

A new operational interpretation of relative entropy and trace distance between quantum states

Anurag Anshu

Centre for Quantum Technologies,
National University of Singapore
a0109169@nus.edu.sg

Rahul Jain

Centre for Quantum Technologies
and Department of Computer
Science, National University of
Singapore
rahul@comp.nus.edu.sg

Priyanka Mukhopadhyay

Centre for Quantum Technologies,
National University of Singapore
a0109168@nus.edu.sg

Ala Shayeghi

Institute for Quantum
Computing, University of
Waterloo
ashayeghi@uwaterloo.ca

Penghui Yao

Centre for Quantum Technologies,
National University of Singapore
phyao1985@gmail.com

April 9, 2014

Abstract

In this paper we present a new operational interpretation of relative-entropy between quantum states in the form of the following protocol.

$\mathcal{P}$: Alice gets to know the eigen-decomposition of a quantum state ρ . Bob gets to know the eigen-decomposition of a quantum state σ . Both Alice and Bob know $S(\rho\|\sigma) \stackrel{\text{def}}{=} \text{Tr} \rho \log \rho - \rho \log \sigma$, the relative entropy between ρ and σ and an error parameter ε . Alice and Bob use shared entanglement and after communication of $\mathcal{O}((S(\rho\|\sigma) + 1)/\varepsilon^4)$ bits from Alice to Bob, Bob ends up with a quantum state $\tilde{\rho}$ such that $F(\rho, \tilde{\rho}) \geq 1 - \varepsilon$, where $F(\cdot)$ represents fidelity.

This result can be considered as a non-commutative generalization of a result due to Braverman and Rao [BR11] where they considered the special case when ρ and σ are classical probability distributions. We use $\mathcal{P}$ to obtain an alternate proof of a direct-sum result for entanglement assisted quantum one-way communication complexity for all relations, which was first shown by Jain, Radhakrishnan and Sen [JRS05, JRS08].

Our second result provides a new operational meaning to trace distance between quantum states in the form of a protocol which can be viewed as a quantum analogue of the classical correlated-sampling protocol, which is widely used, for example by Holenstein [Hol07] in his proof of a parallel-repetition theorem for two-player one-round games. Recently Dinur, Steurer and Vidick [DSV14] have shown another version of a quantum correlated sampling protocol different from our protocol, and used it in their proof of a parallel-repetition theorem for two-prover one-round entangled projection games.

1 Introduction

Relative entropy is a widely used quantity of central importance in both classical and quantum information theory. In this paper we present a new operational meaning to quantum relative entropy in the form of the following protocol.

$\mathcal{P}$: Alice gets to know the eigen-decomposition of a quantum state ρ . Bob gets to know the eigen-decomposition of a quantum state σ . Both Alice and Bob know $S(\rho\|\sigma) \stackrel{\text{def}}{=} \text{Tr} \rho \log \rho - \rho \log \sigma$, the relative entropy between ρ and σ and an error parameter ε . Alice and Bob use shared entanglement and after communication of $\mathcal{O}((S(\rho\|\sigma) + 1)/\varepsilon^4)$ bits from Alice to Bob, Bob ends up with a quantum state $\tilde{\rho}$ such that $F(\rho, \tilde{\rho}) \geq 1 - \varepsilon$, where $F(\cdot)$ represents *fidelity*.

This result can be considered as a non-commutative generalization of a result due to Braverman and Rao [BR11] where they considered the special case when ρ and σ are classical probability distributions. Their protocol, and slightly modified versions of it, were widely used to show several *direct sum* and *direct product* results in communication complexity, for example a direct sum theorem for all relations in the bounded-round public-coin communication model [BR11], direct product theorems for all relations in the public-coin one-way and public-coin bounded-round communication models [Jai13, JPY12, BRWY13]. A direct sum result for a relation f in a model of communication (roughly) states that in order to compute k independent instances of f simultaneously, if we provide communication less than k times the communication required to compute f with the constant success probability $p < 1$, then the success probability for computing all the k instances of f correctly is at most a constant $q < 1$. A direct product result, which is a stronger result, states that in such a situation the success probability for computing all the k instances of f correctly is at most $p^{-\Omega(k)}$.

Protocol $\mathcal{P}$ allows for compressing the communication in one-way entanglement-assisted quantum communication protocols to the *internal information* about the inputs carried by the message. Using this we obtain a direct-sum result for *entanglement assisted quantum one-way communication complexity* for all relations. This direct-sum result was shown previously by Jain, Radhakrishnan and Sen [JRS05, JRS08] and they obtained this result via a protocol that allowed them compression to *external information* carried in the message¹. Their arguments were quite specific to one-way protocols and do not seem to generalize to multi-round communication protocols. Our proof however, is along the lines of a proof which has been generalized to bounded-round classical protocols [BR11] and hence it presents hope that our direct-sum result can also be generalized to bounded-round quantum protocols. The protocol of Braverman and Rao [BR11] was also used by Jain [Jai13] to obtain a direct-product for all relations in the model of one-way public-coin classical communication and later extended to multiple round public-coin classical communication [JPY12, BRWY13]. Hence protocol $\mathcal{P}$ also presents a hope of obtaining similar results for quantum communication protocols.

Our second result provides a new operational meaning to *trace distance* between quantum states in the form of the following protocol.

$\mathcal{P}_1$: Alice gets to know the eigen-decomposition of a quantum state ρ . Bob gets to know the eigen-decomposition of a quantum state σ . Alice and Bob use shared entanglement, do local measurements (no communication) and at the end Alice outputs registers AA_1 and Bob outputs registers BB_1 such that the following holds:

¹Compression to external and internal information can be thought of as one-shot communication analogues of the celebrated results by Shannon [Sha48] and Slepian-Wolf [SW73] exhibiting compression of source to entropy and conditional entropy respectively.

1. The marginal state in A is ρ and the marginal state in B is σ .
2. For any projective measurement $M = \{M_1, \dots, M_w\}$ in the support of the state in AA_1 , the following holds. Let Alice perform M on AA_1 and Bob perform M on BB_1 and obtain outcome $I \in [w], J \in [w]$ respectively. Then,

$$\Pr[I = J] \geq \left(1 - \sqrt{\|\rho - \sigma\|_1 - \frac{1}{4} \|\rho - \sigma\|_1^2}\right)^3.$$

The protocol above can be viewed as a quantum analogue of the *classical correlated-sampling protocol*, which is widely used for example by Holenstein [Hol07] in his proof of a *parallel-repetition theorem* for two-player one-round games. Recently Dinur, Steurer and Vidick [DSV14] have shown another version of a quantum correlated sampling protocol different from ours, and used it in their proof of a parallel-repetition theorem for two-prover one-round entangled projection games.

Our techniques

Our protocol $\mathcal{P}$ is inspired by the protocol of Braverman and Rao [BR11], which as we mentioned, applies to the special case when inputs to Alice and Bob are classical probability distributions P, Q respectively. Let us assume the simpler case first when Alice and Bob know $c = S_\infty(P\|Q) \stackrel{\text{def}}{=} \min\{\lambda \mid P \leq 2^\lambda Q\}$, the *relative min-entropy* between P and Q . In the protocol of [BR11], Alice and Bob share (as public coins) $\{(M_i, R_i) \mid i \in \mathbb{N}\}$ where each (M_i, R_i) is independently and identically distributed uniformly over $\mathcal{U} \times [0, 1]$, where $\mathcal{U}$ is the support of P and Q . Alice accepts index i iff $R_i \leq P(M_i)$ and Bob accepts index i iff $R_i \leq 2^c Q(M_i)$. It is easily argued that for the first index j accepted by Alice, M_j is distributed according to P . Braverman and Rao argue that Alice can communicate this index j to Bob, with high probability, using communication $\mathcal{O}(c)$ bits (for constant ε), using crucially the fact that $P \leq 2^c Q$.

