Broadcast Channel Coding: Algorithmic Aspects and Non-Signaling Assistance

Omar Fawzi* Paul Fermé†

Abstract

We address the problem of coding for classical broadcast channels, which entails maximizing the success probability that can be achieved by sending a fixed number of messages over a broadcast channel. For point-to-point channels, Barman and Fawzi found in [1] an $(1-e^{-1})$-approximation algorithm running in polynomial time, and showed that it is NP-hard to achieve a strictly better approximation ratio. Furthermore, these algorithmic results were at the core of their proof that non-signaling assistance does not change the capacity of point-to-point channels. It is natural to if similar results hold for broadcast channels, exploiting links between approximation algorithms of the channel coding problem and the non-signaling assisted capacity region.

In this work, we make several contributions on algorithmic aspects and non-signaling assisted capacity regions of broadcast channels. We first show that when non-signaling assistance is shared only between the decoders, the capacity region does not change. For the class of deterministic broadcast channels, we describe a $(1-e^{-1})^2$-approximation algorithm running in polynomial time, and we show that the capacity region for that class is the same with or without non-signaling assistance. Finally, we show that in the value query model, we cannot achieve a better approximation ratio than $\Omega\left(\frac{1}{\sqrt{m}}\right)$ in polynomial time for the general broadcast channel coding problem, with m the size of one of the outputs of the channel.

1 Introduction

Broadcast channels, introduced by Cover in [2], describe the simple network communication setting where one sender aims to transmit individual messages to two receivers. Contrary to one-way channels [3] or multiple-access channels [4, 5], the capacity region of broadcast channels is known only for particular classes such as the degraded [6, 7, 8], deterministic [9, 10] and semi-deterministic [11]. Only inner bounds [12, 13, 14] and outer bounds [15, 14, 16, 17] on the capacity region are known in the general setting.

On the one hand, from the point of view of quantum information, it is natural to ask whether additional resources, such as quantum entanglement or more generally non-signaling correlations between the parties, changes the capacity region. A non-signaling correlation is a multipartite input-output box shared between parties that, as the name suggests, cannot by itself be used to send information between parties. However, non-signaling correlations such as the ones generated by measurements of entangled quantum particles, can provide an advantage for various information processing tasks and nonlocal games. The study of such correlations has given rise to the quantum information area known as nonlocality [18]. For example, in the context of channel coding, there exists classical point-to-point channels for which quantum entanglement between the sender and the receiver can increase the optimal success probability for sending one bit of information with a single use of the channel [19, 1]. However,

*Univ Lyon, ENS Lyon, UCBL, CNRS, Inria, LIP, F-69342, Lyon Cedex 07, France. omar.fawzi@ens-lyon.fr
†Univ Lyon, ENS Lyon, UCBL, CNRS, Inria, LIP, F-69342, Lyon Cedex 07, France. paul.ferme@ens-lyon.fr

a well-known result [20] states that for classical point-to-point channels, entanglement and even more generally non-signaling correlations do not change the capacity of the channel; see also [21, 1].

In the network setting, behavior is different. Quek and Shor showed in [22] the existence of two-sender two-receiver interference channels with gaps between their classical, quantum-entanglement assisted and non-signaling assisted capacity regions. Following this result, Leditzky et al. [23, 24] showed that quantum entanglement shared between the two senders of a multiple access channel can strictly enlarge the capacity region. More specifically, a general investigation of non-signaling resources on multiple-access channel coding was done in [25], where it was notably proved that non-signaling advantage occurs even for a simple textbook multiple-access channel: the binary adder channel. A single-letter formula characterization of the quantum-entanglement assisted capacity region of multiple-access channels was later found in [26]. The influence of nonlocal resources on broadcast channel coding has been comparably less studied, the main known result being that quantum entanglement shared between the decoders does not change the capacity region [27].

On the other hand, from an algorithmic point of view, a crucial question in information theory is the complexity of the channel coding problem, which entails maximizing the success probability that can be achieved by sending a fixed number of messages over a channel. However, solving exactly this problem is out of reach, as it is NP-hard to find optimal codes. Therefore, the natural question that arises is the approximability of such a task. For point-to-point channels, Barman and Fawzi found in [1] an $(1-e^{-1})$-approximation algorithm running in polynomial time. They showed also that it is NP-hard to approximate the channel coding problem for any strictly better ratio. For ℓ-list-decoding, where the decoder is allowed to output a list of ℓ guesses, a polynomial-time approximation algorithm achieving a $1-\frac{\ell^\ell e^{-\ell}}{\ell!}$ ratio was found in [28], and it was shown to be NP-hard to do better in [29]. For multiple-access channel coding, the complexity of the problem can be linked to the bipartite densest κ-subgraph problem [30], which is conjectured to be NP-hard to approximate within any constant ratio [31]. However, so far, the approximability of channel coding has not been addressed for broadcast channels.

In the point-to-point scenario studied in [1], the existence of a constant-ratio approximation algorithm is linked to the equality of the capacity regions with and without non-signaling assistance. This is due to the fact the channel coding problem with non-signaling assistance becomes a linear program, thus computable in polynomial time. It is even equal to the natural linear relaxation of the channel coding problem, which is very common strategy towards approximating an integer linear program. Therefore, showing that this approximation strategy guarantees a constant ratio is the main ingredient towards showing the equality of the capacity regions with and without non-signaling assistance. This link has not been studied yet for broadcast channels. This raises the following questions: Does the capacity region of the broadcast channel change when non-signaling resources between parties are allowed? What is the best approximability ratio of the broadcast channel coding problem? How those two questions are related?

Our Results In the present paper, we first extend the result by [27] in a natural way. We prove that non-signaling resource shared between the decoders does not change the capacity region; see Theorem 3.3. More significantly, we study the influence of sharing a non-signaling resource between the three parties. Contrary to the previous case, this turns the broadcast channel coding problem, which is NP-hard, into a linear program, thus solvable in polynomial time.

We describe a $(1-e^{-1})^2$-approximation algorithm running in polynomial time for the broadcast channel coding problem limited to the class of deterministic channels. This is achieved through a graph interpretation of the problem, where one aims at partitioning a bipartite graph into k_1 and k_2 parts, such that the resulting quotient graph is the densest possible; see Proposition 4.3 and Theorem 4.4. Using the ideas coming from this algorithm, we show that for the class of deterministic channels, non-signaling resource shared between the three parties does not change the capacity region; see Theorem 4.6 and Corollary 4.7.

On the other direction, we consider the subproblem of broadcast channel coding where the number of messages one decoder is responsible of is maximum. This subproblem can be interpreted as a social welfare maximization problem. In the theory of fair division [32, 33], social welfare maximization entails partitioning a set of goods among agents in order to maximize the sum of their utilities. The social welfare problem has been extensively studied through black box approach [34], which led to a precise analysis of achievable approximation ratio as well as hardness results [35, 36], depending on the class of utility functions considered and the type of black box access to them. We refine the hardness result for the class of fractionally sub-additive utility functions to the subclass coming from the broadcast channel coding subproblem interpretation. Specifically, we show that in the value query model, we cannot achieve a better approximation ratio than $\Omega\left(\frac{1}{\sqrt{m}}\right)$ in polynomial time, with m the size of one of the outputs of the channel: see Theorem 5.4. Following the previous discussion on the links between constant ratio approximation algorithms and non-signaling capacity regions, this hardness result is a first step towards showing that sharing a non-signaling resource between the three parties of a broadcast channel can enlarge its capacity region.

Organization In Section 2, we define precisely the different versions of the broadcast channel coding problem depending on the choice of objective value, and show that they all lead to the same capacity region. In Section 3, we define the different non-signaling assisted versions of the broadcast channel coding problem. In particular, when shared only between the decoders, we show that the capacity region does not change. In Section 4, we address both algorithmic aspects and capacity considerations of deterministic broadcast channels. Specifically, we describe a $(1-e^{-1})^2$-approximation algorithm running in polynomial time for that class, and we show that the capacity region for that class is the same with or without non-signaling assistance. Finally, in Section 5, we show that in the value query model, we cannot achieve a better approximation ratio than $\Omega\left(\frac{1}{\sqrt{m}}\right)$ in polynomial time for the general broadcast channel coding problem, with m the size of one of the outputs of the channel.

2 Broadcast Channel Coding

2.1 Broadcast Channels

Formally, a broadcast channel is given by a conditional probability distribution on input $\mathcal{X}$ and two outputs $\mathcal{Y}_1$ and $\mathcal{Y}_2$, so $W := (W(y_1, y_2|x))_{y_1 \in \mathcal{Y}_1, y_2 \in \mathcal{Y}_2, x \in \mathcal{X}}$, where $W(y_1 y_2|x) \geq 0, \sum_{y_1 \in \mathcal{Y}_1, y_2 \in \mathcal{Y}_2} W(y_1 y_2|x) = 1$. We define its marginals W_1 and W_2 by $W_1(y_1|x) := \sum_{y_2 \in \mathcal{Y}_2} W(y_1 y_2|x)$ and $W_2(y_2|x) := \sum_{y_1 \in \mathcal{Y}_1} W(y_1 y_2|x)$. We will denote such a broadcast channel by $W : \mathcal{X} \to \mathcal{Y}_1 \times \mathcal{Y}_2$. The tensor product of two broadcast channels $W : \mathcal{X} \to \mathcal{Y}_1 \times \mathcal{Y}_2$ and $W' : \mathcal{X}' \to \mathcal{Y}'_1 \times \mathcal{Y}'_2$ is denoted by $W \otimes W' : \mathcal{X} \times \mathcal{X}' \to (\mathcal{Y}_1 \times \mathcal{Y}'_1) \times (\mathcal{Y}_2 \times \mathcal{Y}'_2)$ and defined by $(W \otimes W')(y_1 y'_1 y_2 y'_2|xx') := W(y_1 y_2|x) \cdot W'(y'_1 y'_2|x')$. We define $W^{\otimes n}(y_1^n y_2^n|x^n) := \prod_{i=1}^n W(y_{1,i} y_{2,i}|x_i)$, for $y_1^n := y_{1,1} \dots y_{1,n} \in \mathcal{Y}_1^n$ and $y_2^n := y_{2,1} \dots y_{2,n} \in \mathcal{Y}_2^n$ and $x^n := x_1 \dots x_n \in \mathcal{X}^n$. We will use the notation $[k] := \{1, \dots, k\}$.

The coding problem for a broadcast channel $W : \mathcal{X} \to \mathcal{Y}_1 \times \mathcal{Y}_2$ is the following: one wants to encode a couple of messages belonging to $[k_1] \times [k_2]$ into $\mathcal{X}$, which will be given as input to the channel W. This results in two random outputs in $\mathcal{Y}_1$ and $\mathcal{Y}_2$, which one needs to decode back into the corresponding messages in $[k_1]$ and $[k_2]$. We will call $e : [k_1] \times [k_2] \to \mathcal{X}$ the encoder, $d_1 : \mathcal{Y}_1 \to [k_1]$ the first decoder and $d_2 : \mathcal{Y}_2 \to [k_2]$ the second decoder. This is depicted in Figure 2.1.

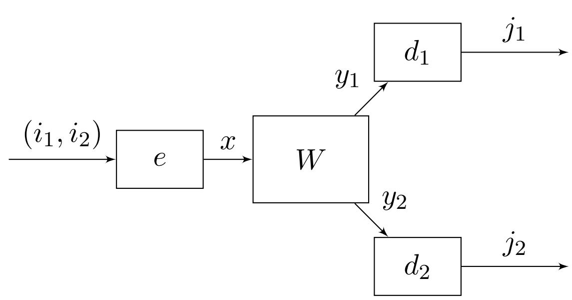

Figure 1: Coding for a broadcast channel W.

We will call $\mathrm{p}_1(W, e, d_1)$ (resp. $\mathrm{p}_2(W, e, d_2)$) the probability of successfully decoding the first (resp. second) message, ie. that $j_1 = i_1$ (resp. $j_2 = i_2$), given that the encoder is e and the decoder is d_1 (resp. d_2). We will also consider $\mathrm{p}(W, e, d_1, d_2)$, the probability of successfully decoding both messages, ie. that $j_1 = i_1$ and $j_2 = i_2$, given that the encoder is e and the decoders are d_1, d_2.

We aim to find the best encoder and decoders according to some figure of merit. However, to do so, we need a one dimensional real value objective to optimize. This leads to two different quantities.

2.2 The Sum Success Probability $\mathrm{S}_{\mathrm{sum}}(W, k_1, k_2)$

We will focus first on maximizing $\frac{\mathrm{p}_1(W,e,d_1)+\mathrm{p}_2(W,e,d_2)}{2}$ over all encoders e and decoders d_1, d_2. We will call $\mathrm{S}_{\mathrm{sum}}(W, k_1, k_2)$ the resulting maximum sum probability of successfully encoding and decoding the messages through W, given that the input message couple is taken uniformly in $[k_1] \times [k_2]$. $\mathrm{S}_{\mathrm{sum}}(W, k_1, k_2)$ is the solution of the following optimization program:

$$
\begin{aligned}
\mathrm{S}_{\mathrm{sum}}(W, k_1, k_2) := \quad & \underset{e, d_1, d_2}{\text{maximize}} && \frac{1}{k_1 k_2} \sum_{i_1, i_2, x, y_1, y_2} W(y_1 y_2|x) e(x|i_1 i_2) \frac{d_1(i_1|y_1) + d_2(i_2|y_2)}{2} \\
& \text{subject to} && \sum_{x \in \mathcal{X}} e(x|i_1 i_2) = 1, \forall i_1 \in [k_1], i_2 \in [k_2] \\
& && \sum_{i_1 \in [k_1]} d_1(y_1|i_1) = 1, \forall y_1 \in \mathcal{Y}_1 \\
& && \sum_{i_2 \in [k_2]} d_2(y_2|i_2) = 1, \forall y_2 \in \mathcal{Y}_2 \\
& && e(x|i_1 i_2), d_1(y_1|i_1), d_2(y_2|i_2) \geq 0
\end{aligned} \tag{1}
$$

Proof. One should note that we allow in fact non-deterministic encoders and decoders for generality reasons, although the value of the program is not changed as it is convex. Besides that remark, let us name $I_1, I_2, J_1, J_2, X, Y_1, Y_2$ the random variables corresponding to $i_1, i_2, j_1, j_2, x, y_1, y_2$ in the coding and decoding process. Then, given e, d_1, d_2 and W, the objective value of the previous program comes from:

$$
\begin{aligned}
\mathrm{p}_1(W, e, d_1) = \mathbb{P}(J_1 = I_1) &= \frac{1}{k_1 k_2} \sum_{i_1, i_2} \mathbb{P}(J_1 = i_1 | I_1 = i_1, I_2 = i_2) \\
&= \frac{1}{k_1 k_2} \sum_{i_1, i_2, x} e(x|i_1 i_2) \mathbb{P}(J_1 = i_1 | I_1 = i_1, I_2 = i_2, X = x) \\
&= \frac{1}{k_1 k_2} \sum_{i_1, i_2, x, y_1, y_2} W(y_1 y_2|x) e(x|i_1 i_2) \mathbb{P}(J_1 = i_1 | I_1 = i_1, I_2 = i_2, X = x, Y_1 = y_1, Y_2 = y_2) \\
&= \frac{1}{k_1 k_2} \sum_{i_1, i_2, x, y_1, y_2} W(y_1 y_2|x) e(x|i_1 i_2) d_1(i_1|y_1) \ ,
\end{aligned} \tag{2}
$$

and symmetrically for $\mathrm{p}_2(W, e, d_2)$, which leads to the announced objective value. □

One can rewrite this optimization program in a more convenient way, proving that $\mathrm{S}_{\mathrm{sum}}(W, k_1, k_2)$ depends only on the marginals of W:

Proposition 2.1.

$$
\begin{aligned}
\mathrm{S}_{\mathrm{sum}}(W,k_1,k_2) = \quad \underset{e,d_1,d_2}{\text{maximize}} \quad & \frac{1}{2k_1k_2}\sum_{i_1,x,y_1} W_1(y_1|x)d_1(i_1|y_1)\sum_{i_2} e(x|i_1i_2) \\
+ \quad & \frac{1}{2k_1k_2}\sum_{i_2,x,y_2} W_2(y_2|x)d_2(i_2|y_2)\sum_{i_1} e(x|i_1i_2) \\
\text{subject to} \quad & \sum_{x\in\mathcal{X}} e(x|i_1i_2) = 1, \forall i_1 \in [k_1], i_2 \in [k_2] \\
& \sum_{i\in[k_1]} d_1(y_1|i_1) = 1, \forall y_1 \in \mathcal{Y}_1 \\
& \sum_{i\in[k_2]} d_2(y_2|i_2) = 1, \forall y_2 \in \mathcal{Y}_2 \\
& e(x|i_1i_2), d_1(y_1|i_1), d_2(y_2|i_2) \geq 0
\end{aligned}
\tag{3}
$$

Proof. It follows directly from the definitions $W_1(y_1|x) := \sum_{y_2} W(y_1y_2|x)$ and $W_2(y_2|x) := \sum_{y_1} W(y_1y_2|x)$. □

Since broadcast channels are more general than one-way channels (by defining $W_1(y_1|x) := \hat{W}(y_1|x)$ for $\hat{W}$ a one-way channel and taking $W_2(y_2|x) = \frac{1}{|\mathcal{Y}_2|}$ a completely trivial channel), computing a single value $\mathrm{S}_{\mathrm{sum}}(W,k_1,k_2)$ is NP-hard, and it is even NP-hard to approximate $\mathrm{S}_{\mathrm{sum}}(W,k_1,k_2)$ within a better factor than $\left(1-e^{-1}\right)$, as a consequence of the hardness result on $\mathrm{S}(W,k)$ proved in [1].

