Smooth min-max relative entropy based bounds for one-shot classical and quantum state redistribution

Anurag Anshu∗ Rahul Jain† Naqueeb Ahmad Warsi‡

Abstract

We study the problem of state redistribution both in the classical (shared randomness assisted) and quantum (entanglement assisted) one-shot settings and provide new upper bounds on the communication required. Our bounds are in terms of smooth-min and max relative entropies. We also consider a special case of this problem in the classical setting, previously studied by Braverman and Rao (2011). We show that their upper bound is optimal. In addition we provide an alternate protocol achieving a priori different looking upper bound. However, we show that our upper bound is essentially the same as their upper bound and hence also optimal.

1 Introduction

Quantum state redistribution is a well studied task in both in the asymptotic and the one-shot settings [1, 2, 3, 4, 5, 6, 7, 8, 9]. In this work we provide a new upper bound on the communication needed to accomplish this task (with a small error) using entanglement assisted one-shot protocols. We also consider classical analogue of this task which we call as classical state redistribution. We provide optimal (upper and lower) bounds on the amount of communication needed for this task using shared randomness assisted one-shot protocols.

To arrive at our result, we consider a simpler communication task: Alice possesses random variables $X, M$, Bob possesses random variable $Y$, where $Y \rightarrow X \rightarrow M$ is a Markov chain. Alice wants to transfer $M$ to Bob with a small error. This task was first considered by [10], wherein they provided an upper bound on the communication (without an optimality proof). We show that their upper bound is optimal. In addition we provide an alternate protocol achieving a priori different looking upper bound. However, we show that our upper bound is essentially the same as the upper bound by [10] and hence also optimal. We provide a quantum protocol for quantum state redistribution based on our classical protocol mentioned above.

To design our classical protocol, we use the techniques of rejection sampling and hypothesis testing. For our quantum protocol we use the techniques of convex split [9] (which can be viewed as a coherent quantum analogue of rejection sampling technique) and quantum hypothesis testing.

∗Centre for Quantum Technologies, National University of Singapore, Singapore. a0109169@u.nus.edu
†Centre for Quantum Technologies, National University of Singapore and MajuLab, UMI 3654, Singapore. rahul@comp.nus.edu.sg
‡Centre for Quantum Technologies, National University of Singapore and School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore and IIITD, Delhi. warsi.naqueeb@gmail.com
Organisation of the paper

We introduce our notations and notions used throughout the paper in Section 2. In Section 3 we describe our achievability and converse bounds for the task of classical state redistribution. These results are heavily based on a series of results shown in [4] where we prove the optimality of Braverman–Rao protocol and present a new protocol using extension states. Furthermore, in this section, we prove the equivalence of bounds obtained by these two different techniques. In Section 5 we present a new protocol for quantum state redistribution, inspired by the classical protocols developed in earlier sections.

2 Preliminaries

Classical information theory

Given a random variable $X$ taking values in a set $\mathcal{X}$, we will represent the probability of an $x \in \mathcal{X}$ as $\Pr_X(x)$. We will represent the logarithm of support size of $X$ with $|X|$. Given two random variables $X, Y$ we represent by $X \otimes Y$ the joint random variable such that $\Pr_{X \otimes Y}(x, y) = \Pr_X(x) \cdot \Pr_Y(y)$. If the random variable $XYZ$ forms a Markov chain with $Y, Z$ being independent conditioned on $X$, we shall represent it, notationally, as $X \rightarrow Y \rightarrow Z$. Given a random variable $X$, we define $\text{supp}(X)$ as the set of all $x \in \mathcal{X}$ such that $\Pr_X(x) > 0$. Given a joint random variable $XY \in \mathcal{X} \otimes \mathcal{Y}$, we denote the random variable $X$ conditioned on $Y$ taking the value $y$ as $X_y$. The distance measure that we use is the trace distance, which is defined as: for random variables $X_1, X_2 \in \mathcal{X}, \|X_1 - X_2\|_1 := \sum_{x \in \mathcal{X}} |\Pr_X(x) - \Pr_{X'}(x)|$. We also define fidelity: $F(X_1, X_2) = \sum_x \sqrt{\Pr_{X_1}(x) \Pr_{X_2}(x)}$.

Fact 1. For random variables $X_1, X_2$ over the same set $\mathcal{X}$, it holds that

$$1 - F(X_1, X_2) \leq \frac{1}{2} \|X_1 - X_2\|_1 \leq \sqrt{1 - F^2(X_1, X_2)}.$$

Given two random variables $X, X'$, we define

- $D_\infty(X||X') := \max_x \log \frac{\Pr_X(x)}{\Pr_{X'}(x)}$.
- $D_\infty^\varepsilon(X||X') := \min_{\|X'' - X\|_\infty \leq \varepsilon} D_\infty(X''||X').$
- $D_\varepsilon(X||X') := \inf \{ a : \Pr_{X \rightarrow X}(\frac{\Pr_X(x)}{\Pr_{X'}(x)} > 2^a) < \varepsilon \}$.
- $D_{\text{KL}}(X||X') := \max_A \{ \log \frac{1}{\Pr_X(A)} : \Pr_X(A) \geq 1 - \varepsilon \}$.

Fact 2. Given a random variable $X$ and an event $A$ such that $\Pr_X(A) = p$. Let the random variable $X_1$ represent $X$ conditioned on event $A$. Then for all $x$, $\Pr_{X_1}(x) \leq \frac{\Pr_X(x)}{p}$.

Proof. We use Bayes’ theorem to proceed as follows.

$$\Pr_{X_1}(x) = \Pr_{X_1}(x|A) = \frac{\Pr_X(x, A)}{\Pr_X(A)} \leq \frac{\Pr_X(x)}{\Pr_X(A)}.$$

Fact 3. For $(0 \leq x \leq 1, y \geq 0)$, it holds that

$$ (1 - x)^y \leq e^{-xy} \quad \text{and} \quad e^{-xy} \leq 1 - x + e^{-y}. $$

Fact 4. Let $XY \in \mathcal{X} \times \mathcal{Y}$ be a joint random variable and $\varepsilon > 0$ be a constant. Then it holds that $D_\varepsilon^a(XY||X \otimes Y) \leq |Y| + \log \frac{1}{\varepsilon}$. 

2
exists a randomness assisted one-way protocol with communication cost $Q$ be given a probability distribution $P$ and $Q$ be given a probability distribution $Q$. For any $\epsilon > 0$, it is shown (Proposition 13, \cite{11}) that

$$\inf_{X'' : P(x,y) \geq 1 - \epsilon} D_\infty(X'' \| X') \geq \tilde{D}_s^2(X \| X') + 2 \log \epsilon - 2 - \log(1 - \epsilon).$$

Fact will follow if we show that $\tilde{D}_s^2(X \| X')$ is a feasible value for the infimum in the definition of $D_\infty(X \| X')$. Thus, by definition, we obtain that $D_\infty(X \| X') \leq |B| + \log \frac{1}{\epsilon}$.

Fact 5 (\cite{11}). Let $X, X'$ be probability distributions. It holds that

$$D_\infty(X \| X') \geq \tilde{D}_s^2(X \| X') - 2 \log \left(\frac{2}{\epsilon}\right).$$

Proof. In \cite{11}, the following entropic quantity is defined

$$\tilde{D}_s^2(X \| X') := \sup\{\alpha : \Pr_{X \leftarrow X}(\Pr_X(x) \leq 2^\alpha) \leq \epsilon\}.$$

For any $\epsilon > 0$, it is shown (Proposition 13, \cite{11}) that

$$\inf_{X'' : P(x,y) \geq 1 - \epsilon} D_\infty(X'' \| X') \geq \tilde{D}_s^2(X \| X') + 2 \log \epsilon - 2 - \log(1 - \epsilon).$$

Now, $D_\infty(X \| X') = \inf_{X'' : \|X - X''\|_1 \leq \epsilon} D_\infty(X'' \| X') \geq \inf_{X'' : P(x,y) \geq 1 - \epsilon} D_\infty(X'' \| X')$, since using Fact 1, \cite{11} \(F^2(X, X'') \geq (1 - \epsilon/2)^2 \geq 1 - \epsilon\). Thus, we have that

$$D_\infty(X \| X') \geq \tilde{D}_s^2(X \| X') + 2 \log \epsilon - 2 - \log(1 - \epsilon).$$

The following protocol was given by Braverman and Rao \cite{10}.

Lemma 1 (Braverman and Rao protocol, \cite{10}). Let Alice be given a probability distribution $P$ and Bob be given a probability distribution $Q$. Suppose Alice and Bob know an upper bound $c$ on the value of $D_\infty(P \| Q)$. Then for every $\epsilon > 0$, there exists a randomness assisted one-way protocol with communication cost $c + \log(\frac{1}{\epsilon})$ such that Bob correctly outputs the distribution $P$ with probability at least $1 - \epsilon$.

An immediate corollary to this is the following.

Corollary 1. Let Alice be given a probability distribution $P$ and Bob be given a probability distribution $Q$. Fix an $\epsilon > 0$. Suppose Alice and Bob know an upper bound $c$ on the value of $D_\infty(P \| Q)$. Then there exists a randomness assisted one-way protocol with communication cost $c + \log(\frac{1}{\epsilon})$ such that Bob outputs a distribution $P'$ such that $\|P' - P\|_1 \leq 3\epsilon$.

