Quantum Log-Approximate-Rank Conjecture is also False

Anurag Anshu*†, Naresh Goud Boddu‡ and Dave Touchette*†§
*Institute for Quantum Computing and Department of Combinatorics and Optimization, University of Waterloo
†Perimeter Institute for Theoretical Physics, Waterloo ON
‡Center for Quantum Technologies, National University of Singapore
§Département d'informatique & Institut Quantique, Université de Sherbrooke
Email: aanshu@uwaterloo.ca, e0169905@u.nus.edu, touchette.dave@gmail.com

***Abstract*—In a recent breakthrough result, Chattopadhyay, Mande and Sherif [ECCC TR18-17] showed an exponential separation between the log approximate rank and randomized communication complexity of a total function f, hence refuting the log approximate rank conjecture of Lee and Shraibman [2009]. We provide an alternate proof of their randomized communication complexity lower bound using the information complexity approach. Using the intuition developed there, we derive a polynomially-related quantum communication complexity lower bound using the quantum information complexity approach, thus providing an exponential separation between the log approximate rank and quantum communication complexity of f. Previously, the best known separation between these two measures was (almost) quadratic, due to Anshu, Ben-David, Garg, Jain, Kothari and Lee [CCC, 2017]. This settles one of the main question left open by Chattopadhyay, Mande and Sherif, and refutes the quantum log approximate rank conjecture of Lee and Shraibman [2009]. Along the way, we develop a Shearer-type protocol embedding for product input distributions that might be of independent interest.**

***Keywords*-quantum communication complexity; logarithm of approximate rank; exponential separation; information theory; information complexity;**

I. Introduction

Communication complexity concerns itself with characterizing the minimum number of bits that distributed parties need to exchange in order to accomplish a given task (such as computing a function F). Over the years, it has established striking connections with various areas of complexity theory and information theory, providing tools for solving central problems in such domains. Since it is in general hard to pin down precisely the communication cost of a task, various lower bound methods have been developed over the years. One such method is the logarithm of the rank of the matrix M_F that encodes the values the function F takes on various inputs. More precisely, this matrix is defined as $M_F(x, y) = F(x, y)$. The following well known conjecture posits that this lower bound is polynomially tight for the deterministic communication complexity of F.

Conjecture 1 (Log-Rank Conjecture, [1])**.** *There exists a universal constant α such that the deterministic communication complexity of every total Boolean function F is $\mathcal{O}(\log^{\alpha}(\mathrm{rk}(M_F)))$.*

See Ref. [2] and reference therein for more details about this and the other conjectures discussed in this work. A natural randomized analogue of Conjecture 1 is the following, comparing randomized communication complexity to the logarithm of the approximate rank rather than actual rank of M_F. (See Section II-B for definitions.)

Conjecture 2 (Log-Approximate-Rank Conjecture, [3])**.** *There exists a universal constant α such that the randomized communication complexity (with error $\frac{1}{3}$) of every total Boolean function F is $O(\log^{\alpha}(\mathrm{rk}_{1/3}(M_F)))$.*

In a recent breakthrough work [2], Chattopadhyay, Mande and Sherif established that Conjecture 2 is false by exhibiting a function with an exponential separation between the randomized communication complexity (with constant error) and Log-Approximate-Rank. Their function is a composition of the 2-bit Xor function and a function that they call Sink. The work [2] asked if their function had implications for the following quantum version of Conjecture 2.

Conjecture 3 (Quantum Log-Approximate-Rank Conjecture, [3])**.** *There exists a universal constant α, such that the quantum communication complexity of every total Boolean function F is $O(\log^{\alpha}(\mathrm{rk}_{1/3}(M_F)))$.*

Here we prove that Conjecture 3 is false as well. Before proceeding to the statement of our main result, we define the Sink function.

Definition 4 (Sink, [2])**.** *Sink function is defined on a complete directed graph of m vertices, using $\binom{m}{2}$ variables $z_{i,j}, i < j \in [m]$, in the following way. Let $z_{i,j} = 1$ if there is a directed edge from vertex v_i to v_j and $z_{i,j} = 0$ if there is a directed edge from vertex v_j to v_i. The function Sink computes whether or not there is a sink in the graph. In other words, Sink$(z) = 1$ iff $\exists i \in [m]$ such that all edges adjacent to v_i are incoming.*

The function of interest for communication complexity is $\mathsf{Sink} \circ \mathsf{Xor}^{\otimes\binom{m}{2}}$, where each Xor takes as input one bit from Alice and one from Bob. For simplicity of notation, we will denote this function as Sink ∘ Xor. Our main theorem is as follows, which lower bounds the quantum information complexity (QIC) of Sink ∘ Xor.

2575-8454/19/$31.00 ©2019 IEEE
DOI 10.1109/FOCS.2019.00063

Theorem 5. *Any t-round entanglement assisted protocol for* $\mathsf{Sink} \circ \mathsf{Xor}$ *achieving error* $1/5$ *satisfies* $\mathrm{QIC}(\Pi, \mu^{\otimes\binom{m}{2}}) \in \Omega(\frac{m}{t^2})$*, with* μ *being the uniform distribution on 1+1 bits* [1].

The desired lower bound on entanglement assisted quantum communication complexity ($\mathrm{Q}^*_{\frac{1}{3}}$) of $\mathsf{Sink} \circ \mathsf{Xor}$ follows by optimizing $\max(t, m/t^2)$ over the number of round t.

Corollary 6. *It holds that* $\mathrm{Q}^*_{1/3}(\mathsf{Sink} \circ \mathsf{Xor}) \in \Omega(m^{1/3})$.

Hence, combining with the following upper bound on the log-approximate-rank due to Ref. [2], the $\mathsf{Sink} \circ \mathsf{Xor}$ function witnesses an exponential separation between log-approximate-rank and quantum communication, and refutes the quantum log-approximate-rank conjecture of Lee and Shraibman [3].

Theorem 7 ([2])**.** *It holds that*

1) $\log \mathrm{rk}_{1/3}(M_{\mathsf{Sink}\circ\mathsf{Xor}}) \leq 4 \log m + o(\log m)$
2) $\log \mathrm{rk}^+_{1/3}(M_{\mathsf{Sink}\circ\mathsf{Xor}}) = O(\log^2 m)$.

In a subsequent version of [2], Chattopadhyay et. al. improved the upper bound on $\log \mathrm{rk}^+_{1/3}(M_{\mathsf{Sink}\circ\mathsf{Xor}})$ to $O(\log m)$.

A. Independent work

Sinha and de Wolf [4] used the fooling distribution method, in independent and simultaneous work, to obtain the same $\Omega(m^{1/3})$ lower bound on the quantum communication complexity of $\mathsf{Sink} \circ \mathsf{Xor}$. This differs from our techniques which we describe below.

B. Proof overview

At a high-level, our argument follows the well-established *information complexity* approach [5], [6], [7], [8], [9]. We view a given function f as some composition of many instances of a simpler component function g, and argue through a direct sum property a reduction from g to f. This is achieved by embedding inputs to g into inputs to f, where the remaining inputs to f are sampled from some suitable distribution in order to achieve the desired direct sum property. Following this, we show a lower bound on the information complexity for g.

In the present context, $\mathsf{Sink}\circ\mathsf{Xor}$ is a composition of many instances of the Equality function, in a way that the input bits are shared across the instances. In Ref. [2], the authors use Shearer's lemma to handle such overlap between the inputs across the instances and derive a corruption lower bound. For the reduction from $\mathsf{Sink} \circ \mathsf{Xor}$ to Equality, we also wish to use a Shearer-type inequality. We further argue that a lower bound on information complexity of Equality (for protocols that make small error in the worst case) under uniform distribution implies a lower bound on information complexity of $\mathsf{Sink} \circ \mathsf{Xor}$. But it is not clear, a priori, that Equality should have high information cost under that distribution, as this function has trivial communication complexity under the uniform distribution. It turns out that the cut-and-paste argument of Anshu, Belovs, Ben-David, Göös, Jain, Kothari, Lee and Santha [10] yields a constant lower bound on information complexity of good protocols for Equality, even under the uniform distribution.

Broadly, our quantum lower bound proceeds along lines similar to above. The quantum cut-and-paste argument of Anshu, Ben-David, Garg, Jain, Kothari and Lee [11] in the quantum setting yields a round dependent lower bound on the *quantum information complexity* (QIC) [5], [12], [13], [14], [15] of good protocols for Equality, even under the uniform distribution. But the quantum version of the embedding argument requires new methods. In the classical setting, using classical information cost IC, as soon as we have Alice and Bob privately sample the remaining inputs, the Shearer-type embedding follows almost directly from a Shearer like inequality for information [16]. In the quantum setting, we would similarly like to use a Shearer-type inequality for quantum information [17]. However, it is not immediately clear how to make the protocol embedding work for quantum information cost QIC. We instead settle on an alternate notion of quantum information cost (variants of which have appeared before [18], [13], [19], [17]) that works well only for product input distributions. The argument then goes through by carefully using this notion, and it is equivalent to QIC up to a round-dependent factor. What we get is a Shearer-type embedding protocol for product input distributions that allows some specific pre-processing of the inputs. We provide such a general version in Section IV-A in the quantum setting, while we give a more direct proof in the classical setting.

Hence, overall we get a round dependent lower bound on the quantum information complexity of $\mathsf{Sink} \circ \mathsf{Xor}$, and the round independent lower bound on quantum communication complexity follows by optimizing over the number of rounds in any good protocol.

II. Preliminaries and notation

For integer $n \geq 1$, let $[n]$ represent the set $\{1, 2, ..., n\}$. Let $\mathcal{X}$ and $\mathcal{Y}$ be finite sets and k be a natural number. Let $\mathcal{X}^k$ be the set $\mathcal{X} \times ... \times \mathcal{X}$, the cross product of $\mathcal{X}$, k times. Let μ be a probability distribution on $\mathcal{X}$. Let $\mu(x)$ represent the probability of $x \in \mathcal{X}$ according to μ. We write $X \sim \mu$ to denote that the random variable X is distributed according to μ. We use the same symbol to represent a random variable and its distribution whenever it is clear from the context. The expectation value of function f on X is defined as $\mathbb{E}_{x \leftarrow X}[f(x)] = \sum_{x \in \mathcal{X}} \Pr(X = x) f(x)$ where $x \leftarrow X$ means that x is drawn according to the distribution of X. We say X and Y are independent iff for each $x \in \mathcal{X}, y \in \mathcal{Y} : \Pr(XY = xy) = \Pr(X = x) \cdot \Pr(Y = y)$. For joint

[1] A random variable on $a + b$ bits takes values over a bits on Alice's side and b bits on Bob's side.

random variables XY, Y^x will denote the distribution of $Y|X=x$.

