Outer bounds on the quantum interference channel

Ivan Savov
ECSE 612 (multiuser communications) project, McGill University

April 15, 2010

Abstract

title says it all

1 Introduction

The interference channel is a communication network where two transmitters try to send information to two receivers using a shared medium. This communication scenario is one of the most general possible in multiuser information theory; it includes the multiple access channel (MAC) and the broadcast channel (BC) as special cases.

The interference channel is an excellent model for many practical communication scenarios where medium contention is an issue. This is why obtaining an expression for the capacity region $\mathcal{C}_{IC}$ was a central in the 70 ties and more recently. Despite the practical importance of the problem, there are very few results about discrete memoryless interference channels (DMIC) to date. Even the special case of Gaussian channels there the capacity region is not known.

Recently, quantum information theory ...

This paper will try to review classic results on the interference channel and define the problem in the quantum setting. Specifically we will try to focus on the

2 Preliminaries

Definition 1 (Interference network)**.** A two party interference network $(\mathcal{X}_1 \otimes \mathcal{X}_2, p(y_1, y_2|x_1, x_2), \mathcal{Y}_1 \otimes \mathcal{Y}_2)$ is general model for communication networks with two inputs, two outputs and a probability transistion matrix $p(y_1, y_2|x_1, x_2)$.

Definition 2 (Interference channel)**.** A two party interference channel is a particular use of an interference network $(\mathcal{X}_1 \otimes \mathcal{X}_2, p(y_1, y_2|x_1, x_2), \mathcal{Y}_1 \otimes \mathcal{Y}_2)$ where a message M_1 is encoded into a codeword X_1 and is to be extracted at the receiver Y_1.

another version – II

Definition 3 (Achievable rate pair)**.** We say that a rate (R_1, R_2) is achievable for a channel $(\mathcal{X}_1 \otimes \mathcal{X}_2, p(y_1, y_2|x_1, x_2), \mathcal{Y}_1 \otimes \mathcal{Y}_2)$ if there exists a code for n uses of the channel where the messages

taken from respective sets $\{1, 2, \ldots, 2^{nR_1}\}$ and $\{1, 2, \ldots, 2^{nR_2}\}$ are transmitted with vanishing error probability.

Definition 4 (Capacity)**.** The capacity region $\mathcal{C}$ of is the closure of the set of achievable rates (R_1, R_2).

Definition 5 (Degraded channel)**.** Let $Y \sim p_y(y|x_1, x_2)$ and $Z \sim p_z(z|x_1, x_2)$ be two conditional distributions. We say Y is a *degraded version* of Z ... copy from [Car83] II A.

Both only depend on marginals $p(y_1|x_1x_2)$ and $p(y_2|x_1x_2)$ since if we manage to get both decoding errors low then we manage to get the error of the AND of the two events also. [*more details needed...*]

Definition 6 (Independent channel)**.** Let $p(y_1|x_1, x_2)$ and $p(y_2|x_1, x_2)$ be the marginals of the associated with some interference network $p(y_1, y_2|x_1, x_2)$. We call the independent channel the interference network

$$p_i(y_1, y_2|x_1, x_2) = p(y_1|x_1x_2)p(y_2|x_1x_2). \tag{1}$$

2.1 Relation to multiple access and broadcast channels

We can think of the IC as either two multiple access channels or two two broadcast channels.

It is therefore important to review briefly the known capacity results for these simpler channels.

Another generalization that we will not be considered here is when the two inputs sources X_1, X_2 can be correlated. To study this problem we must also be familiar with Slepian-Wolf coding of correlated sources. In general the joint source-channel coding is a very interesting problem that remains... [*What can I say more precisely?*]

3 Literature review

3.1 Early papers

Ahlswede 1978: The capacity region of a channel with two senders and two receivers One of the earliest papers to appear on the topic of interference channels is by Rudolf Ahlswede [Ahl74]. While the notions of entropy and mutual information were already well established by those days, the author prefers to use the definitions of entropic quantities from first principles:

$$R_1(p, q, Y_1) = \sum_{x_1,x_2,y_1} p(x_1)p(x_2)p(y_1|x_1, x_2) \log \frac{p(y_1|x_1, x_2)}{\sum_{x_1} p(x_1)p(y_1|x_1, x_2)} = I(X_1; Y_1|X_2). \tag{2}$$

The paper defines three communication problems that can be studied on the interference network $p(y_1, y_2|x_1, x_2)$.

