Unconditionally Secure Commitment Problem

Project-II report submitted to

Indian Institute of Technology Kharagpur

in partial fulfilment for the award of the degree of

Bachelor of Technology

in

Electronics and Electrical Communication Engineering

by

Manideep Mamindlapally

(17EC34003)

Under the supervision of

Professor Amitalok Jayant Budkuley

Department of Electronics and Electrical Communication Engineering

Indian Institute of Technology Kharagpur

Autumn Semester, 2020-21

May 2, 2021

DEPARTMENT OF ELECTRONICS AND ELECTRICAL COMMUNICATION ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR

KHARAGPUR - 721302, INDIA

CERTIFICATE

This is to certify that the project report entitled "**Unconditionally Secure Commitment Problem**" submitted by **Manideep Mamindlapally** (Roll No. 17EC34003) to Indian Institute of Technology Kharagpur towards partial fulfilment of requirements for the award of degree of Bachelor of Technology in Electronics and Electrical Communication Engineering is a record of bona fide work carried out by him under my supervision and guidance during Autumn Semester, 2020-21.

Date: May 2, 2021

Place: Kharagpur

Professor Amitalok Jayant Budkuley
Department of Electronics and Electrical
Communication Engineering
Indian Institute of Technology Kharagpur
Kharagpur - 721302, India

Abstract

Name of the student: **Manideep Mamindlapally** Roll No: **17EC34003**

Degree for which submitted: **Bachelor of Technology**

Department: **Department of Electronics and Electrical Communication Engineering**

Thesis title: **Unconditionally Secure Commitment Problem**

Thesis supervisor: **Professor Amitalok Jayant Budkuley**

Month and year of thesis submission: **May 2, 2021**

Commitment is a widely studied cryptographic primitive, where two mutually distrustful parties, say Alice and Bob, interact over two phases of a protocol, viz., *commit phase* followed by *reveal phase*, to achieve *commitment* on a bit string available to Alice. Commitment (over the string) is said to occur if (i) Alice commits to the string which remains securely hidden from Bob at the end of the commit phase involving Alice's transmission to Bob, and (ii) Alice reveals a string to Bob and Bob is able to successfully detect whether the string is the committed one or not. When Alice and Bob are computationally unbounded, i.e., under the *information-theoretic* setting, it is well known that even a single bit commitment is impossible when the channel available to Alice and Bob is noiseless. Noisy channels, however, offer the potential of non-zero commitment rate, and thus, are a valuable resource.

First, we study information-theoretically secure commitment over general noisy discrete memoryless channels (DMCs). The largest commitment throughput over noisy channels is called the *commitment capacity* or simply *capacity.* We completely characterize via a single-letter expression, the commitment capacity of DMCs under general cost constraints; this generalizes the previously known result in the absence of such cost constraints. We show that cost constrained commitment capacity of any given DMC can significantly differ from its unconstrained value. We also present a *dual* capacity characterization in terms of output distributions. Interestingly, we show that every input distribution achieving the capacity results in the same output distribution; the latter is the unique optimizer of our dual capacity expression.

We then also define and study a special noisy channel called the *compound-binary symmetric channel (compound-BSC).* It models the scenario when a BSC is imprecisely known or poorly characterised. A compound-BSC is a BSC whose transition probability is fixed but unknown to either party; the set of potential values which this transition probability can take, though, is known to both parties *a priori.* We provide an optimal, computationally-efficient scheme for our achievability, and we derive a converse for general alphabet compound DMCs, which is then specialized for compound-BSCs.

Acknowledgements

I am grateful to my supervisor Prof. Amitalok J Budkuley(IIT Kharagpur) for providing this project opportunity. He has been very helpful throughout by clarifying doubts, conducting weekly meetings online and in many other ways. I would also thank Anuj K Yadav(UG student IIT Patna) and Prof. Manoj Mishra(NISER Bhubaneshwar) who have also participated actively in suggesting ideas which have been useful to the project and are also co-authors of our submitted conference papers.

Contents

Chapter 1

Introduction

1.1 The Rock Paper Scissor Commitment Problem

Two friends, Alice and Bob, wish to settle an argument by playing the age old game Rock-Paper-Scissors. To observe social distancing norms, they are both in their respective houses and talking to each other over the phone. Alice and Bob decide to simultaneously yell out their chosen hand. A bad phone signal, however, often causes unpredictable delays in the call; so that when Bob yells ”paper” and hears Alice yell back ”scissors” a second or two later, he cannot tell whether Alice's voice had been delayed by the poor connection, or whether she deliberately spoke her choice only after hearing Bob's - making it seemingly impossible to fairly play the game.

A solution to this problem would be to have Alice write down her choice on a piece of paper, lock it in a safe and then courier the safe over to Bob. Bob now knows Alice cannot change her mind, so Bob tells Alice his chosen word. Once he does that, Alice tells Bob which word she chose and they decide the winner. To verify whether or not Alice was telling the truth, Bob has Alice send him the key to the safe via speed-post; after which he opens the safe and ensures that Alice has not cheated. Alice, of course, is confident that Bob couldn't have taken a premature peek at her slip of paper, because the key to the safe remained with her while Bob made his choice.

This sort of protocol is called a **Commitment Protocol**, where one party (say, Alice) *commits* to sharing a message with the other (Bob). It comes with two guarantees: that Bob will not be able to see this message until Alice reveals it to him (*concealing* from Bob), and that Alice will not be able to reveal a different message than the one she committed (*binding* on Alice).

Though you may not often face the situation of having to play Rock-Paper-Scissors over a faulty phone line with your friend, being able to realise the bit commitment protocol arms you with a fundamental primitive in cryptography; you can build up from this protocol and realise widespread practical applications like sealed-bid auctions (Nojoumian and Stinson, 2010), coin flipping (Naor, 1991a), zero knowledge proofs (Brassard et al., 1988; Goldreich et al., 1991), contract signing (Even et al., 1985) and secure multiparty computation (Chaum et al., 1988; Goldreich et al., 1987).

1.2 State of the Art

Commitment was first studied by Blum (Blum, 1983). The earliest commitment schemes were realised classically in a *computationally secure* sense over noiseless channels with computational limitations on the two parties. More precisely, these schemes were mostly either *computationally hiding*(Naor, 1991b; Ostrovsky et al., 1992) (resp. *computationally binding* (Blum, 1983; Brassard et al., 1988; Halevi, 1999; Halevi and Micali, 1996)), where Bob (resp. Alice) is computationally bounded and the protocol is computationally secure from Alice's (resp. Bob's) point of view. However, in the absence of such computational limitations on the two parties, it was shown that even one-bit commitment is impossible over noiseless channels (cf. (Damgård et al., 1999) for a simple proof).