A. Classical information theory

We start with the following fundamental information theoretic quantities. We refer the reader to the excellent sources for information theory [20] for further study.

Definition 8 (Entropy)**.** *Let random variable X take values in $\mathcal{X}$. Then entropy of random variable X is defined as*

$$H(X) = \sum_{x \in \mathcal{X}} \Pr(X=x) \log\left(\frac{1}{\Pr(X=x)}\right).$$

Definition 9 (Mutual Information)**.** *Let random variable X take values in $\mathcal{X}$ and Y take values in $\mathcal{Y}$. Then the mutual information between random variables X and Y is defined as*

$$I(X:Y) = H(X) + H(Y) - H(XY).$$

The conditional entropy of X given Y is

$$H(X|Y) = \sum_{y \in \mathcal{Y}} \Pr(Y=y) H(X|Y=y).$$

The conditional mutual information between X and Y conditioned on Z is

$$I(X:Y|Z) = \sum_{z \in \mathcal{Z}} \Pr(Z=z) I(X:Y|Z=z).$$

Let μ and ν be two distributions over the set $\mathcal{X}$. The Kullback-Leibler divergence between μ and ν is defined as

$$\mathrm{D}(\mu\|\nu) = -\sum_{x \in \mathcal{X}} \mu(x) \log\left(\frac{\nu(x)}{\mu(x)}\right).$$

The following holds:

$$I(X:Y) = \sum_{y \in \mathcal{Y}} \Pr(Y=y) \mathrm{D}(X^y\|X).$$

Definition 10 (Distance measures)**.** *Let μ and ν be probability distributions over $\mathcal{X}$. We define the following distance measures between distributions.*

Total variation distance: $\Delta(\mu,\nu) = \frac{1}{2}||\mu-\nu||_1 = \max_{T \subseteq \mathcal{X}}[\mu(T)-\nu(T)] = \frac{1}{2}\sum_{x \in \mathcal{X}} |\mu(x)-\nu(x)|$.

Bures metric (Hellinger distance):

$$B(\mu,\nu) := \sqrt{\sum_{x \in \mathcal{X}} \left(\frac{\mu(x)+\nu(x)}{2} - \sqrt{\mu(x)\nu(x)}\right)}.$$

Fact 11 (Relationship between Hellinger distance and Total variation distance)**.** *Let μ, ν be two distributions over $\mathcal{X}$. It holds that,*

$$\frac{1}{\sqrt{2}}\Delta(\mu,\nu) \leq B(\mu,\nu) \leq \sqrt{\Delta(\mu,\nu)}.$$

Fact 12 (Triangle Inequality)**.** *Let μ, ν, η be 3 distributions over $\mathcal{X}$. It holds that,*

$$\Delta(\mu,\eta) \leq \Delta(\mu,\nu) + \Delta(\nu,\eta).$$

Fact 13 (Pinsker's inequality([20], Lemma 11.6.1, p. 370))**.** *Let μ and ν be two distributions over the set $\mathcal{X}$. It holds that,*

$$\Delta^2(\mu,\nu) \leq \frac{\ln 2}{2} \cdot \mathrm{D}(\mu\|\nu).$$

We now review some properties of the Bures metric.

Fact 14 (Facts about Bures metric)**.**

Fact 14.A (Triangle inequality [21])**.** *The following triangle inequality and a weak triangle inequality hold for the Bures metric and the square of the Bures metric.*

1) *Let μ, ν, η be 3 distributions over $\mathcal{X}$. It holds that,*

$$\mathrm{B}(\mu,\eta) \leq \mathrm{B}(\mu,\nu) + \mathrm{B}(\nu,\eta).$$

2) *Let μ^1, μ^2, ..., μ^{t+1} be $t+1$ distributions over $\mathcal{X}$. It holds that,*

$$\mathrm{B}^2(\mu^1,\mu^{t+1}) \leq t \cdot \sum_{i=1}^{t} \mathrm{B}^2(\mu^i,\mu^{i+1}).$$

Fact 14.B (Averaging over register)**.** *For 2 joint random variables $\theta_{XB}, \theta'_{XB}$ with same marginal distribution $\theta_X = \theta'_X$, we have*

$$\mathrm{B}^2(\theta_{XB},\theta'_{XB}) = \mathbb{E}_{x \leftarrow X}[\mathrm{B}^2(\theta_B^x,\theta_B'^x)].$$

Finally, an important property of both these distance measures is monotonicity.

Fact 15 (Monotonicity)**.** *For any 2 joint random variables μ_{XY}, ν_{XY} over $\mathcal{X} \times \mathcal{Y}$, we have*

$$\Delta(\mu_{XY},\nu_{XY}) \geq \Delta(\mu_X,\nu_X) \quad \text{and}$$
$$\mathrm{B}(\mu_{XY},\nu_{XY}) \geq \mathrm{B}(\mu_X,\nu_X).$$

We state the following classical version of the Shearer-type inequality for information.. (See Lemma 36.)

Lemma 16. *Let XZ be a correlated random variable such that $X = X_1 \otimes X_2 \otimes .... \otimes X_n$. If S is an independant random variable distributed on subsets of the coordinates $[n]$, such that for every $i \in [n]$, $\Pr[i \in S] \leq \frac{1}{k}$, then*

$$\mathbb{E}_S[I(X_S : Z|S)] \leq \frac{1}{k} I(X;Z)$$

where X_S is the random variable $(X_i : i \in S)$.

B. Classical communication complexity

Let $f : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ be a total function (that is, its value is defined on every input) and $\varepsilon \in (0,1)$. In a two-party communication task, Alice is given an input $x \in \mathcal{X}$, Bob is given $y \in \mathcal{Y}$ and the task is to compute $f(x,y)$ by exchanging as few bits as possible. The parties are allowed to possess pre-shared randomness (R) and private randomness (R_A, R_B). Without loss of generality, we can assume that Alice communicates first and also gives the final output. The communication cost of a protocol Π, denoted by $CC(\Pi)$, is the maximum number of bits the parties have to communicate over all possible inputs and values of the shared and private randomness. Let $R_\varepsilon(f)$ represent the two-party randomized communication complexity of f with worst case error ε, i.e., the communication of the best two-party randomized protocol for f with error at most ε over any input (x,y). Worst-case error of the protocol Π over the inputs is denoted by $\mathrm{err}(\Pi)$. With some abuse of notation, let Π also denote the random variable composed of all the messages of the protocol. We will not include the public randomness R in the protocol and will account for it separately.

Definition 17 (XOR function)**.** *A function $F : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ is called an XOR function if there exists a function $f : \{0,1\}^n \to \{0,1\}$ such that $F(x_1,...,x_n,y_1,...,y_n) = f(x_1 \oplus y_1,...,x_n \oplus y_n)$ for all $x,y \in \{0,1\}^n$. We denote $F = f \circ XOR$.*

Definition 18 (Rank)**.** *The rank of a matrix M, denoted by $\mathrm{rk}(M)$ is the minimum integer k for which there exist k rank 1 matrices such that $M = \sum_{i=1}^{k} M_i$.*

Definition 19 (Non-negative Rank)**.** *The non-negative rank of a matrix M, denoted by $\mathrm{rk}^+(M)$ is the minimum integer k for which there exist k rank 1 matrices with non-negative entries such that $M = \sum_{i=1}^{k} M_i$.*

Definition 20 (Approximate rank)**.** *Let $\varepsilon \in [0,1/2)$ and M be an $|\mathcal{X}| \times |\mathcal{Y}|$ matrix. The ε-approximate rank of M is defined as*

$$\mathrm{rk}_\varepsilon(M) = \min_{\tilde{M}} \{\mathrm{rk}(\tilde{M})\},$$

in which the minimum ranges over all $\tilde{M}$ satisfying

$$\forall x \in \mathcal{X}, y \in \mathcal{Y},\ |\tilde{M}(x,y) - M(x,y)| \leq \varepsilon.$$

Definition 21 (Approximate non-negative rank)**.** *Let $\varepsilon \in [0,1/2)$ and M be an $|\mathcal{X}| \times |\mathcal{Y}|$ matrix. The ε-approximate non-negative rank of M is defined as*

$$\mathrm{rk}^+_\varepsilon(M) = \min_{\tilde{M}} \{\mathrm{rk}^+(\tilde{M})\},$$

in which the minimum ranges over all $\tilde{M}$ satisfying

$$\forall x \in \mathcal{X}, y \in \mathcal{Y},\ |\tilde{M}(x,y) - M(x,y)| \leq \varepsilon.$$

Definition 22 (Distributional Information Complexity)**.** *Distributional information complexity of a randomized protocol Π with respect to a distribution $XY \sim \mu$ is defined as*

$$\mathrm{IC}(\Pi,\mu) = \mathrm{I}(X : \Pi | YRR_B) + \mathrm{I}(Y : \Pi | XRR_A).$$

Definition 23 (Max Distributional Information Complexity)**.** *Max-distributional information complexity of a randomized protocol Π is defined as*

$$\mathrm{IC}(\Pi) = \max_{\mu} \mathrm{IC}(\Pi,\mu).$$

Definition 24 (Information Complexity of a function)**.** *Information complexity of a function f is defined as*

$$\mathrm{IC}(f) = \inf_{\Pi:\mathrm{err}(\Pi)\leq\varepsilon} \mathrm{IC}(\Pi).$$

Note that since one bit of communication can hold at most one bit of information, for any protocol Π and distribution μ we have $\mathrm{IC}(\Pi,\mu) \leq \mathrm{CC}(\Pi)$. This implies that information complexity of a function is a lower bound on the randomized communication complexity of a function.

Lemma 25 (Cut-and-paste lemma (Lemma 6.3 in [7]))**.** *Let (x,y) and (x',y') be two inputs to a randomized protocol Π. Then*

$$\mathrm{B}((R\Pi)^{x,y},(R\Pi)^{x',y'}) = \mathrm{B}((R\Pi)^{x,y'},(R\Pi)^{x',y}).$$

Fact 26 (Pythagorean property (Lemma 6.4 in [7]))**.** *Let (x,y) and (x',y') be two inputs to a randomized protocol Π. Then*

$$\mathrm{B}^2((R\Pi)^{x,y'},(R\Pi)^{x',y'}) + \mathrm{B}^2((R\Pi)^{x,y},(R\Pi)^{x',y}) \leq 2\mathrm{B}^2((R\Pi)^{x',y'},(R\Pi)^{x,y}).$$

C. Quantum information theory

We now introduce some quantum information theoretic notation. We assume the reader is familiar with standard concepts in quantum computing [22], [23], [24].

