Super-quantum and quantum enhancements of two-sender channels

by

Yihui Quek

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Bachelor of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2016

Signature redacted

Author .....
Department of Physics
May 6, 2016

Signature redacted

Certified by.... .....
Peter W. Shor
Henry Adams Morss Jr. Professor of Applied Mathematics
Thesis Supervisor

Signature redacted

Accepted by..... .....
Nergis Mavalvala
Chairman, Department Committee on Undergraduate Theses

Super-quantum and quantum enhancements of two-sender channels

by

Yihui Quek

Submitted to the Department of Physics
on May 6, 2016, in partial fulfillment of the
requirements for the degree of
Bachelor of Science

Abstract

This thesis studies the consequences of 'super-quantum non-local correlations', which are hypothetical violations of Bell/CHSH inequalities that are stronger – more non-local – than quantum mechanics allows, yet weak enough to respect special relativity in prohibiting faster-than-light communication. Understanding the power of such correlations will yield insight into the non-locality of quantum mechanics. Whereas previous studies of super-quantum correlations have demonstrated enhancements in cryptography and computation of distributed functions, this work opens up a new direction of research by showing that they can also enhance the capacity of classical communication over a noisy channel. Our results exhibit a trifecta of proof-of-concept channels: first, we show an interference channel between two sender-receiver pairs where the senders are not allowed to communicate, for which a shared super-quantum bit allows perfect classical communication. This feat is not achievable with the best classical (senders share no resources) or quantum-assisted (senders share entanglement) strategies. We next show two examples that are conjectured to demonstrate the following capacity separations: an interference channel that strictly separates super-quantum from quantum-assisted strategies, and quantum-assisted from classical strategies; and, lastly, a multiple-access channel that strictly separates super-quantum-assisted strategies from classical ones. At the heart of some of these examples is a novel connection between multi-sender channels and multi-player XOR and pseudo-telepathy games.

Thesis Supervisor: Peter W. Shor
Title: Henry Adams Morss Jr. Professor of Applied Mathematics

Acknowledgments

Quantum information grew out of classical information. My journey of thesis-writing has been much the same. Its intellectual foundations are rooted in coursework for two of my MIT classes, forming an unusual but potent classical-quantum combination: 18.424, Seminar in (Classical) Information Theory; and 8.371/6.445 Quantum Information Theory II. One unifying figure has supported me through both classical and quantum phases of education: Professor Peter Shor. His mastery of all aspects of Applied Mathematics and propensity for ingenious connections from foundations to Physics have contributed indelibly to my education and this thesis.

The quantum-classical division also nicely partitions the set of people who formed my academic and moral support system at MIT. For teaching me **everything quantum**: I thank Professors Barton Zwiebach, Aram Harrow, Seth Lloyd and Isaac Chuang. Barton Zwiebach, having taught 8.05 and recitations for 8.04 and 8.06, witnessed my transformation from QM greenhorn in freshman year to 8.05x TA and fledgling QI researcher. My attitude towards Physics was honed by the many hours I spent in his office in each of these capacities. I also thank my supervisors during my past three summer research experiences who made sure that work was always as fascinating as the touristic charms of Paris, Dresden and Waterloo/Toronto: Fréderic Pascal, Steffen Werner and Joshua Combes, all of whom can take credit for part of my current interdisciplinary research thrusts. Thank you to the researchers who gave valuable comments on this work even through brief chats at conferences and summer schools: Debbie Leung, Ike (again) and Elton Zhu; and to those who so warmly welcomed a curious undergraduate into their community: Krysta Svore, Hideo Mabuchi, Patrick Hayden, K. Birgitta Whaley, my 8.371 classmates, Ted Yoder, Guanghao Low and Dax Koh. Aram Harrow, who falls into many of the above categories, has been a role model of a young and brilliant trailblazer in QI, as well as an absurdly kind professor and all-around super-human.

In keeping with the theme of this thesis, I must also acknowledge my **superquantum** teachers; professors whose lucid explanations imbued into me an appreciа-

tion of the world beyond quantum information – Robert Jaffe, Michael Artin, Nergis Mavalvala, Hung Cheng, Scott Hughes, Thomas Greytak, David Pesetsky, Martin M. Marks, Theresa Neff.

In the **classical realm**, I am the fortunate beneficiary of friendship, support and heroic amounts of tolerance for my puns. No two people deserve more credit for this than Boon Siong Neo and Angela Leong, my long-suffering high school-to-MIT compatriots and aiders and abetters of my nefarious schemes. While they would probably rather eat grass than admit it, they are the most sibling-like non-siblings I know. I would be a strictly lesser person without their, and Stanislav Fořt's, moral support. My MIT experience has also been enriched by the excellent company of Herng Yi Cheng, Minh Tue Vo Thanh, Amyas Chew, Lynn Chua, Diptarka Hait, Konstantin Martynov and Thipok Rak-Amnouykit, to name a few.

Two more sets of people deserve special mention: Prof Hui Khoon Ng, who took me under her wing when I was a high school student and introduced me to academic Physics research. In many ways she is responsible for my career path. Finally, I would like to thank my parents for being my beacons as a child, chief strategists in my early adulthood, and at all times, the people I wish to make proud.

This thesis is dedicated to the memory of my grandmother, Madam TAN Peck Hong (1930-2016), of whose determination and spirit I am a proud heir.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Is Nature local? Must particles be in contact with each other to directly interact? If we take quantum mechanics to be the fundamental theory of nature, the answer seems to be no. In 1964, John Bell showed that quantum theory is not a local hidden-variable theory [5]; a local theory would not be able to reproduce the observed statistical outcomes of quantum measurements. Already, the puzzling non-locality of quantum mechanics is manifest in the Aharanov-Bohm effect, where a charged particle is affected (for instance, by acquiring a phase shift) by an electromagnetic field despite being confined to a region where the field is zero.

But any theory of Nature must also play by the rules of special relativity, which imposes a speed limit on communication. Quantum theory is the only theory we know that allows two spacelike-separated observers to influence each other *(non-locality)*, and yet, prohibits instantaneous communication between them – the above-mentioned 'influence' must not allow for information transfer *(relativistic causality)*. But is it the *unique* theory that checks both these boxes? Popescu and Rohrlich's [6] search for theories that can be deduced from these two criteria alone, led to the discovery that nature could be *even more nonlocal* than quantum mechanics predicts, yet be fully consistent with relativity [7]. A *super-quantum* theory could produce even stronger nonlocal correlations than quantum theory.

Quantum information is the study of how to cleverly utilize quantum resources, such as entanglement, to aid communication tasks – quantum key distribution, quan-

tum bit commitment and so on. This begs the question: could we use the maximally non-local correlations of *super-quantum* theories as a resource, and could they make our lives easier/more efficient?

This thesis is the first survey of how super-quantum resources could be used to aid channel communication. In Chapter 1, we explain how to quantify non-locality via the CHSH inequality, and introduce the abstraction of PR-boxes (first introduced in [6]) to represent super-quantum non-local theories. In Chapter 2, we introduce some preliminaries on classical and quantum channels, including common notation. In Chapter 3, we present our original result of a two-sender, two-receiver communication channel for which the classical joint capacity is strictly higher with the aid of a PR-box, than with entanglement and/or a classical strategy. In Chapter 4, we present two more original channels which show two conjectured separations: an interference channel that strictly separates these three classes of resources (classical, quantum-assisted and PR-box assisted); and a multiple-access channel that has its capacity enhanced by a PR-box. In doing so we draw a new connection between the fields of XOR games and network quantum information theory. We finally conclude with a summary of results and future outlook for related research in Chapter 5.

1.1 Non-signaling theories

Quantum Mechanics is non-local, and therefore permits 'interaction at a distance' by space-like separated parties. How to reconcile this with special relativity, which forbids communication between those same parties (the non-signaling property)? In fact, the two requirements are not mutually exclusive; there exist interactions that do not constitute a *communication*, or transfer of information. Following the convention (which exists for good reason – see footnote[1]), all super-quantum theories we consider

[1]There are merits to this viewpoint: Aharonov [7] pointed out that taking non-locality and non-signaling as fundamental axioms of QM implies non-determinism, another property of quantum mechanics. The reader is referred to [8, 6] to find out more, but Popescu provides a good, intuitive paraphrase of the argument: "If by moving something here, something else instantaneously wiggles there, the only way in which this doesn't lead to instantaneous communication is if that 'wiggling thing' is uncertain and the wiggling can be only spotted a posteriori." This is exactly a description of non-determinism.

shall obey these two properties.

Mathematically, this non-signaling property amounts to the situation where Alice cannot gain any information about Bob's input by altering her input. That is, the sum over Bob's outputs (b) of the joint probability distribution is independent of Bob's inputs (y), and equal to Alice's marginal distribution – and vice versa.

Definition 1.

$$\sum_b P(a,b|x,y) = \sum_b P(a,b|x,y') = P^A(a|x) \forall a,x,y,y'$$
$$\sum_a P(a,b|x,y) = \sum_b P(a,b|x',y) = P^B(b|y) \forall b,y,x,x'. \tag{1.1}$$

We may interpret the above equation as meaning that no subset of the two parties interacting with the theory should be able to find out anything about the inputs of any of the others by looking at their own inputs and outputs. It is not hard to generalize our definition above to a definition of a non-signaling theory involving multiple parties, and the reader is referred to [9] for the details. For our purposes, it suffice to consider the simplest class of non-signaling theories, the ones that involve two parties. Within this set of non-signaling theories, then, we single out the *non-local* ones, the ones where Alice's inputs and outputs affect Bob's inputs and outputs even if Alice and Bob are space-like separated.

1.2 Nonlocality of quantum theory

Bell's paper in 1964 [5] brought to light the existence of correlations that can be obtained from measurements of a bipartite quantum state that could not be reproduced by systems that do not communicate (ie. are restricted, by classical mechanics or otherwise, to be local). Quantum mechanics has therefore been termed a *non-local theory.*

What, exactly, is locality (or the lack thereof)? The term pertains to theories that underlie states of physical objects, or more pragmatically, regimes in which ex-

perimental effects on those physical objects can be observed. All such experiments to elucidate the underlying theory involve some sort of measurement, such as spin measurements on a spin-half particle. A local theory would *prohibit* physical measurements in one place from affecting the measurement outcomes of another experimenter who is spacelike-separated from the first one, if there is no field between them – or, to borrow an analogy from Popescu [7], a nonlocal theory is one in which 'moving something here, something else instantaneously wiggles there'.

Given access to the observations of two spacelike-separated experimenters, one might guess that *locality* of the underlying theory can be inferred from looking at the distribution of their measurement outcomes. This intuition is formalized by the notion of the CHSH value, named after its founders, Clause, Horne, Shimony and Holt [10]. In 1969, they put forth an inequality that bounds the statistics of spatially-separated measurements in local hidden-variable models:

$$|\langle A_0B_0\rangle + \langle A_0B_1\rangle + \langle A_1B_0\rangle - \langle A_1B_1\rangle| \leq 2 \qquad \text{Local HV theory} \tag{1.2}$$

where A_0 and A_1 are local measurement operators corresponding to spin up and spin down on experimenter A's spin-half particle, and B_0 and B_1 the analogous measurement operators for Bob, and $\langle . \rangle$ denotes expectation value, and the quantity on the left hand side of expression 1-3 is sometimes referred to as the 'CHSH value' of a theory and has become a measure of the non-locality of that theory.

Quantum theory is non-local because it can violate this inequality by measurements on an entangled state, such as the state $\frac{|0\rangle_A|1\rangle_B - |1\rangle_A|0\rangle_B}{\sqrt{2}}$. It is easy to verify that this state satisfies the following:

$$|\langle A_0B_0\rangle + \langle A_0B_1\rangle + \langle A_1B_0\rangle - \langle A_1B_1\rangle| = 2\sqrt{2} \qquad \text{Quantum theory} \tag{1.3}$$

which clearly violates the CHSH inequality (1-3).