In our protocol, Alice and Bob share infinite copies of the following quantum state

$$|\psi\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{NK}} \sum_{i=1}^N |i\rangle^A |i\rangle^B \otimes \left(\sum_{m=1}^K |m\rangle^{A_1} |m\rangle^{B_1} \right),$$

where registers A, B serve to sample a maximally mixed state in the support of ρ, σ and the registers A_1, B_1 serve to sample uniformly in the interval $[0, 1]$ (in the limit $K \rightarrow \infty$). Again let us assume the simpler case first when Alice and Bob know $c = S_\infty(\rho\|\sigma) \stackrel{\text{def}}{=} \min\{\lambda \mid \rho \leq 2^\lambda \sigma\}$ (here $\leq$ represent the Löwner order), the *relative min-entropy* between ρ and σ . Let $\rho = \sum_i a_i |a_i\rangle \langle a_i|$ and $\sigma = \sum_i b_i |b_i\rangle \langle b_i|$. Alice performs the following projection on registers AA_1 on each copy of $|\psi\rangle$ and accepts the index of a copy iff the projection succeeds.

$$P_A = \sum_i |a_i\rangle \langle a_i| \otimes \left(\sum_{m=1}^{Ka_i} |m\rangle \langle m| \right).$$

Similarly Bob performs the following projection (for appropriately chosen δ) on registers BB_1 on each copy of $|\psi\rangle$ and accepts the index of a copy iff the projection succeeds.

$$P_B = \sum_i |b_i\rangle \langle b_i| \otimes \left(\sum_{m=1}^{K \cdot \min\{2^c b_i / \delta, 1\}} |m\rangle \langle m| \right).$$

Again it is easily argued that (in the limit $K \rightarrow \infty$), the marginal state in B (and also in A) in the first copy of $|\psi\rangle$, with index i , in which Alice succeeds is ρ . Using crucially the fact

that $\rho \leq 2^c \sigma$, we argue that after Alice's measurement succeeds in a copy, Bob's measurement also succeeds with high probability and hence (by the *gentle measurement lemma*) does not disturb the state much in the register B , conditioned on success. We also argue that Alice can communicate the index of this copy to Bob with communication of $\mathcal{O}(c)$ bits (for constant ε).

As can be seen, our protocol is a natural quantum analogue of the protocol of Braverman and Rao [BR11]. However, since ρ and σ may not commute, our analysis deviates significantly from the analysis of [BR11]. We are required to show several new facts related to the non-commuting case while arguing that the protocol still works fine.

We then consider the case in which $S(\rho\|\sigma)$ (instead of $S_\infty(\rho\|\sigma)$) is known to Alice and Bob. The *quantum substate theorem* [JRS02, JN12] implies that there exists a quantum state ρ' , having high fidelity with ρ such that $S_\infty(\rho'\|\sigma) = \mathcal{O}(S(\rho\|\sigma))$. We argue that our protocol is robust with respect to small perturbations in Alice's input and hence works well for the pair (ρ, σ) as well, and uses communication $\mathcal{O}(S(\rho\|\sigma))$ bits. Again this requires us to show new facts related to the non-commuting case.

Related work

Much progress has been made in the last decade towards proving direct sum and direct product conjectures in various models of communication complexity and information theory has played a crucial role in these works. Most of the proofs have build upon elegant one-shot protocols for interesting information theoretic tasks. For example, consider the following task.

T1: Alice gets to know the eigen-decomposition of a quantum state ρ . Alice and Bob get to know the eigen-decomposition of a quantum state σ . They also know $c \stackrel{\text{def}}{=} S(\rho\|\sigma)$, the relative entropy between ρ and σ and an error parameter ε . They use shared entanglement and communication and at the end of the protocol, Bob ends up with a quantum state $\tilde{\rho}$ such that $F(\rho, \tilde{\rho}) \geq 1 - \varepsilon$.

Jain, Radhakrishnan and Sen in [JRS05, JRS08], showed that this task (for constant ε) can be achieved with communication $\mathcal{O}(S(\rho\|\sigma) + 1)$ bits, and this led to direct sum theorems for all relations in entanglement-assisted quantum one-way and entanglement-assisted quantum simultaneous message-passing communication models. They also considered the special case when the inputs to Alice and Bob are probability distributions P, Q respectively and showed that $\mathcal{O}(S(P\|Q) + 1)$ bits of communication can achieve this task (for constant ε). Later an improved result was obtained by Harsha, Jain, Mc. Allester and Radhakrishnan [HJMR10], where they presented a protocol in which Bob is able to sample exactly from P with expected communication $S(P\|Q) + 2 \log S(P\|Q) + \mathcal{O}(1)$. This led to direct sum theorems for all relations in the public-coin randomized one-way, public-coin simultaneous message passing [JRS05, JRS08] and public-coin randomized bounded-round communication models [HJMR10].

Now let us consider the following task.

T2: Alice gets to know functions $o_A, o_B, e_A : \mathcal{U} \rightarrow [0, 1]$ and Bob gets to know functions $o_B, e_B, e_A : \mathcal{U} \rightarrow [0, 1]$, such that the following functions form probability distributions on $\mathcal{U}$: $P(m) \stackrel{\text{def}}{=} o_A(m)e_B(m)$, $Q(m) \stackrel{\text{def}}{=} o_B(m)e_B(m)$ and $R(m) \stackrel{\text{def}}{=} o_A(m)e_A(m)$. They also receive error parameter $\varepsilon > 0$ as common input. They use shared randomness, communication and at the end of the protocol Bob should sample from a distribution P' such that $F(P, P') \geq 1 - \varepsilon$.

Jain, Radhakrishnan and Sen in [JRS05, JRS08], showed that this task (for constant ε) can be achieved with a single message from Alice to Bob consisting of $\mathcal{O}((S(P\|Q) + 1)2^{(S(P\|R) + 1)})$ bits. This was used by them to provide a round-independent direct-sum theorem for the *distributional* two-way communication complexity of all relations under product distributions. This result

was strengthened by Braverman [Bra12] where they considered the case when o_B is not known to Alice and e_A is not known to Bob. They showed that in this case as well the task can be achieved using same communication. This helped in generalizing the round-independent direct-sum result of [JRS05, JRS08] to non-product distributions. Modified versions of Braverman's protocol were later extensively used for example by Braverman and Weinstein [BW12] to show that *information complexity* is lower bounded by the *discrepancy* bound; by Kerenidis, Laplante, Lerays, Roland, and Xiao [KLL⁺12] to show that information complexity is lower bounded by *smooth-rectangle* bound and by Jain and Yao [JY12] to show a direct-product result for all relations in terms of the smooth-rectangle bound.

Jain, Radhakrishnan and Sen in [JRS05, JRS08] showed that the appropriate quantum version of the task **T2** can also be achieved using similar communication. This implied a round-independent direct-sum result for the distributional two-way entanglement-assisted communication complexity of all relations under product distributions. Recently, using a claim obtained in their result, Jain, Pereszlényi and Yao [JPY14] showed a parallel repetition theorem for *two-player one-round entangled free-games*.

Organization

In the next section we present some preliminaries that are needed for our proofs. In section 3 we present the operational interpretation of quantum relative entropy. In section 4 we present our quantum correlated sampling result. We present our direct sum result in section A.

2 Preliminaries

In this section we present some notations, definitions, facts and lemmas that we will use later in our proofs.

