Outer bounds on the quantum interference channel

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

April 22, 2010

Abstract

In the first part of this work, we present a comprehensive review of results about discrete memoryless interference channels. In the second part, we define the quantum multiple access and broadcast channels and state some of the known results about them. Finally we attempt to use these MAC and BC results to derive an outer bound on the quantum interference channel.

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 central importance to theoretical and practical information theory. Despite the numerous efforts, there are very few results about discrete memoryless interference channels (DMIC) known to date. The capacity region is not known even for the special case of Gaussian interference channels.

Recently, quantum information theory (QIT) has been a popular research area in physics, computer science and mathematics. This new field of information theory is firmly based in Shannon's principle of information-as-statistics, but studies systems that behave according to the laws of quantum mechanics. The current state of QIT is similar to that of classical information theory in the 80s, when most single user problems had been solved and researchers were exploring new grounds in multiuser scenarios. It is therefore an appropriate time to attack the quantum interference channel.

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 classical interference channels. Then in section 3 we will reproduce the main results from a series of seminal papers on the interference channel. Results relating to outer bounds on the IC capacity are collected together in section 4. The following section (section 5) deals with the quantum generalization of the multiple access and broadcast channels. In order to make this work more self-contained, we have included in appendix A a detailed introduction to the ideas of quantum information theory.

Readers not familiar with quantum information theory are encouraged to consult appendix A before reading section 5. Finally we have a section of discussion and pointers to open problems in section 6.

2 Preliminaries

In this section 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*

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

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 $(X_1, X_2) \rightarrow Z \rightarrow Y'$ form a Markov chain.

2.2 Convex sets, closures and optimization constraints

Rate regions in multiuser information theory are 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 it with a well known result [Sha48].

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

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

The $p(x)$ 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 equalities $R \leq I(X;Y)$ for our rate constraints.

2.3 Multiple access channel

The multiple access channel (MAC) is a special case of the interference channel which has 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.

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

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

The coding strategy that makes these rates possible is *successive decoding*. The 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. Note that the rates $R_1 = I(X_1;Y|X_2)$ and $R_2 = I(X_2;Y|X_1)$ are not simultaneously achievable.

2.4 Broadcast channel

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

The first thing to note is that the error criterion in equation (1) of definition 3 only depends on the marginals $p_1(y_1|x_1x_2)$ and $p_2(y_2|x_1x_2)$. This is because if we manage to get both decoding errors low, then we manage to get the error of the union of the two events low also. This observation leads us to the fact 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 the same marginals as $p(y_1, y_2|x_1x_2)$:

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

Thus the capacity region $\mathcal{C}_{IC}$ for an IC channel with probability transition matrix 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 analyze

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. 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 (section 6) at the end of this document.

3 Classical results on the Interference Channel

In this section we present a literature review of important papers on the interference channel.

3.1 Ahlswede 74: 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 in those days, the notation had not yet been standardized so the author uses the definitions of entropic quantities from first principles like this one

$$R_1(p, q, Y_1) \triangleq \sum_{x_1,x_2,y_1} p(x_1)q(x_2)p(y_1|x_1, x_2) \log \frac{p(y_1|x_1, x_2)}{\sum_{x_1} p(x_1)p(y_1|x_1, x_2)}, \tag{7}$$

which stands for $I(X_1; Y_1|X_2)$ in modern notation. Seeing such notation makes me glad I live in the modern era of information theory.

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

- The **MAC** to receiver Y_1 with probability distribution is $\sum_{y_2} p(y_1, y_2|x_1, x_2)$. Ahlswede calls this the (p, T_{21}, I) *communication situation*.
- The **IC** problem which we have defined above. In Ahlswede's nomenclature this is a (p, T_{22}, I) communication 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 first result in this paper is an alternate proof of the Y_1-MAC capacity region. We already stated capacity region $\mathcal{C}_{MAC}$ in subsection 2.3.

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

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

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

Note that the optimization is done over joint distributions $p(x_1, x_2)$ and thus does not quite fit into the paradigm of independent messages which we are interested in. Indeed, the author makes a remark in the conclusion of the paper, that knowing the capacity of the MMAC problem does not help us directly to find the capacity of the IC channel.

3.2 Sato 77: Two-user communication channels

Perhaps the most important of the early papers on the interference channel is by Hiroshi Sato [Sat77]. In this paper the IC problem is defined and the relations between it and the MAC and BC are outlined. Using these relations, the author proves a several progressively stronger outer bounds the best of which is presented in equation (17) in the next section.

This paper is where degraded IC make their first appearance. The author gives an inner bound G_D for the degraded IC, but this bound is not equal to the outer bound. The Sato outer bound and the inner bound G_D do coincide, however, for a special case of degraded channels called *twin channels*.

3.3 Carleial 78: Interference channels

The paper by Carleial is yet another comprehensive discussion on interference channels [Car78]. This paper proves the naive outer bound on the IC (see equation (13) below), and restates the Sato outer bound (17) which is tighter.

