Outer bounds on the quantum interference channel

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

April 21, 2010

abstract ?

1 Introduction

The interference channel (IC) 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 $\mathcal{C}_{IC}$ is of great central importance to theoretical and practical information theory. 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 (QIT) has been a popular research area in physics, computer science and mathematics. This new field of informaiton theory is firmly based in Shannon's principle of informaiton-as-statistics, but studies underlying systems that behave according to the laws of quantum mechanics. The current state of QIT is similar to that of classical informaiton theory in the 80s, when most single user problems had been solved and researchers were exploring new grounds in multiuser informaiton theory.

This paper will try to review classic results on the interference channel and define the problem in the quantum setting. In section 2 we will define the basic concepts related to the classical interference channels. Then in section 3 we will reproduce the main results from a series of seminal classical papers on the interference channel. Results relating to outer bounds on the IC capacity are collected together in section 4. The next seciton (seciton 5) deals with the quantum generalization of the multiple access channel and broadcast channel. Readers not faimilar with quantum information theory are encouraged to consult appendix A, which gives a brief introduction to the subject. Finally we have a section of discussion and pointers to open problems in seciton .

2 Preliminaries

In this seciton we introduce important notation and nomenclature which will be used throughout this paper. We also state some important results from multiuser information theory, which we will need as building blocks.

2.1 Definitions

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 transition 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 messages M_1, M_2 are independently encoded into a codewords X_1, X_2 at rates (R_1, R_2) with Y_1, Y_2 being the intended receiver.*

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*

$$\Pr\left\{M_1 \neq \hat{M}_1 or M_2 \neq \hat{M}_2\right\} \leq \epsilon. \tag{1}$$

Definition 4 (Capacity). *The capacity region $\mathcal{C}_{IC}$ of is the convex hull of the closure of the set of achievable rate pairs (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 random variables defined in terms of X_1, X_2. We say Y is a* degraded version *of Z with respect to (X_1, X_2) if there exits Y', defined on the same sample space as Y, such that $p_{y'}(y|x_1, x_2, z) = p_{y'}(y|z)$ and $p_{y'}(y'|x_1, x_2) = p_y(y|x_1, x_2)$.*

Another way to say the above is that Y' is statistically equivalent to Y and that $(X_1, X_2) \rightarrow Z \rightarrow Y'$ for a Markov chain.

2.2 Convex sets, closures and optimization constraints

Rate regions in multiuser information theory often stated in terms of constrained optimization problems over information theoretic quantities. Furthermore, because of the *time-sharing* principle, we can achieve any rate in the *convex hull* of any other achievable rates.

To rigorously state each capacity result can become overly wordy if we do not find a compact notation to express the above notions. Luckily, smarter people than us have already found such a compact notation [Sat77]. We illustrate is with a well known result [Sha48].

The classical capacity of a single user channel $(\mathcal{X}, p(y|x), \mathcal{Y})$ is given by

$$\mathcal{C} = \operatorname{conv} \bigcup_{p(x_1)} \{R \mid R \leq I(X;Y)\ \}. \tag{2}$$

Where the $p(x_1)$ which we are optimizing over should not be confused with the $p(y|x)$ associated with the channel. The conv is not really necessary in this case.

Technically speaking, the rate $R = I(X;Y)$ is not achievable for the single user channel, but any rate $R - \epsilon$ is. The rigorous way of saying this would be to refer to "the closure of" the set or rates $R < I(X;Y)$, but we will avoid this extra mathematical kung fu and simply use the equalities $R \leq I(X;Y)$ for our rate inequalities.

2.3 Multiple access channel

The multiple access channel (MAC) is a special cases of the interference channel with a single receiver. It is therefore important to review briefly the known capacity result for the MAC.

Definition 6 (MAC)**.** *A multiple access channel is defined as* $(\mathcal{X}_1 \otimes \mathcal{X}_2, p(y|x_1,x_2), \mathcal{Y})$*, and models a communication network with two inputs and one output with the probability transition matrix* $p(y|x_1,x_2)$*.*

The MAC channel is one of the few problems multiuser information theory for which we know precisely the capacity region.

$$\mathcal{C}_{MAC} = \mathrm{conv} \bigcup_{p(x_1)p(x_2)} \{(R_1,R_2)|\text{ Eqn. (4) }\} \tag{3}$$

$$\begin{aligned} R_1 &\leq I(X_1;Y|X_2), \\ R_2 &\leq I(X_2;Y|X_1), \\ R_1+R_2 &\leq I(X_1X_2;Y). \end{aligned} \tag{4}$$

The coding strategy that makes these rates possible is *successive decoding*. Each receiver first decodes one of the two messages, then decodes the second message using the first as side information. The better rate is obtained for the second message, since we have side information.