Information theory

For integer $n \geq 1$, let $[n]$ represent the set $\{1, 2, \dots, n\}$. We use $\log$ to represent $\log_2$. Let $\mathcal{X}$ and $\mathcal{Y}$ be finite sets and k be a natural number. We let $\mathcal{X}^k$ denote the set $\mathcal{X} \times \dots \times \mathcal{X}$, the cross product of $\mathcal{X}$, k times. Let μ be a probability distribution on $\mathcal{X}$. We let $\mu(x)$ represent the probability of $x \in \mathcal{X}$ 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 $\mathcal{X}$ is defined as $\mathbb{E}_{x \leftarrow X}[f(x)] \stackrel{\text{def}}{=} \sum_{x \in \mathcal{X}} \Pr[X = x] \cdot f(x)$, where $x \leftarrow X$ means that x is drawn according to distribution X .

A quantum state (or just a state) ρ is a positive semi-definite matrix with trace equal to 1. It is called *pure* if and only if the rank is 1. Let $|\psi\rangle$ be a unit vector. With slight abuse of notation, we use ψ to represent the state and also the density matrix $|\psi\rangle\langle\psi|$, associated with $|\psi\rangle$. Let $\overline{|\psi\rangle}$ represent the complex conjugation of $|\psi\rangle$, taken in the computational basis. A classical distribution μ can be viewed as a quantum state with diagonal entries $\mu(x)$ and non-diagonal entries 0. For two quantum states ρ and σ , $\rho \otimes \sigma$ represents the tensor product (Kronecker product) of ρ and σ . A quantum super-operator $\mathcal{E}(\cdot)$ is a completely positive and trace preserving (CPTP) linear map from states to states. Readers can refer to [CT91, NC00, Wat11] for more details.

Definition 2.1. The ℓ_1 -distance (a.k.a trace-distance) between quantum states ρ and σ is given by $\|\rho - \sigma\|_1$, where $\|X\|_1 \stackrel{\text{def}}{=} \text{Tr}\sqrt{X^\dagger X}$ is the sum of the singular values of X . We say that ρ is ε -close to σ in ℓ_1 if $\|\rho - \sigma\|_1 \leq \varepsilon$. The ℓ_2 -distance between them is given by $\|\rho - \sigma\|_2$, where $\|X\|_2 \stackrel{\text{def}}{=} \sqrt{\text{Tr}XX^\dagger}$.

Definition 2.2. The fidelity between quantum states ρ and σ is given by $F(\rho, \sigma) \stackrel{\text{def}}{=} \|\sqrt{\rho}\sqrt{\sigma}\|_1$. For two classical distributions P and Q on set $\mathcal{X}$, $F(P, Q) = \sum_{x \in \mathcal{X}} \sqrt{P(x)Q(x)}$.

The following fact relates the ℓ_1 -distance and the fidelity between two states.

Fact 2.3 ([NC00] page 416). For quantum states ρ and σ , it holds that

$$2(1 - F(\rho, \sigma)) \leq \|\rho - \sigma\|_1 \leq 2\sqrt{1 - F(\rho, \sigma)^2}.$$

For two pure states $|\phi\rangle$ and $|\psi\rangle$, we have

$$\|\phi - \psi\|_1 = 2\sqrt{1 - F(\phi, \psi)^2} = 2\sqrt{1 - |\langle\phi|\psi\rangle|^2}.$$

Let ρ^{AB} be a bipartite quantum state in registers AB . We sometimes use the same symbol to represent a quantum register and the Hilbert space associated with it. We define

$$\rho^B \stackrel{\text{def}}{=} \text{Tr}_A(\rho^{AB}) \stackrel{\text{def}}{=} \sum_i (\langle i| \otimes \mathbb{1}_B) \rho^{AB} (|i\rangle \otimes \mathbb{1}_B),$$

where $\{|i\rangle\}_i$ is an orthonormal basis for the Hilbert space A and $\mathbb{1}_B$ is the identity matrix in space B . The state ρ^B is referred to as the marginal state of ρ^{AB} in register B .

The following proposition states that the distance between two states cannot be increased by quantum operations.

Fact 2.4 ([NC00], page 406, 414). For states ρ, σ , and quantum operation $\mathcal{E}(\cdot)$, it holds that

$$\|\mathcal{E}(\rho) - \mathcal{E}(\sigma)\|_1 \leq \|\rho - \sigma\|_1 \quad \text{and} \quad F(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \geq F(\rho, \sigma).$$

In particular, for bipartite states ρ^{AB} and σ^{AB} , it holds that

$$\|\rho^{AB} - \sigma^{AB}\|_1 \geq \|\rho^A - \sigma^A\|_1 \quad \text{and} \quad F(\rho^{AB}, \sigma^{AB}) \leq F(\rho^A, \sigma^A).$$

Fact 2.5 ([Wat11] Lemma 4.41.). Let A, B be two positive semidefinite operators on Hilbert space $\mathcal{X}$. Then

$$\|A - B\|_1 \geq \|\sqrt{A} - \sqrt{B}\|_2^2.$$

Fact 2.6. Given two quantum states ρ and σ ,

$$\text{Tr}\sqrt{\rho}\sqrt{\sigma} \geq 1 - \frac{1}{2} \|\rho - \sigma\|_1 \geq 1 - \sqrt{1 - F(\rho, \sigma)^2}.$$

Proof. By Facts 2.5 and 2.3,

$$2\sqrt{1 - F(\rho, \sigma)^2} \geq \|\rho - \sigma\|_1 \geq \|\sqrt{\rho} - \sqrt{\sigma}\|_2^2 = 2 - 2 \cdot \text{Tr}(\sqrt{\rho}\sqrt{\sigma}). \quad \square$$

The entropy of a quantum state ρ (in register X) is defined as $S(\rho) \stackrel{\text{def}}{=} -\text{Tr}\rho \log \rho$. We also let $S(X)_\rho$ represent $S(\rho)$. The relative entropy between quantum states ρ and σ is defined as $S(\rho\|\sigma) \stackrel{\text{def}}{=} \text{Tr}\rho \log \rho - \text{Tr}\rho \log \sigma$. The relative min-entropy between them is defined as $S_\infty(\rho\|\sigma) \stackrel{\text{def}}{=} \min\{\lambda : \rho \leq 2^\lambda \sigma\}$. Since logarithm is operator-monotone, we have $S(\rho\|\sigma) \leq S_\infty(\rho\|\sigma)$.

Fact 2.7 ([JRS09, JN12]). (**Quantum substate theorem**) Suppose ρ and σ are two quantum states in the same Hilbert space. Then for any $\varepsilon > 0$, there exists ρ' such that

$$F(\rho, \rho') \geq 1 - \varepsilon \quad \text{and} \quad S_\infty(\rho'\|\sigma) \leq \frac{S(\rho\|\sigma) + 1}{\varepsilon} + \log \frac{1}{1 - \varepsilon}.$$

Fact 2.8 ([Win99, ON02]). (**Gentle measurement lemma**) Let ρ be a density operator and Π be a projector. Then,

$$F(\rho, \frac{\Pi\rho\Pi}{\text{Tr}\Pi\rho}) \geq \sqrt{\text{Tr}\Pi\rho}.$$

Proof. Let $|\phi\rangle$ be a purification of ρ . Then $(\Pi \otimes I)|\phi\rangle$ is a purification of $\Pi\rho\Pi$. Hence (using Fact 2.4)

$$F(\rho, \frac{\Pi\rho\Pi}{\text{Tr}\Pi\rho}) \geq \frac{|\langle\phi|(\Pi \otimes I)|\phi\rangle|}{\|(\Pi \otimes I)|\phi\rangle\|} = \sqrt{\text{Tr}(\Pi\rho)}. \quad \square$$

Communication complexity

Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation. In this work, we are concerned with quantum one-way communication complexity. In this model, Alice holds input $x \in \mathcal{X}$ and Bob holds input $y \in \mathcal{Y}$. They may share a prior quantum state independent of the inputs. Alice makes an arbitrary unitary transformation on her qubits and sends part of her qubits to Bob. Bob makes a unitary operation and measures the last few qubits (answer registers) in the computational basis to get the answer $z \in \mathcal{Z}$. The answer is declared correct if $(x, y, z) \in f$. Let $Q_{\varepsilon}^{\text{ent}, A \rightarrow B}(f)$ represent the quantum one-way communication complexity of f with worst case error ε , that is the communication of the best such protocol computing f with error at most ε on any input (x, y) .