- The **mariginal Y_1-MAC channel** where two inputs X_1 and X_2 encode independent messages M_1, M_2 at rates R_1, R_2, which must be decoded at Y_1. The relevant probability distribution is $\sum_{y_2} p(y_1, y_2|x_1, x_2)$. Ahlswede calls this the (p, T_{21}, I) *communication situaiton*.

- The **IC** problem which we have defined above. In Ahlswede nomenclature this is a (p, T_{22}, I) comunication situation.
- Finally there is the **multiple MAC** (MMAC) problem, where messages M_1 and M_2 are to be decoded at both receivers. This is called the (p, T_{22}, II) problem in the paper.

The paper then proves the capacity of the Y_1-MAC problem.

$$\begin{aligned} R_1 &\leq I(X_1; Y_1|X_2), && (3)\\ R_2 &\leq I(X_2; Y_1|X_1), && (4)\\ R_1 + R_2 &\leq I(X_1X_2; Y_1) \qquad \text{for some } p(x_1)p(x_2)\ . && (5) \end{aligned}$$

The author then uses the same argument to prove the capacity of the MMAC channel when the two inputs X_1 and X_2 are allowed to be correlated.

$$\begin{aligned} R_1 &\leq \min\{I(X_1; Y_1|X_2), I(X_1; Y_2|X_2)\}, && (6)\\ R_2 &\leq \min\{I(X_2; Y_1|X_1), I(X_2; Y_2|X_1)\}, && (7)\\ R_1 + R_2 &\leq \min\{I(X_1X_2; Y_1), I(X_1X_2; Y_2)\} \qquad \text{for some } p(x_1, x_2)\ . && (8) \end{aligned}$$

In the paper [*Why need correlated sources?*]
Let $\mathcal{R}_{p(x_1,x_2)}$ be the closure [*of the convex hull?*] of rates (R_1, R_2) which satisfy equations (6) through (6). And let $\mathcal{C}(p, T_{22}, II)$ be the capacity of for the MMAC problem with correlated sources. We want to show that $\mathcal{R}_{p(x_1,x_2)} = \mathcal{C}(p, T_{22}, II)$.

[*Is there an intuitive explanation why this is true?*]
For the part $\mathcal{R}_{p(x_1,x_2)} \subseteq \mathcal{C}(p, T_{22}, II)$ we have the following argument... if you have some rates R_1 and R_2 that are feasibly for both the Y_1-MAC and the Y_2-MAC then you must be able to do the MMAC problem too.

The other direction $\mathcal{R}_{p(x_1,x_2)} \supseteq \mathcal{C}(p, T_{22}, II)$

[*NOT CLEAR ...*]

This problem is connected to correlated sources coding. Slepian-Wolf etc...

The author then makes a remark that knowing the capacity of the MMAC problem does not help us to find the IC problem.

Sato 77: Two-user communication channels

The next paper, and perhaps the most important paper on the interference channel is by Hiroshi Sato [Sat77].

The

three special relations to

Cheng Motani Garg ? best inner bound

3.2 Outer bounds

3.3 Quantum communicaiton channels

4 Known outer bounds

4.1 Naive outer bound

Both broadcast channels and multiple access channel are special cases of the interference channel. In particular we can think of the interference channel as two separate multiple access channels.

We know that the region defined by

$$\begin{aligned} R_1 &\leq I(X_1; Y_2|X_2) && (9)\\ R_2 &\leq I(X_2; Y_2|X_1) && (10) \end{aligned}$$

contains the capacity region $\mathcal{C}$.

This corresponds to a very loose rectangular bound on the true capacity region.

4.2 Sato bound

We can describe a more precises outer bound to the capacity region by spcifying an inequality on the sum rate $R_1 + R_2$. This was done by Sato [Sat78].

