New One Shot Quantum Protocols With Application to Communication Complexity

Anurag Anshu, Rahul Jain, Priyanka Mukhopadhyay, Ala Shayeghi, and Penghui Yao

***Abstract*— In this paper, we present the following quantum compression protocol '$\mathcal{P}$': Let ρ, σ be quantum states, such that $S(\rho\|\sigma) \stackrel{\text{def}}{=} \mathrm{Tr}(\rho\log\rho - \rho\log\sigma)$, the relative entropy between ρ and σ, is finite. Alice gets to know the eigendecomposition of ρ. Bob gets to know the eigendecomposition of σ. Both Alice and Bob know $S(\rho\|\sigma)$ 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 $F(\rho, \tilde{\rho}) \geq 1-5\varepsilon$, where $F(\cdot)$ represents fidelity. This result can be considered as a non-commutative generalization of a result due to Braverman and Rao where they considered the special case when ρ and σ are classical probability distributions (or commute with each other) and use shared randomness instead of shared entanglement. 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 *et al.*. We also present a variant of protocol $\mathcal{P}$ in which Bob has some side information about the state with Alice. We show that in such a case, the amount of communication can be further reduced, based on the side information that Bob has. Our second result provides a quantum analog of the widely used classical correlated-sampling protocol. For example, Holenstein used the classical correlated-sampling protocol in his proof of a parallel-repetition theorem for two-player one-round games.**

***Index Terms*— Quantum information theory, quantum communication complexity, compression protocols, correlated sampling, direct sum results.**

I. Introduction

RELATIVE entropy is a widely used quantity of central importance in both classical and quantum information theory. In this paper we consider the following task. The notations used below are described in section II.

Manuscript received December 9, 2014; revised September 28, 2015; accepted September 22, 2016. Date of publication October 10, 2016; date of current version November 18, 2016. This work was supported in part by the Singapore Ministry of Education Academic Research Fund Tier 3 under Grant MOE2012-T3-1-009 and also in part by the Core Grants of the Center for Quantum Technologies, Singapore. This paper was presented as a poster at the 18th Conference on Quantum Information Processing, QIP 2015.

A. Anshu and P. Mukhopadhyay are with the Centre for Quantum Technologies, National University of Singapore, Singapore 117543 (e-mail: a0109169@u.nus.edu; a0109168@nus.edu.sg).

R. Jain is with the Centre for Quantum Technologies, Department of Computer Science, National University of Singapore, Singapore 117543, and also with the MajuLab, CNRS-UNS-NUS-NTU International Joint Research Unit, UMI 3654, Singapore (e-mail: rahul@comp.nus.edu.sg).

A. Shayeghi is with the Institute for Quantum Computing, Combinatorics and Optimization Department, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: ashayeghi@uwaterloo.ca).

P. Yao is with the Joint Center for Quantum Information and Computer Science, University of Maryland, College Park, MD 20742-2420 USA (e-mail: phyao1985@gmail.com).

Communicated by S. Wolf, Associate Editor for Complexity and Cryptography.

Digital Object Identifier 10.1109/TIT.2016.2616125

$\mathcal{P}$: Given a register A, Alice gets to know the eigen-decomposition of a quantum state $\rho \in \mathcal{D}(A)$. Bob gets to know the eigen-decomposition of a quantum state $\sigma \in \mathcal{D}(A)$ such that $\mathrm{supp}(\rho) \subset \mathrm{supp}(\sigma)$. Both Alice and Bob know $S(\rho\|\sigma) \stackrel{\text{def}}{=} \mathrm{Tr}\rho\log\rho - \rho\log\sigma$, 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 $F(\cdot,\cdot)$ represents *fidelity*.

This result can be considered as a non-commutative generalization of a result due to Braverman and Rao [1] where they considered the special case when ρ and σ are classical probability distributions and the two parties only share public random coins. 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 [1], direct product theorems for all relations in the public-coin one-way and public-coin bounded-round communication models [2]–[4]. 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 *distributional entanglement assisted quantum one-way communication complexity* for all relations. This direct-sum result was shown previously by Jain *et al.* [5] and they obtained this result via a protocol that allowed them compression to *external information* carried in the message.[1] Their arguments are 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 [1] and hence it presents hope that our direct-sum result can also be

[1]Compression to external and internal information can be thought of as one-shot communication analogues of the celebrated results by Shannon [6] and Slepian and Wolf [7] exhibiting compression of source to entropy and conditional entropy respectively.

generalized to bounded-round quantum protocols. The protocol of Braverman and Rao [1] was also used by Jain [2] 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 [3], [4]. Hence protocol $\mathcal{P}$ also presents a hope of obtaining similar results for quantum communication protocols.

We also present a variant of protocol $\mathcal{P}$, with Bob possessing some side information about Alice's input. In such a case, the communication can be further reduced.

$\mathcal{P}'$: Given two registers A and B, Alice and Bob know the description of a quantum channel $\mathcal{E} : \mathcal{L}(A) \to \mathcal{L}(B)$. Alice is given the eigen-decomposition of a state $\rho \in \mathcal{D}(A)$. Bob is given the eigen-decomposition of a state $\sigma \in \mathcal{D}(A)$ (such that $\mathrm{supp}(\rho) \subset \mathrm{supp}(\sigma)$) and the state $\rho' = \mathcal{E}(\rho)$. Let $\mathrm{S}(\rho\|\sigma) - \mathrm{S}(\mathcal{E}(\rho)\|\mathcal{E}(\sigma))$ and $\varepsilon > 0$ be known to Alice and Bob. There exists a protocol, in which Alice and Bob use shared entanglement and Alice sends $\mathcal{O}((\mathrm{S}(\rho\|\sigma) - \mathrm{S}(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) + 1)/\varepsilon^4)$ 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 $\mathrm{F}(\rho, \tilde{\rho}) \geq 1 - \varepsilon$, where $\mathrm{F}(\cdot, \cdot)$ represents *fidelity* .

In the second part of our paper, we present the following protocol, which can be considered as a quantum analogue of the widely used *classical correlated sampling* protocol. For example, Holenstein [8] has used the classical correlated sampling protocol in his proof of a *parallel-repetition theorem* for two-player one-round games.

$\mathcal{P}_1$: Given a register A_1, Alice gets to know the eigen-decomposition of a quantum state $\rho \in \mathcal{D}(A_1)$. Bob gets to know the eigen-decomposition of a quantum state $\sigma \in \mathcal{D}(A_1)$. Alice and Bob use shared entanglement, do local measurements (no communication) and at the end Alice outputs registers A_1A_2 and Bob outputs registers B_1B_2 such that the following holds:

1) $B_1 \equiv A_1$ and $B_2 \equiv A_2$.
2) The marginal state in register A_1 is ρ and the marginal state in register B_1 is σ.
3) For any projective measurement $M = \{M_1, \ldots, M_w\}$ such that $M_i \in \mathcal{L}(A_1A_2)$, the following holds. Let Alice perform M on A_1A_2 and Bob perform M on B_1B_2 and obtain outcomes $I \in [w]$, $J \in [w]$ respectively. Then,

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

Recently, Dinur *et al.* [9] 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 [1], which as we mentioned, applies to the special case when inputs to Alice and Bob are classical probability distributions P, Q respectively. Let us first assume the case when Alice and Bob know $c = \mathrm{S}_\infty(P\|Q) \stackrel{\mathrm{def}}{=} \min\{\lambda |\ P \leq 2^\lambda Q\}$, the *relative max-entropy* between P and Q. In the protocol of [1], Alice and Bob share (as public coins) $\{(M_i, R_i)|\ i \in \mathbb{N}\}$, where each (M_i, R_i) is independently and identically distributed uniformly over $\mathcal{U} \times [0, 1]$, $\mathcal{U}$ being 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 of $\mathcal{O}(c)$ 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

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

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 uniform distribution in the interval $[0, 1]$ (in the limit $K \to \infty$). Again let us first assume the case when Alice and Bob know $c = \mathrm{S}_\infty(\rho\|\sigma) \stackrel{\mathrm{def}}{=} \min\{\lambda |\ \rho \leq 2^\lambda \sigma\}$ (here $\leq$ represents the Löwner order), the relative max-entropy between ρ and σ. Let eigen-decomposition of ρ be $\sum_{i=1}^N a_i |a_i\rangle\langle a_i|$ and eigen-decomposition of σ be $\sum_{i=1}^N b_i |b_i\rangle\langle b_i|$. Consider a projection P_{AA_1} as defined below and I_{AA_1} the identity operator on registers A, A_1. Alice performs a measurement $\{P_{AA_1}, I_{AA_1} - P_{AA_1}\}$, on the register AA_1 of each copy of $|\psi\rangle$ and *accepts* the index of a copy iff outcome of measurement corresponds to P_{AA_1} (which we refer to as a success for Alice).

$$P_{AA_1} = \sum_{i=1}^{N} |a_i\rangle\langle a_i|_A \otimes \left(\sum_{m=1}^{\lceil Ka_i\rceil} |m\rangle\langle m|_{A_1}\right).$$

Similarly, consider a projection P_{BB_1} as defined below (for an appropriately chosen δ) and I_{BB_1} the identity operator on register BB_1. Bob performs a measurement $\{P_{BB_1}, I_{BB_1} - P_{BB_1}\}$ on registers BB_1 on each copy of $|\psi\rangle$ and *accepts* the index of a copy iff the outcome of measurement corresponds to P_{BB_1} (which we refer to as a success for Bob).

$$P_{BB_1} = \sum_{i=1}^{N} |b_i\rangle\langle b_i|_B \otimes \left(\sum_{m=1}^{\min\{\lceil 2^c Kb_i/\delta\rceil, K\}} |m\rangle\langle m|_{B_1}\right).$$

Again it is easily argued that (in the limit $K \to \infty$) the marginal state in B (and also in A), in the first copy of $|\psi\rangle$ on 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. Hence, by *gentle measurement lemma* ([10], [11]), the marginal state in register B is not disturbed much, conditioned on success of both Alice and Bob. 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 [1]. However, since ρ and σ may not commute, our analysis deviates significantly from the analysis of [1]. We are required to show several new facts related to the non-commuting case while arguing that the protocol still works correctly.

We then consider the case in which $\mathrm{S}(\rho\|\sigma)$ (instead of $\mathrm{S}_\infty(\rho\|\sigma)$) is known to Alice and Bob. The *quantum substate theorem* [12], [13] implies that there exists a quantum state ρ', having high fidelity with ρ such that $\mathrm{S}_\infty(\rho'\|\sigma) = \mathcal{O}(\mathrm{S}(\rho\|\sigma))$. 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 $\mathcal{O}(\mathrm{S}(\rho\|\sigma))$ 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 which is a special case of the task we consider in the protocol $\mathcal{P}$.