Variants of the following lemma have appeared in many other works such as [BR11, KLL⁺12].

Lemma 2.9. *Let $N, r > 0, \delta \in (0, 1)$. Let Alice and Bob perform multiple identical independent trials such that for each trial,*

1. $\Pr[\text{Alice succeeds}] = \frac{1}{N}; \Pr[\text{Bob succeeds}] \leq \frac{2^r}{N};$
2. $\Pr[\text{Bob succeeds} \mid \text{Alice succeeds}] \geq 1 - \delta.$

There exists a protocol with $\lceil r + 3 \log \frac{1}{\delta} \rceil$ bits of communication from Alice to Bob such that with probability at least $1 - 4\delta$, Alice and Bob choose the same trial.

Proof of this lemma and the description of protocol is deferred to Appendix B.

3 An operational interpretation of quantum relative entropy

Following is our main result in this section.

Theorem 3.1. *Alice is given the spectral decomposition of $\rho = \sum_{i=1}^N a_i |a_i\rangle\langle a_i|$ and Bob is given the spectral decomposition of $\sigma = \sum_{i=1}^N b_i |b_i\rangle\langle b_i|$. Let $S(\rho\|\sigma)$ and $\varepsilon > 0$ be known to Alice and Bob. There exists a protocol, in which Alice and Bob use shared entanglement and Alice sends $\mathcal{O}(S(\rho\|\sigma) + 1)/\varepsilon^4$ bits of communication to Bob such that with probability at least $1 - 4\varepsilon$, the state $\tilde{\rho}$ that Bob gets at the end of the protocol satisfies $F(\rho, \tilde{\rho}) \geq 1 - \varepsilon$.*

Proof. Follows immediately by combining Lemma 2.9 and Lemma 3.2 below. $\square$

Following is our key lemma.

Lemma 3.2. *Fix $\varepsilon > 0$. Let $\delta = (\varepsilon/3)^4$. Let (ρ, σ) be the input for the trial in Figure 1 with $c = S(\rho\|\sigma)$. Let ρ' be a state, guaranteed by the quantum substate theorem (Fact 2.7), such that $F(\rho', \rho) = 1 - \delta$ and $\rho' \leq 2^{c'}\sigma$, with $c' = \frac{c+2}{\delta} \geq \frac{c+1}{\delta} + \log \frac{1}{1-\delta}$. Then,*

1. $\Pr[\text{Alice succeeds}] = \frac{1}{N}; \Pr[\text{Bob succeeds}] \leq \frac{2^{c'}}{\delta N};$

Input: Alice is given the spectral decomposition of $\rho = \sum_{i=1}^N a_i |a_i\rangle\langle a_i|$ and Bob is given the spectral decomposition of $\sigma = \sum_{i=1}^N b_i |b_i\rangle\langle b_i|$. Let $S(\rho||\sigma)$ and $\varepsilon > 0$ be known to both Alice and Bob.

Let $\delta = (\varepsilon/3)^4$ and $c' = (c+2)/\delta$. Let $\{|1\rangle, |2\rangle, \dots, |N\rangle\}$ be an orthonormal basis for $\mathcal{H}$, the support of ρ, σ . We assume that $a_1, a_2, \dots, a_N, 2^{2c'}b_1, 2^{2c'}b_2, \dots, 2^{2c'}b_N$ are rounded to nearest multiple of $1/K$. The error due to this assumption goes to 0 as $K \rightarrow \infty$.

Alice and Bob share the following state where registers AA_1 belong to Alice and registers BB_1 belong to Bob.

$$|S\rangle = \frac{1}{\sqrt{KN}} \sum_{i=1}^N |i, i\rangle^{AB} \otimes \left(\sum_{m=1}^K |m, m\rangle^{A_1 B_1} \right)$$

1. Alice performs the measurement $\{P_A, I - P_A\}$ on the registers AA_1 where,

$$P_A = \sum_i |a_i\rangle\langle a_i| \otimes \left(\sum_{m=1}^{Ka_i} |m\rangle\langle m| \right).$$

She declares success if P_A succeeds.

2. Bob performs the measurement $\{P_B, I - P_B\}$ on the registers BB_1 , where

$$P_B = \sum_i |b_i\rangle\langle b_i| \otimes \left(\sum_{m=1}^{K \cdot \min\{\frac{1}{8}2^{c'}b_i, 1\}} |m\rangle\langle m| \right).$$

He declares success if P_B succeeds.

Figure 1: Trial

2. $\Pr[\text{Bob succeeds} | \text{Alice succeeds}] \geq 1 - \delta - 2\delta^{1/4} \geq 1 - \varepsilon;$
3. *Given that both Alice and Bob succeed, fidelity between ρ and the state of the register B is at least $\sqrt{1 - \delta - 2\delta^{1/4}} \geq 1 - \varepsilon$.*

Proof. 1. Easily verified by direct calculations.

2. We start with the following claim which is of independent interest as well.

Claim 3.3. Let ρ' have the spectral decomposition $\rho' = \sum_i g_i |g_i\rangle\langle g_i|$. For any $p > 0$ and every $|g_i\rangle\langle g_i|$, we have $\sum_{j|b_j \leq p \cdot g_i} |\langle b_j | g_i \rangle|^2 \leq 2^{c'} \cdot p$.

Proof. Since $\rho' \leq 2^{c'}\sigma$, it implies $g_i |g_i\rangle\langle g_i| \leq 2^{c'}\sigma$. Let Π be the projection onto the eigenspace of σ with eigenvalues less than or equal to $p \cdot g_i$. We have $\Pi\sigma\Pi \leq p \cdot g_i \cdot \Pi$. After applying Π on both sides of the equation $g_i |g_i\rangle\langle g_i| \leq 2^{c'}\sigma$ and taking operator norm on both sides, we get $g_i \sum_{j|b_j \leq p \cdot g_i} |\langle b_j | g_i \rangle|^2 \leq 2^{c'} \cdot p \cdot g_i$. This implies the lemma. $\square$

Define

$$|S_A(\rho)\rangle = \frac{1}{\sqrt{K}} \sum_{i=1}^N |a_i\rangle |\bar{a}_i\rangle \otimes \left(\sum_{m=1}^{Ka_i} |m, m\rangle \right);$$

$$|S_A(\rho')\rangle = \frac{1}{\sqrt{K}} \sum_{i=1}^N |g_i\rangle |\bar{g}_i\rangle \otimes \left(\sum_{m=1}^{Kg_i} |m, m\rangle \right).$$

Here $|\bar{a}_i\rangle$ (similarly $|\bar{g}_i\rangle$) is the state obtained by taking complex conjugate of $|a_i\rangle$, with respect to the basis $\{|1\rangle, |2\rangle \dots |N\rangle\}$.

The following claim asserts that $|S_A(\rho)\rangle$ and $|S_A(\rho')\rangle$ are close if ρ and ρ' are close.

Claim 3.4. $|\langle S_A(\rho) | S_A(\rho') \rangle| \geq 1 - 2(1 - F(\rho, \rho'))^{1/4}$.