Proof. From the definition of $D_\infty(P \| Q)$, it holds that $\Pr_{P \leftarrow P}(\Pr_{P \leftarrow P}(p) \geq 2^c) \leq \epsilon$. Let the set of $p$ be Good if $\Pr_{P \leftarrow P}(p) \geq 2^c$. Define a distribution $P''$ as $\Pr_{P''}(p) = \Pr_{P \leftarrow P}(p)$ if $p \in$ Good, and $0$ otherwise. Then it holds that $\|P'' - P\|_1 \leq 2\epsilon$ and

$$2D_\infty(P'' \| Q) = \frac{1}{\Pr_{P \leftarrow P}(\text{Good})} \cdot \max_{p \in \text{Good}} \Pr_{P \leftarrow P}(p) \Pr_{P \leftarrow P}(p) \leq \frac{2^c}{1 - \epsilon}.$$

Suppose Alice were given the distribution $P''$ and Bob were given the distribution $Q$ and they knew $c + \log(\frac{1}{1 - \epsilon})$. Consider the protocol $Q$ as defined by Lemma 1. With probability at least $1 - \epsilon$ it correctly outputs the distribution $P''$. The protocol $Q$ when run on the distribution $P$, thus produces a distribution $P'$ such that $\|P' - P\|_1 \leq 3\epsilon$. The communication cost is at most $c + \log(\frac{1}{\epsilon}) + \log(\frac{1}{1 - \epsilon}) \leq c + \log(\frac{2}{\epsilon})$. This proves the corollary.
Quantum information theory

Consider a finite dimensional Hilbert space $\mathcal{H}$ endowed with an inner product $\langle \cdot, \cdot \rangle$ (In this paper, we only consider finite dimensional Hilbert spaces). The $\ell_1$ norm of an operator $X$ on $\mathcal{H}$ is $\|X\|_1 := \text{Tr} \sqrt{X^\dagger X}$ and $\ell_2$ norm is $\|X\|_2 := \sqrt{\text{Tr} XX^\dagger}$. A quantum state (or a density matrix or a state) is a positive semi-definite matrix on $\mathcal{H}$ with trace equal to 1. It is called pure if and only if its rank is 1. A sub-normalized state is a positive semi-definite matrix on $\mathcal{H}$ 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 $\psi$ to represent the state and also the density matrix $|\psi\rangle\langle \psi|$, associated with $|\psi\rangle$. Given a quantum state $\rho$ on $\mathcal{H}$, support of $\rho$, called supp$(\rho)$ is the subspace of $\mathcal{H}$ spanned by all eigen-vectors of $\rho$ with non-zero eigenvalues.

A quantum register $A$ is associated with some Hilbert space $\mathcal{H}_A$. Define $|A| := \text{dim}(\mathcal{H}_A)$. 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 $\rho$ 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 $\rho$ and $\sigma$. The identity operator on $\mathcal{H}_A$ (and associated register $A$) is denoted $I_A$.

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

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

where $\{|i\rangle\}$ 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 $\rho_{AB}$. Unless otherwise stated, a missing register from subscript in a state will represent the state and also the density matrix $|\psi\rangle\langle \psi|$, associated with $|\psi\rangle$. Given a quantum state $\rho$ on $\mathcal{H}$, support of $\rho$, called supp$(\rho)$ is the subspace of $\mathcal{H}$ spanned by all eigen-vectors of $\rho$ with non-zero eigenvalues.

A quantum map $\mathbb{E} : \mathcal{L}(A) \rightarrow \mathcal{L}(B)$ 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 $U_A : \mathcal{H}_A \rightarrow \mathcal{H}_A$ is such that $U_A^\dagger U_A = U_A U_A^\dagger = I_A$. An isometry $V : \mathcal{H}_A \rightarrow \mathcal{H}_B$ 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. Reader is referred to [12, 13, 14, 15] for many of these definitions. We consider only normalized states in the definitions below. Let $\varepsilon \geq 0$.

1. **Fidelity** For $\rho_A, \sigma_A \in \mathcal{D}(A)$,

$$F(\rho_A, \sigma_A) \overset{\text{def}}{=} \|\sqrt{\rho_A} \sqrt{\sigma_A}\|_1.$$ For classical probability distributions $P = \{p_i\}, Q = \{q_i\}$,

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

2. **Purified distance** For $\rho_A, \sigma_A \in \mathcal{D}(A)$,

$$P(\rho_A, \sigma_A) = \sqrt{1 - F^2(\rho_A, \sigma_A)}.$$ 

3. **$\varepsilon$-ball** For $\rho_A \in \mathcal{D}(A)$,

$$\mathcal{B}_\varepsilon(\rho_A) \overset{\text{def}}{=} \{\rho_A' \in \mathcal{D}(A) | P(\rho_A, \rho_A') \leq \varepsilon\}.$$ 

4. **Von-neumann entropy** For $\rho_A \in \mathcal{D}(A)$,

$$S(\rho_A) \overset{\text{def}}{=} -\text{Tr}(\rho_A \log \rho_A).$$
5. **Relative entropy** For $\rho_A, \sigma_A \in \mathcal{D}(A)$ such that $\text{supp}(\rho_A) \subset \text{supp}(\sigma_A)$,

$$D(\rho_A||\sigma_A) \overset{\text{def}}{=} \text{Tr}(\rho_A \log \rho_A) - \text{Tr}(\rho_A \log \sigma_A).$$

6. **Max-relative entropy** For $\rho_A, \sigma_A \in \mathcal{D}(A)$ such that $\text{supp}(\rho_A) \subset \text{supp}(\sigma_A)$,

$$D_{\text{max}}(\rho_A||\sigma_A) \overset{\text{def}}{=} \inf\{\lambda \in \mathbb{R} : 2^\lambda \sigma_A \geq \rho_A\}.$$

7. **Smooth min-relative entropy** For $\rho_A, \sigma_A \in \mathcal{D}(A)$,

$$D_{\text{H}}(\rho_A||\sigma_A) \overset{\text{def}}{=} \sup_{0 < \Pi < I, \text{Tr}(\Pi \rho_A) \geq 1 - \varepsilon} \log\left(\frac{1}{\text{Tr}(\Pi \sigma_A)}\right).$$

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

$$I(A : B)_\rho \overset{\text{def}}{=} S(\rho_A) + S(\rho_B) - S(\rho_{AB}) = D(\rho_{AB}||\rho_A \otimes \rho_B).$$

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

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

10. **Max-information** For $\rho_{AB} \in \mathcal{D}(AB)$,

$$I_{\text{max}}(A : B)_\rho \overset{\text{def}}{=} \inf_{\sigma_B \in \mathcal{D}(B)} D_{\text{max}}(\rho_{AB}||\rho_A \otimes \sigma_B).$$

11. **Smooth max-information** For $\rho_{AB} \in \mathcal{D}(AB)$,

$$I_{\text{max}}^\varepsilon(A : B)_\rho \overset{\text{def}}{=} \inf_{\rho' \in \mathcal{B}(\rho)} I_{\text{max}}(A : B)_{\rho'}.$$

We will use the following facts.

**Fact 6** (Triangle inequality for purified distance, [14]). For states $\rho_A, \sigma_A, \tau_A \in \mathcal{D}(A)$,

$$P(\rho_A, \sigma_A) \leq P(\rho_A, \tau_A) + P(\tau_A, \sigma_A).$$

**Fact 7** ([15]). **(Stinespring representation)** Let $\mathbb{E}(\cdot) : \mathcal{L}(A) \to \mathcal{L}(B)$ be a quantum operation. There exists a register $C$ and an unitary $U \in \mathcal{U}(ABC)$ such that $\mathbb{E}(\omega) = \text{Tr}_{A,C} \left(U (\omega \otimes |0\rangle\langle 0|) U^\dagger \right)$. Stinespring representation for a channel is not unique.

**Fact 8** (Monotonicity under quantum operations, [17],[18]). For quantum states $\rho, \sigma \in \mathcal{D}(A)$, and quantum operation $\mathbb{E}(\cdot) : \mathcal{L}(A) \to \mathcal{L}(B)$, it holds that

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

In particular, for bipartite states $\rho_{AB}, \sigma_{AB} \in \mathcal{D}(AB)$, it holds that

$$\|\rho_{AB} - \sigma_{AB}\|_1 \geq \|\rho_A - \sigma_A\|_1 \quad \text{and} \quad F(\rho_{AB}, \sigma_{AB}) \leq F(\rho_A, \sigma_A) \quad \text{and} \quad D(\rho_{AB}||\sigma_{AB}) \geq D(\rho_A||\sigma_A).$$

**Fact 9** (Uhlmann’s theorem, [19]). Let $\rho_A, \sigma_A \in \mathcal{D}(A)$. Let $\rho_{AB} \in \mathcal{D}(AB)$ be a purification of $\rho_A$ and $\sigma_{AC} \in \mathcal{D}(AC)$ be a purification of $\sigma_A$. There exists an isometry $V : C \to B$ such that,

$$F(\langle \theta|\rho_{AB}, |\rho|\rho_{AB}) = F(\rho_A, \sigma_A),$$

where $|\theta\rangle_{AB} = (I_A \otimes V)|\sigma\rangle_{AC}$. 

5
**Fact 10** (Gentle measurement lemma, [20, 21]). Let $\rho$ be a quantum state and $0 < A < I$ be an operator. Then

$$F(\rho, \frac{A\rho A}{\text{Tr}(A^2\rho)}) \geq \sqrt{\frac{\text{Tr}(A^2\rho)}{\text{Tr}(A^2\rho)}}.$$  

**Proof.** Let $|\rho\rangle$ be a purification of $\rho$. Then $(I \otimes A)|\rho\rangle$ is a purification of $A\rho A$. Now, applying monotonicity of fidelity under quantum operations (Fact 8), we find

$$F(\rho, \frac{A\rho A}{\text{Tr}(A^2\rho)}) \geq F(|\rho\rangle\langle\rho|, (I \otimes A)|\rho\rangle\langle\rho|(I \otimes A^\dagger)) = \sqrt{\frac{\text{Tr}(A\rho)^2}{\text{Tr}(A^2\rho)}} > \sqrt{\text{Tr}(A^2\rho)}.$$  

In the last inequality, we have used $A > A^2$.  

**Fact 11** (Hayashi-Nagaoka inequality, [22]). Let $0 < S < I, T$ be positive semi-definite operators. Then

$$(S + T)^{-\frac{1}{2}}T(S + T)^{-\frac{1}{2}} \leq 2(I - S) + 4T.$$  

### 3 Classical state redistribution

**Task:** There are three parties: Reference, Alice and Bob. Reference possesses the random variable $R$, Alice possesses the random variables $(A, C)$ and Bob possesses the random variable $B$. Alice wants to communicate her random variable $C$ to Bob. To accomplish the task of communicating the random variable $C$ to Bob, Alice sends a message $M$ to Bob. Bob on receiving the message $M$ constructs the random variable $C'$ such that $||RABC - RABC'||_1 \leq \varepsilon$.

We define the following quantity

$$\text{Opt}^\varepsilon(RABC) \overset{\text{def}}{=} \inf_{V, R \sim AC - V, RA - BV - C', \|RBAC - RBAC'||_1 \leq \varepsilon} (D_H^V(AC'V \| AC' \otimes V) - D_H^V(BV \| B \otimes V)).$$

We prove the following results:

**Theorem 1.** (Achievability) Let $RABC$ be quadruplet of random variables taking values in $R \times A \times C \times B$ and let $\varepsilon \in (0, 1)$ be given. Let $\ell = \text{Opt}^\varepsilon(RABC) + 2 \log \frac{1}{\varepsilon}$. There exists a shared randomness assisted protocol communicating $\ell$ bits such that Bob outputs a random variable $C''$ satisfying $||RABC - RABC''||_1 \leq 7\varepsilon$.