2.4 Broadcast channel

While the capacity of the broadcast channel (BC) is not known in general, there are several insights learned from the BC problem that also apply to the study of interference channels.

The first thing to notice is that the error criterion in equation (1) of definition 3 only depend on marginals $p(y_1|x_1x_2)$ and $p(y_2|x_1x_2)$. This is because if we manage to get both decoding errors low, then we manage to get the error of the OR of the two events also.

This observation leads us to the realization that all channels which have the same marginals will have the same capacity. More formally, we will define the class $K(p)$ of interference channel probability matrices that have marginals as $p(y_1,y_2|x_1x_2)$:

$$K(p) \triangleq \{q(y_1,y_2|x_1,x_2) \mid q_1(y_1|x_1,x_2) = p_1(y_1|x_1,x_2) \texttt{ and } q_2(y_2|x_1,x_2) = p_2(y_2|x_1,x_2)\}. \tag{5}$$

Thus the capacity region $\mathcal{C}_{IC}$ for an IC channel with probability p, is the same for all channels in $K(p)$. In particular, we can define a particular member of $K(p)$ which might be simpler to analyse

Definition 7 (Independent channel). *Let $p_1(y_1|x_1, x_2)$ and $p_2(y_2|x_1, x_2)$ be the marginals of the associated with some interference network $p(y_1, y_2|x_1, x_2)$. We define the* independent channel *the interference network with product probability transition matrix*

$$p_i(y_1, y_2|x_1, x_2) = p_1(y_1|x_1x_2)p_2(y_2|x_1x_2). \tag{6}$$

2.5 Other uses of the interference network

The interference channel is not the most general communication scenario that we can consider using the the interference network. One possibility is to require transmitters to send messages to both receivers, which will make the channel more similar to a multiple-broadcast channel.

Another generalization of the IC problem is to allow for correlations between the channel inputs X_1, X_2. To study this problem we must also be familiar with Slepian-Wolf coding of correlated sources.

We will briefly mention these and other generalization of the IC in the discussion section at the end of this document.

3 Classical results on the Interference Channel

In this section we present a literature review of important papers which deal with the interference channel and related problems.

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:

$$R_1(p, q, Y_1) = \sum_{x_1, x_2, y_1} p(x_1)p(x_2)p(y_1|x_1, x_2) \log \frac{p(y_1|x_1, x_2)}{\sum_{x_1} p(x_1)p(y_1|x_1, x_2)} = I(X_1; Y_1|X_2). \tag{7}$$

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, which we stated in the introduction (3).

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.

$$G_{MMAC} = \text{conv} \bigcup_{\mathbf{p(x_1,x_2)}} \{(R_1, R_2)| \text{ Eqn. (9) } \} \tag{8}$$

$$\begin{aligned} R_1 &\leq \min\{I(X_1;Y_1|X_2), I(X_1;Y_2|X_2)\}, \\ R_2 &\leq \min\{I(X_2;Y_1|X_1), I(X_2;Y_2|X_1)\}, \\ R_1 + R_2 &\leq \min\{I(X_1X_2;Y_1), I(X_1X_2;Y_2)\} \end{aligned} \tag{9}$$

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 (**??**) through (**??**). 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

Carleil 78: Interference channels

3.2 Han-Kobayashi / Cheng-Motani-Garg inner bound

First appeared in ...

$$G_{HK} = G_{CMG} \triangleq \text{conv} \cup_{p(x_1|u_1)p(x_2|u_2)p(u_1)p(u_2)} \{(R_1, R_2)| \text{ Eqn. (11) } \} \tag{10}$$

$$
\begin{aligned}
R_1 &\leq I(X_1;Y_1|U_2Q) \\
R_2 &\leq I(X_2;Y_2|U_1Q) \\
R_1+R_2 &\leq I(X_1U_2;Y_1|Q)+I(X_2;Y_2|U_1U_2Q) \\
R_1+R_2 &\leq I(X_1;Y_1|U_1U_2Q)+I(X_2U_1;Y_2|Q) \\
R_1+R_2 &\leq I(X_1U_2;Y_1|U_1Q)+I(X_2U_1;Y_2|U_2Q) \\
2R_1+R_2 &\leq I(X_1U_2;Y_1|Q)+I(X_1;Y_1|U_1U_2Q)+I(X_2U_1;Y_2|U_2Q) \\
R_1+2R_2 &\leq I(X_2;Y_2|U_1U_2Q)+I(X_2U_1;Y_2|Q)+I(X_1U_2;Y_1|U_1Q)
\end{aligned} \tag{11}
$$

4 Outer bounds on $\mathcal{C}_{IC}$

In this section we present the best known outer bounds on the capacity region, $\mathcal{C}_{IC}$, of the classical inteference channel.