Proof. Define $R_{ij} \stackrel{\text{def}}{=} a_i |\langle a_i | g_j \rangle|^2$ and $R'_{ij} \stackrel{\text{def}}{=} g_i |\langle a_i | g_j \rangle|^2$. Note that both $\{R_{ij}\}$ and $\{R'_{ij}\}$ form probability distributions over $[N^2]$. Also note that $F(R, R') = \text{Tr}(\sqrt{\rho}\sqrt{\rho'})$. Consider (using Facts 2.6 and 2.3),

$$\begin{aligned} |\langle S_A(\rho) | S_A(\rho') \rangle| &= \sum_{i,j} \min(R_{ij}, R'_{ij}) = 1 - \frac{1}{2} \|R - R'\|_1 \\ &\geq 1 - \sqrt{1 - F(R, R')^2} = 1 - \sqrt{1 - (\text{Tr} \sqrt{\rho} \sqrt{\rho'})^2} \\ &\geq 1 - \sqrt{2(1 - \text{Tr} \sqrt{\rho} \sqrt{\rho'})} \geq 1 - \sqrt{2\sqrt{1 - F(\rho, \rho')}} \\ &\geq 1 - 2(1 - F(\rho, \rho'))^{1/4}. \end{aligned} \quad \square$$

Consider,

$$(I_A \otimes P_B) |S_A(\rho')\rangle = \frac{1}{\sqrt{K}} \sum_{i,j=1}^N |\bar{g}_j\rangle |b_i\rangle \langle b_i | g_j \rangle \left(\sum_{m=1}^{K \cdot \min\{g_j, \frac{1}{\delta} 2^{c'} b_i\}} |m, m\rangle \right).$$

Therefore,

$$\begin{aligned} \|(I_A \otimes P_B) |S_A(\rho')\rangle\|^2 &= \sum_{i,j=1}^N |\langle b_i | g_j \rangle|^2 \min\{g_j, \frac{1}{\delta} 2^{c'} b_i\} \\ &\geq \sum_{j=1}^N g_j \left(\sum_{i | b_i \geq \delta 2^{-c'} g_j} |\langle b_i | g_j \rangle|^2 \right) \geq \sum_{j=1}^N g_j (1 - \delta) = 1 - \delta. \quad (\text{using Claim 3.3}) \end{aligned} \quad (1)$$

Using the above,

$$\begin{aligned} \text{Pr}[\text{Bob succeeds} | \text{Alice succeeds}] &= \text{Tr}(I_A \otimes P_B) |S_A(\rho)\rangle \langle S_A(\rho)| \\ &\geq \text{Tr}(I_A \otimes P_B) |S_A(\rho')\rangle \langle S_A(\rho')| - \frac{1}{2} \|S_A(\rho) - S_A(\rho')\|_1 \\ &= \text{Tr}(I_A \otimes P_B) |S_A(\rho')\rangle \langle S_A(\rho')| - \sqrt{1 - |\langle S_A(\rho) | S_A(\rho') \rangle|^2} \quad (\text{Fact 2.3}) \\ &\geq 1 - \delta - 2\sqrt{(1 - F(\rho, \rho'))^{1/2}}. \quad (\text{Claim 3.4 and Eq. (1)}) \end{aligned}$$

3. From Fact 2.8,

$$F(S_A(\rho), \frac{(I_A \otimes P_B) |S_A(\rho)\rangle \langle S_A(\rho)| (I_A \otimes P_B)}{\text{Tr}(I_A \otimes P_B) |S_A(\rho)\rangle \langle S_A(\rho)|}) \geq \sqrt{\text{Tr}(I_A \otimes P_B) |S_A(\rho)\rangle \langle S_A(\rho)|}.$$

Since the marginal of $|S_A(\rho)\rangle$ on register B is ρ and partial trace does not decrease fidelity (Fact 2.4), using item 2. above, the desired follows. $\square$

As a consequence of Theorem 3.1 we obtain the following direct sum result for all relations in the model of entanglement-assisted one-way communication complexity. Its proof is deferred to Appendix A.

Theorem 3.5. *Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be finite sets, $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation, $0 < \varepsilon, \delta < \frac{1}{10}$ and $k > 1$ be an integer. We have*

$$Q_{\varepsilon}^{ent, A \rightarrow B}(f^k) \geq \Omega\left(\delta^4 \cdot k \left(Q_{\varepsilon+\delta}^{ent, A \rightarrow B}(f) - 1\right)\right).$$

4 Quantum correlated sampling

Theorem 4.1. *Alice is given the spectral decomposition of $\rho = \sum_{i=1}^N a_i |a_i\rangle\langle a_i|$ and Bob is given the spectral decomposition of $\sigma = \sum_{i=1}^N b_i |b_i\rangle\langle b_i|$. There exists a zero-communication protocol (Figure 2) satisfying the following.*

1. *Alice outputs registers A, A_1 and Bob outputs registers B, B_1 respectively, such that state in A is ρ and the state in B is σ .*
2. *Let $M = \{M_1, M_2 \dots M_w\}$ be a projective measurement, in the support of AA_1 . Let M be performed by Alice on the joint system AA_1 with outcome $I \in [w]$ and by Bob on the joint system BB_1 with outcome $J \in [w]$. Then $\Pr[I = J] \geq \left(1 - \sqrt{\|\rho - \sigma\|_1} - \frac{1}{4} \|\rho - \sigma\|_1^2}\right)^3$.*

Input: Alice is given the spectral decomposition of $\rho = \sum_{i=1}^N a_i |a_i\rangle\langle a_i|$ and Bob is given the spectral decomposition of $\sigma = \sum_{i=1}^N b_i |b_i\rangle\langle b_i|$.

Let $\{|1\rangle, |2\rangle \dots |N\rangle\}$ be an orthonormal basis for Hilbert space $\mathcal{H}$, support of ρ, σ . We assume that $a_1, \dots, a_N, b_1, \dots, b_N$ are rounded to nearest multiple of $1/K$. The error due to this assumption goes to 0 as $K \rightarrow \infty$.

Alice and Bob share infinite copies of the following state:

$$|S\rangle = \frac{1}{\sqrt{KN}} \sum_{i=1}^N |i, i\rangle^{AB} \otimes \left(\sum_{m=1}^K |m, m\rangle^{A_1 B_1} \right)$$

1. Alice performs the measurement $\{P_A, I - P_A\}$ on the registers AA_1 in each copy of $|S\rangle$ where,

$$P_A = \sum_i |a_i\rangle\langle a_i| \otimes \left(\sum_{m=1}^{Ka_i} |m\rangle\langle m| \right)$$

For the first copy in which P_A succeeds, she outputs the registers AA_1 .

2. Bob performs the measurement $\{P_B, I - P_B\}$ on the registers BB_1 in each copy of $|S\rangle$ where,

$$P_B = \sum_i |b_i\rangle\langle b_i| \otimes \left(\sum_{m=1}^{Kb_i} |m\rangle\langle m| \right)$$

For the first copy in which P_B succeeds, he outputs the registers BB_1 .

Figure 2: Quantum correlated sampling

Proof. Let the joint state in the registers output by Alice and Bob at the end of the protocol be τ . The following claim shows the first part of the theorem.

Claim 4.2. $\text{Tr}_{A_1 B B_1}(\tau) = \rho$ and $\text{Tr}_{B_1 A A_1}(\tau) = \sigma$.