Let $\mathcal{H}$ be a finite-dimensional complex Euclidean space, i.e., $\mathbb{C}^n$ for some positive integer n with the usual complex inner product $\langle\cdot,\cdot\rangle$, which is defined as $\langle u,v\rangle = \sum_{i=1}^{n} u_i^* v_i$. We will also refer to $\mathcal{H}$ as an Hilbert space. We will usually denote vectors in $\mathcal{H}$ using bra-ket notation, e.g., $|\psi\rangle \in \mathcal{H}$.

The ℓ_1 norm (also called the trace norm) of an operator X on $\mathcal{H}$ is $\|X\|_1 := \mathrm{Tr}(\sqrt{X^\dagger X})$, which is also equal to (vector) ℓ_1 norm of the vector of singular values of X. A *quantum state* (or a *density matrix* or simply a *state*) ρ is a positive semidefinite matrix on $\mathcal{H}$ with $\mathrm{Tr}(\rho) = 1$. The state ρ is said to be a *pure state* if its rank is 1, or equivalently if $\mathrm{Tr}(\rho^2) = 1$, and otherwise it is called a *mixed state*. Let $|\psi\rangle$ be a unit vector on $\mathcal{H}$, that is $\langle\psi|\psi\rangle = 1$. With some abuse of notation, we use ψ to represent the vector $|\psi\rangle$ and

also the density matrix $|\psi\rangle\langle\psi|$, associated with $|\psi\rangle$. Given a quantum state ρ on $\mathcal{H}$, the *support of* ρ, denoted $\text{supp}(\rho)$, is the subspace of $\mathcal{H}$ spanned by all eigenvectors of ρ with nonzero eigenvalues.

A *quantum register* A is associated with some Hilbert space $\mathcal{H}_A$. Define $|A| \coloneqq \log \dim(\mathcal{H}_A)$. Let $\mathcal{L}(A)$ represent the set of all linear operators on $\mathcal{H}_A$. We denote by $\mathcal{D}(A)$ the set of density matrices on the Hilbert space $\mathcal{H}_A$. We use subscripts (or superscripts according to whichever is convenient) to denote the space to which a state belongs, e.g, ρ with subscript A indicates $\rho_A \in \mathcal{H}_A$. If two registers A and B are associated with the same Hilbert space, we represent this relation by $A \equiv B$. For two registers A and B, we denote the combined register as AB, which is associated with Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. For two quantum states $\rho \in \mathcal{D}(A)$ and $\sigma \in \mathcal{D}(B)$, $\rho \otimes \sigma \in \mathcal{D}(AB)$ represents the tensor product (or Kronecker product) of ρ and σ. The identity operator on $\mathcal{H}_A$ is denoted $\mathbb{I}_A$.

Let $\rho_{AB} \in \mathcal{D}(AB)$. We define the *partial trace with respect to* A of ρ_{AB} as

$$\rho_B \coloneqq \text{Tr}_A(\rho_{AB}) \coloneqq \sum_i (\langle i| \otimes \mathbb{I}_B)\rho_{AB}(|i\rangle \otimes \mathbb{I}_B),$$

where $\{|i\rangle\}_i$ is an orthonormal basis for the Hilbert space $\mathcal{H}_A$. The state $\rho_B \in \mathcal{D}(B)$ is referred to as a *reduced density matrix* or a *marginal state*. Unless otherwise stated, a missing register from subscript in a state will represent partial trace over that register. Given $\rho_A \in \mathcal{D}(A)$, a *purification* of ρ_A is a pure state $\rho_{AB} \in \mathcal{D}(AB)$ such that $\text{Tr}_B(\rho_{AB}) = \rho_A$. Any quantum state has a purification using a register B with $|B| \leq |A|$. The purification of a state, even for a fixed B, is not unique as any unitary applied on register B alone does not change ρ_A.

An important class of states that we will consider are the *classical quantum states*. They are of the form $\rho_{AB} = \sum_a \mu(a) |a\rangle\langle a|_A \otimes \rho_B^a$, where μ is a probability distribution. In this case, ρ_A can be viewed as a probability distribution and we shall continue to use the notations that we have introduced for probability distribution, for example, $\mathbb{E}_{a \leftarrow A}$ to denote the average $\sum_a \mu(a)$.

A quantum *super-operator* (or a *quantum channel* or a *quantum operation*) $\mathcal{E} : A \to B$ is a completely positive and trace preserving (CPTP) linear map (mapping states from $\mathcal{D}(A)$ to states in $\mathcal{D}(B)$). The identity operator in Hilbert space $\mathcal{H}_A$ (and associated register A) is denoted $\mathbb{I}_A$. A *unitary* operator $\mathcal{U}_A : \mathcal{H}_A \to \mathcal{H}_A$ is such that $\mathcal{U}_A^\dagger \mathcal{U}_A = \mathcal{U}_A \mathcal{U}_A^\dagger = \mathbb{I}_A$. The set of all unitary operations on register A is denoted by $\mathcal{U}(A)$.

A 2-outcome quantum measurement is defined by a collection $\{M, \mathbb{I} - M\}$, where $0 \preceq M \preceq \mathbb{I}$ is a positive semidefinite operator, where $A \preceq B$ means $B - A$ is positive semidefinite. Given a quantum state ρ, the probability of getting outcome corresponding to M is $\text{Tr}(\rho M)$ and getting outcome corresponding to $\mathbb{I} - M$ is $1 - \text{Tr}(\rho M)$.

1) Distance measures for quantum states: We now define the distance measures we use and some properties of these measures. Before defining the distance measures, we introduce the concept of *fidelity* between two states, which is not a distance measure but a similarity measure. Note that all the notions introduced below also apply to classical random variables, when viewed as diagonal quantum states in some basis.

Definition 27 (Fidelity)**.** *Let $\rho_A, \sigma_A \in \mathcal{D}(A)$ be quantum states. The fidelity between ρ and σ is defined as*

$$\text{F}(\rho_A, \sigma_A) \coloneqq \|\sqrt{\rho_A}\sqrt{\sigma_A}\|_1 .$$

For two pure states $|\psi\rangle$ and $|\phi\rangle$, we have $\text{F}(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|) = |\langle\psi|\phi\rangle|$. We now introduce the two distance measures we use.

Definition 28 (Distance measures)**.** *Let $\rho_A, \sigma_A \in \mathcal{D}(A)$ be quantum states. We define the following distance measures between these states.*

$$\textit{Trace distance:} \quad \Delta(\rho_A, \sigma_A) \coloneqq \frac{1}{2}\|\rho_A - \sigma_A\|_1.$$
$$\textit{Bures metric:} \quad \text{B}(\rho_A, \sigma_A) \coloneqq \sqrt{1 - \text{F}(\rho_A, \sigma_A)}.$$

Note that for any two quantum states ρ_A and σ_A, these distance measures lie in $[0, 1]$. The distance measures are 0 if and only if the states are equal, and the distance measures are 1 if and only if the states have orthogonal support, i.e., if $\rho_A \sigma_A = 0$.

Conveniently, these measures are closely related.

Fact 29. *For all quantum states $\rho_A, \sigma_A \in \mathcal{D}(A)$, we have*

$$\text{B}^2(\rho_A, \sigma_A) \leq \Delta(\rho_A, \sigma_A) \leq \sqrt{2} \cdot \text{B}(\rho_A, \sigma_A).$$

Proof: The Fuchs-van de Graaf inequalities [25], [24] state that

$$1 - \text{F}(\rho_A, \sigma_A) \leq \Delta(\rho_A, \sigma_A) \leq \sqrt{1 - \text{F}^2(\rho_A, \sigma_A)}.$$

Our fact follows from this and the relation

$$1 - \text{F}^2(\rho_A, \sigma_A) \leq 2 - 2\text{F}(\rho_A, \sigma_A)$$

. ■

We now review some properties of the Bures metric.

Fact 30 (Facts about Bures metric)**.**

Fact 30.A (Triangle inequality [21])**.** *The following triangle inequality and a weak triangle inequality hold for the Bures metric and the square of the Bures metric.*

1) $\text{B}(\rho_A, \sigma_A) \leq \text{B}(\rho_A, \tau_A) + \text{B}(\tau_A, \sigma_A)$.
2) $\text{B}^2(\rho_A^1, \rho_A^{t+1}) \leq t \cdot \sum_{i=1}^t \text{B}^2(\rho_A^i, \rho_A^{i+1})$.

Fact 30.B (Averaging over classical registers)**.** *For classical-quantum states $\theta_{XB}, \theta'_{XB}$ with $\theta_X = \theta'_X$, we have*

$$\text{B}^2(\theta_{XB}, \theta'_{XB}) = \mathbb{E}_{x \leftarrow X}[\text{B}^2(\theta_B^x, \theta_B'^x)].$$

Finally, an important property of both these distance measures is monotonicity under quantum operations [26], [27].

Fact 31 (Monotonicity under quantum operations)**.** *For quantum states ρ_A, $\sigma_A \in \mathcal{D}(A)$, and a quantum operation $\mathcal{E}(\cdot) : \mathcal{L}(A) \to \mathcal{L}(B)$, it holds that*

$$\Delta(\mathcal{E}(\rho_A), \mathcal{E}(\sigma_A)) \leq \Delta(\rho_A, \sigma_A) \quad \textit{and}$$
$$\mathrm{B}(\mathcal{E}(\rho_A), \mathcal{E}(\sigma_A)) \leq \mathrm{B}(\rho_A, \sigma_A),$$

with equality if $\mathcal{E}$ is unitary. In particular, for bipartite states $\rho_{AB}, \sigma_{AB} \in \mathcal{D}(AB)$, it holds that

$$\Delta(\rho_{AB}, \sigma_{AB}) \geq \Delta(\rho_A, \sigma_A) \quad \textit{and}$$
$$\mathrm{B}(\rho_{AB}, \sigma_{AB}) \geq \mathrm{B}(\rho_A, \sigma_A).$$

2) Mutual information: We start with the following fundamental information theoretic quantities. We refer the reader to the excellent sources for quantum information theory [23], [24] for further study.

Definition 32. *Let $\rho_A \in \mathcal{D}(A)$ be a quantum state. We then define the following.*

von Neumann entropy: $\mathrm{H}(\rho_A) \coloneqq -\mathrm{Tr}(\rho_A \log \rho_A)$.