The author also proves an inner bound counterpart to the naive outer bound, which we will call the naive inner bound

$$G_{ach} \triangleq \{(R_1, R_2) \in \text{conv}\{P_1, P_2, O\}\}, \tag{10}$$

where the three points are defined as

$$\begin{aligned} P_1 &= (I(X_1; Y_2|X_2),\ 0), \\ P_2 &= (0,\ I(X_2; Y_2|X_1)), \\ O &= (0, 0). \end{aligned}$$

Theorem 1 (Theorem 1 in [Car78])**.** *The capacity region $\mathcal{C}_{IC}$ satisfies $G_{ach} \subseteq \mathcal{C}_{IC}$.*

Proof. Consider the Y_1-MAC sub-channel of the IC. We know from the result on the MAC channel (equation (3)) that the rate $R_1 = I(X_1; Y_2|X_2)$ is achievable for the Y_1-MAC. Thus the point $P_1 = (I(X_1; Y_2|X_2),\ 0)$ is achievable for the IC. A similar argument is applied to show P_2 is achievable using the Y_2-MAC sub-channel. By the time sharing principle, we can achieve any rate in the convex hull of these points. □

Carleial also generalizes the notion of *twin channels* which he calls "interference channels with statistically equivalent outputs" for which he calculates the capacity region. These channels are simple to analyze because both receivers Y_1 and Y_2 obtain the same information and thus the IC problem reduces to the MAC problem.

Using his result on the twin channel he gives loose inner and outer bounds for degraded ICs. To get an inner bound, he creates a twin channel from the degraded channel by discarding part of the information at the better receiver. To obtain the outer bound, he uses the reasoning that all degraded outputs have less information than the best channel output, thus the capacity of the twin channel with copies of the best output forms an outer bound on the degraded IC capacity.

The paper then discusses some examples: binary ICs and Gaussian ICs. Even for these simplistic examples of ICs we cannot calculate the capacity region. In the section dealing with Gaussian ICs, the author sows the seeds of two important ideas which would later be taken up by others and proven for general ICs. The first of these ideas is the notion of successive decoding which would later be used by Han and Kobayashi for the general DMIC [HK81]. The second idea is the notion of channels with *strong interference* where the interference in the channel is so bad that it is easier for a receiver to decode the interfering message (intended for the other receiver) than the one intended for him.

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

The best achievable region for the interference channel is due to Han and Kobayashi [HK81]. By splitting the message M_1 intended for Rx1 into public and personal parts, M_{10} and M_{11} respectively and using successive decoding at the receivers we can achieve a rate region G_{HK} described by 14 inequalities.

Recently Cheng, Motani and Garg simplified the H-K coding strategy and relaxed some of the decoding requirements in order to obtain a simpler achievable rate region where the rates are bounded by only 7 inequalities [CMGEG08]. A further contribution by Han and Kobayashi shows that the two regions are in fact equivalent [HK07].

Without further ado, we present the HK-CMG rate region, which we will call this way to give credit to all.

$$G_{HK} = G_{CMG} \triangleq \text{conv} \bigcup_{p(x_1|u_1)p(x_2|u_2)p(u_1)p(u_2)} \{(R_1, R_2) | \text{ Eqn. (12) } \} \tag{11}$$

$$\begin{aligned}
R_1 &\leq I(X_1; Y_1|U_2) \\
R_2 &\leq I(X_2; Y_2|U_1) \\
R_1 + R_2 &\leq I(X_1U_2; Y_1) + I(X_2; Y_2|U_1U_2) \\
R_1 + R_2 &\leq I(X_1; Y_1|U_1U_2) + I(X_2U_1; Y_2) \\
R_1 + R_2 &\leq I(X_1U_2; Y_1|U_1) + I(X_2U_1; Y_2|U_2) \\
2R_1 + R_2 &\leq I(X_1U_2; Y_1) + I(X_1; Y_1|U_1U_2) + I(X_2U_1; Y_2|U_2) \\
R_1 + 2R_2 &\leq I(X_2; Y_2|U_1U_2) + I(X_2U_1; Y_2) + I(X_1U_2; Y_1|U_1)
\end{aligned} \tag{12}$$

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

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

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:

$$G_{Naive} \triangleq \text{conv} \bigcup_{p(x_1)p(x_2)} \{(R_1, R_2) | \text{ Eqn. (14) } \} \tag{13}$$

$$\begin{aligned}
R_1 &\leq I(X_1; Y_2|X_2), \\
R_2 &\leq I(X_2; Y_2|X_1).
\end{aligned} \tag{14}$$