Tsirelson[11] proved that with quantum mechanics, $2\sqrt{2}$ is the maximal achievable

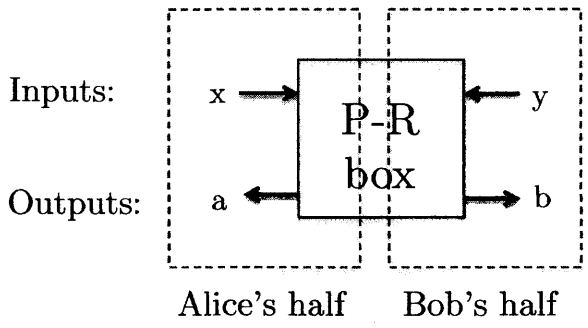

Figure 1-1: A PR, or non-local box, whose inputs and outputs are governed by the distribution in Equation 1.4.

violation of the inequality in equation 1-3. This falls short of its algebraic maximum, 4. In 1994, Popescu and Rohrlich[6], pursuing the line of inquiry 'Why isn't quantum theory more non-local?', found that there do exist theories that are more non-local than quantum mechanics, and achieve a CHSH value of 4! To unify these theories, they proposed an abstraction to represent the probability distribution that they induce on measurement outcomes: a *non-local box*, visualized in figure 1-1. This is a bipartite correlated box with two ends, one of which is held by Alice and the other by Bob. Alice inputs x (respectively Bob inputs y) and the box outputs a (respectively b) according to the probability distribution $P(a,b|x,y)$ (where $x,y,a,b \in \{0,1\}$):

$$P^{PR}(a,b|x,y) \begin{cases} 1/2 & \text{if } a \oplus b = xy \\ 0 & \text{otherwise} \end{cases} \tag{1.4}$$

To put ourselves on the same footing with our earlier discussion, we now interpret A_0, A_1 (respectively B_0, B_1) as the expected value of the box's output when Alice (respectively Bob) puts in 0, 1 into their ends of the PR-box respectively. This information-theoretic formulation of Alice and Bob's interaction with the theory is completely analogous with our previous language of measurements when construed within the measurement-operator formalism: in making measurements of a two-level system, Alice and Bob apply a set of measurement operators $\{\Pi_0, \Pi_1\}$ corresponding to the two possible outcomes, which correspond exactly to the set of inputs $\{0,1\}$ of

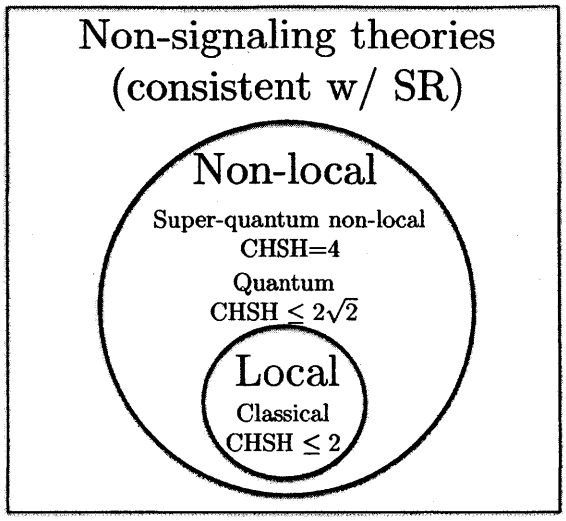

Figure 1-2: Types of theories grouped by their locality properties (they must all not permit space-like separated observers to communicate and hence all fall under the banner of non-signaling)

both experimenters to the PR-box.

Thus, with such a PR box, we achieve the following super-quantum correlations:

$$|\langle A_0 B_0\rangle + \langle A_0 B_1\rangle + \langle A_1 B_0\rangle - \langle A_1 B_1\rangle| = 4 \qquad \text{Super-quantum nonlocality} \quad (1.5)$$

Our discussion of theories in the past two sections, in terms of their locality properties (as measured by their CHSH value) and the non-signaling requirement, is summarized in Figure 1-2, which the reader is encouraged to familiarize herself with before proceeding.

Much of this paper aims to identify communication tasks that show a separation in efficiency given three classes of quantum resources: classical, quantum and super-quantum (in the form of PR-boxes).

1.3 XOR games

Many of the most striking applications of quantum information come from the fields of communication and cryptography – one famous example is quantum key distribution, an unconditionally secure method to transmit cryptographic keys. In QKD and many such tasks, it is quantum non-locality (entanglement) that provides the speedup over the best known classical strategy.

However, entanglement alone cannot be used for communication – if Alice and Bob share an entangled state $|\Psi\rangle$ and Alice alone performs some measurement on her part of the state, the results are completely determined by her reduced density matrix, $\mathrm{Tr}(|\Psi\rangle\langle\Psi|)_B$ and are independent of any operation that Bob might perform. It is only when Alice and Bob perform measurements *together* that interesting (ie. not replicable classically) correlations between their measurement outcomes might appear. We give this a name, non-locality.

The fact that there exist ways to use non-local correlations to defeat best-known classical protocols is but one reason to study entanglement. An even more compelling one is that *we may view entanglement as a resource*, which we can exploit to design new protocols to achieve tasks previously thought impossible. Entanglement is commonly studied through multiplayer games for which a quantum strategy wins with higher probability than the best classical one. Such a game is called an XOR game if the players always answer with a single $\{0/1\}$ bit each, and the validity of a given pair of answers only depends on their parity. For an excellent survey of XOR games and their generalizations, the reader is referred to [12].

Perhaps the most illuminating (also, canonical) example of these is the CHSH game, which we turn to in the next section. Far from being a mere intellectual curiosity, this game has turned into a primitive for many quantum communication protocols – including ours – that exploit some facet of its non-local winning strategy.

1.3.1 CHSH game

The CHSH game is named after Clauser, Horne, Shimony and Holt, who also lay claim to the eponymous inequality from an earlier section. Much of this section is an abridged version of the wonderfully concise summary of nonlocal games in [1]. It is one of a class of games with the following structure, illustrated in figure 1-3: there are two (or more) players, Alice and Bob, who cooperate but do not communicate with each other. Instead, each of the players is allowed to interact only with a referee. The referee randomly selects a question for each player (represented by a number – suppose r for Alice and s for Bob), and each player must respond to the referee with an answer (another number – suppose a from Alice and b from Bob). The players are considered to have won the game if some predicate $f(r, s, a, b)$ computed (by the referee) on their questions and answers evaluates to a desired value, and lose otherwise.

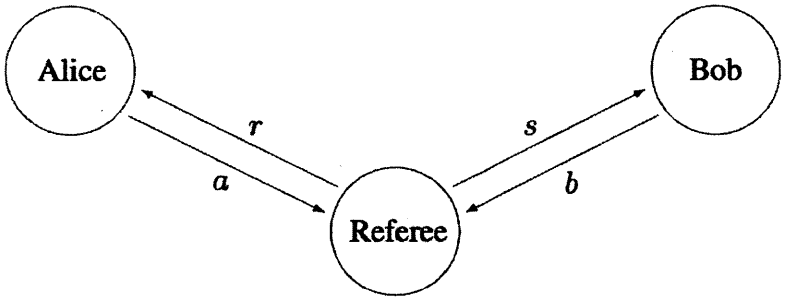

Figure 1-3: The structure of a two-player game. Figure and captions taken from [1]. Here, the referee chooses a pair of questions (r, s) (according to some prespecified distribution), sends r to Alice and s to Bob, and Alice and Bob answer with a and b, respectively. The referee evaluates some predicate on (r, s, a, b) to determine if they win or lose.

In the CHSH game, the referee chooses questions $rs \in \{00, 01, 10, 11\}$ for Alice and Bob uniformly, and each of them must answer a single bit, a and b respectively. They win if

$$a \oplus b = r \wedge s \tag{1.6}$$

ie. the winning conditions are as given in Table 1.1. It is not a coincidence that this is identical to the condition on inputs and outputs of the PR-box which must be satisfied for said input-output combination to be produced by the box with non-zero probability. The PR-box's inputs and outputs are always maximally non-locally correlated – in exactly the same manner that Alice and Bob's questions and answers should have to be correlated if they are to win the game. However, as we shall see soon, Alice and Bob's best quantum strategy produces outputs correlated like so with probability of at most 0.85 (so they win with *at most* that probability), whereas if they had a PR-box, they could win with probability 1. This is precisely a restatement of the fact that quantum mechanics is not maximally non-local (a notion defined in terms of its CHSH value, which, of course, arises from the CHSH game), while a PR-box is.

(r, s) (Questions)	(a, b) (Answers)
(0,0)	(1,1), (0,0)
(0,1)	(1,1), (0,0)
(1,0)	(1,1), (0,0)
(1,1)	(0,1), (1,0)

Table 1.1: Winning conditions for CHSH game, as well as input-output condition on PR-box: $a \oplus b = r \wedge s$

Classically, the maximum winning probability is $\frac{3}{4}$, and it can be shown as follows: we consider only deterministic strategies, that is, the answer (a, b) given by each player is a fixed function of the question $\in \{00, 01, 10, 11\}$ they get, and is denoted by a_0, a_1, b_0, b_1. Then we may re-write the winning condition 1.6:

$$a_0 \oplus b_0 = 0$$

$$a_0 \oplus b_1 = 0$$

$$a_1 \oplus b_0 = 0$$

$$a_1 \oplus b_1 = 1$$

Adding all of these modulo 2 on the left-hand side and right-hand-side give $0 = 1$,

a contradiction. Since at least one out of the four predicates above cannot be satisfied by any assignment of $\{a_0, b_0, a_1, b_1\}$, the maximal winning probability is $\frac{3}{4}$, and this is true even if we permit a probabilistic classical strategy (which is but an ensemble of deterministic classical strategies).

Does our winning probability improve if we allow the players to exploit entanglement? Remarkably, the answer is yes. Let Alice and Bob share an entangled state

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}}$$

. The idea is for each of the players to make a measurement, in a basis that depends on the question they are given. They report the answer '0' if their measurement outcome is '1' and '1' if their measurement outcome is '-1'. The strategy is as follows:

If Alice receives the question '0', she measures her qubit with respect to a basis of z-eigenstates,

$$\{|0\rangle, |1\rangle\}.$$

If she receives the question 1, she measures with respect to the x-axis,

$$\left\{\frac{|0\rangle + |1\rangle}{\sqrt{2}}, \frac{|0\rangle - |1\rangle}{\sqrt{2}}\right\}.$$

Bob uses the same strategy, but with his choices of measurement axes rotated by $\frac{\pi}{4}$ with respect to Alice's. That is, question 0 corresponds to the basis

$$\left\{\cos\left(\frac{\pi}{8}\right)|0\rangle + \sin\left(\frac{\pi}{8}\right)|1\rangle, -\sin\left(\frac{\pi}{8}\right)|0\rangle + \cos\left(\frac{\pi}{8}\right)|1\rangle\right\}$$

while question 1 corresponds to the basis

$$\left\{\cos\left(\frac{\pi}{8}\right)|0\rangle - \sin\left(\frac{\pi}{8}\right)|1\rangle, \sin\left(\frac{\pi}{8}\right)|0\rangle + \cos\left(\frac{\pi}{8}\right)|1\rangle\right\}.$$

With this strategy, the players always win. To see this in one specific instance (it is

easy to enumerate the other three), suppose Alice receives question 1 and Bob receives question 0. Referring back to Table 1.1 for the winning conditions and plugging in the above prescription, we may calculate their winning probability as:

$$\begin{aligned}
\Pr(\text{win}) &= \Pr(\text{Both get measurement outcome 1}) + \Pr(\text{Both get measurement outcome -1}) \\
&= \left[\left(\frac{\langle 0| + \langle 1|}{\sqrt{2}}\right)_A \left(\cos\left(\frac{\pi}{8}\right)\langle 0| - \sin\left(\frac{\pi}{8}\right)\langle 1|\right)_B \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right)_{AB}\right]^2 \\
&+ \left[\left(\frac{\langle 0| - \langle 1|}{\sqrt{2}}\right)_A \left(\cos\left(\frac{\pi}{8}\right)\langle 0| + \sin\left(\frac{\pi}{8}\right)\langle 1|\right)_B \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right)_{AB}\right]^2 \\
&= \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}}\right) = \cos^2\left(\frac{\pi}{8}\right) \qquad (1.7)
\end{aligned}$$

One can verify that this strategy produces the same winning probability ($\approx 85\%$) for the other three question combinations as well. This is noticeably better than the best classical strategy, which yields a winning probability of ($\approx 75\%$). Tsirelson's bound [11] shows that no quantum strategy does better than $\cos^2\left(\frac{\pi}{8}\right)$ even if the two senders are allowed to share a state on many qubits. A proof of this, provided in the Appendix, abstracts away from any explicit representation of an entangled state, and follows from operator theory. But let us return to the key physical takeaway from this thought experiment, which is: we have defined a sense in which quantum correlations are more non-local than classical correlations, and furthermore, found a simple experimental demonstration of the fact (the CHSH game). The rest of this thesis explores what we can do with quantum, and super-quantum, correlations.