2.3 The Joint Success Probability $\mathrm{S}(W,k_1,k_2)$

We will now focus on maximizing $\mathrm{p}(W,e,d_1,d_2)$ over all encoders e and decoders d_1,d_2. We will call $\mathrm{S}(W,k_1,k_2)$ the resulting maximum probability of successfully encoding and decoding the messages through W, given that the input message couple is taken uniformly in $[k_1]\times[k_2]$. $\mathrm{S}(W,k_1,k_2)$ is the solution of the following optimization program:

$$
\begin{aligned}
\mathrm{S}(W,k_1,k_2) := \quad \underset{e,d_1,d_2}{\text{maximize}} \quad & \frac{1}{k_1k_2}\sum_{i_1,i_2,x,y_1,y_2} W(y_1y_2|x)e(x|i_1i_2)d_1(i_1|y_1)d_2(i_2|y_2) \\
\text{subject to} \quad & \sum_{x\in\mathcal{X}} e(x|i_1i_2) = 1, \forall i_1 \in [k_1], i_2 \in [k_2] \\
& \sum_{i_1\in[k_1]} d_1(y_1|i_1) = 1, \forall y_1 \in \mathcal{Y}_1 \\
& \sum_{i_2\in[k_2]} d_2(y_2|i_2) = 1, \forall y_2 \in \mathcal{Y}_2 \\
& e(x|i_1i_2), d_1(y_1|i_1), d_2(y_2|i_2) \geq 0
\end{aligned}
\tag{4}
$$

The proof is the same as in the sum probability scenario. The objective values of those two optimization programs are not the same, but $\mathrm{S}(W,k_1,k_2)$ and $\mathrm{S}_{\mathrm{sum}}(W,k_1,k_2)$ still characterize the same capacity region [37]. Let us recall first the definition of their capacity regions:

Definition 2.2 (Capacity Region $\mathcal{C}(W)$ (resp. $\mathcal{C}_{\mathrm{sum}}(W)$) of a broadcast channel W)**.** A rate pair (R_1,R_2) is achievable (resp. sum-achievable) if:

$$\lim_{n\to+\infty} \mathrm{S}(W^{\otimes n}, \lceil 2^{R_1n}\rceil, \lceil 2^{R_2n}\rceil) = 1 \ .$$

$$(\text{resp. } \lim_{n\to+\infty} \mathrm{S}_{\mathrm{sum}}(W^{\otimes n}, \lceil 2^{R_1n}\rceil, \lceil 2^{R_2n}\rceil) = 1) \ .$$

We define the joint (resp. sum) (classical) capacity region $\mathcal{C}(W)$ (resp. $\mathcal{C}_{\mathrm{sum}}(W)$) as the closure of the set of all achievable (resp. sum-achievable) rate pairs.

Proposition 2.3. For any broadcast channel W, $\mathcal{C}(W) = \mathcal{C}_{\text{sum}}(W)$.

Proof. Let us focus on error probabilities rather than success ones. Call them respectively $\mathrm{E}(W,k_1,k_2) := 1 - \mathrm{S}(W,k_1,k_2)$ and $\mathrm{E}_{\text{sum}}(W,k_1,k_2) := 1 - \mathrm{S}_{\text{sum}}(W,k_1,k_2)$. Let us fix a solution e, d_1, d_2 of the optimization program computing $\mathrm{S}(W,k_1,k_2)$. Let us remark first that:

$$\sum_{i_1,i_2,x,y_1,y_2} W(y_1y_2|x)e(x|i_1i_2) = k_1k_2 \ ,$$

thus, the value of the maximum error for those encoder and decoders is:

$$\begin{aligned}
\mathrm{E}(W,k_1,k_2,e,d_1,d_2) &:= 1 - \frac{1}{k_1k_2}\left(\sum_{i_1,i_2,x,y_1,y_2} W(y_1y_2|x)e(x|i_1i_2)d_1(i_1|y_1)d_2(i_2|y_2)\right) \\
&= \frac{1}{k_1k_2}\left(\sum_{i_1,i_2,x,y_1,y_2} W(y_1y_2|x)e(x|i_1i_2) - \sum_{i_1,i_2,x,y_1,y_2} W(y_1y_2|x)e(x|i_1i_2)d_1(i_1|y_1)d_2(i_2|y_2)\right) \\
&= \frac{1}{k_1k_2}\left(\sum_{i_1,i_2,x,y_1,y_2} W(y_1y_2|x)e(x|i_1i_2)\left[1 - d_1(i_1|y_1)d_2(i_2|y_2)\right]\right) \ .
\end{aligned} \tag{5}$$

Similarly, the value of the sum error for those encoder and decoders is:

$$\begin{aligned}
\mathrm{E}_{\text{sum}}(W,k_1,k_2,e,d_1,d_2) &:= 1 - \frac{1}{k_1k_2}\left(\sum_{i_1,i_2,x,y_1,y_2} W(y_1y_2|x)e(x|i_1i_2)\frac{d_1(i_1|y_1)+d_2(i_2|y_2)}{2}\right) \\
&= \frac{1}{k_1k_2}\left(\sum_{i_1,i_2,x,y_1,y_2} W(y_1y_2)e(x|i_1i_2)\left[1 - \frac{d_1(i_1|y_1)+d_2(i_2|y_2)}{2}\right]\right) \ .
\end{aligned} \tag{6}$$

However, for $x,y \in [0,1]$, we have that:

$$1 - xy \geq \max(1-x, 1-y) \geq 1 - \frac{x+y}{2} \ ,$$

and:

$$1 - xy \leq (1-x) + (1-y) = 2\left(1 - \frac{x+y}{2}\right) \ .$$

This means that, for all e, d_1, d_2, we have:

$$\mathrm{E}_{\text{sum}}(W,k_1,k_2,e,d_1,d_2) \leq \mathrm{E}(W,k_1,k_2,e,d_1,d_2) \leq 2\mathrm{E}_{\text{sum}}(W,k_1,k_2,e,d_1,d_2) \ ,$$

so, maximizing over all e, d_1, d_2, we get:

$$\mathrm{E}_{\text{sum}}(W,k_1,k_2) \leq \mathrm{E}(W,k_1,k_2) \leq 2\mathrm{E}_{\text{sum}}(W,k_1,k_2) \ .$$

Thus, up to a multiplicative factor 2, the error is the same. In particular, when one of those errors tends to zero, the other one tends to zero as well. This implies that the capacity regions are the same. □

3 Non-Signaling Assistance

In this section, we will consider the broadcast channel coding problem with additional resources, in order to determine how these resources affect its success probabilities as well as the capacity regions that can be defined from them.

3.1 Non-Signaling Assistance Between the Decoders

Here, we consider the case where the receivers are given non-signaling assistance. This resource, which is a theoretical but easier to manipulate generalization of quantum entanglement, can be characterized as follows. A non-signaling box $d(j_1j_2|y_1y_2)$ is any joint conditional probability distribution such that the marginal from one party is independent from the other party's input, ie. we have:

$$\begin{aligned}\forall j_1, y_1, y_2, y_2', \quad \sum_{j_2} d(j_1j_2|y_1y_2) = \sum_{j_1} d(j_1j_2|y_1y_2') \ ,\\ \forall j_2, y_1, y_2, y_1', \quad \sum_{j_1} d(j_1j_2|y_1y_2) = \sum_{j_1} d(j_1j_2|y_1'y_2) \ .\end{aligned} \tag{7}$$

Thus, when receivers are given non-signaling assistance, it means that the product $d_1(j_1|y_1)d_2(j_2|y_2)$ is replaced by the non-signaling box $d(j_1j_2|y_1y_2)$. We define the joint and sum success probabilities $\mathrm{S}^{\mathrm{NS}_{\mathrm{decs}}}(W, k_1, k_2)$ and $\mathrm{S}^{\mathrm{NS}_{\mathrm{decs}}}_{\mathrm{sum}}(W, k_1, k_2)$ according to this:

$$\begin{aligned}
\mathrm{S}^{\mathrm{NS}_{\mathrm{decs}}}(W, k_1, k_2) := \quad & \underset{e,d_1,d_2}{\text{maximize}} \quad \frac{1}{k_1k_2} \sum_{i_1,i_2,x,y_1,y_2} W(y_1y_2|x)e(x|i_1i_2)d(i_1i_2|y_1y_2)\\
\Big(\text{resp. } \mathrm{S}^{\mathrm{NS}_{\mathrm{decs}}}_{\mathrm{sum}}(W, k_1, k_2) := \quad & \underset{e,d_1,d_2}{\text{maximize}} \quad \frac{1}{k_1k_2} \sum_{i_1,i_2,x,y_1,y_2} W(y_1y_2|x)e(x|i_1i_2)\frac{\sum_{j_2} d(i_1j_2|y_1y_2) + \sum_{j_1} d(j_1i_2|y_1y_2)}{2}\Big)\\
& \text{subject to} \quad \sum_x e(x|i_1i_2) = 1\\
& \sum_{j_2} d(j_1j_2|y_1y_2) = \sum_{j_1} d(j_1j_2|y_1y_2')\\
& \sum_{j_1} d(j_1j_2|y_1y_2) = \sum_{j_1} d(j_1j_2|y_1'y_2)\\
& \sum_{j_1,j_2} d(j_1j_2|y_1y_2) = 1\\
& e(x|i_1i_2), d(j_1j_2|y_1y_2) \geq 0
\end{aligned} \tag{8}$$

The corresponding capacity regions $\mathcal{C}^{\mathrm{NS}_{\mathrm{decs}}}(W)$ and $\mathcal{C}^{\mathrm{NS}_{\mathrm{decs}}}_{\mathrm{sum}}(W)$ are defined as before:

Definition 3.1 (Capacity Region $\mathcal{C}^{\mathrm{NS}_{\mathrm{decs}}}(W)$ (resp. $\mathcal{C}^{\mathrm{NS}_{\mathrm{decs}}}_{\mathrm{sum}}(W)$) of a broadcast channel W)**.** A rate pair (R_1, R_2) is achievable (resp. sum-achievable) with non-signaling assistance between the decoders if:

$$\lim_{n\to+\infty} \mathrm{S}^{\mathrm{NS}_{\mathrm{decs}}}(W^{\otimes n}, \lceil 2^{R_1n}\rceil, \lceil 2^{R_2n}\rceil) = 1 \ .$$

$$(\text{resp. } \lim_{n\to+\infty} \mathrm{S}^{\mathrm{NS}_{\mathrm{decs}}}_{\mathrm{sum}}(W^{\otimes n}, \lceil 2^{R_1n}\rceil, \lceil 2^{R_2n}\rceil) = 1) \ .$$

We define the joint (resp. sum) capacity region with non-signaling assistance between the decoders $\mathcal{C}^{\mathrm{NS}_{\mathrm{decs}}}(W)$ (resp. $\mathcal{C}^{\mathrm{NS}_{\mathrm{decs}}}_{\mathrm{sum}}(W)$) as the closure of the set of all rate pairs achievable (resp. sum-achievable) with non-signaling assistance between the decoders.

The objective of this section is to show that sharing non-signaling assistance between the decoders does not change the capacity regions of a broadcast channel. In order to do so, we will also show that sum and joint capacity regions with non-signaling assistance between the decoders are the same. This is an extension of the result by [27] where it has been proved to be the case for the weaker resource that is quantum entanglement.

Proposition 3.2. For any broadcast channel W, $\mathcal{C}^{\mathrm{NS}_{\mathrm{decs}}}_{\mathrm{sum}}(W) = \mathcal{C}^{\mathrm{NS}_{\mathrm{decs}}}(W)$

Proof. Given an encoder e and a non-signaling decoding box d, the maximum success probability of encoding and decoding correctly with those is given by:

$$\mathrm{S}^{\mathrm{NS}_{\mathrm{decs}}}(W, k_1, k_2, e, d) := \frac{1}{k_1k_2} \sum_{i_1,i_2,x,y_1,y_2} W(y_1y_2|x)e(x|i_1i_2)d(i_1i_2|y_1y_2) \ .$$

This should be compared to the sum success probability of encoding and decoding correctly with those:

$$\mathrm{S}_{\mathrm{sum}}^{\mathrm{NS_{decs}}}(W,k_1,k_2,e,d) := \frac{1}{k_1k_2}\sum_{i_1,i_2,x,y_1,y_2} W(y_1y_2|x)e(x|i_1i_2)\left[\frac{\sum_{j_2} d(i_1j_2|y_1y_2)+\sum_{j_1} d(j_1i_2|y_1y_2)}{2}\right] .$$

Similarly to what was done in Proposition 2.3, we focus on the error probabilities rather than success probabilities. This leads again to:

$$\mathrm{E}^{\mathrm{NS_{decs}}}(W,k_1,k_2,e,d) = \frac{1}{k_1k_2}\sum_{i_1,i_2,x,y_1,y_2} W(y_1y_2|x)e(x|i_1i_2)\left[1-d(i_1i_2|y_1y_2)\right] ,$$

and:

$$\mathrm{E}_{\mathrm{sum}}^{\mathrm{NS_{decs}}}(W,k_1,k_2,e,d) = \frac{1}{k_1k_2}\sum_{i_1,i_2,x,y_1,y_2} W(y_1y_2|x)e(x|i_1i_2)\left[\frac{1-\sum_{j_2} d(i_1j_2|y_1y_2)}{2}+\frac{1-\sum_{j_1} d(j_1i_2|y_1y_2)}{2}\right] .$$

But we have that:

$$1-d(i_1i_2|y_1y_2) \geq \max\left(1-\sum_{j_2} d(i_1j_2|y_1y_2), 1-\sum_{j_1} d(j_1i_2|y_1y_2)\right) \geq \frac{1-\sum_{j_2} d(i_1j_2|y_1y_2)}{2}+\frac{1-\sum_{j_1} d(j_1i_2|y_1y_2)}{2} ,$$

since $d(j_1j_2|y_1y_2) \in [0,1]$, and we have that:

$$1-\sum_{j_2} d(i_1j_2|y_1y_2)+1-\sum_{j_1} d(j_1i_2|y_1y_2) = 1-d(i_1i_2|y_1y_2)+1-\sum_{(j_1,j_2)\in S} d(j_1j_2|y_1y_2) \geq 1-d(i_1i_2|y_1y_2) ,$$

with $S := \{(i_1,j_2) : j_2 \in [k_2]-\{i_2\}\} \sqcup \{(j_1,i_2) : j_1 \in [k_1]-\{i_1\}\}$.