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

Theorem 2. *The region G_{Naive} defined in* (13) *is an outer bound on $\mathcal{C}_{IC}$, i.e.*

$$\mathcal{C}_{IC} \subset G_{Naive}. \tag{15}$$

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 is applied for the second inequality using the Y_2-MAC sub-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

$$p(y_1, y_2|x_1, x_2) = p_i(y_1, y_2|x_1, x_2) = p(y_1|x_1, x_2)p(y_2|x_1, x_2). \tag{16}$$

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.

$$G_{Sato}(p) \triangleq \text{conv} \bigcup_{p(x_1)p(x_2)} \{(R_1, R_2)| \text{ Eqn. (18) } \} \tag{17}$$

$$\begin{aligned} R_1 &\leq I(X_1; Y_1|X_2), \\ R_2 &\leq I(X_2; Y_2|X_1), \\ R_1 + R_2 &\leq I(X_1X_2; Y_1Y_2). \end{aligned} \tag{18}$$

We now state two adaptations of the BC capacity outer bounds which apply directly to the interference channel.

Theorem 3 (Theorem 1 in [Sat77])**.** *The region G_{Sato} defined in* (17) *is an outer bound on $\mathcal{C}_{IC}$, i.e.*

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

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

$$\mathcal{C}_{IC} \subset \bigcap_{q \in K(p)} G_{Sato}(q). \tag{20}$$

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

$$\begin{aligned} Y_1 &\quad \text{is a degraded version of} \quad Z_1, && (21)\\ Y_2 &\quad \text{is a degraded version of} \quad Z_2, && (22)\\ Y_2 &\quad \text{is a degraded version of} \quad (X_1, Z_1), && (23)\\ Y_1 &\quad \text{is a degraded version of} \quad (X_2, Z_2), && (24) \end{aligned}$$

then we have the following outer bound

$$G_{Carl} \triangleq \text{conv} \cup_{p(x_1)p(x_2)} \{(R_1, R_2) | \text{ Eqn. (26) } \} \tag{25}$$

$$\begin{aligned} R_1 &\leq I(X_1; Y_1|X_2) \\ R_2 &\leq I(X_2; Y_2|X_1) \\ R_1 + R_2 &\leq \min\{\, I(X_1X_2; Z_1)\ ,\ I(X_1X_2; Z_2)\, \} \end{aligned} \tag{26}$$

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

5 Quantum multiuser information theory

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

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 as their classical counterpart. Often times, the rates obtained in a quantum protocol will be bounded by the same expressions in terms of mutual information as in the classical result (up to a constant factor of $\frac{1}{2}$).

This resemblance at the "formula level" is misleading, however, since very often 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 classical information.

Here is a short list with 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. This is in fact not strictly a quantum property, even classical laws of physics are reversible at the micro-level.
- Entanglement can be used to increase the rates of certain protocols despite being provably useless as a communication resource by itself (no-signalling principle from Special Relativity). (classically shared randomness cannot increase the rate of IT protocols)
- von Neumann conditional entropy can be negative, $H(A|B)_\rho < 0$, and this has an operational interpretation related to quantum Slepian-Wolf coding.
- Monogamy of entanglement: If we have some tripartite pure state on ABC, then the more entangled A is with B, the less entangled it can be with C.
- 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.
- Quantum Error Correcting Codes exist. Despite the fact that quantum information is in some sense continuous, we can still *stabilize* it using a finite set of operations just like classical ECC.

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 see [NC00].

In this section we will review some of the known results in quantum information theory 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.

5.1 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 for 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: $\mathcal{Q}_{\texttt{E-A}}(\mathcal{N})$

Suppose we have coding strategy that makes n uses of the channel $\mathcal{N}$. A common but not universal convention is to denote rates for the communication of classical information by the letter R (number of bits sent per channel use), and the quantum rates by the letter Q. Quantum rates measure the dimension of the system ψ that we managed to send across the channel so $Q = \frac{1}{n} \log \dim \psi$. 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.

5.2 Quantum multiple access channels

A quantum multiple access channel is a map $\mathcal{M}^{A'B' \to C}$ that takes two quantum systems as inputs and produces a third quantum system.

The quantum MAC was first studied in [HZH00], where the authors investigated the classical capacity. The authors, however, studied the classical capacity in a paradigm which allows for correlations between the input states to the channel. This approach is similar to Ahlswede's correlated input approach to the MMAC problem [Ahl74].

The classical capacity $\mathcal{C}_{QMAC}(\mathcal{M})$ of the quantum MAC was not proven until one year later by Winter [Win01].

