z\mathcal{L}^+_\Sigma(Z)\big].$$

The quantity $\chi_{\mathrm{SLD}}/2$ is both the measured quantum triangular discrimination of ρ and σ [Liu25] and the Bures quantum χ^2 divergence from either state to their average [TKR+10]. We use the doubled quantity because Lemma 5 gives the trace-norm bounds needed below in this scaling. The other quantum triangular discriminations considered in [Liu25] do not give the required upper bound.

Lemma 5 (Trace-norm bounds for the SLD quadratic form)**.** *For all density operators* ρ, σ*,*

$$\frac{\|\rho - \sigma\|_1^2}{2} \le \chi_{\mathrm{SLD}}(\rho,\sigma) \le \|\rho - \sigma\|_1.$$

Proof. This is a result of [Liu25] in the displayed scaling. There $T = \|\rho - \sigma\|_1/2$, and the measured quantum triangular discrimination is $\mathrm{Tr}[\rho_- \mathcal{L}^+_{\rho_+}(\rho_-)]$, where $\rho_\pm = (\rho \pm \sigma)/2$. Substituting

$\rho_+ = \Sigma/2$ and $\rho_- = Z/2$ shows that this quantity is $\chi_{\mathrm{SLD}}/2$. The bound $T^2 \le \chi_{\mathrm{SLD}}/2 \le T$ from [Liu25] is therefore the displayed bound. $\square$

For each input pair, abbreviate $\chi(x,y) = \chi_{\mathrm{SLD}}(\rho^0_{x,y}, \rho^1_{x,y})$. Since $\rho^0+\rho^1 = 2\tau$ and $\rho^0-\rho^1 = 2\Delta$, the quantity satisfies $\chi(x,y)/2 = \mathrm{Tr}[\Delta_{x,y}\mathcal{L}^+_{\tau_{x,y}}(\Delta_{x,y})]$. Combining Proposition 4 and Lemma 5 at $\epsilon = 0.09$ gives

$$f(x,y) = 0 \Longrightarrow \chi(x,y) \le 1.2, \qquad f(x,y) = 1 \Longrightarrow \chi(x,y) \ge \frac{1.82^2}{2} = 1.6562. \tag{2}$$

4. Approximate rank of the statistic

This section is the third part of the proof of Theorem 1. The values $\chi(x,y)$ form a matrix with one row per x and one column per y. We show that this matrix is close, entry by entry, to a low-rank matrix. By Equation (2), each entry is above or below a fixed threshold according to $f(x,y)$. An error below half the gap does not change which side of the threshold contains an entry. Thus the low-rank matrix still determines f, and a function that requires high rank gives a lower bound on the shared state.

The inverse of $\mathcal{L}_{\tau_{x,y}}$ requires a separate step. The average state $\tau_{x,y}$ can have arbitrarily small eigenvalues, which make the inverse large. We address this by mixing a small amount of a product state into the states for every input pair. The new average lies between two multiples of that product state, with ratio $O(d^2)$ (Lemma 6). After preconditioning by the product state, the resulting operator therefore has uniform lower and upper spectral bounds. The gap decreases only slightly (Lemma 7).

Let $\overline{\mathcal{E}}_x = (\mathcal{E}^0_x + \mathcal{E}^1_x)/2$, $\alpha_x = \overline{\mathcal{E}}_x(I_L/d)$, and $\beta_y = \mathcal{F}_y(\rho_R)$, where $\rho_R = \sum_r \lambda_r |r\rangle\langle r|$. Both α_x and β_y are density operators. Define the product state $\Omega_{x,y} = \alpha_x \otimes \beta_y$.

Lemma 6 (Product-state upper bound)**.** *On the Schmidt support, $|\psi\rangle\langle\psi| \preceq dI_L \otimes \rho_R$. Consequently, $\tau_{x,y} \preceq d^2\Omega_{x,y}$ for every input pair.*

Proof. The first inequality follows from the pinching inequality [Hay02]. In the form used here, the constant is the number of blocks [OH04]. We include the short direct proof. For a vector $|v\rangle = \sum_{r,s} v_{r,s}|r,s\rangle$,

$$|\langle\psi \mid v\rangle|^2 = |\sum_r \sqrt{\lambda_r} v_{r,r}|^2 \le d\sum_r \lambda_r |v_{r,r}|^2 \le d\langle v|I_L \otimes \rho_R|v\rangle.$$

This proves the operator inequality. The map $\overline{\mathcal{E}}_x \otimes \mathcal{F}_y$ preserves operator inequalities. Applying it gives $\tau_{x,y} \preceq d\,\overline{\mathcal{E}}_x(I_L) \otimes \mathcal{F}_y(\rho_R) = d^2\alpha_x \otimes \beta_y$. $\square$

Fix $s = 0.05$ and define $\rho^{q,s}_{x,y} = (1-s)\rho^q_{x,y} + s\Omega_{x,y}$, $\tau^s_{x,y} = (1-s)\tau_{x,y} + s\Omega_{x,y}$, and $\Delta^s_{x,y} = (1-s)\Delta_{x,y}$. Lemma 6 implies

$$a\Omega_{x,y} \preceq \tau^s_{x,y} \preceq b\Omega_{x,y}, \qquad a = 0.05, \qquad b = 0.95d^2 + 0.05. \tag{3}$$

Let $\chi_s(x,y) = \chi_{\mathrm{SLD}}(\rho^{0,s}_{x,y}, \rho^{1,s}_{x,y})$. For Hermitian operators U, V, write $\langle U, V\rangle_{\mathsf{HS}} = \mathrm{Tr}(UV)$ for their Hilbert–Schmidt inner product.