**Proof.** Let $V, C'$ be the random variables achieving the infimum in the definition of $\text{Opt}^\varepsilon(RABC)$. The main ingredient for the proof is Lemma 2. We now invoke Lemma 2 with $AC \leftarrow X, B \leftarrow Y$ and $V \leftarrow M$ to generate a random variable $V'$ at Charlie’s end such that

$$||ACBV - ACBV'||_1 \leq 6\varepsilon. \quad (1)$$

Bob on getting the random variables pair $(V', B)$ generates a random variable $C''$, as guaranteed by the relation $RA - BV - C'$. We will now show that $||RABC - RABC''||_1 \leq 6\varepsilon$. Towards this notice the following set of inequalities:

$$6\varepsilon \overset{a}{=} \|ACVB - ACBV'\|_1 \overset{b}{=} \|RACVB - RACBV'\|_1 \overset{c}{=} \|RACVBC' - RACBV'C''\|_1 \overset{d}{=} \|RABC' - RABC''\|_1.$$
where \( a \) follows from (1) and the relation \( \|AC - AC'\| \leq \epsilon \), \( b \) follows because \( R - ABC - V \) and \( R - ABC - V' \), \( c \) and \( d \) follow from the monotonicity property of the distance between two distributions. The proof now follows by using the relation \( \|RBAC - RBAC'\| \leq \epsilon \).

**Theorem 2.** (Converse) In any one-way communication protocol for the task, the number of bits communicated is at least \( \text{Opt}^\epsilon(RABC) - \log \frac{1}{\epsilon} \).

**Proof.** In any one-way communication protocol for this task with a shared randomness \( S \), Alice produces a message \( V \) using \( ACS \), and communicates the message to Bob. Thus, communication cost is \( |V| \). Using the message \( V \), and random variables \( B, S \), Bob produces a random variable \( C' \). By correctness of the protocol, we have that \( \|RACB - RAC'B\|_1 \leq \epsilon \). Using Fact 4 we conclude that \( D_2^\epsilon(AC'V\|AC'\otimes V) \leq D_2^\epsilon(AC'SV\|AC'S \otimes V) \leq |V| + \log \frac{1}{\epsilon} \). Thus the item follows.

## 4 Optimality of Braverman-Rao protocol and equivalent formulations

Consider the following setting. There are two parties Alice and Bob: Alice possesses the random variable \((XM)\) and Bob possesses the random variable \(Y\). Furthermore let us assume that \(Y = X - M\). Alice wants to communicate the random variable \(M\) to Bob. To accomplish this task Alice sends a message to Bob. Bob on receiving these bits constructs a random variable \(M'\) such that \(\|XYM - XYM'\|_1\) is small.

Let \(\mathcal{P}\) be any shared randomness assisted communication protocol in which Alice and Bob work on input \(XY\), and Bob outputs a distribution \(M'\) correlated with \(XY\). Let \(\mathcal{P}(XY) := XYM'\) represent the resulting output distribution. We define \(\text{err}(\mathcal{P}) := \|\mathcal{P}(XY) - XYM\|_1\) as the error incurred by the protocol and \(\text{C}(\mathcal{P})\) as the communication cost of the protocol.

Our main result is the following theorem, which gives the near optimality of the protocol given due to Braverman and Rao [10]. In addition, we also give a new protocol using extensions of the random variable \(XYM\). In this direction, we define the two main quantities below: \(\text{BR}^\epsilon\) (which captures the worst case communication cost of the Braverman-Rao protocol, and we refer to it as the Braverman-Rao quantity) and \(\text{Ext}^\epsilon\) (which captures the worst case communication cost of the protocol that uses extensions of the input).

**Theorem 3.** Consider joint random variables \((X, Y, M)\) such that \(Y = X - M\) holds. Define the following three quantities associated to these random variables.

- \(\text{Opt}\epsilon := \min_{\mathcal{P}:\text{err}(\mathcal{P}) \leq \epsilon} \text{C}(\mathcal{P})\).
- \(\text{Opt}_1^\epsilon := \min_{XYS:M'} D_{\infty}(XS\bigcup XS \bigcup U), \text{such that } S \text{ is independent of } XY, Y - XS - V \text{ and } X - YVS - M' \text{ and } \|XYM' - XYM\|_1 \leq \epsilon\). Here \(U\) is the uniformly distributed random variable.
- \(\text{BR}^\epsilon := \min_{XY'M':Y' - X' - M'} D_{\epsilon}(X'M'|X'\otimes N_y)\max_y \in \text{supp}(Y') \inf_{N_y} D^\epsilon(S(X'M'|X'\otimes N_y))\).
- \(\text{Ext}^\epsilon := \min_{XY'M'E:Y' - X' - M' E \text{ and } \|XY'M'E - XYM\|_1 \leq \epsilon} (D^\epsilon(X'M'E||X'\otimes U) - D^\epsilon(Y'M'E||Y'\otimes U))\), where \(U\) is random variable distributed according to uniform distribution.

It holds that

1. \(\text{Opt}^\epsilon \geq \text{Opt}_1^\epsilon\).
2. \(\text{Opt}_1^\epsilon \geq \text{BR}^3\sqrt{\epsilon} - \log(\frac{1}{\epsilon})\).
3. \(\text{BR}^\epsilon + \log(\frac{1}{\epsilon}) \geq \text{Opt}^6\epsilon\).
4. \(\text{Ext}^\epsilon + 2\log(\frac{1}{\epsilon}) + 2\log \log(1/\epsilon) \geq \text{Opt}^{4\sqrt{\epsilon}}\).
5. \(\text{BR}^\epsilon > \text{Ext}^\epsilon\).

**Proof.** We will prove the results in sequence.
1. In any one-way communication protocol with a shared randomness \( S \), Alice produces a message \( V \) using \( XS \), and communicates the message to Bob. Thus, communication cost is \( |V| \). Using the message \( V \), shared randomness \( S \) and his input \( Y \), Bob produces a register \( M' \). By correctness of the protocol, we have that \( \|XYM' - XYM\|_1 \leq \varepsilon \). Since \( D_{\infty}(XS^V\|XS \otimes U) \leq |V| \), this item follows.

2. Given the distribution \( XYM' \), we recall that \( \|XYM' - XYM\|_1 \leq \varepsilon \). This can be re-written as \( \sum_y \Pr_Y(y)\|XM'_y - XM_y\|_1 \leq \varepsilon \). By Markov inequality, probability that \( \|XM'_y - XM_y\|_1 > \sqrt{\varepsilon} \) is less than \( \sqrt{\varepsilon} \). Let all such \( y \) be in the set \( \text{Bad} \) and rest of \( y \) be in the set \( \text{Good} \). Then for all \( y \in \text{Good} \) we have that \( \mathbb{E}_{X \leftarrow X} \|M'_y - M\|_1 \leq \sqrt{\varepsilon} \). Now, we construct a new random variable \( X_1 Y_1 M_1 \) as follows. We set \( \Pr_{X_1 Y_1 M_1}(x, y, m) = \frac{\Pr_Y(y)\Pr_{X M}(xm|y)}{\sum_{y \notin \text{Bad}} \Pr_Y(y)\Pr_{X M}(xm|y)} \) if \( y \) is not in \( \text{Bad} \). Else we set it 0. Then we construct the random variable \( X' YM'' \) as \( \Pr_{X' YM''}(x, y, m) = \Pr_{X_1 Y_1}(x, y) \cdot \Pr_M(m|x) \). Following properties are satisfied by these random variables.

   (a) \( \text{supp}(Y_1) = \text{supp}(Y') = \text{Good} \).
   (b) \( X_1 Y_1 \sim X' Y' \) and for all \( y \in \text{Good} \), \( X'_y \sim (X_1)_y \sim X_y \).
   (c) It holds that \( Y' - X' - M'' \).
   (d) For all \( y \in \text{Good} \),
   \[
   \|X'M''_y - XM'_y\|_1 = \mathbb{E}_{x \leftarrow X} \|X'M''_y - XM'_y\|_1 = \mathbb{E}_{x \leftarrow X} \|M_x - M'_xy\|_1 \leq \sqrt{\varepsilon}.
   \]
   (e) Using the fact that \( \|X_1 Y_1 M_1 - XYM\|_1 \leq 2\Pr_Y(\text{Bad}) \leq 2\sqrt{\varepsilon} \),
   \[
   \|X'YM'' - XYM\|_1 \leq \|X'YM'' - X_1 Y_1 M_1\|_1 + \|X_1 Y_1 M_1 - XYM\|_1 \leq \max_{y \in \text{Good}} \mathbb{E}_{x \leftarrow X} \|M''_y - (M_1)_y\|_1 + 2\sqrt{\varepsilon} = \max_{y \in \text{Good}} \mathbb{E}_{x \leftarrow X} \|M_x - M'_xy\|_1 + 2\sqrt{\varepsilon} \leq 3\sqrt{\varepsilon}.
   \]

Now, we proceed as follows.

\[
D_{\infty}(XS^V\|X \otimes S \otimes \mu) \overset{b}{=} D_{\infty}(YXS^V\|YX \otimes S \otimes U) = \max_{y \in \text{supp}(Y)} D_{\infty}(XS^V_y\|X_y \otimes S_y \otimes U) \overset{c}{=} \max_{y \in \text{supp}(Y)} \min_{S'V'} D_{\infty}(XS^V_y\|X_y \otimes S'V') \overset{d}{=} \max_{y \in \text{supp}(Y)} \min_{N_y} D_{\infty}(XM'_y\|X_y \otimes N_y) \overset{e}{=} \max_{y \in \text{supp}(Y_1)} \min_{N_y} D_{\infty}(X'M''_y\|X'_y \otimes N_y) \overset{f}{=} \max_{y \in \text{supp}(Y_1)} \min_{N_y} D_{\infty}(X'M''_y\|X'_y \otimes N_y) \overset{g}{=} \max_{y \in \text{supp}(Y')} \min_{N_y} D_2^{\sqrt{\varepsilon}}(X'M''_y\|X'_y \otimes N_y) - \log\left(\frac{4}{\varepsilon}\right).
\]

Above, (b) follows from the fact that \( Y - X - SV \), (c) follows by minimizing over all random variables \( S'V' \), (d) follows from the fact that conditioned on \( y \), Bob produces the random variable \( M' \) independent of \( X \), (e) follows from the fact that for all \( y \in \text{supp}(Y_1) \), we have that \( \|X'M''_y - XM'_y\|_1 \leq \sqrt{\varepsilon} \), (f) follows from Property 2 listed above and (g) follows from Fact 5.