Theorem 5 (Theorem 10 in [Win01])**.** *The capacity of the quantum MAC $\mathcal{M}$ to carry classical information is $\mathcal{C}_{QMAC}(\mathcal{M}) = \frac{1}{n} \bigcup_{n=1}^{\infty} \mathcal{C}^{(1)}_{QMAC}(\mathcal{M}^{\otimes n})$ where*

$$\mathcal{C}^{(1)}_{QMAC} = \text{conv} \bigcup_{p(x_1)p(x_2)\sigma_{x_1}\sigma_{x_2}} \{(R_1, R_2) | \textit{ Eqn. (28) } \} \tag{27}$$

$$\begin{aligned} R_1 &\leq I(A;C|B)_\theta, \\ R_2 &\leq I(B;C|A)_\theta, \\ R_1 + R_2 &\leq I(AB;C)_\theta, \end{aligned} \tag{28}$$

where mutual informations are calculated with respect to

$$\theta = \sum_x p(x_1)p(x_2)|x_1\rangle\langle x_1|_A \otimes |x_2\rangle\langle x_2|_B \otimes \mathcal{M}(\sigma_{x_1} \otimes \sigma_{x_2})_C. \tag{29}$$

This paper by Winter is also interesting for our purposes because it is the only place in literature where the quantum interference channel is mentioned. The author writes that he is not sure how to extend his coding strategy to the case of multiple decoders.

The next set of contributions on the quantum MAC problem are by Yard et al [YHD08] and the conference version [YDH05]. In these papers, the authors give regularized capacity regions for the classical-quantum capacity region $\mathcal{CQ}(\mathcal{M})$, where one of the inputs to the MAC is classical and the other is quantum.

The authors also give a multi-letter formula for $\mathcal{Q}(\mathcal{M})$, the quantum capacity region for the MAC.

Theorem 6 (Theorem 2 in [YDH05]). *The capacity of the quantum MAC $\mathcal{M}^{A'B'\to C}$ to carry quantum information is $\mathcal{Q}_{QMAC}(\mathcal{M}) = \frac{1}{n}\bigcup_{n=1}^{\infty}\mathcal{Q}^{(1)}_{QMAC}(\mathcal{M}^{\otimes n})$ where*

$$\mathcal{Q}^{(1)}_{QMAC}(\mathcal{M}) = \operatorname{conv}\bigcup_{p(x_1)p(x_2)\sigma_{x_1}\sigma_{x_2}}\{(Q_1,Q_2)|\ \textit{Eqn. (31)}\ \} \tag{30}$$

$$\begin{aligned} Q_1 &\leq I_c(A\rangle BC)_\omega, \\ Q_2 &\leq I_c(B\rangle AC)_\omega, \\ Q_1+Q_2 &\leq I_c(AB\rangle C)_\omega, \end{aligned} \tag{31}$$

where mutual informations are calculated with respect to the state

$$\omega^{ABC} = \mathrm{id}^{AB}\otimes\mathcal{M}(\Psi_1\otimes\Psi_2), \tag{32}$$

and the states $|\Psi_1\rangle^{AA'}$ and $|\Psi_1\rangle^{BB'}$ are pure.

Recently, the entanglement assisted capacity of the QMAC channel was proved by Hsieh, Devetak and Winter in [HDW08].

Theorem 7 (Theorem 2 in [HDW08]). *Consider the quantum MAC $\mathcal{M}^{A'B'\to C}$. For some states $\rho_1^{A'}$ and $\rho_2^{B'}$ define the output of the channel*

$$\theta^{ABC} = \mathrm{id}^{AB}\otimes\mathcal{M}(\varphi_1^{AA'}\otimes\varphi_1^{BB'}), \tag{33}$$

where $|\varphi_1\rangle^{AA'}$ and $|\varphi_1\rangle^{BB'}$ are purifications of $\rho_1^{A'}$ and $\rho_2^{B'}$ respectively. Define the two dimensional region

$$\mathcal{C}^{(1)}_{\text{E-A}}(\mathcal{M}) = \operatorname{conv}\bigcup_{\rho_1,\rho_2}\{(R_1,R_2)|\ \textit{Eqn. (35)}\ \} \tag{34}$$

$$\begin{aligned} R_1 &\leq I(A;C|B)_\theta, \\ R_2 &\leq I(B;C|A)_\theta, \\ R_1+R_2 &\leq I(AB;C)_\theta. \end{aligned} \tag{35}$$

Then the entanglement-assisted capacity region $\mathcal{C}_{\text{E-A}}(\mathcal{M})$ is given by the regularized expression

$$\mathcal{C}_{\text{E-A}}(\mathcal{M}) = \overline{\bigcup_{n=1}^{\infty}\mathcal{C}^{(1)}_{\text{E-A}}(\mathcal{M}^{\otimes n})}. \tag{36}$$

Note that by using the teleportation protocol we can obtain the quantum entanglement assisted rate region $\mathcal{Q}_{\text{E-A}}(\mathcal{M})$ which will be of the same form but each of the inequalities in (35) will have a $\frac{1}{2}$ in front of it.

5.3 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 message 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.

The first result on the quantum BC is by Yard, Hayden and Devetak [YHD06]. They define three problems and prove a multi-letter capacity region for each of them. The first of these problems involves a special kind of channel that does not really fit well with the rest of QIT. This channel has a classical input and a quantum output.