Lemma 7 (Bounds after mixing)**.** *For every input pair, $(1-s)\chi(x,y) - 2s \le \chi_s(x,y) \le (1-s)\chi(x,y)$.*

Proof. If L is a positive definite linear map on Hilbert–Schmidt space and v is Hermitian, then the following identity holds [BV04]:

$$\langle v, L^{-1}v\rangle_{\mathsf{HS}} = \max_X \left(2\langle v, X\rangle_{\mathsf{HS}} - \langle X, LX\rangle_{\mathsf{HS}}\right), \tag{4}$$

where the maximum may be restricted to Hermitian X.

For the upper bound, apply Equation (4) to $\mathcal{L}_{\tau^s}$ and Δ^s on $\operatorname{supp}\Omega$. Since $\mathcal{L}_{\tau^s} \succeq (1-s)\mathcal{L}_\tau$ and $\Delta^s = (1-s)\Delta$, the expression for every Hermitian X is at most $(1-s)$ times the corresponding expression for $\mathcal{L}_\tau$ and Δ. The maximum of the latter expression is $\chi/2$. Indeed, $-\tau \preceq \Delta \preceq \tau$ implies that Δ belongs to the range of $\mathcal{L}_\tau$.

For the lower bound, set $X = \mathcal{L}_\tau^+(\Delta)$ and extend it by zero from $\operatorname{supp}\tau$ to $\operatorname{supp}\Omega$. On operators supported on $\operatorname{supp}\tau$, the inverse of $\mathcal{L}_\tau$ has a positive integral representation [HP14]. Together with $-\tau \preceq \Delta \preceq \tau$, this gives $-I \preceq X \preceq I$ on $\operatorname{supp}\Omega$. Insert this X into Equation (4) for $\mathcal{L}_{\tau^s}$ and Δ^s. Since $\langle X, \mathcal{L}_\Omega(X)\rangle_{\mathsf{HS}} = \operatorname{Tr}(\Omega X^2) \le 1$,

$$\frac{\chi_s}{2} \ge 2(1-s)\langle \Delta, X\rangle_{\mathsf{HS}} - (1-s)\langle X, \mathcal{L}_\tau(X)\rangle_{\mathsf{HS}} - s\langle X, \mathcal{L}_\Omega(X)\rangle_{\mathsf{HS}} \ge (1-s)\frac{\chi}{2} - s.$$

Multiplication by two proves the lower bound. □

The bounds in Equation (2) now give

$$f(x,y) = 0 \Longrightarrow \chi_s(x,y) \le 1.14, \qquad f(x,y) = 1 \Longrightarrow \chi_s(x,y) \ge 1.47339. \tag{5}$$

Set

$$\theta = 1.306695, \qquad g = 0.166695. \tag{6}$$

Thus $\chi_s(x,y) \le \theta - g$ when $f(x,y) = 0$, and $\chi_s(x,y) \ge \theta + g$ when $f(x,y) = 1$.

We next show that each matrix q_j defined below is close, entry by entry, to a low-rank matrix. A polynomial approximates the inverse in χ_s. Expanding the polynomial writes the approximation to χ_s as a linear combination of the q_j, so we can study one q_j at a time.

We use two related preconditionings. First, $A = \mathcal{L}_\Omega^{-1/2}\mathcal{L}_{\tau^s}\mathcal{L}_\Omega^{-1/2}$ has eigenvalues in $[a,b]$. Second, define $C = \Omega^{-1/2}E\Omega^{-1/2}$ and $D = \Omega^{-1/2}\Delta^s\Omega^{-1/2}$. Each is still a sum of tensor products. In each product, one factor depends only on x and the other only on y. Both the number of terms and the operator-norm bounds depend only on d.

The remaining step concerns $\mathcal{L}_\Omega^{-1}$. In an eigenbasis of Ω, this map introduces factors $1/(\omega_u+\omega_v)$. These factors do not separate into an x-dependent factor and a y-dependent factor. The Fourier expansion below writes them as integrals of complex powers. Since $\Omega_{x,y} = \alpha_x \otimes \beta_y$, these powers separate as $\alpha_x^z \otimes \beta_y^z$. This gives the rank bound in Lemma 12. All linear operators on Hilbert–Schmidt space below act on $\operatorname{supp}\Omega_{x,y}$.

Define $A = \mathcal{L}_\Omega^{-1/2}\mathcal{L}_{\tau^s}\mathcal{L}_\Omega^{-1/2}$ and $u = \mathcal{L}_\Omega^{-1/2}(\Delta^s)$; the indices x,y are suppressed. From Equation (3),

$$aI \preceq A \preceq bI. \tag{7}$$

Moreover,

$$\frac{\chi_s}{2} = \langle u, A^{-1}u\rangle_{\mathsf{HS}} \le 1, \tag{8}$$

where the last inequality follows from Lemma 5 and the bound $\|\rho^{0,s} - \rho^{1,s}\|_1 \le 2$. Put $E = \tau^s - \Omega$ and $B = A - I = \mathcal{L}_\Omega^{-1/2}\mathcal{L}_E\mathcal{L}_\Omega^{-1/2}$. Also define the Ω-preconditioned operators $C = \Omega^{-1/2}E\Omega^{-1/2}$ and $D = \Omega^{-1/2}\Delta^s\Omega^{-1/2}$.