Wyner's seminal work on wire-tap channels (Wyner, 1975) first explored the potential of noisy channels for security; he showed that a noisy random channel can be a great asset to realise various *information-theoretically secure* cryptographic protocols (where security is guaranteed against computationally unbounded parties). Wyner's results have subsequently spawned a wide area of research on information-theoretic security; his results have been extended and strengthened in a multitude of works. See, for instance, (Bloch and Barros, 2011; Csiszár and Korner, 1978), and the references therein. These results were then further extended for secret-key distillation by (Maurer, 1993), who proved that the ability of generating a secret key could be improved through public communication, followed by (Ahlswede and Csiszár, 1993), and others.

Specific to commitment, Crépeau and Kilian (Crépeau and Kilian, 1988) first studied commitment (along with another closely related cryptographic primitive called oblivious transfer, see (Mishra et al., 2017) for more details) over noisy channels. These results were subsequently improved in (Crépeau, 1997; Damgård et al., 1999). Winter *et al.* characterized the largest throughput possible or the *commitment capacity* over general DMCs (Winter et al., 2003). The commitment capacity over continuous Gaussian channels was explored in (Nascimento et al., 2008), where Nascimento showed that the capacity is infinite.

In (Damgård et al., 1999), Damgård et al. proposed the *unfair noisy channel (UNC)*, which is a channel with two parameters γ, δ, where $0 < \gamma < \delta < 1/2$. Damgård et al.'s UNC$(\gamma, \delta)$ is essentially a BSC where the transition probability p can belong to $[\gamma, \delta]$ interval. Damgård et al. (Damgård et al., 1999) characterized the threshold (in terms of the parameters γ and δ) for positive commitment throughput; the commitment capacity of the UNC was subsequently characterized recently in (Crépeau et al., 2020). Khurana et al. (Khurana et al., 2016) recently studied another closely related variant of the compound-BSC called the *elastic* channel. The elastic channel with parameters γ, δ, where $0 < \gamma < \delta < 1/2$, can be seen as a UNC under *relaxation*, where only a dishonest receiver is allowed to know and control the transition probability p of the BSC in the range $[\gamma, \delta]$. Furthermore, when both parties are honest, the channel specializes to a classic BSC(δ).

1.3 Our contribution

Our contributions in this project are two fold. The first of them focuses on the effect of channel input costs on the commitment capacity. Given that noisy channels are an important resource for realizing commitment, it is pertinent to understand this. We build upon the result in (Winter et al., 2003) for DMCs under unconstrained input costs and study the *commitment capacity* (henceforth also called *capacity*) when non-trivial costs are incurred for using specific input symbols on the channel and there is an overall budget on the total transmission cost. These are some of the contributions of the project pertaining to this const constraint characterisation (Discussed in Chapter 3).

- We completely characterize in Theorem 3.8 the commitment capacity of DMCs under general input constraints. We also specify conditions (in terms of cost functions and cost constraints) which result in zero capacity (cf. Section 3.2).

- In Theorem 3.16, we present a *dual* capacity characterization in terms of channel output distributions. Interestingly, we show that every capacity achieving input distribution results in the same channel output distribution. Thus, the capacity achieving output distribution in the dual is unique.

- We determine the capacity of a binary symmetric channel under input Hamming weight constraints (cf. Example 3.4.2) using both our characterizations.

The second of our contributions is realising commitment over a special Discrete Memoryless Channel, the compound-binary symmetric channel (compound-BSC). We do that in Chapter 4. For the purpose of this study we do not consider any cost constraints. The focus would be on realising a maximum rate commitment protocol on the compound BSC channel.
In the classic BSC(p), where $p \in (0, 1/2)$[1] is a fixed value known to both parties, has been widely studied for commitment (as well as other cryptographic primitives); the commitment capacity of the BSC(p) is given by $H(p)$ (this result also follows from (Winter et al., 2003), where commitment capacity of general DMCs was characterized). A compound-BSC is a BSC where the transition probability p is fixed but unknown to the parties; what the parties do know, however, is that p takes values from a set $\mathcal{S}$. [2] The compound-BSC specializes to a classic BSC when $\mathcal{S}$ is a singleton.

The compound-BSC sits at the heart of several channels models of wide interest; surprisingly, however, and to the best of our knowledge, compound-BSCs have not been studied in literature previously. Our work aims to fill in this gap and initiate a systematic study of compound-BSCs for commitment. We formally define compound-BSCs in Chapter 2 and completely characterize their commitment capacity in Chapter 5. The following are the key contributions:

- A formal definition of a compound-BSC channel.
- A converse proved for a general DMC, and subsequently, specialized for the compound-BSC.
- An optimal, computationally-efficient commitment scheme using the compound-BSC.

[1] Given the symmetry of the BSC, we consider the range $p \leq 1/2$ without loss of generality. As commitment is impossible over a noiseless channel (Blum, 1983) and when the channel input and output are independent, we rule out the two cases when $p = 0$ and $p = 1/2$.

[2] In this work, we assume that $\mathcal{S}$ is bounded, although this can be relaxed by approximation techniques (cf. (El Gamal and Kim, 2011)).

1.4 Organisation

We introduce certain preliminaries and Notation in Chapter 2. Chapter 3 formulates the commitment problem over a general DMC channel and states and proves the primal and dual commitment capacity expressions for the same. In Chapter 4 we define a compound BSC channel and propose a capacity achieving commitment scheme. Chapter 5 concludes the work with some important takeaways.