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, such that $\mathrm{supp}(\rho) \subset \mathrm{supp}(\sigma)$. They also know $c \stackrel{\mathrm{def}}{=} \mathrm{S}(\rho\|\sigma)$, 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 $\mathrm{F}(\rho, \tilde{\rho}) \geq 1 - \varepsilon$.

Jain *et al.* [5], showed that this task (for constant ε) can be achieved with communication $\mathcal{O}(\mathrm{S}(\rho\|\sigma) + 1)$ 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 sharing public random coins and $\mathcal{O}(\mathrm{S}(P\|Q) + 1))$ bits of communication can achieve this task (for constant ε). Later an improved result was obtained by Harsha *et al.* [14], where they presented a protocol in which Bob is able to sample exactly from P with expected communication $\mathrm{S}(P\|Q) + 2\log \mathrm{S}(P\|Q) + \mathcal{O}(1)$. This led to direct sum theorems for all relations in the public-coin randomized one-way, public-coin simultaneous message passing [5] and public-coin randomized bounded-round communication models [14].

Our work strengthens their results by showing that $O(\mathrm{S}(\rho\|\sigma))$ bits of communication is enough even if σ is not known to Alice.

Very recently, Touchette [15] introduced the notion of *quantum information cost* which generalizes the internal information cost in the classical communication to the quantum setting. Moreover, he showed that in *bounded-round entanglement assisted quantum communication tasks*, the communication can be compressed to the quantum information cost based on the *state redistribution* protocol [16], [17]. Using such a compression protocol, he showed a direct sum theorem for bounded round entanglement assisted quantum communication model.

Organization

In section II, we discuss our notations and relevant notions needed for our proofs. In Section III we describe our one shot quantum protocol $\mathcal{P}$. The direct sum result follows in Section IV. In Section V we present quantum correlated sampling. We conclude in Section VI.

II. 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 $\{1, 2, \ldots, n\}$. We let log represent logarithm to the base 2 and ln represent logarithm to the base e. Let $\mathcal{X}$ and $\mathcal{Y}$ be finite sets. $\mathcal{X} \times \mathcal{Y}$ represents the cross product of $\mathcal{X}$ and $\mathcal{Y}$. For a natural number k, we let $\mathcal{X}^k$ denote the set $\mathcal{X} \times \cdots \times \mathcal{X}$, 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 $\mathbb{E}_{x \leftarrow X}[f(x)] \stackrel{\mathrm{def}}{=} \sum_{x \in \mathcal{X}} \Pr[X = x] \cdot f(x)$, where $x \leftarrow X$ means that x is drawn according to distribution X.

Consider a Hilbert space $\mathcal{H}$ endowed with an inner product $\langle \cdot, \cdot \rangle$. The ℓ_1 norm of an operator X on $\mathcal{H}$ is $\|X\|_1 \stackrel{\mathrm{def}}{=} \mathrm{Tr}\sqrt{X^\dagger X}$ and ℓ_2 norm is $\|X\|_2 \stackrel{\mathrm{def}}{=} \sqrt{\mathrm{Tr} X X^\dagger}$. A quantum state (or a density matrix or just a state) is a positive semi-definite matrix with trace equal to 1. It is called *pure* if and only if the rank is 1. A sub-normalized state is a positive semi-definite matrix with trace less than or equal to 1. 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 state and also the density matrix $|\psi\rangle\langle\psi|$, associated with $|\psi\rangle$.

Fix an orthonormal basis on $\mathcal{H}$, referred to as *computational basis*. 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 non-diagonal entries 0. Given a quantum state ρ on $\mathcal{H}$, *support of* ρ, called $\mathrm{supp}(\rho)$ is the subspace of $\mathcal{H}$ spanned by all eigen-vectors of ρ with non-zero eigenvalues.

A *quantum register* A is associated with some Hilbert space $\mathcal{H}_A$. Define $|A| \stackrel{\mathrm{def}}{=} \dim(\mathcal{H}_A)$. Let $\mathcal{L}(A)$ represent the set of all linear operators on $\mathcal{H}_A$. We denote by $\mathcal{D}(A)$, the set of quantum states on the Hilbert space $\mathcal{H}_A$. State ρ with subscript A indicates $\rho_A \in \mathcal{D}(A)$. If two registers A, B are associated with the same Hilbert space, we shall represent the relation by $A \equiv B$. Composition of two registers A and B, denoted AB, 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 (Kronecker product) of ρ and σ. The identity operator on $\mathcal{H}_A$ (and associated register A) is denoted I_A.

Let $\rho_{AB} \in \mathcal{D}(AB)$. We define

$$\rho_B \stackrel{\mathrm{def}}{=} \mathrm{Tr}_A(\rho_{AB}) \stackrel{\mathrm{def}}{=} \sum_i (\langle i| \otimes I_B)\rho_{AB}(|i\rangle \otimes I_B),$$

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 the marginal state of ρ_{AB}. Unless otherwise stated, a missing register from subscript in a state will represent partial trace over that register. Given a $\rho_A \in \mathcal{D}(A)$, a *purification* of ρ_A is a pure state $\rho_{AB} \in \mathcal{D}(AB)$ such that $\mathrm{Tr}_B(\rho_{AB}) = \rho_A$. A purification of a quantum state is not unique.

A quantum map $\mathcal{E} : \mathcal{L}(A) \to \mathcal{L}(B)$ is a completely positive and trace preserving (CPTP) linear map (mapping states in $\mathcal{D}(A)$ to states in $\mathcal{D}(B)$). A *unitary* operator $U_A : \mathcal{H}_A \to \mathcal{H}_A$ is such that $U_A^\dagger U_A = U_A U_A^\dagger = I_A$. An *isometry* $V : \mathcal{H}_A \to \mathcal{H}_B$ is such that $V^\dagger V = I_A$ and $VV^\dagger = I_B$. The set of all unitary operations on register A is denoted by $\mathcal{U}(A)$.

Definition 1: We shall consider the following information theoretic quantities. Let A be a quantum register. Let $\varepsilon \geq 0$.

1) *Fidelity:* For $\rho, \sigma \in \mathcal{D}(A)$,

$$\mathrm{F}(\rho, \sigma) \stackrel{\text{def}}{=} \left\| \sqrt{\rho}\sqrt{\sigma} \right\|_1 .$$

For classical probability distributions $P = \{p_i\}$, $Q = \{q_i\}$,

$$\mathrm{F}(P, Q) \stackrel{\text{def}}{=} \sum_i \sqrt{p_i \cdot q_i}.$$

2) *Entropy:* For $\rho \in \mathcal{D}(A)$,

$$S(\rho) \stackrel{\text{def}}{=} -\mathrm{Tr}(\rho \log \rho).$$

3) *Relative entropy:* For $\rho, \sigma \in \mathcal{D}(A)$ such that $\mathrm{supp}(\rho) \subset \mathrm{supp}(\sigma)$,

$$\mathrm{S}(\rho \| \sigma) \stackrel{\text{def}}{=} \mathrm{Tr}(\rho \log \rho) - \mathrm{Tr}(\rho \log \sigma).$$

4) *Relative max-entropy:* For $\rho, \sigma \in \mathcal{D}(A)$ such that $\mathrm{supp}(\rho) \subset \mathrm{supp}(\sigma)$,

$$\mathrm{S}_\infty(\rho \| \sigma) \stackrel{\text{def}}{=} \inf\{\lambda \in \mathbb{R} : 2^\lambda \sigma \geq \rho\}.$$

5) *Mutual information:* For $\rho_{AB} \in \mathcal{D}(AB)$,

$$\begin{aligned}\mathrm{I}(A:B)_\rho &\stackrel{\text{def}}{=} S(\rho_A) + S(\rho_B) - S(\rho_{AB}) \\ &= \mathrm{S}(\rho_{AB} \| \rho_A \otimes \rho_B) .\end{aligned}$$

6) *Conditional mutual information:* For $\rho_{ABC} \in \mathcal{D}(ABC)$,

$$\mathrm{I}(A:B|C)_\rho \stackrel{\text{def}}{=} \mathrm{I}(A:BC)_\rho - \mathrm{I}(A:C)_\rho .$$

We will use the following facts.

Fact 2 ([19, p. 416]): For quantum states $\rho, \sigma \in \mathcal{D}(A)$, it holds that

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

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

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

Fact 3 [19]: (**Stinespring Representation**) Let $\mathcal{E}(\cdot) : \mathcal{L}(A) \to \mathcal{L}(B)$ be a quantum operation. There exists a Hilbert space C and an unitary $U : A \otimes B \otimes C \to A \otimes B \otimes C$ such that $\mathcal{E}(\omega) = \mathrm{Tr}_{A,C}\left(U(\omega \otimes |0\rangle\langle 0|^{B,C})U^\dagger\right)$. Stinespring representation for a channel is not unique.

Fact 4 [20], [21]: For states $\rho, \sigma \in \mathcal{D}(A)$, and quantum operation $\mathcal{E}(\cdot) : \mathcal{L}(A) \to \mathcal{L}(B)$, it holds that

$$\begin{aligned}\|\mathcal{E}(\rho) - \mathcal{E}(\sigma)\|_1 &\leq \|\rho - \sigma\|_1 , \\ \mathrm{F}(\mathcal{E}(\rho), \mathcal{E}(\sigma)) &\geq \mathrm{F}(\rho, \sigma), \\ \mathrm{S}(\rho\|\sigma) &\geq \mathrm{S}(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) .\end{aligned}$$

In particular, for bipartite states $\rho^{AB}, \sigma^{AB} \in \mathcal{D}(AB)$, it holds that

$$\begin{aligned}&\left\|\rho^{AB} - \sigma^{AB}\right\|_1 \geq \left\|\rho^A - \sigma^A\right\|_1 \\ &\mathrm{F}(\rho^{AB}, \sigma^{AB}) \leq \mathrm{F}(\rho^A, \sigma^A), \\ &\mathrm{S}(\rho_{AB}\|\sigma_{AB}) \geq \mathrm{S}(\rho_A\|\sigma_A) .\end{aligned}$$

Fact 5 ([23, Lemma 4.41.]): Let A, B be two positive semidefinite operators on Hilbert space $\mathcal{H}$. Then

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

Fact 6: Given two quantum states ρ and σ,

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

Proof: By Facts 5 and 2,

$$\begin{aligned}2\sqrt{1 - \mathrm{F}(\rho, \sigma)^2} &\geq \|\rho - \sigma\|_1 \geq \left\|\sqrt{\rho} - \sqrt{\sigma}\right\|_2^2 \\ &= 2 - 2 \cdot \mathrm{Tr}\left(\sqrt{\rho}\sqrt{\sigma}\right) .\end{aligned}$$