4.1 Naive bound

The first bound, comes from the relationship to the multiple access channel. In particular we can think of the interference channel as two separate MACs – one with Y_1 as output and another one with Y_2 as output. The probability distributions for these MACs are $p(y_1|x_1,x_2) = \sum_{y_2} p(y_1,y_2|x_1,x_2)$ and $p(y_2|x_1,x_2) = \sum_{y_1} p(y_1,y_2|x_1,x_2)$ respectively.

Note that if some rate R_1 is achievable for the IC, then it must also be achievable for the Y_1-MAC channel since it is a special case of the IC. We define the following "rectangle" region:

$$G_{Naive} \triangleq \text{conv} \cup_{p(x_1)p(x_2)} \{(R_1,R_2)| \text{ Eqn. (13) }\} \tag{12}$$

$$
\begin{aligned}
R_1 &\leq I(X_1;Y_2|X_2) \\
R_2 &\leq I(X_2;Y_2|X_1)
\end{aligned} \tag{13}
$$

The region defined above forms our first outer bound on the IC rate region.

Theorem 1. *The region G_{Naive} defined in* (12) *is an outer bound on $\mathcal{C}_{IC}$, i.e.*

$$\mathcal{C}_{IC} \subset G_{Naive}. \tag{14}$$

Proof. Suppose, for contradiction, that some rate $R_1^* > I(X_1;Y_2|X_2)$ is achievable for the IC. As pointed out above, this means that it must also be achievable for the Y_1-MAC channel.

But we know from equation (3) the exact capacity region for the MAC channel, and rate R_1^* lies clearly outside of that region, so it must not have been achievable for the IC in the first place.

A similar argument can be applied for the second inequality in relation to the Y_2-MAC channel.

□

This bound is in general very loose. Note however that this bound is achievable for an IC that is independent, i.e. one with probability distribution

$$p(y_1,y_2|x_1,x_2) = p_i(y_1,y_2|x_1,x_2) = p(y_1|x_1,x_2)p(y_2|x_1,x_2). \tag{15}$$

4.2 Sato's BC bound

We can describe a more precise outer bound to the capacity region by specifying an inequality on the sum rate $R_1 + R_2$. This was done by Hiroshi Sato [Sat77] by adapting his results on the broadcast channel [Sat78].

Consider a IC with probability transition matrix $p(y_1, y_2|x_1, x_2)$, we define the Sato region associated with that channel as follows.

$$G_{Sato}(p) \triangleq \operatorname{conv} \cup_{p(x_1)p(x_2)} \{(R_1, R_2)| \text{ Eqn. (17) } \} \tag{16}$$

$$\begin{aligned} R_1 &\leq I(X_1; Y_1|X_2), \\ R_2 &\leq I(X_2; Y_2|X_1), \\ R_1 + R_2 &\leq I(X_1X_2; Y_1Y_2). \end{aligned} \tag{17}$$

We now state two adaptations of the BC capacity outer bounds which apply directly to the interference channel as pointed out in [Sat77].

Theorem 2 (Theorem 1 in [Sat77])**.** *The region G_{Sato} defined in* (16) *is an outer bound on $\mathcal{C}_{IC}$, i.e.*

$$\mathcal{C}_{IC} \subset G_{Sato}(p). \tag{18}$$

Proof. To obtain the first two inequalities, we use the same argument as in the proof of G_{Naive}. The third inequality follows from the consideration of a MAC channel where the output is $Y = (Y_1, Y_2)$ which gives us the bound $R_1 + R_2 \leq I(X_1X_2; Y)$.

Combining the two outputs Y_1 and Y_2 is equivalent to allowing joint decoding for the interference channel. Since the joint-decoding channel is more powerful than the original IC, any outer bound on the joint-decoding IC is also an outer bound on $\mathcal{C}_{IC}$. □

Furthermore, using the fact that the IC capacity $\mathcal{C}_{IC}$ is the same for all channels in the same-marginals class $K(p)$ defined in equation (5), we can take the intersection of G_{Sato} over all channels in that class.