Chapter 2

Preliminaries and Notation

2.1 Notations

We denote random variables by upper case letters (eg. X), the values they take by lower case letters (eg., x), and their alphabets by calligraphic letters (eg. $\mathcal{X}$). Unless stated otherwise, all sets are assumed to be finite. We denote random vectors and the concomitant values they take by boldface letters (e.g., $\mathbf{X} = (X_1, X_2, \cdots, X_n)$, $\mathbf{x} = (x_1, x_2, \cdots, x_n)$, resp.). The set of real numbers, non-negative real numbers and real vectors (of length n) are denoted by $\mathbb{R}$, $\mathbb{R}_+$, and $\mathbb{R}^n$ respectively. The set of natural numbers is denoted by $\mathbb{N}$. For $a \in \mathbb{N}$, let $[a] := \{1, 2, \cdots, a\}$. We denote the Hamming distance between two vectors, say $\mathbf{x}, \mathbf{x}' \in \mathcal{X}^n$ by $d_H(\mathbf{x}, \mathbf{x}')$.

$$d_H(\mathbf{x}, \mathbf{x}') = \sum_{i=1}^{n} \mathbf{1}_{\{x_i \neq x_i'\}}$$

where $\mathbf{1}_A$ denotes the indicator of A. Let $\mathcal{P}(\mathcal{X})$ denote the simplex of probability distributions on set $\mathcal{X}$. Let $\mathcal{P}(\mathcal{X}|\mathcal{Y})$ denote the set of all conditional probability distributions induced by random variable $X \in \mathcal{X}$ conditioned on events generated by random variable $Y \in \mathcal{Y}$. We denote by P_X, $P_{X|Y}$ and $P_{X,Y}$ the probability distribution of random variable $X \in \mathcal{X}$, the conditional probability distribution induced by random variable $X \in \mathcal{X}$ conditioned on events generated by random variable $Y \in \mathcal{Y}$ and the joint probability distribution on the pair of random variables $(X, Y) \in \mathcal{X} \times \mathcal{Y}$ respectively. For the latter, we denote the marginal distribution on random variable X by $[P_{X,Y}]_X$. Given P_X, $P_X^{(n)}$ denotes the n-fold memoryless extension of P_X. Let $\mathbb{P}(A)$ denote the probability of event A. Deterministic and random functions will be denoted by lower case letters (eg. f) and by upper case letters (e.g., F) respectively. Given $P_X, Q_X \in \mathcal{P}(\mathcal{X})$, let $D(P_X||Q_X)$ denote the KL divergence between P_X and Q_X and let $||P_X - Q_X||_1$ denote the ℓ_1 distance between P_X and Q_X. Given P_X and $\delta > 0$, let $\mathcal{T}_\delta^{(n)}(P_X) = \{\mathbf{x} : |T_{\mathbf{x}}(x) - P_X(x)| \leq \delta, \, \forall x \in \mathcal{X}\}$ denote the set of typical $\mathbf{x}$-sequences, where $T_{\mathbf{x}}$ denotes the type of a sequence $\mathbf{x} \in \mathcal{X}^n$.

2.2 Some Preliminaries

We state here some useful information measures (see, for instance, (Bloch and Barros, 2011) for details). Given a discrete random variable X and $\alpha \in [0, 1)$, the *Renyi entropy of order* α is defined by:

$$H_\alpha(X) = \frac{1}{1-\alpha} \log_2 \sum_x (P_X(x))^\alpha.$$

The Renyi entropy specializes to the *Shannon entropy* as specified below:

$$H(X) = \lim_{\alpha \to 1} H_\alpha(X) = \sum_{x \in \mathcal{X}} P_X(x) \log \left(\frac{1}{P_X(x)} \right)$$

The *min-entropy* is the Renyi entropy for order $\alpha \to \infty$, viz.,

$$H_\infty(X) = \lim_{\alpha \to \infty} H_\alpha(X) = \min_x \log \left(\frac{1}{P_X(x)} \right)$$

It's conditional version is given by:

$$H_\infty(X|Y) = \min_y H_\infty(X|Y = y)$$

The *max-entropy* (i.e., Renyi entropy of order $\alpha \to 0$) and its conditional version are defined as:

$$\begin{aligned}H_0(X) &= \lim_{\alpha\to 0} H_\alpha(X) = \log|\{x \in \mathcal{X}|P_X(x) > 0\}| \\ H_0(X|Y) &= \max_y H_0(X|Y = y).\end{aligned}$$

Given two distributions $P_X, Q_X \in \mathcal{P}(\mathcal{X})$, let the total variation distance between P_X and Q_X be defined as

$$\|P_X - Q_X\| := \frac{1}{2}\sum_{x\in\mathcal{X}} |P_X(x) - Q_X(x)|.$$

For $\epsilon \in [0, 1)$, the ϵ-smooth entropies and their conditional versions are given by:

$$\begin{aligned}H^\epsilon_\infty(X) &= \max_{X':||P_{X'}-P_X||\le\,\epsilon} H_\infty(X') \\ H^\epsilon_\infty(X|Y) &= \max_{X',Y':||P_{X',Y'}-P_{X,Y}||\le\,\epsilon} H_\infty(X',Y') \\ H^\epsilon_0(X) &= \min_{X':||P_{X'}-P_X||\le\,\epsilon} H_0(X') \\ H^\epsilon_0(X|Y) &= \min_{X',Y':||P_{X',Y'}-P_{X,Y}||\le\,\epsilon} H_0(X'|Y').\end{aligned}$$

The following chain rules for these ϵ-smooth versions of entropy hold. For any $0 \le \epsilon, \epsilon', \epsilon_1, \epsilon_2 < 1$ and any random variables (X, Y, W), we have

$$H^{\epsilon+\epsilon'}_\infty(XY|W) - H^{\epsilon'}_\infty(Y|W) \ge H^\epsilon_\infty(X|YW) \ge H^{\epsilon_1}_\infty(XY|W) - H^{\epsilon_2}_0(Y|W) - \log\left(\frac{1}{\epsilon - \epsilon_1 - \epsilon_2}\right) \tag{2.1}$$

$$H_0^{\epsilon+\epsilon^{'}}(XY|W)-H_0^{\epsilon^{'}}(Y|W) \leq H_0^{\epsilon}(X|YW) \leq H_0^{\epsilon_1}(XY|W)-H_\infty^{\epsilon_2}(Y|W)+\log\left(\frac{1}{\epsilon-\epsilon_1-\epsilon_2}\right) \tag{2.2}$$

We also require general universal hash functions and randomness extractors in our proof of achievability in second part of our results (Chapter 5). We define them next (see (Bloch and Barros, 2011; Nisan and Zuckerman, 1996) for more details).

Definition 2.1 ($\beta-$Univeral Hash functions (Carter and Wegman, 1977, 1979))**.** Let $\mathcal{H}$ be a class of functions from $\mathcal{X}$ to $\mathcal{Y}$. We say that $\mathcal{H}$ is $\beta-$universal, where $\beta \in \mathbb{N}$, if for all distinct $x_1, x_2, ...x_\beta \in \mathcal{X}$ and $h \in \mathcal{H}$ chosen uniformly at random from $\mathcal{H}$, $(h(x_1), h(x_2), ...h(x_\beta))$ is distributed uniformly over $\mathcal{Y}^\beta$.