Definition 1. A family of operators $O_{x,y}$ on $X \otimes Y$ has a product expansion of length at most r if there are operator families $O^X_{x,k}$ and $O^Y_{y,k}$ such that

$$O_{x,y} = \sum_{k=1}^{r} O^X_{x,k} \otimes O^Y_{y,k}$$

for every (x,y). No bound on the individual summands is part of this definition.

Lemma 8 (Preconditioned product decomposition)**.** *The operators $C_{x,y}$ and $D_{x,y}$ have product expansions with lengths r_C and r_D, where $r_C \le d^2+1$ and $r_D \le d^2$. They satisfy $\|C_{x,y}\|_{\mathrm{op}} \le d^2$ and $\|D_{x,y}\|_{\mathrm{op}} \le d^2$. These bounds depend only on d. They do not depend on the dimensions of the messages, local outputs, work registers, or environments, or on the Kraus ranks.*

Proof. For each q, the range of every Kraus operator of $\mathcal{E}_x^q$ lies in $\operatorname{supp}\alpha_x$. Therefore, for every T,

$$P_{\alpha_x}\mathcal{E}_x^q(T)P_{\alpha_x} = \mathcal{E}_x^q(T).$$

The same identity holds for $\overline{\mathcal{E}}_x$ and $\mathcal{G}_x$. Since ρ_R is positive definite on the Schmidt support, $\mathcal{F}_y(\rho_R)$ and $\mathcal{F}_y(I)$ have the same support. Thus $P_{\beta_y}\mathcal{F}_y(T)P_{\beta_y} = \mathcal{F}_y(T)$ for every T on that support. Every term in Equation (1) is therefore supported on $\operatorname{supp}\alpha_x \otimes \operatorname{supp}\beta_y$. The conjugations by the Moore–Penrose inverse square roots below are well-defined.

Let $\mathcal{G}_x = (\mathcal{E}_x^0 - \mathcal{E}_x^1)/2$. Using Equation (1) and $\Omega^{-1/2} = \alpha_x^{-1/2} \otimes \beta_y^{-1/2}$ on the support yields

$$D_{x,y} = (1-s) \sum_{r,s=1}^{d} \sqrt{\lambda_r \lambda_s}\, \alpha_x^{-1/2} \mathcal{G}_x(|r\rangle\langle s|) \alpha_x^{-1/2} \otimes \beta_y^{-1/2} \mathcal{F}_y(|r\rangle\langle s|) \beta_y^{-1/2},$$

which has length at most d^2. Since $E = (1-s)(\tau - \Omega)$, the same expansion with $\overline{\mathcal{E}}_x$ gives d^2 terms for $\Omega^{-1/2}\tau\Omega^{-1/2}$, followed by the additional product term $-(1-s)P_{\alpha_x} \otimes P_{\beta_y}$. Thus $r_C \le d^2 + 1$. The inequalities $-\tau \preceq \Delta \preceq \tau$ and $\tau \preceq d^2\Omega$ imply $-d^2 P_\Omega \preceq D \preceq d^2 P_\Omega$. The two inequalities in Equation (3) give $(a-1)P_\Omega \preceq C \preceq (b-1)P_\Omega$. With the values of a, b, both operator norms are at most d^2 for every $d \ge 1$. □

For $j \ge 0$, define the real matrix

$$q_j(x,y) = \langle \Delta^s, \mathcal{L}_\Omega^{-1}(\mathcal{L}_E \mathcal{L}_\Omega^{-1})^j(\Delta^s)\rangle_{\mathsf{HS}}.$$

Lemma 9 (Moment identity). *For every $j \ge 0$, $q_j(x,y) = \langle u, B^j u\rangle_{\mathsf{HS}}$.*

Proof. Substitute $B = \mathcal{L}_\Omega^{-1/2}\mathcal{L}_E\mathcal{L}_\Omega^{-1/2}$ and $u = \mathcal{L}_\Omega^{-1/2}(\Delta^s)$. Each pair of adjacent factors $\mathcal{L}_\Omega^{-1/2}$ multiplies to $\mathcal{L}_\Omega^{-1}$. Therefore $\langle u, B^j u\rangle_{\mathsf{HS}} = \langle \Delta^s, \mathcal{L}_\Omega^{-1}(\mathcal{L}_E\mathcal{L}_\Omega^{-1})^j(\Delta^s)\rangle_{\mathsf{HS}}$. □

The next lemma chooses the positive real exponents used in the Fourier expansion. It is stated for a tree because the expansion below always produces one.

Lemma 10 (Edge weights on a tree). *Let $T = (V, E)$ be a tree with $m \ge 1$ edges. Set $\varepsilon_m = 1/(2m)$. For each endpoint v of each edge e, there is a number $\theta_{e,v}$ such that $\theta_{e,u} + \theta_{e,v} = 1$ for $e = \{u, v\}$, $\varepsilon_m \le \theta_{e,v} \le 1 - \varepsilon_m$, and $z_v := \sum_{e \ni v} \theta_{e,v} \le 1$ for every v.*

Proof. Assign an initial value ε_m to each endpoint of every edge. Each edge then needs an additional total $r = 1 - 2\varepsilon_m = 1 - 1/m$, and vertex v can receive at most an additional $c_v = 1 - \varepsilon_m \deg_T(v) \ge 0$. Build a flow network. Connect the source to each edge-node with capacity r. Connect each edge-node to its two endpoint vertices with infinite capacity, and connect each vertex to the sink with capacity c_v.