The outer bound becomes:

$$\begin{aligned} R_1 &\leq I(X_1; Y_2|X_2) && (11)\\ R_2 &\leq I(X_2; Y_2|X_1) && (12)\\ R_1 + R_2 &\leq I(X_1X_2; Y_1Y_2) && (13) \end{aligned}$$

Proof. b □

4.3 Carleial

A further development concerning an outer bound was obtained by Carleilal [Car83].

Consider the two random variables Z_1, Z_2 such that

$$\begin{aligned} Y_1 &\quad \text{is a degraded version of} \quad Z_1 && (14)\\ Y_2 &\quad \text{is a degraded version of} \quad Z_2 && (15)\\ Y_2 &\quad \text{is a degraded version of} \quad (X_1, Z_1) && (16)\\ Y_1 &\quad \text{is a degraded version of} \quad (X_2, Z_2) && (17) \end{aligned}$$

If we think of Z_1, Z_2 as the outputs of the channel, we can interpret the above conditions as the following. Knowing the input and output of the other person's communication we can recover our own signal. If we manage to decode our message correctly, then I can also simulate the output that the other person has received – i.e. I have a joint decoder on (Y_1, Y_2).

then we have the following outer bound

$$\begin{aligned} R_1 &\leq I(X_1;Y_1|X_2) && (18)\\ R_2 &\leq I(X_2;Y_2|X_1) && (19)\\ R_1+R_2 &\leq \min\{I(X_1X_2;Z_1), I(X_1X_2;Z_2)\} && (20) \end{aligned}$$

4.4 Nair - El Gamal outer bound

[*is this useful?*]

In a recent paper [NEG07], Nair and El Gamal give the following outer bound on the BC with independent messages M_1, M_2 encoded into U, V respectively, which are later *jointly encoded* into an input symbol X for the BC.

$$\begin{aligned} R_1 &\leq I(U;Y_1) && (21)\\ R_2 &\leq I(V;Y_2) && (22)\\ R_1+R_2 &\leq I(U;Y_1)+I(V;Y_2|U) && (23)\\ R_1+R_2 &\leq I(V;Y_2)+I(U;Y_1|V) && (24) \end{aligned}$$

for some choice of input distribution $p(u,v,x) = p(u,v)p(x|u,v)$.

The region defined above is an outer bound on the BC, which involves joint encoding of the two sources U, V. If we are dealing with the IC, we have a more restrictive scenario since $U \to X_1$ and $V \to X_2$, i.e. we don't have joint encoding of the two sources.

It follows that the above region is also an outer bound on the IC cap region.

5 Quantum interference channel

6 Discussion and conclusion

Most general use of the interference network The difference between the IC and the MMAC is that we guarantee that an extra ressource of "cross communication" is available. Wouldn't it be best to represent rates then as 4-tuples?

$$\begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \quad (25)$$

The IC problem is basically a promise about R_{11} and R_{22}, and no statement about the cross rates.

Are the cross rates not useful? These extra rates could be used to convert some other information and I feel they should be taken into account in general. (entanglement between receivers?)

open problems...

A Introduction to QIT

A.1 Informal introduction

A mixed quantum state, the most general kind of quantum state, is a probability distributions over probability distributions. The "outer" probability is the same as the probability $p(x)$ in describing an "information source" in classical information theory. It reflects the probability that the source S outputs symbol x.

The inner probabilities are *pure* quantum states which are just vectors in some normed complex vector spaces. The two most important vector spaces to know about are

- The set of complex functions: $\psi \colon \mathbb{R} \to \mathbb{C}$,
- The set of n dimensional vectors over the complex numbers $\mathbb{C}$.

The continuous *wavefunction representaiton* of quantum mechanics is useful in optics, atomic physics and solid state physics, where people calculate things by integrals in the position basis (Dirac $\delta(x)$ functions), and its fourier transform ($sin(\omega x), cos(\omega x)$) the momentum basis.