Thus, this implies that:

$$\mathrm{E}_{\mathrm{sum}}^{\mathrm{NS_{decs}}}(W,k_1,k_2,e,d) \leq \mathrm{E}^{\mathrm{NS_{decs}}}(W,k_1,k_2,e,d) \leq 2\mathrm{E}_{\mathrm{sum}}^{\mathrm{NS_{decs}}}(W,k_1,k_2,e,d) ,$$

and by maximizing over all e and d:

$$\mathrm{E}_{\mathrm{sum}}^{\mathrm{NS_{decs}}}(W,k_1,k_2) \leq \mathrm{E}^{\mathrm{NS_{decs}}}(W,k_1,k_2) \leq 2\mathrm{E}_{\mathrm{sum}}^{\mathrm{NS_{decs}}}(W,k_1,k_2) .$$

As before, this implies that the capacity regions are the same. □

Theorem 3.3. For any broadcast channel W, $\mathcal{C}(W) = \mathcal{C}^{\mathrm{NS_{decs}}}(W)$

Proof. In the sum scenario, since the objective function does not depend on the product $d_1(j_1|y_1)d_2(j_2|y_2)$ but only on the marginals $d_1(j_1|y_1)$ and $d_2(j_2|y_2)$, the non-signaling box won't give additional decoding power. Indeed, for any encoder e and non-signaling decoding box d, we have that:

$$\begin{aligned}\mathrm{S}_{\mathrm{sum}}^{\mathrm{NS_{decs}}}(W,k_1,k_2,e,d) := &\frac{1}{2k_1k_2}\sum_{i_1,x,y_1} W_1(y_1|x)\left(\sum_{j_2} d(i_1j_2|y_1y_2)\right)\sum_{i_2} e(x|i_1i_2)\\ &+\frac{1}{2k_1k_2}\sum_{i_2,x,y_2} W_2(y_2|x)\left(\sum_{j_1} d(j_1i_2|y_1y_2)\right)\sum_{i_1} e(x|i_1i_2) .\end{aligned} \tag{9}$$

Thus, by choosing $d_1(j_1|y_1) := \sum_{j_2} d(j_1j_2|y_1y_2)$ and $d_2(j_2|y_2) := \sum_{j_1} d(j_1j_2|y_1y_2)$, which are well-defined since d is a non-signaling box, one gets that $\mathrm{S}_{\mathrm{sum}}(W,k_1,k_2,e,d_1,d_2) = \mathrm{S}_{\mathrm{sum}}^{\mathrm{NS_{decs}}}(W,k_1,k_2,e,d)$. By optimizing over all e and d, one gets $\mathrm{S}_{\mathrm{sum}}^{\mathrm{NS_{decs}}}(W,k_1,k_2) \leq \mathrm{S}_{\mathrm{sum}}(W,k_1,k_2)$. Since the inequality is obvious in the other direction, as $d(j_1j_2|y_1y_2) := d_1(j_1|y_1)d_2(j_2|y_2)$ is always a non-signaling box, we have that $\mathrm{S}_{\mathrm{sum}}(W,k_1,k_2) = \mathrm{S}_{\mathrm{sum}}^{\mathrm{NS_{decs}}}(W,k_1,k_2)$. This implies in particular that the capacity regions are the same, ie. $\mathcal{C}_{\mathrm{sum}}(W) = \mathcal{C}_{\mathrm{sum}}^{\mathrm{NS_{decs}}}(W)$

Finally, since $\mathcal{C}(W) = \mathcal{C}_{\mathrm{sum}}(W)$ by Proposition 2.3 and $\mathcal{C}_{\mathrm{sum}}^{\mathrm{NS_{decs}}}(W) = \mathcal{C}^{\mathrm{NS_{decs}}}(W)$ by Proposition 3.2, we get that $\mathcal{C}(W) = \mathcal{C}^{\mathrm{NS_{decs}}}(W)$. □

3.2 Full Non-Signaling Assistance

In this section, we will consider the case where the sender and the receivers are given non-signaling assistance. This means that a three-party non-signaling box $P(xj_1j_2|(i_1i_2)y_1y_2)$ will replace the product $e(x|i_1i_2)d_1(j_1|y_1)d_2(j_2|y_2)$ in the previous objective values. A joint conditional probability $P(xj_1j_2|(i_1i_2)y_1y_2)$ is a non-signaling box if the marginal from any two parties is independent from the removed party's input:

$$\begin{aligned}\forall j_1, j_2, i_1, i_2, y_1, y_2, i_1', i_2', \quad & \sum_x P(xj_1j_2|(i_1i_2)y_1y_2) = \sum_x P(xj_1j_2|(i_1'i_2')y_1y_2) \ , \\ \forall x, j_2, i_1, i_2, y_1, y_2, y_1', \quad & \sum_{j_1} P(xj_1j_2|(i_1i_2)y_1y_2) = \sum_{j_1} P(xj_1j_2|(i_1i_2)y_1'y_2) \ , \\ \forall x, j_1, i_1, i_2, y_1, y_2, y_2', \quad & \sum_{j_2} P(xj_1j_2|(i_1i_2)y_1y_2) = \sum_{j_2} P(xj_1j_2|(i_1i_2)y_1y_2') \ .\end{aligned} \tag{10}$$

The scenario is depicted in Figure 2.

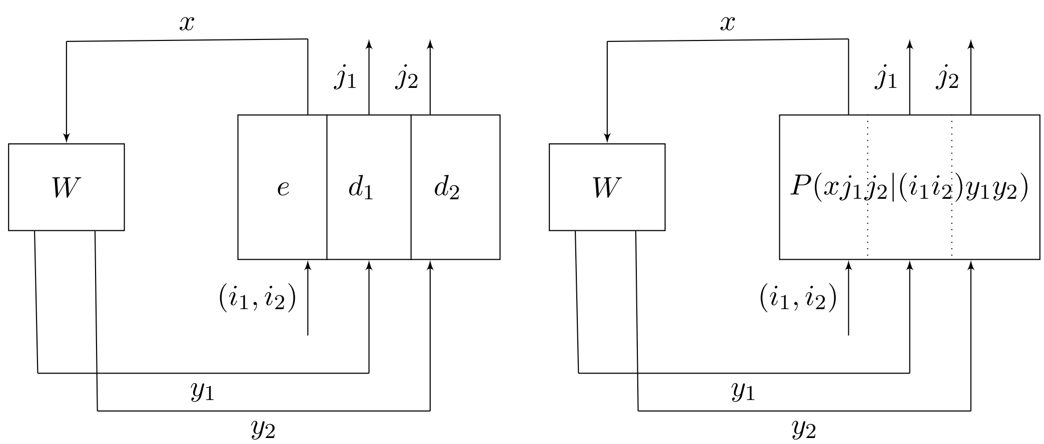

Figure 2: A non-signaling box P replacing e, d_1, d_2 in the coding problem for the broadcast channel W.

The cyclicity of Figure 2 is at first sight counter-intuitive. Note first that P being a non-signaling box is completely independent from W: in particular, the variables y_1, y_2 do not need to follow any laws in the definition of P being a non-signaling box. Therefore, the remaining ambiguity is the apparent need to encode and decode at the same time. However, since P is a non-signaling box, we won't need to do both at the same time, although the global correlation between the sender and the receivers will be characterized only by $P(xj_1j_2|(i_1i_2)y_1y_2)$. Indeed, $\forall y_1, y_2, P(x|(i_1i_2)) = P(x|(i_1i_2)y_1y_2)$ by the non-signaling property of P. Thus, one can get the output x on input (i_1i_2) without access to y_1, y_2, as that knowledge won't affect the law of x. Then (y_1, y_2) follows the law given by W given that x. Finally, given access to y_1, y_2, the decoding process is described by:

$$P(j_1j_2|(i_1i_2)y_1y_2x) = \frac{P(xj_1j_2|(i_1i_2)y_1y_2)}{P(x|(i_1i_2)y_1y_2)} = \frac{P(xj_1j_2|(i_1i_2)y_1y_2)}{P(x|(i_1i_2))} \ ,$$

so we recover $P(j_1j_2|(i_1i_2)y_1y_2x)) \times P(x|(i_1i_2)) = P(xj_1j_2|(i_1i_2)y_1y_2)$, and therefore, the process is not cyclic. Non-signaling boxes define exactly the conditional probability distributions where it is possible to consider the conditional probabilities of each party independently. This clarifies how one can effectively encode and then decode messages through a non-signaling box.

We will call the maximum sum success probability $\mathrm{S}^{\mathrm{NS}}_{\mathrm{sum}}(W, k_1, k_2)$, which is given by the following linear program, where the constraints translate precisely the fact that P is a non-signaling box:

$$
\begin{aligned}
\mathrm{S}^{\mathrm{NS}}_{\mathrm{sum}}(W,k_1,k_2) := \quad & \underset{P}{\text{maximize}} && \frac{1}{2k_1k_2}\sum_{i_1,x,y_1} W_1(y_1|x)\sum_{i_2,j_2} P(xi_1j_2|(i_1i_2)y_1y_2)\\
& \qquad + && \frac{1}{2k_1k_2}\sum_{i_2,x,y_2} W_2(y_2|x)\sum_{i_1,j_1} P(xj_1i_2|(i_1i_2)y_1y_2)\\
& \text{subject to} && \sum_{x} P(xj_1j_2|(i_1i_2)y_1y_2) = \sum_{x} P(xj_1j_2|(i_1'i_2')y_1y_2)\\
& && \sum_{j_1} P(xj_1j_2|(i_1i_2)y_1y_2) = \sum_{j_1} P(xj_1j_2|(i_1i_2)y_1'y_2)\\
& && \sum_{j_2} P(xj_1j_2|(i_1i_2)y_1y_2) = \sum_{j_2} P(xj_1j_2|(i_1i_2)y_1y_2')\\
& && \sum_{x,j_1,j_2} P(xj_1j_2|(i_1i_2)y_1y_2) = 1\\
& && P(xj_1j_2|(i_1i_2)y_1y_2) \geq 0
\end{aligned} \tag{11}
$$

Since it is given as a linear program, the complexity of computing $\mathrm{S}^{\mathrm{NS}}_{\mathrm{sum}}(W,k_1,k_2)$ is polynomial in the number of variables and constraints (see for instance Section 7.1 of [38]), which is a polynomial in $|\mathcal{X}|, |\mathcal{Y}_1|, |\mathcal{Y}_2|, k_1$ and k_2.

Similarly, we define the maximum joint success probability $\mathrm{S}^{\mathrm{NS}}(W,k_1,k_2)$ in the following way:

$$
\begin{aligned}
\mathrm{S}^{\mathrm{NS}}(W,k_1,k_2) := \quad & \underset{P}{\text{maximize}} && \frac{1}{k_1k_2}\sum_{i_1,i_2,x,y_1,y_2} W(y_1y_2|x)P(xi_1i_2|(i_1i_2)y_1y_2)\\
& \text{subject to} && \sum_{x} P(xj_1j_2|(i_1i_2)y_1y_2) = \sum_{x} P(xj_1j_2|(i_1'i_2')y_1y_2)\\
& && \sum_{j_1} P(xj_1j_2|(i_1i_2)y_1y_2) = \sum_{j_1} P(xj_1j_2|(i_1i_2)y_1'y_2)\\
& && \sum_{j_2} P(xj_1j_2|(i_1i_2)y_1y_2) = \sum_{j_2} P(xj_1j_2|(i_1i_2)y_1y_2')\\
& && \sum_{x,j_1,j_2} P(xj_1j_2|(i_1i_2)y_1y_2) = 1\\
& && P(xj_1j_2|(i_1i_2)y_1y_2) \geq 0
\end{aligned} \tag{12}
$$

We can rewrite both these programs in more convenient and smaller linear programs:

Proposition 3.4.

$$
\begin{aligned}
\mathrm{S}^{\mathrm{NS}}_{\mathrm{sum}}(W,k_1,k_2) = \quad & \underset{p,r,r^1,r^2}{\text{maximize}} && \frac{1}{2k_1k_2}\left(\sum_{x,y_1} W_1(y_1|x)r^1_{x,y_1} + \sum_{x,y_2} W_2(y_2|x)r^2_{x,y_2}\right)\\
& \text{subject to} && \sum_{x} r_{x,y_1,y_2} = 1\\
& && \sum_{x} r^1_{x,y_1} = k_2\\
& && \sum_{x} r^2_{x,y_2} = k_1\\
& && \sum_{x} p_x = k_1k_2\\
& && 0 \leq r_{x,y_1,y_2} \leq r^1_{x,y_1}, r^2_{x,y_2} \leq p_x\\
& && p_x - r^1_{x,y_1} - r^2_{x,y_2} + r_{x,y_1,y_2} \geq 0
\end{aligned} \tag{13}
$$

$$
\begin{aligned}
\mathrm{S}^{\mathrm{NS}}(W,k_1,k_2) = \quad & \underset{p,r,r^1,r^2}{\text{maximize}} && \frac{1}{k_1k_2}\sum_{x,y_1,y_2} W(y_1y_2|x)r_{x,y_1,y_2} \\
& \text{subject to} && \sum_x r_{x,y_1,y_2} = 1 \\
& && \sum_x r^1_{x,y_1} = k_2 \\
& && \sum_x r^2_{x,y_2} = k_1 \\
& && \sum_x p_x = k_1k_2 \\
& && 0 \leq r_{x,y_1,y_2} \leq r^1_{x,y_1}, r^2_{x,y_2} \leq p_x \\
& && p_x - r^1_{x,y_1} - r^2_{x,y_2} + r_{x,y_1,y_2} \geq 0
\end{aligned}
\tag{14}
$$

Proof. One can check that given a solution of the original program, the following choice of variables is a valid solution of the second program achieving the same objective value:

$$
\begin{aligned}
r_{x,y_1,y_2} &:= \sum_{i_1,i_2} P(xi_1i_2|(i_1i_2)y_1y_2) \ , \\
r^1_{x,y_1} &:= \sum_{j_2,i_1,i_2} P(xi_1j_2|(i_1i_2)y_1y_2) \ , \\
r^2_{x,y_2} &:= \sum_{j_1,i_1,i_2} P(xj_1i_2|(i_1i_2)y_1y_2) \ , \\
p_x &:= \sum_{j_1,j_2,i_1,i_2} P(xj_1j_2|(i_1i_2)y_1y_2) \ .
\end{aligned}
\tag{15}
$$

For the other direction, given those variables, a non-signaling probability distribution $P(xj_1j_2|(i_1i_2)y_1y_2)$ is given by, for $j_1 \neq i_1$ and $j_2 \neq i_2$:

$$
\begin{aligned}
P(xi_1i_2|(i_1i_2)y_1y_2) &= \frac{r_{x,y_1,y_2}}{k_1k_2} \ , \\
P(xj_1i_2|(i_1i_2)y_1y_2) &= \frac{r^2_{x,y_2} - r_{x,y_1,y_2}}{k_1k_2(k_1-1)} \ , \\
P(xi_1j_2|(i_1i_2)y_1y_2) &= \frac{r^1_{x,y_1} - r_{x,y_1,y_2}}{k_1k_2(k_2-1)} \ , \\
P(xj_1j_2|(i_1i_2)y_1y_2) &= \frac{p_x - r^1_{x,y_1} - r^2_{x,y_2} + r_{x,y_1,y_2}}{k_1k_2(k_1-1)(k_2-1)} \ .
\end{aligned}
\tag{16}
$$

□

As before, one can define the capacity regions of broadcast channels with non-signaling assistance:

Definition 3.5 (Capacity Region $\mathcal{C}^{\mathrm{NS}}(W)$ (resp. $\mathcal{C}^{\mathrm{NS}}_{\mathrm{sum}}(W)$) of a broadcast channel W)**.** A rate pair (R_1,R_2) is achievable with (full) non-signaling assistance if:

$$\lim_{n\to+\infty} \mathrm{S}^{\mathrm{NS}}(W^{\otimes n}, \lceil 2^{R_1n}\rceil, \lceil 2^{R_2n}\rceil) = 1 \ .$$

$$(\text{resp. } \lim_{n\to+\infty} \mathrm{S}^{\mathrm{NS}}_{\mathrm{sum}}(W^{\otimes n}, \lceil 2^{R_1n}\rceil, \lceil 2^{R_2n}\rceil) = 1) \ .$$

We define the joint (resp. sum) non-signaling assisted capacity region $\mathcal{C}^{\mathrm{NS}}(W)$ (resp. $\mathcal{C}^{\mathrm{NS}}_{\mathrm{sum}}(W)$) as the closure of the set of all rate pairs achievable with (full) non-signaling assistance.

Proposition 3.6. For any broadcast channel W, $\mathcal{C}^{\mathrm{NS}}(W) = \mathcal{C}^{\mathrm{NS}}_{\mathrm{sum}}(W)$.