It remains to check the following cut inequality. Let $F \subseteq E$ be nonempty, and let $V(F)$ be the set of its endpoints. The graph $(V(F), F)$ is a forest with $k \ge 1$ components, so $|V(F)| = |F| + k$. Also $\sum_{v \in V(F)} \deg_T(v) \le 2m$. Therefore

$$\sum_{v \in V(F)} c_v \ge |F| + k - 2m\varepsilon_m = |F| + k - 1 \ge |F| \ge r|F|.$$

Every cut has enough capacity. The max-flow min-cut theorem [FF56] therefore gives a flow in which every edge-node sends its required amount. Add this flow to the initial endpoint values. Each $\theta_{e,v}$ then lies in $[\varepsilon_m, 1-\varepsilon_m]$, and the sum at each vertex is at most one. □

Use the Fourier convention

$$\widehat{g}(t) = \int_{\mathbb{R}} g(z) e^{-itz}\, dz, \qquad g(z) = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{g}(t) e^{itz}\, dt.$$

Lemma 11 (Fourier transform of g_θ). *For $0 < \theta < 1$, let $g_\theta(z) = e^{\theta z}/(1 + e^z)$. Then*

$$\widehat{g_\theta}(t) = \Gamma(\theta - it)\Gamma(1 - \theta + it) = \frac{\pi}{\sin(\pi(\theta - it))},$$

and there is an absolute constant C_F such that $\frac{1}{2\pi}\|\widehat{g_\theta}\|_1 \le C_F\big(1 + \log \frac{1}{\min\{\theta, 1-\theta\}}\big)$.

Proof. With $u = e^z$, $\widehat{g_\theta}(t) = \int_0^\infty u^{\theta - it - 1}(1+u)^{-1}\, du = \Gamma(\theta - it)\Gamma(1 - \theta + it)$: the integral is the beta integral, and the reflection formula gives the sine expression [DLMF]. Moreover, $|\sin(\pi(\theta - it))|^2 = \sin^2(\pi\theta) + \sinh^2(\pi t)$. Let $\delta = \min\{\theta, 1-\theta\}$. For $0 \le \delta \le 1/2$, $\sin(\pi\delta) \ge 2\delta$. On $0 \le t \le 1$, $\sinh(\pi t) \ge \pi t$, while on $t \ge 1$ it grows exponentially. Hence

$$\frac{1}{2\pi}\|\widehat{g_\theta}\|_1 \le C \int_0^1 \frac{dt}{\sqrt{\delta^2 + t^2}} + C \int_1^\infty e^{-\pi t}\, dt \le C_F \left(1 + \log \frac{1}{\delta}\right).$$

□

For a real matrix M, define its entrywise approximate rank by $\operatorname{rank}_\eta(M) = \min\{\operatorname{rank}_{\mathbb{R}} \widetilde{M} : \|M - \widetilde{M}\|_{\max} \le \eta\}$, where $\|A\|_{\max} = \max_{x,y}|A_{x,y}|$.

Lemma 12 (Moment rank bound)**.** *Let the row and column input sets both have size N. Suppose $C_{x,y}$ and $D_{x,y}$ have product expansions of length at most r_C and r_D, respectively. Let $K_C = \max_{x,y}\|C_{x,y}\|_{\mathrm{op}}$ and $K_D = \max_{x,y}\|D_{x,y}\|_{\mathrm{op}}$. For $j \ge 0$, put $R_j = r_D^2 r_C^j$, $M_j = K_D^2 K_C^j$, and $V_j = 2^{j+1}\big(C_F(1+\log(2j+4))\big)^{j+1}$. Then for every $\eta > 0$,*

$$\operatorname{rank}_\eta(q_j) \le 2R_j \left\lceil 6V_j^2 M_j^2 \eta^{-2} \log(2N) \right\rceil . \tag{9}$$

Proof. Fix one input pair and diagonalize Ω on its support, with eigenvalues $\omega_a > 0$. In this basis, $\big(\mathcal{L}_\Omega^{-1}(X)\big)_{a,b} = 2X_{a,b}/(\omega_a+\omega_b)$. Each occurrence of $\mathcal{L}_E$ is the average of left and right multiplication by E. Expand the j occurrences into words $w \in \{L,R\}^j$.

Fix a word $w \in \{L,R\}^j$. Start with two vertices A_0, B_0. At step ℓ, if $w_\ell = L$, introduce a new vertex A_ℓ and set $B_\ell = B_{\ell-1}$. If $w_\ell = R$, introduce a new vertex B_ℓ and set $A_\ell = A_{\ell-1}$. The pairs $e_\ell = \{A_\ell, B_\ell\}$, for $0 \le \ell \le j$, are the edges of a tree T_w. Each step keeps one endpoint of the preceding edge and adds one new vertex.

Define a second multigraph on the same vertices. Its first edge is the initial D matrix element between A_0 and B_0. Each of its next j edges is the C matrix element between a replaced vertex and the new vertex that replaces it. Its last edge is the final D matrix element between A_j and B_j. Every vertex has degree two, and the multigraph is connected. It is therefore a cycle. When $j = 0$, this cycle consists of two parallel D edges.