The matrix representation, which we will focus on here, is useful for describing phenomena like light polarizaiton, electron spin, angular momentum and other finite dimensional degrees of freedom. It is also the basis of choice for the entire field of Quantum Information Science, which comprises quantum information theory, quantum crypography, quantum error correcting codes, quantum computation and many others.

The basic informaiton carriers in QIT are *qbits*, two dimensional quantum degrees of freedom analogous to the classical bits.

Definition 7 (qubit)**.** A qubit is a two dimensitonal quantum state $\vec{s}$ is a unit length vector in $\mathbb{C}^2 = \{(a,b)^T | a, b \in \mathbb{C}\}$.

For example, we can prepare a electron is a spin state which has spin up $\vec{s} = (1,0)^T$, down $\vec{s} = (0,1)^T$ or a 50–50 *superposition* of the up and down $\vec{s} = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^T$.

The description in words "50 50 superposition" is not rich enough to describe all possible superpositions since the coefficients of the voctor are in general complex numbers. Here is another 50–50 quantum state: $\vec{s} = (\frac{1}{\sqrt{2}}, \frac{-i}{\sqrt{2}})^T$, in fact there is a whole continuum of 50–50 states $\vec{s} = (\frac{1}{\sqrt{2}}, \frac{e^{i\theta}}{\sqrt{2}})^T$.

Note however that the global complex phase is not important $(1,0)^T = e^{i\theta}(1,0)^T$ so without loss of generality we can always assume that the first component of any quantum vector is real.

An ensemble $\mathcal{E} = \{p_i, |\psi_i\rangle\}$ is a set of quantum states $|\psi_i\rangle$ which occur with probability p_i. One way to describe a quantum source is to specify the states $|\psi_i\rangle$ and the corresponding probabilities p_i associated with this source.

A.2 Informal introduction

A.3 Quantum states

The fundamental principles of quantum mechanics are simple enough to be explained in the space available on the back of an envelope, but to truly understand the implications of these principles takes years of training and effort. We assume the reader is familiar with basic notions of quantum mechanics [Sak94, NC00]. This section will focus on specific notions and notation that are used in quantum information theory.

We will denote quantum systems by uppercase roman letters like A, B, R and the corresponding Hilbert spaces as $\mathcal{H}^A, \mathcal{H}^B, \mathcal{H}^R$ with respective dimensions d_A, d_B, d_R. We denote pure states of the system A by *ket*s: $|\varphi\rangle^A$ and *density matrices* as φ^A. Because of the probabilistic interpretation of quantum mechanics, all kets have unit norm and all density matrices are positive and with unit trace. We will refer to both kets and density matrices as *states*.

We use the partial trace operator to model partial knowledge of a state. Given a bipartite state ρ^{AB} shared between Alice and Bob, we say that Alice holds in her lab the reduced density matrix: $\rho^A = \mathrm{Tr}_B \rho^{AB}$, where Tr_B denotes a partial trace over Bob's degrees of freedom. In general the state produced in this manner will be *mixed* – a classical probability distribution over states.

Conversely, any mixed state $\sigma^A \in {}^1\mathcal{H}^A$ can be *purified* to a fictitious larger Hilbert space. That is, we imagine a corresponding pure state $|\sigma\rangle^{AR} \in \mathcal{H}^A \otimes \mathcal{H}^R$ such that taking the partial trace over the R system gives the original state: $\mathrm{Tr}_R \left(|\sigma\rangle\langle\sigma|^{AR} \right) = \sigma^A$. The purification procedure is often referred to as escaping to the *Church of the larger Hilbert space* in literature.

A.4 Quantum information theory

The fundamental ideas of quantum information theory are analogous to those of classical information theory.

A.4.1 von Neumann entropy

Analogously to classical information theory, we quantify the information content of quantum systems by using an entropy function.

Definition 8 (von Neumann Entropy)**.** Given the density matrix $\rho^A \in \mathcal{H}^A$, the expression

$$S(A)_\rho = -\mathrm{Tr}\left(\rho^A \log \rho^A\right) \tag{26}$$

is known as the *von Neumann entropy* of the state ρ^A.