□

Fact 7 (Joint concavity of fidelity [23, Proposition 4.7]): Given states $\rho_1, \rho_2 \ldots \rho_k, \sigma_1, \sigma_2 \ldots \sigma_k$ and positive numbers $p_1, p_2 \ldots p_k$ such that $\sum_i p_i = 1$. Then

$$\mathrm{F}(\sum_i p_i\rho_i, \sum_i p_i\sigma_i) \geq \sum_i p_i \mathrm{F}(\rho_i, \sigma_i).$$

Fact 8 ([13], [23]): (**Quantum Substate Theorem**) Given $\rho, \sigma \in \mathcal{D}(A)$, such that $\mathrm{supp}(\rho) \subset \mathrm{supp}(\sigma)$. For any $\varepsilon > 0$, there exists $\rho' \in \mathcal{D}(A)$ such that

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

Fact 9 [10], [11]: (**Gentle Measurement Lemma**) Let $\rho \in \mathcal{D}(A)$ and Π be a projector. Then,

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

Proof: Introduce a register B, such that $|B| \geq |A|$. Let $\phi \in \mathcal{D}(AB)$ be a purification of ρ. Then $(\Pi \otimes I_B)\phi(\Pi \otimes I_B)$ is a purification of $\Pi\rho\Pi$. Hence (using monotonicity of fidelity under quantum operation, Fact 4)

$$\begin{aligned}&\mathrm{F}(\rho, \frac{\Pi\rho\Pi}{\mathrm{Tr}\Pi\rho})\mathrm{F}\left(\phi, (\Pi \otimes I_B)\phi(\Pi \otimes I_B)\right) \\ &= \frac{|\langle\phi|\,(\Pi \otimes I)\,|\phi\rangle|}{\|(\Pi \otimes I)\,|\phi\rangle\|} = \sqrt{\mathrm{Tr}(\Pi\rho)}.\end{aligned}$$

□

Fact 10: Given quantum states $\sigma_{AB} \in \mathcal{D}(AB)$, $\rho_A \in \mathcal{D}(A)$, such that $\mathrm{supp}(\rho_A) \subset \mathrm{supp}(\sigma_A)$, it holds that

$$\mathrm{Tr}(e^{\log(\sigma_{AB}) - \log(\sigma_A \otimes I_B) + \log(\rho_A \otimes I_B)}) \leq 1.$$

Proof: Consider,

$$
\begin{aligned}
&\mathrm{Tr}(e^{\log(\sigma_{AB})-\log(\sigma_A\otimes I_B)+\log(\rho_A\otimes I_B)}) \\
&< \int_0^\infty du \mathrm{Tr}(\sigma_{AB}\frac{1}{\sigma_A+uI_A}\rho_A\frac{1}{\sigma_A+uI_A}) \\
&\quad \text{(Theorem5, [25])} \\
&= \int_0^\infty du \mathrm{Tr}(\frac{1}{\sigma_A+uI_A}\sigma_A\frac{1}{\sigma_A+uI_A}\rho_A) \\
&= \mathrm{Tr}(\sigma_A\int_0^\infty du\frac{1}{(\sigma_A+uI_A)^2}\rho_A) \\
&= \mathrm{Tr}(\sigma_A\sigma_A^{-1}\rho_A) = 1.
\end{aligned}
$$

□

Fact 11: [25], [26](**Strong Subadditivity Theorem**) For any tripartite quantum state $\rho \in \mathcal{D}(ABC)$, it holds that $\mathrm{I}(A:C|B)_\rho \geq 0$.

Fact 12 [19, p. 515], [27]: For a quantum state $\rho_{AB} \in \mathcal{D}(AB)$, it holds that $|\mathrm{S}(\rho_A)-\mathrm{S}(\rho_B)| \leq \mathrm{S}(\rho_{AB}) \leq \mathrm{S}(\rho_A)+\mathrm{S}(\rho_B)$. Furthermore,

$$\mathrm{I}(A:B)_\rho = \mathrm{S}(\rho_A)+\mathrm{S}(\rho_B)-\mathrm{S}(\rho_{AB}) \leq 2\mathrm{S}(\rho_A).$$

Fact 13: Let $\rho_{A_1A_2\ldots A_kBC} \in \mathcal{D}(A_1\cdots A_kBC)$ such that $\rho_{A_1A_2\ldots A_k} = \rho_{A_1}\otimes\rho_{A_2}\otimes\ldots\rho_{A_k}$. Then,

$$\mathrm{I}(A_1A_2\ldots A_k:B|C)_\rho \geq \sum_{i=1}^k \mathrm{I}(A_i:B|C)_\rho.$$

Proof: Consider,

$$
\begin{aligned}
&\mathrm{I}(A_1A_2\ldots A_k:B|C)_\rho \\
&= \mathrm{I}(A_1:B|C)_\rho+\mathrm{I}(A_2A_3\ldots A_k:B|A_1C)_\rho \\
&= \mathrm{I}(A_1:B|C)_\rho+\mathrm{I}(A_2A_3\ldots A_k:A_1BC)_\rho \\
&\quad -\mathrm{I}(A_1:A_2A_3\ldots A_k)_\rho \\
&= \mathrm{I}(A_1:B|C)_\rho+\mathrm{I}(A_2A_3\ldots A_k:A_1BC)_\rho \\
&\geq \mathrm{I}(A_1:B|C)_\rho+\mathrm{I}(A_2A_3\ldots A_k:B|C)_\rho
\end{aligned}
$$

The first and second equalities follow from the definition of the conditional mutual information. The third equality is from the independence between A_1 and $A_2A_3\ldots A_k$. The last inequality is from strong subadditivity (Fact 11). Proof follows by induction. □

For the facts appearing below, the proofs can be obtained by direct calculations and hence have been skipped.

Fact 14: Given $\rho_{AB},\sigma_{AB} \in \mathcal{D}(AB)$, such that $\mathrm{supp}(\sigma_{AB}) \subset \mathrm{supp}(\rho_{AB})$, $\rho_{AB} = \sum_a \mu(a)|a\rangle\langle a|_A \otimes \rho_B^a$ and $\sigma_{AB} = \sum_a \mu'(a)|a\rangle\langle a|_A\otimes\sigma_B^a$, where $\rho_B^a,\sigma_B^a \in \mathcal{D}(B)$, $\mu(a),\mu'(a)\geq 0$ and $\sum_a\mu(a)=1, \sum_a\mu'(a)=1$. It holds from the definition of relative entropy that

$$\mathrm{S}(\sigma_{AB}\|\rho_{AB}) = \mathrm{S}(\mu\|\mu') + \mathop{\mathbb{E}}_{a\leftarrow\mu'}\left[\mathrm{S}(\sigma_B^a\|\rho_B^a)\right].$$

Fact 15: Given a classical-quantum state $\rho_{AB}\in\mathcal{D}(AB)$ of the form $\rho_{AB} = \sum_a\mu(a)|a\rangle\langle a|_A\otimes\rho_B^a$, where $\rho_B^a\in\mathcal{D}(B)$ and $\sum_a\mu(a)=1$, $\mu(a)\geq 0$, we have

$$\mathrm{I}(A:B)_\rho = \mathrm{S}\left(\sum_a\mu(a)\rho_a\right)-\sum_a\mu(a)\mathrm{S}(\rho_a),$$

Fact 16: Let ρ_{ABC} be a state of the form $\rho_{ABC} = \sum_c\mu(c)|c\rangle\langle c|_C \otimes \rho_{AB}^c$, where $\rho_{AB}^c \in \mathcal{D}(AB)$ and $\sum_c\mu(c)=1$, $\mu(c)\geq 0$. Then

$$\mathrm{I}(A:B|C)_\rho = \sum_c\mu(c)\mathrm{I}(A:B)_{\rho^c}.$$

Communication Complexity

In this section we briefly describe entanglement assisted quantum one-way communication complexity. A mathematically detailed definition has been given by Touchette in [28]. Let $f \subseteq \mathcal{X}\times\mathcal{Y}\times\mathcal{Z}$ be a relation. Alice holds input $x\in\mathcal{X}$ and Bob holds input $y\in\mathcal{Y}$. They may share prior quantum states independent of the inputs. Alice makes a unitary transformation on her qubits, based on her input x, and sends part of her qubits to Bob. Bob makes a unitary operation, based on his input y, 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 $\mathrm{Q}_\varepsilon^{\mathrm{ent,A\to B}}(f)$ represent the quantum one-way communication complexity of f with worst case error ε, that is minimum number of qubits Alice needs to send to Bob, over all protocols computing f with error at most ε on *any* input (x,y).

We let $\mathrm{Q}_\varepsilon^{\mathrm{ent,A\to B},\mu}(f)$ represent distributional quantum one-way communication complexity of f under distribution μ over $\mathcal{X}\times\mathcal{Y}$ with distributional error at most ε. This is the communication cost of the best protocol computing f with maximum error ε averaged over distribution μ. Following is Yao's min-max theorem connecting the worst case error and the distributional error settings.

Fact 17: [29] $\mathrm{Q}_\varepsilon^{\mathrm{ent,A\to B}}(f) = \max_\mu \mathrm{Q}_\varepsilon^{\mathrm{ent,A\to B},\mu}(f)$.

III. A Quantum Compression Protocol

Following is our main result in this section.

Theorem 18: Given quantum states ρ,σ on a Hilbert space $\mathcal{H}$ with dimension N, such that $supp(\rho)\subset supp(\sigma)$. Alice is given the eigen-decomposition of ρ and Bob is given the eigen-decomposition of σ. Let $\mathrm{S}(\rho\|\sigma)$ and $\varepsilon>0$ be known to Alice and Bob. There exists an entanglement assisted quantum one-way communication protocol, with Alice sending $\mathcal{O}(\mathrm{S}(\rho\|\sigma)+1)/\varepsilon^4)$ bits of communication to Bob, such that the state $\tilde{\rho}$ that Bob outputs at the end of the protocol satisfies $\mathrm{F}(\rho,\tilde{\rho})\geq 1-5\varepsilon$.