Orient each operator edge from the row index to the column index of its matrix element. Every vertex then has one incoming and one outgoing operator edge, so the operator cycle is directed. Choose $v_0, v_1, \dots, v_{j+1}$ in this directed order, and let $O_{w,k}$ be the operator on the edge from v_k to v_{k+1}, with indices read modulo $j+2$. Exactly two of the $O_{w,k}$ are D, and the remaining j are C. If I_Ω is the eigenbasis index set and $\iota : V(T_w) \to I_\Omega$ is an assignment, direct expansion of the matrix products gives

$$q_j = 2 \sum_{w \in \{L,R\}^j} \sum_{\iota : V(T_w) \to I_\Omega} K_w(\omega_\iota) \prod_{k=0}^{j+1} (O_{w,k})_{\iota(v_k),\iota(v_{k+1})}. \tag{10}$$

The $j+1$ factors $\mathcal{L}_\Omega^{-1}$ contribute 2^{j+1}, while the j left-right averages contribute 2^{-j}. This gives the factor two in Equation (10). Substituting $E = \Omega^{1/2} C \Omega^{1/2}$ and $\Delta^s = \Omega^{1/2} D \Omega^{1/2}$ shows that the scalar eigenvalue factor for a fixed assignment of eigenbasis indices to the vertices is

$$K_w(\omega) = \frac{\prod_{v \in V(T_w)} \omega_v}{\prod_{e=\{u,v\} \in E(T_w)} (\omega_u + \omega_v)}. \tag{11}$$

Indeed, the operator cycle has degree two at every vertex, so the square-root factors from its incident operators multiply to one power of ω_v.

Apply Lemma 10 to T_w, with $m = j+1$. For $e = \{u,v\}$, write $\theta_{e,u} = \theta_e$ and $\theta_{e,v} = 1 - \theta_e$ after choosing an order of its endpoints, and put $z_v = \sum_{e \ni v} \theta_{e,v}$ and $h_v = 1 - z_v$. Then $h_v \ge 0$ and $\sum_v h_v = |V(T_w)| - |E(T_w)| = 1$. Equation (11) becomes

$$K_w(\omega) = \prod_v \omega_v^{h_v} \prod_{e=\{u,v\}} \frac{\omega_u^{\theta_{e,u}} \omega_v^{\theta_{e,v}}}{\omega_u + \omega_v}. \tag{12}$$

The formula remains valid when two vertices are assigned the same eigenvalue index.

For each ordered edge $e = (u,v)$,

$$\frac{\omega_u^{\theta_e} \omega_v^{1-\theta_e}}{\omega_u + \omega_v} = g_{\theta_e}(\log \omega_u - \log \omega_v).$$

Apply Lemma 11 to each edge of T_w. The allowed range of θ_e bounds the total variation of the Fourier measure for one edge by $C_F(1+\log(2j+4))$.

Let t_e be the Fourier frequency assigned to edge e. For each vertex v, define $t_v = \sum_{e \ni v} \sigma_{v,e} t_e$, where $\sigma_{u,e} = 1$ and $\sigma_{v,e} = -1$ for the chosen order $e = (u, v)$. The Fourier factors then combine to $\prod_v \omega_v^{it_v}$, and $\sum_v t_v = 0$.

Insert these Fourier integrals into Equation (10). For each word w, the sum over eigenbasis indices becomes the cyclic trace in Equation (13). Combine the measures for all $j + 1$ tree edges, sum over the 2^j words, and include the remaining factor two. This gives a finite complex measure μ_j with $\|\mu_j\|_{\mathrm{TV}} \le V_j$. Here $\|\mu_j\|_{\mathrm{TV}}$ is the total mass of the variation measure $|\mu_j|$. The integration variables are the word and all of its frequencies.

For each choice ζ of a word and its Fourier frequencies, define

$$F_\zeta(x, y) = \mathrm{Tr}\left[\Omega^{h_0+it_0} O_0 \Omega^{h_1+it_1} O_1 \cdots \Omega^{h_{j+1}+it_{j+1}} O_{j+1}\right], \tag{13}$$

where the O_k consist of two copies of D and j copies of C, and the nonnegative exponents satisfy $\sum_{k=0}^{j+1} h_k = 1$. When $h_i = 0$, functional calculus defines Ω^{it_i} as a unitary operator on $\mathrm{supp}\,\Omega$. Noncommutative Hölder gives a bound that does not depend on the dimension. If $h_i > 0$, then $\|\Omega^{h_i+it_i}\|_{1/h_i} = 1$ because $\mathrm{Tr}\,\Omega = 1$. If $h_i = 0$, then $\|\Omega^{it_i}\|_{\mathrm{op}} = 1$. Using the operator norm for every O_i gives $|F_\zeta(x, y)| \le \|D\|_{\mathrm{op}}^2 \|C\|_{\mathrm{op}}^j \le M_j$.

Now use $\Omega_{x,y} = \alpha_x \otimes \beta_y$. Every complex power separates as $\Omega_{x,y}^z = \alpha_x^z \otimes \beta_y^z$. Expand the product expansions of the two D factors and the j C factors in Equation (13). Enumerate the $R_j = r_D^2 r_C^j$ tuples of product indices by $\ell = 1, \ldots, R_j$. Then

$$F_\zeta(x, y) = \sum_{\ell=1}^{R_j} F_{\zeta,\ell}^X(x) F_{\zeta,\ell}^Y(y),$$

where the two factors are the corresponding traces on X and Y. Hence the complex matrix $[F_\zeta(x, y)]_{x,y}$ has rank at most R_j.