We often use the notation H for entropy even in the quantum case because it is essentially the same funciton; the von Neumann entropy of quantum state ρ^A (density matrix) with spectral

[1]Strictly speaking, we should say $\sigma^A \in D(\mathcal{H}^A)$ where $D(\mathcal{H}^A)$ is the set of density matrices over $\mathcal{H}^A$. We will use this economy of notation consistently.

decomposition $\rho^A = \sum_i \lambda_i |e_i\rangle\langle e_i|$, we can calculate $H(A)_\rho = -\text{Tr}\left(\rho^A \log \rho^A\right) = -\sum_i \lambda_i \log \lambda_i$. The von Neumann entropy of a pure state is zero, since it has only a single eigenvalue.

For bipartite states ρ^{AB} we can also define the quantum conditional entropy

$$H(A|B)_\rho := H(AB)_\rho - H(B)_\rho \tag{27}$$

where $H(B)_\rho = -\text{Tr}\left(\rho^B \log \rho^B\right)$ is the entropy of the reduced density matrix $\rho^B = \text{Tr}_A\left(\rho^{AB}\right)$. In the same fashion we can also define the quantum mutual information information

$$I(A;B)_\rho := H(A)_\rho + H(B)_\rho - H(AB)_\rho \tag{28}$$

and in the case of a tripartite system ρ^{ABC} we define the conditional mutual information as

$$\begin{aligned} I(A;B|C)_\rho &:= H(A|C)_\rho + H(B|C)_\rho - H(AB|C)_\rho & (29)\\ &= H(AC)_\rho + H(BC)_\rho - H(ABC)_\rho - H(C)_\rho. & (30) \end{aligned}$$

It can be shown that $I(A;B|C)$ is strictly non negative for any state ρ^{ABC}. The formula $I(A;B|C) \geq 0$ can also be written in the form

$$H(AC) + H(BC) \geq H(C) + H(ABC). \tag{31}$$

This inequality, originally proved in [LR73], is called the *strong subadditivity* of von Neumann entropy and forms an important building block of quantum information theory.

On the surface, it may appear to the reader that quantum information theory has nothing new to offer except a rewriting of the classical formulas in a new context. This observation is highly misleading. We present the following example to illustrate some of the new aspects of quantum information theory.

Example 1. Consider the Φ^+ Bell state

$$|\Phi\rangle^{AB} = \tfrac{1}{\sqrt{2}}(|00\rangle^{AB} + |11\rangle^{AB}). \tag{32}$$

This state exhibits a form of quantum correlation called *entanglement* that is fundamentally different from classical correlation. The associated density matrix is $\Phi^{AB} = |\Phi\rangle\langle\Phi|^{AB}$, which has the reduced density matrices $\Phi^A = \Phi^B = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|)$.

Next we calculate the entropy of the two subsystems A, B and the system as a whole

$$H(A)_\Phi = 1, \qquad H(B)_\Phi = 1, \qquad H(AB)_\Phi = 0, \tag{33}$$

since Φ^A, Φ^B are maximally mixed and $|\Phi\rangle^{AB}$ is pure. Using these results, it is now simple to calculate the conditional entropy

$$H(A|B) = H(AB) - H(B) = -1 \text{ [bits]}, \tag{34}$$

and the mutual information

$$I(A;B) = H(A) + H(B) - H(AB) = 2 \text{ [bits]}. \tag{35}$$

Equation (34) illustrates one of the key differences between classical information theory and quantum information theory: the fact that conditional entropy can be negative.

In classical information theory, the mutual information between two binary sources attains its maximal value of 1 when the two sources are perfectly correlated. As we can see from equation (35), in the quantum world two qubits can be, in some sense, *more than perfectly correlated* and have mutual information as much as 2 bits!

A.5 Quantum operations

Quantum operations are mappings that take quantum states as inputs and produce quantum states as outputs. We usually denote them by calligraphic letters like $\mathcal{E}$ and $\mathcal{D}$:

$$\rho \xrightarrow{\mathcal{E}} \rho' \qquad \text{or} \qquad \mathcal{E}(\rho) = \rho' \tag{36}$$

Unitary transformations are a type of quantum operation

$$\mathcal{E}(\rho) = U\rho U^\dagger \tag{37}$$

where U is a unitary matrix. Unitary operations coorespond to the evolution of isolated systems that do not interact with the world.