We now define mutual information and conditional mutual information.

Definition 33 (Mutual information)**.** *Let $\rho_{ABC} \in \mathcal{D}(ABC)$ be a quantum state. We define the following measures.*

Mutual information:

$$\mathrm{I}(A:B)_\rho \coloneqq \mathrm{H}(\rho_A) + \mathrm{H}(\rho_B) - \mathrm{H}(\rho_{AB})$$

Conditional mutual information:

$$\mathrm{I}(A:B \mid C)_\rho \coloneqq \mathrm{I}(A:BC)_\rho - \mathrm{I}(A:C)_\rho.$$

We will need the following basic properties.

Fact 34 (Properties of I)**.** *Let $\rho_{ABC} \in \mathcal{D}(ABC)$ be a quantum state. We have the following.*

Fact 34.A (Nonnegativity)**.**

$$\mathrm{I}(A:B)_\rho \geq 0 \textit{ and } \mathrm{I}(A:B \mid C)_\rho \geq 0.$$

If $\rho_{AB} = \rho_A \otimes \rho_B$ is a product state, then

$$\mathrm{I}(A:B) = 0.$$

Fact 34.B (Chain rule)**.**

$$\mathrm{I}(A:BC)_\rho = \mathrm{I}(A:B)_\rho + \mathrm{I}(A:C \mid B)_\rho.$$

Fact 34.C (Monotonicity)**.** *For a quantum operation $\mathcal{E}(\cdot) : \mathcal{L}(A) \to \mathcal{L}(B)$, $\mathrm{I}(A:\mathcal{E}(B)) \leq \mathrm{I}(A:B)$ with equality when $\mathcal{E}$ is unitary. In particular $\mathrm{I}(A:BC)_\rho \geq \mathrm{I}(A:B)_\rho$.*

Fact 34.D (Averaging over conditioning register)**.** *For classical-quantum state (register X is classical) ρ_{XAB}:*

$$\mathrm{I}(A:B|X)_\rho = \mathbb{E}_{x \leftarrow X} \mathrm{I}(A:B)_{\rho^x}.$$

The following lemma, known as the Average Encoding Theorem [5], formalizes the intuition that if a classical and a quantum registers are weakly correlated, then they are nearly independent.

Lemma 35. *For any $\rho_{XA} = \sum_x p_X(x) \cdot |x\rangle\langle x|_X \otimes \rho_A^x$ with a classical system X and states ρ_A^x,*

$$\sum_x p_X(x) \cdot \mathrm{B}^2(\rho_A^x, \rho_A) \quad \leq \quad \mathrm{I}(X:A)_\rho \ . \tag{1}$$

The following Shearer-type inequality for quantum information was shown in Ref. [17]. Classical variants appeared in [16], [28].

Lemma 36. *Consider registers $U_1, U_2, \ldots U_m, V$ and define $U \coloneqq U_1 U_2 \ldots U_m$. Consider a quantum state Ψ_{UV} such that $\Psi_{U_1 U_2 \ldots U_m} = \Psi_{U_1} \otimes \Psi_{U_2} \otimes \ldots \otimes \Psi_{U_m}$. Let $S = \{i_1, \ldots, i_{|S|}\} \subseteq [m]$ be a random set picked independently of Ψ_{UV} satisfying $\Pr[i \in S] \leq \frac{1}{k}$ for all i and $U_S \coloneqq U_{i_1} U_{i_2} \ldots U_{i_{|S|}}$. Then it holds that*

$$\mathrm{I}(U_S : V \mid S)_\Psi \leq \frac{\mathrm{I}(U:V)_\Psi}{k},$$

D. Quantum communication complexity

In quantum communication complexity, two players wish to compute a classical function $F \colon \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ for some finite sets $\mathcal{X}$ and $\mathcal{Y}$. The inputs $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ are given to two players Alice and Bob, and the goal is to minimize the quantum communication between them required to compute the function.

While the players have classical inputs, the players are allowed to exchange quantum messages. Depending on whether or not we allow the players arbitrary shared entanglement, we get $\mathrm{Q}(F)$, bounded-error quantum communication complexity without shared entanglement and $\mathrm{Q}^*(F)$, for the same measure with shared entanglement. Obviously $\mathrm{Q}^*(F) \leq \mathrm{Q}(F)$. In this paper we will only work with $\mathrm{Q}^*(F)$, which makes our results stronger since we prove lower bounds in this work.

An entanglement assisted quantum communication protocol Π for a function is as follows. Alice and Bob start with preshared entanglement $|\Theta_0\rangle_{A_0 B_0}$. Upon receiving inputs (x, y), where Alice gets x and Bob gets y, they exchange quantum messages. At the end of the protocol, Alice applies a two outcome measurement on her qubits and correspondingly outputs 1 or 0. Let $O(x, y)$ be the random variable corresponding to the output produced by Alice in Π, given input (x, y).

Let μ be a distribution over $\mathrm{dom}(F)$. Let inputs to Alice and Bob be given in registers X and Y in the state

$$\rho_\mu \coloneqq \sum_{x,y} \mu(x,y) \, |x\rangle\langle x|_X \otimes |y\rangle\langle y|_Y \, . \tag{2}$$

Let these registers be purified by R_X and R_Y respectively, which are not accessible to either players. Denote

$$|\mu\rangle_{XR_XYR_Y} := \sum_{x,y} \sqrt{\mu(x,y)}\,|xxyy\rangle_{XR_XYR_Y}\,. \tag{3}$$

Let Alice and Bob initially hold register A_0, B_0 with shared entanglement Θ_{0,A_0B_0}. Then the initial state is

$$|\Psi_0\rangle_{XYR_XR_YA_0B_0} := |\mu\rangle_{XYR_XR_Y}\,|\Theta_0\rangle_{A_0B_0}\,. \tag{4}$$

Alice applies a unitary $U^1 : XA_0 \to XA_1C_1$ such that the unitary acts on A_0 conditioned on X. She sends C_1 to Bob. Let $B_1 \equiv B_0$ be a relabeling of Bob's register B_0. He applies $U^2 : YC_1B_1 \to YC_2B_2$ such that the unitary acts on C_1B_0 conditioned on Y. He sends C_2 to Alice. Players proceed in this fashion for t messages, for t even, until the end of the protocol. At any round r, let the registers be $A_rC_rB_r$, where C_r is the message register, A_r is Alice's register and B_r is Bob's register. If r is odd, then $B_r \equiv B_{r-1}$ and if r is even, then $A_r \equiv A_{r-1}$. On input x, y, let the joint state in registers $A_rC_rB_r$ be $\Theta^{x,y}_{r,A_rC_rB_r}$. Then the global state at round r is

$$|\Psi_r\rangle_{XYR_XR_YA_rC_rB_r} := \sum_{x,y} \sqrt{\mu(x,y)}\,|xxyy\rangle_{XR_XYR_Y}\,|\Theta^{x,y}_r\rangle_{A_rC_rB_r}\,. \tag{5}$$

We define the following quantities.

Worst-case error: $\mathrm{err}(\Pi) := \max_{(x,y)}\{\Pr[O(x,y) \neq F(x,y)]\}$.

Quantum CC of a protocol: $\mathrm{QCC}(\Pi) := \sum_i |C_i|$.

Quantum CC of F: $\mathrm{Q}^*_\varepsilon(F) := \min_{\Pi:\mathrm{err}(\Pi)\leq\varepsilon} \mathrm{QCC}(\Pi)$.

Our first fact links $\mathrm{err}(\Pi)$ with the distance Δ between a pair of final states corresponding to inputs with different outputs.

Fact 37 (Error vs. distance)**.** *Consider a non-constant function f, and let x, y and y' be inputs such that $f(x,y) \neq f(x,y')$. For any protocol Π with t rounds, it holds that*

$$\Delta(\Theta^{x,y}_{t,A_tC_t}, \Theta^{x,y'}_{t,A_tC_t}) \geq 1 - 2\mathrm{err}(\Pi).$$

In below, let A'_r, B'_r represent Alice and Bob's registers after reception of the message C_r at round r. That is, at even round r, $A'_r = A_rC_r, B'_r = B_r$ and at odd r, $A'_r = A_r, B'_r = B_rC_r$. We will need the following version of the quantum-cut-and-paste lemma from [29] (also see [12], [13] for similar arguments). This is a special case of [29, Lemma 7] and we have rephrased it using our notation.

Lemma 38 (Quantum cut-and-paste)**.** *Let Π be a quantum protocol with classical inputs and consider distinct inputs u, u' for Alice and v, v' for Bob. Let $|\Psi_{0,A_0B_0}\rangle$ be the initial shared state between Alice and Bob. Also let $\left|\Psi^{u'',v''}_{k,A'_kB'_k}\right\rangle$ be the shared state after round k of the protocol when the inputs to Alice and Bob are (u'', v'') respectively. For k odd, let*

$$h_k = \mathrm{B}\left(\Psi^{u,v}_{k,B'_k}, \Psi^{u',v}_{k,B'_k}\right)$$

and for even k, let

$$h_k = \mathrm{B}\left(\Psi^{u,v}_{k,A'_k}, \Psi^{u,v'}_{k,A'_k}\right).$$

Then

$$\mathrm{B}\left(\Psi^{u',v}_{r,A'_r}, \Psi^{u',v'}_{r,A'_r}\right) \leq 2\sum_{k=1}^{r} h_k.$$

As discussed in the introduction, approximate rank lower bounds bounded-error quantum communication complexity with shared entanglement [30]:

Fact 39. *For any two-party function $F : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ and $\varepsilon \in [0, 1/3]$, we have $\mathrm{Q}^*_\varepsilon(F) = \Omega(\log \mathrm{rk}_\varepsilon(F)) - O(\log\log(|\mathcal{X}| \cdot |\mathcal{Y}|))$.*

E. Quantum information complexity

Definition 40. *Given a quantum protocol Π with classical inputs distributed as μ, the* quantum information cost *is defined as*

$$\begin{aligned}\mathrm{QIC}(\Pi,\mu) \quad &= \textstyle\sum_{i\ \mathrm{odd}} \mathrm{I}(R_XR_Y : C_i \,|\, YB_i) \\ &+ \quad \textstyle\sum_{i\ \mathrm{even}} \mathrm{I}(R_XR_Y : C_i \,|\, XA_i)\ .\end{aligned}$$