Now, recall from Property 3 above that \( Y' - X' - M'' \). Since \( \|X'YM'' - XYM\|_1 \leq 3\sqrt{\varepsilon} \) (Property 5), we find that

\[
\max_{y \in \text{supp}(Y')} \min_{N_y} D_2^{\sqrt{\varepsilon}}(X'M''_y\|X'_y \otimes N_y) \geq \max_{y \in \text{supp}(Y')} \min_{N_y} D_2^{3\sqrt{\varepsilon}}(X'M''_y\|X'_y \otimes N_y) \geq \text{BR}^{3\sqrt{\varepsilon}}.
\]

This proves the item.
3. Consider a distribution \( X'Y'M' \), along with the distributions \( N_y \) for all \( y \in \text{supp}(Y) \), that achieves the minimum in \( \text{BR}^\varepsilon \). Recall that \( Y' - X' - M' \) holds, which implies that the expression for \( \text{BR}^\varepsilon \) is as follows.

\[
\text{BR}^\varepsilon = \max_{y \in \text{supp}(Y')} \inf(a_y : \Pr_x \Pr_{N_y}(m | x, m \in X'M'_y) > 2^{-a_y}) \leq \varepsilon).
\]

Alice and Bob start with the distribution \( X'Y' \), Alice has knowledge of the distribution \( M'_y \), Bob has the knowledge of \( N_y \) and both parties know the value of \( \text{BR}^\varepsilon \). Both run the protocol \( Q \) as defined in Corollary 1 with the value of \( c = \text{BR}^\varepsilon \). Let the random variable output by \( Q \) be \( X'Y'M'' \). By definition of \( \text{BR}^\varepsilon \), we have that for all \( y \in \text{supp}(Y') \), \( \Pr_{x \leftarrow X'_y} (D_y (M_x || N_y) > \text{BR}^\varepsilon) \leq \varepsilon \). Thus for all \( y \in \text{supp}(Y') \), with probability at least \( 1 - \varepsilon \) according to \( X'_y \), it holds that \( \|M''_y - M'_y\|_1 \leq \varepsilon \). This implies that \( \|X'Y'M'' - X'Y'M'\|_1 \leq 4\varepsilon \). In other words, \( \|Q(X'Y') - X'Y'M'\|_1 \leq 4\varepsilon \).

Since \( \|X'Y' - X'Y\|_1 \leq \|X'Y'M' - X'Y\|_1 \leq \varepsilon \) and \( Q \) is a protocol, it holds that

\[
\|Q(X'Y) - X'YM\|_1 \leq \|Q(X'Y) - Q(X'Y')\|_1 + \|Q(X'Y') - X'YM'\|_1 + \|X'Y'M' - X'YM\|_1 \leq \varepsilon + 4\varepsilon + \varepsilon.
\]

The communication cost is \( \text{BR}^\varepsilon + \log(\frac{3}{\varepsilon}) \). This completes the proof.

4. As shown in Lemma 2, for any distribution \( X'Y'M'E' \) such that \( Y' - X' - M'E' \), there exists a protocol \( P \) with \( \|P(X'Y'M') - X'YM'\|_1 \leq 3\sqrt{\varepsilon} \) and communication cost

\[
D_\varepsilon(X'M'E' || X' \otimes U) - D_\varepsilon(Y'M'E' || Y' \otimes U) + 2 \log \frac{1}{\varepsilon} + \log \frac{1}{\varepsilon}.
\]

Thus, choosing \( X'Y'M' \) such that \( \|X'Y'M'-X'YM\|_1 \leq \varepsilon \), we have that \( \|P(X'YM) - X'YM\|_1 \leq 4\sqrt{\varepsilon} \). This proves the item.

5. Let \( X'Y'M' \) and the distributions \( N_y \) be as obtained from the definition of \( \text{BR}^\varepsilon \). From Theorem 4 in Section 4, it holds that there exists a random variable \( E \) such that \( X'Y'M'E \) satisfies \( Y' - X' - M'E \) and

\[
\max_y D_\varepsilon(X'M'_y || X'_y \otimes N_y) \geq D_\varepsilon(X'M'E || X' \otimes U) - D_\varepsilon(Y'M'E || Y' \otimes U).
\]

The item follows by observing that \( \text{Ext}^\varepsilon \) is obtained by minimizing right hand side over all \( X'Y'M'E \) and \( U \), such that \( Y' - X' - M'E \) and \( \|X'Y'M' - X'YM\|_1 \leq \varepsilon \).

\[ \Box \]

**A new protocol using extensions**

In this section, we construct a new communication protocol for sending a random variable from Alice to Bob. We consider the following scenario. Alice possesses the random variable \( (XME) \) (with \( M, E \) taking values on the sets \( M, E \) respectively) and Bob possesses the random variable \( Y \). Furthermore let us assume that \( Y - X - ME \). Alice wants to communicate the random variable \( ME \) to Bob. To accomplish this task Alice sends a message to Bob. Bob on receiving the message constructs a pair of random variables \( (M', E') \) such that \( \|XYME - XYM'E'\|_1 \) is small. The following lemma makes the above statement more precise.

**Lemma 2.** (Achievability) Let \( XYME \) be as defined above. Further, let \( U \) be a uniformly distributed over the set \( U := M \otimes E \) and be independent of \( XYME \). Let \( R \) and \( r \) be such that

Then there exists a shared randomness assisted communication protocol in which Bob outputs random variable pair \( (M', E') \) satisfying \( \|XYME - XYM'E'\|_1 \leq 6\varepsilon \).

In particular, the communication cost of the protocol is upper bounded by

\[
D_\varepsilon(XME || X \otimes U) - D_\varepsilon(YME || Y \otimes U) + 2 \log \frac{1}{\varepsilon}.
\]
Proof. Define $R, r$ as follows.

\[ R + r = D_f^\epsilon(XME\| X \otimes U) + \log \left( \frac{-\ln(\epsilon)}{(1 - \epsilon)} \right); \]

\[ r = D_H^\epsilon(YME\| Y \otimes U) + \log \left( \frac{\epsilon (1 - \epsilon)}{-\ln(\epsilon)} \right) \text{ (if } D_f^\epsilon(XME\| X \otimes U) > 1); \]

\[ r = 0 \quad \text{if } D_f^\epsilon(XME\| X \otimes U) \leq 1 \] (4)

Let $U$ be as in the lemma, let $Z$ be uniformly distributed over the set $[0, 1]$. Also, let $b : X \times U \times [0, 1] \to \{0, 1\}$ be defined by $b(x, m, e, z) = 1$ iff $z \leq \min(2^{-D_f^\epsilon(XME\| X \otimes U)} \frac{\Pr_{ME}(m, e|x)}{\Pr_U(m, e)}, 1)$. Let $B = b(X \otimes U \otimes Z)$ be the corresponding random variable. We now make the following claims which are easy to prove and will help us to prove the claim of the lemma:

Claim 1. $\Pr_B(1 \mid X \otimes U = (x, m, e)) = \min(2^{-D_f^\epsilon(XME\| X \otimes U) \frac{\Pr_{ME}(m, e|x)}{\Pr_U(m, e)}}, 1)$

Proof. It follows by direct calculation. \qed

Claim 2. It holds that

\[ 2^{-D_f^\epsilon(XME\| X \otimes U)} \geq \Pr_B(1|x) \geq 2^{-D_f^\epsilon(XME\| X \otimes U)} \Pr_{(m, e) \sim ME_x} \left( \frac{\Pr_{ME}(m, e|x)}{\Pr_U(m, e)} \right) < 2^{D_f^\epsilon(XME\| X \otimes U)}. \]

In particular,

\[ 2^{-D_f^\epsilon(XME\| X \otimes U)} \geq \Pr(1) \geq 2^{-D_f^\epsilon(XME\| X \otimes U)}(1 - \epsilon). \]

Proof. We proceed as follows.

\[ \Pr_B(1|x) = \sum_{m, e} \Pr_U(m, e) \Pr_B(1|x) X \otimes U = (x, m, e) \]

\[ = \sum_{m, e} \Pr_U(m, e) \min(2^{-D_f^\epsilon(XME\| X \otimes U)} \frac{\Pr_{ME}(m, e|x)}{\Pr_U(m, e)}, 1) \geq \]

\[ 2^{-D_f^\epsilon(XME\| X \otimes U)} \sum_{m, e} \Pr_{ME}(m, e|x) \text{IND}(2^{-D_f^\epsilon(XME\| X \otimes U)} \frac{\Pr_{ME}(m, e|x)}{\Pr_U(m, e)} < 1) \]

where IND is the indicator function. Thus,

\[ \Pr_B(1|x) \geq 2^{-D_f^\epsilon(XME\| X \otimes U)} \Pr_{(m, e) \sim ME_x} \left( \frac{\Pr_{ME}(m, e|x)}{\Pr_U(m, e)} \right) < 2^{D_f^\epsilon(XME\| X \otimes U)}. \]

This also leads to

\[ \Pr_B(1) \geq 2^{-D_f^\epsilon(XME\| X \otimes U)} \sum_x P_X(x) \Pr_{(m, e) \sim ME_x} \left( \frac{\Pr_{ME}(m, e|x)}{\Pr_U(m, e)} \right) < 2^{D_f^\epsilon(XME\| X \otimes U)} \]

\[ = 2^{-D_f^\epsilon(XME\| X \otimes U)} \Pr_{(x, m, e) \sim ME} \left( \frac{\Pr_{ME}(m, e|x)}{\Pr_U(m, e)} < 2^{D_f^\epsilon(XME\| X \otimes U)} \right) \geq 2^{-D_f^\epsilon(XME\| X \otimes U)}(1 - \epsilon). \]

For the upper bound, we proceed as

\[ \Pr_B(1|x) = \sum_{m, e} \Pr_U(m, e) \min(2^{-D_f^\epsilon(XME\| X \otimes U)} \frac{\Pr_{ME}(m, e|x)}{\Pr_U(m, e)}, 1) \]

\[ \leq \sum_{m, e} \Pr_U(m, e) \cdot 2^{-D_f^\epsilon(XME\| X \otimes U)} \frac{\Pr_{ME}(m, e|x)}{\Pr_U(m, e)} \]

\[ = 2^{-D_f^\epsilon(XME\| X \otimes U)} \sum_{m, e} \Pr_U(m, e) = 2^{-D_f^\epsilon(XME\| X \otimes U)} \]

10
Claim 3. \( \Pr_{X \otimes U}(x, m, e \mid B = 1) \leq \Pr_{XME}(x, m, e)/(1 - \varepsilon) \). In particular, let \( YXU_1 \) be the random variable \( YX \otimes U \) conditioned on the event \( B = 1 \). Then

\[
\|YXU_1 - YXME\|_1 \leq \varepsilon.
\]

\textbf{Proof.} Using Bayes’ theorem and Claims 1 and 2, \( \Pr_{X \otimes U}(x, m, e \mid B = 1) \) can be written as

\[
\frac{\Pr_B(1 | X \otimes U = (x, m, e)) \cdot \Pr_{X \otimes U}(x, m, e)}{\Pr_B(1)} \leq \min(\Pr_{XME}(x, m, e), 2^{D_Y(YXME \| X \otimes U)} \cdot \Pr_U(m, e) \Pr_X(x)) .
\]

First part of the claim follows since \( \min(a, b) \leq a \).