In real life, however, no system is perfectly isolated from its environment ρ_{env} and when we account for it we obtain the more general form of operation

$$\mathcal{E}(\rho) = \text{Tr}_{\text{env}} U \left(\rho \otimes \rho_{\text{env}}\right) U^\dagger \tag{38}$$

where the unitary operation now acts on the enlarged Hilbert space, but we trace out over the environment degrees of freedom in the end.

More generally, the laws of physics require all quantum operations to be: trace preserving (TP), hermitian preserving and completely positive (CP). Thus, another name for quantum operations is CPTP maps.

Measurements are the second fundamental class of quantum operations. They are our only means to relate the quantum world to classical variables we can observe. A measurement operation $\mathcal{E}$:$\mathcal{H}^A \rightarrow (\mathcal{H}^A, \mathbb{N})$ acts on density matrices to produce a classical output as well as a possibly modified quantum state. It is modeled by a set of projection operators $\{M_i\}$, which sum up to the identity operator $\sum_i M_i^\dagger M_i = I$. The probability of outcome m occurring when the input system is ρ is given by

$$p(m) = \text{Tr}\Big(M_m^\dagger \rho M_m\Big)\,. \tag{39}$$

The output quantum state associated with this outcome is

$$\widetilde{\rho}_m = \frac{M_m^\dagger \rho M_m}{\text{Tr}\Big(M_m^\dagger \rho M_m\Big)}. \tag{40}$$

A.6 Quantum resources

The current trend in quantum information theory is to look at communication tasks as interconversions between clearly defined information resources. To render the resource picture generic, we always imagine a scenario in which two localized parties, usually called Alice and Bob, want to perform a certain communication task. Local computation will be regarded as free of cost in order to focus on the communication aspects of the task.

An example of a classical communication resource is the *noiseless channel* from Alice to Bob, denoted $[c \to c]$. The symbol $[c \to c]$ represents the ability to send one bit of information from Alice to Bob. A related classical resources is the *noisy channel*, denoted $\{c \to c\}$ which is usually modeled as a probabilistic mapping $\mathcal{N}^{X \to Y}$ with probability $p(Y = y|X = x)$ where X is the input variable sent by Alice and Y the random variable received by Bob. The noiseless channel $[c \to c]$ is, therefore, a special case of the general channel $\{c \to c\}$ with the identity mapping $\mathcal{N} = \mathrm{id}^{X \to Y}$ from X to Y. Another classical resource denoted $[cc]$ represents a random bit shared between Alice and Bob.

Quantum information theory introduces a new set of resources. In analogy to the classical case, we have the *noiseless quantum channel* $[q \to q]$ which represents the ability to transfers one *qubit*, a generic two dimensional quantum system, from Alice to Bob. A *noisy quantum channel*, $\{q \to q\}$, is modeled by a mapping $\mathcal{N}^{A \to B}$ which takes density matrices in $\mathcal{H}^A$ to density matrices in $\mathcal{H}^B$. The mapping $\mathcal{N}$ is a *quantum operation*: a completely positive trace preserving (CPTP) operator [NC00].

Once we have defined the different classical and quantum communicaiton resources, we can state information theoretic results in a very consise form. A *resource inequalities*, $2 * a + b \geq c$ is a statement which indicates that the resources on the left hand side (two uses of a and one of b), can be used to simulate the resource on the right hand side (c).

To illustrate the new notation we will state the famous channel capacity formula **[?]**:

$$\{c \to c\} \geq \max_{p(x)} I(X;Y)[c \to c], \tag{41}$$

which states that a noisy classical channel $\mathcal{N}$ can be used as a noiseless channel at the "conversion rate" equal to the capacity $\mathcal{C} = \max_{p(x)} I(X;Y)$.