Definition 41. *Given a quantum protocol Π with classical inputs distributed as μ, the cumulative* Holevo information cost *is defined as*

$$\begin{aligned}\mathrm{HQIC}(\Pi,\mu) \quad &= \textstyle\sum_{i\ \mathrm{odd}} \mathrm{I}(X : B_iC_i \,|\, Y) \\ &+ \quad \textstyle\sum_{i\ \mathrm{even}} \mathrm{I}(Y : A_iC_i \,|\, X)\ .\end{aligned}$$

Definition 42. *Given a quantum protocol Π and a product distribution μ over the classical inputs, the cumulative* superposed-Holevo information cost *is defined as*

$$\begin{aligned}\mathrm{SQIC}(\Pi,\mu) \quad &= \textstyle\sum_{i\ \mathrm{odd}} \mathrm{I}(X : YR_YB_iC_i)_{\rho_i} \\ &+ \quad \textstyle\sum_{i\ \mathrm{even}} \mathrm{I}(Y : XR_XA_iC_i)_{\rho_i}\ .\end{aligned}$$

Note that for product input distributions on XY and for each i,

$$\mathrm{I}(X : B_iC_i|Y)_{\rho_i} = \mathrm{I}(X : YB_iC_i)_{\rho_i} \leq \mathrm{I}(X : YR_YB_iC_i)_{\rho_i}, \tag{6}$$

$$\mathrm{I}(Y : A_iC_i|X)_{\rho_i} = \mathrm{I}(Y : XA_iC_i)_{\rho_i} \leq \mathrm{I}(Y : XR_XA_iC_i)_{\rho_i}. \tag{7}$$

Combining with other results in Ref. [19], we get the following for any t round protocol Π and any product

distribution μ:

$$2\mathrm{QCC}(\Pi) \geq \mathrm{QIC}(\Pi,\mu) \tag{8}$$
$$\geq \frac{1}{t}\,\mathrm{SQIC}(\Pi,\mu) \tag{9}$$
$$\geq \frac{1}{t}\,\mathrm{HQIC}(\Pi,\mu) \tag{10}$$
$$\geq \frac{1}{2t}\,\mathrm{QIC}(\Pi,\mu). \tag{11}$$

.

III. Lower bound on the information complexity of Sink ∘ Xor

A. Reducing Equality to *Sink* ∘ *Xor*

We define the Equality function as

$$\mathrm{EQ}(x,y) = \begin{cases} 1 & \textit{if } x = y, \\ 0 & \textit{otherwise.} \end{cases}$$

Recall the Sink function from Definition 4. Following [2] we use projections of the inputs in our proof to analyze the input of the Sink function. Let $w \in \{0,1\}^{\binom{m}{2}}$. Let E_{v_i} be the set of $m-1$ input coordinates that correspond to the edges incident to v_i. We use the notation w_{v_i} to denote the input projected to the coordinates in E_{v_i}. Note that w_{v_i} decides whether or not v_i is a sink. By z_{v_i}, we refer to the $m-1$ bit string such that v_i is a sink if and only if $w_{v_i} = z_{v_i}$. Sink can be written as

$$\mathsf{Sink}(w) = \vee_{i=1}^{m} \mathrm{EQ}(w_{v_i}, z_{v_i})$$

since only one of the vertex can be a sink in the complete directed graph. Our communication function is Sink ∘ Xor : $\{0,1\}^{\binom{m}{2}} \times \{0,1\}^{\binom{m}{2}} \to \{0,1\}$. Similar to Sink, Sink ∘ Xor can be represented as

$$\mathsf{Sink} \circ \mathsf{Xor}(x,y) = \vee_{i=1}^{m} \mathrm{EQ}(x_{v_i}, y_{v_i} \oplus z_{v_i}).$$

Our first result is as follows.

Theorem 43. *Suppose $m \geq 10$. Let Π be a protocol for Sink ∘ Xor which makes a worst case error of at most $\frac{1}{4}$. There exists a protocol Π' for* EQ *that makes a worst case error of at most $\frac{1}{4} + \frac{m-1}{2^{m-2}} \leq \frac{1}{3}$. Furthermore, it holds that*

$$\mathrm{IC}(\Pi',\nu) \leq \frac{2}{m}\,\mathrm{IC}(\Pi,\mu),$$

where ν is the uniform distribution over inputs to EQ *and μ is uniform over the inputs to Sink ∘ Xor.*

Proof: We have

$$\mathrm{IC}(\Pi,\mu) = \mathrm{I}(X : \Pi | YRR_B) + \mathrm{I}(Y : \Pi | XRR_A) \tag{12}$$
$$= \mathrm{I}(X : \Pi Y RR_B) + \mathrm{I}(Y : \Pi X RR_A), \tag{13}$$

where the information quantities are evaluated on μ and the associated Π. Let S be a random variable which takes values in $\{E_{v_1}, E_{v_2}, \ldots, E_{v_m}\}$ with uniform probability. Let $X_{E_{v_i}}$ (similarly $Y_{E_{v_i}}$) be the restriction of X (similarly Y) to coordinates in E_{v_i}. Since each coordinate j appears in exactly two sets in $\{E_{v_1}, E_{v_2}, \ldots, E_{v_m}\}$, we have $\Pr[j \in S] = \frac{2}{m}$. Thus, from Lemma 36, we have

$$\frac{2}{m}\,\mathrm{IC}(\Pi,\mu) \geq \mathbb{E}_s[\mathrm{I}(X_S : Y\Pi RR_B | S = s) + \mathrm{I}(Y_S : X\Pi RR_A | S = s)] \tag{14}$$
$$= \mathbb{E}_s[\mathrm{I}(X_S : \Pi | YRR_B, S = s) + I(Y_S : \Pi | XRR_A, S = s)]. \tag{15}$$

The protocol Π' for EQ is now as follows, for inputs $c, d \in \{0,1\}^{m-1}$ (we use c, d as inputs here to avoid confusion with x, y for Sink ∘ Xor).

- Alice and Bob take a sample s from S using shared randomness. Let i be such that $E_{v_i} = s$.
- They set $x_s = c$ and $y_s = d \oplus z_{v_i}$. Alice samples $x_{\bar{s}}$ uniformly at random from private randomness and Bob samples $y_{\bar{s}}$ uniformly at random from private randomness. Here $\bar{s}$ is the complement of s. This specifies the input x, y for Sink ∘ Xor.
- They run the protocol Π and output accordingly.

Observe that x_s and y_s are distributed uniformly if c and d are. Thus,

$$\mathrm{IC}(\Pi',\nu) = \mathbb{E}_s[\mathrm{I}(X_S : \Pi | Y_S R Y_{\bar{S}} R_B, S = s) + \mathrm{I}(Y_S : \Pi | X_S R X_{\bar{S}} R_A, S = s)]$$
$$= \mathbb{E}_s[\mathrm{I}(X_S : \Pi | YRR_B, S = s) + \mathrm{I}(Y_S : \Pi | XRR_A, S = s)],$$

where the information quantities are evaluated on μ and the associated Π, and the desired information bound follows by (14).

To bound the worst case error of Π', we argue as follows. Fix some input c, d to Π'. If $c = d$, then $x_s = y_s \oplus z_{v_i}$ which implies that error of Π' on this input is same as the error of Π on the corresponding x, y, hence at most $\mathrm{err}(\Pi)$. Now consider the case where $c \neq d$. The function Sink ∘ Xor evaluates to 1 only if $x_{E_{v_j}} = y_{E_{v_j}} \oplus z_{v_j}$ for some $j \in [m]$. Since, $c \neq d$, we conclude that j (if it exists) cannot be equal to i. Moreover, the edge adjacent to i is already fixed by c, d, and if it is not consistent with the corresponding value in z_{v_i}, then j is not a sink. Hence, similar to the argument in [2, Claim 5.6], the probability that j is a sink is at most $\frac{1}{2^{m-2}}$, as all $m-1$ edges must be incoming and the edge adjacent to i is already fixed. Hence by a union bound, the probability for an x, y (that satisfy $x_{v_i} = c, y_{v_i} = d \oplus z_{v_i}, c \neq d$) to form a 1 input at some other coordinate j is at most $\frac{m-1}{2^{m-2}}$. This implies that $\mathrm{err}(\Pi') \leq \mathrm{err}(\Pi) + \frac{m-1}{2^{m-2}}$. This completes the proof. ∎

B. Lower bound on information complexity of Equality

To complete the argument, we use the following lemma (that uses a cut and paste argument) implicit in [10] and

obtain a lower bound on the information complexity of EQ. We repeat its proof for completeness (and consistency with our notation).

Lemma 44. *Let Π be a protocol for* EQ *that makes a worst case error of at most $\frac{1}{3}$. Then it holds that $\mathrm{IC}(\Pi,\nu) \geq \frac{1}{432}$, where ν is uniform over inputs to* EQ.

Proof: Let R_A and R_B be private randomness of Alice and Bob (respectively) in the protocol and R be the public randomness. We have

$$\mathrm{IC}(\Pi,\nu) = \mathrm{I}(Y:\Pi|XR_AR) + \mathrm{I}(X:\Pi|YR_BR).$$

By the average-encoding theorem (Fact 35), it holds that added R below

$$\begin{aligned}\mathrm{I}(X:\Pi \mid YR_BR) &= \mathrm{I}(X:RR_B\Pi \mid Y)\\ &\geq \mathrm{I}(X:R\Pi \mid Y)\\ &\geq \mathbb{E}_{x,y\leftarrow XY}\mathrm{B}^2((R\Pi)^{x,y},(R\Pi)^y).\end{aligned}$$

Similarly,

$$\begin{aligned}\mathrm{I}(Y:\Pi \mid XR_AR) &= \mathrm{I}(Y:XR_AR\Pi)\\ &\geq \mathrm{I}(Y:R\Pi)\\ &\geq \mathbb{E}_{y\leftarrow Y}\mathrm{B}^2((R\Pi)^y,R\Pi).\end{aligned}$$

Using the weak triangle inequality (Fact 30.A), the above two inequalities imply

$$\begin{aligned}&\mathbb{E}_{x,y\leftarrow XY}\mathrm{B}^2((R\Pi)^{x,y},R\Pi) \leq\\ &\qquad 2\mathbb{E}_{x,y\leftarrow XY}(\mathrm{B}^2((R\Pi)^{x,y},(R\Pi)^y)\\ &\qquad\qquad + \mathrm{B}^2((R\Pi)^y,R\Pi))\\ &\leq 2(\mathrm{I}(X:\Pi \mid YR_BR)\\ &\qquad\qquad + \mathrm{I}(Y:\Pi \mid XR_AR))\\ &= 2\,\mathrm{IC}(\Pi,\nu).\end{aligned}$$