Proof: Let the eigen-decomposition of ρ be $\sum_{i=1}^N a_i|a_i\rangle\langle a_i|$ and that of σ be $\sum_{i=1}^N b_i|b_i\rangle\langle b_i|$. Define $c \stackrel{\mathrm{def}}{=} \mathrm{S}(\rho\|\sigma)$, $\delta\stackrel{\mathrm{def}}{=}(\varepsilon/3)^4$ and $c'\stackrel{\mathrm{def}}{=}(c+2)/\delta$. Without loss of generality, assume $a_1,a_2\ldots a_N, \frac{2^{c'}}{\delta}b_1,\frac{2^{c'}}{\delta}b_2\ldots\frac{2^{c'}}{\delta}b_N$ to be rational numbers, and define K be the least common multiple of their denominators. The error due to this assumption can be made arbitrarily close to 0, for large enough K.

Let $\{|1\rangle,|2\rangle\ldots|N\rangle\}$ be an orthonormal basis for $\mathcal{H}$. Introduce registers A_1,B_1 associated to $\mathcal{H}$ and registers A_2,B_2 associated to some Hilbert space $\mathcal{H}'$ with an orthonormal basis $\{|1\rangle,|2\rangle\ldots|K\rangle\}$.

Consider the following state on A_1, A_2, B_1, B_2.

$$|S\rangle_{A_1A_2B_1B_2} \stackrel{\text{def}}{=} \frac{1}{\sqrt{KN}}\sum_{i=1}^{N}|i,i\rangle_{A_1B_1}\otimes\left(\sum_{m=1}^{K}|m,m\rangle_{A_2B_2}\right) \quad (1)$$

For brevity, define registers A, B such that $A \stackrel{\text{def}}{=} A_1A_2$ and $B \stackrel{\text{def}}{=} B_1B_2$.

The protocol is described below.

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

Shared resources: Alice and Bob hold $\lceil N\log(\frac{1}{\delta})\rceil$ registers $A_1^iA_2^iB_1^iB_2^i$ ($i \in [\lceil N\log(\frac{1}{\delta})\rceil]$), such that $A_1^i \equiv A_1, A_2^i \equiv A_2, B_1^i \equiv B_1, B_2^i \equiv B_2$. The shared state in register $A_1^iA_2^iB_1^iB_2^i$ is $|S\rangle_{A_1^iA_2^iB_1^iB_2^i}$. Let i refer to the 'index' of corresponding registers.

They also share infinitely many random hash functions $h_1, h_2, \cdots$, where each $h_l : \{0,\cdots,N-1\} \to \{0,1\}$.

1) **For** $i = 1$ to $\lceil N\log(\frac{1}{\delta})\rceil$,
   a) Alice performs the measurement $\{P_A, I_A - P_A\}$ on each register $A_1^iA_2^i$ where,
   $$P_A \stackrel{\text{def}}{=} \sum_i |a_i\rangle\langle a_i|_{A_1}\otimes\left(\sum_{m=1}^{Ka_i}|m\rangle\langle m|_{A_2}\right) (2)$$
   On each index i, she declares *success* if her outcome corresponds to P_A.
   b) Bob performs the measurement $\{P_B, I_B - P_B\}$ on each register $B_1^iB_2^i$ where,
   $$P_B \stackrel{\text{def}}{=} \sum_i |b_i\rangle\langle b_i|_{B_1} \otimes\left(\sum_{m=1}^{\min\{\frac{K}{\delta}2^{c'}b_i,K\}}|m\rangle\langle m|_{B_2}\right) \quad (3)$$
   On each index i, he declares *success* if his outcome corresponds to P_B.

   Endfor
2) If Alice does not succeed on any index, she aborts.
3) Else, Alice selects the first index m where she succeeds and sends to Bob the binary encoding of $k = \lceil m/N\rceil$ using $\lceil\log\log\frac{1}{\delta}\rceil$ bits.
4) Alice sends $\{h_l(m \mod N) |\ l \in [\lceil c' + \log(\frac{1}{\delta}) + 2\log\frac{1}{\varepsilon}\rceil]\}$ to Bob.
5) Define $S_B \stackrel{\text{def}}{=} \{t|\ \text{Bob succeeds on index t}\} \cap \{(\text{k-1})N,\cdots,kN-1\}$. If S_B is empty, he outputs $|0\rangle\langle 0|$. Bob selects the first index n in S_B such that $\forall l \in [\lceil c' + \log(\frac{1}{\delta}) + 2\log\frac{1}{\varepsilon}\rceil] : h_l(n \mod N) = h_l(m \mod N)$ and outputs the state in B_1^n (if no such index exists, he outputs $|0\rangle\langle 0|$).

We analyze the protocol through a series of claims. Following claim computes the probability of success for Alice and Bob.

Claim 19: For each index i, $\Pr[\text{Alice succeeds}] = \frac{1}{N}$; $\Pr[\text{Bob succeeds}] \leq \frac{2^{c'}}{\delta N}$

Proof: Follows from direct calculation. □

From quantum substate theorem (Fact 8), there exists a state ρ' which satisfies $\mathrm{F}(\rho,\rho') \geq 1-\delta$ and

$$\mathrm{S}_\infty(\rho'\|\sigma) \leq \frac{\mathrm{S}(\rho\|\sigma)+1}{\delta} + \log\frac{1}{1-\delta} \leq \frac{\mathrm{S}(\rho\|\sigma)+2}{\delta} = c'.$$

We prove the following claim which is of independent interest as well.

Claim 20: Let ρ' have the eigen-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 $\sum_{j|\ b_j \leq p\cdot g_i} \left|\langle b_j|g_i\rangle\right|^2 \leq 2^{c'}\cdot p$.

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 eigen-space 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 $g_i\sum_{j:\ b_j\leq p\cdot g_i}\left|\langle b_j|g_i\rangle\right|^2 \leq 2^{c'}\cdot p\cdot g_i$. This implies the lemma. □

Define

$$|S_A(\rho)\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{K}}\sum_{i=1}^{N}|a_i\rangle\,|\overline{a_i}\rangle\otimes\left(\sum_{m=1}^{Ka_i}|m,m\rangle\right);$$
$$\left|S_A(\rho')\right\rangle \stackrel{\text{def}}{=} \frac{1}{\sqrt{K}}\sum_{i=1}^{N}|g_i\rangle\,|\overline{g_i}\rangle\otimes\left(\sum_{m=1}^{\lceil Kg_i\rceil}|m,m\rangle\right).$$

Here $|\overline{a_i}\rangle$ (similarly $|\overline{g_i}\rangle$) is the state obtained by taking complex conjugate of $|a_i\rangle$ ($|g_i\rangle$), with respect to the basis $\{|1\rangle, |2\rangle \ldots |N\rangle\}$ in $\mathcal{H}$.

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

Claim 21: $\left|\langle S_A(\rho)|S_A(\rho')\rangle\right| \geq 1 - 2(1-\mathrm{F}(\rho,\rho'))^{1/4}$.

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

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

where the first equality is from the definitions of $|S_A(\rho)\rangle$ and $\left|S_A(\rho')\right\rangle$; the second equality is from the definition of ℓ_1 distance; the first inequality is from 2; the second inequality is from the fact that $\mathrm{Tr}\sqrt{\rho}\sqrt{\rho'} \leq 1$; the third inequality is from Facts 6. □

We use these claims to prove the following.

Claim 22: For each index i, $\Pr[\text{Bob succeeds}|\text{ Alice succeeds}] \geq 1 - \delta - 2\delta^{1/4} \geq 1 - \varepsilon$.

Proof: Consider,

$$
\begin{aligned}
&(I_A \otimes P_B) \left|S_A(\rho')\right\rangle \\
&= \frac{1}{\sqrt{K}} \sum_{i,j=1}^{N} \left|\overline{g_j}\right\rangle |b_i\rangle \langle b_i|g_j\rangle \left(\sum_{m=1}^{\min\{\lceil K g_j \rceil, \frac{K}{\delta} 2^{c'} b_i\}} |m, m\rangle \right).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\left\| (I_A \otimes P_B) \left|S_A(\rho')\right\rangle \right\|^2 &\geq \sum_{i,j=1}^{N} \left|\langle b_i|g_j\rangle\right|^2 \min\{g_j, \frac{1}{\delta} 2^{c'} b_i\} \\
&\geq \sum_{j=1}^{N} g_j \left(\sum_{i|\ b_i \geq \delta 2^{-c'} g_j} \left|\langle b_i|g_j\rangle\right|^2 \right) \\
&\geq \sum_{j=1}^{N} g_j (1 - \delta) \\
&= 1 - \delta. \quad \text{(using Claim 20)} \qquad (4)
\end{aligned}
$$

Using the above,

$$
\begin{aligned}
&\Pr[\text{Bob succeeds}|\text{ Alice succeeds}] \\
&= \mathrm{Tr}(I_A \otimes P_B) \left|S_A(\rho)\right\rangle\langle S_A(\rho)| \\
&\geq \mathrm{Tr}(I_A \otimes P_B) \left|S_A(\rho')\right\rangle\!\left\langle S_A(\rho')\right| \\
&\quad - \frac{1}{2} \left\| S_A(\rho) - S_A(\rho') \right\|_1 \\
&= \mathrm{Tr}(I_A \otimes P_B) \left|S_A(\rho')\right\rangle\!\left\langle S_A(\rho')\right| \\
&\quad - \sqrt{1 - |\langle S_A(\rho)|S_A(\rho')\rangle|^2} \quad \text{(Fact 2)} \\
&\geq 1 - \delta - 2\sqrt{(1 - \mathrm{F}(\rho, \rho'))^{1/2}} \\
&\quad \text{(Claim 21 and Eq. (4))}
\end{aligned}
$$

□

Finally, we show that if Alice and Bob succeed together on an index, the state in register B with Bob is close to ρ.

Claim 23: 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: From gentle measurement lemma (Fact 9),

$$
\begin{aligned}
&\mathrm{F}(S_A(\rho), \frac{(I_A \otimes P_B) \left|S_A(\rho)\right\rangle\langle S_A(\rho)| (I_A \otimes P_B)}{\mathrm{Tr}(I_A \otimes P_B) \left|S_A(\rho)\right\rangle\langle S_A(\rho)|}) \\
&\quad \geq \sqrt{\mathrm{Tr}(I_A \otimes P_B) \left|S_A(\rho)\right\rangle\langle S_A(\rho)|}.
\end{aligned}
$$

Since the marginal of $|S_A(\rho)\rangle$ on register B is ρ and partial trace does not decrease fidelity (Fact 4), using item 2. above, the desired result follows.

□

Let j be the first index where Alice and Bob both succeed. As described in the protocol, m is the first index where Alice succeeds and n is the index such that Bob outputs the state in B_1^n. We have the following claim,

Claim 24: With probability at least $1 - 4\varepsilon$, $m = n = j$.

Before proving Claim 24, let us define the following "bad" events.