The Fourier expansion gives the exact identity $q_j(x, y) = \int F_\zeta(x, y)\, d\mu_j(\zeta)$. Write $d\mu_j = \xi\, d|\mu_j|$, where $|\xi| = 1$. Draw the word and all of its frequencies jointly from the probability measure $|\mu_j|/\|\mu_j\|_{\mathrm{TV}}$. Let $\zeta_1, \ldots, \zeta_T$ be independent samples, and set

$$\widetilde{q}_j = \mathrm{Re}\left[\frac{\|\mu_j\|_{\mathrm{TV}}}{T} \sum_{k=1}^{T} \xi(\zeta_k) F_{\zeta_k}\right].$$

For each entry, the real summands have absolute value at most $V_j M_j$. Hoeffding's inequality and a union bound over the N^2 entries show that $\|q_j - \widetilde{q}_j\|_{\max} \le \eta$ with positive probability whenever $T \ge 6 V_j^2 M_j^2 \eta^{-2} \log(2N)$. Finally, the real part of a complex rank-R_j matrix has real rank at most $2R_j$. The sampled sum therefore has real rank at most $2R_j T$, which proves Equation (9).

All formulas were written on $\mathrm{supp}\,\Omega$. If Ω is singular, the support restriction removes zero eigenvalues before the Fourier expansion. The powers still separate into functions of α_x and β_y, and the Hölder bound is unchanged. Thus no continuity argument or extra factor depending on the smallest nonzero eigenvalue is needed. □

Corollary 13 (Moment rank in terms of d)**.** *For the routing construction above, there is an absolute constant C_0 such that for every $j \ge 0$ and every $0 < \eta \le 1$,*

$$\mathrm{rank}_\eta(q_j) \le \exp\left(C_0(j + 1)\big(\log(2d) + \log\log(j + 3)\big)\right) \eta^{-2} \log(2N).$$

In particular, the bound is $\exp(O(d \log(2d)))\eta^{-2} \log(2N)$ uniformly for $j = O(d)$.

Proof. Use Lemma 8 in Lemma 12: $r_D \le d^2$, $r_C \le d^2 + 1$, and $K_D, K_C \le d^2$. The logarithms of R_j and M_j are $O((j + 1) \log(2d))$, while $\log V_j = O((j + 1) \log\log(j + 3))$. Increasing the absolute constant C_0 covers the ceiling and all fixed factors. □

The last step of this part approximates A^{-1} by a polynomial in A. The eigenvalues of A lie in $[a, b]$ by Equation (7). A Chebyshev polynomial gives the approximation on that interval, and the bounds above apply to the resulting terms. The outcome is Proposition 15: the matrix of the statistic is within $g/2$, entry by entry, of a matrix of rank $\exp(O(d \log(2d))) \log(2N)$.

Lemma 14 (Chebyshev approximation of the reciprocal)**.** *For every $0 < \delta < 1$, there is a polynomial p_m of degree $m = O(\sqrt{b/a}\log\frac{2}{\delta}) = O(d\log\frac{2}{\delta})$ such that $\sup_{\lambda\in[a,b]}|1-\lambda p_m(\lambda)| \le \delta$. It can be chosen as follows: for $k = m+1$, set*

$$r_k(\lambda) = \frac{T_k((b+a-2\lambda)/(b-a))}{T_k((b+a)/(b-a))}, \qquad p_m(\lambda) = \frac{1-r_k(\lambda)}{\lambda}.$$

If $p_m(1+z) = \sum_{j=0}^m c_j z^j$, then $\sum_{j=0}^m |c_j| \le \exp(O(m))$, with an absolute implied constant.

Proof. This is the standard Chebyshev construction [Saa03, Vis12]. Among degree-k polynomials that equal one at zero, r_k minimizes the largest absolute value on $[a,b]$. Its degree bound follows from the growth of T_k outside $[-1,1]$. The coefficient bound is proved here because no reference states it after the required shift by one.

The affine map $t(\lambda) = (b+a-2\lambda)/(b-a)$ sends $[a,b]$ to $[1,-1]$. Put $\kappa = b/a = 19d^2+1$ and $c = (b+a)/(b-a) = (\kappa+1)/(\kappa-1) > 1$. Since $r_k(0) = 1$, the numerator $1 - r_k(\lambda)$ is divisible by λ. Thus p_m is a polynomial of degree at most $k-1 = m$, and $1-\lambda p_m(\lambda) = r_k(\lambda)$. On $[a,b]$, $|r_k(\lambda)| \le 1/T_k(c)$. The growth estimate for Chebyshev polynomials gives $1/T_k(c) \le \delta$ once $k \ge (\sqrt{\kappa}/2)\log(2/\delta)$. This proves the approximation bound.

It remains to bound the coefficients after shifting by one. For a polynomial H, let $\|H\|_{\ell_1}$ denote the sum of the absolute values of its coefficients. Write $t(1+z) = \xi - \nu z$ with $\xi = 1 - 2/d^2$ and $\nu = 2/(0.95d^2)$. For every $d \ge 1$, this affine polynomial has coefficient norm less than four. Let $H_\ell(z) = T_\ell(t(1+z))$. The recurrence $T_{\ell+1} = 2tT_\ell - T_{\ell-1}$ gives $\|H_{\ell+1}\|_{\ell_1} \le 8\|H_\ell\|_{\ell_1} + \|H_{\ell-1}\|_{\ell_1}$. Starting from $H_0 = 1$ and $\|H_1\|_{\ell_1} < 4$, induction gives $\|H_k\|_{\ell_1} \le 9^k$.