For second part, we observe that \( Y = X - U_1 \) holds, as the event \( B = 1 \) is defined conditioned on \( x \). Also recall that \( Y - X \sim ME \). Thus, \( \|YXU_1 - YXME\|_1 \leq \|XU_1 - XME\|_1 \). Second part follows as first part of the claim implies that \( \|XU_1 - XME\|_1 \leq \varepsilon \). □

Suppose Alice and Bob share \( T = \{(U(1), Z(1)), (U(2), Z(2)), \ldots, (U(2^{R+r}), Z(2^{R+r}))\} \) of \( 2^{R+r} \) i.i.d. pairs each with distribution \( (U, Z) \). Further, let us divide these \( 2^{R+r} \) shared copies of \( (U, Z) \) into \( 2^R \) bands each having \( 2^r \) copies of \( (U, Z) \). This subdivision of \( T \) is known to both Alice and Bob. Further, for \( i \in [1: 2^{(R+r)}] \) let \( B(i) \) be the analogous random variable (as defined above).

We now discuss the protocol which ensures the statement of the lemma. Alice on receiving \( (x, m, e) \leftarrow XME \) tries to find the first \( i^* \) such that \( B(i) = 1 \). Alice then conveys the band number \( \text{Band}(i^*) \) of this index to Bob. Bob on receiving the band number \( \text{Band}(i^*) \) tries to infer the correct \( i^* \). We will now show that Bob is indeed able to infer the correct \( i^* \) with high probability. We will then further show that our protocol will accomplish the task as claimed in the lemma. Towards this let \( \mathcal{A} \) be such that

\[
\Pr_{YME}(\mathcal{A}) \geq 1 - \varepsilon
\]

and

\[
D_H(YME \| Y \otimes U) = -\log \Pr_{Y \otimes U}(\mathcal{A}) .
\]

The decoding strategy is simply that Bob on receiving \( Y \) and the band number \( \text{Band}(i^*) \), tries to find a unique \( \hat{i} \in \text{Band}(i^*) \) such that \( (U(\hat{i}), Y) \in \mathcal{A} \). To find the error probability with respect to this decoding strategy by Bob let us define the following events

\[
\mathcal{E}_1 = \{ \forall i \in [1: 2^{R+r}] : B(i) = 0 \} ; \\
\mathcal{E}_2 = \{ (U(i^*), Y) \notin \mathcal{A} \} ; \\
\mathcal{E}_3 = \{ (U(i'), Y) \in \mathcal{A} \text{ for some } i' \neq i^* \} .
\]

Let \( \hat{i} \) be the decoded index and let \( \hat{\mathcal{E}} \) be the event that \( \hat{i} \neq i^* \). We now bound the overall error probability of the protocol as follows:

\[
\Pr\{ \hat{\mathcal{E}} \} \leq \Pr(\mathcal{E}_1) + \Pr(\mathcal{E}_2 \cap \mathcal{E}_1) + \Pr(\mathcal{E}_3 \cap \mathcal{E}_1),
\]

where the probability is computed over all the randomness in the protocol.
We now calculate each of the terms on the R.H.S of (5). Our strategy is to bound each of the error conditioned on $x$ and then average over $x$. For the first error, we proceed as follows.

$$\Pr\{E_1\} = \sum_x \Pr(x) \Pr[B(i) = 0, \forall i \in [1 : 2^{R+r}] | x]$$

$$= \sum_x \Pr(x) \left(1 - \Pr_B(1|x)\right)^a \leq \sum_x \Pr(x) e^{-\Pr_B(1|x)2^{R+r}}$$

$$\leq b \sum_x \Pr(x) (1 - \Pr_{(m,e) \sim ME_x} \frac{\Pr_{ME}(m,e|x)}{\Pr_U(m,e)} < 2^{D'_s(XME||X \otimes U)}) + e^{-2^{R+r}2^{-D'_s(XME||X \otimes U)}}$$

$$\leq 1 - \Pr_{XME}(\frac{\Pr_{ME}(m,e|x)}{\Pr_U(m,e)} < 2^{D'_s(XME||X \otimes U)}) + 2\varepsilon \leq 3\varepsilon.$$ (8)

where (a) and (b) follow from Fact 3 and last inequality follows from our choice of $R + r$ and definition of $D'_s(XME||X \otimes U)$.

Consider $\Pr\{E_1^c \cap E_2\}$:

$$\Pr\{E_1^c \cap E_2\} \leq \Pr\{E_2 | E_1^c\}$$

$$= \sum_{x \in X} \Pr(x) \sum_{(m,e) \notin A} \Pr_{U(i^*)} \{m, e \mid B(i^*) = 1, x\} \Pr_Y(y \mid x, m, e, B(i^*) = 1)$$

$$\leq \sum_{x \in X} \sum_{(m,e) \notin A} \Pr_{XME}(x,m,e) \frac{\varepsilon}{1 - \varepsilon} \Pr_Y(y \mid x)$$

$$\leq \sum_{(m,e) \notin A} \Pr_{MEY}(m,e,y) \frac{\varepsilon}{1 - \varepsilon} \leq \frac{\varepsilon}{(1 - \varepsilon)}.$$ (9)

where $a$ follows from Claim 5 and the fact that conditioned on $x, Y_x$ is independent of $U(i^*)$, $b$ follows because $Y - X - ME$ and $c$ follows from the definitions of the set $A$.

Consider $\Pr\{E_1^c \cap E_3\}$. If $D'_s(XME||X \otimes U) \leq 1$, then this event is empty, as the size of each band is 1. So we assume that $D'_s(XME||X \otimes U) > 1$ and proceed as follows.

Fix an $x$ and let $U_x(i')$ (for $i' < i^*$) be the distribution over $U$ conditioned on $B(i') = 0$. From Fact 2 and the fact that $U(i')$ and $X$ are independent of each other before the event $B(i') = 0$, we have

$$\Pr_{U_x(i')} (m,e) \leq \Pr_{U(i')}(m,e) \frac{\Pr_U(0|x)}{\Pr_B(0|x)} \leq \Pr_{U(i')}(m,e) 1 - 2^{-D'_s(XME||X \otimes U)},$$

where last inequality follows from Claim 2.

On the other hand, for $i' > i^*$, the distribution on $U(i')$ remains unchanged. Thus, we can use union bound, and upper bound $\Pr\{E_1^c \cap E_3\}$ as follows.

$$\Pr\{E_1^c \cap E_3\} \leq \frac{2^r}{1 - 2^{-D'_s(XME||X \otimes U)}} \sum_x \Pr(x) \sum_{y,m,e \in A} \Pr_U(m,e) \cdot \Pr_Y(y|x)$$

$$= \frac{2^r}{1 - 2^{-D'_s(XME||X \otimes U)}} \sum_{y,m,e \in A} \Pr_U(m,e) \cdot \sum_x \Pr(x) \Pr_Y(y|x)$$

$$= \frac{2^r}{1 - 2^{-D'_s(XME||X \otimes U)}} \sum_{y,m,e \in A} \Pr_U(m,e) \Pr_Y(y)$$

$$\leq \frac{2^r}{1 - 2^{-D'_s(XME||X \otimes U)}} \cdot 2^{-D'_s(YME||Y \otimes U)}$$

12
In the first inequality, we have divided the expression by $1 - 2^{-D_r^*(X^ME\|X \otimes U)}$ to include the upper bound that may hold for the distribution $\hat{U}_i(i^*)$.

By our assumption, $D_r^*(X^ME\|X \otimes U) > 1$ implies $2^{-D_r^*(X^ME\|X \otimes U)} < 1/2$. Thus, we find that $\Pr\{E_1 \cap E_2\} \leq 2 \cdot 2^{-D_r^*(Y^ME\|Y \otimes U) - \log(1 - \varepsilon)} \leq 2\varepsilon$ for the choice of $r$. Collectively, we find that

$$\Pr\{\hat{E}\} \leq \Pr\{E_1\} + \Pr\{E_1^c \cap E_2\} + \Pr\{E_1^c \cap E_3\} \leq 3\varepsilon + 3\varepsilon \leq 6\varepsilon.$$

To conclude the proof, we need to upper bound $\|XYM'E' - XYME\|_1$. We note that conditioned on the event that Bob outputs from index $i^*$, Claim 3 ensures that $\|XYU_{B(i^*)} - XYME\|_1 \leq \varepsilon$. Since probability that the index output is not $i^*$ is equal to $\Pr\{\hat{E}\} \leq 5\varepsilon$, we conclude that the final output $XYM'E'$ satisfies $\|XYM'E' - XYME\|_1 \leq 6\varepsilon$. The communication cost of the protocol is equal to $R$. In both the cases of $D_r^*(X^ME\|X \otimes U) \leq 1$ and $D_r^*(X^ME\|X \otimes U) > 1$, the upper bound

$$R \leq D_r^*(X^ME\|X \otimes U) - D_r^H(Y^ME\|Y \otimes U) + 2 \log \log \frac{1}{\varepsilon} + \log \frac{1}{\varepsilon}$$

is easily seen to hold. This completes the proof.

\[\Box\]

**Relating Braverman-Rao quantity and extension quantity**

Recall the quantity $BR^*$ as defined in Theorem 3. Let $X'Y'M'$ be the optimal Markov chain and $N_y$ (for all $y \in \text{supp}(Y')$) be the set of distributions from the definition. We will define an extension $X'Y'M'E'$ such that $Y' - X' - M'E$ holds. Towards this let $P = \{(x, m) : P_{M'}(m | x) > 0\}$. Let $K$ be the smallest integer such that $K \cdot P_{M'}(m|x)$ is an integer. This can be assumed to hold with arbitrarily small error.