Proof. Let us show that:

$$2\mathrm{S}^{\mathrm{NS}}_{\mathrm{sum}}(W,k_1,k_2) - 1 \leq \mathrm{S}^{\mathrm{NS}}(W,k_1,k_2) \leq \mathrm{S}^{\mathrm{NS}}_{\mathrm{sum}}(W,k_1,k_2) \ .$$

This will imply in particular that:

$$\lim_{n\to+\infty} \mathrm{S}^{\mathrm{NS}}(W^{\otimes n}, \lceil 2^{R_1 n}\rceil, \lceil 2^{R_2 n}\rceil) = 1 \iff \lim_{n\to+\infty} \mathrm{S}^{\mathrm{NS}}_{\mathrm{sum}}(W^{\otimes n}, \lceil 2^{R_1 n}\rceil, \lceil 2^{R_2 n}\rceil) = 1 \ ,$$

thus define the same capacity region.

Let us consider an optimal solution $p_x, r_{x,y_1,y_2}, r^1_{x,y_1}, r^2_{x,y_2}$ of the program computing $\mathrm{S}^{\mathrm{NS}}_{\mathrm{sum}}(W,k_1,k_2)$. We have:

$$\mathrm{S}^{\mathrm{NS}}_{\mathrm{sum}}(W,k_1,k_2) = \frac{1}{k_1k_2}\left(\sum_{x,y_1,y_2} W(y_1y_2|x)\frac{r^1_{x,y_1}+r^2_{x,y_2}}{2}\right) \ .$$

However $r^1_{x,y_1}+r^2_{x,y_2} \leq p_x + r_{x,y_1,y_2}$ so we get that:

$$\begin{aligned}\mathrm{S}^{\mathrm{NS}}_{\mathrm{sum}}(W,k_1,k_2) &\leq \frac{1}{2k_1k_2}\left(\sum_{x,y_1,y_2} W(y_1y_2|x)\left(p_x + r_{x,y_1,y_2}\right)\right) = \frac{1}{2} + \frac{1}{2}\left[\frac{1}{k_1k_2}\left(\sum_{x,y_1,y_2} W(y_1y_2|x) r_{x,y_1,y_2}\right)\right] \\ &\leq \frac{1}{2} + \frac{1}{2}\mathrm{S}^{\mathrm{NS}}(W,k_1,k_2) \ ,\end{aligned} \tag{17}$$

since $p_x, r_{x,y_1,y_2}, r^1_{x,y_1}, r^2_{x,y_2}$ is a valid solution of the program computing $\mathrm{S}^{\mathrm{NS}}(W,k_1,k_2)$.

On the other hand, consider now $p_x, r_{x,y_1,y_2}, r^1_{x,y_1}, r^2_{x,y_2}$ an optimal solution of the program computing $\mathrm{S}^{\mathrm{NS}}(W,k_1,k_2)$. We have that $r_{x,y_1,y_2} \leq r^1_{x,y_1}, r^2_{x,y_2}$ so we have that $r_{x,y_1,y_2} \leq \frac{r^1_{x,y_1}+r^2_{x,y_2}}{2}$ and thus:

$$\begin{aligned}\mathrm{S}^{\mathrm{NS}}(W,k_1,k_2) &= \frac{1}{k_1k_2}\left(\sum_{x,y_1,y_2} W(y_1y_2|x) r_{x,y_1,y_2}\right) \leq \frac{1}{k_1k_2}\left(\sum_{x,y_1,y_2} W(y_1y_2|x)\frac{r^1_{x,y_1}+r^2_{x,y_2}}{2}\right) \\ &\leq \mathrm{S}^{\mathrm{NS}}_{\mathrm{sum}}(W,k_1,k_2) \ ,\end{aligned} \tag{18}$$

since $p_x, r_{x,y_1,y_2}, r^1_{x,y_1}, r^2_{x,y_2}$ is a valid solution of the program computing $\mathrm{S}^{\mathrm{NS}}_{\mathrm{sum}}(W,k_1,k_2)$. This prove the inequalities $2\mathrm{S}^{\mathrm{NS}}_{\mathrm{sum}}(W,k_1,k_2) - 1 \leq \mathrm{S}^{\mathrm{NS}}(W,k_1,k_2) \leq \mathrm{S}^{\mathrm{NS}}_{\mathrm{sum}}(W,k_1,k_2)$, and thus concludes the proof. □

4 Approximation of Deterministic Broadcast Channel Coding

In this section, we will address the question of the approximability of $\mathrm{S}(W,k_1,k_2)$, in the restricted scenario of a deterministic broadcast channel W.

We say that W is deterministic if $\forall x, y_1, y_2, W(y_1y_2|x) \in \{0,1\}$. We can then define $(W_1(x), W_2(x))$ as the only pair (y_1,y_2) such that $W(y_1y_2|x) = 1$, which exists uniquely as W is a conditional probability distribution. Thus we have:

$$\begin{aligned}\mathrm{S}(W,k_1,k_2) = \quad & \underset{e,d_1,d_2}{\text{maximize}} && \frac{1}{k_1k_2}\sum_{x,i_1,i_2} e(x|i_1i_2)d_1(i_1|W_1(x))d_2(i_2|W_2(x)) \\ & \text{subject to} && \sum_{x\in\mathcal{X}} e(x|i_1i_2) = 1, \forall i_1 \in [k_1], i_2 \in [k_2] \\ &&& \sum_{j_1\in[k_1]} d_1(j_1|y_1) = 1, \forall y_1 \in \mathcal{Y}_1 \\ &&& \sum_{j_2\in[k_2]} d_2(j_2|y_2) = 1, \forall y_2 \in \mathcal{Y}_2 \\ &&& e(x|i_1i_2), d_1(j_1|y_2), d_2(j_2|y_2) \geq 0\end{aligned} \tag{19}$$

4.1 Reformulation as a Graph Problem

A deterministic channel W, up to a permutation of elements of $\mathcal{X}$, is characterized by the following bipartite graph:

$$G_W := (\mathcal{Y}_1 \sqcup \mathcal{Y}_2, E = \{(y_1, y_2) \in \mathcal{Y}_1 \times \mathcal{Y}_2 : \exists x \in \mathcal{X}, y_1 = W_1(x) \text{ and } y_2 = W_2(x)\}) \ .$$

Indeed, permuting the elements of $\mathcal{X}$ won't change G_W, and permuting elements of $\mathcal{Y}_1$ or $\mathcal{Y}_2$ will result in an isomorphic graph. These permutations do not change $\mathrm{S}(W, k_1, k_2)$, so we can ignore them in order to compute this quantity.

As a consequence, the quantity $k_1 k_2 \mathrm{S}(W, k_1, k_2)$ can be seen as the solution of the following bipartite graph problem on G_W, k_1 and k_2, which we call DensestQuotientGraph: given a bipartite graph G and integers k_1, k_2, maximize, over all quotient graphs of G in k_1 parts on the left and k_2 parts on the right, the resulting number of edges. Formally, this gives:

Definition 4.1 (Graph notations)**.** Consider a bipartite graph $G = (V_1 \sqcup V_2, E \subseteq V_1 \times V_2)$:

1. $G^{\mathcal{P}_1,\mathcal{P}_2}$, the quotient of G by partitions $\mathcal{P}_1, \mathcal{P}_2$ of respectively V_1, V_2, is defined by:
$$G^{\mathcal{P}_1,\mathcal{P}_2} := (\mathcal{P}_1 \sqcup \mathcal{P}_2, \{(p_1, p_2) \in \mathcal{P}_1 \times \mathcal{P}_2 : E \cap (p_1 \times p_2) \neq \emptyset\}) \ .$$
2. $e_G(\mathcal{P}_1, \mathcal{P}_2) := |E^{G^{\mathcal{P}_1,\mathcal{P}_2}}|$ is the number of edges of $G^{\mathcal{P}_1,\mathcal{P}_2}$.
3. $N_G^{\mathcal{P}_1,\mathcal{P}_2}(p) := N_{G^{\mathcal{P}_1,\mathcal{P}_2}}(p)$ is the set of neighbors of p in the graph $G^{\mathcal{P}_1,\mathcal{P}_2}$.
4. Similarly, $\deg_G^{\mathcal{P}_1,\mathcal{P}_2}(p) := \deg_{G^{\mathcal{P}_1,\mathcal{P}_2}}(p)$ is the degree, ie. the number of neighbors, of p in the graph $G^{\mathcal{P}_1,\mathcal{P}_2}$.
5. We will use V_1, V_2 in previous notations when we do not partition on the left and right part respectively (or identify those to trivial partitions in singletons). For instance, $G^{V_1,V_2} = G$.
6. We will use the notations $e(\mathcal{P}_1, \mathcal{P}_2)$, $N_{\mathcal{P}_1,\mathcal{P}_2}(p)$ and $\deg_{\mathcal{P}_1,\mathcal{P}_2}(p)$ when the graph G considered is clear from the context.

Definition 4.2.

$$\textsc{DensestQuotientGraph}(G, k_1, k_2) := \underset{\mathcal{P}_1 \text{ in } k_1 \text{ parts}, \mathcal{P}_2 \text{ in } k_2 \text{ parts}}{\text{maximize}} e_G(\mathcal{P}_1, \mathcal{P}_2) \ .$$

Let us show the equivalence with the broadcast channel coding problem:

Proposition 4.3. We have $k_1 k_2 \mathrm{S}(W, k_1, k_2) = \textsc{DensestQuotientGraph}(G_W, k_1, k_2)$.

Proof. An optimal solution achieving $\mathrm{S}(W, k_1, k_2)$ can be obtained with extremal points as the optimization program defining $\mathrm{S}(W, k_1, k_2)$ is convex. We can thus assume that e, d_1, d_2 are functions. Therefore, d_1 defines a partition of $\mathcal{Y}_1$, d_2 defines a partition of $\mathcal{Y}_2$, and then e denotes the choice of which edge between partition i_1 and i_2 one choose. Indeed, with $\mathcal{P}_1, \mathcal{P}_2$ corresponding partitions of $\mathcal{Y}_1, \mathcal{Y}_2$ in k_1, k_2 parts:

$$\begin{aligned}
k_1 k_2 \mathrm{S}(W, k_1, k_2) &= \underset{e, d_1, d_2}{\text{maximize}} \sum_{x, i_1, i_2} \mathbb{1}_{x = e(i_1, i_2)} \mathbb{1}_{i_1 = d_1(W_1(x))} \mathbb{1}_{i_2 = d_2(W_2(x))} \\
&= \underset{e, d_1, d_2}{\text{maximize}} \sum_{i_1, i_2} \mathbb{1}_{i_1 = d_1(W_1(e(i_1, i_2)))} \mathbb{1}_{i_2 = d_2(W_2(e(i_1, i_2)))} \\
&= \underset{e, \mathcal{P}_1, \mathcal{P}_2}{\text{maximize}} \sum_{i_1, i_2} \mathbb{1}_{W_1(e(i_1, i_2)) \in \mathcal{P}_1^{i_1}} \mathbb{1}_{W_2(e(i_1, i_2)) \in \mathcal{P}_2^{i_2}} \\
&= \underset{\mathcal{P}_1, \mathcal{P}_2}{\text{maximize}} \sum_{i_1, i_2} \max_{x \in \mathcal{X}} \mathbb{1}_{W_1(x) \in \mathcal{P}_1^{i_1}} \mathbb{1}_{W_2(x) \in \mathcal{P}_2^{i_2}} \\
&= \underset{\mathcal{P}_1, \mathcal{P}_2}{\text{maximize}} \sum_{i_1, i_2} \mathbb{1}_{\exists (y_1, y_2) \in E^{G_W} : y_1 \in \mathcal{P}_1^{i_1} \text{ and } y_2 \in \mathcal{P}_2^{i_2}} \\
&= \underset{\mathcal{P}_1, \mathcal{P}_2}{\text{maximize}} \sum_{i_1, i_2} \mathbb{1}_{E^{G_W} \cap \left(\mathcal{P}_1^{i_1} \times \mathcal{P}_2^{i_2}\right) \neq \emptyset} \\
&= \underset{\mathcal{P}_1, \mathcal{P}_2}{\text{maximize}} \ e_{G_W}(\mathcal{P}_1, \mathcal{P}_2) \ .
\end{aligned} \tag{20}$$

□

4.2 Approximation Algorithm for DensestQuotientGraph

In this section, we will sort out how hard is DensestQuotientGraph, and thanks to Proposition 4.3, how hard is it to solve the broadcast channel coding problem for deterministic channels.

Theorem 4.4. There exists a polynomial-time $(1-e^{-1})^2$-approximation algorithm for DensestQuotientGraph. Furthermore, it is NP-hard to solve exactly DensestQuotientGraph.

Corollary 4.5. There exists a polynomial-time $(1-e^{-1})^2$-approximation algorithm for the deterministic broadcast channel coding problem.

The approximation algorithm is a two-step process. First, we consider the problem of maximizing $\sum_{i_2=1}^{k_2} \min\left(k_1, \deg_{V_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})\right)$ over all partitions $\mathcal{P}_2$ of V_2 in k_2 parts. We will show that this is a special case of the submodular welfare problem, which can be approximated within a factor $1-e^{-1}$ [40]. We then choose the partition $\mathcal{P}_1$ on V_1 in k_1 parts uniformly at random. This partition couple will give an objective value $e(\mathcal{P}_1, \mathcal{P}_2)$ within a $(1-e^{-1})^2$ factor from the optimal solution in expectation.

Proof of Theorem 4.4. Consider first the hardness result. Let us show that the decision version of DensestQuotientGraph is NP-complete. It is in NP, the certificate being the two partitions and the selection of edges between those partitions. It is NP-hard as one of its particular cases is the SetSplitting problem (see for instance [39]), in the case where $k_1 = 2$ and $k_2 = |V_2|$, by interpreting the neighbors of $v_2 \in V_2$ as a set covering elements of V_1.

We will show nonetheless that this problem can be approximated within a factor $(1-e^{-1})^2$ in polynomial time. First we consider the case where $k_2 = |V_2|$. We can then always assume that the right partition is $\mathcal{P}_2 := \{\{v_2\} : v_2 \in V_2\}$, which leads necessarily to a greater or equal number of edges in the quotient graph that with any other right partition. So, in that setting, we need only to find a partition of V_1 in k_1 parts maximizing the number of edges between vertices in the right part and the quotient of the left vertices.