Theorem 3 (Theorem 2 in [Sat77])**.**

$$\mathcal{C}_{IC} \subset \bigcap_{q \in K(p)} G_{Sato}(q). \tag{19}$$

Proof. The IC capacity is the same for all channels in $K(p)$, so we might as well pick the worst case in $K(p)$, which is exactly what the intersection operation accomplishes. □

4.3 Carleial

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

Consider the two random variables Z_1, Z_2 such that

$$\begin{aligned} &Y_1 \quad \text{is a degraded version of} \quad Z_1, && (20)\\ &Y_2 \quad \text{is a degraded version of} \quad Z_2, && (21)\\ &Y_2 \quad \text{is a degraded version of} \quad (X_1, Z_1), && (22)\\ &Y_1 \quad \text{is a degraded version of} \quad (X_2, Z_2), && (23) \end{aligned}$$

then we have the following outer bound

$$G_{Carl} \triangleq \text{conv} \cup_{p(x_1)p(x_2)} \{(R_1, R_2)| \text{ Eqn. (25) } \} \qquad (24)$$

$$\begin{aligned} R_1 &\leq I(X_1; Y_1|X_2) \\ R_2 &\leq I(X_2; Y_2|X_1) \qquad (25) \\ R_1 + R_2 &\leq \min\{ I(X_1X_2; Z_1) \ , \ I(X_1X_2; Z_2) \ \} \end{aligned}$$

This bound can be shown to be strictly tighter than the Sato bound and is the best known outer bound on the interference channel.

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.

$$\begin{aligned} R_1 &\leq I(U; Y_1) && (26)\\ R_2 &\leq I(V; Y_2) && (27)\\ R_1 + R_2 &\leq I(U; Y_1) + I(V; Y_2|U) && (28)\\ R_1 + R_2 &\leq I(V; Y_2) + I(U; Y_1|V) && (29) \end{aligned}$$

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 $U \rightarrow X_1$ and $V \rightarrow X_2$, 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 multiuser information theory

Quantum information theory (QIT) has been one area of very active research in the past decade. Ever since the B. Schumacher discovery that quantum information can be compressed[Sch95], information theorists and physicists have been working hard to answer the basic questions about source coding, channel capacity and efficient coding schemes.

QIT can be considered in some sense a direct extension of classical information theory. It is not surprising then if many QIT results carry the same flavour the classical strategies. Often times, the rates obtained in a quantum protocol will be of the same form, and in fact the same expressions in terms of mutual information as the classical counterpart (or at least up to a constant $\frac{1}{2}$ factor).

This resemblance at the "formula level" is misleading, however, since the equations come about for very different reasons. Indeed, the properties of quantum information and the von Neumann entropy used to measure it, have some very different characteristics compared to the classical data.

Here is a short list of a few of the particularities quantum information has to offer:

- Quantum information cannot be copied. This is called the *no-cloning* principle.
- Quantum information cannot be destroyed. A quantum system only "loses" information in the sense that it gets mixed up in the environment.
- Entanglement can be used to increase the rates of certain protocols despite being provably useless for communication resource by itself (no-signalling principle from Relativity). (note that classically shared randomness cannot increase the rate of classical protocols)
- von Neumann conditional entropy can be negative. $H(A|B)_\rho < 0$ also has an operational interpretation related to quantum Slepian-Wolf coding.
- Monogamy of entanglement: If we have some tripartite pure state on ABC and we can prove that the A-entanglement is not in system B, then it must be in system C because it must be somewhere.
- Polygamy of purification: In a pure tripartite state on ABC, some subset $\hat{B} \subsetneq B$ can, with hight probability, contain the purification of two *both* A and C.

For a crash-course introduction to the subject, and in particular the notation used the reader can consult Appendix A at this point. For a more in depth introduction to the subject an excellent book is [NC00].

In this section we will review some of the known results in quantum information theory that about the multiple access and broadcast channels. Before we get to that, we must discuss the kinds of rates that one might want to study.

6 Communication tasks

We can use a quantum channel to transmit classical or quantum information. There are therefore two different communication tasks and different capacity regions associated with each task on the same quantum channel $\mathcal{N}$. To complicate things further, we often consider coding strategies, which use an additional communication resource of shared entanglement between sender and receiver.

Thus, there are 4 different capacities:

- Classical data: $\mathcal{C}(\mathcal{N})$
- Quantum data: $\mathcal{Q}(\mathcal{N})$
- Entanglement-assisted classical data: $\mathcal{C}_{\texttt{E-A}}(\mathcal{N})$
- Entanglement-assisted quantum data: $\mathcal{Q}_{\texttt{E-A}}(\mathcal{N})$

A common but not universal convention is to denote rates for the communication of classical information by the R-rates $R_1, R_2, \ldots$. Quantum rates will be written as Q-rates. In cases where we do not consider entanglement to be "free", we will denote by E the rate at which entanglement is being used-up or generated by the protocol.