Further, let $E$ be a random variable taking values over the set $K := \{1, \cdots, K\}$ and defined as follows. For every $(m, e, x) \in \mathcal{M} \otimes K \otimes X$ we have

$$\Pr\{M'X'E \mid e, m, x\} = \begin{cases} \frac{\Pr_{X'}(x)}{K} & \text{if } e \leq K \Pr_{M'(m \mid x)}, \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

It is easy to see that the property $Y' - X' - M'E$ holds. Let $U$ be a uniform random variable distributed over the set $\mathcal{M} \otimes K$. We now have the following theorem.

**Theorem 4.** Consider $Y' - X' - M'$, $\{N_y\}_y$ and $(E, U)$ be as defined above. Then,

$$\max_y D_r^*(X'M'_y\|X'_y \otimes N_y) \geq D_r^*(X'M'E\|X' \otimes U) - D_r^H(Y'M'E\|Y \otimes U).$$

**Proof.** Let us first calculate $D_\infty(X'M'E\|X' \otimes U)$. Towards this notice the following set of inequalities

$$D_\infty(X'M'E\|X' \otimes U) = \max_{m, x, e} \log \frac{\Pr_{X'M'E}(x, m, e)}{\Pr_{X'}(x) \Pr_U(u)} = b \log \frac{|M|K}{\Pr_U(u)} + b \log |M|, \quad (11)$$

where $a$ follows from the definition of $D_\infty(X'M'E\|X' \otimes U)$; $b$ follows from (10) and the fact that $U$ is uniform over the set $\mathcal{M} \otimes K$. Let us define the following set

$$\mathcal{A} := \left\{(y, m, e) \in \mathcal{Y} \otimes \mathcal{M} \otimes K : e \leq K^{2\max_y D_r^*(X'M'_y\|X'_y \otimes N_y)} \Pr(m)\right\}. \quad (12)$$
We will prove the following

\[ \Pr_{Y'U} \{ A \} = 2^{-\left(\log |M| - \max_y D^*_M(X'M'_y \| X'_y \otimes N_y)\right)}, \]  

(13)

\[ \Pr_{M'Y'E} \{ A \} \geq 1 - \varepsilon. \]

(14)

The theorem now follows since by definition of \( D^*_H(Y'M'E \| Y' \otimes U) \) and equation 11

\[ D^*_H(Y'M'E \| Y' \otimes U) \geq \log |M| - \max_y D^*_M(X'M'_y \| X'_y \otimes M'_y) \]

\[ = D_s(X'M'E \| X' \otimes U) - \max_y D^*_M(X'M'_y \| X'_y \otimes N_y). \]

Let us first prove (13). Towards this notice the following

\[ \Pr_{Y'U} \{ A \} = \sum_{(y,m,e) \in A} \Pr(y) \Pr(m, e) \]

\[ = \sum_{y \in Y'} \Pr(y) \sum_{(m,e) | (y,m,e) \in A} \frac{1}{|M|K} \]

\[ = \sum_{y \in Y'} \Pr(y) \sum_{(m,e) | (y,m,e) \in A} \frac{K2^{\max_y D^*_M(X'M'_y \| X'_y \otimes N_y)}}{|M|K} \]

\[ = \sum_{y \in Y'} \Pr(y) \sum_{(m,e) | (y,m,e) \in A} \frac{2^{\max_y D^*_M(X'M'_y \| X'_y \otimes N_y)}}{|M|} \]

\[ = 2^{-\left(\log |M| - \max_y D^*_M(X'M'_y \| X'_y \otimes N_y)\right)}. \]

This completes the proof for (13). Let us now prove the claim in (14). Towards this we have the following set of inequalities:

\[ \Pr_{Y'M'E} \{ A \} = \sum_x \Pr(x) \sum_{(y,m,e) \in A} \Pr(y \mid x) \Pr_{M'E} (me \mid x) \]

\[ = \sum_x \Pr(x) \sum_{y} \Pr(y \mid x) \sum_{m : e \leq K} \Pr_{M'E}(m \mid x) \sum_{(y,m,e) \in A} \frac{1}{K} \]

\[ \geq \sum_{(x,y)} \Pr_{X'Y'}(x,y) \sum_{m : Pr_{M'E}(m \mid x) \leq 2^{\max_y D^*_M(X'M'_y \| X'_y \otimes N_y)}} \Pr_{M'E}(m \mid x) \]

\[ = \Pr_{Y'X'M'} \left\{ \Pr_{M'E}(m \mid x) \leq 2^{\max_y D^*_M(X'M'_y \| X'_y \otimes N_y)} \Pr_{N_y}(m) \right\} \]

\[ \geq 1 - \varepsilon, \]

where \( a \) follows from (10), \( b \) follows if we restrict to \( m \) such that

\[ \Pr_{M'E}(m \mid x) \leq 2^{\max_y D^*_M(X'M'_y \| X'_y \otimes N_y)} \Pr_{N_y}(m), \]

as for such \( m \) and \( e \) such that \( e \leq K \Pr_{M'E}(m \mid x), (y, m, e) \) automatically belongs to the set \( A \). The step \( c \) follows from the definition of \( \max_y D^*_M(X'M'_y \| X'_y \otimes N_y) \). This completes the proof.
5 An achievability bound on quantum state redistribution

Following lemma was shown in [9], which the authors refer to as convex-split lemma.

**Lemma 3 (Convex-split lemma).** [9] Let $\rho_{PQ} \in \mathcal{D}(PQ)$ and $\sigma_Q \in \mathcal{D}(Q)$ be quantum states such that $\text{supp}(\rho_{PQ}) \subset \text{supp}(\sigma_Q)$. Let $k \stackrel{\text{def}}{=} \max \{ \rho_{PQ} \| \rho_P \otimes \sigma_Q \}$. Define the following state

$$\tau_{PQ,Q_2\ldots Q_n} \stackrel{\text{def}}{=} \frac{1}{n}\sum_{j=1}^{n} \rho_{PQ_j} \otimes \sigma_{Q_1} \otimes \sigma_{Q_2} \ldots \otimes \sigma_{Q_{j-1}} \otimes \sigma_{Q_{j+1}} \ldots \otimes \sigma_{Q_n}$$

(15)

on $n+1$ registers $P,Q_1,Q_2,\ldots,Q_n$, where $\forall j \in [n] : \rho_{PQ_j} = \rho_{PQ}$ and $\sigma_{Q_j} = \sigma_Q$. Then,

$$D(\tau_{PQ_1Q_2\ldots Q_n} \| \tau_P \otimes \sigma_{Q_1} \otimes \sigma_{Q_2} \ldots \otimes \sigma_{Q_n}) \leq \log(1 + \frac{2^k}{n})$$

and

$$F^2(\tau_{PQ_1Q_2\ldots Q_n}, \tau_P \otimes \sigma_{Q_1} \otimes \sigma_{Q_2} \ldots \otimes \sigma_{Q_n}) \geq \frac{1}{1 + \frac{2^k}{n}}.$$  

In particular, for $\delta \in (0,1/3)$ and $n = \lceil \frac{2^k}{\delta} \rceil$,

$$D(\tau_{PQ_1Q_2\ldots Q_n} \| \tau_P \otimes \sigma_{Q_1} \otimes \sigma_{Q_2} \ldots \otimes \sigma_{Q_n}) \leq \log(1 + \delta)$$

and

$$F^2(\tau_{PQ_1Q_2\ldots Q_n}, \tau_P \otimes \sigma_{Q_1} \otimes \sigma_{Q_2} \ldots \otimes \sigma_{Q_n}) \geq 1 - \delta.$$  

Using this lemma, the authors showed the following result, which says that given a quantum state $\Phi_{RA'MB'}$ shared between Alice (registers $A'$), Bob (register $B'$) and Referee (register $R$), the message $M$ can be sent from Alice to Bob with communication cost close to $I_{\max} RB'M\Phi$.

**Theorem 5 (9).** There exists an entanglement-assisted one-way protocol $\mathcal{P}$, which takes as input $|\Phi\rangle_{RA'MB'}$ shared between three parties Referee ($R$), Bob ($B'$) and Alice ($A'$) and outputs a state $\Phi'_{RA'MB'} \in \mathcal{B}C(\Psi_{RA'MB'})$ and the number of qubits communicated by Alice to Bob in $\mathcal{P}$ is upper bounded by:

$$\frac{1}{2} I_{\max} (RB' : M)_{\Phi} + \log \left( \frac{1}{\epsilon} \right).$$

In this section, we improve upon this theorem, showing that further compression is possible. We use this to derive a new achievability bound on the task of quantum state redistribution, to be discussed later in this section.

**New compression result**

We prove the following theorem.

**Theorem 6 (Quantum compression).** There exists an entanglement-assisted one-way protocol $\mathcal{P}$, which takes as input $|\Phi\rangle_{RA'MB'}$ shared between three parties Referee ($R$), Bob ($B'$) and Alice ($A'$) and outputs a state $\Phi'_{RA'MB'}$ shared between Referee ($R$), Bob ($B'M$) and Alice ($A'$) such that $\Phi'_{RA'MB'} \in \mathcal{B}C(\Psi_{RA'MB'})$ and the number of qubits communicated by Alice to Bob in $\mathcal{P}$ is upper bounded by:

$$\frac{1}{2} \inf_{\sigma_M} (D_{\max}(\Phi_{RB'M:MB'} \| \Phi_{RB'} \otimes \sigma_M) - D_{\mathbb{H}}(\Phi_{B'M:MB'} \| \Phi_{B'} \otimes \sigma_M)) + \log \left( \frac{1}{\epsilon} \right).$$

15
Proof. The proof is similar in spirit to the proof of Theorem 5 except that the decoding operation at Bob’s side is now made more non-trivial.

Let $\sigma_M$ be an arbitrary state in $D(M)$. Let $k \overset{\text{def}}{=} D_{\max}(\Phi_{RB}'M_1 \otimes \Phi_{RB}' \otimes \sigma_M), \delta \overset{\text{def}}{=} \varepsilon^2$ and $n \overset{\text{def}}{=} \lceil \frac{2k}{\delta} \rceil$.