The second result concerns the classical-quantum capacity $\mathcal{CQ}(\mathcal{N})$ for a quantum channel $\mathcal{N}^{A'\to BC}$ where the sender (Alice) wants to send quantum information to Rx1 (Bob) at a rate Q while at the same time broadcasting a common message $M_0 \in \{1, 2, \ldots, 2^{nR_0}\}$ to both receivers. The authors give a proof of a multi-letter capacity region for this problem.

The third result of that paper considers a true quantum broadcast scenario. In order to circumvent the difficulties of copying quantum information outlined above, the authors give a protocol which establishes GHZ states of the form

$$|GHZ\rangle^{ABC} = \frac{1}{\sqrt{2}}\left(|000\rangle^{ABC} + |111\rangle^{ABC}\right), \tag{37}$$

which are a form of tripartite entanglement. At a later point in time, this communication resource can be used to establish quantum communication between any two of three parties. Thus, generating this GHZ is equivalent to being able to communicate with either of the two receivers (but not both) and we can decide whom we wish to communicate with later on.

The second paper on the quantum broadcast channel is by Dupuis[1] and Hayden [DH06]. In this paper the authors reproduce the classical Marton achievable rates for the entanglement assisted quantum capacity $\mathcal{Q}_{\texttt{E-A}}(\mathcal{N})$.

Theorem 8 (Theorem 2 in [DH06])**.** *Let $\mathcal{N}^{A'\to B_1B_2}$ be a quantum broadcast channel. Then the following rate region is achievable for $|\psi\rangle^{A_1A_2B_1B_2DE} = U_{\mathcal{N}}^{A'\to B_1B_2E}|\phi\rangle^{A_1A_2A'D}$, where $|\phi\rangle$ is any pure state:*

$$\begin{aligned} Q_1 &\leq \frac{1}{2}I(A_1;B_1)_\psi, \\ Q_2 &\leq \frac{1}{2}I(A_2;B_2)_\psi, \\ Q_1+Q_2 &\leq \frac{1}{2}\left[I(A_1;B_1)_\psi + I(A_2;B_2)_\psi - I(A_1;A_2)_\psi\right]. \end{aligned} \tag{38}$$

Q_1 is the rate at which Alice sends qubits to Bob 1, and likewise Q_2 for Bob 2.

[1] Frédéric Dupuis used to sit just across the room from me until last summer.

The authors also prove a multi-letter converse for the same rate region and show how this result can be adapted to give an inner bound for the quantum capacity $\mathcal{Q}(\mathcal{N})$ without entanglement assistance.

5.4 Quantum interference channel

In this section we will define clearly the quantum interference channel problem. The quantum IC is an open problem in quantum information theory and has not been attempted yet. Hopefully, this project will be the first setting stone in that direction.

Definition 8. *The quantum interference network $\mathcal{N}^{A_1'A_2'\to B_1B_2}$ is a completely-positive trace-preserving map that takes two quantum inputs and produces two quantum outputs.*

Definition 9. *The quantum interference channel problem with entanglement assistance is a particular use of the interference network $\mathcal{N}^{A_1'A_2'\to B_1B_2}$ such that two independent quantum states ψ^{A_1}, ψ^{A_2} are transmitted faithfully to their respective receiver at rates (Q_1, Q_2), possibly using entanglement between sender and receiver in the process.*

We will now outline a general protocol for entanglement-assisted transmission of quantum information over the quantum interference network $\mathcal{N}^{A_1'A_2'\to B_1B_2}$. Two spatially separated senders control the respective input systems A_1' and A_2', and two spatially separated receivers will be given the respective output systems B_1 and B_2. The channel has an extension to an isometry $U_{\mathcal{N}}^{A_1'A_2'\to B_1B_2E}$, where E models the environment. 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) can be described in five steps: the inputs, the encoding, the transmission, the decoding and finally the error analysis. We detail each of these steps below.

1. **Input.** Each sender given the quantum state which is to be sent in register A_i alongside an auxiliary 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 input state is
$$\psi^{R_1A_1} \otimes \Phi^{T_{A_1}T_{B_1}} \otimes \psi^{R_2A_2} \otimes \Phi^{T_{A_2}T_{B_2}}.$$

2. **Encoding.** Sender i encodes the registers A_i and T_{A_i} according to some CPTP encoding map $\mathcal{E}_i^{A_iT_{A_i}\to A_i'}$. The state after the encoding maps is
$$\sigma^{R_iA_i'^nT_{B_i}} \equiv \mathcal{E}_i^{A_iT_{A_i}\to A_i'}(\psi^{R_iA_i} \otimes \Phi^{T_{A_i}T_{B_i}}).$$
The overall state after encoding is
$$\rho^{R_1A_1'^nT_{B_1}R_2A_2'^nT_{B_2}} \equiv \sigma^{R_1A_1'^nT_{B_1}} \otimes \sigma^{R_2A_2'^nT_{B_2}}.$$
Observe that the state between $R_1A_1'^nT_{B_1}$ and $R_2A_2'^nT_{B_2}$ is a product state.