6.1 Quantum multiple access channels

The quantum MAC problem was first approached in [HZH00]

This work was followed by [Win01]

Theorem 4 (Theorem 10 in [Win01]). *The capacity of the quantum MAC $\mathcal{M}$ to carry classical information is $\mathcal{C}_{QMAC}(\mathcal{M}) = \frac{1}{n}\cup_{n=1}^{\infty}\mathcal{C}_{QMAC}^{(1)}(\mathcal{M}^{\otimes n})$ where*

$$\mathcal{C}_{QMAC}^{(1)} = \text{conv} \cup_{p(x_1)p(x_2)\sigma_x} \{(R_1, R_2) | \textit{ Eqn. } (31)\ \} \tag{30}$$

$$\begin{aligned} R_1 &\leq I(A;C|B)_\theta, \\ R_2 &\leq I(B;C|A)_\theta, \\ R_1 + R_2 &\leq I(AB;C)_\theta, \end{aligned} \tag{31}$$

where mutual informations are calculated with respect to

$$\theta = \sum_x p(x_1)p(x_2)|x_1\rangle\langle x_1|_A \otimes |x_2\rangle\langle x_2|_B \otimes \mathcal{M}(\sigma_{x_1} \otimes \sigma_{x_2})_C. \tag{32}$$

Another set of contributions [YHD08] and the conference version [YDH05].

Recently, the entanglement assisted capacity of the QMAC channel was proved in [HDW08].

Theorem 5 (Hsieh, Devetak, Winter 05). *The capacity of the quantum MAC $\mathcal{M}$ to carry classical information* when an additional resource of entanglement is available is

$$\mathcal{C}_{\texttt{E-A}} = \text{conv} \cup_{p(x_1)p(x_2)\sigma_x} \{(R_1, R_2) | \textit{ Eqn. } (34)\ \} \tag{33}$$

$$\begin{aligned} R_1 &\leq I(A;C|B)_\theta, \\ R_2 &\leq I(B;C|A)_\theta, \\ R_1 + R_2 &\leq I(AB;C)_\theta, \end{aligned} \tag{34}$$

where mutual informations are calculated with respect to

$$WRONG\theta = \sum_x p(x_1)p(x_2)|x_1\rangle\langle x_1|_A \otimes |x_2\rangle\langle x_2|_B \otimes \mathcal{M}(\sigma_x)_C. \tag{35}$$

6.2 Quantum broadcast channels

The quantum BC problem has received comparatively little attention. The problems with the quantum BC are not only technical (proving theorems), but also conceptual (defining the problem). The classical broadcast problem defines three rates. Two personal rates R_1, R_2 for messages intended for Rx1, Rx2 respectively and a rate R_0 for messages M_0 recoverable at both receivers.

In the quantum BC, we can define rates Q_1, Q_2 in analogy with their classical counterparts but since we are not allowed to copy quantum information, how could we possibly deliver a common quantum message $|M_0\rangle$ to both receivers?

This problem of not being able to copy quantum information is not unique in the BC case. In the study of quantum network coding, it is equally difficult to find a quantum analog of the multicast scenario which requires a common message to to be delivered to multiple receivers.

[YHD06]

[DH06].

6.3 Quantum interference channel

The quantum version of the DMIC has not been studied yet in literature.

6.3.1 Definitions and Notation

quantum interference channel $\mathcal{N}^{A_1' A_2' \to B_1 B_2}$ is a completely-positive trace-preserving map. Let $\mathcal{N}^{A_1'^n A_2'^n \to B_1^n B_2^n}$ denote the IID version of this channel:

$$\mathcal{N}^{A_1'^n A_2'^n \to B_1^n B_2^n} \equiv (\mathcal{N}^{A_1' A_2' \to B_1 B_2})^{\otimes n}.$$

trace norm $\|A\|_1 \equiv \text{Tr}\left\{\sqrt{A^\dagger A}\right\}$. Fidelity $F(\rho, \sigma) \equiv \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1^2$.

We begin by outlining a general protocol for entanglement-assisted transmission of quantum information over a quantum interference channel $\mathcal{N}^{A_1' A_2' \to B_1 B_2}$. Two spatially separated senders control the respective input systems A_1' and A_2', and two spatially separated receivers control the respective output systems B_1 and B_2. The channel has an extension to an isometry $U_{\mathcal{N}}^{A_1' A_2' \to B_1 B_2 E}$, where another party Eve has access to system E. The task of sender i is to transmit a 2^{nQ_i}-dimensional quantum system by exploiting some large number n uses of the channel $\mathcal{N}$ and entanglement in the form of 2^{nE_i} ebits shared with receiver i, where $i \in \{1, 2\}$. The goal for receiver i is to decode with high fidelity the quantum state that sender i transmits.