Definition 2.2 (Strong Randomness Extractors (Dodis et al., 2008; Nisan and Zuckerman, 1996))**.** A probabilistic polynomial time function of the form Ext: $\{0,1\}^n \times \{0,1\}^d \rightarrow \{0,1\}^m$ is an (n, k, m, ϵ)-strong extractor if for every probability distribution P_Z on $\mathcal{Z} = \{0,1\}^n$, and $H_\infty(Z) \geq k$, for random variables D (called 'seed') and M, distributed uniformly in $\{0,1\}^d$ and $\{0,1\}^m$ respectively, we have $||P_{Ext(Z;D),D} - P_{M,D}|| \leq \epsilon$.

i

Chapter 3

Commitment Capacity under Cost Constraints

This chapter starts with introducing the commitment problem over general DMCS. We then look at Trivial channels and proceed to define the commitment problem over such channels with cost constraints. We later find the primal and dual expressions for the commitment capacity.

3.1 Commitment Problem Setup over Discrete Memoryless channels with Cost Constraint

Refer to the bit commitment setup depicted in Fig 4.1. Here two mutually distrustful parties Alice and Bob aim to *commit* on a random bit string $C \in [2^{nR}]$ (where $R > 0$ is specified later) available at Alice. The two parties use a discrete memoryless channel (DMC) resource, specified by the conditional probability law $W_{Y|X} \in \mathcal{P}(\mathcal{Y}|\mathcal{X})$. Alice uses the DMC $W_{Y|X}$ through n rounds of one-way communication. Let $\mathbf{X}$ denote Alice's transmitted vector or *codeword* on the channel $W_{Y|X}$. Alice's set of feasible codewords is given by $\mathcal{S}(\Gamma) := \{\mathbf{x} \in \mathcal{X}^n : \sum_{i=1}^{n} \rho_X(x_i) \leq n\Gamma\}$,

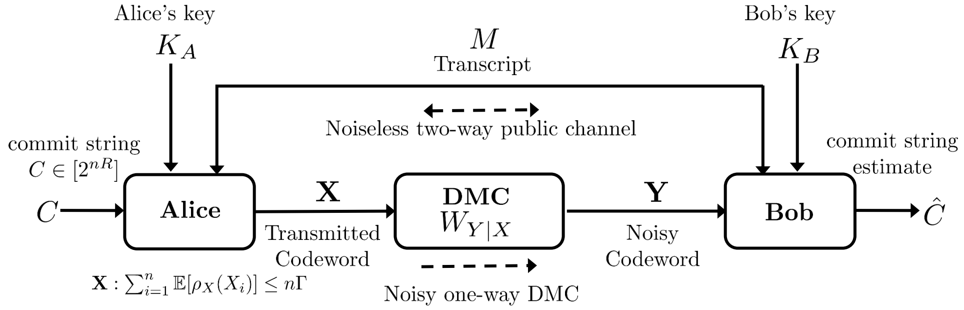


FIGURE 3.1: Commitment over DMC $W_{Y|X}$ with cost function $\rho_X : \mathcal{X} \to \mathbb{R}^+$ and cost constraint $\Gamma > 0$.

where $\rho_X : \mathcal{X} \to \mathbb{R}^+$ denotes the *channel cost function* and $\Gamma > 0$ denotes the *cost constraint*. Bob observes a noisy version $\mathbf{Y}$ of Alice's transmitted codeword $\mathbf{X}$. Alice and Bob are allowed to use private randomness strings $K_A \in \mathcal{K}_A$ and $K_B \in \mathcal{K}_B$ respectively. In addition, both Alice and Bob can also utilize *with no cost* a bi-directional *noiseless* (also, authenticated) link. Any message transmitted by a user during a protocol is a function of what the user observes up to that time. We now define a bit commitment protocol and its concomitant parameters, viz., *soundness*, *concealment* and *bindingness*.

3.1.1 Key features of a commitment protocol

We now define three key features of an (n, R) protocol. Let $\epsilon > 0$ be any arbitrary constant.

Definition 3.1 (ϵ-sound)**.** An (n, R) protocol is said to be ϵ-sound[1] if, when *both* parties Alice and Bob are *honest* and execute the protocol,

$$\mathbb{P}\left(T(C, \mathbf{X}, V_B) = 0\right) \leq \epsilon. \tag{3.1}$$

[1] In our definition of an ϵ-sound protocol, we average over the (uniformly) random commit strings C. As such, this is an *average* soundness criterion. If instead, we drop the averaging over C and demand that $\mathbb{P}\left(T(c, \mathbf{X}, \mathbf{Y}, M) = 0\right) \leq \epsilon$, $\forall c \in [2^{nR}]$, then the resulting criterion is a *maximum* soundness criterion.

Definition 3.2 (ϵ-concealing)**.** An (n, R) protocol is said to be ϵ-concealing if, under *any* strategy of Bob,

$$I(C; V_B) \leq \epsilon.$$

Definition 3.3 (ϵ-binding)**.** An (n, R) protocol is said to be ϵ-binding if under *any* strategy of Alice with an accompanying choice of $\mathbf{X} \in \mathcal{S}(\Gamma)$ during the commit phase and for any two pairs $(\bar{c}, \bar{\mathbf{X}})$, $(\hat{c}, \hat{\mathbf{X}})$, where $\bar{c} \neq \hat{c}$, and $\bar{\mathbf{X}}, \hat{\mathbf{X}} \in \mathcal{S}(\Gamma)$,

$$\mathbb{P}\Big(T(\bar{c}, \bar{\mathbf{X}}, V_B) = 1 \quad \& \quad T(\hat{c}, \hat{\mathbf{X}}, V_B) = 1\Big) \leq \epsilon.$$

A rate $R > 0$ is said to be *achievable* if for every $\epsilon > 0$, there exists an (n, R)-commitment protocol for every n sufficiently large such that the the protocol is ϵ*-sound*, ϵ*-binding* and ϵ*-concealing.* We define the *commitment capacity* or *capacity* under input constraint Γ, viz., $\mathbb{C}(\Gamma)$, as the supremum of all achievable rates.

Definition 3.4 (Commitment protocol)**.** An (n, R)-commitment protocol (also called protocol) is an exchange of messages between Alice and Bob over two phases called the *commit* phase followed by the *reveal* phase towards commitment over a random string $C \in [2^{nR}]$. Here $R > 0$ is called the *rate* of this protocol.