First, one can note that the maximum value we can get is upper bounded by $\sum_{v_2\in V_2} \min\left(k_1, \deg(v_2)\right)$. Indeed, each vertex of v_2 van be connected at most to the k_1 parts of V_1, so its contribution is bounded by k_1, and there needs to be an edge to each part it is connected, so its contribution is also bounded by $\deg(v_2)$. Let us show that if we take a partition $\mathcal{P}_1$ of V_1 uniformly at random, we get:

$$\mathbb{E}_{\mathcal{P}_1}[e(\mathcal{P}_1, V_2)] \geq \left(1 - \left(1 - \frac{1}{k_1}\right)^{k_1}\right) \sum_{v_2\in V_2} \min\left(k_1, \deg(v_2)\right) \geq (1-e^{-1}) \max_{\mathcal{P}_1} e(\mathcal{P}_1, V_2) \ .$$

We have that $e(\mathcal{P}_1, V_2) = \sum_{v_2\in V_2} \deg_{\mathcal{P}_1,V_2}(v_2)$, so by linearity of expectation $\mathbb{E}_{\mathcal{P}_1}[e(\mathcal{P}_1, V_2)] = \sum_{v_2\in V_2} \mathbb{E}_{\mathcal{P}_1}[\deg_{\mathcal{P}_1,V_2}(v_2)]$. However $\deg_{\mathcal{P}_1,V_2}(v_2) = |\{i_1 \in [k_1] : N(v_2)\cap \mathcal{P}_1^{i_1} \neq \emptyset\}|$. Recall also that for any v_1, $\mathbb{P}\left(v_1 \in \mathcal{P}_1^{i_1}\right) = \frac{1}{k_1}$ since the partition is taken uniformly at random. Thus, we get:

$$\begin{aligned}
\mathbb{E}_{\mathcal{P}_1}[\deg_{\mathcal{P}_1,V_2}(v_2)] &= \mathbb{E}_{\mathcal{P}_1}\left[|\{i_1 \in [k_1] : N(v_2)\cap \mathcal{P}_1^{i_1} \neq \emptyset\}|\right] = \mathbb{E}_{\mathcal{P}_1}\left[\sum_{i_1=1}^{k_1} \mathbb{1}_{N(v_2)\cap\mathcal{P}_1^{i_1}\neq\emptyset}\right] \\
&= \sum_{i_1=1}^{k_1} \mathbb{E}_{\mathcal{P}_1}\left[\mathbb{1}_{N(v_2)\cap\mathcal{P}_1^{i_1}\neq\emptyset}\right] = \sum_{i_1=1}^{k_1} \mathbb{P}\left(N(v_2)\cap\mathcal{P}_1^{i_1}\neq\emptyset\right) \\
&= \sum_{i_1=1}^{k_1} \left(1 - \mathbb{P}\left(N(v_2)\cap\mathcal{P}_1^{i_1}=\emptyset\right)\right) = \sum_{i_1=1}^{k_1}\left(1 - \prod_{v_1\in N(v_2)} \mathbb{P}\left(v_1 \notin \mathcal{P}_1^{i_1}\right)\right) \\
&= \sum_{i_1=1}^{k_1}\left(1 - \prod_{v_1\in N(v_2)} \mathbb{P}\left(v_1 \notin \mathcal{P}_1^{i_1}\right)\right) = k_1\left(1 - \left(1-\frac{1}{k_1}\right)^{\deg(v_2)}\right) ,
\end{aligned} \tag{21}$$

since $\mathbb{P}\left(v_1 \notin \mathcal{P}_1^{i_1}\right) = 1 - \frac{1}{k_1}$ and $|N(v_2)| = \deg(v_2)$. So, in all:

$$\mathbb{E}_{\mathcal{P}_1}[e(\mathcal{P}_1, V_2)] = \sum_{v_2 \in V_2} \mathbb{E}_{\mathcal{P}_1}[\deg_{\mathcal{P}_1, V_2}(v_2)] = k_1 \sum_{v_2 \in V_2} \left(1 - \left(1 - \frac{1}{k_1}\right)^{\deg(v_2)}\right) .$$

However, the function $f : x \mapsto 1 - \left(1 - \frac{1}{k_1}\right)^x$ is nondecreasing concave with $f(0) = 0$, so $\frac{f(x)}{x} \geq \frac{f(y)}{y}$ for $x \leq y$. In particular, we have that:

$$f(\min(k_1, \deg(v_2))) \geq \frac{\min(k_1, \deg(v_2))}{k_1} f(k_1) ,$$

and thus:

$$\begin{aligned}\mathbb{E}_{\mathcal{P}_1}[e(\mathcal{P}_1, V_2)] &\geq k_1 \sum_{v_2 \in V_2} \left(1 - \left(1 - \frac{1}{k_1}\right)^{\min(k_1, \deg(v_2))}\right) \geq k_1 \frac{\sum_{v_2 \in V_2} \min(k_1, \deg(v_2))}{k_1} \left(1 - \left(1 - \frac{1}{k_1}\right)^{k_1}\right) \\ &\geq \left(1 - \left(1 - \frac{1}{k_1}\right)^{k_1}\right) \sum_{v_2 \in V_2} \min\left(k_1, \deg(v_2)\right) \geq (1 - e^{-1}) \max_{\mathcal{P}_1} e(\mathcal{P}_1, V_2) ,\end{aligned} \tag{22}$$

Let us now consider the general case with k_2 unconstrained. We apply the previous discussion on the graph $G^{V_1, \mathcal{P}_2}$ for some fixed partition $\mathcal{P}_2$ of V_2. Since $e_{G^{V_1, \mathcal{P}_2}}(\mathcal{P}_1, \mathcal{P}_2) = e(\mathcal{P}_1, \mathcal{P}_2)$, we have the upper bound:

$$\max_{\mathcal{P}_1} e(\mathcal{P}_1, \mathcal{P}_2) \leq \sum_{i_2=1}^{k_2} \min\left(k_1, \deg_{V_1, \mathcal{P}_2}(\mathcal{P}_2^{i_2})\right) ,$$

and the previous algorithm gives us a partition $\mathcal{P}_1$ of V_1 such that:

$$e(\mathcal{P}_1, \mathcal{P}_2) \geq (1 - e^{-1}) \sum_{i_2=1}^{k_2} \min\left(k_1, \deg_{V_1, \mathcal{P}_2}(\mathcal{P}_2^{i_2})\right) .$$

Therefore, let us focus on the following optimization problem:

$$\max_{\mathcal{P}_2} \sum_{i_2=1}^{k_2} \min\left(k_1, \deg_{V_1, \mathcal{P}_2}(\mathcal{P}_2^{i_2})\right) ,$$

We will give a $(1 - e^{-1})$ approximation algorithm running in polynomial time for this problem. In all, this will allow us to get in polynomial time a partition pair $(\mathcal{P}_1, \mathcal{P}_2)$ such that:

$$\begin{aligned}e(\mathcal{P}_1, \mathcal{P}_2) &\geq (1 - e^{-1}) \sum_{i_2=1}^{k_2} \min\left(k_1, \deg_{V_1, \mathcal{P}_2}(\mathcal{P}_2^{i_2})\right) \\ &\geq (1 - e^{-1})^2 \max_{\mathcal{P}_2} \sum_{i_2=1}^{k_2} \min\left(k_1, \deg_{V_1, \mathcal{P}_2}(\mathcal{P}_2^{i_2})\right) \\ &\geq (1 - e^{-1})^2 \max_{\mathcal{P}_1, \mathcal{P}_2} e(\mathcal{P}_1, \mathcal{P}_2) .\end{aligned} \tag{23}$$

The problem $\max_{\mathcal{P}_2} \sum_{i_2=1}^{k_2} \min\left(k_1, \deg_{V_1, \mathcal{P}_2}(\mathcal{P}_2^{i_2})\right)$ is a particular instance of the submodular welfare problem discussed in [40]. Note first that $\deg_{V_1, \mathcal{P}_2}(\mathcal{P}_2^{i_2}) = \deg_{V_1, \{\mathcal{P}_2^{i_2}, V_2 - \mathcal{P}_2^{i_2}\}}(\mathcal{P}_2^{i_2})$, as the degree of $\mathcal{P}_2^{i_2}$ does not depend on the rest of the partition $\mathcal{P}_2$. Then, $h(S_2) := \min\left(k_1, \deg_{V_1, \{S_2, V_2 - S_2\}}(S_2)\right)$, for $S_2 \subseteq V_2$, is a nondecreasing submodular function, as $S_2 \mapsto \deg_{V_1, \{S_2, V_2 - S_2\}}(S_2)$ is a nondecreasing submodular function on V_2 and $\min(k_1, \cdot)$ is nondecreasing concave. Thus, we want to maximize $\sum_{i_2=1}^{k_2} h(S_{i_2})$ where $(S_{i_2})_{i_2 \in [k_2]}$ is a partition of items in V_2 among k_2 bidders. It is a particular case of the submodular welfare problem where each nondecreasing submodular utility weight is the same for all bidders and equal to h. Thus, thanks to [40], there exists a $(1 - e^{-1})$ polynomial-time approximation of $\max_{\mathcal{P}_2} \sum_{i_2=1}^{k_2} \min\left(k_1, \deg_{V_1, \mathcal{P}_2}(\mathcal{P}_2^{i_2})\right)$. □

4.3 Deterministic Non-Signaling Assisted Capacity Region

Thanks to Theorem 4.4 and Proposition 4.3, there exists a constant factor approximation algorithm for the broadcast channel coding problem running in polynomial time. We aim to show here that that the non-signaling assisted value is linked by a constant factor to the unassisted one. Indeed, the hope is that the non-signaling assisted program is linked to the linear relaxation of the unassisted problem, thus is likely a good approximation since the broadcast channel coding problem can be approximated in polynomial time.

This turns out to be true, and will be proved through the following theorem:

Theorem 4.6. If W is a deterministic broadcast channel, then for all $\ell_1 \leq k_1$ and $\ell_2 \leq k_2$:

$$\left(1-\frac{k_1^{k_1}e^{-k_1}}{k_1!}\right)\left(1-\left(1-\frac{1}{\ell_1}\right)^{k_1}\right)\left(1-\left(1-\frac{1}{\ell_2}\right)^{k_2}\right)\mathrm{S}^{\mathrm{NS}}(W,k_1,k_2) \leq \mathrm{S}(W,\ell_1,\ell_2) \ .$$

Corollary 4.7. For any deterministic broadcast channel W, $\mathcal{C}^{\mathrm{NS}}(W) = \mathcal{C}(W)$.

Proof. We apply Theorem 4.6 on the deterministic broadcast channel $W^{\otimes n}$.

We fix $k_1 = 2^{nR_1}, k_2 = 2^{nR_2}$ and $\ell_1 = \frac{2^{nR_1}}{n}, \ell_2 = \frac{2^{nR_2}}{n}$. Since $1-\left(1-\frac{1}{\ell}\right)^k \geq 1-e^{-\frac{k}{\ell}}$, we get:

$$\left(1-\frac{k_1^{k_1}e^{-k_1}}{k_1!}\right)\left(1-e^{-n}\right)^2\mathrm{S}^{\mathrm{NS}}(W^{\otimes n},2^{nR_1},2^{nR_1}) \leq \mathrm{S}\left(W^{\otimes n},\frac{2^{nR_1}}{n},\frac{2^{nR_2}}{n}\right) \ .$$

As $\left(1-\frac{k_1^{k_1}e^{-k_1}}{k_1!}\right)\left(1-e^{-n}\right)^2$ tends to 1 when n tends to infinity, we get that $\forall\varepsilon, \exists N\in\mathbb{N}, \forall n\geq N$:

$$(1-\varepsilon)\mathrm{S}^{\mathrm{NS}}(W^{\otimes n},2^{nR_1},2^{nR_1}) \leq \mathrm{S}(W^{\otimes n},2^{n(R_1-\frac{\log(n)}{n})},2^{n(R_2-\frac{\log(n)}{n})}) \ .$$

Thus, if $\lim\limits_{n\to+\infty}\mathrm{S}^{\mathrm{NS}}(W^{\otimes n},2^{nR_1},2^{nR_1}) = 1$, we have that for all $R_1' < R_1$ and $R_2' < R_2$:

$$\lim_{n\to+\infty}\mathrm{S}(W^{\otimes n},2^{nR_1'},2^{nR_1'}) \geq 1-\varepsilon \ .$$

Since this is true for all $\varepsilon > 0$, we get in fact that $\lim\limits_{n\to+\infty}\mathrm{S}(W^{\otimes n},2^{nR_1'},2^{nR_1'}) = 1$. This implies that $\mathcal{C}^{\mathrm{NS}}(W) \subseteq \mathcal{C}(W)$, and thus that the capacity regions are equal as the other inclusion is always satisfied. □

Let us now prove the main result:

Proof of Theorem 4.6. The proof will be done in three parts. We will work on the graph G_W:

1. First, we prove that for any partition $\mathcal{P}_2$ of $\mathcal{Y}_2$ in ℓ_2 parts:

$$\mathrm{S}(W,\ell_1,\ell_2) \geq \left(1-\left(1-\frac{1}{\ell_1}\right)^{k_1}\right)\frac{\sum_{i_2=1}^{\ell_2}\min\left(k_1,\deg_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})\right)}{k_1\ell_2} \ ,$$

2. Then, we show that there exists a partition $\mathcal{P}_2$ such that:

$$\frac{\sum_{i_2=1}^{\ell_2}\min\left(k_1,\deg_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})\right)}{k_1\ell_2} \geq \left(1-\frac{k_1^{k_1}e^{-k_1}}{k_1!}\right)\left(1-\left(1-\frac{1}{\ell_2}\right)^{k_2}\right)\frac{\min\left(k_1k_2,\sum_{y_1}\min(k_2,\deg(y_1))\right)}{k_1k_2} \ ,$$

3. Finally, we prove that:

$$\frac{\min\left(k_1k_2,\sum_{y_1}\min(k_2,\deg(y_1))\right)}{k_1k_2} \geq \mathrm{S}^{\mathrm{NS}}(W,k_1,k_2) \ .$$

By combining these three inequalities, we get precisely the claimed result.

1. This part shares a lot of similarities with the proof of Theorem 4.4, which we will adapt to this particular situation. Let us show that if we take a partition $\mathcal{P}_1$ of $\mathcal{Y}_1$ of size ℓ_1 uniformly at random, we get, for some fixed $\mathcal{P}_2$ of size ℓ_2:

$$\mathbb{E}_{\mathcal{P}_1}[e_{G_W}(\mathcal{P}_1,\mathcal{P}_2)] \geq \frac{\ell_1}{k_1}\left(1-\left(1-\frac{1}{\ell_1}\right)^{k_1}\right)\sum_{i_2=1}^{\ell_2}\min\left(k_1, \deg_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})\right) \ .$$

Then, since $\ell_1\ell_2\mathrm{S}(W,\ell_1,\ell_2) = \underset{\mathcal{P}_1 \text{ in } \ell_1 \text{ parts}, \mathcal{P}_2 \text{ in } \ell_2 \text{ parts}}{\text{maximize}} e_{G_W}(\mathcal{P}_1,\mathcal{P}_2)$ by Proposition 4.3, this will imply that:

$$\mathrm{S}(W,\ell_1,\ell_2) \geq \frac{1}{\ell_1\ell_2}\mathbb{E}_{\mathcal{P}_1}[e_{G_W}(\mathcal{P}_1,\mathcal{P}_2)] \geq \left(1-\left(1-\frac{1}{\ell_1}\right)^{k_1}\right)\frac{\sum_{i_2=1}^{\ell_2}\min\left(k_1,\deg_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})\right)}{k_1\ell_2} \ .$$

We have that $e_{G_W}(\mathcal{P}_1,\mathcal{P}_2) = \sum_{i_2=1}^{\ell_2}\deg_{\mathcal{P}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})$, so by linearity of expectation, we have that $\mathbb{E}_{\mathcal{P}_1}[e_{G_W}(\mathcal{P}_1,\mathcal{P}_2)] = \sum_{i_2=1}^{\ell_2}\mathbb{E}_{\mathcal{P}_1}[\deg_{\mathcal{P}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})]$, so we will focus on the contribution of one particular $\mathcal{P}_2^{i_2}$.
Then, we have that $\deg_{\mathcal{P}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2}) = |\{i_1 \in [\ell_1] : N_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})\cap\mathcal{P}_1^{i_1} \neq \emptyset\}|$. Recall that $\mathbb{P}\left(v_1 \in \mathcal{P}_1^{i_1}\right) = \frac{1}{\ell_1}$ for any v_1 since the partition is taken uniformly at random. Thus:

$$\begin{aligned}
\mathbb{E}_{\mathcal{P}_1}[\deg_{\mathcal{P}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})] &= \mathbb{E}_{\mathcal{P}_1}\left[|\{i_1 \in [\ell_1] : N_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})\cap\mathcal{P}_1^{i_1} \neq \emptyset\}|\right] = \mathbb{E}_{\mathcal{P}_1}\left[\sum_{i_1=1}^{\ell_1}\mathbb{1}_{N_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})\cap\mathcal{P}_1^{i_1}\neq\emptyset}\right] \\
&= \sum_{i_1=1}^{\ell_1}\mathbb{E}_{\mathcal{P}_1}\left[\mathbb{1}_{N_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})\cap\mathcal{P}_1^{i_1}\neq\emptyset}\right] = \sum_{i_1=1}^{\ell_1}\mathbb{P}\left(N_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})\cap\mathcal{P}_1^{i_1}\neq\emptyset\right) \\
&= \sum_{i_1=1}^{\ell_1}\left(1-\mathbb{P}\left(N_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})\cap\mathcal{P}_1^{i_1}=\emptyset\right)\right) = \sum_{i_1=1}^{\ell_1}\left(1-\prod_{v_1\in N(\mathcal{P}_2^{i_2})}\mathbb{P}\left(v_1\notin\mathcal{P}_1^{i_1}\right)\right) \\
&= \ell_1\left(1-\left(1-\frac{1}{\ell_1}\right)^{\deg_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})}\right) \ .
\end{aligned} \tag{24}$$

So, in all we have that:

$$\mathbb{E}_{\mathcal{P}_1}[e_{G_W}(\mathcal{P}_1,\mathcal{P}_2)] = \sum_{i_2=1}^{\ell_2}\mathbb{E}_{\mathcal{P}_1}[\deg_{\mathcal{P}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})] = \ell_1\sum_{i_2=1}^{\ell_2}\left(1-\left(1-\frac{1}{\ell_1}\right)^{\deg_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})}\right) \ .$$

However the function $f : x \mapsto 1-\left(1-\frac{1}{\ell_1}\right)^x$ is nondecreasing concave with $f(0)=0$, so $\frac{f(x)}{x} \geq \frac{f(y)}{y}$ for $x \leq y$. In particular, we have that:

$$f(\min(k_1,\deg_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2}))) \geq \frac{\min(k_1,\deg_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})))}{k_1}f(k_1) \ ,$$

and thus:

$$\begin{aligned}
\mathbb{E}_{\mathcal{P}_1}[e_{G_W}(\mathcal{P}_1,\mathcal{P}_2)] &\geq \ell_1\sum_{i_2=1}^{\ell_2}\left(1-\left(1-\frac{1}{\ell_1}\right)^{\min(k_1,\deg_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2}))}\right) \\
&\geq \ell_1\frac{\sum_{i_2=1}^{\ell_2}\min(k_1,\deg_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2}))}{k_1}\left(1-\left(1-\frac{1}{\ell_1}\right)^{k_1}\right) \\
&= \frac{\ell_1}{k_1}\left(1-\left(1-\frac{1}{\ell_1}\right)^{k_1}\right)\sum_{i_2=1}^{\ell_2}\min\left(k_1,\deg_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})\right) \ ,
\end{aligned} \tag{25}$$

which concludes the first part of the proof.