Since x,y are uniform, we can write the above relation as

$$\mathbb{E}_{t\leftarrow Y}\mathbb{E}_{x\leftarrow X}\mathrm{B}^2((R\Pi)^{x,x\oplus t},R\Pi) \leq 2\,\mathrm{IC}(\Pi,\nu).$$

Since $\Pr[t=0] = \frac{1}{2^{m-1}}$, this implies that there exists an $t\neq 0$ such that

$$\mathbb{E}_{x\leftarrow X}\mathrm{B}^2((R\Pi)^{x,x\oplus t},R\Pi) \leq 3\,\mathrm{IC}(\Pi,\nu).$$

An equivalent way to write the above inequality, by relabeling $x\to x\oplus t$, is

$$\mathbb{E}_{x\leftarrow X}\mathrm{B}^2((R\Pi)^{x\oplus t,x},R\Pi) \leq 3\,\mathrm{IC}(\Pi,\nu).$$

By the weak triangle inequality (Fact 30.A), we conclude

$$\mathbb{E}_{x\leftarrow X}\mathrm{B}^2((R\Pi)^{x\oplus t,x},(R\Pi)^{x,x\oplus t}) \leq 12\,\mathrm{IC}(\Pi,\nu).$$

The pythagorean property (Fact 26) now implies that

$$\mathbb{E}_{x\leftarrow X}\mathrm{B}^2((R\Pi)^{x,x},(R\Pi)^{x,x\oplus t}) \leq 24\,\mathrm{IC}(\Pi,\nu).$$

Thus, there exists some x for which $\mathrm{B}^2((R\Pi)^{x,x},(R\Pi)^{x,x\oplus t}) \leq 24\,\mathrm{IC}(\Pi,\nu)$. Since Π makes an error of at most $\frac{1}{3}$, we require (using relation between Bures metric and triangle inequality, Fact 29)

$$\begin{aligned}\mathrm{B}^2((R\Pi)^{x,x},(R\Pi)^{x,x\oplus t}) &\geq \frac{1}{2}\Delta^2((R\Pi)^{x,x},(R\Pi)^{x,x\oplus t})\\ &\geq \frac{1}{2}\Delta^2((\Pi)^{x,x},(\Pi)^{x,x\oplus t})\\ &\geq \frac{1}{18}.\end{aligned}$$

Thus, $\mathrm{IC}(\Pi,\nu)\geq\frac{1}{432}$, which completes the proof. ∎

Theorem 43 and Lemma 44 jointly imply that $\mathrm{IC}(\Pi,\mu)\geq\frac{m}{864}$, for any protocol Π that makes an error of at most $\frac{1}{4}$ on $\mathsf{Sink}\circ\mathsf{Xor}$. This establishes the desired lower bound.

IV. Reducing Equality to Sink for quantum information

A. Shearer-type embedding

We begin by showing a general embedding result based on the Shearer-type lemma for quantum information (Lemma 36). Consider a protocol Π acting on input registers $X_1,X_2,\ldots,X_m$ and $Y_1,Y_2,\ldots Y_m$, with $X_1\equiv X_2\equiv\ldots\equiv X_m$ and $Y_1\equiv Y_2\equiv\ldots\equiv Y_m$. Define $X = X_1X_2\ldots X_m$, $Y=Y_1Y_2\ldots Y_m$. Consider a product input distribution $\mu=\mu_1\otimes\mu_2$ on X_iY_i. Consider $t\in[m]$ and let $S=\{i_1,i_2,\ldots,i_t\}\subseteq[m]$ be a random set of size t picked independently of the input on XY and satisfying $\Pr[i\in S]\leq\frac{1}{k}$ for all i. Let $X_S = X_{i_1}X_{i_2}\ldots X_{i_t}$, $Y_S = Y_{i_1}Y_{i_2}\ldots Y_{i_t}$. We define the following protocol Π_S acting on input $A_{in}B_{in}$, with $A_{in}\equiv X_S$, $B_{in}\equiv Y_S$.

Protocol Π_S on input $\sigma_{A_{in}B_{in}}$

1) Alice privately sample X_i for each $i\notin S$ as $|\mu_1\rangle_{X_iR_{X_i}}$.
2) Bob privately sample Y_i for each $i\notin S$ as $|\mu_2\rangle_{Y_iR_{Y_i}}$.
3) Alice embeds A_{in} into X_S.
4) Bob embeds B_{in} into Y_S.
5) They run Π, and output Π's output.

Lemma 45.

$$\begin{aligned}\Pi_S(\sigma_{A_{in}B_{in}}) &= \Pi(\sigma_{X_SY_S}\otimes(\rho_\mu^{\otimes m-t})_{X_{\bar S}Y_{\bar S}}),\\ \mathrm{SQIC}(\Pi_S,\mu^{\otimes t}) &= \sum_{i\ odd}\mathrm{I}(X_S{:}YR_YB_iC_i)_{\rho_i} \quad +\\ &\qquad\qquad \sum_{i\ even}\mathrm{I}(Y_S{:}XR_XA_iC_i)_{\rho_i},\end{aligned}$$

with ρ_i the state in round i when Π is run on input distribution $\mu^{\otimes m}$.

Proof: By the definition of protocol Π_S, the channel it implements is $\Pi(\sigma_{X_SY_S}\otimes(\rho_\mu^{\otimes m-t})_{X_{\bar S}Y_{\bar S}})$ (see (2) in Section II-D for definition of ρ_μ) on input $\sigma_{A_{in}B_{in}}$.

For the information cost when Π_S is run on input distribution $\mu^{\otimes t}$, first notice that for a given S, we can rewrite $YR_Y = Y_S R_{Y_S} Y_{\bar{S}} R_{Y_{\bar{S}}}$. After embedding $A_{in}B_{in}$ into $X_S Y_S$, the $X_S Y_S$ registers correspond to the input of Π_S while $R_{X_S} R_{Y_S}$ correspond to the purification of the input registers. The $X_{\bar{S}} R_{X_{\bar{S}}}$ and $Y_{\bar{S}} R_{Y_{\bar{S}}}$ registers correspond to the part privately sampled according to $\mu = \mu_1 \otimes \mu_2$ by Alice and Bob, respectively, in order to run Π. Hence, for a given S, the terms in SQIC look like

$$\mathrm{I}(X_S : Y_S R_{Y_S} Y_{\bar{S}} R_{Y_{\bar{S}}} B_i C_i) = \mathrm{I}(X_S : Y R_Y B_i C_i),$$
$$\mathrm{I}(Y_S : X_S R_{X_S} X_{\bar{S}} R_{X_{\bar{S}}} A_i C_i) = \mathrm{I}(Y_S : X R_X A_i C_i).$$

The result follows. ∎

Let $|\phi_S\rangle_{S_A S_B}$ be a quantum state shared between Alice and Bob and encoding the distribution on S. Given S, let P_A^S and P_B^S be permutations (over the computational basis) acting on A_{in} and B_{in}, respectively, and such that μ is invariant under their action, i.e.

$$(P_A^S \otimes P_B^S)(\rho_\mu^{\otimes t}) = \rho_\mu^{\otimes t}. \tag{16}$$

We define the following protocol $\hat{\Pi}$ also acting on $A_{in}B_{in}$.

Protocol $\hat{\Pi}$ on input $\sigma_{A_{in}B_{in}}$

1) Alice and Bob share $|\phi_S\rangle_{S_A S_B}$.
2) Conditioned on the value of S shared in $|\phi_S\rangle$, Alice and Bob apply P_A^S and P_B^S to their inputs, respectively.
3) Conditioned on value of S shared in $|\phi_S\rangle$, Alice and Bob run Π_S, and output Π_S's output.

Lemma 46.

$$\hat{\Pi}(\sigma_{A_{in}B_{in}}) = \mathbb{E}_S[\Pi_S \circ (P_A^S \otimes P_B^S)(\sigma_{A_{in}B_{in}})],$$
$$\mathrm{SQIC}(\hat{\Pi}, \mu^{\otimes t}) = \mathbb{E}_S \mathrm{SQIC}(\Pi_S, \mu^{\otimes t}) \leq \mathrm{SQIC}(\Pi, \mu^{\otimes m})/k.$$

Proof:

By the definition of protocol $\hat{\Pi}$, the channel it implements is $\mathbb{E}_S[\Pi_S \circ (P_A^S \otimes P_B^S)]$.

For the information cost, let $\hat{\rho}_i$ be the state in round i when $\hat{\Pi}$ is run on input distribution $\mu^{\otimes t}$. Similar comments in the proof of Lemma 45 hold regarding XY vs. $X_S Y_S X_{\bar{S}} Y_{\bar{S}}$ and the corresponding R purification registers. Hence the terms for SQIC look like

$$\mathrm{I}(X_S : S_B Y R_Y B_i C_i)_{\hat{\rho}_i} = \mathrm{I}(X_S : Y R_Y B_i C_i | S)_{\hat{\rho}_i} \tag{17}$$
$$= \mathbb{E}_S \mathrm{I}(X_S : Y R_Y B_i C_i)_{\hat{\rho}_i^S}, \tag{18}$$

where $\hat{\rho}_i^S$ is the state on registers other than $S_A S_B$, conditioned on S. Let $P_{A,X_S}^S, P_{A,R_{X_S}}^S$ (similarly $P_{B,Y_S}^S, P_{B,R_{Y_S}}^S$) be the operator P_A^S (similarly P_B^S) acting on the registers X_S, R_{X_S} (similarly Y_S, R_{Y_S}) respectively. Then, for any S, Equation 16 implies that

$$(P_{A,X_S}^S \otimes P_{B,Y_S}^S)(P_{A,R_{X_S}}^S \otimes P_{B,R_{Y_S}}^S) \left|\mu^{\otimes t}\right\rangle_{X_S R_{X_S} Y_S R_{Y_S}} = \left|\mu^{\otimes t}\right\rangle_{X_S R_{X_S} Y_S R_{Y_S}}.$$

Recall that ρ_i is the state in round i when Π is run on input distribution $\mu^{\otimes m}$. Thus

$$(P_{A,R_{X_S}}^S \otimes P_{B,R_{Y_S}}^S)(\hat{\rho}_i^S) = \rho_i \tag{19}$$

is independent of S, since the operations on the R registers commute with the operations in protocol Π. By invariance of mutual information under local unitaries, we get

$$\mathbb{E}_S \mathrm{I}(X_S : Y R_Y B_i C_i)_{\hat{\rho}_i^S} = \mathbb{E}_S \mathrm{I}(X_S : Y R_Y B_i C_i)_{\rho_i} \tag{20}$$
$$= \mathrm{I}(X_S : Y R_Y B_i C_i | S)_{\rho_i}, \tag{21}$$

in which we also used that S is picked independently of the input and thus stays independent of ρ_i throughout. Similar results hold for the terms accounting for Alice's information about Bob's input in SQIC. It follows that

$$\mathrm{SQIC}(\hat{\Pi}, \mu^{\otimes t}) = \mathbb{E}_S \mathrm{SQIC}(\Pi_S, \mu^{\otimes t}).$$

To relate this to $\mathrm{SQIC}(\Pi, \mu^{\otimes m})$, we apply the Shearer type lemma for quantum information (Lemma 36) to get

$$\mathrm{I}(X_S : Y R_Y B_i C_i | S)_{\rho_i} \leq \frac{1}{k} \mathrm{I}(X : Y R_Y B_i C_i)_{\rho_i},$$
$$\mathrm{I}(Y_S : X R_X A_i C_i | S)_{\rho_i} \leq \frac{1}{k} \mathrm{I}(Y : X R_X A_i C_i)_{\rho_i},$$

and the result follows. ∎

B. From *Sink ∘ Xor* to EQ

We get the following theorem relating SQIC for Sink∘Xor and EQ.