Since $T_k(c) \ge 1$, the polynomial $R(z) = 1 - r_k(1+z)$ has coefficient norm at most $1+9^k$. Also $R(-1) = 0$, so $1+z$ divides $R(z)$ and $R(z) = (1+z)P(z)$ with $P(z) = p_m(1+z)$. If $R(z) = \sum_{\ell=0}^k a_\ell z^\ell$, the quotient coefficients satisfy $c_j = \sum_{\ell=0}^j (-1)^{j-\ell} a_\ell$. Therefore $\sum_{j=0}^{k-1}|c_j| \le k\sum_{\ell=0}^k |a_\ell| \le k(1+9^k) = \exp(O(k))$. Since $k = m+1$, this proves the coefficient estimate. □

Proposition 15 (Approximate rank of the routing statistic)**.** *There is an absolute constant C_1 such that, for every $N = 2^n$ and every protocol resource of smaller marginal rank d,*

$$\mathrm{rank}_{g/2}([\chi_s(x,y)]_{x,y}) \le \exp(C_1 d\log(2d))\log(2N),$$

where g is the constant in Equation (6)*.*

Proof. Choose $\delta = g/8$ in Lemma 14. The resulting degree satisfies $m = O(d)$. Since the eigenvalues of A lie in $[a,b]$, Equation (8) gives $|\langle u, A^{-1}u\rangle_{\mathsf{HS}} - \langle u, p_m(A)u\rangle_{\mathsf{HS}}| \le \delta\langle u, A^{-1}u\rangle_{\mathsf{HS}} \le \delta$. With $A = I+B$, Lemma 9 gives $\langle u, p_m(A)u\rangle_{\mathsf{HS}} = \sum_{j=0}^m c_j q_j$. Thus the polynomial approximation contributes at most $2\delta = g/4$ to the error in χ_s.

For each j, set $\eta_j = g/(8(m+1)\max\{1,|c_j|\})$, and choose a real matrix $\widetilde{q}_j$ with $\|q_j - \widetilde{q}_j\|_{\max} \le \eta_j$ and rank bounded by Corollary 13. The matrix $K = 2\sum_{j=0}^m c_j\widetilde{q}_j$ satisfies

$$\|[\chi_s(x,y)] - K\|_{\max} \le 2\delta + 2\sum_{j=0}^m |c_j|\eta_j \le \frac{g}{4} + \frac{g}{4} = \frac{g}{2}.$$

Lemma 14 gives $\max_j|c_j| \le \exp(O(m))$. Hence $\eta_j^{-2} \le C_g(m+1)^2\exp(O(m))$, where C_g depends only on the fixed constant g. Apply Corollary 13 to every $j \le m = O(d)$ and add the ranks of the $m+1$ matrices. This gives $\mathrm{rank}\, K \le \exp(C_1 d\log(2d))\log(2N)$ for an absolute constant C_1. □

5. Inner product modulo 2

This section completes the proof. The preceding sections show that every correct protocol gives a low-rank matrix. Its entries lie on opposite sides of a fixed threshold according to whether $f(x,y) = 0$ or $f(x,y) = 1$. We now use the inner product modulo 2. Its sign matrix has sign rank at least $\sqrt{N}$. Comparing this lower bound with the earlier rank upper bound proves Theorem 1.

A sign matrix has entries in $\{-1,+1\}$. Its sign rank is the smallest rank of a real matrix H whose entries have the same signs:

$$\mathrm{signrank}(S) = \min\{\mathrm{rank}_{\mathbb{R}} H : S_{x,y}H_{x,y} > 0 \text{ for all } x,y\}.$$

Let $N = 2^n$, and label the rows and columns by n-bit strings. We use $f_n(x,y) = \mathrm{IP}_n(x,y) = \bigoplus_{i=1}^n x_i y_i$. Its sign matrix is the Walsh–Hadamard matrix $W_N(x,y) = (-1)^{\mathrm{IP}_n(x,y)}$.

Lemma 16 (Sign rank of inner product modulo 2)**.** *For every $n \geq 1$,* $\mathrm{signrank}(W_N) \geq \sqrt{N} = 2^{n/2}$*.*

Proof. The rows of W_N are pairwise orthogonal, and each has squared length N. Hence $W_N W_N^{\mathsf{T}} = NI$ and its spectral norm is $\|W_N\|_{\mathrm{op}} = \sqrt{N}$. The bound in [For02] states that $\mathrm{signrank}(S) \geq N/\|S\|_{\mathrm{op}}$ for every $N \times N$ sign matrix S. Applying it to W_N proves the claim. ☐

The target sign matrix is $S_N(x,y) = 2f_n(x,y) - 1 = -W_N(x,y)$. Multiplying every entry by -1 does not change sign rank. Therefore $\mathrm{signrank}(S_N) \geq \sqrt{N}$.

Proof of Theorem 1. Fix a protocol for f_n with error at most 0.09. Let d be the smaller marginal rank of its shared state. By Proposition 15, there is a real matrix K with $\|K - [\chi_s(x,y)]\|_{\max} \leq g/2$. It also satisfies $\mathrm{rank}\, K \leq \exp(C_1 d \log(2d)) \log(2N)$. The bounds in Equations (5) and (6) show that $K_{x,y} < \theta$ when $f_n(x,y) = 0$ and $K_{x,y} > \theta$ when $f_n(x,y) = 1$. Let J be the all-ones matrix. Then $K - \theta J$ has the same entrywise signs as S_N. Consequently,

$$2^{n/2} \leq \mathrm{signrank}(S_N) \leq \mathrm{rank}(K - \theta J) \leq \exp(C_1 d \log(2d)) \log(2N) + 1.$$

For all sufficiently large n, the left side is at least two, so $2^{n/2-1} \leq \exp(C_1 d \log(2d)) \log(2^{n+1})$. Taking natural logarithms gives $\frac{n \log 2}{2} + O(1) \leq C_1 d \log(2d) + \log(n+1) + O(1)$. The term $\log(n+1)$ and the bounded terms grow more slowly than n. Hence, for all sufficiently large n, $d \log_2(2d) \geq cn$ for an absolute constant $c > 0$.