Proof. It is easily seen that the state in register A in $(P_A \otimes I) |S\rangle$ is ρ . Similarly the state in register B in $(I \otimes P_B) |S\rangle$ is σ . $\square$

Claim 4.3.

$$\tau \geq \frac{(P_A \otimes P_B) |S\rangle \langle S| (P_A \otimes P_B)}{1 - \langle S| (I - P_A) \otimes (I - P_B) |S\rangle}.$$

Proof. Consider the event that Alice and Bob succeed at the same index. The resulting state in $AA_1 B B_1$ is

$$\frac{(P_A \otimes P_B) |S\rangle \langle S| (P_A \otimes P_B)}{\langle S| (P_A \otimes P_B) |S\rangle},$$

and this event occurs with probability

$$\sum_{i=0}^{\infty} \langle S| (I - P_A) \otimes (I - P_B) |S\rangle^i \cdot \langle S| (P_A \otimes P_B) |S\rangle = \frac{\langle S| (P_A \otimes P_B) |S\rangle}{1 - \langle S| (I - P_A) \otimes (I - P_B) |S\rangle}.$$

Since the cases of Bob succeeding before Alice and vice versa add positive operators to τ , we get the desired. $\square$

Claim 4.4. Let $|\theta\rangle \stackrel{\text{def}}{=} \frac{(P_A \otimes P_A) |S\rangle}{\|(P_A \otimes P_A) |S\rangle\|}$. Then

$$\langle \theta | \tau | \theta \rangle \geq \frac{\left(1 - \sqrt{\|\rho - \sigma\|_1 - \frac{1}{4} \|\rho - \sigma\|_1^2}\right)^2}{1 + \sqrt{\|\rho - \sigma\|_1 - \frac{1}{4} \|\rho - \sigma\|_1^2}} \geq \left(1 - \sqrt{\|\rho - \sigma\|_1 - \frac{1}{4} \|\rho - \sigma\|_1^2}\right)^3.$$

Proof. Consider,

$$\begin{aligned} \langle \theta | \tau | \theta \rangle &\geq \frac{|\langle \theta | P_A \otimes P_B |S\rangle|^2}{1 - \langle S| (I - P_A) \otimes (I - P_B) |S\rangle} && \text{(Claim 4.3)} \\ &= \frac{|\langle \theta | P_A \otimes P_B |S\rangle|^2}{2/N - \langle S| P_A \otimes P_B |S\rangle} && \text{(using } \langle S| P_A \otimes I |S\rangle = \langle S| I \otimes P_B |S\rangle = 1/N). \end{aligned}$$

By direct calculation, we get

$$\begin{aligned} (P_A \otimes P_B) |S\rangle &= \frac{1}{\sqrt{KN}} \sum_{i,j} |\bar{a}_i\rangle \langle b_j | a_i \rangle |b_j\rangle \sum_{m=1}^{K \min(a_i, b_j)} |m, m\rangle; \\ |\theta\rangle &= \frac{1}{\sqrt{K}} \sum_i |\bar{a}_i\rangle |a_i\rangle \sum_{m=1}^{K a_i} |m, m\rangle. \end{aligned}$$

Hence,

$$\langle \theta | \tau | \theta \rangle \geq \frac{\left(\sum_{i,j} \min(a_i, b_j) |\langle a_i | b_j \rangle|^2\right)^2}{2 - \sum_{i,j} \min(a_i, b_j) |\langle a_i | b_j \rangle|^2}. \quad (2)$$

Define $R_{ij} \stackrel{\text{def}}{=} a_i |\langle a_i | b_j \rangle|^2$ and $R'_{ij} \stackrel{\text{def}}{=} b_i |\langle a_i | b_j \rangle|^2$. Note that both $\{R_{ij}\}$ and $\{R'_{ij}\}$ form probability distributions over $[N^2]$. Also note that $F(R, R') = \text{Tr}(\sqrt{\rho}\sqrt{\sigma})$. Consider (using Facts 2.6 and 2.3),

$$\begin{aligned} \sum_{i,j} \min(R_{ij}, R'_{ij}) &= 1 - \frac{1}{2} \|R - R'\|_1 \geq 1 - \sqrt{1 - F(R, R')^2} \\ &= 1 - \sqrt{1 - (\text{Tr}\sqrt{\rho}\sqrt{\sigma})^2} \geq 1 - \sqrt{\|\rho - \sigma\|_1 - \frac{1}{4} \|\rho - \sigma\|_1^2}. \end{aligned} \quad (3)$$

Combining Equations (2) and (3) we get the desired. $\square$

Claim 4.5. Let $M = \{M_1, M_2 \dots M_w\}$ be a projective measurement in the support of AA_1 . Let $E = \sum_{i=1}^w M_i \otimes M_i$. Then $\text{Tr} E |\theta\rangle\langle\theta| = 1$.

Proof. Since M_i is a projector in the support of AA_1 , we have $(M_i \otimes M_i) |\theta\rangle = (M_i \otimes I) |\theta\rangle$. Hence,

$$\langle\theta| E |\theta\rangle = \sum_i \langle\theta| M_i \otimes M_i |\theta\rangle = \sum_i \langle\theta| M_i \otimes I |\theta\rangle = 1. \quad \square$$

Finally using Fact 2.4 and Claim 4.4 we get the second part of the theorem as follows.

$$\sqrt{\text{Tr}(E\tau)} \geq F(\tau, |\theta\rangle\langle\theta|) = \sqrt{\langle\theta| \tau |\theta\rangle} \geq \left(1 - \sqrt{\|\rho - \sigma\|_1 - \frac{1}{4} \|\rho - \sigma\|_1^2}\right)^{3/2}. \quad \square$$

Open questions

Some interesting open questions related to this work are as follows.

1. Can we show a direct product result for all relations in the one-way entanglement assisted communication model ?
2. Can we show a direct sum (and also possibly direct product) result for all relations in the bounded-round entanglement assisted communication model ?
3. Can we find other interesting applications of the protocols appearing in this work ?

Acknowledgment

We thank Ashwin Nayak for helpful discussions. This work is supported by the internal grants of the Center for Quantum Technologies (CQT), Singapore. Work of A.S. done while visiting CQT, Singapore.

References

- [BR11] Mark Braverman and Anup Rao. Information equals amortized communication. In *Proceedings of the 52nd Symposium on Foundations of Computer Science, FOCS '11*, pages 748–757, Washington, DC, USA, 2011. IEEE Computer Society.
- [Bra12] Mark Braverman. Interactive information complexity. In *Proceedings of the 44th annual ACM Symposium on Theory of Computing, STOC '12*, pages 505–524, New York, NY, USA, 2012. ACM.

- [BRWY13] Mark Braverman, Anup Rao, Omri Weinstein, and Amir Yehudayoff. Direct product via round-preserving compression. In *Proceedings of the 40th international conference on Automata, languages and programming, ICALP'13*, Berlin, Heidelberg, 2013. Springer-Verlag.
- [BW12] Mark Braverman and Omri Weinstein. A discrepancy lower bound for information complexity. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, Lecture Notes in Computer Science, pages 459–470. Springer Berlin Heidelberg, 2012.
- [CT91] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley Series in Telecommunications. John Wiley & Sons, New York, NY, USA, 1991.
- [DSV14] Irit Dinur, David Steurer, and Thomas Vidick. A parallel repetition theorem for entangled projection games. In *Proceedings of the 29th IEEE Annual Conference on Computational Complexity, CCC '14*, to appear, Washington, DC, USA, 2014. IEEE Computer Society.
- [HJMR10] P. Harsha, R. Jain, D. McAllester, and J. Radhakrishnan. The communication complexity of correlation. *IEEE Transactions on Information Theory*, 56(1):438–449, Jan 2010.
- [Hol07] Thomas Holenstein. Parallel repetition: simplifications and the no-signaling case. In *Proceedings of the thirty-ninth annual ACM Symposium on Theory of Computing, STOC '07*, pages 411–419, New York, NY, USA, 2007.
- [Jai13] Rahul Jain. New strong direct product results in communication complexity. *Journal of the ACM*, to appear, 2013.
- [JN12] Rahul Jain and Ashwin Nayak. Short proofs of the quantum substate theorem. *IEEE Transactions on Information Theory*, 58(6):3664–3669, 2012.
- [JPY12] Rahul Jain, Attila Pereszlényi, and Penghui Yao. A direct product theorem for the two-party bounded-round public-coin communication complexity. In *Proceedings of the 2012 IEEE 53rd Annual Symposium on Foundations of Computer Science, FOCS '12*, pages 167–176, Washington, DC, USA, 2012.
- [JPY14] Rahul Jain, Attila Pereszlényi, and Penghui Yao. A parallel repetition theorem for entangled two-player one-round games under product distributions. In *Proceedings of the 29th IEEE Annual Conference on Computational Complexity, CCC '14*, to appear, Washington, DC, USA, 2014. IEEE Computer Society.
- [JRS02] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. Privacy and interaction in quantum communication complexity and a theorem about the relative entropy of quantum states. In *Proceedings of the 43rd Symposium on Foundations of Computer Science, FOCS '02*, pages 429–438, Washington, DC, USA, 2002. IEEE Computer Society.
- [JRS05] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. Prior entanglement, message compression and privacy in quantum communication. In *Proceedings of the 20th Annual IEEE Conference on Computational Complexity*, pages 285–296, Washington, DC, USA, 2005. IEEE Computer Society.