3. **Transmission.** Both senders input their respective systems $A_1'^n$ and $A_2'^n$ into the quantum interference channel $\mathcal{N}^{A_1'^n A_2'^n \to B_1^n B_2^n}$. The resulting state is

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

   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 necessarily a product state.

4. **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 $\mathcal{D}_i^{B_i^n T_{B_i} \to \hat{A}_i}$ 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

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

5. **Error analysis.** The conditions for a good quantum code are that both receivers be able to decode the quantum states from the respective senders with high fidelity:

$$F\left(\theta^{R_1 \hat{A}_1 R_2 \hat{A}_2},\ \psi^{R_1 A_1} \otimes \psi^{R_2 A_2}\right) \geq 1 - \epsilon.$$

Figure 1 depicts all of the above steps as they would occur in any entanglement-assisted quantum interference channel code.

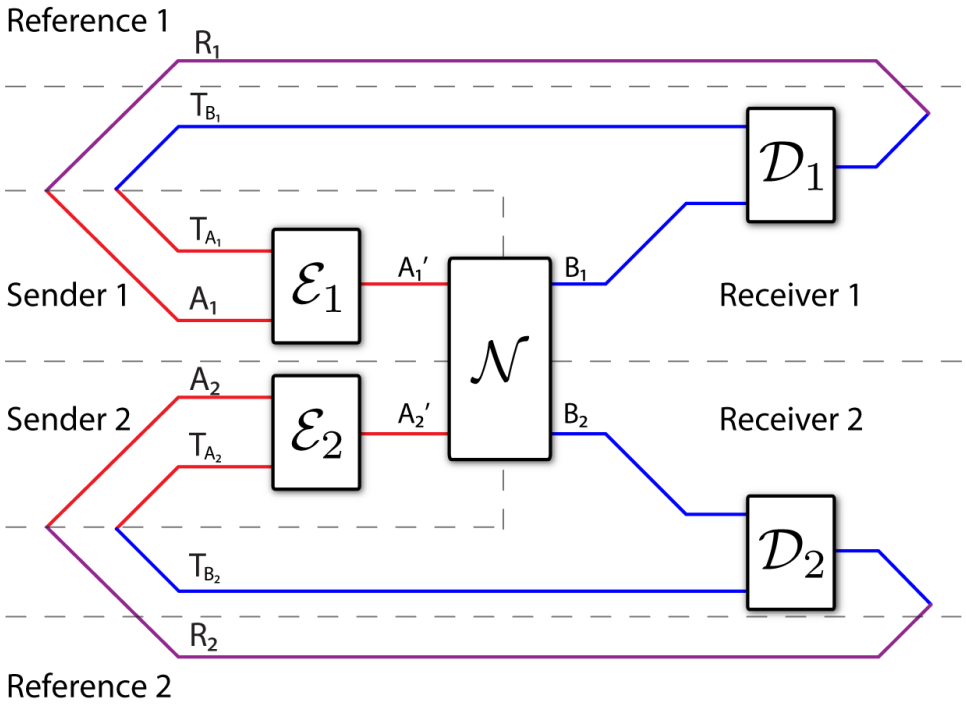

Figure 1: The circuit diagram for an entanglement-assisted quantum interference channel code.

Definition 10. *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.*

Definition 11. *The capacity region $\mathcal{C}_{\text{E-A}}^{QIC}(\mathcal{N})$ is the closure of the set of all achievable rates (Q_1, Q_2).*

While the current page length of this document prohibits me from writing them down with details, it should be noted that we can easily port two of the classical bounds to obtain bounds on $\mathcal{C}_{\text{E-A}}^{QIC}(\mathcal{N})$.

Using the results of Theorem 7 for the entanglement-assisted classical capacity of the MAC $\mathcal{C}_{\text{E-A}}(\mathcal{M})$, we can derive a quantum "naive" inner bound (Theorem 1) and the quantum Sato outer bound (Theorem 3). Both of these theorems simply use the MAC result as a black box, so they can be ported.

6 Discussion and conclusion

In this section we will make references 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? The most general classical problem to consider involves six independent messages: 3 starting at Tx1 and 3 starting at Tx2 with respective rates

$$\vec{R} = (R_{11}, R_{12}, R_{10},\ R_{21}, R_{22}, R_{20}), \tag{39}$$

where R_{ij} ($i, j \in \{1, 2\}$) is the rate for the message from sender i to receiver j and R_{k0} ($k \in \{1, 2\}$) is a broadcast rate from sender k to both receivers. The interference channel problem cuts out the two dimensional R_{11}, R_{22} plane and ignores all the remaining rates. The MMAC problem cuts out the (R_{10}, R_{20}) two dimensional region.