If $d \geq n$, the claimed bound follows from $\log_2 d \geq \log_2 n$. Now suppose that $d < n$. For $n \geq 2$, $\log_2(2d) \leq 1 + \log_2 n \leq 2\log_2 n$, so $d \geq cn/(2\log_2 n)$. Taking logarithms gives $\log_2 d \geq \log_2 n - \log_2 \log_2 n - O(1)$. This is the claimed lower bound on E_{dim}. ☐

6. Conclusion

For f-routing with the inner product modulo 2, every protocol with worst-case error at most 0.09 satisfies

$$d \log_2(2d) = \Omega(n), \qquad d = \min\{\mathrm{rank}\, \rho_L, \mathrm{rank}\, \rho_R\}, \qquad E_{\mathrm{dim}}(\rho_{LR}) \geq \log_2 n - \log_2 \log_2 n - O(1).$$

Hence the smaller marginal rank is $\Omega(n/\log n)$. The result allows arbitrary mixed shared states, unrestricted message lengths in the single simultaneous exchange, and local systems of arbitrary finite dimension. Error is allowed in both routing cases.

The central technical step is an approximate-rank bound of $\exp(O(d \log(2d))) \log(2N)$. Its resource dependence is only through d, and it is independent of message and workspace dimensions. Only the final sign-rank step uses the inner product modulo 2. The cost applies to the entire shared state prepared before the inputs arrive. For mixed states, it is not an entanglement measure and may charge separable correlations or shared randomness. The result does not give a superlogarithmic lower bound on E_{dim}. Proving such a bound, and in particular a bound polynomial in n, remains open.

References

[ABM+24] R. Allerstorfer, H. Buhrman, A. May, F. Speelman, and P. Verduyn Lunel, *Relating non-local quantum computation to information theoretic cryptography*, Quantum **8** (2024), 1387.

[ACCM26] V. R. Asadi, R. Cleve, E. Culf, and A. May, *Linear gate bounds against natural functions for position-verification*, arXiv:2402.18648v4, 2026.

[ACM25] V. R. Asadi, E. Culf, and A. May, *Rank lower bounds on non-local quantum computation*, in *16th Innovations in Theoretical Computer Science Conference*, LIPIcs **325** (2025), 11:1–11:18.

[AKL+25] V. R. Asadi, K. Kuroiwa, D. Leung, A. May, S. Pasterski, and C. Waddell, *Conditional disclosure of secrets with quantum resources*, Quantum **9** (2025), 1885.

[BCS22] A. Bluhm, M. Christandl, and F. Speelman, *A single-qubit position verification protocol that is secure against multi-qubit attacks*, Nature Physics **18** (2022), 623–626.

[BHM+26] A. Bluhm, S. Höfer, A. May, M. Stasiuk, P. Verduyn Lunel, and H. Yuen, *A complexity theory for non-local quantum computation*, Quantum **10** (2026), 2152.

[BV04] S. Boyd and L. Vandenberghe, *Convex optimization*, Cambridge University Press, 2004.

[CM26] R. Cleve and A. May, *Lower bounds on non-local computation from controllable correlation*, arXiv:2602.00255, 2026.

[FF56] L. R. Ford Jr. and D. R. Fulkerson, *Maximal flow through a network*, Canadian Journal of Mathematics **8** (1956), 399–404.

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

[GMOW26] U. Girish, A. May, L. Orshansky, and C. Waddell, *Comparing classical and quantum conditional disclosure of secrets*, Quantum **10** (2026), 2049.

[Hay02] M. Hayashi, *Optimal sequence of quantum measurements in the sense of Stein's lemma in quantum hypothesis testing*, Journal of Physics A **35** (2002), 10759–10773.

[HP14] F. Hiai and D. Petz, *Introduction to matrix analysis and applications*, Universitext, Springer, 2014.

[KSW08] D. Kretschmann, D. Schlingemann, and R. F. Werner, *The information-disturbance tradeoff and the continuity of Stinespring's representation*, IEEE Transactions on Information Theory **54** (2008), 1708–1717.

[Liu25] Y. Liu, *Quantum state testing beyond the polarizing regime and quantum triangular discrimination*, Computational Complexity **34** (2025), 11.

[May26] A. May, *Entanglement cost in non-local quantum computation*, arXiv:2605.02840, 2026.

[OH04] T. Ogawa and M. Hayashi, *On error exponents in quantum hypothesis testing*, IEEE Transactions on Information Theory **50** (2004), 1368–1372.

[DLMF] F. W. J. Olver et al. (eds.), *NIST Digital Library of Mathematical Functions*, `https://dlmf.nist.gov`.

[Saa03] Y. Saad, *Iterative methods for sparse linear systems*, second edition, SIAM, 2003.

[TKR+10] K. Temme, M. J. Kastoryano, M. B. Ruskai, M. M. Wolf, and F. Verstraete, *The χ^2-divergence and mixing times of quantum Markov processes*, Journal of Mathematical Physics **51** (2010), 122201.

[Vis12] N. K. Vishnoi, *$Lx = b$: Laplacian solvers and their algorithmic applications*, Foundations and Trends in Theoretical Computer Science **8**, nos. 1–2 (2012), 1–141, `https://doi.org/10.1561/0400000054`.