Let $b \overset{\text{def}}{=} \lceil \varepsilon^2 \cdot 2^{D_H(\Phi_{B}'M_1 \otimes \Phi_{B}' \otimes \sigma_M)} \rceil$. By definition of $D_H(\Phi_{B}'M_1 \otimes \Phi_{B}' \otimes \sigma_M)$, there exists a projector $\Pi_{B'M_1}$ such that $\text{Tr}(\Pi_{B'M_1} \Phi_{B'} \otimes \sigma_M) \geq 1 - \varepsilon^2$ and $\text{Tr}(\Pi_{B'M_1} \Phi_{B'} \otimes \sigma_M) \leq \varepsilon^2/b$.

Consider the state,

$$\mu_{RB'M_1...M_n} = \frac{1}{n} \sum_{j=1}^{n} \Phi_{RB'M_j} \otimes \sigma_{M_1} \otimes \cdots \otimes \sigma_{M_{j-1}} \otimes \sigma_{M_{j+1}} \otimes \cdots \otimes \sigma_{M_n}.$$ 

Note that $\Phi_{RB'} = \mu_{RB'}$. Consider the following purification of $\mu_{RB'M_1...M_n}$,

$$|\mu\rangle_{RB'JL_1...L_nM_1...M_n}$$

$$= \frac{1}{\sqrt{n}} \sum_{j=1}^{n} |j\rangle |\tilde{\Phi}_{RB'A'M_j} \otimes |\sigma\rangle_{L_1M_1} \otimes \cdots \otimes |\sigma\rangle_{L_{j-1}M_{j-1} \otimes |0\rangle_{L_j} \otimes |\sigma\rangle_{L_{j+1}M_{j+1} \otimes \cdots \otimes |\sigma\rangle_{L_nM_n}}$$

Here, $\forall j \in [n] : |\sigma\rangle_{L_jM_j}$ is a purification of $\sigma_{M_j}$ and $|\tilde{\Phi}_{RB'L_jM_j}\rangle$ is a purification of $\Phi_{RB'M_j}$. Consider the following protocol $\mathcal{P}_1$.

1. Alice, Bob and Referee start by sharing the state $|\mu\rangle_{RB'JL_1...L_nM_1...M_n}$ between themselves where Alice holds registers $JL_1...L_n$, Referee holds the register $R$ and Bob holds the registers $B'M_1M_2...M_n$.

2. Alice measures the register $J$ and obtains the measurement outcome $j \in [n]$. She sends the integer $[(j - 1)/b]$ to Bob using $\log(n/b)$ qubits of quantum communication. Alice and Bob employ superdense coding (23) using fresh entanglement to achieve this.

3. Bob swaps registers $M_j |(j-1)/b + 1\rangle, M_j |(j-1)/b + 2\rangle, \ldots, M_j |(j-1)/b + b\rangle$ with the set of registers $M_1, M_2, \ldots, M_b$ in that order.

   - At this step of the protocol, the joint state in the registers $RB'A'M_1M_2...M_b$ is

   $$\mu^2_{RB'A'M_1M_2...M_b} = \frac{1}{b} \sum_{j=1}^{b} |\Phi\rangle_{|RB'A'M_j} \otimes |\sigma\rangle_{M_1} \otimes \cdots \otimes |\sigma\rangle_{M_{j-1}} \otimes |\sigma\rangle_{M_{j+1}} \otimes \cdots \otimes |\sigma\rangle_{M_n}.$$ 

4. $\Pi_j \overset{\text{def}}{=} \Pi_{B'M_j} \otimes I_{M_1} \otimes \cdots \otimes I_{M_{j-1}} \otimes I_{M_{j+1}} \otimes \cdots \otimes I_{M_n}$ and $\Pi \overset{\text{def}}{=} \sum_j \Pi_j$. Bob applies the measurement

   $$A_{good}(X) = \sum_j \sqrt{\Pi_j - \frac{1}{2} } \Pi_j - \frac{1}{2} X \sqrt{\Pi_j - \frac{1}{2} } \Pi_j - \frac{1}{2} \otimes |j\rangle \langle j|_O,$$

   where $O$ is the outcome register, and upon obtaining the outcome $j$, swaps $M_j, M_1$.

5. Final state is obtained in the registers $RA'B'M_1$. We call it $\Phi_{RB'A'M_1}$

We have the following claim.

Claim 4. It holds that $F^2(\Phi_{RB'A'M_1}^3, \Phi_{RB'A'M_1}^3) \geq 1 - 18\varepsilon^2$.

Proof. Applying the measurement $A_{good}$ to the state $\mu_{RB'A'M_1M_2...M_b}$, we obtain the following state (we will ignore the register labels on some of the states for convenience)

$$A_{good}(\mu^2) = \frac{1}{b} \sum_{i,j} \sqrt{\Pi_j - \frac{1}{2} } \Pi_j - \frac{1}{2} |\Phi\rangle_{|RB'A'M_j} \otimes |\sigma\rangle_{M_1} \otimes \cdots \otimes |\sigma\rangle_{M_{j-1}} \otimes |\sigma\rangle_{M_{j+1}} \otimes \cdots \otimes |\sigma\rangle_{M_n} \rangle \langle \Pi_j - \frac{1}{2} } \Pi_j - \frac{1}{2} \otimes |i\rangle \langle i|_O.$$ 

Let $p_{i,j}$ be defined as follows:

$$p_{i,j} \overset{\text{def}}{=} \text{Tr}(\Pi_j - \frac{1}{2} ) \Pi_j - \frac{1}{2} |\Phi\rangle_{|RB'A'M_j} \otimes |\sigma\rangle_{M_1} \otimes \cdots \otimes |\sigma\rangle_{M_{j-1}} \otimes |\sigma\rangle_{M_{j+1}} \otimes \cdots \otimes |\sigma\rangle_{M_n} \rangle \langle i|_O.$$
The overall probability of error is
\[
\frac{1}{b} \sum_{i \neq j} p_{i,j} = \frac{1}{b} \sum_j \text{Tr}(\Pi_j^{1/2} (\sum_{i \neq j} \Pi_i \Pi_j^{1/2}) \Phi_{B'M_j} \otimes \sigma_{M_i} \otimes \ldots \sigma_{M_{j-1}} \otimes \sigma_{M_{j+1}} \otimes \ldots \sigma_{M_b}).
\]

Now using Hayashi-Nagaoka inequality (Fact \ref{fact:Nagaoka}), we obtain
\[
\frac{1}{b} \sum_{i \neq j} p_{i,j} \leq \frac{1}{b} \sum_j 4 \text{Tr}((I - \Pi_j) \Phi_{B'M_j} \otimes \sigma_{M_i} \otimes \ldots \sigma_{M_{j-1}} \otimes \sigma_{M_{j+1}} \otimes \ldots \sigma_{M_b}) + 4 \sum_j \text{Tr}((\sum_{i \neq j} \Pi_i) \Phi_{B'} \otimes \sigma_{M_i}) \leq 2\epsilon^2 + 4(b-1)\frac{\epsilon^2}{b} \leq 6\epsilon^2.
\]

This implies that \( \frac{1}{b} \sum_i p_{i,i} = \frac{1}{b} \sum_{i,j} p_{i,j} \geq 1 - 6\epsilon^2. \)

Define the states
\[
\mu^{3} \overset{\text{def}}{=} \frac{1}{\sum_{i} p_{i,i}} \sum_{i} \sqrt{\Pi^{1/2} \Pi_i \Pi^{1/2}} |\Phi\rangle \langle \Phi|_{RB'A'M_i} \otimes \sigma_{M_i} \otimes \ldots \sigma_{M_{i-1}} \otimes \sigma_{M_{i+1}} \otimes \ldots \sigma_{M_b}) \sqrt{\Pi^{1/2} \Pi_i \Pi^{1/2}} |i\rangle \langle i|_O
\]
and
\[
\mu^{4} \overset{\text{def}}{=} \frac{1}{\sum_{i} p_{i,i}} \sum_{i} |\Phi\rangle \langle \Phi|_{RB'A'M_i} \otimes \sigma_{M_i} \otimes \ldots \sigma_{M_{i-1}} \otimes \sigma_{M_{i+1}} \otimes \ldots \sigma_{M_b} \otimes |i\rangle \langle i|_O.
\]

It holds that there exists a quantum state \( \mu_{\text{other}} \) such that \( A_{\text{good}}(\mu^2) = \sum_i p_{i,i} \mu^3 + \sum_{i \neq j} p_{i,j} \mu_{\text{other}}. \) Thus, we have,
\[
F(\mu^4, A_{\text{good}}(\mu^2)) = F(\mu^4, \frac{1}{b} \sum_{i} p_{i,i} \mu^3 + \sum_{i \neq j} p_{i,j} \mu_{\text{other}}) \geq \frac{1}{b} \sum_{i} p_{i,i} F(\mu^4, \mu^3) \geq (1 - 6\epsilon^2) F(\mu^4, \mu^3).
\]

Now, using gentle measurement lemma (Fact \ref{fact:Gentle}), we find
\[
F(\mu^4, \mu^3) = \frac{1}{\sqrt{b} \sum_i p_{i,i}} \sum_i \sqrt{p_{i,i}} \cdot F(\langle \Phi|_{RB'A'M_i} \otimes \sigma_{M_i} \otimes \ldots \sigma_{M_{i-1}} \otimes \sigma_{M_{i+1}} \otimes \ldots \sigma_{M_b} | \Phi\rangle \langle \Phi|)
\leq \frac{1}{\sqrt{b} \sum_i p_{i,i}} \sum_i \text{Tr}(\Pi^{1/2} \Pi_i \Pi^{1/2} |\Phi\rangle \langle \Phi|_{RB'A'M_i} \otimes \sigma_{M_i} \otimes \ldots \sigma_{M_{i-1}} \otimes \sigma_{M_{i+1}} \otimes \ldots \sigma_{M_b}) \sqrt{\Pi^{1/2} \Pi_i \Pi^{1/2}}
\geq \frac{1}{\sqrt{b} \sum_i p_{i,i}} \sum_i \sqrt{\sum_{i,j} p_{i,j}} \geq \sqrt{1 - 6\epsilon^2}
\]
Thus, \( F^2(\mu^4, A_{\text{good}}(\mu^2)) \geq (1 - 6\epsilon^2)^2 \geq 1 - 18\epsilon^2. \) Since swapping register \( M_j \) and \( M_1 \), controlled on value \( j \) inn register \( O \), on the state \( \mu^4 \) gives the state \( \Phi_{RB'A'M_1} \) in registers \( RB'A'M_1 \), the claim follows.