Theorem 47. *Fix a t round quantum communication protocol Π making worst-case error ε on function Sink ∘ Xor for inputs of size $\binom{m}{2}$ bits. Then there exists a t round quantum communication protocol Π_E making worst case error $\varepsilon + o(1)$ on* EQ *with inputs of size $m-1$ bits and satisfying the following for ν the uniform distribution on $1+1$ bits :*

$$\mathrm{SQIC}(\Pi_E, \nu^{\otimes m-1}) \leq \frac{2}{m} \mathrm{SQIC}(\Pi, \nu^{\otimes \binom{m}{2}}).$$

Proof: Recall the sets E_{v_i}, for $i \in [m]$, as defined in Subsection III-A. In the setting of the Shearer-type embedding above (Lemma 46), pick $S = E_{v_i}$ with probability $1/m$ for each $i \in [m]$. Let $P_A^{S_i}$ be the map that performs bit-wise addition $\oplus z_{v_i}$, and $P_B^{S_i}$ is the identity. Notice that each pair (k, l), for $k < l$, appears for exactly two choices of i: once for $i = k$, and once for $i = l$. Hence, $\Pr[l \in S] \leq 2/m$ for all $l \in [m]$, and $2/m$ is the probability we use in the Shearer-type embedding. By using ν the uniform distribution on $1+1$ bits as the product distribution μ in the Shearer-type embedding, the SQIC bound follows.

It is left to argue that the resulting protocol Π_E taken to be $\hat{\Pi}$ of the embedding is good at solving EQ. But this follows as in the classical embedding argument (see the proof of Theorem 43) since the probability that Alice and Bob privately sampled inputs to Π on $\bar{S}$ that already make $\mathsf{Sink} \circ \mathsf{Xor}$ evaluate to 1 on $\bar{S}$ is exponentially small in m, hence the additional error is $o(1)$. ∎

C. Quantum information cost of Equality function

We use the following lemma about the quantum information cost of the equality function EQ on the uniform distribution, which was implicitly shown via a quantum cut and paste argument in Ref. [11].

Lemma 48. *Fix a t round quantum communication protocol Π making worst-case error at most $\frac{1}{3}$ on* EQ. *Let $|\Psi_r\rangle_{XYR_XR_YA_rC_rB_r}$ be the quantum state in r-th round, as defined in (5) in Section II-D, when Π is run on the uniform distribution $\mu^{\otimes k}$ on $k+k$ bits. It holds that*

$$\mathrm{HQIC}(\Pi, \mu^{\otimes k}) \geq \frac{1}{40000t}.$$

The proof of our main result, Theorem 5, follows.

Proof of Theorem 5: Let Π be a t-round protocol for $\mathsf{Sink} \circ \mathsf{Xor}$ making worst-case error at most $1/5$ on input graphs of size m, for m large enough. Then by Theorem 47 there exists a t-round protocol Π_E for EQ making error at most $1/3$ and with information cost satisfying

$$\mathrm{SQIC}(\Pi, \mu^{\otimes\binom{m}{2}}) \geq \frac{m}{2}\,\mathrm{SQIC}(\Pi_E, \mu^{\otimes m-1}),$$

with μ the uniform distribution on $1+1$ bits. Combining with Lemma 48 and (9), the following chain of inequality gives the result:

$$\begin{aligned}\frac{2t}{m}\,\mathrm{QIC}(\Pi, \mu^{\otimes\binom{m}{2}}) &\geq \frac{2}{m}\,\mathrm{SQIC}(\Pi, \mu^{\otimes\binom{m}{2}}) \\ &\geq \mathrm{SQIC}(\Pi_E, \mu^{\otimes m-1}) \\ &\geq \mathrm{HQIC}(\Pi_E, \mu^{\otimes m-1}) \\ &\geq \frac{1}{40000t}.\end{aligned}$$

∎

We add the proof of Lemma 48 for completeness.

Proof of Lemma 48: By averaging over the conditioning register and then applying the average encoding theorem (Fact 34.D and Lemma 35), we conclude that

$$\begin{aligned}\mathrm{HQIC}(\Pi, \mu^{\otimes k}) &\coloneqq \sum_{r=odd} \mathrm{I}(X : B_rC_r|Y)_{\Psi_r} \\ &\quad + \sum_{r=even} \mathrm{I}(Y : A_rC_r|X)_{\Psi_r} \\ &\geq \mathbb{E}_{x,y\leftarrow\mu}\{ \sum_{r=odd} \mathrm{B}\left(\Psi^{x,y}_{r,B_rC_r}, \Psi^{y}_{r,B_rC_r}\right)^2 \\ &\quad + \sum_{r=even} \mathrm{B}\left(\Psi^{x,y}_{r,A_rC_r}, \Psi^{x}_{r,A_rC_r}\right)^2\} \\ &\geq \frac{1}{t}\{\mathbb{E}_{x,y\leftarrow\mu}\{ \sum_{r=odd} \mathrm{B}\left(\Psi^{x,y}_{r,B_rC_r}, \Psi^{y}_{r,B_rC_r}\right) \\ &\quad + \sum_{r=even} \mathrm{B}\left(\Psi^{x,y}_{r,A_rC_r}, \Psi^{x}_{r,A_rC_r}\right)^2\}\}. \end{aligned} \tag{22}$$

Let x_1, x_2, y_2 be drawn uniformly from $\{0,1\}^k$ and let $y_1 \coloneqq x_1$. Observe that, taken separately, (x_1, y_2), (x_2, y_1) and (x_2, y_2) are distributed uniformly. Thus, (22) ensures that

$$\begin{aligned}&\sqrt{t\,\mathrm{HQIC}(\Pi, \mu^{\otimes k})} \geq \\ &\quad \mathbb{E}_{x_1,y_2\leftarrow\mu}\{ \sum_{r=odd} \mathrm{B}\left(\Psi^{x_1,y_2}_{r,B_rC_r}, \Psi^{y_2}_{r,B_rC_r}\right) \\ &\quad + \sum_{r=even} \mathrm{B}\left(\Psi^{x_1,y_2}_{r,A_rC_r}, \Psi^{x_1}_{r,A_rC_r}\right)\} \\ &\sqrt{t\,\mathrm{HQIC}(\Pi, \mu^{\otimes k})} \geq \\ &\quad \mathbb{E}_{x_2,y_1\leftarrow\mu}\{ \sum_{r=odd} \mathrm{B}\left(\Psi^{x_2,y_1}_{r,B_rC_r}, \Psi^{y_1}_{r,B_rC_r}\right) \\ &\quad + \sum_{r=even} \mathrm{B}\left(\Psi^{x_2,y_1}_{r,A_rC_r}, \Psi^{x_2}_{r,A_rC_r}\right)\} \\ &\sqrt{t\,\mathrm{HQIC}(\Pi, \mu^{\otimes k})} \geq \\ &\quad \mathbb{E}_{x_2,y_2\leftarrow\mu}\{ \sum_{r=odd} \mathrm{B}\left(\Psi^{x_2,y_2}_{r,B_rC_r}, \Psi^{y_2}_{r,B_rC_r}\right) \\ &\quad + \sum_{r=even} \mathrm{B}\left(\Psi^{x_2,y_2}_{r,A_rC_r}, \Psi^{x_2}_{r,A_rC_r}\right)\}. \end{aligned}$$

Moreover, it holds that $\Pr\left(\mathrm{EQ}(x_1, y_2) = 1\right) = \Pr\left(\mathrm{EQ}(x_2, y_1) = 1\right) = \Pr\left(\mathrm{EQ}(x_2, y_2) = 1\right) = \frac{1}{2^k}$. Thus, by first conditioning (separately) on $\mathrm{EQ}(x_1, y_2) = \mathrm{EQ}(x_2, y_1) = \mathrm{EQ}(x_2, y_2) = 0$ and then applying Markov's inequality, we find that there exists a choice of x_1, x_2, y_2 satisfying the non-equality conditions and such that

$$\begin{aligned}&5\sqrt{t\,\mathrm{HQIC}(\Pi, \mu^{\otimes k})} \geq \\ &\quad \sum_{r=odd} \mathrm{B}\left(\Psi^{x_1,y_2}_{r,B_rC_r}, \Psi^{y_2}_{r,B_rC_r}\right) \\ &\quad + \sum_{r=even} \mathrm{B}\left(\Psi^{x_1,y_2}_{r,A_rC_r}, \Psi^{x_1}_{r,A_rC_r}\right), \\ &5\sqrt{t\,\mathrm{HQIC}(\Pi, \mu^{\otimes k})} \geq \\ &\quad \sum_{r=odd} \mathrm{B}\left(\Psi^{x_2,y_1}_{r,B_rC_r}, \Psi^{y_1}_{r,B_rC_r}\right) \\ &\quad + \sum_{r=even} \mathrm{B}\left(\Psi^{x_2,y_1}_{r,A_rC_r}, \Psi^{x_2}_{r,A_rC_r}\right), \\ &5\sqrt{t\,\mathrm{HQIC}(\Pi, \mu^{\otimes k})} \geq \end{aligned}$$

$$\sum_{r=odd} \mathrm{B}\left(\Psi^{x_2,y_2}_{r,B_rC_r}, \Psi^{y_2}_{r,B_rC_r}\right) + \sum_{r=even} \mathrm{B}\left(\Psi^{x_2,y_2}_{r,A_rC_r}, \Psi^{x_2}_{r,A_rC_r}\right). \tag{23}$$