The key resource that differentiates quantum information theory from its classical counterpart are the maximally entangled states shared between Alice and Bob

$$|\Phi\rangle^{AB} = \tfrac{1}{\sqrt{2}}(|00\rangle^{AB} + |11\rangle^{AB}), \tag{42}$$

which we denote $[qq]$ in resource inequalities. Entanglement is a fundamental quantum resource because it cannot be generated by local operations and classical communication (LOCC). The precise characterization of entanglement has been a great focal point of research in the last decade. For an in depth review of the subject we refer the readers to the excellent papers [VP98, HHHH07]. Entanglement forms a crucial building block for quantum information theory because it can be used to perform or assist with many communication tasks.

In particular, two of the first quantum protocols that ever appeared involve *ebits*, or entangled bits. The *quantum teleportation* protocol [BBC$^+$93] uses entanglement and two bits of classical communication to send a quantum state from Alice to Bob

$$[qq] + 2[c \to c] \quad \geq \quad [q \to q], \tag{TP}$$

while the *superdense coding* protocol [BW92] uses entanglement to send two classical bits of information with only a single use of a quantum channel

$$[qq] + [q \to q] \quad \geq \quad 2[c \to c]. \tag{SC}$$

The two protocols (TP) and (SC) are only the tip of the iceberg: there are many more protocols and fundamental results in quantum information theory that can be written as resource inequalities.

A.7 Error analysis

In the error criterion for most classical information theory protocols is of the form $P_e = Pr\{M_1 \neq \hat{M}_1\}$. We need an analogous criterion for comparing quantum states.

The fidelity between two pure quantum states is the square of their inner product

$$F(|\varphi\rangle, |\psi\rangle) = |\langle\varphi|\psi\rangle|^2 . \tag{43}$$

The natural generalization of this notion to mixed states ρ, σ is the formula

$$F(\rho, \sigma) = \mathrm{Tr}\left(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)^2 . \tag{44}$$

Two states that are very similar have fidelity close to 1 whereas states with little similarity will have low fidelity.

Let $\mathcal{N}^{A\to\widehat{A}}$ with input $|\psi\rangle^A \in \mathcal{H}^A$ and output $\sigma^{\widehat{A}} \in \mathcal{H}^{\widehat{A}}$ be the quantum operation associated with the protocol:

$$\mathcal{N}(|\psi\rangle\langle\psi|) = \sigma^{\widehat{A}}. \tag{45}$$

To measure how faithfully the input state has been reproduced at the output we calculate the input-output fidelity $F(|\psi\rangle^A, \sigma^{\widehat{A}})$. In order to measure how faithfully the source as a whole is reproduced at the output, we have to average over the input-output fidelities of the ensemble

$$\overline{F}(\{p_i, |\psi_i\rangle\}, \mathcal{N}) := \sum_i p_i F(|\psi_i\rangle, \sigma_i), \qquad \sigma_i = \mathcal{N}(|\psi_i\rangle\langle\psi_i|). \tag{46}$$

If we want the source to be preserved perfectly then we require $\overline{F}(\mathcal{E}, \mathcal{N}) = 1$. In general, however, we will be content with approximate transmission where

$$\overline{F}(\mathcal{E}, \mathcal{N}) \geq 1 - \epsilon \tag{47}$$

for arbitrary small ϵ.

For technical reasons, it turns out that the better way of judging the success of a quantum protocol that relies on the idea of the *Church of the larger Hilbert space*. Let $|\psi\rangle^{AR}$ be a purification of ρ^A to some reference system R. This reference system is entirely fiducial and does not participate in the protocol. In the larger Hilbert space $\mathcal{H}^A \otimes \mathcal{H}^R$ the $\mathcal{N}^{A\to\widehat{A}}$ operation acts as

$$\mathcal{N}^{A\to\widehat{A}} \otimes \mathrm{id}^R\big(|\psi\rangle\langle\psi|^{AR}\big) = \sigma^{\widehat{A}R}. \tag{48}$$

For approximate transmission, we now require the fidelity between the pure input state $|\psi\rangle^{AR}$ and the possibly mixed output state $\sigma^{\widehat{A}R}$ to be high

$$F(|\psi\rangle^{AR}, \sigma^{\widehat{A}R}) = \langle\psi^{AR}|\sigma^{\widehat{A}R}|\psi^{AR}\rangle \geq 1 - \epsilon. \tag{49}$$

Equation (49) measures the *entanglement fidelity* of the operation: how well the protocol manages to transfers the R-entanglement from the A system to the $\widehat{A}$ system. In other words, not only are we guaranteeing that the particular sustem A was successfully transmitted to system $\widehat{A}$, but that all possible correlations of A with the outside world were also faithfully transmitted.