1.4 Chapter summary

In this chapter, we introduced the ongoing debate about non-locality and non-signaling theories as well as two important tools for the analysis in the rest of this paper: PR-boxes and the CHSH game. Violations of the CHSH inequality give us a sense in which we may classify theories according to their non-locality properties: classical, quantum and super-quantum. Having gotten a grasp on measures of non-locality, we

now explore how we can exploit it for channel communication.

Chapter 2

Classical and Quantum Channels

Just as Clauser et al [10] brought to light the importance of entanglement as a resource by translating the abstruse mathematical formalism of Bell's inequalities into a physically realizable experiment demonstrating the power of quantum mechanics, we, too, would like to find an application for PR-boxes that beats the best classical or quantum strategy. To prime the reader for our main result – the demonstration of channels that prove such a separation in super-quantum, quantum and classical capacities – we spend some time in this chapter developing foundations of classical channel coding. We conclude with brief remarks about its quantum analogue.

Channel coding is arguably the founding work in the field of information theory. In 1949, Claude Shannon published his model of a communication system [2], along with two theorems that quickly gained prominence for placing limits on achievable rates of communication in real systems. As if its seminal impact on modern communications systems were not reason enough to study this model, Shannon's second theorem will pave the way for our understanding of channel capacity, a concept referenced heavily in the next chapter.

It behooves us right now to define the concepts of *entropy*, *conditional entropy* and *mutual information*, for, not only were they invented by Shannon in the same paper as measures to quantify the information content of probability distributions, but will also crop up repeatedly in what follows. Let X be a discrete random variable with alphabet $\mathcal{X}$ and probability mass function $p(x) = \Pr\{X = x\}, x \in X$ (correspond-

ingly: $Y, \mathcal{Y}, p(y)$). Let their joint probability mass functions be $p(x, y)$, and whenever two variables are considered, let $p(x), p(y)$ denote their marginal distributions.

Definition 2 (Entropy of a single random variable). *The entropy $H_n(X)$ of a discrete random variable X is defined by*

$$H_n(X) \equiv -\sum_{x \in \mathcal{X}} p(x) \log_n p(x) \tag{2.1}$$

(For simplicity's sake, in future we will always be using $n = 2$, and we will drop the subscript on $H(X)$.) Entropy is sometimes termed 'surprisal', for it is always maximized on the uniform distribution of X: since all values are equally likely, one has no clue which will be drawn next. With two random variables, we may define three other measures:

Definition 3 (Two-variable information measures).

1. *The joint entropy $H(X, Y)$ is the entropy of the joint probability distribution of a pair of discrete random variables (X, Y).*

$$H(X, Y) \equiv -\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(x, y) \tag{2.2}$$

2. *The conditional entropy of X given Y is the entropy of X given full knowledge of Y.*

$$\begin{aligned} H(X|Y) &\equiv -\sum_{x \in \mathcal{X}} p(x) \log p(x) \\ &= -\sum_{x \in \mathcal{X}} p(x) \sum_{y \in \mathcal{Y}} p(y|x) \log p(y|x) \end{aligned} \tag{2.3}$$

3. *The mutual information $I(X; Y)$ is the reduction in the uncertainty of X due*

to knowledge of Y.

$$\begin{aligned} I(X;Y) &\equiv \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\log\frac{p(x,y)}{p(x)p(y)} \\ &= -\sum_{x,y} p(x,y)\log p(x) + \sum_{x,y} p(x,y)\log p(x|y) \qquad \textit{by Bayes' Rule} \\ &= H(X) - H(X|Y) \end{aligned} \tag{2.4}$$

and we easily verify that $I(X;Y) = I(Y;X)$.

Much of the classical part of the following review is taken from Shannon's original paper [2] and *Elements of Information Theory* [3], while the quantum part is compiled from [13] and [14].

2.1 Classical channel coding

The problem of sending a message through a noisy channel may be summarized as in figure 2-1. Whereas the description below is for a single sender-receiver pair, it is not difficult to see how the problem generalizes for multiple such pairs.

The following three broad steps define an **(M, n) code** for the channel:

1. **(Encoding)** The sender would like to communicate the message W to the receiver, drawn from the index set $\{1, 2, ...M\}$. This message is encoded by the sender (deterministically) as the length-n **codeword** $X^n(W) \equiv x^1x^2...x^n$. The set of all codewords for all possible W is called the *codebook*.

2. **(Transmission)** The sender inputs each letter of the codeword into a single use of the channel, leading to outputs (Y_1^n, Y_2^n) $p(y_1^n, y_2^n|x_1^n, x_2^n)$ where p is the probability distribution of the channel, to the receiver.

3. **(Decoding)** The receiver then guesses the index W_i by an appropriate decoding rule $(\hat{W}_i) = g(Y_i^n)$. He makes an error if $\hat{W}$ is not the same as the index W that was transmitted.

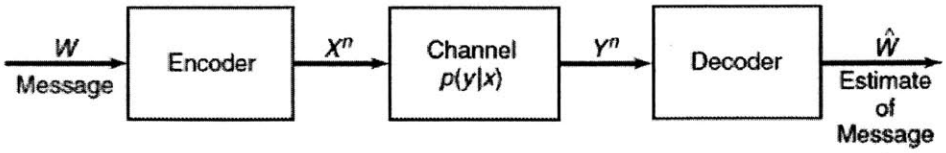

Figure 2-1: The Shannon-Weaver model[2] of a communication system (Figure taken from [3]).

All channels that we consider will be discrete and memoryless. Shannon's contribution was to represent such a channel, as well as the communication system outlined above, mathematically: the **channel** $(\mathcal{X}, p(y|x), \mathcal{Y})$ consists of input and output alphabets $\mathcal{X}$ and $\mathcal{Y}$ as well as a probability transition matrix $p(y|x)$ such that $\Sigma_y p(y|x) = 1$. That is, the entry in the y^{th} column and x^{th} row of this matrix is the probability that the receiver gets the output symbol y given that the sender sent the symbol x.

Definition 4. *The information capacity of a channel is*

$$C = \max_{p(x)} I(X;Y) \tag{2.5}$$

where the maximum is taken over all input distributions $p(x)$, and $I(X;Y)$ denotes the mutual information (Equation 2.4) between X and Y.

How good a code is depends on the probability that $\hat{W}$ is close to W, which we formalize[1] with the following quantities. These will eventually help us define the notion of the 'rate' of a code:

1. (Conditional probability of error) $\lambda_i = P(\hat{W} \neq i | W = i)$. Here, i represents the sender's intended message.

2. (Maximal probability of error) $\lambda^{(n)} = \max_{i \in \{1,2,...M\}} \lambda_i$ – that is, the worst-possible performance of the sender-receiver pair.

[1] Alternatively, *codify*

For a particular coding strategy, the **rate**, R, measures how much information can be communicated via that strategy such that the error vanishes as the input length goes to infinity. The maximum rate for a given channel is the *capacity* of that channel.

Definition 5. *The rate of a code is given by*

$$R = \frac{\log_2 M}{n} \tag{2.6}$$

and is the ratio of the number of message bits that can be sent to the number of codeword bits required to encode the full set of messages, and is achievable only if, for all n, there exists a code with $M = 2^{nR} s.t. \lambda^{(n)} \to 0 \, as \, n \to \infty$.

Definition 6. *The channel capacity is the supremum, over all coding strategies, of all achievable rates.*

2.1.1 Shannon's Noisy-Channel Coding theorem and some examples

Shannon's Noisy-Channel Coding theorem may be stated in a beautifully simple manner:

For a channel, the channel capacity is equal to the information capacity.

We may state this more precisely as follows:

Theorem 1 (Shannon's Noisy-Channel Coding theorem). *For a discrete memoryless channel, all rates below information capacity C are achievable. Conversely, any sequence of $(2^{nR}, n)$ codes with $\lambda^{(n)} \to 0$ must have $R \leq C$.*

This should astonish us. For any given degree of noise contamination of a communication channel, it is possible to communicate *nearly error-free* across the channel at any rate below its information capacity, which has the simple expression given in Equation 2.5! Beautiful as the proof of this theorem is, it would take us too far afield. Instead, we turn to calculating the information capacities of some canonical examples of noisy channels.

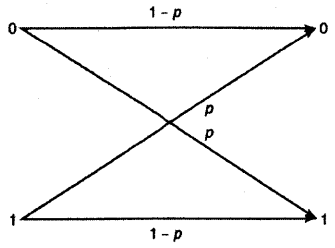

Figure 2-2: Binary Symmetric Channel – the input is inverted with probability p.

Example 1 (Binary Symmetric Channel). *On a Binary Symmetric Channel as shown in Figure 2-2, the input is inverted with probability p. So, some of the time, a* 0 *is transmitted as a* 1 *(and vice versa), and the receiver has no way of telling which of her bits were corrupted. Yet, Shannon's theorem guarantees error-free communication over this channel – up to the rate C! From Eqn 2.5, to find C we upper bound the mutual information:*

$$\begin{aligned} I(X;Y) &:= H(Y) - H(Y|X) \\ &= H(Y) - \Sigma p(x) \underbrace{H(Y|X=x)}_{H(p)} \qquad H(p) \text{ is a constant} \\ &= H(Y) - \underbrace{\Sigma p(x)}_{1} H(p) \\ &\leq \boxed{1 - H(p)} \end{aligned} \tag{2.7}$$

The capacity of a binary symmetric channel is $1 - H(p)$ *bits, and it is achieved on the uniform distribution on both input symbols.*

Example 2. *On a Binary Erasure Channel, shown in figure 2-3, bits are* lost *with some probability, rather than corrupted. We calculate the capacity of the binary erasure channel as follows:*

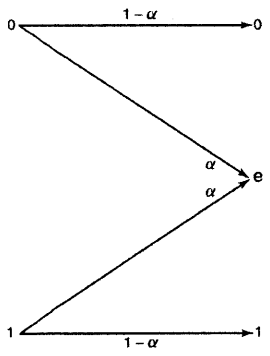

Figure 2-3: Binary Erasure Channel – the bits are erased with probability α.

$$C \equiv \max_{p(x)} I(X;Y) = \max_{p(x)} H(Y) - H(Y|X)$$
$$= \max_{p(x)} H(Y) - H(\alpha)$$

$H(Y)$, the entropy of the output symbol, depends both on the erasure probability, α, and the probability of each input symbol (let $\Pr(X = 1) = \pi$). Letting E be the event that Y gets erased, $\{Y = e\}$. Using the expansion $H(Y) = H(Y, E) = H(E) + H(Y|E)$, we have

$$H(Y) = H((1-\pi)(1-\alpha)), \alpha, \pi(1-\alpha)) = H(\alpha) + (1-\alpha)H(\pi) \tag{2.8}$$

Hence

$$C = \max_{p(x)} H(Y) - H(\alpha)$$
$$= \max_{\pi} (1-\alpha)H(\pi) \qquad \textit{from 2.8}$$
$$= \boxed{1-\alpha} \tag{2.9}$$

The capacity of a binary erasure channel depends only on the probability of erasure and is $1 - \alpha$ bits.

2.2 Quantum Channel Coding

We conclude this chapter with some remarks on Quantum Channel Coding, which generalizes classical Shannon theory by allowing inputs to be *quantum states.*

The basic model of a quantum channel does not differ too much from Shannon's classical one; there are the same steps of encoding, transmission (down a channel subject to quantum sources of noise) and decoding. A real-life example of a quantum channel is an optical fiber – a photon in some quantum state goes in, suffers noise and distortion in passing through the fiber, and if it is not absorbed and does not tunnel out, emerges in a transformed quantum state.

However, the fact that optical fibers are still not commonly used for quantum key distribution or communication of quantum information testifies to both practical and theoretical gaps in our understanding of quantum channels and how to use them. On the experimental side, it is a challenge to find a robust source of entangled photon pairs with high spectral brightness, broad wavelength coverage and a single-mode spatial output that is compatible with fiber networks or free-space operation [15]. On the theoretical side, the capacity region of a quantum channel is poorly characterized. Whereas a classical channel is completely characterized by a single capacity, quantum channels have four different capacities [16]:

1. A classical capacity, C, for transmitting classical information

2. A quantum capacity, Q, for transmitting intact quantum states

3. A classically-assisted quantum capacity, Q_2, for transmitting intact quantum states with the help of a two-way classical side-channel, and

4. An entanglement-assisted classical capacity, C_E, a quantum channel's classical capacity with the benefit of unlimited prior pure entanglement shared between the sender and receiver.