Are these cross rates not useful? These extra rates could maybe be converted into some other information resource. In the classical case, these extra messages could be used for some cooperative behaviour. In the quantum case, with the addition of entanglement between receivers they could perhaps be used with some post-processing to improve the direct transmission rates R_{11}, R_{22}. I feel these cross rates should be taken into account in general.

As mentioned several times in this paper, there exists an even more general model for the MAC and the IC in which the inputs are allowed to be correlated. I think that understanding joint source-channel coding in the multiuser case is a very interesting problem. Some early work on the subject [CGS80] reveals that it depends on a polymatroid structure which is very interesting since it connects with deeper notions about polytopes [Edm69, Zie95].

The capacity of the classical interference channel remains an open problem. Furthermore we have added the new problem of finding the capacity of the quantum interference channel. Which problem will be solved first?

The fact that after so many years we still don't know $\mathcal{C}_{IC}$ begs an even more fundamental question. Are the techniques of information theory general enough to study such problems? Perhaps an even more pressing question for quantum information theory is why are we only able to prove multi-letter formulas for these channels. Is there a fundamental paradigm shift that needs to happen in order to go further in information theory?

A Introduction to quantum information theory

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. For a good book on quantum mechanics see [Sak94]. For a book specifically on quantum information theory the reader can consult [NC00], which is the *bible* of the field.

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 polarization, electron spin, angular momentum and other finite dimensional degrees of freedom. It is also the representation of choice for the entire field of Quantum Information Science, which comprises quantum information theory, quantum cryptography, quantum error correcting codes, quantum computation and many others.

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

Definition 12 (qubit)**.** *A qubit is a two dimensional quantum state $\vec{s}$ which is normalized, i.e. it is a unit length vector in $\mathbb{C}^2 = \{(a,b)^T | a,b \in \mathbb{C}\}$.*

For example, we can prepare an electron in a spin up state $\vec{s} = (1,0)^T$, spin down state $\vec{s} = (0,1)^T$ or a 50–50 *superposition* of the up and down states $\vec{s} = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^T$. 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 overall 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 $\mathcal{E} = \{p(i), \vec{\psi}_i\}$ is a set of quantum states $\vec{\psi}_i$ which occur with probability $p(i)$. One way to describe a quantum source is to specify the states $\vec{\psi}_i$ and the corresponding probabilities p_i associated with this source. Quantum information deals with the compression, transmission and otherwise manipulation of quantum ensembles.

A.2 Notation

This section will focus on specific notions and notation that are used in quantum information theory. While quantum states really *are* vectors, the customary notation for quantum states is using bra's and ket's instead of the arrow on top.

$$|\psi\rangle \triangleq \vec{\psi}. \tag{40}$$

We will use this notation from here on to denote quantum states.

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 $\sigma^A \in$[2]$\mathcal{H}^A$ 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.3 Quantum information theory

The fundamental ideas of quantum information theory are analogous to those of classical information theory. We quantify the information content of quantum systems by using their entropy.

Definition 13 (von Neumann Entropy)**.** *Given the density matrix $\rho^A \in \mathcal{H}^A$, the expression*

$$S(A)_\rho = -\mathrm{Tr}\left(\rho^A \log \rho^A \right) \tag{41}$$

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 function. The von Neumann entropy of quantum state ρ^A with spectral decomposition $\rho^A = \sum_i \lambda_i |e_i\rangle\langle e_i|$, is equal to the Shannon entropy of its eigenvalues.

$$S(A)_\rho = -\mathrm{Tr}\left(\rho^A \log \rho^A \right) = -\sum_i \lambda_i \log \lambda_i = H(\{\lambda_i\}) \tag{42}$$

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

$$H(A|B)_\rho \triangleq H(AB)_\rho - H(B)_\rho \tag{43}$$

[2]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.

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

$$I(A;B)_\rho \triangleq H(A)_\rho + H(B)_\rho - H(AB)_\rho \tag{44}$$

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

$$\begin{aligned} I(A;B|C)_\rho &\triangleq H(A|C)_\rho + H(B|C)_\rho - H(AB|C)_\rho & (45)\\ &= H(AC)_\rho + H(BC)_\rho - H(ABC)_\rho - H(C)_\rho. & (46)\end{aligned}$$

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

$$H(AC) + H(BC) \geq H(C) + H(ABC). \tag{47}$$

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.

Furthermore, we define a purely quantum information theoretic quantity called *coherent information* which depending on the context will be expressed in one of two ways. For a fixed joint state σ^{AB}, we write $I_c(A\rangle B) \equiv H(B) - H(AB) = -H(A|B)$. Otherwise, if we are given a density matrix $\rho^{A'}$ and a channel $\mathcal{N}^{A'\to B}$ which give rise to a joint state $(1^A \otimes \mathcal{N})(\Phi_\rho)$, where $|\Phi_\rho\rangle^{AA'}$ is any purification of ρ, we will often use the notation

$$I_c(A\rangle B) = I_c(\rho, \mathcal{N}) = H(\mathcal{N}(\rho)) - H((1 \otimes \mathcal{N})(\Phi_\rho)).$$

It can be shown that this latter expression is independent of the particular purification $|\Phi_\rho\rangle$ that is chosen for ρ.