This shows that protocol \( \mathcal{P}_1 \) succeeds with fidelity squared as given in the claim. Now we proceed to construct the actual protocol.

Consider the state,
\[
\xi_{RB'M_1\ldots M_n} \overset{\text{def}}{=} \Phi_{RB'} \otimes \sigma_{M_1} \ldots \otimes \sigma_{M_n}.
\]
Let $|\theta\rangle_{L_1 \ldots L_n M_1 \ldots M_n} = |\sigma\rangle_{L_1 M_1} \otimes |\sigma\rangle_{L_2 M_2} \ldots |\sigma\rangle_{L_n M_n}$ be a purification of $\sigma_{M_1} \otimes \ldots \sigma_{M_n}$. Let

$$|\xi\rangle_{RA'B'ML_1 \ldots L_n M_1 \ldots M_n} \overset{\text{def}}{=} |\Phi\rangle_{RA'B'M} \otimes |\theta\rangle_{L_1 \ldots L_n M_1 \ldots M_n}.$$  

Using convex-split lemma (Lemma [3]) and choice of $n$ we have,

$$F^2(\xi_{RB'M_1 \ldots M_n}, \mu_{RB'M_1 \ldots M_n}) \geq 1 - \varepsilon^2.$$  

Let $|\xi\rangle_{RB'JL_1 \ldots L_n M_1 \ldots M_n}$ be a purification of $\xi_{RB'M_1 \ldots M_n}$ (guaranteed by Uhlmann’s theorem, Fact [9]) such that,

$$F^2(|\xi\rangle_{RB'JL_1 \ldots L_n M_1 \ldots M_n}, |\mu\rangle_{RB'JL_1 \ldots L_n M_1 \ldots M_n}) = F^2(\xi_{RB'M_1 \ldots M_n}, \mu_{RB'M_1 \ldots M_n}) \geq 1 - \varepsilon^2.$$  

Let $V' : A'ML_1 \ldots L_n \rightarrow JL_1 \ldots L_n$ be an isometry (guaranteed by Uhlmann’s theorem, Fact [9]) such that,

$$V' |\xi\rangle_{RA'B'ML_1 \ldots L_n M_1 \ldots M_n} = |\xi\rangle_{RB'JL_1 \ldots L_n M_1 \ldots M_n}.$$  

Consider the following protocol $P$.

1. Alice, Bob and Referee start by sharing the state $|\xi\rangle_{RA'B'ML_1 \ldots L_n M_1 \ldots M_n}$ between themselves where Alice holds registers $A'ML_1 \ldots L_n$, Referee holds the register $R$ and Bob holds the registers $B'M_1 \ldots M_n$. Note that $|\Psi\rangle_{RA'B'M}$ is provided as input to the protocol and $|\theta\rangle_{L_1 \ldots L_n M_1 \ldots M_n}$ is additional shared entanglement between Alice and Bob.

2. Alice applies isometry $V'$ to obtain state $|\xi\rangle_{RB'JL_1 \ldots L_n M_1 \ldots M_n}$, where Alice holds registers $JL_1 \ldots L_n$, Referee holds the register $R$ and Bob holds the registers $B'M_1 \ldots M_n$.

3. Alice and Bob simulate protocol $P_1$ from Step 2. onwards.

Let $\Phi'_{RA'B'M}$ be the output of protocol $P$. Since quantum maps (the entire protocol $P_1$ can be viewed as a quantum map from input to output) do not decrease fidelity (monotonicity of fidelity under quantum operation, Fact [8]), we have,

$$F^2(\Phi'_{RA'B'M}, \Phi'_{RA'B'M}) \geq F^2(|\xi\rangle_{RB'JL_1 \ldots L_n M_1 \ldots M_n}, |\mu\rangle_{RB'JL_1 \ldots L_n M_1 \ldots M_n}) \geq 1 - \varepsilon^2.$$  

This implies by claim [4] and triangle inequality for purified distance [6] that $F^2(\Phi'_{RA'B'M}, \Phi_{RA'B'M}) \geq 1 - 36\varepsilon^2$. That is, $\Phi'_{RA'B'M} \in B^{ec}(\Phi_{RA'B'M})$.

The number of qubits communicated by Alice to Bob in $P$ is equal to the number of qubits communicated in $P_1$ and is upper bounded by:

$$\frac{\log(n/b)}{2} \leq \frac{1}{2} D_{\max}(\Phi_{RB'M} || \Phi_{RB'} \otimes \sigma_M) - D_H(\Phi_{B'M} || \Phi_{B'} \otimes \sigma_M) + 2 \log \left( \frac{1}{\varepsilon} \right).$$

\[\square\]

**An achievability bound for quantum state redistribution**

Quantum state redistribution is the following coherent quantum task.

**Quantum state redistribution task**: Alice, Bob and Referee share a pure state $|\Psi\rangle_{RABC}$, with $AC$ belonging to Alice, $B$ to Bob and $R$ to Referee. Alice wants to transfer the register $C$ to Bob, such that the final state $\Phi_{RABC}$ satisfies $F(\Phi_{RABC}, \Psi_{RABC}) \geq \sqrt{1 - \varepsilon^2}$, for a given $\varepsilon \geq 0$.

Using the Theorem [5] the authors in [9] showed that the following quantity tightly captures the quantum communication cost of quantum state redistribution.
Definition 2 (9). Let $\varepsilon \geq 0$ and $|\Psi\rangle_{RABC}$ be a pure state. Define,

$$Q^R_{\langle\psi\rangle_{RABC}} = \inf_{T,U_{BCT},\sigma_T,K_{RBC}} I_{\max}(RB:CT)_{\kappa}$$

with the condition that $U_{BCT}$ is a unitary on registers $BCT$, $\sigma_T \in \mathcal{D}(T)$, $K_{RB} = \Psi_{RB}$ and

$$(I_R \otimes U_{BCT}) K_{RBC} (I_R \otimes U_{BCT}^\dagger) \in \mathcal{B}^\varepsilon(\Psi_{RBC} \otimes \sigma_T).$$

One of their main results is as follows.

Theorem 7 (9). Let $\varepsilon \in (0,1/3)$ and $\Psi_{RABC} \in \mathcal{D}(RABC)$ be a pure state. There exists an entanglement-assisted one-way protocol $P$, which takes as input $|\Psi\rangle_{RABC}$ shared between three parties Referee (R) Bob (B) and Alice (AC) and outputs a state $\Phi_{RABC}$ shared between Referee (R), Bob (BC) and Alice (A) such that $\Phi_{RABC} \in \mathcal{B}^{2\varepsilon}(\Psi_{RABC})$. The number of qubits communicated by Alice to Bob in $P$ is upper bounded by:

$$\frac{1}{2} Q^R_{\langle\psi\rangle_{RABC}} + \log \left( \frac{1}{\varepsilon} \right).$$

Furthermore, for any one way quantum protocol $P$ which takes as input $|\Psi\rangle_{RABC}$ shared between three parties Referee (R) Bob (B) and Alice (AC) and outputs a state $\Phi_{RABC}$ shared between Referee (R), Bob (BC) and Alice (A) such that $\Phi_{RABC} \in \mathcal{B}^{2\varepsilon}(\Psi_{RABC})$, the number of qubits communicated must be at least

$$\frac{1}{2} Q^R_{\langle\psi\rangle_{RABC}}.$$

Using our compression result Theorem 6, we could improve upon the achievability in terms of $Q^R_{\langle\psi\rangle_{RABC}}$. But this would not yield a substantially better bound, as the quantity $Q^R_{\langle\psi\rangle_{RABC}}$ is already nearly optimal. On the other hand, this quantity has the drawback of being a complex optimization problem. In below, we present a new achievability bound using Theorem 6 which is simpler to understand and gives the correct characterization of quantum communication cost of quantum state redistribution in asymptotic setting.

Theorem 8. Let $\varepsilon \in (0,1/3)$ and $\Psi_{RABC} \in \mathcal{D}(RABC)$ be a pure state. There exists an entanglement-assisted one-way protocol $P$, which takes as input $|\Psi\rangle_{RABC}$ shared between three parties Referee (R) Bob (B) and Alice (AC) and outputs a state $\Phi_{RABC}$ shared between Referee (R), Bob (BC) and Alice (A) such that $\Phi_{RABC} \in \mathcal{B}^{2\varepsilon}(\Psi_{RABC})$. The number of qubits communicated by Alice to Bob in $P$ is upper bounded by:

$$\inf_{\Psi'_{RABC} \in \mathcal{B}^{2\varepsilon}(\Psi_{RABC})} \frac{1}{2} \left( D_{\text{max}}(\Psi'_{RABC}\|\Psi'_{RB}\otimes\sigma_C) - D_{\text{H}}^\varepsilon(\Psi'_{BC}\|\Psi'_{B}\otimes\sigma_C) \right) + \log \left( \frac{1}{\varepsilon} \right).$$

Proof. Consider any state $\Psi'_{RABC}$ such that $\Psi'_{RABC} \in \mathcal{B}^{2\varepsilon}(\Psi_{RABC})$ and an arbitrary state $\sigma_C$. Suppose Alice, Bob and Referee started with the state $\Psi_{RABC}$ and followed the protocol as presented in the proof of Theorem 6. The resulting state $\Phi'_{RABC}$ satisfies $\Phi'_{RABC} \in \mathcal{B}^{2\varepsilon}(\Psi'_{RABC})$ and the communication cost of the protocol is

$$\frac{1}{2} \left( D_{\text{max}}(\Psi'_{RABC}\|\Psi'_{RB}\otimes\sigma_C) - D_{\text{H}}^\varepsilon(\Psi'_{BC}\|\Psi'_{B}\otimes\sigma_C) \right) + \log \left( \frac{1}{\varepsilon} \right).$$

The same protocol, when run on the input $\Psi_{RABC}$, thus produces a state $\Phi_{RABC}$ such that $P(\Phi_{RABC};\Phi'_{RABC}) \leq \varepsilon$ (as $P(\Psi_{RABC},\Psi'_{RABC}) \leq \varepsilon$ and protocol is a quantum operation). Since $P(\Psi_{RABC},\Psi'_{RABC}) \leq 6\varepsilon$, we conclude that $P(\Phi_{RABC},\Psi_{RABC}) \leq 7\varepsilon$. This completes the proof. 

Acknowledgment

This work is supported by the Singapore Ministry of Education and the National Research Foundation, also through the Tier 3 Grant Random numbers from quantum processes MOE2012-T3-1-009.
References