It can be shown [**?**] that if the channel $\mathcal{N}$ has high entanglement fidelity then the average fidelity $\overline{F}(\mathcal{E}, \mathcal{N})$ will also be high for any ensemble $\mathcal{E}$ such that $\rho^A = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. In other words, equation (49) implies equation (47).

B Compression

The *Typical Subspace* is the support of $P_\epsilon^{(n)}$ or equivalently, SPAN$\{|j^n\rangle, j^n \in T_\epsilon^{(n)}\}$.

$$\text{Tr}\left[\rho^{\otimes n} P_\epsilon^{(n)}\right] > 1 - \delta \qquad \forall \delta, \epsilon > 0, \text{and } n \text{ sufficiently large.} \tag{50}$$

$$\text{rank} P_\epsilon^{(n)} = |T_\epsilon^{(n)}| \tag{51}$$

$$\forall \epsilon > 0, \quad |T_\epsilon^{(n)}| \leq 2^{n[H(\rho)+\epsilon]} \tag{52}$$

$$\forall \epsilon, \delta > 0, \quad |T_\epsilon^{(n)}| \geq (1-\delta) 2^{n[H(\rho)-\epsilon]} \tag{53}$$

Proof. The proof of (51) is easy.
For (52) we note $H(\rho) = H(r)$ and use the Typical Sequence Theorem **??** [*From Punit's part*]. The only thing one that is more interesting is (50) for which we note:

$$\text{Tr}\left[\rho^{\otimes n} P_\epsilon^{(n)}\right] \quad = \quad \sum_{j^n} \langle j^n| \left(\sum_{k^n} r_{k^n} |k^n\rangle\langle k^n| \right) \sum_{l^n \in T_\epsilon^{(n)}} |l^n\rangle\langle l^n| |j^n\rangle \tag{54}$$

□

References

[Ahl74] R. Ahlswede. The capacity region of a channel with two senders and two receivers. *The Annals of Probability*, 2(5):805–814, 1974.

[BBC+93] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters. Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels. *Phys. Rev. Lett.*, 70:18951899, 1993.

[BW92] C. H. Bennett and S. J. Wiesner. Communication via one- and two-particle operators on einstein-podolsky-rosen states. *Phys. Rev. Lett.*, 69:28812884, 1992.

[Car83] A. Carleial. Outer bounds on the capacity of interference channels (Corresp.). *IEEE transactions on information theory*, 29(4):602–606, 1983.

[HHHH07] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. Quantum entanglement. 2007. arXiv:quant-ph/0702225.

[LR73] E. H. Lieb and M. B. Ruskai. Proof of the strong subaddivity of quantum-mechanical entropy. *J. Math. Phys.*, 14:1938–1941, 1973.

[NC00] M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, 2000.

[NEG07] C. Nair and A. El Gamal. An outer bound to the capacity region of the broadcast channel. *IEEE Transactions on Information Theory*, 53(1):350–355, 2007.

[Sak94] J.J. Sakurai. *Modern quantum mechanics*. Addison-Wesley, 1994.

[Sat77] H. Sato. Two-user communication channels. *IEEE transactions on information theory*, 23(3):295–304, 1977.

[Sat78] H. Sato. An outer bound to the capacity region of broadcast channels (Corresp.). *IEEE Transactions on Information Theory*, 24(3):374–377, 1978.

[VP98] V. Vedral and M. B. Plenio. Entanglement measures and purification procedures. *Phys. Rev. A*, 57:1619, 1998. arXiv:quant-ph/9707035.