In most cases, only upper and lower bounds on these capacities are known, not the capacities themselves [16, 17]. Even the first of the four quantities above – C,

the capacity for classical information of a noisy quantum channel – has not been completely solved. What we do know is how to calculate the classical capacity for a channel ϵ assuming Alice encodes her messages using *product states* of the form $\rho_1 \otimes \rho_2 \otimes ...$ where each of the ρ_1, ρ_2 are potential inputs for one use of the channel ϵ. We call this the *product state capacity* and denote it by $C^{(1)}(\epsilon)$, and it is given by the following theorem:

Theorem 2 (Holevo-Schumacher-Westmoreland (HSW) theorem). *Let ϵ be a trace-preserving quantum operation. Define*

$$\chi(\epsilon) \equiv \max_{\rho_i, \rho_j} \left[S\left(\epsilon\left(\Sigma_j p_j \rho_j\right)\right) - \Sigma_j p_j S(\epsilon(\rho_j)) \right] \tag{2.10}$$

where the maximum is over all ensembles $\{p_j, \rho_j\}$ of possible input states ρ_j to the channel. Then $\chi(\epsilon)$ is the product state capacity for the channel ϵ.

(Here, $S(\rho)$ denotes the von Neumann, or 'quantum entropy', defined on density operators. It turns out that all the classical information theoretic measures in Equations 2.1 to 2.4 have classical analogs.) χ gives the product state capacity, but it was recently proven that, contrary to popular belief, using entangled states it is sometimes possible to exceed χ [18], and this is due to the non-additivity of χ, meaning that in general $(1/n)\chi(\epsilon^{\otimes n}) \geq \chi(\epsilon)$, where $\epsilon^{\otimes n}$ is the n-fold tensor product representing parallel uses of the channel. Hence, even the most basic of questions about quantum channels – 'Does allowing entangled signals improve the capacity?' – remains an open question to date.

Fortunately, we do not have to wade into this quicksand. We will see that we can get quite far just by proposing channels that only accept classical states. Therefore, whenever we use a metric for channel capacity in future we will use just the first of the above items – C, the classical capacity – and consider how it changes under different classes of resources. Since a closed-form expression for even the classical capacity of interference channels is still a work in progress (see [4] for discussions), and we are even further from this goal in the quantum setting (see [19], [20] for bounds), we leave it to future generations of readers to characterize a fully general version of our

channel.

2.3 Chapter Summary

This chapter covers rudiments of information theory for the purpose of understanding channel coding. Key concepts introduced include the Shannon-Weaver model of communication and Shannon's noisy-channel coding theorem, which provides a closed-form expression for the channel capacity. The power of this theorem to simplify the calculation of channel capacities was evidenced in the ensuing examples, which discussed the capacities of the binary symmetric channel and binary erasure channel. In the final section, we made the transition to *quantum* channel coding – the full characterization of which is still very much a work in progress – and glimpsed one of the founding theorems of the field, the HSW theorem. This is the extent to which we will treat this topic because all future channels will involve only classical inputs and outputs.

There is no *super-quantum* channel coding section, because the field does not exist. I hope that the results from the rest of this thesis will be the first of many works that could motivate a future student to write one.

Chapter 3

Channel I: Super-Quantum Superiority on Interference Channels

This thesis characterizes the power of PR boxes with respect to channel coding, but it is just one chapter of an ongoing story. We believe that understanding the power of PR boxes will yield insight into the non-locality of quantum mechanics. In 2005, Cerf, Gisin, Massar and Popescu demonstrated a sense in which super-quantum non-locality subsumes quantum non-locality – they showed that a PR box could simulate the correlations obtained from any bipartite measurement of a maximally entangled pair of qubits without communication[21]. Clearly, the reverse direction of simulation is impossible – correlations due to PR-boxes are provably more non-local than those from entanglement. A natural question, then, is whether this non-locality separation has any import on the efficiency of physical tasks.

Previous studies have answered in the affirmative, by identifying two domains in which PR boxes give an edge over quantum strategies: cryptographic functionalities such as unconditionally secure bit commitment and oblivious transfer[22, 23]) and two-party computation of functions (see [24]). In particular, van Dam proved an astonishing result that super-quantum correlations reduce all distributed computations between two parties, no matter how complex, to procedures that require only one bit of communication! There has also been some work by Broadbent, Méthot and Brassard ([25, 26]) on how non-local boxes can yield an advantage in pseudo-

telepathy games – games that can only be won with (sometimes large amounts of) entanglement.

The interference channel whose analysis forms the crux of this chapter therefore brings to light a new field where super-quantum non-locality yields an advantage over classical and quantum (entanglement-aided) strategies: channel coding. In future, all references to 'classical' strategies shall imply strategies where the senders are allowed to share no communication but may discuss a strategy before-hand.

3.1 Interference channel preliminaries

We intend to exhibit a two-sender, two-receiver interference channel for which a super-quantum strategy is provably better than a classical strategy and entanglement-aided strategies. The theory of two sender-receiver pair interference channels is a simple extension of the one sender-receiver pair case which was discussed in detail in the previous chapter. Here we will provide only essential details:

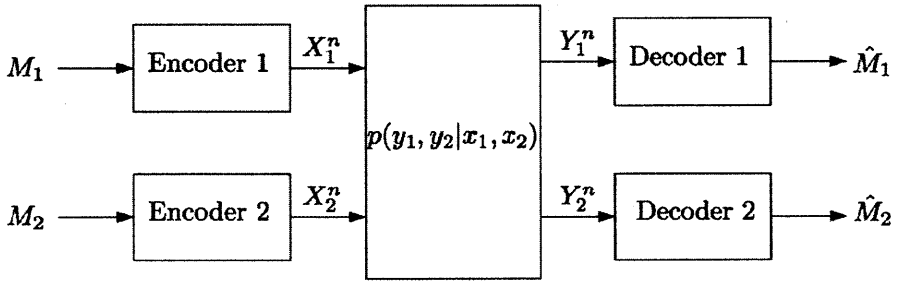

Figure 3-1: General model of a two sender-receiver pair communication system. Figure taken from [4].

The basic model of a two sender-receiver pair interference channel is depicted in figure 3-1, and the following is a condensed version of the discussion in [4] relating to such channels. Such a channel is denoted $(\mathcal{X}_1 \times \mathcal{X}_2, p(y_1, y_2|x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$. A $(2^{nR_1}, 2^{nR_2}, n)$ code for this channel consists of:

- Two message sets $[1 : 2^{nR_1}]$ and $[1 : 2^{nR_2}]$

- Two encoders, where encoder 1 assigns a codeword $x_1^n(m_1)$ to each message $m_1 \in [1 : 2^{nR_1}]$ (respectively encoder 2 assigns $x_2^n(m_2)$ for $m_2 \in [1 : 2^{nR_2}]$).

- Two decoders, where decoder 1 uses a *decoding rule* to assign an estimate $\hat{m}_1$ or an error message e to each received sequence y_1^n, and decoder 2 does the same (ie. assigns $\hat{m}_2$ or e).

A rate pair (R_1, R_2) is said to be achievable for this channel if there exist a sequence of $(2^{nR_1}, 2^{nR_2}, n)$ codes such that $\lim_{n\to\infty}[P_e^{(n)} \equiv P\{(\hat{M}_1, \hat{M}_2) \neq (M_1, M_2)\}] = 0$. The capacity region, $\mathscr{C}$, is the closure of the set of achievable rate pairs (R_1, R_2).

For this channel, the relevant figures of merit are the capacity region, defined above, and the sum-capacity, $C_{\text{sum}} = \max\{R_1+R_2 : (R_1, R_2) \text{ simultaneously achievable}\}$. Note that $C_{\text{sum}} \neq \max_{p(x_1)} R_1 + \max_{p(x_2)} R_2$ in general, because C_{sum} is the sum of rates attainable *simultaneously*. Whenever we speak of the 'capacity' of a channel, we shall refer to its sum-capacity.

3.2 Channel I

Consider the following channel, for which, on each use of the channel, the senders send two bits and the channel outputs one bit to each receiver.

$X_1 \backslash X_2$	**00**	**01**	**10**	**11**
00	00	11	01	10
01	11	00	10	01
10	10	01	00	11
11	01	10	11	00

Table 3.1: Channel I: The senders each send two-bit codewords, and the two-bit entries in the table correspond to what is received – the first bit goes to Y_1 and the second, Y_2.

The maximum possible sum-capacity (no matter what resources are used) is 2. The receivers each receive one bit, so the senders cannot possibly communicate more than one bit each, and in fact, $C_{\text{sum}} = 2$ only if there exists a strategy where the receiver always gets *exactly* the bit that the sender intends to send.

We will show in Lemma1 that the following is the classical capacity region of this interference channel:

$$\begin{aligned} R_1 &\leq 1 \\ R_2 &\leq 1 \\ R_1 + R_2 &\leq 1. \end{aligned} \tag{3.1}$$

If we allow the senders to share a PR box, the sum-capacity equation, 3.1, becomes

$$R_1 + R_2 \leq 2 \tag{3.2}$$

and we provide a coding strategy that exactly achieves this capacity.

Lastly, Lemma 3 shows that even if the senders are allowed to share $2 \times n$ entanglement, the sum-capacity is strictly outer-bounded by 2, that is $R_1 + R_2 < 2$. We do not know the exact capacity under this strategy (ie. 2 may not be the best outer-bound), but these three facts taken together suffice to show the following separation on this channel:

$$\mathbf{C_{classical}}, \mathbf{C_{quantum}} < \mathbf{C_{super\text{-}quantum}}.$$

3.2.1 Capacity of Channel I with no assistance

Channel I is quite puzzling at first blush – it takes *two*-bit inputs, so if the senders can ultimately communicate only one bit, the second bit seems redundant. Might the redundancy improve communication? It seems reasonable to search for a good code that exploits this redundancy in such a way that both receivers may communicate simultaneously at a high rate. Two facts about classical coding are immediately evident:

- A uniform probability distribution over the entire input alphabet, for both senders, is a recipe for disaster: taking the marginal probability distribution for the first pair (by summing up over all input-outputs for the second pair that

correspond to the same input-output for the first) results in what is effectively a binary symmetric channel for both pairs with bit-flip probability $p = 0.5$. Plugging this into the expression 2.7 for the capacity of the binary symmetric channel and noting that $H_2(0.5) = 1$, the joint rate is 0.

- More optimistically, there is at least one obvious strategy that gives a joint rate of 1 (and therefore 1 is an inner bound on the sum capacity). If the pairs are allowed to elaborate a strategy beforehand, then A_2 may just send 00 while A_1 encodes message bit 0 as 00 and message bit 1 as 01; then B_1 receives exactly the bit that A_1 intended to send. Thus the joint capacity is lower bounded by 1.

Is it possible to do *better* than $R_1 + R_2 = 1$? Surprisingly, Lemma 1 combined with some numerical simulations shows the answer is no.

Lemma 1 (Capacity of Channel I with no resources). *If the senders are limited to a classical strategy with no aid from communication, entanglement or PR boxes, on the given channel the sum-capacity is strictly outer-bounded:*

$$R_1 + R_2 < 2.$$

Proof. We show that $R_1 := I(X_1 : Y_1) = 1$ implies $R_2 := I(X_2 : Y_2) < 1$.

Suppose $I(X_1 : Y_1) = 1$. Using the chain rule for mutual information shows that $I(X_2 : Y_1|X_1) = 0$.

$$I(X_1 : Y_1) = 1 = \underbrace{I(X_1, X_2 : Y_1)}_{\text{takes on maximal value, 1}} - \underbrace{I(X_2 : Y_1|X_1)}_{=0} \tag{3.3}$$

This condition $I(X_2 : Y_1|X_1) = 0$ is a rather powerful one. Interpreted in the information-theoretic sense, it states that the first receiver's output, Y_1, cannot possibly distinguish between the possibilities for the second sender's message, X_2 – which implies that the alphabet that sender 2 was using must have been a special subset of all the symbols available. This gives a restriction on the second sender's alphabet

set. For instance, examining the first row of the table, if Sender 1 sends 00 (i.e. if 00 is part of Sender 1's alphabet), then one of the following must be true:

1. Sender 2 sends either 00 or 10 with equal probability (both resulting in the output $Y_1 = 0$) but never sends 01 or 11.

2. Sender 2 sends either 01 or 11 with equal probability (both resulting in the output $Y_1 = 1$) but never sends 00 or 10.