Applying the triangle inequality (Fact 30.A) to (23), we conclude that

$$\begin{aligned} 10\sqrt{t\,\mathrm{HQIC}(\Pi,\mu^{\otimes k})} &\geq \sum_{r=odd} \mathrm{B}\left(\Psi^{x_1,y_2}_{r,B_rC_r}, \Psi^{x_2,y_2}_{r,B_rC_r}\right) \\ 10\sqrt{t\,\mathrm{HQIC}(\Pi,\mu^{\otimes k})} &\geq \sum_{r=even} \mathrm{B}\left(\Psi^{x_2,y_1}_{r,A_rC_r}, \Psi^{x_2,y_2}_{r,A_rC_r}\right). \end{aligned}$$

Assume that t is even and Alice produces the output, we use the quantum cut-and-paste Lemma (Lemma 38) to conclude that

$$\begin{aligned} \mathrm{B}\left(\Psi^{x_1,y_2}_{t,A_tC_t}, \Psi^{x_1,y_1}_{t,A_tC_t}\right) &\leq 2(10\sqrt{t\,\mathrm{HQIC}(\Pi,\mu^{\otimes k})} \\ &\quad + 10\sqrt{t\,\mathrm{HQIC}(\Pi,\mu^{\otimes k})}) \\ &= 40\sqrt{t\,\mathrm{HQIC}(\Pi,\mu^{\otimes k})}. \end{aligned}$$

If $\mathrm{HQIC}(\Pi,\mu^{\otimes k}) \leq \frac{1}{40000t}$, we conclude that $40\sqrt{t\,\mathrm{HQIC}(\Pi,\mu^{\otimes k})} \leq \frac{1}{5}$, and then

$$\begin{aligned} 1 - 2\mathrm{err}(\Pi) &\leq \Delta(\Psi^{x_1,y_2}_{t,A_tC_t}, \Psi^{x_1,y_1}_{t,A_tC_t}) \\ &\leq \sqrt{2}B(\Psi^{x_1,y_2}_{t,A_tC_t}, \Psi^{x_1,y_1}_{t,A_tC_t}) \\ &\leq \sqrt{2}/5 \\ &< 1/3, \end{aligned}$$

which leads to contradiction with the fact that protocol Π makes an error of at most $\frac{1}{3}$. This completes the proof. ■

V. Conclusion and open problems

Our main result exhibits that the function introduced in [2] witnesses an exponential separation between quantum communication complexity and log-approximate rank. A consequence of our lower bound is that the randomized and quantum communication complexities of this function are polynomially related. Thus, the long-standing problem of finding a total function, that provides an exponential separation between randomized communication complexity and quantum communication complexity, remains open.

An interesting question that our techniques do not resolve is if we can show a round independent exponential separation between log-approximate rank and QIC. We believe that it would be surprising if the log-approximate rank and QIC were polynomially related. Known functions witnessing exponential separation between QIC and QCC have a completely different structure [16], [28], [17].

Further, we would like to understand if the Shearer-type embedding can go beyond product input distributions, and if it can be improved for QIC. Finally, it would be interesting if the lower bound in Corollary 6 could be improved to $\Omega(m^{1/2})$, matching the achievable protocol using distributed Grover search (up to logarithmic terms; see [2, Conclusion]).

Acknowledgements

We thank Ashwin Nayak for detailed discussions about the proof. N.G.B thanks Shalev Ben-David for pointing out [2], A.A. thanks N.G.B. for pointing out [2] and finally, D.T. thanks A.A. for pointing out [2]. We also thank Ronald de Wolf and Makrand Sinha for helpful correspondence.

This work was done when N.G.B was visiting the Institute for Quantum Computing, University of Waterloo and supported by a Queen Elizabeth Scholarship. N.G.B was also supported by the National Research Foundation, Prime Ministers Office, Singapore and the Ministry of Education, Singapore under the Research Centres of Excellence programme. A.A. and D.T. were supported in part by NSERC, CIFAR, Industry Canada. D.T. was also supported by an NSERC PDF. IQC and PI are supported in part by the Government of Canada and the Province of Ontario.

References

[1] L. Lovasz and M. Saks, "Lattices, mobius functions and communications complexity," in *Proceedings of the 29th Annual Symposium on Foundations of Computer Science (FOCS 1988)*, 1988, pp. 81–90.

[2] A. Chattopadhyay, N. Mande, and S. Sherif, "The log-approximate-rank conjecture is false," in *Proceedings of the 51st Annual ACM on Symposium on Theory of Computing (STOC 2019)*, 2019.

[3] T. Lee and A. Shraibman, "Lower bounds in communication complexity," *Foundations and Trends in Theoretical Computer Science*, vol. 3, no. 4, pp. 263–399, 2009.

[4] M. Sinha and R. de Wolf, "Exponential separation between quantum communication and logarithm of approximate rank," in *Proceedings of the 60th IEEE Symposium on Foundations of Computer Science (FOCS 2019)*, 2019.

[5] H. Klauck, A. Nayak, A. Ta-Shma, and D. Zuckerman, "Interaction in quantum communication," *IEEE Transactions on Information Theory*, vol. 53, no. 6, pp. 1970–1982, 2007.

[6] A. Chakrabarti, Y. Shi, A. Wirth, and A. Yao, "Informational complexity and the direct sum problem for simultaneous message complexity," in *Proceedings of the 42nd IEEE Symposium onFoundations of Computer Science (FOCS 2001)*. IEEE, 2001, pp. 270–278.

[7] Z. Bar-Yossef, T. Jayram, R. Kumar, and D. Sivakumar, "An information statistics approach to data stream and communication complexity," *Journal of Computer and System Sciences*, vol. 68, no. 4, pp. 702–732, 2004.

[8] R. Jain, J. Radhakrishnan, and P. Sen, "A direct sum theorem in communication complexity via message compression," in *Proceedings of the International Colloquium on Automata, Languages, and Programming (ICALP 2003)*. Springer, 2003, pp. 300–315.

[9] B. Barak, M. Braverman, X. Chen, and A. Rao, "How to compress interactive communication," in *Proceedings of the 42nd ACM Symposium on Theory of Computing (STOC 2010)*, 2010, pp. 67–76.

[10] A. Anshu, A. Belovs, S. Ben-David, M. Göös, R. Jain, R. Kothari, T. Lee, and M. Santha, "Separations in communication complexity using cheat sheets and information complexity," *Proceedings of the 57h IEEE Symposium on Foundations of Computer Science (FOCS 2016)*, pp. 555–564, 2016.

[11] A. Anshu, S. Ben-David, A. Garg, R. Jain, R. Kothari, and T. Lee, "Separating quantum communication and approximate rank," in *Proceedings of the 32nd Computational Complexity Conference (CCC 2017)*, 2017, pp. 24:1–24:33.

[12] R. Jain, J. Radhakrishnan, and P. Sen, "A lower bound for the bounded round quantum communication complexity of set disjointness," in *Proceedings of the 44th IEEE Symposium on Foundations of Computer Science (FOCS 2003)*, 2003, pp. 220–229.

[13] R. Jain and A. Nayak, "The space complexity of recognizing well-parenthesized expressions in the streaming model: The index function revisited," *IEEE Transactions on Information Theory*, vol. 60, no. 10, pp. 6646–6668, 2014.

[14] D. Touchette, "Quantum information complexity," in *Proceedings of the 47th Annual ACM on Symposium on Theory of Computing (STOC 2015)*, 2015, pp. 317–326.

[15] I. Kerenidis, M. Lauriere, F. Le Gall, and M. Rennela, "Information cost of quantum communication protocols." *Quantum Information & Computation*, vol. 16, no. 3&4, pp. 181–196, 2016.

[16] A. Ganor, G. Kol, and R. Raz, "Exponential separation of information and communication for boolean functions," in *Proceedings of the 47th Annual ACM on Symposium on Theory of Computing (STOC 2015)*, 2015, pp. 557–566.

[17] A. Anshu, D. Touchette, P. Yao, and N. Yu, "Exponential separation of quantum communication and classical information," in *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing (STOC 2017)*, 2017, pp. 277–288.

[18] R. Jain, J. Radhakrishnan, and P. Sen, "Prior entanglement, message compression and privacy in quantum communication," in *20th Annual IEEE Conference on Computational Complexity (CCC'05)*, 2005, pp. 285–296.

[19] M. Laurière and D. Touchette, "The Flow of Information in Interactive Quantum Protocols: the Cost of Forgetting," in *8th Innovations in Theoretical Computer Science Conference (ITCS 2017)*, 2017, pp. 47:1–47:1.

[20] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, 2006.

[21] D. Bures, "An extension of Kakutani's theorem on infinite product measures to the tensor product of semifinite ω^*-algebras," *Transactions of the American Mathematical Society*, vol. 135, pp. 199–212, 1969.

[22] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[23] M. M. Wilde, *Quantum Information Theory*. Cambridge: Cambridge University Press, 12 2012.

[24] J. Watrous, *The theory of quantum information*. Cambridge University Press, 2018.

[25] C. A. Fuchs and J. van de Graaf, "Cryptographic distinguishability measures for quantum-mechanical states," *IEEE Transactions on Information Theory*, vol. 45, no. 4, pp. 1216–1227, 1999.

[26] G. Lindblad, "Completely positive maps and entropy inequalities," *Communications in Mathematical Physics*, vol. 40, no. 2, pp. 147–151, 1975.

[27] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schmacher, "Noncommuting mixed states cannot be broadcast," *Phys. Rev. Lett.*, vol. 76, no. 15, pp. 2818–2821, 1996.

[28] A. Rao and M. Sinha, "Simplified separation of information and communication," *Theory of Computing*, vol. 14, no. 1, pp. 1–29, 2018.

[29] A. Nayak and D. Touchette, "Augmented index and quantum streaming algorithms for DYCK(2)," in *Proceedings of the 32nd Computational Complexity Conference (CCC 2017)*, 2017, pp. 23:1–23:21.

[30] T. Lee and A. Shraibman, "An approximation algorithm for approximation rank," in *Proceedings of the 24th IEEE Conference on Computational Complexity (CCC 2008)*, 2008, pp. 351–357.