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 9. *Consider the Φ^+ Bell state*

$$|\Phi\rangle^{AB} = \tfrac{1}{\sqrt{2}}(|00\rangle^{AB} + |11\rangle^{AB}). \tag{48}$$

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

$$H(A)_\Phi = 1, \qquad H(B)_\Phi = 1, \qquad H(AB)_\Phi = 0, \tag{49}$$

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

$$H(A|B) = H(AB) - H(B) = -1 \text{ [bits]}, \tag{50}$$

and the mutual information

$$I(A;B) = H(A) + H(B) - H(AB) = 2 \text{ [bits]}. \tag{51}$$

Equation (50) 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 (51), 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.4 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 $\mathcal{D}$:

$$\rho \xrightarrow{\mathcal{E}} \rho' \qquad \text{or} \qquad \mathcal{E}(\rho) = \rho' \tag{52}$$

Unitary transformations are a type of quantum operation

$$\mathcal{E}(\rho) = U\rho U^\dagger \tag{53}$$

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

$$\mathcal{E}(\rho) = \text{Tr}_{\text{env}} U\left(\rho \otimes \rho_{\text{env}}\right) U^\dagger \tag{54}$$

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

$$p(m) = \text{Tr}\Big(M_m^\dagger \rho M_m\Big). \tag{55}$$

The output quantum state associated with this outcome is

$$\widetilde{\rho}_m = \frac{M_m^\dagger \rho M_m}{\text{Tr}\Big(M_m^\dagger \rho M_m\Big)}. \tag{56}$$

A.5 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 problem.

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]:

$$\{c \to c\} \geq \max_{p(x)} I(X;Y)[c \to c], \tag{57}$$

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

$$|\Phi\rangle^{AB} = \tfrac{1}{\sqrt{2}}(|00\rangle^{AB} + |11\rangle^{AB}), \tag{58}$$

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 reader 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.6 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

$$F(|\varphi\rangle, |\psi\rangle) = |\langle\varphi|\psi\rangle|^2 . \tag{59}$$

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

$$F(\rho, \sigma) = \mathrm{Tr}\left(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)^2 . \tag{60}$$

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:

$$\mathcal{N}(|\psi\rangle\langle\psi|) = \sigma^{\widehat{A}}. \tag{61}$$

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

$$\overline{F}(\{p_i, |\psi_i\rangle\}, \mathcal{N}) := \sum_i p_i F(|\psi_i\rangle, \sigma_i), \qquad \sigma_i = \mathcal{N}(|\psi_i\rangle\langle\psi_i|). \tag{62}$$

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

$$\overline{F}(\mathcal{E}, \mathcal{N}) \geq 1 - \epsilon \tag{63}$$

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

$$\mathcal{N}^{A\to\widehat{A}}\otimes \mathrm{id}^R\big(|\psi\rangle\langle\psi|^{AR}\big) = \sigma^{\widehat{A}R}. \tag{64}$$

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

$$F(|\psi\rangle^{AR}, \sigma^{\widehat{A}R}) = \langle\psi^{AR}|\sigma^{\widehat{A}R}|\psi^{AR}\rangle \geq 1-\epsilon. \tag{65}$$

Equation (65) 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 (65) implies equation (63).

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.

[Car78] A. Carleial. Interference channels. *IEEE Transactions on Information Theory*, 24(1):60–70, 1978.

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

[CGS80] T.M. Cover, AE Gamal, and M. Salehi. Multiple access channels with arbitrarily correlated sources. *IEEE Transactions on Information Theory*, 26(6):648–657, 1980.

[CMGEG08] H-F. Chong, M. Motani, H. K. Garg, and H. El Gamal. On the Han-Kobayashi region for the interference channel. *IEEE Transactions on Information Theory*, 54(7):3188–3195, 2008.

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

[Edm69] J. Edmonds. Submodular functions, matroids, and certain polyhedra. *Proc. Calgary Int. Conf. Combinatorial Structures and Algorithms*, pages 69–87, June 1969. (Reprinted in *LNCS* 2570:11–26, 2003).

[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.

[HK81] Te Han and K. Kobayashi. A new achievable rate region for the interference channel. *Information Theory, IEEE Transactions on*, 27(1):49–60, Jan 1981.

[HK07] TS Han and K. Kobayashi. A further consideration on the HK and the CMG regions for the interference channel. In *Proc. Inf. Theory Applicat. Workshop*, 2007.

[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.

[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.

[Zie95] G. M. Ziegler. *Lectures on polytopes*. Springer-Verlag, New York, 1995.