Definition 25:
- T_1 is the event that Alice does not succeed on any of the indices.
- T_2 is the event that $m \notin S_B$ conditioned on $\neg T_1$.
- T_3 represents the event that $n \neq m$ conditioned on $\neg T_1$.

Notice that if none of above events occur, then both Alice and Bob output the same index $n = m$, and since m is the first index at which Alice succeeds, $n = m = j$.

We have the following claim.

Claim 26: It holds that: 1. $\Pr[T_1] \leq \varepsilon$; 2. $\Pr[T_2] \leq \varepsilon$; 3. $\Pr[T_3] \leq 3\varepsilon$.

Proof:

1) $\Pr[T_1] \leq \left(1 - \frac{1}{N}\right)^{\lceil N \cdot \log \frac{1}{\varepsilon} \rceil} \leq \exp^{-\lceil \log \frac{1}{\varepsilon} \rceil} \leq \varepsilon$.
2) Follows from Claim 22.
3) For this argument we condition on $\neg T_1$ for all events below. From Claim 19 and the fact that Bob independently measures each index, we have $\mathbb{E}[|S_B|] = N \cdot \Pr[\text{Bob succeeds}] \leq \frac{2^{c'}}{\delta}$. Using Markov's inequality,

$$
\Pr\left[|S_B| \geq \frac{2^{c'}}{\delta\varepsilon} \right] \leq \frac{\delta\varepsilon}{2^{c'}} \cdot \mathbb{E}[|S_B|] \leq \varepsilon. \qquad (5)
$$

Thus

$$
\begin{aligned}
\Pr[T_3] &\leq \Pr\left[|S_B| \geq \frac{2^{c'}}{\delta\varepsilon} \text{ or } m \notin S_B \right] \\
&\quad + \Pr\left[T_3 \mid m \in S_B \text{ and } |S_B| \leq \frac{2^{c'}}{\delta\varepsilon} \right] \\
&\leq \Pr\left[|S_B| \geq \frac{2^{c'}}{\delta\varepsilon} \right] + \Pr[T_2] \\
&\quad + \Pr\left[T_3 \mid m \in S_B \text{ and } |S_B| \leq \frac{2^{c'}}{\delta\varepsilon} \right] \\
&\leq 2\varepsilon + \Pr\left[T_3 \mid m \in S_B \text{ and } |S_B| \leq \frac{2^{c'}}{\delta\varepsilon} \right] \\
&\quad \text{(Eq. (5) and item 2. of this claim)} \\
&\leq 2\varepsilon + 2^{-\lceil c' + \log \frac{1}{\delta} + 2 \log \frac{1}{\varepsilon} \rceil} \cdot \frac{2^{c'}}{\delta\varepsilon} \quad \leq \quad 3\varepsilon.
\end{aligned}
$$

□

We bound the probability that $m \neq n$. If $m = n$, then m being the first index on which Alice succeeds, we have $m = n = j$.

Proof of Claim24: We conclude the claim since,

$$
\Pr[n \neq m] \leq \Pr[T_1] + \Pr[\neg T_1] \cdot \Pr[T_3] \leq 4\varepsilon.
$$

□

From claims 19,22 and 24, the probability that Bob learns the index j is atleast $1 - 4\varepsilon$. Conditioned on this event, Claim 23, implies that the state $\rho' \in \mathcal{D}(B^j)$ that Bob outputs satisfies $\mathrm{F}(\rho', \rho) \geq 1 - \varepsilon$. Conditioned on the event that Bob learns the wrong index or the protocol is aborted, let the state output by Bob be μ. Then Bob outputs the state $\tilde{\rho} = \alpha\rho' + (1 - \alpha)\mu$, where $\alpha \geq 1 - 4\varepsilon$. Using concavity of fidelity

(Fact 7), we have $\mathrm{F}(\tilde{\rho}, \rho) \geq \alpha \mathrm{F}(\rho', \rho) + (1-\alpha)\mathrm{F}(\mu, \rho) \geq (1-4\varepsilon)(1-\varepsilon) \geq 1-5\varepsilon$.

The communication cost of above protocol is

$$\lceil \log\log\frac{1}{\delta}\rceil + \lceil c' + \log\frac{1}{\delta} + 2\log\frac{1}{\varepsilon}\rceil \leq \lceil 3^4\frac{c+2}{\varepsilon^4} + 7\log\frac{1}{\varepsilon}\rceil.$$

This completes the proof of theorem. □

It may be noted that variants of the part of protocol that uses hash functions, have appeared in many other works such as [1] and [30].

Remark 27: Note that if Alice and Bob get a real number $r > \mathrm{S}(\rho\|\sigma)$, instead of $\mathrm{S}(\rho\|\sigma)$ (all other inputs remaining the same), the protocol above works in the same fashion, with the communication upper bounded by $O((r+1)/\varepsilon^4)$.

A. Compression With Side Information

Here we present a variant of our protocol with side information. We start with the following.

Lemma 28: Let A, B be two registers. Alice is given the eigen-decomposition of a bipartite state $\rho_{AB} \in \mathcal{D}(AB)$. Bob is given the eigen-decompositions of a bipartite state $\sigma_{AB} \in \mathcal{D}(AB)$ and the state $\rho_A \stackrel{\mathrm{def}}{=} \mathrm{Tr}_B(\rho_{AB})$, such that supp$(\rho_{AB}) \subset$ supp(σ_{AB}). Define $\sigma_A \stackrel{\mathrm{def}}{=} \mathrm{Tr}_B(\sigma_{AB})$. Let $\mathrm{S}(\rho_{AB}\|\sigma_{AB}) - \mathrm{S}(\rho_A\|\sigma_A)$ and $\varepsilon > 0$ be known to Alice and Bob. There exists a protocol, in which Alice and Bob use shared entanglement and Alice sends $\mathcal{O}((\mathrm{S}(\rho_{AB}\|\sigma_{AB}) - \mathrm{S}(\rho_A\|\sigma_A) + 1)/\varepsilon^4)$ bits of communication to Bob such that the state $\tilde{\rho}_{AB}$ that Bob outputs at the end of the protocol satisfies $\mathrm{F}(\rho_{AB}, \tilde{\rho}_{AB}) \geq 1-5\varepsilon$.

Proof: Following equality follows from definitions.

$$\begin{aligned}&\mathrm{S}(\rho_{AB}\|\sigma_{AB}) - \mathrm{S}(\rho_A\|\sigma_A)\\ &= \mathrm{S}\left(\rho_{AB}\middle\| e^{\log(\sigma_{AB})-\log(\sigma_A\otimes I_B)+\log(\rho_A\otimes I_B)}\right).\end{aligned}$$

Define,

$$\begin{aligned}Z &= \mathrm{Tr}(e^{\log(\sigma_{AB})-\log(\sigma_A\otimes I_B)+\log(\rho_A\otimes I_B)});\\ \tau_{AB} &= e^{\log(\sigma_{AB})-\log(\sigma_A\otimes I_B)+\log(\rho_A\otimes I_B)}/Z.\end{aligned}$$

It holds that $Z \leq 1$ (from Fact 10) and hence $\mathrm{S}(\rho_{AB}\|\tau_{AB}) \leq \mathrm{S}(\rho_{AB}\|\sigma_{AB}) - \mathrm{S}(\rho_A\|\sigma_A)$. Bob computes the eigen-decomposition of τ_{AB} using his input. They run the protocol given by Theorem 18 with the following setting: Alice knows a state ρ_{AB}, Bob knows a state τ_{AB} and both know a number $(= \mathrm{S}(\rho_{AB}\|\sigma_{AB}) - \mathrm{S}(\rho_A\|\sigma_A))$ greater than $\mathrm{S}(\rho_{AB}\|\tau_{AB})$. They also know the error parameter $\varepsilon > 0$. By the virtue of Remark 27, at the end of the protocol, Bob obtains a state $\tilde{\rho}_{AB}$, such that $\mathrm{F}(\rho_{AB}, \tilde{\rho}_{AB}) \geq 1-5\varepsilon$. Communication from Alice is upper bounded by $\mathcal{O}((\mathrm{S}(\rho_{AB}\|\sigma_{AB}) - \mathrm{S}(\rho_A\|\sigma_A) + 1)/\epsilon^4)$. □

We now present the protocol $\mathcal{P}'$ as mentioned in the Introduction.

Theorem 29: Let A, B be two registers associated to Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B$ respectively. Alice and Bob know a Stinespring representation (Fact 3) of a quantum channel $\mathcal{E} : \mathcal{L}(A) \to \mathcal{L}(B)$. Alice is given the eigen-decomposition of a state $\rho \in \mathcal{D}(A)$. Bob is given the eigen-decompositions of a state $\sigma \in \mathcal{D}(A)$ (such that supp$(\rho) \subset$ supp(σ)) and the state $\rho' = \mathcal{E}(\rho)$. Let $\mathrm{S}(\rho\|\sigma) - \mathrm{S}(\mathcal{E}(\rho)\|\mathcal{E}(\sigma))$ and $\varepsilon > 0$ be known to Alice and Bob. There exists a protocol, in which Alice and Bob use shared entanglement and Alice sends $\mathcal{O}((\mathrm{S}(\rho\|\sigma) - \mathrm{S}(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) + 1)/\varepsilon^4)$ bits of communication to Bob, such that the state $\tilde{\rho}$ that Bob outputs at the end of the protocol satisfies $\mathrm{F}(\rho, \tilde{\rho}) \geq 1-5\varepsilon$.

Proof: Let a Stinespring representation of $\mathcal{E}$ be $\mathcal{E}(\omega) = \mathrm{Tr}_{A,C}\left(V(\omega\,|0\rangle\langle 0|_{BC})V^\dagger\right)$, where $V : \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C \to \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ is a unitary operation (Fact 3). Alice and Bob compute the states $V(\rho \otimes |0\rangle\langle 0|_{BC})V^\dagger$ and $V(\sigma \otimes |0\rangle\langle 0|_{BC})V^\dagger$, respectively. From Lemma 28 and the equality $\mathrm{S}\left(V(\rho \otimes |0\rangle\langle 0|_{BC})V^\dagger \middle\| V(\sigma \otimes |0\rangle\langle 0|_{BC})V^\dagger\right) = \mathrm{S}(\rho\|\sigma)$, there exists a protocol, in which Alice and Bob use shared entanglement and Alice sends $\mathcal{O}(\mathrm{S}(\rho\|\sigma) - \mathrm{S}(\mathcal{E}(\rho)\|\mathcal{E}(\sigma)) + 1)/\varepsilon^4$ bits of communication to Bob, such that the state $\tilde{\rho}_{ABC}$ that Bob gets at the end of the protocol satisfies $\mathrm{F}(V(\rho \otimes |0\rangle\langle 0|_{BC})V^\dagger, \tilde{\rho}_{ABC}) \geq 1-5\varepsilon$. Bob outputs $\tilde{\rho} = \mathrm{Tr}_{BC}V^\dagger(\tilde{\rho}_{ABC})V$. From monotonicity of fidelity under quantum operation (Fact 4), $\mathrm{F}(\rho, \tilde{\rho}) \geq 1-5\varepsilon$. □

IV. A Direct Sum Theorem for Quantum One-Way Communication Complexity

As a consequence of Theorem 18 we obtain the following direct sum result for all relations in the model of entanglement-assisted one-way communication complexity.