2. Let us take $\mathcal{P}_2$ a partition of $\mathcal{Y}_2$ of size ℓ_2 uniformly at random, and let us prove that

$$\mathbb{E}\left[\sum_{i_2=1}^{\ell_2} \min\left(k_1, \deg_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})\right)\right]$$

is greater than or equal to

$$\frac{\ell_2}{k_2}\left(1-\frac{k_1^{k_1}e^{-k_1}}{k_1!}\right)\left(1-\left(1-\frac{1}{\ell_2}\right)^{k_2}\right)\min\left(k_1k_2, \sum_{y_1}\min(k_2,\deg(y_1))\right) .$$

First, $\sum_{i_2=1}^{\ell_2}\min\left(k_1, \deg_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})\right) = \sum_{i_2=1}^{\ell_2}\varphi(\deg_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2}))$ with $\varphi(j) := \min(k_1, j)$ which is a concave function. We will use several tools introduced in [41] to handle those concave functions in approximation algorithms. Specifically, we have that the Poisson concavity ratio $\alpha_\varphi := \inf_{x\in\mathbb{R}^+}\frac{\mathbb{E}[\varphi(\mathrm{Poi}(x))]}{\varphi(x))} = 1-\frac{k_1^{k_1}e^{-k_1}}{k_1!}$ for that particular function. We will also need the following lemma, whose proof is based on convex order:

Lemma 4.8 (Lemma 2.2 of [41]). For φ concave, and $p\in[0,1]^m$, we have:

$$\mathbb{E}\Big[\varphi\Big(\sum_{i=1}^m \mathrm{Ber}(p_i)\Big)\Big] \geq \mathbb{E}\Big[\varphi\Big(\mathrm{Poi}\left(\sum_{i=1}^m p_i\right)\Big)\Big] .$$

Let us find the law of $\deg_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})$:

$$\begin{aligned}\deg_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2}) &= \sum_{y_1}\mathbb{1}_{N(y_1)\cap\mathcal{P}_2^{i_2}\neq\emptyset} = \sum_{y_1}\left(1-\mathbb{1}_{N(y_1)\cap\mathcal{P}_2^{i_2}=\emptyset}\right) = \sum_{y_1}\left(1-\mathbb{1}_{\forall y_2\in N(y_1), y_2\notin\mathcal{P}_2^{i_2}}\right)\\ &= \sum_{y_1}\mathrm{Ber}\left(1-\left(1-\frac{1}{\ell_2}\right)^{\deg_(y_1)}\right)\end{aligned} \tag{26}$$

Thus:

$$\begin{aligned}\mathbb{E}\left[\varphi(\deg_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2}))\right] &= \mathbb{E}\left[\varphi\left(\sum_{y_1}\mathrm{Ber}\left(1-\left(1-\frac{1}{\ell_2}\right)^{\deg(y_1)}\right)\right)\right]\\ &\geq \mathbb{E}\left[\varphi\left(\mathrm{Poi}\left(\sum_{y_1}\left(1-\left(1-\frac{1}{\ell_2}\right)^{\deg(y_1)}\right)\right)\right)\right] \text{ by Lemma 4.8}\\ &\geq \alpha_\varphi\varphi\left(\sum_{y_1}\left(1-\left(1-\frac{1}{\ell_2}\right)^{\deg(y_1)}\right)\right) \text{ by definition of } \alpha_\varphi.\end{aligned} \tag{27}$$

But:

$$\begin{aligned}\sum_{y_1}\left(1-\left(1-\frac{1}{\ell_2}\right)^{\deg(y_1)}\right) &\geq \sum_{y_1}\left(1-\left(1-\frac{1}{\ell_2}\right)^{\min(k_2,\deg(y_1)))}\right)\\ &\geq \left(1-\left(1-\frac{1}{\ell_2}\right)^{k_2}\right)\frac{1}{k_2}\sum_{y_1}\min\left(k_2,\deg(y_1)\right) ,\end{aligned} \tag{28}$$

as before. Since φ is concave and $\varphi(0)=0$, we have in particular that for all $0\leq c\leq 1$ and $x\in\mathbb{R}, \varphi(cx)\geq c\varphi(x)$. We know also that φ is nondecreasing. This implies that:

$$\varphi\left(\sum_{y_1}\left(1-\left(1-\frac{1}{\ell_2}\right)^{\deg(y_1)}\right)\right) \geq \varphi\left(\left(1-\left(1-\frac{1}{\ell_2}\right)^{k_2}\right)\frac{1}{k_2}\sum_{y_1}\min\left(k_2,\deg(y_1)\right)\right)$$
$$\geq \left(1-\left(1-\frac{1}{\ell_2}\right)^{k_2}\right)\varphi\left(\frac{1}{k_2}\sum_{y_1}\min\left(k_2,\deg(y_1)\right)\right)\ , \tag{29}$$

as $0 \leq 1-\left(1-\frac{1}{\ell_2}\right)^{k_2} \leq 1$. Thus:

$$\mathbb{E}\left[\varphi(\deg_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2}))\right] \geq \alpha_\varphi\left(1-\left(1-\frac{1}{\ell_2}\right)^{k_2}\right)\min\left(k_1,\frac{1}{k_2}\sum_{y_1}\min\left(k_2,\deg(y_1)\right)\right)$$
$$= \frac{1}{k_2}\left(1-\frac{k_1^{k_1}e^{-k_1}}{k_1!}\right)\left(1-\left(1-\frac{1}{\ell_2}\right)^{k_2}\right)\min\left(k_1k_2,\sum_{y_1}\min\left(k_2,\deg(y_1)\right)\right) \tag{30}$$

since $\alpha_\varphi = 1-\frac{k_1^{k_1}e^{-k_1}}{k_1!}$.

Finally, $\mathbb{E}\left[\sum_{i_2=1}^{\ell_2}\min\left(k_1,\deg_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})\right)\right] = \sum_{i_2=1}^{\ell_2}\mathbb{E}\left[\varphi(\deg_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2}))\right]$, so we get that

$$\mathbb{E}\left[\sum_{i_2=1}^{\ell_2}\min\left(k_1,\deg_{\mathcal{Y}_1,\mathcal{P}_2}(\mathcal{P}_2^{i_2})\right)\right]$$

is larger than

$$\frac{\ell_2}{k_2}\left(1-\frac{k_1^{k_1}e^{-k_1}}{k_1!}\right)\left(1-\left(1-\frac{1}{\ell_2}\right)^{k_2}\right)\min\left(k_1k_2,\sum_{y_1}\min\left(k_2,\deg(y_1)\right)\right)\ .$$

Thus, in particular, there exists some partition $\mathcal{P}_2$ that satisfies the same inequality, which concludes the second part of the proof.

3. Let us consider an optimal solution $r_{x,y_1,y_2}, p_x, r^1_{x,y_1}, r^2_{x,y_2}$ of the program computing $\mathrm{S}^{\mathrm{NS}}(W,k_1,k_2)$, so that $\mathrm{S}^{\mathrm{NS}}(W,k_1,k_2) = \frac{1}{k_1k_2}\sum_x r_{x,W_1(x),W_2(x)}$.
   (a) It comes directly from $r_{x,y_1,y_2} \leq p_x$ that:
   $$\sum_x r_{x,W_1(x),W_2(x)} \leq \sum_x p_x = k_1k_2\ .$$
   (b) $\sum_x r_{x,W_1(x),W_2(x)} = \sum_{y_1}\sum_{x:W_1(x)=y_1} r_{x,y_1,W_2(x)}$ and we have that:
      i. $\sum_{x:W_1(x)=y_1} r_{x,y_1,W_2(x)} \leq \sum_{x:W_1(x)=y_1} 1 = \deg(y_1)$,
      ii. $\sum_{x:W_1(x)=y_1} r_{x,y_1,W_2(x)} \leq \sum_{x:W_1(x)=y_1} r^1_{x,y_1} \leq \sum_x r^1_{x,y_1} = k_2$,

   so $\sum_{x:W_1(x)=y_1} r_{x,y_1,W_2(x)} \leq \min(k_2,\deg(y_1))$, and thus:
   $$\sum_x r_{x,W_1(x),W_2(x)} \leq \sum_{y_1}\min(k_2,\deg(y_1))\ .$$

In all, we get that:

$$\mathrm{S}^{\mathrm{NS}}(W,k_1,k_2) = \frac{1}{k_1k_2}\sum_x r_{x,W_1(x),W_2(x)} \leq \frac{\min\left(k_1k_2,\sum_{y_1}\min(k_2,\deg(y_1))\right)}{k_1k_2}\ ,$$

which concludes the third and last part of the proof.

□

5 Hardness of Approximation of Broadcast Channel Coding

The goal of this section is to show that the general broadcast channel coding problem cannot be approximated in polynomial time within a $\Omega(1)$ factor, under reasonable hardness assumptions. It will be a good insight that non-signaling assistance will enlarge the capacity region of the channel as discussed in the introduction.

Formally, one would want to show that it is NP-hard to approximate this problem within a $\Omega(1)$ factor in polynomial time. As a first step towards this goal, we will prove a $\Omega\left(\frac{1}{\sqrt{m}}\right)$-approximation hardness in the value query model.

5.1 Social Welfare Reformulation

The social welfare maximization problem is defined as follows: given a set M of m items as well as k bidders with their associated utilities $\left(v_i : 2^M \to \mathbb{R}_+\right)_{i\in[k]}$, the goal is to partition M between the bidders to maximize the sum of their utilities. Formally, we want to compute:

$$\underset{\mathcal{P}\text{ partition in }k\text{ parts of }M}{\text{maximize}} \sum_{i=1}^{k} v_i\left(\mathcal{P}^i\right) .$$

Let us show that computing $\mathrm{S}(W, k_1, |\mathcal{Y}_2|)$ can be reformulated as a particular instance of the social welfare maximization problem. First, we can write:

$$\mathrm{S}(W,k_1,k_2) = \underset{e,d_1,d_2}{\text{maximize}} \ \frac{1}{k_1k_2}\sum_{y_1,y_2,i_1,i_2} W(y_1y_2|e(i_1,i_2))\mathbb{1}_{d_1(y_1)=i_1,d_2(y_2)=i_2} , \tag{31}$$

with $e : [k_1]\times[k_2] \to \mathcal{X}$, $d_1 : \mathcal{Y}_1 \to [k_1]$, $d_2 : \mathcal{Y}_2 \to [k_2]$. Indeed, this comes from the fact that a linear program is maximized on its extremal points. Then:

$$\begin{aligned}
\mathrm{S}(W,k_1,k_2) &= \underset{e,d_1,d_2}{\text{maximize}} \ \frac{1}{k_1k_2}\sum_{i_1,i_2}\ \sum_{y_1\in d_1^{-1}(i_1),y_2\in d_2^{-1}(i_2)} W(y_1y_2|e(i_1,i_2)) \\
&= \underset{\mathcal{P}_1,\mathcal{P}_2}{\text{maximize}} \ \frac{1}{k_1k_2}\sum_{i_1,i_2}\max_{e(i_1,i_2)}\ \sum_{y_1\in\mathcal{P}_1^{i_1},y_2\in\mathcal{P}_1^{i_2}} W(y_1y_2|e(i_1,i_2)) \\
&\qquad \text{with } \mathcal{P}_1,\mathcal{P}_2 \text{ partitions of } \mathcal{Y}_1,\mathcal{Y}_2 \text{ respectively} \\
&= \underset{\mathcal{P}_1,\mathcal{P}_2}{\text{maximize}} \ \frac{1}{k_1k_2}\sum_{i_1,i_2}\max_{x}\ \sum_{y_1\in\mathcal{P}_1^{i_1},y_2\in\mathcal{P}_1^{i_2}} W(y_1y_2|x) .
\end{aligned} \tag{32}$$

Thus, in all, we have that:

$$\mathrm{S}(W,k_1,k_2) = \underset{\mathcal{P}_1,\mathcal{P}_2}{\text{maximize}} \ \frac{1}{k_1k_2}\sum_{i_1,i_2} f_W(\mathcal{P}_1^{i_1},\mathcal{P}_1^{i_2}) \text{ with } f_W(S_1,S_2) := \max_x \sum_{y_1\in S_1,y_2\in S_2} W(y_1y_2|x) . \tag{33}$$

We will now focus on the particular case of $k_2 = |\mathcal{Y}_2|$. In that case, it is easy to see that $\mathcal{P}_2 = (\{y_2\})_{y_2\in\mathcal{Y}_2}$ is always an optimal solution. Indeed, for any partition $\mathcal{P}_2$, we have:

$$\begin{aligned}
\frac{1}{k_1k_2}\sum_{i_1,i_2}\max_x \sum_{y_1\in\mathcal{P}_1^{i_1},y_2\in\mathcal{P}_2^{i_2}} W(y_1y_2|x) &\leq \frac{1}{k_1k_2}\sum_{i_1,i_2}\sum_{y_2\in\mathcal{P}_2^{i_2}}\max_x\sum_{y_1\in\mathcal{P}_1^{i_1}} W(y_1y_2|x) \\
&= \frac{1}{k_1k_2}\sum_{i_1}\sum_{y_2\in\mathcal{Y}_2}\max_x\sum_{y_1\in\mathcal{P}_1^{i_1}} W(y_1y_2|x) .
\end{aligned} \tag{34}$$

We can now rewrite our program in the following way:

$$\mathrm{S}^1(W,k_1) := k_1k_2\mathrm{S}(W,k_1,|\mathcal{Y}_2|) = \underset{\mathcal{P}_1}{\text{maximize}} \quad \sum_{i_1}\sum_{y_2}\max_x \sum_{y_1\in\mathcal{P}_1^{i_1}} W(y_1y_2|x)\ . \tag{35}$$

Finally, we can rewrite this as:

$$\mathrm{S}^1(W,k_1) = \underset{\mathcal{P}_1}{\text{maximize}} \quad \sum_{i_1} f_W^1(\mathcal{P}_1^{i_1}) \text{ with } f_W^1(S_1) := \sum_{y_2}\max_x \sum_{y_1\in S_1} W(y_1y_2|x) \tag{36}$$

Hence, computing $\mathrm{S}^1(W,k_1)$ is a particular case of the social welfare maximization problem with a common utility f_W^1 for all k_1 bidders.

5.2 Value Query Hardness

Let us first introduce the value query model. As described in [35, 36], a value query to a utility v asks for the value of some input set $S\subseteq M$, and gets as response $v(S)\in\mathbb{R}_+$. In the value query model, we aim at solving the social welfare maximization problem accessing the data only through value queries to $(v_i)_{i\in[k]}$.