- *commit phase:* Given $C \in [2^{nR}]$, Alice transmits a feasible codeword $\mathbf{X} \in \mathcal{S}(\Gamma)$ on the DMC. Before and after every transmission (over time instants $i \in [n]$) over the DMC $W_{Y|X}$, both Alice and Bob take turns to exchange messages (these may be arbitrarily many but are finite in number) over the public noiseless channel. Let M denote all the messages exchanged over the public channel, i.e., the *transcript* of the protocol, at the end of the commit phase. Let the collection of all random variables generated and/or observed through the (n, R) protocol till the end of the commit phase, i.e., the *views* of Alice and Bob, be denoted by V_A and V_B, respectively. In particular, Alice's view $V_A := (C, K_A, \mathbf{X}, M)$ and Bob's view $V_B := (\mathbf{Y}, K_B, M)$ at the end of commit phase.

- *reveal phase:* In this phase, Alice and Bob exchange messages only over the public channel. To begin, Alice announces or *reveals* to Bob a pair of commit string and codeword[2] $(\bar{c}, \bar{\mathbf{x}})$, where $\bar{c} \in [2^{nR}]$ and $\bar{\mathbf{x}} \in \mathcal{S}(\Gamma)$. Thereafter, Bob runs a test $T = T(\bar{c}, \bar{\mathbf{x}}, V_B)$. Here the test output $T \in \{0, 1\}$, where 0 indicates that Bob *rejects* the commit string $\bar{c}$ and 1 indicates that Bob *accepts* the commit string $\bar{c}$.

We will now see a class of Discrete Memoryless Channels, called the Trivial Channels.

3.2 Trivial Channels

We know that even single-bit commitment (defined in Chapter 3) is impossible over noiseless DMCs (cf. (Damgård et al., 1999, pg. 9)); in fact, we show that such noiseless channels belong to a larger class of so-called (ρ_X, Γ)-*trivial channels* over which single-bit commitment is impossible. To specify this class, we introduce the following useful definitions[3].

Definition 3.5 ((ρ_X, Γ)-non-redundant channel)**.** Let P_X be such that $\mathbb{E}_{P_X}[\rho_X(X)] \leq \Gamma$. Then, $W_{Y|X}$ is said to be P_X-non-redundant if

$$W_{Y|X}(y|x) \neq \sum_{x' \in \mathcal{X},\ x' \neq x} P_X(x') W_{Y|X}(y|x'),$$

for every $y \in \mathcal{Y}$, $x \in \mathcal{X}$ such that $P_X(x) = 0$. If $W_{Y|X}$ is P_X-non-redundant for every feasible P_X, then we say that $W_{Y|X}$ is a (ρ_X, Γ)-non-redundant channel.

[2] Note that the pair $(\bar{c}, \bar{\mathbf{x}})$ may be the same pair that Alice used in the commit phase (if Alice is honest) or a different one (if Alice is dishonest).

[3] Our definition generalizes the corresponding notion in (Winter et al., 2003) for channels with no input constraints

Definition 3.6 ((ρ_X, Γ)-trivial channel)**.** A (ρ_X, Γ)-non-redundant channel $W_{Y|X}$ is said to be (ρ_X, Γ)-trivial if

$$W_{Y|X}(y|x) \cdot W_{Y|X}(y|x') = 0, \quad \forall y \in \mathcal{Y},$$

for every non-trivial and distinct $x, x' \in \mathcal{X}$.

Remark 3.7. For (ρ_X, Γ)-trivial channels, commitment fails due to the impossibility of satisfying the information-theoretic concealment requirement (cf. Definition 3.2) at Bob. This is because for (ρ_X, Γ)-trivial channels, the effective support of the conditional distributions $W_{Y|X}(y|x), W_{Y|X}(y|x') \in \mathcal{P}(\mathcal{Y})$ for any two distinct, non-trivial[4] symbols $x, x' \in \mathcal{X}$ are disjoint. As such, upon observing the output, Bob can effectively infer Alice's input string thereby making concealment impossible.

Having shown that the commitment capacity of (ρ_X, Γ)-trivial DMCs is zero, we now specify the the commitment capacity of (ρ_X, Γ)-non-trivial DMCs over some cost constraints.

3.3 Commitment Capacity over Non Redundant Channels

Theorem 3.8. *Let $W_{Y|X}$ be a (ρ_X, Γ)-non-trivial discrete memoryless channel. Then, the commitment capacity of $W_{Y|X}$ under the input constraint Γ, where $\Gamma \geq \min_x \rho_X(x)$, is given by*

$$\mathbb{C}(\Gamma) = \max_{P_X : \mathbb{E}[\rho_X(X) \leq \Gamma]} H(X|Y). \tag{3.2}$$

[4] In case, we have a (ρ_X, Γ)-redundant channel $W_{Y|X}$, then we can expurgate symbols in $\mathcal{X}$ which result in the redundancy. The resulting channel then is a (ρ_X, Γ)-non-redundant channel. The symbols in $\mathcal{X}$ retained after removing the redundant ones are the so-called 'non-trivial' symbols in $\mathcal{X}$.

Remark 3.9. (i) The commitment capacity specializes to that of the (input) unconstrained capacity (Winter et al., 2003); note that $\mathcal{S}(\Gamma) = \mathcal{X}^n$ when input is unconstrained, i.e., all $\mathbf{x}$ are feasible vectors.

(ii) $\mathbb{C}(\Gamma)$ is invariant to the specific nature of the ϵ-sound criterion, i.e., whether an *average* (as in (3.1)) or a *maximum* over the commit strings C is considered. We analyse the *maximum* (resp. *average*) criterion in the achievability (resp. converse) to establish this result.

The proof of this theorem is broken into an Achievability part and a Converse part

3.3.1 Converse Proof

To begin, we first state some useful claims which will be needed in the proof of the converse.

Claim 3.10 (Concavity of $H(X|Y)$). $H(X|Y)$ is a concave function of P_X, when $P_{Y|X}$ is fixed.

The proof uses the log-sum inequality.