Theorem 30: Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be finite sets, $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation, $0 < \varepsilon, \delta$ be error parameters and $k > 1$ be an integer. We have

$$\mathrm{Q}_{\varepsilon}^{\mathrm{ent,A\to B}}\left(f^k\right) \geq \Omega\left(k\left(\delta^9 \cdot \mathrm{Q}_{\varepsilon+\delta}^{\mathrm{ent,A\to B}}(f) - 1\right)\right).$$

Proof: Let μ be any distribution over $\mathcal{X}\times\mathcal{Y}$. We show the following, which combined with Fact 17 implies the desired:

$$\mathrm{Q}_{\varepsilon}^{\mathrm{ent,A\to B},\mu^k}\left(f^k\right) \geq \Omega\left(k\left(\delta^9 \cdot \mathrm{Q}_{\varepsilon+\delta}^{\mathrm{ent,A\to B},\mu}(f) - 1\right)\right).$$

Let $\mathcal{Q}$ 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 inputs to Alice and Bob be given in registers $X_1, X_2 \ldots X_k$ and $Y_1, Y_2 \ldots Y_k$. For brevity, we define $X \stackrel{\mathrm{def}}{=} X_1, X_2 \ldots X_k$ and $Y \stackrel{\mathrm{def}}{=} Y_1, Y_2 \ldots Y_k$. Thus, the state $\sum_{xy}\mu^k(x,y)\,|xy\rangle\langle xy|_{XY}$ represents the joint input, where x is drawn from X and y is drawn from Y.

Let σ_{E_A,E_B} be the shared entanglement between Alice and Bob where register E_A is with Alice and E_B with Bob. Alice applies unitary $U : \mathcal{H}_X \otimes \mathcal{H}_{E_A} \to \mathcal{H}_X \otimes \mathcal{H}_A \otimes \mathcal{H}_M$, where $E_A \equiv AM$, sends the message register M to Bob, and then Bob applies the unitary $V : \mathcal{H}_Y \otimes \mathcal{H}_M \otimes \mathcal{H}_{E_B} \to \mathcal{H}_Y \otimes \mathcal{H}_{B'} \otimes \mathcal{H}_Z$, where $ME_B \equiv B'Z$. Since unitary operations by Alice and Bob are conditioned on their respective inputs, the unitaries U, V are of the form $U = \sum_x |x\rangle\langle x|_X \otimes U_x$ and $V = \sum_y |y\rangle\langle y|_Y \otimes V_y$, where $U_x : \mathcal{H}_{E_A} \to \mathcal{H}_A \otimes \mathcal{H}_M$ and $V_y : \mathcal{H}_M \otimes \mathcal{H}_{E_B} \to \mathcal{H}_{B'} \otimes \mathcal{H}_Z$. Let the following be the global state before Alice applies her unitary:

$$\theta_{XYE_AE_B} = \sum_{xy}\mu^k(x,y)\,|xy\rangle\langle xy|_{XY} \otimes \sigma_{E_AE_B}.$$

Let $D = D_1 \cdots D_k$ be a random variable uniformly distributed over $\{0,1\}^k$ and independent of the input XY. Define random variables $U_1, U_2 \ldots U_k$ such that $U_i = X_i$ if $D_i = 0$

and $U_i = Y_i$ if $D_i = 1$. Let $U = U_1, U_2 \dots U_k$. Consider the state $\theta_{XYE_AE_BDU}$, with registers D, U as defined above.

Let $\rho_{XYAME_BDU} \stackrel{\text{def}}{=} U\theta_{XYE_AE_BDU}U^\dagger$ be the state after Alice applies her unitary and sends M to Bob. Since

$$\mathrm{I}(XE_AE_B : Y|DU)_\theta = 0,$$

it holds that

$$\mathrm{I}(XAE_BM : Y|DU)_\rho = 0.$$

From the definition of DU, we thus have (below $-i$ represents the set $\{1, 2 \dots i-1, i+1 \dots k\}$),

$$\begin{aligned}&\mathrm{I}(X_{-i}AE_BM : Y|X_iD_{-i}U_{-i})_\rho\\ &= \mathrm{I}(XAE_BM : Y_{-i}|Y_iD_{-i}U_{-i})_\rho = 0.\end{aligned}$$

Since $\log|M| \le ck$ and register E_B is independent of registers $XYDU$ in the state ρ_{E_BXYDU}, we have

$$\begin{aligned}\mathrm{I}(XYDU : ME_B)_\rho &= \mathrm{I}(XYDU : E_B)_\rho\\ &\quad + \mathrm{I}(XYDU : M|E_B)_\rho\\ &= \mathrm{I}(XYDU : M|E_B)_\rho\\ &\le 2\log|M| \le 2ck.\end{aligned}$$

where the second last inequality is from Fact 12. Consider

$$\begin{aligned}2ck &\ge \mathrm{I}(XYDU : ME_B)_\rho \ge \mathrm{I}(XY : ME_B|DU)_\rho\\ &\ge \sum_{i=1}^{k} \mathrm{I}(X_iY_i : ME_B|DU)_\rho \quad \text{(Fact 13)}\\ &= \sum_{i=1}^{k} \mathrm{I}(X_iY_i : ME_B|D_iU_iD_{-i}U_{-i})_\rho\\ &= \frac{1}{2}\sum_{i=1}^{k} \mathrm{I}(X_i : ME_B|Y_iD_{-i}U_{-i})_\rho\\ &\quad + \mathrm{I}(Y_i : ME_B|X_iD_{-i}U_{-i})_\rho\\ &\ge \frac{1}{2}\sum_{i=1}^{k} \mathrm{I}(X_i : ME_B|Y_iD_{-i}U_{-i})_\rho .\end{aligned}$$

where the last equality is from the definition of DU and the last inequality is from Fact 11. Hence there exists $j \in [k]$ such that

$$\mathrm{I}\big(X_j : ME_B \big| Y_jD_{-j}U_{-j}\big)_\rho \le 4c. \tag{6}$$

Furthermore, we have

$$\mathrm{I}\big(X_jY_j : D_{-j}U_{-j}\big)_\rho = \mathrm{I}\big(X_jY_j : D_{-j}U_{-j}\big)_\theta = 0. \tag{7}$$

since the unitary by Alice does not change the state on registers $DUXY$.

For brevity, set $B \stackrel{\text{def}}{=} ME_B$. Define the following states, which are obtained by conditioning on various classical registers:

$$\begin{aligned}\rho_B^{x_jy_jd_{-j}u_{-j}} &\stackrel{\text{def}}{=} \frac{\langle x_jy_jd_{-j}u_{-j}|\,\rho_{BXYDU}\,|x_jy_jd_{-j}u_{-j}\rangle}{\langle x_jy_jd_{-j}u_{-j}|\,\rho_{XYDU}\,|x_jy_jd_{-j}u_{-j}\rangle},\\ \rho_B^{x_jd_{-j}u_{-j}} &\stackrel{\text{def}}{=} \frac{\langle x_jd_{-j}u_{-j}|\,\rho_{BXDU}\,|x_jd_{-j}u_{-j}\rangle}{\langle x_jd_{-j}u_{-j}|\,\rho_{XDU}\,|x_jd_{-j}u_{-j}\rangle}\\ \rho_B^{y_jd_{-j}u_{-j}} &\stackrel{\text{def}}{=} \frac{\langle y_jd_{-j}u_{-j}|\,\rho_{BYDU}\,|y_jd_{-j}u_{-j}\rangle}{\langle y_jd_{-j}u_{-j}|\,\rho_{YDU}\,|y_jd_{-j}u_{-j}\rangle}\end{aligned}$$

From (6), we have

$$\mathrm{I}\big(Y : B \big| X_jD_{-j}U_{-j}\big)_\rho = 0.$$

which is equivalent to, using Fact 16 and the fact that registers X, Y, U, D are classical in ρ_B:

$$\mathop{\mathbb{E}}_{x_jy_jd_{-j}u_{-j}}\Big[\mathrm{S}\Big(\rho_B^{x_jy_jd_{-j}u_{-j}}\Big\|\rho_B^{x_jd_{-j}u_{-j}}\Big)\Big] = 0,$$

where $x_jy_jd_{-j}u_{-j}$ are drawn from the distribution $X_jY_jD_{-j}U_{-j}$.

This implies $\rho_B^{x_jy_jd_{-j}u_{-j}} = \rho_B^{x_jd_{-j}u_{-j}}$ for all $x_j, y_j, d_{-j}u_{-j}$.

From (6), and Fact 16,

$$\mathop{\mathbb{E}}_{x_jy_jd_{-j}u_{-j}}\Big[\mathrm{S}\Big(\rho_B^{x_jy_jd_{-j}u_{-j}}\Big\|\rho_B^{y_jd_{-j}u_{-j}}\Big)\Big] \le 4c,$$

where $x_jy_jd_{-j}u_{-j}$ are drawn from the distribution $X_jY_jD_{-j}U_{-j}$.

Let $G \stackrel{\text{def}}{=}$

$$\left\{(x_j, y_j, d_{-j}, u_{-j}) : \mathrm{S}\Big(\rho_B^{x_jy_jd_{-j}u_{-j}}\Big\|\rho_B^{y_jd_{-j}u_{-j}}\Big) \le \frac{4c}{\delta}\right\}.$$

By Markov's inequality,

$$\Pr\big[X_jY_jD_{-j}U_{-j} \in G\big] \ge 1 - \delta.$$

Now, we exhibit an entanglement-assisted one-way protocol $\mathcal{Q}'$ for f with communication less than c and distributional error ε under distribution μ.