- [JRS08] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. Optimal direct sum and privacy trade-off results for quantum and classical communication complexity. *CoRR*, abs/0807.1267, 2008.
- [JRS09] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. A new information-theoretic property about quantum states with an application to privacy in quantum communication. *Journal of the ACM*, 56(6), September 2009. Article no. 33.
- [JY12] Rahul Jain and Penghui Yao. A strong direct product theorem in terms of the smooth rectangle bound. *CoRR*, abs/1209.0263, 2012.
- [KLL⁺12] Iordanis Kerenidis, Sophie Laplante, Virginie Lerays, Jeremie Roland, and David Xiao. Lower bounds on information complexity via zero-communication protocols and applications. In *Proceedings of the 2012 IEEE 53rd Annual Symposium on Foundations of Computer Science, FOCS '12*, pages 500–509, Washington, DC, USA, 2012. IEEE Computer Society.
- [NC00] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, UK, 2000.
- [ON02] T. Ogawa and H. Nagaoka. A new proof of the channel coding theorem via hypothesis testing in quantum information theory. In *Information Theory, 2002. Proceedings. 2002 IEEE International Symposium on*, pages 73–, 2002.
- [Sha48] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948.
- [SW73] David Slepian and Jack K. Wolf. Noiseless coding of correlated information sources. *IEEE Transactions on Information Theory*, 19(4):471–480, 1973.
- [Wat11] John Watrous. Theory of Quantum Information, lecture notes, <https://cs.uwaterloo.ca/~watrous/LectureNotes.html>, 2011.
- [Win99] Andreas Winter. Coding theorem and strong converse for quantum channels. *IEEE Transactions on Information Theory*, 45(7):2481–2485, 1999.
- [Yao79] Andrew C. Yao. Some complexity questions related to distributive computing (preliminary report). In *Proceedings of the eleventh annual ACM Symposium on Theory of Computing, STOC '79*, pages 209–213, New York, NY, USA, 1979. ACM.

A A direct sum theorem for quantum one-way communication complexity

We start with some preliminaries needed for our proof.

Let ρ^{XY} be a quantum state in space $X \otimes Y$. The *conditional entropy* is defined as $S(X|Y) \stackrel{\text{def}}{=} S(\rho^{XY}) - S(\rho^Y)$. The *mutual information* between registers X and Y is defined as

$$I(X : Y)_\rho \stackrel{\text{def}}{=} S(X) - S(X|Y) = S(X)_\rho + S(Y)_\rho - S(XY)_\rho.$$

It is easy to see that $I(X : Y)_\rho = S(\rho \| \rho^X \otimes \rho^Y)$. If X is a classical register, namely $\rho = \sum_x \mu(x) |x\rangle\langle x| \otimes \rho_x$, where μ is a probability distribution over X , then

$$I(X : Y)_\rho = S(Y)_\rho - S(Y|X)_\rho = S\left(\sum_x \mu(x) \rho_x\right) - \sum_x \mu(x) S(\rho_x),$$

where the conditional entropy is defined as $S(Y|X)_\rho \stackrel{\text{def}}{=} \mathbb{E}_{x \leftarrow \mu}[S(\rho_x)]$. Let ρ^{XYZ} be a quantum state with Y being a classical register. The mutual information between X and Z , conditioned on Y , is defined as

$$I(X : Z|Y)_\rho \stackrel{\text{def}}{=} \mathbb{E}_{y \leftarrow Y} [I(X : Z|Y = y)_\rho] = S(X|Y)_\rho + S(Z|Y)_\rho - S(XZ|Y)_\rho.$$

The following *chain rule* for mutual information follows easily from the definitions, when Y is a classical register.

$$I(X : YZ)_\rho = I(X : Y)_\rho + I(X : Z|Y)_\rho.$$

We will need the following basic facts.

Fact A.1 ([NC00], page 515). For any joint quantum system AB , it holds that $|S(A) - S(B)| \leq S(AB) \leq S(A) + S(B)$. Hence

$$I(A : B) = S(A) + S(B) - S(AB) \leq 2S(A).$$

We have the following chain rules for relative entropy and mutual information.

Fact A.2. Let $\rho = \sum_x \mu(x) |x\rangle\langle x| \otimes \rho_x$ and $\rho^1 = \sum_x \mu^1(x) |x\rangle\langle x| \otimes \rho_x^1$. It holds that

$$S(\rho^1 \| \rho) = S(\mu^1 \| \mu) + \mathbb{E}_{x \leftarrow \mu^1} [S(\rho_x^1 \| \rho_x)].$$

Fact A.3. Let XM be a joint system where $X = X_1 \cdots X_k$ are independent. Then,

$$I(X : M) \geq \sum_{i=1}^k I(X_i : M).$$

Fact A.4. Relative entropy is non-increasing when subsystems are considered. Let ρ^{XY} and σ^{XY} be quantum states, then $S(\rho^{XY} \| \sigma^{XY}) \geq S(\rho^X \| \sigma^X)$.

We let $Q_\varepsilon^{\text{ent}, A \rightarrow B, \mu}(f)$ represent distributional quantum one-way communication complexity of f under μ with distributional error at most ε , that is the communication of the best such protocol computing f with error averaged over μ upper bounded by ε . Following is Yao's min-max theorem connecting the worst case error and the distributional error settings.

Fact A.5. [Yao79] $Q_\varepsilon^{\text{ent}, A \rightarrow B}(f) = \max_\mu Q_\varepsilon^{\text{ent}, A \rightarrow B, \mu}(f)$.

Proof of Theorem 3.5: Let μ be any distribution over $\mathcal{X} \times \mathcal{Y}$. We show the following, which combined with Fact A.5 implies the desired.

$$Q_\varepsilon^{\text{ent}, A \rightarrow B, \mu^k}(f^k) \geq \Omega\left(\delta^4 \cdot k \left(Q_{\varepsilon+\delta}^{\text{ent}, A \rightarrow B, \mu}(f) - 1\right)\right).$$

Let $\mathcal{P}$ be a quantum one-way protocol with communication $c \cdot k$ computing f^k with overall probability of success at least $1 - \varepsilon$ under distribution μ^k . Let the following be the global state after Alice sends the message to Bob and before Bob does any operation.

$$\rho \stackrel{\text{def}}{=} \sum_{xy} \mu^k(x, y) |xy\rangle\langle xy|^{XY} \otimes |\psi_{xy}\rangle\langle\psi_{xy}|^{AB}.$$

Let $D = D_1 \cdots D_k$ be uniformly distributed over $\{0, 1\}^k$ and independent of the input XY . Let $U_i = X_i$ if $D_i = 0$ and $U_i = Y_i$ if $D_i = 1$. Note that $B = MB'$ where M represents the message

Alice sends to Bob and B' be the part Bob holds initially. It holds that $|M| \leq ck$ and $B' = |0\rangle\langle 0|$ is independent of everything. Hence

$$I(XYDU : B)_\rho = I(XYDU : B')_\rho + I(XYDU : M|B')_\rho \leq 2S(M) \leq 2ck,$$

where the last inequality is from Fact A.1. Applying the chain rule, we have (below $-i$ represents $[k] - \{i\}$)

$$\begin{aligned} 2c &\geq I(XY : B|DU)_\rho \geq \sum_{i=1}^k I(X_i Y_i : B|DU)_\rho \\ &= \frac{1}{2} \left(\sum_{i=1}^k \left(I(X_i : B|Y_i D_{-i} U_{-i})_\rho + I(Y_i : B|X_i D_{-i} U_{-i})_\rho \right) \right) = \frac{1}{2} \sum_{i=1}^k I(X_i : B|Y_i D_{-i} U_{-i})_\rho, \end{aligned}$$

where the first equality is from the definition of DU and the second equality is from the fact that conditioned on X_i , Y_i is independent of everything else.