An (n, Q_1, Q_2, ϵ) *entanglement-assisted quantum interference channel code* (EAQIC) consists of four steps: preparation, channel coding, transmission, and channel decoding. We detail each of these steps below.

Preparation. Each sender prepares a general quantum state in register A_i alongside a register T_{A_i} that contains the entanglement shared with receiver i. Let R_i be a quantum register that contains the purification of the state in register A_i. The overall state after preparation is

$$\psi^{R_1 A_1} \otimes \Phi^{T_{A_1} T_{B_1}} \otimes \psi^{R_2 A_2} \otimes \Phi^{T_{A_2} T_{B_2}},$$

Figure 1: The steps in a general entanglement-assisted quantum interference channel code. In the above figure, we employ the shorthands $A_1' \equiv A_1'^n$, $A_2' \equiv A_2'^n$, $B_1 \equiv B_1^n$, $B_2 \equiv B_2^n$.

where it is implicit that the state $\psi^{R_1 A_1}$ can be different from $\psi^{R_2 A_2}$ (the superscript labels uniquely identify the state).

Channel Encoding. Sender i encodes the registers A_i and T_{A_i} according to some CPTP encoding map $\mathcal{E}_i^{A_i T_{A_i} \rightarrow A_i'}$. The state after the encoding maps is

$$\sigma^{R_i A_i'^n T_{B_i}} \equiv \mathcal{E}_i^{A_i T_{A_i} \rightarrow A_i'}(\psi^{R_i A_i} \otimes \Phi^{T_{A_i} T_{B_i}}).$$

The overall state after encoding is

$$\rho^{R_1 A_1'^n T_{B_1} R_2 A_2'^n T_{B_2}} \equiv \sigma^{R_1 A_1'^n T_{B_1}} \otimes \sigma^{R_2 A_2'^n T_{B_2}}.$$

Observe that the state between $R_1 A_1'^n T_{B_1}$ and $R_2 A_2'^n T_{B_2}$ is a product state.

Transmission. Both senders transmit their respective systems $A_1'^n$ and $A_2'^n$ through the IID quantum interference channel $\mathcal{N}^{A_1'^n A_2'^n \rightarrow B_1^n B_2^n}$. The resulting state is

$$\omega^{R_1 B_1^n T_{B_1} R_2 B_2^n T_{B_2}} \equiv \mathcal{N}^{A_1'^n A_2'^n \rightarrow B_1^n B_2^n}(\rho^{R_1 A_1'^n T_{B_1} R_2 A_2'^n T_{B_2}}).$$

Observe that the state between registers $R_1 B_1^n T_{B_1}$ and $R_2 B_2^n T_{B_2}$ is not a product state for general quantum interference channels.

Channel Decoding. The two receivers obtain the respective systems B_1^n and B_2^n from the output of the channel. Receiver i performs a CPTP decoding map $\mathcal{D}_i^{B_i^n T_{B_i} \rightarrow \hat{A}_i}$ that combines the channel output B_i^n with his half T_{B_i} of the entanglement shared with sender i. The state after these decoding maps is

$$\theta^{R_1 \hat{A}_1 R_2 \hat{A}_2} \equiv (\mathcal{D}_1^{B_1^n T_{B_1} \rightarrow \hat{A}_1} \otimes \mathcal{D}_2^{B_2^n T_{B_2} \rightarrow \hat{A}_2})(\omega^{R_1 B_1^n T_{B_1} R_2 B_2^n T_{B_2}}).$$

Figure 1 depicts all of the above steps that occur in a general entanglement-assisted quantum interference channel code.

The conditions for a good entanglement-assisted quantum interference channel code are that both receivers be able to decode the quantum states from the respective senders with small trace distance:

$$\left\| \theta^{R_1 \hat{A}_1 R_2 \hat{A}_2} - \psi^{R_1 A_1} \otimes \psi^{R_2 A_2} \right\|_1 \leq \epsilon.$$

A rate pair (Q_1, Q_2) is *achievable* if there exists an $(n, Q_1 - \delta, Q_2 - \delta, \epsilon)$ entanglement-assisted quantum interference channel code for any $\epsilon, \delta > 0$ and sufficiently large n. The capacity region $C_{\text{QIC}}(\mathcal{N})$ is a two-dimensional region in the (Q_1, Q_2) space, containing all achievable rate pairs (Q_1, Q_2).

There is another way to depict an even more general protocol (that we discussed in the group meeting). Figure 2 depicts this setting.

Figure 2: (Color online) The above figure depicts the alternate setting that we had in our discussion. Each block in the figure is an isometric extension of the corresponding CPTP maps from the previous figure. Thus, the state is pure at each point in the protocol.

7 Discussion and conclusion

In this seciton we will make reference to some of the generalizations and open problems associated with the study of interference channels.