1) Alice and Bob share public coins according to distribution $\rho_{D_{-j}U_{-j}}$, and the shared entanglement needed to run the protocol $\mathcal{P}$ from Theorem 18.
2) Alice and Bob are given the input $(x, y) \sim \mu$. They embed the input to the j-th coordinate X_jY_j. The input is independent of shared randomness, from equation (7).
3) Given input $(x_j, y_j) \equiv (x, y)$ and shared public coins $d_{-j}u_{-j}$, Alice knows the eigen-decomposition of the state $\rho_B^{x_jy_jd_{-j}u_{-j}}$, since $\rho_B^{x_jy_jd_{-j}u_{-j}} = \rho_B^{x_jd_{-j}u_{-j}}$. Bob knows the eigen-decomposition of state $\rho_B^{y_jd_{-j}u_{-j}}$.
4) They run the protocol in Theorem 18 with inputs $\rho_B^{x_jy_jd_{-j}u_{-j}}$, $\frac{4c}{\delta}$ (given to Alice) and $\rho_B^{y_jd_{-j}u_{-j}}$, $\frac{4c}{\delta}$ (given to Bob). After communicating $\mathcal{O}(4c/\delta^9)$ bits to Bob, Bob receives a state $\sigma_B^{x_jy_jd_{-j}u_{-j}}$ satisfying $\|\sigma_B^{x_jy_jd_{-j}u_{-j}} - \rho_B^{x_jy_jd_{-j}u_{-j}}\|_1 \le \delta$ if $(x_j, y_j, d_{-j}, u_{-j}) \in G$.
5) Bob samples the distribution from $\rho_{Y_{-j}}$, since he has the registers $D_{-j}U_{-j}Y_j$. This is possible from equation 6, which states that register Y_{-j} is independent of registers A, B, X conditioned on registers $D_{-j}U_{-j}Y_j$.
6) Bob applies the unitary V, as in the protocol $\mathcal{Q}$, on registers $BY \equiv E_BMY$ and then measures the register Z. He outputs the outcome.

From the protocol, it is clear that overall distributional error in $\mathcal{Q}'$ is at most $2\delta + \varepsilon$. The error 2δ occurs since the state $\sigma_B^{x_j y_j d_{-j} u_{-j}}$ satisfies $\|\sigma_B^{x_j y_j d_{-j} u_{-j}} - \rho_B^{x_j y_j d_{-j} u_{-j}}\|_1 \le \delta$ and the probability that $(x_j, y_j, d_{-j}, u_{-j}) \notin G$ is at most δ. The error ε is due to the original protocol $\mathcal{Q}$. Hence

$$\mathrm{Q}_{\varepsilon+2\delta}^{\mathrm{ent},\mathrm{A}\to\mathrm{B},\mu}(f) \le \mathcal{O}((c+1)/\delta^9),$$

which implies (changing $\delta \to \frac{\delta}{2}$)

$$\mathrm{Q}_{\varepsilon}^{\mathrm{ent},\mathrm{A}\to\mathrm{B},\mu^k}\left(f^k\right) \ge \Omega\left(k\left(\delta^9 \cdot \mathrm{Q}_{\varepsilon+\delta}^{\mathrm{ent},\mathrm{A}\to\mathrm{B},\mu}(f) - 1\right)\right).$$

□

V. Quantum Correlated Sampling

In this section, we give a quantum analogue to classical correlated sampling. In our framework, Alice and Bob (given quantum states ρ and σ respectively as inputs) create a joint quantum state with marginals ρ and σ on respective sides. The joint state has the property that same projective measurement performed by Alice and Bob gives very correlated outcomes, if ρ and σ are close to each other in ℓ_1 distance. Following theorem makes this sampling task precise.

Theorem 31: Let ρ, σ be quantum states on a Hilbert space $\mathcal{H}$ of dimension N. Alice is given the eigen-decomposition of ρ and Bob is given the eigen-decomposition of σ. There exists a zero-communication protocol satisfying the following.

1) *Alice outputs registers A_1, A_2 and and Bob outputs registers B_1, B_2 respectively, such that state in A_1 is ρ, the state in B_1 is σ and $A_1 \equiv B_1, A_2 \equiv B_2$.*
2) *Let $M = \{M_1, M_2 \dots M_w\}$ be a projective measurement, in the support of A_1A_2. Let M be performed by Alice on the joint system A_1A_2 with outcome $I \in [w]$ and by Bob on the joint system B_1B_2 with outcome $J \in [w]$. Then* $\Pr[I = J] \ge \left(1 - \sqrt{\|\rho - \sigma\|_1 - \frac{1}{4}\|\rho - \sigma\|_1^2}\right)^3$.

Proof: Let eigen-decomposition of ρ be $\sum_{i=1}^N a_i |a_i\rangle\langle a_i|$ and of σ be $\sum_{i=1}^N b_i |b_i\rangle\langle b_i|$. Let $\{|1\rangle, |2\rangle \dots |N\rangle\}$ be an orthonormal basis for $\mathcal{H}$. We assume that $a_1, \dots, a_N, b_1, \dots b_N$ are rational numbers and let K be the least common multiple of their denominators. The error due to this assumption goes to 0 as $K \to \infty$.

Introduce registers A_1, B_1 associated to $\mathcal{H}$ and registers A_2, B_2 associated to some Hilbert space $\mathcal{H}'$ with an orthonormal basis $\{|1\rangle, |2\rangle \dots |K\rangle\}$.

Consider the following state shared in A_1, A_2, B_1, B_2.

$$|S\rangle_{A_1B_1A_2B_2} \stackrel{\mathrm{def}}{=} \frac{1}{\sqrt{KN}} \sum_{i=1}^N |i, i\rangle_{A_1B_1} \otimes \left(\sum_{m=1}^K |m, m\rangle_{A_2B_2}\right)$$

For brevity, define the registers $A \stackrel{\mathrm{def}}{=} A_1A_2$ and $B \stackrel{\mathrm{def}}{=} B_1B_2$. The protocol is described below.

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

Shared resources: Alice and Bob hold infinitely many registers $A_1^iA_2^iB_1^iB_2^i$ $(i > 0)$, such that $A_1^i \equiv A_1, A_2^i \equiv A_2, B_1^i \equiv B_1, B_2^i \equiv B_2$. The shared state in register $A_1^iA_2^iB_1^iB_2^i$ is $|S\rangle_{A_1^iA_2^iB_1^iB_2^i}$. Let $A \equiv A_1A_2$ and $B \equiv B_1B_2$ be used as output registers. Let i refer to the 'index' of corresponding registers.

1) For each $i > 0$, Alice performs the measurement $\{P_A, I - P_A\}$ on the registers $A_1^iA_2^i$, where
$$P_A \stackrel{\mathrm{def}}{=} \sum_i |a_i\rangle\langle a_i|_{A_1} \otimes \left(\sum_{m=1}^{Ka_i} |m\rangle\langle m|_{A_2}\right)$$
She declares *success* if she obtains outcome corresponding to P_A. She stops once she succeeds in some register A^j, and swaps A^j with A.
2) For each $i > 0$, Bob performs the measurement $\{P_B, I - P_B\}$ on the registers $B_1^iB_2^i$, where
$$P_B \stackrel{\mathrm{def}}{=} \sum_i |b_i\rangle\langle b_i|_{B_1} \otimes \left(\sum_{m=1}^{Kb_i} |m\rangle\langle m|_{B_2}\right)$$
He declares *success* if he obtains outcome corresponding to P_B. He stops once he succeeds in some register B^j, and swaps B^j with B.

At the end of above protocol, let the joint state in the register AB be τ. The following claim shows the first part of the theorem.

Claim 32: $\mathrm{Tr}_{A_2B_1B_2}(\tau) = \rho$ and $\mathrm{Tr}_{A_1A_2B_2}(\tau) = \sigma$.

Proof: It is easily seen that the marginal of the state $(P_A \otimes I_B)|S\rangle$ in register A is ρ. Similarly the marginal of the state $(I_A \otimes P_B)|S\rangle$ in register B in is σ. □

Following series of claims establish second part of the theorem.

Claim 33:

$$\tau \ge \frac{(P_A \otimes P_B)|S\rangle\langle S|(P_A \otimes P_B)}{1 - \langle S|(I_A - P_A) \otimes (I_B - P_B)|S\rangle}.$$

Proof: Consider the event that Alice and Bob succeed at the same index. The resulting state in AA_1BB_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

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

Since the cases of Bob succeeding before Alice and Alice succeeding before Bob add positive operators to τ, we get the desired. □

Claim 34: 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,

$$\begin{aligned}\langle\theta|\,\tau\,|\theta\rangle &\geq \frac{|\langle\theta|\,P_A\otimes P_B\,|S\rangle|^2}{1-\langle S|\,(I_A-P_A)\otimes(I_B-P_B)\,|S\rangle} \quad \text{(Claim 33)}\\ &= \frac{|\langle\theta|\,P_A\otimes P_B\,|S\rangle|^2}{2/N-\langle S|\,P_A\otimes P_B\,|S\rangle}\\ &\quad \text{(using } \langle S|\,P_A\otimes I_B\,|S\rangle=\langle S|\,I_A\otimes P_B\,|S\rangle=1/N).\end{aligned}$$

By direct calculation, we get

$$(P_A\otimes P_B)\,|S\rangle = \frac{1}{\sqrt{KN}}\sum_{i,j}|\overline{a_i}\rangle\,\langle b_j|a_i\rangle\,|b_j\rangle\sum_{m=1}^{K\min(a_i,b_j)}|m,m\rangle;$$

$$X\,|\theta\rangle = \frac{1}{\sqrt{K}}\sum_i|\overline{a_i}\rangle\,|a_i\rangle\sum_{m=1}^{Ka_i}|m,m\rangle\,.$$

Hence,

$$\langle\theta|\,\tau\,|\theta\rangle \geq \frac{\left(\sum_{i,j}\min(a_i,b_j)\,\left|\langle a_i|b_j\rangle\right|^2\right)^2}{2-\sum_{i,j}\min(a_i,b_j)\,\left|\langle a_i|b_j\rangle\right|^2}. \tag{8}$$

Define $R_{ij} \stackrel{\text{def}}{=} a_i\left|\langle a_i|b_j\rangle\right|^2$ and $R'_{ij} \stackrel{\text{def}}{=} b_i\left|\langle a_i|b_j\rangle\right|^2$. Note that both $\{R_{ij}\}$ and $\{R'_{ij}\}$ form probability distributions over $[N^2]$. Also note that $\mathrm{F}(R,R')=\mathrm{Tr}(\sqrt{\rho}\sqrt{\sigma})$. Consider (using relation between fidelity and ℓ_1 distance, Facts 6 and 2),

$$\begin{aligned}\sum_{i,j}\min(R_{ij},R'_{i,j}) &= 1-\frac{1}{2}\left\|R-R'\right\|_1\\ &\geq 1-\sqrt{1-\mathrm{F}(R,R')^2}\\ &= 1-\sqrt{1-(\mathrm{Tr}\sqrt{\rho}\sqrt{\sigma})^2}\\ &\geq 1-\sqrt{\|\rho-\sigma\|_1-\frac{1}{4}\|\rho-\sigma\|_1^2}.\end{aligned} \tag{9}$$

Combining Equations (8) and (9) we get the desired. □

Claim 35: Let $M = \{M_1, M_2 \ldots M_w\}$ be a projective measurement in the support of A_1A_2. Let $E=\sum_{i=1}^{w} M_i\otimes M_i$. Then $\mathrm{Tr}(E\,|\theta\rangle\langle\theta|)=1$.