This is more restricted than using any algorithm, but in such a model, it is possible to show unconditional lower bounds on the number of queries needed to solve a given problem within an approximation rate. In the case of the social welfare maximization problem with XOS utility functions, the approximation rate achievable in polynomial time has been proved in [35, 36] to be of the order of $\Theta\left(\frac{1}{\sqrt{m}}\right)$. Specifically, in [35], a $\Omega\left(\frac{1}{m^{\frac{1}{2}}}\right)$-approximation in polynomial time was given, and in [36], it has been shown that any $\Omega\left(\frac{1}{m^{\frac{1}{2}-\varepsilon}}\right)$-approximation for $\varepsilon>0$ requires an exponential number of value queries. We will adapt their proof in the particular case of one common XOS utility function of the form $f_W^1(S_1) := \sum_{y_2}\max_x\sum_{y_1\in S_1} W(y_1y_2|x)$ for some broadcast channel W. But first, let us introduce the definition of XOS functions and prove that f_W^1 is one of those.

Definition 5.1. A linear valuation function (also known as additive) is a set function $a: 2^M\to\mathbb{R}_+$ that assigns a non-negative value to every singleton $\{j\}$ for $j\in M$, and for all $S\subseteq M$ it holds that $a(S)=\sum_{j\in S}a(\{j\})$.

A fractionally sub-additive function (XOS) is a set function $f: 2^M\to\mathbb{R}_+$, for which there is a finite set of linear valuation functions $A=\{a_1,\dots,a_\ell\}$ such that $f(S)=\max_{i\in[\ell]}a_i(S)$ for every $S\subseteq M$.

Remark. Note that the size of A is not bounded in the definition.

Proposition 5.2. f_W^1 is XOS.

Proof.

$$\begin{aligned}
f_W^1(S) &= \sum_{y_2}\max_x \sum_{y_1\in S_1} W(y_1y_2|x) = \max_{\lambda:\mathcal{Y}_2\to\mathcal{X}} a_\lambda(S)\ \text{, where} \\
a_\lambda(S) &= \sum_{y_2}\sum_{y_1\in S} W(y_1y_2|\lambda(y_2)) = \sum_{y_1\in S}\left[\sum_{y_2} W(y_1y_2|m(y_2))\right] \\
&= \sum_{y_1\in S} a_\lambda(\{y_1\})) \text{ with } a_\lambda(\{y_1\})) = \sum_{y_2} W(y_1y_2|\lambda(y_2)) \in\mathbb{R}_+
\end{aligned} \tag{37}$$

So f_W^1 is the maximum of the set of a_λ for $\lambda\in\mathcal{X}^{\mathcal{Y}_1}$, which are linear valuation functions, thus f_W^1 is XOS. $\square$

In order to prove our hardness result, we will need the following version of the Chernoff-Hoeffding bound with negatively associated random variables, which is a weaker notion of independence where tail bounds are still valid (see [42] for further discussion on negatively associated random variables):

Proposition 5.3 (Chernoff-Hoeffding bound)**.** Let $X_1, \dots, X_m$ be negatively associated Bernouilli random variables of parameter p. Then for $0 < \varepsilon \leq \frac{1}{2}$:

$$\mathbb{P}\left(\frac{1}{m}\sum_{i=1}^m X_i > (1+\varepsilon)p\right) \leq e^{-\frac{pm\varepsilon^2}{4}} \ .$$

Proof. Usual proofs of the Chernoff-Hoeffding bound work in the same way with negatively associated variables as pointed out by [42]. So, one obtain as in the original proof [43] that:

$$\mathbb{P}\left(\frac{1}{m}\sum_{i=1}^m X_i > (1+\varepsilon)p\right) \leq e^{-D((1+\varepsilon)p||p)m} \ ,$$

with $D\left(x||y\right) := x\ln\left(\frac{x}{y}\right) + (1-x)\ln\left(\frac{1-x}{1-y}\right)$ the Kullback–Leibler divergence between Bernoulli distributed random variables with parameters x and y. Using the bound $D\left((1+\varepsilon)p||p\right) \geq \frac{\varepsilon^2 p}{4}$ for $0 < \varepsilon < \frac{1}{2}$ concludes the proof. □

Let us now state the value query hardness of approximation of the broadcast channel problem:

Theorem 5.4. In the value query model, for any fixed $\varepsilon > 0$, a $\Omega\left(\frac{1}{m^{\frac{1}{2}-\varepsilon}}\right)$-approximation algorithm computing $\mathrm{S}^1(W, k_1) = k_1 k_2 \mathrm{S}(W, k_1, |\mathcal{Y}_2|)$, with $m = |\mathcal{Y}_1| = k_1^2$ and W a broadcast channel, requires exponentially many value queries to f^1_W.

Remark. As our problem is a particular instance of the social welfare maximization problem with XOS functions, the polynomial-time $\Omega\left(\frac{1}{m^{\frac{1}{2}}}\right)$-approximation from [35] works also here.

Proof. The proof is inspired by Theorem 3.1 of [36]. We will show using probabilistic arguments that any $\Omega\left(\frac{1}{m^{\frac{1}{2}-\varepsilon}}\right)$-approximation algorithm requires an exponential number of value queries. Let us fix a small constant $\delta > 0$. We choose $k_1 \in \mathbb{N}$ as the number of messages (the bidders) and the output space $\mathcal{Y}_1 := [m]$ with $m := k_1^2$ (the items). Then, we choose uniformly at random an equi-partition of $\mathcal{Y}_1$ in k_1 parts of size k_1, which we name $T_1, \dots, T_{k_1}$.

Let us define now $\mathcal{Y}_2 := [m + k_1 + 1]$. We take $\mathcal{X} := \mathcal{Y}_2 = [m + 1 + k_1]$ as well. We can now define our broadcast channel W, with some positive constant C to be fixed later to guarantee that W is a conditional probability distribution. Let us define its value for $y_2 = 1$:

$$W(y_1 1|x) := C \times \begin{cases} m^{2\delta}\mathbb{1}_{y_1 = x} & \text{when } 1 \leq x \leq m \\ \frac{1}{m^{\frac{1}{2}-\delta}} & \text{when } x = m+1 \\ \mathbb{1}_{y_1 \in T_j} & \text{when } 1 \leq j := x - (m+1) \leq k_1 \end{cases}$$

Then, we define other y_2 inputs as translations of $W(y_1 1|x)$. Specifically, we define:

$$W(y_1 y_2|x) := W(y_1 1|t_{y_2 - 1}(x)) \text{ with } t_s(x) := 1 + [(x - 1 + s) \mod (m + k_1 + 1)] \ .$$

All coefficients are nonnegative. So W will be a channel if for all x, $\sum_{y_1, y_2} W(y_1 y_2|x) = 1$. However, one has, for some fixed x_0:

$$\begin{aligned}\sum_{y_1, y_2} W(y_1 y_2|x_0) &= \sum_{y_1}\sum_{y_2} W(y_1 y_2|x_0) = \sum_{y_1}\sum_{y_2} W(y_1 1|t_{y_2-1}(x_0)) = \sum_{y_1}\sum_{x} W(y_1 1|x) \\ &= C\sum_{y_1}\left[\sum_{1\leq i\leq m} m^{2\delta}\mathbb{1}_{y_1 = i} + \frac{1}{m^{\frac{1}{2}-\delta}} + \sum_{1\leq j\leq k_1}\mathbb{1}_{y_1\in T_j}\right] \\ &= C\left[\sum_{1\leq i\leq m} m^{2\delta} + m\times\frac{1}{m^{\frac{1}{2}-\delta}} + \sum_{1\leq j\leq k_1} k_1\right] \\ &= 1 \text{ by choosing } C = \frac{1}{m^{1+2\delta} + m^{\frac{1}{2}+\delta} + m}, \text{ which does not depend on } x_0 \ .\end{aligned} \tag{38}$$

Thus, we have defined a correct instance of our problem. Note that on this instance, we have:

$$
\begin{aligned}
f^1_W(S) &= \sum_{y_2} \max_x \sum_{y_1 \in S} W(y_1 y_2 | x) = \sum_{y_2} \max_x \sum_{y_1 \in S} W(y_1 1 | t_{y_2 - 1}(x)) \\
&= (m + k_1 + 1) \max_x \sum_{y_1 \in S} W(y_1 1 | x) \text{ since } t_{y_2-1} \text{ bijection} \\
&= C(m + k_1 + 1) \times \max \begin{cases} m^{2\delta} |\{i\} \cap S| & \text{for } 1 \leq i \leq m \\ \frac{1}{m^{\frac{1}{2}-\delta}} |S| \\ |T_j \cap S| & \text{for } 1 \leq j \leq k_1 \end{cases}
\end{aligned} \tag{39}
$$

Let us also consider an alternate broadcast channel W', with the only difference that $\mathbb{1}_{y_1 \in T_j}$ is replaced by $\frac{1}{m^{\frac{1}{2}}}$, for $j \in [k_1]$. For that channel, the constant C remains the same (since $\sum_j \sum_{y_1} \mathbb{1}_{y_1 \in T_j} = k_1 \times k_1 = k_1 \times m \times \frac{1}{k_1} = \sum_j \sum_{y_1} \frac{1}{m^{\frac{1}{2}}}$), so we get that:

$$
\begin{aligned}
f^1_{W'}(S) &= C(m + k_1 + 1) \times \max \begin{cases} m^{2\delta} |\{i\} \cap S| & \text{for } 1 \leq i \leq m \\ \frac{1}{m^{\frac{1}{2}-\delta}} |S| \\ \frac{1}{m^{\frac{1}{2}}} |S| & \text{for } 1 \leq j \leq k_1 \end{cases} \\
&= C(m + k_1 + 1) \times \max \begin{cases} m^{2\delta} |\{i\} \cap S| & \text{for } 1 \leq i \leq m \\ \frac{1}{m^{\frac{1}{2}-\delta}} |S| \end{cases}
\end{aligned} \tag{40}
$$

since $\frac{1}{m^{\frac{1}{2}}} |S| \leq \frac{1}{m^{\frac{1}{2}-\delta}} |S|$. We will prove that it takes an exponential number of value queries to distinguish between f^1_W and $f^1_{W'}$. On the one hand, one can easily show that the maximum value of $\frac{\mathrm{S}^1(W', k_1)}{C(m+k_1+1)}$ is $(k_1 - 1)m^{2\delta} + \frac{1}{m^{\frac{1}{2}-\delta}}(m - (k_1 - 1)) = O(m^{\frac{1}{2}+2\delta})$, obtained taking $(k_1 - 1)$ singletons as the first components of the partition (the bidders), giving the rest of $\mathcal{Y}_1$ (the items) to the last. On the other hand, the maximum value of $\frac{\mathrm{S}^1(W, k_1)}{C(m+k_1+1)}$ is $k_1 \times k_1 = m$, obtained with the partition $T_1, \dots, T_{k_1}$. The fact that it requires an exponential number of value queries to distinguish between the two situations will imply that one cannot get an approximation rate better than $\Omega\left(\frac{1}{m^{\frac{1}{2}-2\delta}}\right)$ in less than an exponential number of value queries.

We will now prove that distinguishing between f^1_W and $f^1_{W'}$ requires an exponential number of value queries. We define $v(S) := \frac{f^1_W(S)}{C(m+k_1+1)}$ and $v'(S) := \frac{f^1_{W'}(S)}{C(m+k_1+1)}$, so distinguishing between v and v' is the same as distinguishing between f^1_W and $f^1_{W'}$. We have always $v(\emptyset) = v'(\emptyset) = 0$, so we do not need to consider empty sets.

Let us fix some non-empty set $S \subseteq [m]$. Let us define the random boolean variables $X^i_j := \mathbb{1}_{i \in T_j}$ for $j \in [k_1]$ and $i \in [m]$. By construction of the random equi-partition $T_1, \dots, T_{k_1}$, $(X^i_j)_{i \in [m]}$ is a permutation distribution (see Definition 2.10 of [44]) of $(0, \dots, 0, 1, \dots, 1)$ with $m - k_1$ zeros and k_1 ones, each X^i_j following a Bernouilli law of parameter $p := \frac{1}{k_1}$. Thus it is negatively associated (by Theorem 2.11 of [44]), and the sub-family $(X^i_j)_{i \in S}$ is negatively associated as well. Note in particular that $|T_j \cap S| = \sum_{i \in S} X^i_j$ is a sum of negatively associated Bernouilli variables of the same parameter p, so Chernoff-Hoeffding bound as depicted in Proposition 5.3 holds.

Let us first assume that S is of size $0 < |S| \leq m^{\frac{1}{2}+\delta}$. Then, we have that $\frac{1}{m^{\frac{1}{2}-\delta}} |S| \leq m^{2\delta}$, so we get that $v'(S) = m^{2\delta}$. On the other hand, we have that:

$$
v(S) = \max \begin{cases} m^{2\delta} \\ |T_j \cap S| \text{ for } 1 \leq j \leq k_1 \end{cases} \tag{41}
$$

Thus, $v(S)$ is different from $v'(S)$ if and only if $\exists j \in k_1, |T_j \cap S| > m^{2\delta}$. But, we have:

$$\begin{aligned}
&\mathbb{P}\left(\exists j \in k_1, |T_j \cap S| > m^{2\delta}\right) \leq \sum_{j\in[k_1]} \mathbb{P}\left(|T_j \cap S| > m^{2\delta}\right) \text{ by union bound} \\
&= \sum_{j\in[k_1]} \mathbb{P}\left(\sum_{i\in S} X_j^i > m^{2\delta}\right) = \sum_{j\in[k_1]} \mathbb{P}\left(\frac{1}{|S|}\sum_{i\in S} X_j^i > \left(1 + \frac{\frac{m^{2\delta}}{|S|} - 1}{p}\right) p\right) \\
&\leq \sum_{j\in[k_1]} \mathbb{P}\left(\frac{1}{|S|}\sum_{i\in S} X_j^i > \left(1 + \frac{\frac{m^{2\delta}}{|S|}}{p}\right) p\right) \\
&\leq \sum_{j\in[k_1]} \exp\left(-\frac{p|S|}{4}\left(\frac{m^{2\delta}}{p|S|}\right)^2\right) \text{ by Chernoff-Hoeffding bound as depicted in Proposition 5.3} \\
&= \sum_{j\in[k_1]} \exp\left(-\frac{1}{4p|S|} m^{4\delta}\right) \leq \sum_{j\in[k_1]} \exp\left(-\frac{m^{-\delta}}{4} m^{4\delta}\right) \text{ since } \frac{1}{p|S|} = \frac{k_1}{|S|} \geq m^{-\delta} \\
&= m^{\frac{1}{2}} e^{-\frac{m^{3\delta}}{4}} \;.
\end{aligned} \tag{42}$$

Thus, this event occurs with exponentially small probability (on the choice of the partition $T_1, \ldots, T_{k_1}$).

Let us now study the case of S of size $|S| > m^{\frac{1}{2}+\delta}$. Then, we have that $\frac{1}{m^{\frac{1}{2}-\delta}}|S| > m^{2\delta}$, so we get that $v'(S) = \frac{1}{m^{\frac{1}{2}-\delta}}|S|$. On the other hand, we have that:

$$v(S) = \max \begin{cases} \frac{1}{m^{\frac{1}{2}-\delta}}|S| \\ |T_j \cap S| \text{ for } 1 \leq j \leq k_1 \end{cases} \tag{43}$$

Thus, $v(S)$ is different from $v'(S)$ if and only if $\exists j \in [k_1], |T_j \cap S| > \frac{1}{m^{\frac{1}{2}-\delta}}|S|$. But, we have:

$$\begin{aligned}
&\mathbb{P}\left(\exists j \in [k_1], |T_j \cap S| > \frac{1}{m^{\frac{1}{2}-\delta}}|S|\right) \leq \sum_{j\in[k_1]} \mathbb{P}\left(|T_j \cap S| > \frac{1}{m^{\frac{1}{2}-\delta}}|S|\right) \text{ by union bound} \\
&= \sum_{j\in[k_1]} \mathbb{P}\left(\sum_{i\in S} X_j^i > \frac{1}{m^{\frac{1}{2}-\delta}}|S|\right) = \sum_{j\in[k_1]} \mathbb{P}\left(\frac{1}{|S|}\sum_{i\in S} X_j^i > \left(1 + \frac{\frac{1}{|S|m^{\frac{1}{2}-\delta}}|S| - 1}{p}\right) p\right) \\
&\leq \sum_{j\in[k_1]} \mathbb{P}\left(\frac{1}{|S|}\sum_{i\in S} X_j^i > \left(1 + \frac{\frac{1}{m^{\frac{1}{2}-\delta}}}{p}\right) p\right) \\
&= \sum_{j\in[k_1]} \mathbb{P}\left(\frac{1}{|S|}\sum_{i\in S} X_j^i > \left(1 + m^{\delta}\right) p\right) \text{ since } p = \frac{1}{k_1} = \frac{1}{m^{\frac{1}{2}}} \\
&\leq \sum_{j\in[k_1]} \exp\left(-\frac{p|S|}{4} m^{2\delta}\right) \text{ by Chernoff-Hoeffding bound as depicted in Proposition 5.3} \\
&\leq \sum_{j\in[k_1]} \exp\left(-\frac{m^{\delta}}{4} m^{2\delta}\right) \text{ since } p|S| = \frac{|S|}{k_1} \geq m^{\delta} \\
&= m^{\frac{1}{2}} e^{-\frac{m^{3\delta}}{4}} \;.
\end{aligned} \tag{44}$$

Thus, this event occurs with exponentially small probability as well. We have then that for all set S, $\mathbb{P}\left(v(S) \neq v'(S)\right) \leq p_{\text{leak}} := m^{\frac{1}{2}} e^{-\frac{m^{3\delta}}{4}}$, which is an exponentially small bound that does not depend on S.