What is the most general use of the interference network we can think of?

six rates....

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?

$$\begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \tag{36}$$

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?)

In general the joint source-channel coding is a very interesting problem that remains...

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 (Diract $\delta(x)$ 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 8 (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: $\text{Tr}_R \left(|\sigma\rangle\langle\sigma|^{AR}\right) = \sigma^A$. 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 9 (von Neumann Entropy)**.** *Given the density matrix $\rho^A \in \mathcal{H}^A$, the expression*

$$S(A)_\rho = -\text{Tr}\left(\rho^A \log \rho^A\right) \tag{37}$$

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 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

$$H(A|B)_\rho := H(AB)_\rho - H(B)_\rho \tag{38}$$

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

$$I(A;B)_\rho := H(A)_\rho + H(B)_\rho - H(AB)_\rho \tag{39}$$

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

$$\begin{aligned} I(A;B|C)_\rho &:= H(A|C)_\rho + H(B|C)_\rho - H(AB|C)_\rho && (40) \\ &= H(AC)_\rho + H(BC)_\rho - H(ABC)_\rho - H(C)_\rho. && (41) \end{aligned}$$

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

$$H(AC) + H(BC) \geq H(C) + H(ABC). \tag{42}$$

[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.

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 6. *Consider the Φ^+ Bell state*

$$|\Phi\rangle^{AB} = \tfrac{1}{\sqrt{2}}(|00\rangle^{AB} + |11\rangle^{AB}). \tag{43}$$

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

$$H(A)_\Phi = 1, \qquad H(B)_\Phi = 1, \qquad H(AB)_\Phi = 0, \tag{44}$$

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

$$H(A|B) = H(AB) - H(B) = -1 \text{ [bits]}, \tag{45}$$

and the mutual information

$$I(A;B) = H(A) + H(B) - H(AB) = 2 \text{ [bits]}. \tag{46}$$

Equation (45) 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 (46), 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 $\mathcal{D}$:

$$\rho \xrightarrow{\mathcal{E}} \rho' \qquad \text{or} \qquad \mathcal{E}(\rho) = \rho' \tag{47}$$

Unitary transformations are a type of quantum operation

$$\mathcal{E}(\rho) = U\rho U^\dagger \tag{48}$$

where U is a unitary matrix. Unitary operations cooorespond 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

$$\mathcal{E}(\rho) = \mathrm{Tr}_{\mathrm{env}} U \left(\rho \otimes \rho_{\mathrm{env}}\right) U^{\dagger} \tag{49}$$

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 $\mathcal{E}:\mathcal{H}^A \to (\mathcal{H}^A, \mathbb{N})$ 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

$$p(m) = \mathrm{Tr}\left(M_m^{\dagger} \rho M_m\right). \tag{50}$$

The output quantum state associated with this outcome is

$$\widetilde{\rho}_m = \frac{M_m^{\dagger} \rho M_m}{\mathrm{Tr}\left(M_m^{\dagger} \rho M_m\right)}. \tag{51}$$

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 [Sha48]:

$$\{c \to c\} \geq \max_{p(x)} I(X;Y)[c \to c], \tag{52}$$

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

$$|\Phi\rangle^{AB} = \tfrac{1}{\sqrt{2}}(|00\rangle^{AB} + |11\rangle^{AB}), \tag{53}$$

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

$$F(|\varphi\rangle, |\psi\rangle) = |\langle\varphi|\psi\rangle|^2 . \tag{54}$$

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

$$F(\rho,\sigma) = \mathrm{Tr}\left(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)^2 . \tag{55}$$

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:

$$\mathcal{N}(|\psi\rangle\langle\psi|) = \sigma^{\widehat{A}}. \tag{56}$$

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

$$\overline{F}(\{p_i, |\psi_i\rangle\}, \mathcal{N}) := \sum_i p_i F(|\psi_i\rangle, \sigma_i), \qquad \sigma_i = \mathcal{N}(|\psi_i\rangle\langle\psi_i|). \tag{57}$$

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

$$\overline{F}(\mathcal{E},\mathcal{N}) \geq 1-\epsilon \tag{58}$$

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

$$\mathcal{N}^{A\to\widehat{A}} \otimes \mathrm{id}^R\big(|\psi\rangle\langle\psi|^{AR}\big) = \sigma^{\widehat{A}R}. \tag{59}$$

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

$$F(|\psi\rangle^{AR}, \sigma^{\widehat{A}R}) = \langle\psi^{AR}|\sigma^{\widehat{A}R}|\psi^{AR}\rangle \geq 1-\epsilon. \tag{60}$$

Equation (60) 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 [Sch96] 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 (60) implies equation (58).