Repeating the analysis for the other 3 possible choices of Sender 1's message, we get the following sets of alphabets for the two senders:

1. Sender 1's alphabet is $\{00, 01\}$ and Sender 2's alphabet is either $\{00, 10\}$ or $\{01, 11\}$. ·

2. Sender 1's alphabet is $\{10, 11\}$ and Sender 2's alphabet is either $\{00, 11\}$ or $\{10, 01\}$.

We analyze these four cases individually. Since X_1 and X_2 are not allowed to communicate during the sending of the messages, they must choose an alphabet at the start and stick to it. Consequently, only one of these four cases can hold. Here we show that if X_1 uses the alphabet $\{00, 01\}$ and X_2 uses $\{00, 10\}$ (Table 3.2 depicts this schematically), then the second sender-receiver pair cannot communicate perfectly, and therefore $R_2 < 1$ as we asserted. The same turns out to be true for the other 3 cases.

To get $I(X_1 : Y_1) = H(X_1) - H(X_1|Y_1) = 1$ when there are only two options for X_1, the first term must take its maximal value of 1, which can only happen if X_1 is uniformly distributed over $\{00, 01\}$. Let X_2 send 00 with probability c and 01 with probability $1 - c$. This is shown on the left in Table 3.2. Since we will be interested in calculating $I(X_2 : Y_2)$, we also calculate the input-output probability distribution experienced by sender-receiver pair 2, shown on the right in Table 3.2.

$X_1 \backslash X_2$	**00**	**10**
00	00	01
01	11	10

$X_2 \backslash Y_2$	**0**	**1**
00	$\frac{c}{2}$	$\frac{c}{2}$
10	$\frac{1-c}{2}$	$\frac{1-c}{2}$

Table 3.2: Left: reduced alphabets of senders and resulting output to the receivers (in the format Y_1Y_2). Right: Joint probability distribution experienced by the second sender-receiver pair on this coding scheme.

Referring to the right side of Table 3.2, we obtain

$$\begin{aligned} I(X_2 : Y_2) &= H(X_2) + H(Y_2) - H(X_2, Y_2) \\ &= [-c \log c - (1-c)\log(1-c)] + 1 \\ &\quad - \left[2\left(-\frac{c}{2}\log\frac{c}{2}\right) + 2\left(-\frac{1-c}{2}\log\frac{1-c}{2}\right)\right] \\ &= 0 \end{aligned} \tag{3.4}$$

(Note that to calculate $H(X_2)$ and $H(Y_2)$, we had to use the marginal probability distributions over Y_2 and X_2 respectively, a step that we omitted.)

We have therefore shown that $I(X_1 : Y_1) = 1$ implies that $I(X_2 : Y_2) = 0$, so that $I(X_1 : Y_1) + I(X_2 : Y_2) = 2$ will never be achieved. **Put another way, perfect coding between one pair implies that the other pair can do no better than random guessing.** □

Of course, this proof does not rule out the possibility that if one of the sender-receiver pairs is willing to accept a sub-optimal (less than 1) rate, the other pair will be able to attain a high rate such that $R_1 + R_2 > 1$. However, we ran an algorithm based on modified gradient descent that shows that in fact, the maximum joint rate is *exactly* 1. This algorithm is given in pseudocode here (Algorithm 1). The inputs to the algorithm are two vectors $\vec{x_1} := (a_1, b_1, c_1, d_1), \vec{x_2} := (a_2, b_2, c_2, d_2)$, such that the square of the entries in the first vector $\{a_1^2, b_1^2, c_1^2, d_1^2\}$ represents the probabilities of sender 1 sending $\{00, 01, 10, 11\}$ respectively, and correspondingly $\{a_2^2, b_2^2, c_2^2, d_2^2\}$ for sender 2. The modification to the usual gradient descent algorithm was to respect

the constraints

$$a_1^2 + b_1^2 + c_1^2 + d_1^2 = 1 \ ; \ a_2^2 + b_2^2 + c_2^2 + d_2^2 = 1.$$

To do this, we treated the problem of simultaneous gradient descent where the component vectors had to lie on two 4-D unit spheres. A pseudocode for our algorithm is given below.

Algorithm 1 Finds the maximum value of the function $I(X_1 : Y_1) + I(X_2 : Y_2)$ over all input distributions

Modified-Gradient-Descent($\vec{x}$)

$f(\vec{x_1}, \vec{x_2}) := -I(X_1; Y_1) - I(X_2; Y_2)$ ▷ Define objective function

$\vec{g_1} := \vec{\nabla}_{x_1} f \ ; \ \vec{g_2} := \vec{\nabla}_{x_2} f$ ▷ Define gradient of function

Initialize $x_1, x_2, tol, maxiter$

while $iter < maxiter$ and $dx > tol$ **do**

Evaluate $\vec{g_1}(\vec{x_1}, \vec{x_2}) \ ; \ \vec{g_2}(\vec{x_1}, \vec{x_2})$

$\vec{h_1} \leftarrow \vec{g_1} - (\vec{g_1} \cdot \vec{x_1})\vec{x_1} \ ; \ \vec{h_2} \leftarrow \vec{g_2} - (\vec{g_2} \cdot \vec{x_2})\vec{x_2}$

$\alpha_1 \leftarrow \frac{h_1^2}{h_1^2 + h_2^2} \ ; \ \alpha_2 \leftarrow \frac{h_2^2}{h_1^2 + h_2^2}$

$\vec{n_1} \leftarrow \frac{\vec{h_1}}{|\vec{h_1}|} \ ; \ \vec{n_2} \leftarrow \frac{\vec{h_2}}{|\vec{h_2}|}$

$\phi' \leftarrow \arg\min_{\phi} f(\cos(\alpha_1\phi)\vec{x_1} + \sin(\alpha_1\phi)\vec{n_1}, \cos(\alpha_2\phi)\vec{x_2} + \sin(\alpha_2\phi)\vec{n_2})$

$\vec{x_1'} \leftarrow \cos(\alpha_1\phi')\vec{x_1} + \sin(\alpha_1\phi')\vec{n_1}; \ \vec{x_2'} \leftarrow \cos(\alpha_2\phi')\vec{x_2} + \sin(\alpha_2\phi')\vec{n_2}$

$dx \leftarrow \sqrt{(x_1'^2 - x_1^2 + x_2'^2 - x_2^2)}$

end while

3.2.2 Capacity of Channel I with super-quantum assistance

In the above section, we imposed the requirement $I(X_1 : Y_1) = 1$, found two conditions specifying four allowed combinations of senders' alphabets and concluded that one of four cases must hold. We established that at most one of those two conditions may be satisfied if we limit ourselves to a classical strategy.

Now if we allow ourselves to wonder: 'What must hold true for the sum-capacity to be saturated, that is, $I(X_1 : Y_1) = 1$ AND $I(X_2 : Y_2) = 1$?', we may recycle our above analysis. This gives us two additional conditions to our existing two, and we list all of them below:

1. If $X_1 \in \{00, 01\}$, either $X_2 \in \{00, 10\}$ or $X_2 \in \{01, 11\}$.

2. If $X_1 \in \{10, 11\}$, either $X_2 \in \{01, 10\}$ or $X_2 \in \{00, 11\}$.

3. If $X_2 \in \{00, 01\}$, either $X_1 \in \{00, 10\}$ or $X_1 \in \{01, 11\}$.

4. If $X_2 \in \{10, 11\}$, either $X_1 \in \{01, 10\}$ or $X_1 \in \{00, 11\}$.

Suppose now that we unshackle ourselves from classical intuition and permit the two senders to coordinate their input alphabets in real-time, perhaps by using a non-classical resource. In that case, there are only two strategies that could fulfil *all four* conditions at once, shown in Table 3.3:

(a) Later we will see that there exists a super-quantum strategy that produces only this distribution of inputs.

$X_1 \backslash X_2$	**00**	**01**	**10**	**11**
00	**00**	11	**01**	10
01	11	**00**	10	**01**
10	**10**	01	00	**11**
11	01	**10**	**11**	00

(b) An alternative strategy that always enables the senders to get their messages across.

$X_1 \backslash X_2$	**00**	**01**	**10**	**11**
01	00	**11**	01	**10**
00	**11**	00	**10**	01
11	10	**01**	**00**	11
10	**01**	10	11	**00**

Table 3.3: Each shaded box corresponds to a message pair that sent simultaneously by the senders. These are the only two coding strategies that will allow both $I(X_1 : Y_1) = 1$ and $I(X_2 : Y_2) = 1$.

The following lemma shows that a PR-box is *exactly* such a non-classical resource that would enable both sender-receiver pairs to communicate perfectly:

Lemma 2 (Capacity of Channel I with super-quantum resources). *If the senders are allowed to share a PR-box, the capacity of the given channel is exactly 2. This is the algebraically maximal sum-capacity of the channel.*

We have all but spelt out our super-quantum strategy. It remains to put all of the above together as follows: Our encoding strategy is that on each use of the channel, the senders each encode each message bit into a two-bit codeword by concatenating

that bit with the output of the PR box (i.e. Sender 1 sends $X_1 = (a, x)$ and Sender 2 sends $X_2 = (b, y)$).

However, the PR box only produces certain outputs on a given input (we reiterate that it takes into account both parties' inputs), and these input-output combinations fulfilling $a \oplus b = xy$ are as summarized in Table 1.1, which we reproduce below. Scrutinizing that table carefully and comparing it to Table 3.3a reveals that the resulting encoded message pairs (created using the above procedure) are special for our channel: they are exactly the combinations whereby receiver 1 and receiver 2 respectively receive the original 1-bit messages that sender 1 and sender 2 intended to send. Hence, this super-quantum strategy enables perfect message transmission.

(r, s) (Questions)	(a, b) (Answers)
(0,0)	(1,1), (0,0)
(0,1)	(1,1), (0,0)
(1,0)	(1,1), (0,0)
(1,1)	(0,1), (1,0)

Table 3.4: Winning conditions for CHSH game, as well as input-output condition on PR-box: $a \oplus b = r \wedge s$

Corollary 1 (Capacity of Channel I if senders may share 1 bit of communication). *If senders are allowed to share one bit of communication, they will achieve a joint rate of 2.*

Proof. The technique described in the proof of the previous lemma all but spells out a classical strategy that achieves a joint rate of 2 if the senders are allowed to share 1 bit of communication. The strategy is as follows: Sender 1 and Sender 2 again have their message bits, m_1 and m_2. However, this time, Sender 1 encodes her message bit by duplicating it. She then uses her one bit of communication by sending this message bit to Sender 2. Sender 2, having obtained this information, attempts to replicate the action of the PR box. He encodes his message bit by padding it with the unique bit that ensures that his input to the channel is exactly the input that he would have provided, had the two senders used the PR-box strategy, AND sender 1 received from the PR-box exactly what she had put in.

Since this strategy achieves (deterministically) exactly the same input sets as the PR-box assisted strategy described above, it too achieves a joint rate of 2. □

3.2.3 Capacity of Channel I with quantum assistance

Since the key to our strategy from the previous sub-section was to allow Alice and Bob to share a non-classical resource that would enable them to coordinate their inputs, one might reasonably wonder if perhaps resorting to super-quantum resources is an overkill. Could one achieve the very same effect by restricting Alice and Bob's shared resource to be a purely quantum one? In this section we show that even if we allow the two senders to share an entangled pair – or indeed, any entangled state of dimension $2 \times n$, they could not achieve a perfect rate.

Lemma 3 (Classical sum-capacity with quantum resources). *A classical sum-capacity of 2 on the given channel is not achievable even if the senders are allowed to share an entangled quantum state of dimension* $2 \times n$.

To prove this, one may model a general quantum strategy as follows:

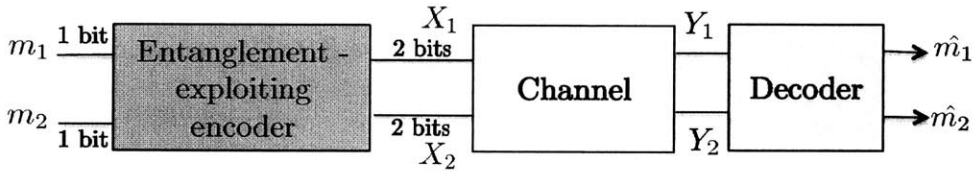

Figure 3-2: Model of a quantum communication system over this channel.

Namely, the two senders share an entangled state $|\Psi\rangle$, independently choose a POVM depending on their message bit m_1, apply that POVM to the share of the entangled state, and apply a rule that maps the measurement outcome of the POVM to a 2-bit input to the channel. These bits go through the channel and the output of the channel is decoded by each of the two receivers.