Hence, for every set S, only with exponentially small probability p_{leak} can one distinguish between v and v'. For some fixed algorithm $\mathcal{A}$, let us consider the sequence L of queries made by $\mathcal{A}$ before it is able to distinguish between v and v': $L := (S_1, \dots, S_n)$, with $v(S_i) = v'(S_i)$ for $i \in [n]$ and $v'(S_{n+1}) > v(S_{n+1})$. L is independent from $T_1, \dots T_{k_1}$ as no information from this partition is leaked before S_{n+1}. Thus, for such an algorithm to be correct, it should work for any equi-partition $T_1, \dots T_{k_1}$. We have:

$$\mathbb{P}\left(\exists i \in [n] : v(S_i) \neq v'(S_i)\right) \leq \sum_{i=1}^{n} \mathbb{P}\left(v(S_i) \neq v'(S_i)\right) = np_{\text{leak}} \text{ by union bound.}$$

In particular, this implies that:

$$\mathbb{P}\left(\forall i \in [n] : v(S_i) = v'(S_i)\right) \geq 1 - np_{\text{leak}} \ .$$

So, if $1 - np_{\text{leak}} > 0$, ie. $n < \frac{1}{p_{\text{leak}}}$, then there exists some equi-partition $T_1, \dots T_{k_1}$ such that our algorithm outputs a sequence L of queries of length n before being able to distinguish between v and v'. In particular, we can take $n = \frac{1}{2p_{\text{leak}}}$ so that L is of exponential size. Hence, for any algorithm $\mathcal{A}$, there exists some equi-partition $T_1, \dots T_{k_1}$ such that $\mathcal{A}$ needs an exponential number of value queries to distinguish between v and v'. This concludes the proof of the theorem for any deterministic algorithm.

Finally, the hardness result holds also for randomized algorithms. Indeed, let us call $\mathcal{A}_s$, the running algorithm conditioned on its random bits being s. $\mathcal{A}_s$ is deterministic so the previous proof holds: with high probability p, the sequence of $\lfloor \frac{1-p}{p_{leak}} \rfloor$ queries does not reveal anything to distinguish between v and v', although it is of exponential size in m. Then, averaging over all its random bitstrings, the same result holds, as p_{leak} is independent from the equi-partition $T_1, \dots, T_{k_1}$. □

5.3 Limitations of the Model

The main weakness of the previous result is that it highly relies on the restriction that one has access to the data only through value queries. Indeed, if one has access to the full data, it is possible to read the partition $T_1, \dots, T_{k_1}$ which gives the optimal solution directly. This weakness comes from the fact that our utility function f_W^1 can be described by polynomial-size data, as it is characterized by a broadcast channel W, whereas one usually consider any XOS functions, in particular with some having inherently an exponential-size defining set of linear valuation functions.

On the other hand, one can also remark that when f_W^1 is written as a maximum of linear valuation functions, then that defining set of linear valuation functions is of exponential size. Hence, we think that it is still relevant to study the value query complexity of such a family of functions, as it is not clear how one could recover the partition in polynomial time from this exponential-size set of linear valuations without any additional information.

6 Conclusion

In this work, we have studied several algorithmic aspects and non-signaling assisted capacity regions of broadcast channels. We have shown that when non-signaling assistance is shared only between the decoders, the capacity region does not change. For the class of deterministic broadcast channels, we have described a $(1 - e^{-1})^2$-approximation algorithm running in polynomial time, and we have shown that the capacity region for that class is the same with or without non-signaling assistance. Finally, we have shown that in the value query model, we cannot achieve a better approximation ratio than $\Omega\left(\frac{1}{\sqrt{m}}\right)$ in polynomial time for the general broadcast channel coding problem, with m the size of one of the outputs of the channel.

Our results suggest that non-signaling assistance could improve the capacity region of general broadcast channels, which is left as a major open question. An intermediate result would be to show that it is NP-hard to approximate the broadcast channel coding problem within any constant, strengthening our hardness result without relying on the value query model. Finally, one could also try to develop approximations algorithms for other sub-classes of broadcast channels, such as semi-deterministic or degraded ones. This could be a crucial step towards showing that the capacity region for those classes is the same with or without non-signaling assistance.

Acknowledgements

We would like to thank Stéphan Thomassé and Siddharth Barman for discussions respectively on graph algorithms and social welfare problems.

References

[1] Siddharth Barman and Omar Fawzi. Algorithmic aspects of optimal channel coding. *IEEE Trans. Inf. Theory*, 64(2):1038–1045, 2018. doi:10.1109/TIT.2017.2696963.

[2] Thomas M. Cover. Broadcast channels. *IEEE Trans. Inf. Theory*, 18(1):2–14, 1972. doi:10.1109/TIT.1972.1054727.

[3] Claude E. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27(3):379–423, 1948. doi:10.1002/j.1538-7305.1948.tb01338.x.

[4] Henry Herng-Jiunn Liao. Multiple access channels (Ph.D. Thesis abstr.). *IEEE Trans. Inf. Theory*, 19(2):253, 1973. doi:10.1109/TIT.1973.1054960.

[5] Rudolf Ahlswede. Multi-way communication channels. In *2nd International Symposium on Information Theory, Tsahkadsor, Armenia, USSR, September 2-8, 1971*, pages 23–52. Publishing House of the Hungarian Academy of Science, 1973.

[6] Patrick P. Bergmans. Random coding theorem for broadcast channels with degraded components. *IEEE Trans. Inf. Theory*, 19(2):197–207, 1973. doi:10.1109/TIT.1973.1054980.

[7] R.G. Gallager. Capacity and coding for degraded broadcast channels. *Problemy Peredachi Informatsii*, 10(3):3–14, 1974.

[8] Rudolf Ahlswede and János Körner. Source coding with side information and a converse for degraded broadcast channels. *IEEE Trans. Inf. Theory*, 21(6):629–637, 1975. doi:10.1109/TIT.1975.1055469.

[9] Katalin Marton. The capacity region of deterministic broadcast channels. In *Trans. Int. Symp. Inform. Theory*, pages 243–248, 1977.

[10] Mark Semenovich Pinsker. Capacity of noiseless broadcast channel. *Problemy Peredachi Informatsii*, 14(2):28–34, 1978.

[11] Sergei Izrail'evich Gel'fand and Mark Semenovich Pinsker. Capacity of a broadcast channel with one deterministic component. *Problemy Peredachi Informatsii*, 16(1):24–34, 1980.

[12] Thomas M. Cover. An achievable rate region for the broadcast channel. *IEEE Transactions on Information Theory*, 21(4):399–404, 1975. doi:10.1109/TIT.1975.1055418.

[13] Edward C. van der Meulen. Random coding theorems for the general discrete memoryless broadcast channel. *IEEE Trans. Inf. Theory*, 21(2):180–190, 1975. doi:10.1109/TIT.1975.1055347.

[14] Katalin Marton. A coding theorem for the discrete memoryless broadcast channel. *IEEE Trans. Inf. Theory*, 25(3):306–311, 1979. doi:10.1109/TIT.1979.1056046.

[15] Hiroshi Sato. An outer bound to the capacity region of broadcast channels (corresp.). *IEEE Trans. Inf. Theory*, 24(3):374–377, 1978. doi:10.1109/TIT.1978.1055883.

[16] Chandra Nair and Abbas El Gamal. An outer bound to the capacity region of the broadcast channel. *IEEE Trans. Inf. Theory*, 53(1):350–355, 2007. doi:10.1109/TIT.2006.887492.

[17] Amin Gohari and Chandra Nair. New outer bounds for the two-receiver broadcast channel. In *IEEE International Symposium on Information Theory, ISIT 2020, Los Angeles, CA, USA, June 21-26, 2020*, pages 1492–1497. IEEE, 2020. doi:10.1109/ISIT44484.2020.9174003.

[18] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. Bell nonlocality. *Rev. Mod. Phys.*, 86:419–478, Apr 2014. doi:10.1103/RevModPhys.86.419.

[19] R. Prevedel, Y. Lu, W. Matthews, R. Kaltenbaek, and K. J. Resch. Entanglement-enhanced classical communication over a noisy classical channel. *Phys. Rev. Lett.*, 106:110505, Mar 2011. doi:10.1103/PhysRevLett.106.110505.

[20] Charles H Bennett, Peter W Shor, John A Smolin, and Ashish V Thapliyal. Entanglement-assisted classical capacity of noisy quantum channels. *Physical Review Letters*, 83(15):3081, 1999. doi:10.1103/PhysRevLett.83.3081.

[21] William Matthews. A linear program for the finite block length converse of Polyanskiy-Poor-Verdú via nonsignaling codes. *IEEE Trans. Inf. Theory*, 58(12):7036–7044, 2012. doi:10.1109/TIT.2012.2210695.

[22] Yihui Quek and Peter W. Shor. Quantum and superquantum enhancements to two-sender, two-receiver channels. *Phys. Rev. A*, 95:052329, May 2017. doi:10.1103/PhysRevA.95.052329.

[23] Felix Leditzky, Mohammad A Alhejji, Joshua Levin, and Graeme Smith. Playing games with multiple access channels. *Nature communications*, 11(1):1–5, 2020. doi:10.1038/s41467-020-15240-w.

[24] Akshay Seshadri, Felix Leditzky, Vikesh Siddhu, and Graeme Smith. On the separation of correlation-assisted sum capacities of multiple access channels. In *IEEE International Symposium on Information Theory, ISIT 2022, Espoo, Finland, June 26 - July 1, 2022*, pages 2756–2761. IEEE, 2022. doi:10.1109/ISIT50566.2022.9834459.

[25] Omar Fawzi and Paul Fermé. Beating the sum-rate capacity of the binary adder channel with non-signaling correlations. In *IEEE International Symposium on Information Theory, ISIT 2022, Espoo, Finland, June 26 - July 1, 2022*, pages 2750–2755. IEEE, 2022. doi:10.1109/ISIT50566.2022.9834699.

[26] Uzi Pereg, Christian Deppe, and Holger Boche. The multiple-access channel with entangled transmitters. *CoRR*, abs/2303.10456, 2023. arXiv:2303.10456, doi:10.48550/arXiv.2303.10456.

[27] Uzi Pereg, Christian Deppe, and Holger Boche. Quantum broadcast channels with cooperating decoders: An information-theoretic perspective on quantum repeaters. In *IEEE International Symposium on Information Theory, ISIT 2021, Melbourne, Australia, July 12-20, 2021*, pages 772–777. IEEE, 2021. doi:10.1109/ISIT45174.2021.9518284.

[28] Siddharth Barman, Omar Fawzi, Suprovat Ghoshal, and Emirhan Gürpinar. Tight approximation bounds for maximum multi-coverage. In Daniel Bienstock and Giacomo Zambelli, editors, *Integer Programming and Combinatorial Optimization - 21st International Conference, IPCO 2020, London, UK, June 8-10, 2020, Proceedings*, volume 12125 of *Lecture Notes in Computer Science*, pages 66–77. Springer, 2020. doi:10.1007/978-3-030-45771-6_6.

[29] Szymon Dudycz, Pasin Manurangsi, Jan Marcinkowski, and Krzysztof Sornat. Tight approximation for proportional approval voting. In Christian Bessiere, editor, *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI 2020*, pages 276–282. ijcai.org, 2020. doi:10.24963/ijcai.2020/39.

[30] Uriel Feige, Guy Kortsarz, and David Peleg. The dense k-subgraph problem. *Algorithmica*, 29(3):410–421, 2001. doi:10.1007/s004530010050.

[31] Noga Alon, Sanjeev Arora, Rajsekar Manokaran, Dana Moshkovitz, and Omri Weinstein. Inapproximability of densest κ-subgraph from average case hardness. *Unpublished manuscript*, 1, 2011. URL: https://www.tau.ac.il/~nogaa/PDFS/dks8.pdf.

[32] Steven J. Brams and Alan D. Taylor. *Fair division - from cake-cutting to dispute resolution.* Cambridge University Press, 1996.

[33] Hervé Moulin. *Fair division and collective welfare.* MIT Press, 2003.

[34] Liad Blumrosen and Noam Nisan. On the computational power of iterative auctions. In John Riedl, Michael J. Kearns, and Michael K. Reiter, editors, *Proceedings 6th ACM Conference on Electronic Commerce (EC-2005), Vancouver, BC, Canada, June 5-8, 2005*, pages 29–43. ACM, 2005. doi:10.1145/1064009.1064013.

[35] Shahar Dobzinski and Michael Schapira. An improved approximation algorithm for combinatorial auctions with submodular bidders. In *Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2006, Miami, Florida, USA, January 22-26, 2006*, pages 1064–1073. ACM Press, 2006. URL: http://dl.acm.org/citation.cfm?id=1109557.1109675.

[36] Vahab S. Mirrokni, Michael Schapira, and Jan Vondrák. Tight information-theoretic lower bounds for welfare maximization in combinatorial auctions. In Lance Fortnow, John Riedl, and Tuomas Sandholm, editors, *Proceedings 9th ACM Conference on Electronic Commerce (EC-2008), Chicago, IL, USA, June 8-12, 2008*, pages 70–77. ACM, 2008. doi:10.1145/1386790.1386805.

[37] F.M.J. Willems. The maximal-error and average-error capacity region of the broadcast channel are identical : A direct proof. *Problems of Control and Information Theory*, 19(4):339–347, December 1990. URL: https://research.tue.nl/nl/publications/the-maximal-error-and-average-error-capacity-region-of-the-broadc.

[38] Bernd Gärtner and Jirí Matousek. *Understanding and using linear programming*. Universitext. Springer, 2007. doi:10.1007/978-3-540-30717-4.

[39] M. R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.

[40] Jan Vondrák. Optimal approximation for the submodular welfare problem in the value oracle model. In Cynthia Dwork, editor, *Proceedings of the 40th Annual ACM Symposium on Theory of Computing, Victoria, British Columbia, Canada, May 17-20, 2008*, pages 67–74. ACM, 2008. doi:10.1145/1374376.1374389.

[41] Siddharth Barman, Omar Fawzi, and Paul Fermé. Tight approximation guarantees for concave coverage problems. In Markus Bläser and Benjamin Monmege, editors, *38th International Symposium on Theoretical Aspects of Computer Science, STACS 2021, March 16-19, 2021, Saarbrücken, Germany (Virtual Conference)*, volume 187 of *LIPIcs*, pages 9:1–9:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. doi:10.4230/LIPIcs.STACS.2021.9.

[42] Devdatt P. Dubhashi and Desh Ranjan. Balls and bins: A study in negative dependence. *Random Struct. Algorithms*, 13(2):99–124, 1998. doi:10.7146/brics.v3i25.20006.

[43] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963. doi:10.1080/01621459.1963.10500830.

[44] Kumar Joag-Dev and Frank Proschan. Negative association of random variables with applications. *The Annals of Statistics*, pages 286–295, 1983. doi:10.1214/aos/1176346079.