B Compression

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

$$\text{Tr}\left[\rho^{\otimes n} P_\epsilon^{(n)}\right] > 1 - \delta \qquad \forall \delta, \epsilon > 0, \text{and } n \text{ sufficiently large.} \tag{61}$$

$$\text{rank} P_\epsilon^{(n)} = |T_\epsilon^{(n)}| \tag{62}$$

$$\forall \epsilon > 0, \quad |T_\epsilon^{(n)}| \leq 2^{n[H(\rho)+\epsilon]} \tag{63}$$

$$\forall \epsilon, \delta > 0, \quad |T_\epsilon^{(n)}| \geq (1-\delta) 2^{n[H(\rho)-\epsilon]} \tag{64}$$

Proof. The proof of (62) is easy.
For (63) 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 (61) for which we note:

$$\text{Tr}\left[\rho^{\otimes n} P_\epsilon^{(n)}\right] = \sum_{j^n} \langle j^n| \left(\sum_{k^n} r_{k^n} |k^n\rangle\langle k^n| \right) \sum_{l^n \in T_\epsilon^{(n)}} |l^n\rangle\langle l^n||j^n\rangle \tag{65}$$

□

Holevo, Schumacher and Westmoreland found a multiletter formula for the classical capacity [H98,SW97].

Theorem 7. *The classical capacity of a quantum channel $\mathcal{N}$ is given by $\mathcal{C}(\mathcal{N}) = \frac{1}{n} \cup_{n=1}^{\infty} C^{(1)}(\mathcal{N}^{\otimes n})$ where*

$$C^{(1)}(\mathcal{N}) = \max_\rho I(A;B) \tag{66}$$

where ρ is the output state

$$\rho = \sum_x p(x)|x\rangle\langle x|_A \otimes \mathcal{N}(\sigma_x)_B. \tag{67}$$

and $\{p(x), \sigma_x\}$ is the input distribution.

quantum capacity Lloyd, Shor and Devetak independently proved a formula for the quantum capacity of a channels.

Theorem 8 (LSD Theorem)**.** *Consider the input state $\rho^{AA'}$, half of which is sent through the channel $\mathcal{N}$ to obtain $w^{AB} = \mathcal{N}^{A' \to B}(\rho^{AA'})$. Define the quantity*

$$Q^{(1)}(\mathcal{N}) = \max_\rho (S(B)_w - S(AB)_w). \tag{68}$$

The quantum capacity Q of the channel $\mathcal{N}$ is given by the regularization of the $Q^{(1)}$ quantity

$$Q(\mathcal{N}) = \lim_{n \to \infty} \frac{Q^{(1)}(\mathcal{N}^{\otimes n})}{n}. \tag{69}$$

[L97, S00, D03]

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.

[DH06] Frédéric Dupuis and Patrick Hayden. A father protocol for quantum broadcast channels. arXiv:quant-ph/0612155, December 2006.

[HDW08] Min-Hsiu Hsieh, Igor Devetak, and Andreas Winter. Entanglement-assisted capacity of quantum multiple-access channels. *IEEE Transactions on Information Theory*, 54(7):3078–3090, 2008.

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

[HZH00] M. Huang, Y. Zhang, and G. Hou. Classical capacity of a quantum multiple-access channel. *Physical Review A*, 62(5):52106, 2000.

[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.

[Sch95] B. Schumacher. Quantum coding. *Phys. Rev. A*, 51:2738–2747, 1995. doi:10.1103/PhysRevA.51.2738.

[Sch96] B. Schumacher. Sending entanglement through noisy quantum channels. *Phys. Rev. A*, 54:2614–2628, 1996. arXiv:quant-ph/9604023.

[Sha48] C. E. Shannon. A mathematical theory of communication. *Bell Sys. Tech. Journal*, 27:379–423,623–656, 1948.

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

[Win01] A. Winter. The capacity of the quantum multiple-access channel. *IEEE Transactions on Information Theory*, 47(7):3059–3065, 2001.

[YDH05] J. Yard, I. Devetak, and P. Hayden. Capacity theorems for quantum multiple access channels. In *Information Theory, 2005. ISIT 2005. Proceedings. International Symposium on*, pages 884–888, 2005.

[YHD06] Jon Yard, Patrick Hayden, and Igor Devetak. Quantum broadcast channels. arXiv:quant-ph/0603098, March 2006.

[YHD08] Jon Yard, Patrick Hayden, and Igor Devetak. Capacity theorems for quantum multiple-access channels: Classical-quantum and quantum-quantum capacity regions. *IEEE Transactions on Information Theory*, 54(7):3091–3113, July 2008.