This is indeed the most general form of a quantum communications strategy; Naimark's theorem guarantees that the POVM formalism is the most general form of measurement, and the most general decoder that the receivers can employ is one

that looks at both the outputs from the channel. This includes as a special case a restricted decoding strategy where receivers do not communicate. In fact, this model (and the proof it inspires) is in very much the same spirit as the model in [25], which was used to prove that there cannot exist a two-player pseudo-telepathy game (that is, a game on which a quantum strategy wins with 100% probability) of dimension 2×2. It turns out that the two proofs are similar because it is possible to map the actions of a quantum winning strategy for the senders in this case to a quantum winning strategy for a pseudo-telepathy game.

Proof. For this proof, we borrow notation from [25]. Let $\mathcal{P}$ denote the set of all POVMs acting on a single qubit. For each $E \in \mathcal{P}$, let $\{E_i\}$ denote the corresponding set of positive matrices that form that POVM, with $\Sigma_i E_i = \mathbb{1}$. Let O_E denote the set of all outcomes for the POVM, which shall be assumed to be a set of integers from which the index i in E_i is drawn. Any quantum strategy for communication can be defined in terms of the following mappings (the left column corresponds to the actions of the first sender-receiver pair, and the right column the second):

$$\begin{array}{cccc} \mathcal{X} : m_1 \to \mathcal{P} & & \mathcal{Y} : m_2 \to \mathcal{P} & \text{Choose POVM \& measure} \\ \mathcal{A} : m_1 \times O \to X_1 & & \mathcal{B} : m_2 \times O \to X_2 & \text{Encoding rule} \\ \mathcal{D}_1 : Y_1 \to \hat{m_1} & & \mathcal{D}_2 : Y_2 \to \hat{m_2} & \text{Decoding rule} \end{array} \quad (3.5)$$

Our goal is to show that that if there exists a quantum strategy that achieves rate 2 (ie. perfect coding), there is a classical strategy that achieves the same rate. However, since the last section established that there is not a classical strategy that achieves perfect coding, there cannot be a quantum one.

The key to the proof is the reasonable assumption that any quantum decoding strategy depends only on the bits that the receivers receive, that is, whenever bits $Y_1 Y_2$ are produced by the channel, each of the 4 possible inputs (X_1, X_2) that correspond to that output have to be produced by the encoder on the *same* pair of message bits m_1, m_2. Furthermore, a rate of 2 rules out any probabilistic decoding strategy,

such as the one employed in Shannon's proof of the noisy coding theorem. These considerations define the map given in Table 3.5 between the original message bits and their allowed encoding.

$(\mathbf{m_1}, \mathbf{m_2})$	$(\mathbf{X_1}, \mathbf{X_2})$
$(m_1, m_2)_a$	$(00, 00), (01, 01), (10, 10), (11, 11)$
$(m_1, m_2)_b$	$(00, 01), (01, 00), (10, 11), (11, 10)$
$(m_1, m_2)_c$	$(00, 10), (01, 11), (10, 01), (11, 00)$
$(m_1, m_2)_d$	$(00, 11), (01, 10), (10, 00), (11, 01)$

Table 3.5: Map between message bits and their encoding. $(m_1, m_2)_a, (m_1, m_2)_b, (m_1, m_2)_c, (m_1, m_2)_d$ must correspond to some permutation of the message set $\{00, 01, 10, 11\}$.

Therefore, all we are asking of our classical strategy is that, for any combination of message bits, it should encode them as some subset of the allowed encodings in the right column of the corresponding row – since any of those encodings, if they were produced by the quantum strategy, would suffice for perfect decoding. That is, our classical strategy should *never* produce an illegal output even though some legal outputs may never occur. We may then use the method described in [25] to devise such a classical strategy. □

Lastly, we will quickly sketch the method given in [25] of designing a classical strategy that never produces an output that would have had zero probability of being produced by the quantum strategy. The method will follow after the subsequent lemmas:

Lemma 4. *For any two-sender-receiver pair communication strategy that relies on the senders sharing some state $|\Phi\rangle$ of dimension 2×2, there exists a communication strategy that achieves the same rate where the senders are restricted to sharing a state of the form $|\Psi\rangle = \alpha|00\rangle + \beta|11\rangle$, where α and β are well-chosen positive real numbers.*

Proof. The key idea is to re-write $|\Phi\rangle$ in terms of its Schmidt decomposition, and then apply a unitary transformation to get $|\Psi\rangle$. Then, the senders may apply the quantum strategy whose existence we have assumed. More precisely, there exist

orthogonal bases $\{|A_0\rangle, |A_1\rangle\}$ for Sender 1 and $\{|B_0\rangle, |B_1\rangle\}$ for Sender 2 such that $|\Phi\rangle$ can be rewritten as

$$|\Phi\rangle = \alpha |A_0\rangle |B_0\rangle + \beta |A_1\rangle |B_1\rangle .$$

From there it is easy to see that Sender 1 may apply the unitary transformation $|A_0\rangle \langle 0| + |A_1\rangle \langle 1|$, and sender 2 may apply the unitary transformation $|B_0\rangle \langle 0| + |B_1\rangle \langle 1|$, to their qubits, to transform $|\Psi\rangle$ into $|\Phi\rangle$. Any such unitary U is completely accounted for in our model of communication in 3.5 by applying it to the POVMs M_i that the senders choose for their states (which preserves its POVM properties), that is, using the property $U|\Phi\rangle = |\Psi\rangle \rightarrow \langle\Phi| M_i |\Phi\rangle = \langle\Phi| U M_i U^\dagger |\Phi\rangle$. □

Since the following two lemmas are almost identical to the ones in [25], we merely cite them and leave the reader to refer to [25] for their proofs.

Lemma 5. *For any two-party quantum communication protocol that uses an entangled state of dimension $d_A \times d_B$, there exists a two-party quantum communication protocol that uses a state of dimension $d \times d$ where $d := \min(d_A, d_B)$.*

This justifies the audaciously general claim made in Lemma 3 that *no* quantum state of dimension $2 \times d$ could possibly enable a perfect joint rate for communication. The proof is similar to the proof of Lemma 4 and relies on the following fact from the Schmidt decomposition: if H_1 and H_2 are Hilbert spaces of dimensions n, m respectively, and we assume without loss of generality that $n \geq m$, for any vector $w \in H_1 \otimes H_2$, there exist orthonormal bases $\{u_i, 1 \leq i \leq n\}$ for H_1 and $\{v_j, 1 \leq j \leq m\}$ for H_2 respectively such that

$$w = \Sigma_{i=1}^{m} \alpha_i u_i \otimes v_i. \tag{3.6}$$

Lemma 6. *Any POVM can be written in a way such that all its elements are proportional to one-dimensional projectors. Each such projector can be re-written in the*

form

$$P = \begin{pmatrix} \cos^2(\theta) & e^{-i\phi}\sin(\theta)\cos(\theta) \\ e^{i\phi}\sin(\theta)\cos(\theta) & \sin^2(\theta) \end{pmatrix} \tag{3.7}$$

for appropriate angles $0 \leq \theta \leq \frac{\pi}{2}$ *and* $0 \leq \phi \leq 2\pi$. *Since this representation is unique, we may associate each such projector with a three-dimensional unit vector* $\vec{v} = (\sin(2\theta)\cos(\phi), \sin(2\theta)\sin(\phi), \cos(2\theta))$.

Finally, the classical strategy promised three lemmas ago is described. Thanks to Lemma 4, we may assume that the two senders are using an entangled state of the form $|\Psi\rangle = \alpha|00\rangle + \beta|11\rangle$, where α and β are strictly positive real numbers.

Suppose a quantum strategy exists and the POVMs applied by the two senders, $M^x := \mathcal{X}(x) = \{\gamma_i^x P_i^x\}$ and $N^y := \mathcal{Y}(y) = \{\gamma_j^y Q_j^y\}$ have been fixed beforehand for each $x, y \in \{0, 1\}$. We will show that any measurement outcome (i, j) on $|\Psi\rangle$ as described in the first row of Equations 3.5 can be replicated perfectly classically. The probability of getting the tuple (i, j) is:

$$\begin{aligned} \Pr[i,j] &= \langle\Psi| (\gamma_i^x P_i^x) \otimes (\gamma_j^y Q_j^y) |\Psi\rangle \\ &= \gamma_i^x \gamma_j^x \left[\alpha^2 \cos^2(\theta_i^x)\cos^2(\theta_j^y) + 2\alpha\beta\left[\cos(\phi_i^x + \phi_j^y)\sin\theta_i^x \cos\theta_i^x \sin\theta_j^y \cos\theta_j^y\right]\right. \\ &\left. +\beta^2 \sin^2(\theta_i^x)\sin^2(\theta_j^y)\right] \\ &= \gamma_i^x \gamma_j^x (a^2 + b^2 + 2abc) \end{aligned} \tag{3.8}$$

where $a := \alpha\cos(\theta_i^x)\cos(\theta_j^y)$, $b := \beta\sin(\theta_i^x)\sin(\theta_j^y)$ and $c := \cos(\phi_i^x + \phi_j^y)$. Using the AM-GM inequality we may show that $\Pr[i, j]$ can only vanish if one of the following two things are true of the POVMs used by the two senders ($\{\gamma_i^x P_i^x\}, \{\gamma_j^y Q_j^y\}$):

- $a = b = 0$

  Attained if $\theta_i^x = 0, \theta_j^y = \pi/2$ or vice versa – that is, either P_i^x or Q_j^y belongs to neither hemisphere.

- $a = b$ and $c = -1$.

  Attained if $\phi_i^x + \phi_j^y = \pi$(both projectors in eastern hemisphere) or $\phi_i^x + \phi_j^y = 3\pi$

(both projectors in western hemisphere).

But all our classical strategy needs to do is to choose a classical tuple, (i, j), such that the corresponding quantum POVM elements, P_i^x and Q_j^y, would not fulfill either of these conditions. To do this, it suffices for Sender 1, knowing $M^x := \{\gamma_i^x P_i^x\}$, to choose an i such that P_i^x belongs to the eastern hemisphere and for Sender 2, knowing $N^y := \{\gamma_j^y Q_j^y\}$, to choose a j such that Q_j^y belongs to the western hemisphere (without actually measuring anything). This is always possible since POVM elements have to sum to the identity. They may then carry out the (classical) mappings $\mathcal{A}$ and $\mathcal{B}$ on their message bits and POVM 'outcomes' as per normal.

This completes the proof of Lemma 1.

3.3 Discussion

In this chapter, we have demonstrated that super-quantum resources are capable of enhancing the capacity of an interference channel.

It is interesting to consider whether the super-quantum strategy outlined above can be efficiently approached even if the senders must share a noisy, instead of perfect, PR-box. A noisy box is one of the form

$$P_\epsilon = \epsilon P^{PR} + (1 - \epsilon) P^C \tag{3.9}$$

where P^C denotes the fully-correlated box $P^C(ab|xy) = \frac{1}{2}$ if $a \oplus b = 0$, ie. the box always produces either 00 or 11 on all inputs. Thanks to the protocol of non-locality distillation due to [27], which increases the CHSH value of a box via n successive operations in such a way that $NL[P_\epsilon^n] > NL[P_\epsilon]$, one can almost certainly approach a perfectly non-local (ie. PR) box in the asymptotic limit; the only question is whether this can be done *efficiently*, that is, with a number of operations that scales efficiently with the desired precision.

Of note is also the observation made in Corollary 1 that the use of a non-local box can be replaced with 1 bit of communication. This, combined with the observation in

[26] that classical strategies with n bits of communication can always be transformed into protocols with n uses of a non-local box, suggests that there might exist a 1 to 1 tradeoff between non-locally correlated bits and bits of communication used as a resource. A partial answer to whether the two resources are interchangeable can be found in [28], where the authors proved that there exist non-signaling correlations that can be generated from a single bit of communication which cannot be simulated with an NLB.

Lastly, although we have not proven explicitly that there exists a quantum strategy on this channel that beats the best classical strategy, we believe that this channel indeed exhibits such behavior and the proof should be within reach quite easily.

3.4 Chapter summary

This channel introduced in this chapter demonstrates the following separations in classical capacities (where C_{resource} is the classical capacity of the channel when the senders are permitted that resource):

$$\mathbf{C_{classical}}, \mathbf{C_{quantum}} < \mathbf{C_{super\text{-}quantum}}.$$

In the following chapter, we will introduce a closely-related channel that demonstrates a strict separation between all three classes.