Proof of claim:

$$H(X|Y) = \sum_{x,y} P_{X,Y}(x,y) \log P_{X|Y}(x|y) \tag{3.3}$$

$$= \sum_{x,y} P_{X,Y}(x,y) \log \frac{P_{X,Y}(x,y)}{P_Y(y)} \tag{3.4}$$

$$\tag{3.5}$$

Let $P_X^{(1)}$ and $P_X^{(2)}$ be two input distributions. Let $P_{X,Y}^{(1)} := P_X^{(1)} P_{Y|X}$ and $P_{X,Y}^{(2)} := P_X^{(2)} P_{Y|X}$. Let $P_X^{(\beta)} := \beta P_X^{(1)} + (1-\beta) P_X^{(2)}$; correspondingly, let $P_{X,Y}^{(\beta)} := \beta P_{X,Y}^{(1)} +$

$(1-\beta)P^{(2)}_{X,Y}$ and $P^{(\beta)}_{Y} := \beta P^{(1)}_{Y} + (1-\beta)P^{(2)}_{Y}$. Also, let $H^{(\beta)}(X|Y)$ be the conditional entropy under the joint distribution $P^{\beta}_{X,Y}$ and $H^{(i)}(X|Y)$, $i=1,2$ denote the conditional entropy under $P^{(i)}_{X,Y}$, $i=1,2$.

Then, from the definition of conditional entropy, we have

$$
\begin{aligned}
&H^{(\beta)}(X|Y) && (3.6)\\
&= -\sum_{x,y} P^{(\beta)}_{X,Y}(x,y)\log P^{(\beta)}_{X|Y}(x|y)\\
&= -\sum_{x,y} P^{(\beta)}_{X,Y}(x,y)\log \frac{P^{(\beta)}_{X,Y}(x,y)}{P^{(\beta)}_{Y}(y)}\\
&\overset{(a)}{=} -\sum_{x,y}(\beta P^{(1)}_{X,Y}(x,y) + (1-\beta)P^{(2)}_{X,Y}(x,y))\log \frac{\beta P^{(1)}_{X,Y}(x,y) + (1-\beta)P^{(2)}_{X,Y}(x,y)}{\beta P^{(1)}_{Y}(y) + (1-\beta)P^{(2)}_{Y}(y)} && (3.7)\\
&\overset{(b)}{\geq} -\sum_{x,y}\beta P^{(1)}_{X,Y}(x,y)\log \frac{\beta P^{(1)}_{X,Y}(x,y)}{\beta P^{(1)}_{Y}(y)} + (1-\beta)P^{(2)}_{X,Y}(x,y)\log \frac{(1-\beta)P^{(2)}_{X,Y}(x,y)}{(1-\beta)P^{(2)}_{Y}(y)} && (3.8)\\
&= -\beta\sum_{x,y} P^{(1)}_{X,Y}(x,y)\log \frac{P^{(1)}_{X,Y}(x,y)}{P^{(1)}_{Y}(y)} - (1-\beta)\sum_{x,y} P^{(2)}_{X,Y}(x,y)\log \frac{P^{(2)}_{X,Y}(x,y)}{P^{(2)}_{Y}(y)}\\
&= \beta H^{(1)}(X|Y) + (1-\beta)H^{(2)}(X|Y).
\end{aligned}
$$

where

(a) follows from the definitions of $P^{(\beta)}_{X,Y}$ and $P^{(\beta)}_{Y}$.

(b) follows from the log-sum inequality (cf. (Cover, 1999, Chapter 2)) applied to each term of the summation.

■

Claim 3.11. $\mathbb{C}(\Gamma)$ is a non-decreasing function of Γ.

Proof: The proof of this claim is straight forward; here are the details for completeness. Let $\mathcal{S}_1 := \{P_X : \mathbb{E}[\rho_X(x)] \leq \Gamma_1\}$ and $\mathcal{S}_2 := \{P_X : \mathbb{E}[\rho_X(X)] \leq \Gamma_2\}$. Now if

$\Gamma_1 \leq \Gamma_2$ are two average power constraint values, then $\mathcal{S}_1 \subseteq \mathcal{S}_2$. Hence, from the definition of capacity (cf. (3.2)), it follows that $\mathbb{C}(\Gamma_1) \leq \mathbb{C}(\Gamma_2)$. ■

Claim 3.12. $\mathbb{C}(\Gamma)$ is a concave function of Γ.

Proof: Let $\Gamma^{(1)}, \Gamma^{(2)} > 0$ denote two average power constraint values. Also, let $\mathbb{C}_i := \mathbb{C}(\Gamma^{(i)})$, $i = 1, 2$, according to (3.2), and let $P_X^{(1)}$ and $P_X^{(2)}$ be corresponding maximizing distributions. For $\beta \in [0,1]$, define $\Gamma^{(\beta)} := \beta\Gamma^{(1)} + (1-\beta)\Gamma^{(2)}$ and $P_X^{\beta} := \beta P_X^{(1)} + (1-\beta)P_X^{(2)}$. Then, if P_X^* denotes an optimizer for $\mathbb{C}(\Gamma^{(\beta)})$ (cf. (3.2)), we have

$$\begin{aligned}
\mathbb{C}(\Gamma^{(\beta)}) &= H_{P_X^*}(X|Y) \\
&\geq H_{P_X^{(\beta)}}(X|Y) \\
&\overset{(a)}{\geq} \beta H_{P_X^{(1)}}(X|Y) + (1-\beta)H_{P_X^{(2)}}(X|Y) \\
&\overset{(b)}{=} \beta\mathbb{C}(\Gamma^{(1)}) + (1-\beta)\mathbb{C}(\Gamma^{(2)}),
\end{aligned}$$

where

(a) follows by noting that $H(X|Y)$ is a concave function of P_X when $P_{Y|X} = W_{Y|X}$ is fixed (cf. Claim 3.10).

(b) follows from definition of $C(\Gamma^{(i)})$, $i = 1, 2$. Hence, it follows that $C(\Gamma)$ is concave in Γ.

■

For the converse, consider any sequence of (n, R)-commitment protocols, say $\{\mathscr{P}_n\}$, $n \geq 1$, where protocol $\mathscr{P}_n$, $\forall n$, satisfies the cost constraint Γ. In addition, let protocol $\mathscr{P}_n$, $\forall n$, be ϵ_n-sound, ϵ_n-concealing and ϵ_n-binding, where $\epsilon_n > 0$ and $\epsilon_n \to 0$ as $n \to \infty$. Then, the following lemma holds.

Lemma 3.13. *For every $\mathscr{P}_n$, $H(C|\mathbf{X}, V_B) \leq n\epsilon'_n$, where $\epsilon'_n \to 0$ as $n \to \infty$.*

Proof of claim: Recall that V_B denotes Bob's view at the end of commit phase of the protocol. We aim to show that given its view V_B and with $\mathbf{X}$ additionally, Bob's uncertainty in C is vanishingly small (as block length n increases).