Proof: Since M_i is a projector in the support of A_1A_2, we have $(M_i\otimes M_i)\,|\theta\rangle=(M_i\otimes I)\,|\theta\rangle$. Hence,

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

□

Finally using monotonicity of fidelity under quantum operation (Fact 4) and Claim 34 we get the second part of the theorem as follows.

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

□

VI. Conclusion and Open Questions

We have described two one shot quantum protocols, one of which has been applied to direct sum problem in quantum communication complexity. Our first protocol is a compression protocol, in which communication of a quantum state ρ (held by Alice) can be made much smaller than $\log(|\mathrm{supp}(\rho)|)$, given a description of an another quantum state σ with Bob. This protocol is then used to obtain a direct sum result for one round entanglement assisted communication complexity. It may be noted that this application has been superseded by a recent result of Touchette [15] for bounded round entanglement assisted communication complexity models.

Our second protocol is a quantum generalization of classical correlated sampling. We show that if Alice and Bob are given descriptions of quantum states ρ and σ, respectively, then they can create a joint state with marginals ρ (on Alice's side) and σ (on Bob's side), such that the joint state is correlated. Any measurement done joint by both parties gives highly correlated outcomes, if ρ and σ are close to each other in ℓ_1 distance.

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 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

The authors thank Mario Berta, Ashwin Nayak, Mark M. Wilde and Andreas Winter for helpful discussions. They also thank anonymous referees for important suggestions for improvement of the manuscript. Work of A.S. was done while visiting CQT, Singapore.

References

[1] M. Braverman and A. Rao, "Information equals amortized communication," in *Proc. 52nd Symp. Found. Comput. Sci. (FOCS)*, Washington, DC, USA, 2011, pp. 748–757.
[2] R. Jain, "New strong direct product results in communication complexity," *J. ACM*, vol. 62, no. 3, 2013, Art. no. 20.
[3] R. Jain, A. Pereszlényi, and P. Yao, "A direct product theorem for the two-party bounded-round public-coin communication complexity," in *Proc. IEEE 53rd Annu. Symp. Found. Comput. Sci. (FOCS)*, Washington, DC, USA, Oct. 2012, pp. 167–176. [Online]. Available: http://dx.doi.org/10.1109/FOCS.2012.42
[4] M. Braverman, A. Rao, O. Weinstein, and A. Yehudayoff, "Direct product via round-preserving compression," in *Proc. 40th Int. Conf. Automata, Lang. Program. (ICALP)*, 2013, pp. 300–315. [Online]. Available: http://dl.acm.org/citation.cfm?id=1759210.1759242
[5] R. Jain, J. Radhakrishnan, and P. Sen, "Prior entanglement, message compression and privacy in quantum communication," in *Proc. 20th Annu. IEEE Conf. Comput. Complex.*, Washington, DC, USA, Jun. 2005, pp. 285–296. [Online]. Available: http://dl.acm.org/citation.cfm?id=1068502.1068658
[6] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul./Oct. 1948. [Online]. Available: http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x

[7] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 471–480, Jul. 1973. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1055037
[8] T. Holenstein, "Parallel repetition: Simplifications and the no-signaling case," in *Proc. 39th Annu. ACM Symp. Theory Comput. (STOC)*, New York, NY, USA, 2007, pp. 411–419. [Online]. Available: http://doi.acm.org/10.1145/1250790.1250852
[9] I. Dinur, D. Steurer, and T. Vidick, "A parallel repetition theorem for entangled projection games," in *Proc. 29th Annu. Conf. Comput. Complex. (CCC)*, Washington, DC, USA, 2014, pp. 201–254.
[10] A. Winter, "Coding theorem and strong converse for quantum channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 7, pp. 2481–2485, Nov. 1999. [Online]. Available: http://dblp.uni-trier.de/db/journals/tit/tit45.html#Winter99
[11] T. Ogawa and H. Nagaoka, "A new proof of the channel coding theorem via hypothesis testing in quantum information theory," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul./Jul. 2002, p. 73.
[12] R. Jain, J. Radhakrishnan, and P. Sen, "Privacy and interaction in quantum communication complexity and a theorem about the relative entropy of quantum states," in *Proc. 43rd Symp. Found. Comput. Sci. (FOCS)*, Washington, DC, USA, 2002, pp. 429–438. [Online]. Available: http://dl.acm.org/citation.cfm?id=645413.652142
[13] R. Jain and A. Nayak, "Short proofs of the quantum substate theorem," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3664–3669, Jun. 2012.
[14] P. Harsha, R. Jain, D. McAllester, and J. Radhakrishnan, "The communication complexity of correlation," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 438–449, Jan. 2010.
[15] D. Touchette, "Quantum information complexity," in *Proc. 47th Annu. ACM Symp. Theory Comput. (STOC)*, 2015, pp. 317–326. [Online]. Available: http://doi.acm.org/10.1145/2746539.2746613
[16] I. Devetak and J. Yard, "Exact cost of redistributing multipartite quantum states," *Phys. Rev. Lett.*, vol. 100, no. 23, p. 230501, 2008. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevLett.100.230501
[17] J. T. Yard and I. Devetak, "Optimal quantum source coding with quantum side information at the encoder and decoder," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5339–5351, Nov. 2009.
[18] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge, U.K.: Cambridge Univ. Press, 2000.
[19] W. F. Stinespring, "Positive functions on c*-algebras," *Proc. Amer. Math. Soc.*, vol. 6, no. 2, pp. 211–216, 1955.
[20] H. Barnum, C. M. Cave, C. A. Fuch, R. Jozsa, and B. Schmacher, "Noncommuting mixed states cannot be broadcast," *Phys. Rev. Lett.*, vol. 76, pp. 2818–2821, Apr. 1996.
[21] G. Lindblad, "Completely positive maps and entropy inequalities," *Commun. Math. Phys.*, vol. 40, no. 2, pp. 147–151, Jun. 1975.
[22] J. Watrous. (2011). *Theory of Quantum Information, Lecture Notes*. [Online]. Available: https://cs.uwaterloo.ca/~/LectureNotes.html
[23] R. Jain, J. Radhakrishnan, and P. Sen, "A new information-theoretic property about quantum states with an application to privacy in quantum communication," *J. ACM*, vol. 56, no. 6, Sep. 2009, Art. no. 33.
[24] M. B. Ruskai, "Inequalities for quantum entropy: A review with conditions for equality," *J. Math. Phys.*, vol. 43, pp. 4358–4375, Sep. 2002. [Online]. Available: http://scitation.aip.org/content/aip/journal/jmp/43/9/10.1063/1.1497701
[25] E. H. Lieb, "Convex trace functions and the Wigner–Yanase–Dyson conjecture," *Adv. Math.*, vol. 11, no. 3, pp. 267–288, 1973. [Online]. Available: http://www.sciencedirect.com/science/article/pii/000187087390011X
[26] E. H. Lieb and M. B. Ruskai, "Proof of the strong subadditivity of quantum-mechanical entropy," *J. Math. Phys.*, vol. 14, no. 12, p. 1938, 1973.
[27] H. Araki and E. H. Lieb, "Entropy inequalities," *Commun. Math. Phys.*, vol. 18, no. 2, pp. 160–170, 1970.
[28] D. Touchette. (2014). "Quantum information complexity and amortized communication." [Online]. Available: https://arxiv.org/abs/1404.3733
[29] A. C. Yao, "Some complexity questions related to distributive computing (preliminary report)," in *Proc. 11th Annu. ACM Symp. Theory Comput. (STOC)*, 1979, pp. 209–213. [Online]. Available: http://doi.acm.org/10.1145/800135.804414
[30] I. Kerenidis, S. Laplante, V. Lerays, J. Roland, and D. Xiao, "Lower bounds on information complexity via zero-communication protocols and applications," in *Proc. IEEE 53rd Annu. Symp. Found. Comput. Sci. (FOCS)*, Oct. 2012, pp. 500–509. [Online]. Available: http://dx.doi.org/10.1109/FOCS.2012.68

Anurag Anshu is pursuing his Ph. D. degree in computer science at Centre for Quantum Technologies, National University of Singapore, Singapore. His research interests are in quantum information theory, communication complexity and quantum hamiltonian complexity.

Rahul Jain obtained his Ph.D. degree in computer science from the Tata Institute of Fundamental Research, Mumbai, India, in 2003. He completed two postdoctoral fellowships: two years at the University of California, Berkeley, CA, USA, followed by two years at the Institute for Quantum Computing at the University of Waterloo, Waterloo, ON, Canada. He joined Centre for Quantum Technologies (CQT), Singapore, as a Principal Investigator and the National University of Singapore (NUS), Singapore, as an Assistant Professor in 2008. He is presently an Associate Professor (starting July 2013) at NUS and Principal Investigator at CQT. His research interests are in the areas of information theory, quantum computation, cryptography, communication complexity, and computational complexity theory.

Priyanka Mukhopadhyay is pursuing her Ph.D. degree with a major in Mathematics at Centre for Quantum Technologies, National University of Singapore, Singapore. Her research interests include computational and algebraic complexity, information theory and quantum computation.

Ala Shayeghi is pursuing his Ph.D. degree in Mathematics in the department of Combinatorics and Optimization and Institute for Quantum Computing, at the University of Waterloo, Canada. His research interests are in quantum computing, classical and quantum information theory and communication complexity.

Penghui Yao obtained his Ph.D. degree in computer science from the Centre for Quantum Technologies (CQT), National University of Singapore in 2013. He spent one year at CQT as a research associate, one year at the Centrum Wiskunde & Informatica in Netherlands as a postdoc, one year at the Institute for Quantum Computing at the University of Waterloo, Waterloo, ON, Canada as a postdoc. He is presently a Hartree postdoctoral fellow at the Joint Center for Quantum Information and Computer Science, University of Maryland, MD, USA. His research interests are in the areas of communication complexity, information theory, computational complexity and quantum tomography.