Chapter 4

Channels II and III: New Conjectured Separations on Interference and Multiple-Access Channels

The previous chapter demonstrated a channel for which there exists a simple strategy using PR-boxes that achieves perfect communication, a feat not possible using just a classical strategy or $2 \times n$ quantum entanglement between the senders. In this chapter, we present two related two sender-receiver pair channels that are closely related to Channel I, but that demonstrate the following new conjectured separations:

$$\text{Channel II: } \mathbf{C_{classical}} < \mathbf{C_{quantum^1}} < \mathbf{C_{super\text{-}quantum}}$$

$$\text{Channel III: } \mathbf{C_{classical}} < \mathbf{C_{super\text{-}quantum^2}}$$

This chapter differs from the previous one in that although the proofs are not as rigorous, the examples provided are inventive and yield significant insight. In fact, the reader may find this channel more interesting for that reason. It is believed that this rigor-to-intuition tradeoff is worthwhile; many of the proof techniques from the previous chapter should carry over, and in exchange for that, we have made inroads into obtaining super-quantum enhancements for a new type of channel, the multiple-access channel. Together with the interference channel, this is one of the building blocks of Network Information Theory.

4.1 Channel II

Channel II is a simple modification to our existing one; it is a channel for which some outputs are erased with probability 1/2, and some with probability 1.

$X_1 \backslash X_2$	**00**	**01**	**10**	**11**
00	00/ee	ee	01/ee	ee
01	ee	00/ee	ee	01/ee
10	10/ee	ee	ee	11/ee
11	ee	10/ee	11/ee	ee

Table 4.1: Channel II: a variation on Channel I in which the channel outputs not corresponding to the PR-box-encoded inputs are erased with probability 1, and the channel outputs corresponding to the PR-box-encoded inputs are erased with probability $p = 0.5$. Erased bits are denoted by 'e'.

Conjecture 1. $C_{classical} < C_{quantum} < C_{super\text{-}quantum}$

4.1.1 A super-quantum assisted strategy on Channel II

Lemma 7. *There exists a super-quantum-assisted strategy on Channel II that achieves* $R_1 + R_2 = 1$.

Proof. The **super-quantum** coding strategy is exactly the same as in the previous section. This is best visualized by comparing our channel in Table 4.3 to the set of encoded messages produced by the PR-box strategy from the previous chapter, summarized in Table 3.3a – the encoding only produces the channel inputs whose outputs are erased with probability $\frac{1}{2}$. Since our strategy guarantees that outputs that are not erased are perfectly decoded by each receiver, this amounts to a binary erasure channel for each sender-receiver pair with erasure parameter 0.5. Using the expression 2.9 for the binary erasure channel's capacity, this amounts to a joint rate of $2 \times (1 - 0.5) = 1$. □

4.1.2 An entanglement-assisted strategy on Channel II

There exists a simple strategy that makes use of an entangled pair in such a way that the senders play a CHSH game to communicate. This enables both sender-receiver

pairs to communicate with rate $\cos^2(\frac{\pi}{8}) \approx 0.854$. This is still less than the joint rate achievable with a PR-box.

Lemma 8. *There exists a super-quantum-assisted strategy on Channel II that achieves* $R_1 + R_2 = \cos^2(\frac{pi}{8}) \approx 0.854$.

Proof. In place of the PR-box from the previous super-quantum strategy, let the two senders share the CHSH entangled pair. We may then map the encoding step to a CHSH game (section 1.3.1), where the senders are the players in the game. The message bits of the senders play an analogous role to the referee's questions in the game, and the outputs that the players would produce using the CHSH strategy correspond to the bits that the senders concatenate with their message bits to form the channel inputs. Just as previously, the decoding step is the identity; the receivers directly use the channel outputs (if they are not erased) as their estimates for the senders' message bits.

More concretely, let the two senders share an entangled pair, $|\Psi\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$, which is the same state that they can use to win the CHSH game. Recall that the winning condition of the CHSH game is that

$$a \oplus b = r \wedge s \tag{4.1}$$

and that there exists a good quantum strategy that makes use of this entangled pair. That is, let r = player 1's question, s = player 2's question, a = player 1's response, b = player 2's response; this strategy fulfills Equation 4.1 with probability $\cos^2(\frac{\pi}{8})$. But observe that eqn: wincond also governs the input-output distribution (inputs: r, s, outputs: a, b) of a PR-box, the corner-piece of our perfect coding strategy from the last chapter. This is a huge hint; with just the entangled pair and no PR-box, we can try to mimic our super-quantum strategy in a way spelt out in the previous paragraph, and this allows for pretty-good communication.

How well does this strategy do? We may observe that we obtain a 'good' encoding (ie. an output for the two senders that wins the CHSH game $\leftrightarrow$ an output produced by the PR-box) with probability ≈ 0.854; such an output is erased with probability

$p = \frac{1}{2}$. Otherwise, our strategy yields a 'bad' encoding (one that gets erased with probability 1 when sent over the channel) with probability ≈ 0.147. Table 4.2 below summarizes the channel outputs for each possible input encoding:

$X_1 \backslash X_2$	**0x**	**1x**
0y	**in: (00, 00), (01, 01) → out: 00/ee** in: (00, 01), (01, 00) → out: ee	**in:(00, 10), (01, 11) → out: 01/ee** in: (00, 11), (01, 10) → out: ee
1y	**in: (10, 00), (11, 01) → out: 10/ee** in: (10, 01), (11, 00) → out: ee	**in: (10, 10), (11, 11) → out: 11/ee** in: (10, 11), (11, 10) → out: ee

Table 4.2: Channel outputs using the quantum coding strategy. Within each box, the outputs in bold (which also allow for perfect decoding) are produced with probability $\cos^2(\frac{\pi}{8}) \approx 0.854$, and the non-bolded outputs, Pr ≈ 0.147.

To summarize, each sender gets his input bit erased with probability $\alpha = \sin^2(\frac{\pi}{8}) + \frac{1}{2}\cos^2(\frac{\pi}{8})$, and transmitted perfectly with probability $\frac{1}{2}\cos^2(\frac{\pi}{8})$. This amounts to each sender-receiver pair experiencing a binary erasure channel (Section 2.1.1 and Figure 2-3) with erasure probability α. Since the capacity of a binary erasure channel is $1 - \alpha$, the joint rate achieved by such a strategy is $2(1 - \alpha) = \cos^2(\frac{\pi}{8}) \approx 0.854$. □

4.1.3 Discussion

The intuition that quantum and super-quantum strategies perform better than classical ones on this channel is as follows: It is clear from looking at Table 4.3 that any classical choice of input alphabets for the two senders results in at least one combination of inputs that is erased with probability 1. Using a PR-box helps us avoid any input combinations that result in deterministic erasure, and using entanglement helps us avoid them with probability 0.854.

Although we have here spelt out two strategies for which the one that utilizes super-quantum resources does better than the one that utilizes shared entanglement, this is insufficient to prove Conjecture 1. However, we think it should not be hard to do so, especially in the limit where the erasure parameter p is close to 1.

As a validation of the proof (from the previous chapter) that quantum strategies fail to achieve a rate of 2 and an exercise, we could consider using the same CHSH game-assisted strategy on Channel I. It turns out that this induces a binary *symmetric*

channel (Section 2.1.1 and Figure 2-2) with bit-flip probability α for both pairs (the strategy ensures that the channel flips either no bits, or both message bits at once!). We may then calculate the joint rate of the channel on this strategy using the equation for the capacity of the BSC, 2.7: it is just

$$R_1 + R_2 = 2 \times \left[1 - H(\cos^2 \frac{\pi}{8})\right] \approx 0.798,$$

a far cry from the super-quantum rate of 2.

4.2 Multiple access channel preliminaries

We have been spending some time now on interference channels and their super-quantum/quantum enhancements. We now change tacks to explore another of the fundamental building blocks of Network Information Theory, multiple-access channels. Since our aim will eventually be to demonstrate that there exists a MAC whose capacity region is conjectured to be enhanced by super-quantum resources, we will take the same tack previously, and briefly introduce the theory of MACs before specifically describing our channel.

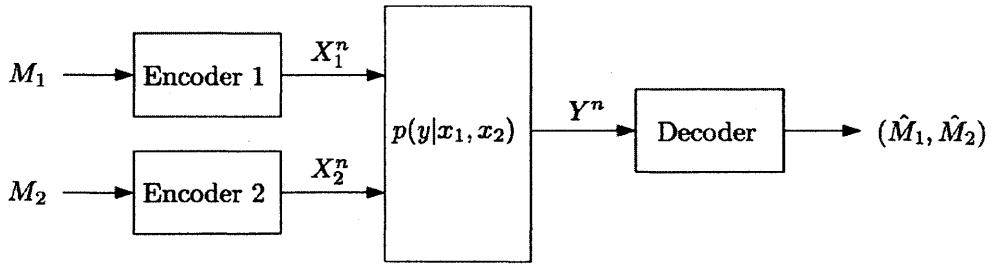

Figure 4-1: General model of a multiple-access communication system with independent messages. Figure taken from [4].

The distinguishing feature of the discrete memoryless multiple access channel (henceforth abbreviated to MAC, since we will only be concerned with discrete, memoryless channels) is that the senders wish to communicate an *independent* message reliably to a *common* receiver. That is, a MAC is denoted by $(\mathcal{X}_1 \times \mathcal{X}_2, p(y|x_1, x_2), \mathcal{Y})$.

A $(2^{nR_1}, 2^{nR_2}, n)$ code for the MAC is characterized by the same message sets and encoders as the interference channel but with a single channel output that goes to a decoder, visible in Figure 4-1. Therefore the decoding step is altered (from the interference channel) such that the receiver must guess both senders' messages:

- A decoder assigns an estimate $(\hat{m}_1, \hat{m}_2) \in [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$ or an error message to each received sequence y^n.

An achievable rate pair (R_1, R_2) is defined analogously, that is, there exists a sequence of $(2^{nR_1}, 2^{nR_2}, n)$ codes such that $\lim_{n\to\infty} [P_e^{(n)} \equiv P\{(\hat{M}_1, \hat{M}_2) \neq (M_1, M_2)\}] = 0$. Again, we will be concerned with the sum-capacity, $C_{\text{sum}} = \max\{R_1 + R_2 : (R_1, R_2)$ simultaneously achievable$\}$. Unlike the interference channel, a single-letter characterization exists of the MAC capacity region. Let $\mathscr{R}(X_1, X_2)$ be set of rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &\leq I(X_1 : Y|X_2) \\ R_2 &\leq I(X_2 : Y|X_1) \\ R_1 + R_2 &\leq I(X_1, X_2 : Y) \end{aligned} \tag{4.2}$$

In general, this is a pentagonal region with a 45° side, pictured in Figure 4-2 below: Then the capacity region $\mathscr{C}$ of the MAC $p(y|x_1, x_2)$ is the convex hull of the union of

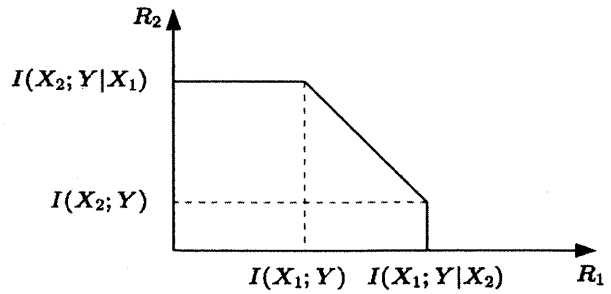

Figure 4-2: The region $\mathscr{R}(X_1, X_2)$ for a typical MAC. Figure taken from [4].

the regions $\mathscr{R}(X_1, X_2)$ over all $p(x_1)p(x_2)$.

4.3 Channel III: Super-quantum superiority on multiple-access channels

We now exhibit a channel on which super-quantum strategies give a conjectured advantage over quantum strategies. This is a modification to our existing channels, but now, let both output bits go to the same receiver, ie. we have a channel that accepts two 2-bit inputs and produces one 2-bit output.

$X_1 \backslash X_2$	**00**	**01**	**10**	**11**
00	00	10/01/11	01	00/10/11
01	01/10/11	00	00/10/11	01
10	10	00/01/11	00/01/10	11
11	00/01/11	10	11	00/01/10

Table 4.3: Channel III: a variation on Channel I and II in which the channel outputs not corresponding to the PR-box encodings are produced probabilistically (suppose, uniformly over the three options), instead of deterministically. Note that, in this case, both output bits go to the same receiver.