To begin, let us define $\tilde{c} := \arg\max_{c \in [2^{nR}]} T(\tilde{c}, \mathbf{X}, V_B)$. We now bound $\mathbb{P}(\hat{C} \neq C)$, viz., the probability of error in the recovery of the commit string, where $\hat{C} = \hat{C}(V_B, \mathbf{X}) = \tilde{c}$. As the code is ϵ_n-binding, it follows that

$$\mathbb{P}\left(T(\bar{c}, \bar{\mathbf{X}}, V_B) = 1 \quad \& \quad T(\hat{c}, \hat{\mathbf{X}}, V_B) = 1\right) \leq \epsilon_n \tag{3.9}$$

for any two distinct $(\bar{c}, \bar{\mathbf{X}})$ and $(\hat{c}, \hat{\mathbf{X}})$. For the given decoder, we have

$$\begin{aligned} \mathbb{P}(\hat{C} \neq C) &= \mathbb{P}(\hat{C} = 0) + \mathbb{P}(\hat{C} \neq C | C \neq 0) &(3.10)\\ &\leq \epsilon_n + \epsilon_n = 2\epsilon_n. &(3.11) \end{aligned}$$

where in the final step the first part follows from noting that the code $\mathcal{C}_n$ is ϵ_n-binding while the second part follows from the fact that conditioned on $\mathcal{C}_n$ being ϵ_n-binding, the probability that $\hat{C}$ is different from C is at most ϵ_n due to the soundness guarantee.

We now invoke Fano's inequality (cf. (Cover, 1999)) to complete the proof.

$$\begin{aligned} H(C|\mathbf{X}, V_B) &\leq 1 + \mathbb{P}(\hat{C} \neq C) nR &(3.12)\\ &\leq n\left(\frac{1}{n} + 2\epsilon_n R\right) &(3.13)\\ &\leq n\epsilon'_n &(3.14) \end{aligned}$$

where $\epsilon'_n \to 0$ as $n \to \infty$. This completes the proof of the claim. ■

We now bound the rate R as follows:

$$\begin{aligned}
nR &= H(C) \\
&= H(C|V_B) + I(C;V_B) \\
&\overset{(a)}{\leq} H(C|V_B) + \epsilon_n \\
&\overset{(b)}{=} H(C|\mathbf{Y}, M, K_B) + \epsilon_n \\
&\overset{(c)}{=} H(C, \mathbf{X}|\mathbf{Y}, M, K_B) - H(\mathbf{X}|\mathbf{Y}, M, K_B, C) + \epsilon_n \\
&\overset{(d)}{\leq} H(C, \mathbf{X}|\mathbf{Y}, M, K_B) + \epsilon_n \\
&\overset{(e)}{=} H(\mathbf{X}|\mathbf{Y}, M, K_B) + H(C|\mathbf{X}, \mathbf{Y}, M, K_B) + \epsilon_n \\
&= H(\mathbf{X}|\mathbf{Y}, M, K_B) + H(C|\mathbf{X}, V_B) + \epsilon_n \\
&\overset{(f)}{\leq} H(\mathbf{X}|\mathbf{Y}) + n\epsilon_n' + \epsilon_n \\
&\leq \sum_{i=1}^{n} H(X_i|Y_i) + n\epsilon_n' + \epsilon_n \\
&= n\left(\sum_{i=1}^{n} \frac{1}{n} H(X_i|Y_i)\right) + n\epsilon_n' + \epsilon_n \\
&\overset{(g)}{\leq} n\left(\sum_{i=1}^{n} \frac{1}{n} \mathbb{C}(\mathbb{E}[\rho_X(X_i)])\right) + n\epsilon'_+\epsilon_n \\
&\overset{(h)}{\leq} n\mathbb{C}\left(\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[\rho_X(X_i)]\right) + n\epsilon_n' + \epsilon_n \\
&\overset{(i)}{\leq} n\mathbb{C}\left(\Gamma\right) + n\epsilon_n' + \epsilon_n.
\end{aligned}$$

Here

(a) follows from the fact that the sequence of coding schemes is ϵ_n-concealing.

(b) follows from noting that $V_B = (\mathbf{Y}, M, K_B)$

(c) follows from the chain rule of joint entropy

(d) follows from the fact that $H(\mathbf{X}|\mathbf{Y}, M, K_B, C) \geq 0$

(e) follows from the chain rule of joint entropy

(f) follows from the fact that conditioning reduces entropy and Lemma 4.8

(g) follows from definition of the commitment capacity $\mathbb{C}(\Gamma)$

(h) follows from the concavity of $\mathbb{C}(\Gamma)$ w.r.t. Γ (see Claim 3.12)

(i) follows from the non-decreasing nature of $\mathbb{C}(\Gamma)$ w.r.t. Γ (see Claim 3.11).

Now taking the limit $n \to \infty$ and noting that $\epsilon_n, \epsilon'_n \to 0$ as $n \to \infty$, we get the bound $R \leq \mathbb{C}(\Gamma)$. This completes the proof of the converse.

3.3.2 Achievability

Outline: We now present a sequence of commitment capacity-achieving protocols. Each protocol in this sequence consists of a codebook, using which we describe the commit and reveal phase of that protocol. We construct a deterministic binned codebook $\mathcal{C}$ (described later) using the random coding argument with expurgation. This is along the lines of approach in (Winter et al., 2003; Wyner, 1975). The codebook properties are specified in Lemma 3.14. Our protocol uses this codebook $\mathcal{C}$ and employs a stochastic encoding strategy by Alice. We show that this protocol satisfies the three requirements of soundness (cf. Definition 3.1), concealment (cf. Definition 3.2) and bindingness (cf. Definition 3.3). Owing to space constraints, we present an outline of the analysis.

The following lemma (proof omitted) guarantees the existence of a deterministic codebook with the requisite properties.