Hence there exists $j \in [k]$ such that

$$I(X_j : B|Y_j D_{-j} U_{-j})_\rho \leq 4c. \quad (4)$$

We also have

$$I(X_j Y_j : D_{-j} U_{-j})_\rho = 0. \quad ; \quad I(Y : B|X_j D_{-j} U_{-j})_\rho = 0. \quad (5)$$

Above holds from definitions and because Y is independent of everything conditioned on $D_{-j} U_{-j} X_j$.

We exhibit an entanglement-assisted one-way protocol $\mathcal{Q}$ for f with communication less than c and distributional error ε under distribution μ . Given input $(x, y) \sim \mu$, Alice and Bob embed the input to the j -th coordinate $X_j Y_j$. They share public coins according to distribution $D_{-j} U_{-j}$, which are independent of the inputs by (5). From (5),

$$\mathbb{E}_{x_j y_j d_{-j} u_{-j} \leftarrow X_j Y_j D_{-j} U_{-j}} \left[S \left(\rho_{x_j y_j d_{-j} u_{-j}}^B \parallel \rho_{x_j d_{-j} u_{-j}}^B \right) \right] = 0,$$

which implies $\rho_{x_j y_j d_{-j} u_{-j}}^B = \rho_{x_j d_{-j} u_{-j}}^B$ for all $x_j, y_j, d_{-j} u_{-j}$. From (4),

$$\mathbb{E}_{x_j y_j d_{-j} u_{-j} \leftarrow X_j Y_j D_{-j} U_{-j}} \left[S \left(\rho_{x_j y_j d_{-j} u_{-j}}^B \parallel \rho_{y_j d_{-j} u_{-j}}^B \right) \right] \leq 4c.$$

Note that given input (x_j, y_j) and shared public coins $d_{-j} u_{-j}$, Alice knows the state $\rho_{x_j y_j d_{-j} u_{-j}}^B = \rho_{x_j d_{-j} u_{-j}}^B$, which is the actual state needed to transmit to Bob. Bob knows the state $\rho_{y_j d_{-j} u_{-j}}^B$. Let

$$G \stackrel{\text{def}}{=} \left\{ (x_j, y_j, d_{-j}, u_{-j}) : S \left(\rho_{x_j y_j d_{-j} u_{-j}}^B \parallel \rho_{y_j d_{-j} u_{-j}}^B \right) \leq 4c/\delta \right\}.$$

By Markov inequality,

$$\Pr[X_j Y_j D_{-j} U_{-j} \in G] \geq 1 - \delta.$$

Using the protocol in Theorem 3.1, for all $(x_j, y_j, d_{-j}, u_{-j}) \in G$, Alice is able to transmit a state δ -close in ℓ_1 distance to $\rho_{x_j y_j d_{-j} u_{-j}}^B$ with $\mathcal{O}(c+1)/\delta^2$ communication. Here we use Fact 2.6 to convert fidelity to trace distance. Bob then creates the pure state corresponding to remaining Y and then acts as in $\mathcal{P}$ to output the answer. The overall distributional error in $\mathcal{Q}$ is therefore at most $2\delta + \varepsilon$. Hence (resetting $\delta \rightarrow 2\delta$)

$$Q_{\varepsilon+\delta}^{\text{ent}, A \rightarrow B, \mu}(f) \leq \mathcal{O}((c+1)/\delta^4),$$

which implies

$$Q_{\varepsilon}^{\text{ent}, A \rightarrow B, \mu^k}(f^k) \geq \Omega \left(\delta^4 \cdot k \left(Q_{\varepsilon+\delta}^{\text{ent}, A \rightarrow B, \mu}(f) - 1 \right) \right).$$

□

B Proof of Lemma 2.9

Proof. The protocol is given in Figure 3. The communication cost is $\lceil \log \log \frac{1}{\delta} \rceil + \lceil r + 2 \log \frac{1}{\delta} \rceil \leq \lceil r + 3 \log \frac{1}{\delta} \rceil$. To prove the probability of success, let us define the following "bad" events.

Definition B.1. Let

- B_1 represents the event that the length of the binary representation of k exceeds $\lceil \log \log \frac{1}{\delta} \rceil$;
- B_2 represents the event that $i \notin S_B$ conditioning on $\neg B_1$;
- B_3 represents the event that $j \neq i$ conditioning on $\neg B_1$.

Alice and Bob share infinitely many random hash functions $h_1, h_2, \dots$, where each $h_l : \{0, \dots, N-1\} \rightarrow \{0, 1\}$.

1. Alice selects the first index i where she succeeds and sends to Bob the binary encoding of $k = \lceil i/N \rceil$ using $\lceil \log \log \frac{1}{\delta} \rceil$ bits (if k is too big, she send 0 meaning abort.) Alice outputs i .
 2. Alice sends $\{h_l(i \bmod N) \mid l \in [\lceil r + 2 \log \frac{1}{\delta} \rceil]\}$ to Bob.
 3. Bob selects the first index j in $S_B \stackrel{\text{def}}{=} \{t \mid \text{Bob succeeds on index } t\} \cap \{(k-1)N, \dots, kN-1\}$ such that $\forall l \in [\lceil r + 2 \log \frac{1}{\delta} \rceil] : h_l(j \bmod n) = h_l(i \bmod n)$ and outputs j (if no such index exists, he outputs 0).
-

Figure 3: Protocol

Claim B.2. It holds that: 1. $\Pr[B_1] \leq \delta$; 2. $\Pr[B_2] \leq \delta$; 3. $\Pr[B_3] \leq 3\delta$.

Proof. 1. $\Pr[B_1] \leq \left(1 - \frac{1}{N}\right)^{N \cdot 2^{\lceil \log \log \frac{1}{\delta} \rceil}} \leq \exp^{-2^{\lceil \log \log \frac{1}{\delta} \rceil}} \leq \delta$.

2. Follows from item 2 in Lemma 2.9.

3. For this argument we condition on $\neg B_1$ for all events below. From item 1 in Lemma 2.9 and Markov's inequality,

$$\Pr\left[|S_B| \geq \frac{2^r}{\delta}\right] \leq \frac{\delta}{2^r} \cdot \mathbb{E}[|S_B|] \leq \delta. \quad (6)$$

Thus

$$\begin{aligned} \Pr[B_3] &\leq \Pr[|S_B| \geq 2^r/\delta \text{ or } i \notin S_B] + \Pr[B_3 \mid i \in S_B \text{ and } |S_B| \leq 2^r/\delta] \\ &\leq 2\delta + \Pr[B_3 \mid i \in S_B \text{ and } |S_B| \leq 2^r/\delta] \quad (\text{Eq. (6) and item 2. of this claim}) \\ &\leq 2\delta + 2^{-\lceil r + 2 \log \frac{1}{\delta} \rceil} \cdot \frac{2^r}{\delta} \leq 3\delta. \end{aligned}$$

□

We conclude the result using above and item 1. of this claim:

$$\Pr[j \neq i] \leq \Pr[B_1] + \Pr[\neg B_1] \cdot \Pr[B_3] \leq 4\delta.$$

□