Lemma 9 (Capacity of Channel III with super-quantum assistance). *If the senders are allowed to share a PR-box, the sum-capacity of the given channel is exactly 2.*

Proof. The PR-box aided strategy that achieves a joint rate of 2 is exactly the same as the strategy for Channel I, and we only need to be careful in specifying the decoding process. This is as follows: each of the senders encodes 1-bit messages into 2 bits using the PR-box, and the receiver applies the following decoding rule to the channel outputs: $y_1 \to \hat{m_1}$, $y_2 \to \hat{m_2}$. □

Conjecture 2. *The classical sum-capacity (with no resources) of this channel is strictly less than 2, and therefore, the channel demonstrates the following separation:*

$$C_{classical} < C_{super\text{-}quantum}$$

4.3.1 Discussion

It might be worthwhile to examine exactly why, if we merely declare Channel I a multiple-access channel by interpreting both output bits as going to one receiver, it

fails to demonstrate the required separation. With a multiple-access channel, the receiver has the power of *global* knowledge of both channel outputs. Therefore, any multiple-access channel created in this way from an interference channel (IC) has *at least* the capacity of the corresponding interference channel; the receiver in this case merely needs to apply both of the decoding strategies that receivers 1 and 2 would have used separately for the corresponding IC. Therefore, whereas MACs are commonly considered more fundamental than ICs, it is paradoxically harder to come up with an example where super-quantum or quantum techniques yield higher communication rates.

Having understood this, it is easier to gain an intuition for why Channel III separates classical and super-quantum rates: we require a channel which disadvantages the receiver in a way that even global knowledge cannot repair. This handicap, in this case, is probabilistic output bits. Furthermore, since the benefit of using PR-box strategies of the form we have considered is that the senders are allowed to send coordinated input sets, we designed our channel such that output bits are probabilistically corrupted *unless* they correspond to coordinated input sets. Furthermore, the options for corrupted bits were chosen such as to maximally obfuscate any classical strategy the senders might agree on that uses a subset of 2 out of the 4 symbols in the alphabet. Again, we believe that the proof of this conjecture should be reasonably within reach.

4.4 Chapter summary

If our conjectures are proven true, we have exhibited a simple example of an interference channel on which the quantum-assisted capacity is strictly greater than the classical capacity, but is in turn superseded by the super-quantum-assisted capacity, as well as a MAC that has its capacity enhanced by PR boxes. Classically, network information theory is a still-growing field, and in the quantum domain, even less is known. There have been some nice recent results in characterizing the capacity regions of quantum extensions of multiple-access channels [29] and interference chan-

nels [30, 19, 20], but the latter derive single-letter capacity bounds only under certain assumptions such as very strong interference between the two sender-receiver pairs or requiring the receivers to decode both messages (simultaneous decoding strategy). What is truly remarkable is that we have arrived at these examples without resorting to any information-theoretic power-tools more complex than the concept of mutual information.

The contribution of this result is then two-fold. Firstly, Channel II and III are novel: II, if proven rigorously, is the only known example in literature of an interference channel that demonstrates the desired two-way separation; and III, if proven rigorously, will be the only known example of a MAC whose capacity is enhanced by a PR box. Secondly, games are not commonly discussed in relation to channels (a rare example is the paper [31], which uses graph coloring games to generate single-sender-receiver pair channels enhanced by entanglement assistance). The connection to multiplayer XOR games and non-locality has hitherto not been made for network channels. In this chapter we have considered only the very simplest of such games, the CHSH game, but we believe that further exploration of XOR games as well as quantum pseudo-telepathy games (see [26]) should yield a treasure trove of results. Ultimately, there could even exist a method to map *all* multiplayer games of a certain class to channel coding strategies that exhibit quantum or super-quantum superiority.

Chapter 5

Conclusion

It is now time to conclude our exploratory tour of Network Information Theory with our newly-acquired super-quantum and quantum information tools. We will use this last chapter to summarize our results and highlight the specific contribution of this thesis. We end off with some suggestions for future research.

5.1 Summary of original contributions

The present work demonstrates clearly that interference channels can have their classical capacity regions enhanced with super-quantum and (it is conjectured) quantum assistance, while multiple-access channels are conjectured to also allow for super-quantum enhancement. In illustrating this, we have laid the foundation stones of a bridge between super-quantum non-locality and multiple-sender communication channels. The interference channel in particular is notoriously hard to analyze even classically; the best known inner bound (the Han-Kobayashi bound [4]) on the general sum-capacity involves simultaneous satisfaction of *five* functions of various mutual informations! One benefit of these simple-to-analyze examples is therefore pedagogical; they build intuition for the types of capacity region enhancements possible with quantum- and super-quantum resources, as well as serve as a check for even better capacity characterizations in the future.

Specifically, we have exhibited a channel that show the following new separations

in classical capacity on the given classes of resources:

- Channel I: $\mathbf{C_{classical}}, \mathbf{C_{quantum}} < \mathbf{C_{super\text{-}quantum}}$
- Channel II: $\mathbf{C_{classical}}, \mathbf{C_{quantum}} < \mathbf{C_{super\text{-}quantum}}$ (conjectured)
- Channel III: $\mathbf{C_{classical}} < \mathbf{C_{super\text{-}quantum}}$ (conjectured)

where the label 'quantum' comes with the caveat that only $2 \times n$ entanglement between senders is considered. Our choice to limit our channel to handling only classical information (as opposed to density matrices representing quantum information) proved fruitful, as it paved the way for proofs that rely on classical information theory, as well as some results from pseudo-telepathy games where the referee, too, accepts only a discrete (albeit distributed) set of outcomes. This novel connection, originally drawn to prove the quantum-to-super-quantum separation in both Channels I and II, perhaps foreshadows greater intermixing between strategies for multiple-sender channels and multiplayer XOR or pseudo-telepathy games.

5.2 Future directions

An immediate goal is to prove our conjectures of the capacity separations in Channel IV, as well as to generalize Channels I, II and III to network channels with more than 2 sender-receiver pairs. In our consideration of quantum-assisted strategies, one could also formally extend the analysis to situations where the senders and receivers are allowed to share entanglement. We would also like to see a rigorous proof that these separations can be maintained even if the senders are provided a noisy PR-box and allowed multiple uses of it for non-locality distillation.

Slightly farther afield is the open question of whether the capacity region of broadcast channels (arguably a third pillar of Network Information Theory) can be enhanced with super-quantum and quantum resources. All of this, of course, paves the way for the million-dollar question of how to replicate the above separations on an *arbitrary* channel, or at the very least, characterize channels and coding strategies in a way that optimizes them for each of the three classes of resources.

Perhaps the most surprising aspect of this work is the connection between multi-sender channels and multi-player games. In hindsight, it seems natural to draw this connection given that pseudo-telepathy games exhibit the twin boons of being *known* to demonstrate super-quantum-to-quantum separations, and having had winning strategies (in a few cases) characterized and generalized to an arbitrarily large number of parties [26, 32]! It would be very satisfying if a general strategy could be found to map all pseudo-telepathy games to channels which demonstrate capacity separations. These facts practically necessitate a sequel to this work in the multi-sender ($n \geq 3$) case. This is almost certainly just the tip of the iceberg.

Bibliography

[1] J. Watrous, "On bell inequalities and nonlocality." CPSC519/619 Course notes, 2006.

[2] C. E. Shannon, "A mathematical theory of communication," tech. rep., ATT Bell Laboratories, 1948.

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* John Wiley and Sons, 2 ed., 2006.

[4] A. El Gamal and Y.-H. Kim, *Network information theory.* Cambridge University Press, 2012.

[5] J. Bell, "On the einstein podolsky rosen paradox," *Physics, 1*, pp. 195–200, 1964.

[6] D. Rohrlich and S. Popescu, "Quantum nonlocality as an axiom," *Foundations of Physics*, vol. 24, no. 3, 1994.

[7] S. Popescu, "Nonlocality beyond quantum mechanics," *Nature*, vol. 10, no. 4, pp. 264–270, 2014.

[8] A. Shimony, "Controllable and uncontrollable non-locality," *Proc. Int. Symp. Foundations of Quantum Mechanics (Kamefuchi, S. et al.(eds.)), Physical Society of Japan*, pp. 225–230, 1983.

[9] E. Hänggi, *Device-Independent Quantum Key Distribution.* PhD thesis, ETH Zurich, 1988.

[10] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, "Proposed experiment to test local hidden-variable theories," *Phys. Rev. Lett.*, vol. 23, p. 880, 13 Oct 1969.

[11] B. S. Cirelson, "Quantum generalizations of bell's inequality," *Letters in Mathematical Physics*, vol. 4, no. 2, pp. 93–100, March 1980.

[12] O. Regev and T. Vidick, "Quantum xor games," *ACM Transactions on Computation Theory (TOCT)*, vol. 7, Sept 2015.

[13] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information.* Cambridge University Press, 10th anniversary edition ed., March 2010.

[14] M. M. Wilde, "From classical to quantum information theory," 2011.

[15] J. Fan, M. D. Eisaman, and A. Migdall, "An optical fiber-based source of polarization-entangled photon pairs," *Proc. of SPIE*, vol. 6780, 2007.

[16] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, "Entanglement-assisted classical capacity of noisy quantum channels," *Phys. Rev. Lett.*, vol. 83, p. 3081, 1999.

[17] P. W. Shor, "The adaptive classical capacity of a quantum channel, or information capacities of three symmetric pure states in three dimensions," *IBM J. Research and Development*, vol. 48, Jan 2004.

[18] M. B. Hastings, "Superadditivity of communication capacity using entangled inputs," *Nature Phys.*, vol. 5, pp. 255–257, 2009.

[19] I. Savov, *Network information theory for classical-quantum channels.* PhD thesis, McGill University, School of Computer Science, 2012.

[20] O. Fawzi, P. Hayden, I. Savov, P. Sen, and M. M. Wilde, "Classical communication over a quantum interference channel," *IEEE Transactions on Information Theory*, vol. 58, no. 6, pp. 3670–3691, 2012.

[21] N. Cerf, N. Gisin, S. Massar, and S. Popescu, "Simulating maximal quantum entanglement without communication," *Physical Review Letters*, vol. 94, 2005.

[22] S. Winkler, J. Wullschleger, and S. Wolf, "Bit commitment from non-signaling correlations," *IEEE Transactions on Information Theory*, vol. 57, no. 3, pp. 1770–1779, 2011.

[23] H. Buhrman, M. Christandl, F. Unger, and S. Wehner, "Implications of superstrong nonlocality for cryptography," *Proc. Royal Soc. A.*, July 2006.

[24] W. van Dam, "Implausible consequences of superstrong nonlocality," *Journal of Natural Computing*, vol. 12, pp. 9–12, March 2013.

[25] G. Brassard, A. A. Methot, and A. Tapp, "Minimum entangled state dimension required for pseudo-telepathy," *Quantum Information & Computation*, vol. 5, pp. 275–284, July 2005.

[26] A. Broadbent and A. A. Méthot, "On the power of non-local boxes," *Journal Theoretical Computer Science archive Volume 358 Issue 1, 31 July 2006 Journal of Theoretical Computer Science*, vol. 358, July 2006.

[27] M. Forster, S. Winkler, and S. Wolf, "Distilling nonlocality," *Phys. Rev. Lett.*, Mar 2009.

[28] N. Brunner, N. Gisin, and V. Scarani, "Entanglement and non-locality are different resources," *New Journal of Physics*, 2005.

[29] M.-H. Hsieh, I. Devetak, and A. Winter, "Entanglement-assisted capacity of quantum multiple-access channels," *IEEE Transactions on Information Theory*, vol. 54, no. 7, 2008.

[30] I. Savov, O. Fawzi, M. M. Wilde, P. Sen, and P. Hayden, "Entanglement-assisted classical capacity of noisy quantum channels," *Proceedings of the 49th Annual Allerton Conference on Communication, Control, and Computing.*, pp. 609–616, 2011.

[31] L. Mančinska, G. Scarpa, and S. Severini, "New separations in zero-error channel capacity through projective kochen?specker sets and quantum coloring," *IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 59, NO. 6, JUNE 2013*, vol. 48, Jun 2013.

[32] A. Arkhipov, "Extending and characterizing quantum magic games," Master's thesis, MIT, 2012.

[33] B. Schumacher and M. D. Westmoreland, "Sending classical information via noisy quantum channels," *Physical Review A*, vol. 56, July 1997.

[34] H. Sato, "Two-user communication channels," *IEEE Transactions on Information Theory*, vol. 23, no. 3, pp. 295–304, 1977.

[35] P. Lamontagne, "Non-local boxes," 2014.