Lemma 3.14 (Binned Codebook construction). *Let P_X be such that $\mathbb{E}_{P_X}[\rho_X(X)] \leq \Gamma$. Let $\varepsilon > 0, \eta > 0$. Let $R_{ov} = R + \tilde{R}$, where R_{ov}, R and $\tilde{R} > 0$. Fix $R = H(X|Y) - \varepsilon$, $\tilde{R} = I(X;Y) + \varepsilon/2$, and $R_{ov} = R + \tilde{R} = H(X) - \varepsilon/2$, where $H(X)$ is evaluated under the distribution P_X, while $H(X|Y)$ and $I(X;Y)$ are evaluated under the joint* *distribution* $P_{X,Y} = P_X W_{Y|X}$ *resp.. Then, there exists a collection of* $2^{nR_{ov}}$ *codewords* $\{\mathbf{x}_{c,k}\}$*, where* $c \in [2^{nR}]$*,* $k \in [2^{n\tilde{R}}]$ *and every vector* $\mathbf{x}_{c,k} \in \mathcal{T}_{\delta}^{(n)}(P_X)$*, for some small enough* $\delta(\varepsilon) > 0$*, where* $\delta \to 0$ *as* $\varepsilon \to 0$*, such that*

(i) $d_H(\mathbf{x}_{c,k}, \mathbf{x}_{c',k'}) \geq 2n\eta$*,* $\forall c \neq c'$*,* $c, c' \in [2^{nR}]$*,* $k, k' \in [2^{n\tilde{R}}]$*,*

(ii) for every $c \in [2^{nR}]$*,*

$$D\left(\frac{1}{2^{n\tilde{R}}}\sum_{k=1}^{2^{n\tilde{R}}} W_{Y|X}^{(n)}(\mathbf{y}|\mathbf{x}_{c,k})\middle\|[P_X W_{Y|X}]_Y^{(n)}(\mathbf{y})\right) \leq e^{-n\alpha}, \tag{3.15}$$

for some $\alpha(\delta) > 0$*, where* $\alpha \to 0$ *as* $\delta \to 0$*.*

Our proof of this lemma analyses an i.i.d. random code (under the distribution P_X). We use the random coding argument along with expurgation of codewords (so as to ensure the property (i) above). A crucial result used in our proof is the so-called 'stronger' soft covering lemma (Cuff, 2015) stated below.

Lemma 3.15 ('Stronger' Soft Covering lemma (Cuff, 2015))**.** *Fix any* $P_X \in \mathcal{P}(\mathcal{X})$*,* $P_{Y|X} \in \mathcal{P}(\mathcal{Y}|\mathcal{X})$*, where* $|\mathcal{Y}| < \infty$*. Let* $R > I(X;Y)$*. Then, there exists* $\beta_1, \beta_2 > 0$ *such that for n sufficiently large,*

$$\mathbb{P}\left[D\left(\frac{1}{2^{nR}}\sum_{k=1}^{2^{nR}} P_{Y|X}^{(n)}(\mathbf{y}|\mathbf{X}_k)\middle\|[P_X P_{Y|X}]_Y^{(n)}(\mathbf{y})\right) > e^{-n\beta_1}\right] \leq e^{-e^{n\beta_2}},$$

where the probability is over the random code $\{\mathbf{X}_i\}_{i=1}^{2^{nR}}$*, and* $\mathbf{X}_i \sim P_X$ *i.i.d. for every* $i \in [2^{nR}]$*.*

We now describe our protocol.

- *Commit phase:* Alice wants to commit on binary string $c \in [2^{nR}]$ with Bob and proceeds as follows:

- Alice picks the bin $c \in [2^{nR}]$ in $\mathcal{C}$ and chooses a random codeword $\mathbf{X}_{c,K}$, where $K \sim \text{Unif}\left([2^{n\tilde{R}}]\right)$.
- Alice transmits $\mathbf{X} := \mathbf{X}_{c,k}$ over the DMC to Bob.
- Bob receives $\mathbf{Y}$ over the channel.

- *Reveal phase:* The reveal phase proceeds as follows :

  - Having received $\mathbf{Y} = \mathbf{y}$, Bob calculates the following list of candidate codewords:

$$\mathcal{L}(\mathbf{y}) := \{\mathbf{x} \in \mathcal{C} : T_{\mathbf{x},\mathbf{y}} \in \mathcal{T}_{\delta'}^{(n)}(P_X W_{Y|X})\},$$

  where $\delta'(\delta) > 0$ small enough and $\delta' \to 0$ as $\delta \to 0$.

  - Alice announces over the noiseless link the pair $(\tilde{c}, \tilde{\mathbf{x}})$.
  - Bob checks if $\tilde{\mathbf{x}} \in \mathcal{L}(\mathbf{y})$. If $\tilde{\mathbf{x}} = \tilde{\mathbf{x}}_{\tilde{c},k}$ for some $k \in [2^{n\tilde{R}}]$ satisfies this test uniquely, then Bob accepts the commit string $\hat{c} = \tilde{c}$. Otherwise, Bob declares error with $\hat{c} = 0$.

- *Analysis:* Here is an outline of our analysis.

(i) *ϵ-sound:* We analyse the event $\{\mathbf{X} \notin \mathcal{L}(\mathbf{Y})\}$. Using standard Chernoff bounds, it can be shown that $\mathbb{P}\left(\mathbf{X} \notin \mathcal{L}(\mathbf{Y})\right)$ is exponentially decreasing as $n \to \infty$[5]; hence, for n sufficiently large, we can show that our protocol is ϵ-sound.

(ii) *ϵ-concealing:* Recall property (ii) in Lemma 3.14 which guarantees that for *every* $c \in [2^{nR}]$, the KL divergence in (3.15) is vanishingly small. Pinsker's inequality (Csiszár and Körner, 2011) then guarantees that the corresponding ℓ_1 distance is exponentially small. This implies (using (Damgard et al., 1998) and some analysis) that the resulting average mutual information $I(C; V_B) \leq \epsilon$ for n sufficiently large. Note that here Bob's view $V_B = \mathbf{Y}$. This follows from $M = \emptyset$ and $K_B = \emptyset$ as public

[5] Here the probability is over Alice's private randomness and $W_{Y|X}$.

channel is not utilized and no randomness K_B at Bob is used in the commit phase. (iii) *ϵ-binding:* Let a potentially dishonest Alice send a feasible vector $\mathbf{x} \in \mathcal{S}(\Gamma)$ over the DMC in the commit phase and let $\mathbf{Y} = \mathbf{y}$ be Bob's received output; without loss of generality, let us assume that Alice intends to violate the bindingness guarantee through two distinct pairs $(\bar{c}, \mathbf{x}_{\bar{c},\cdot})$ and $(\bar{c}', \mathbf{x}_{\bar{c}',\cdot})$, where $\bar{c} \neq \bar{c}'$. However, recall that our code $\mathcal{C}$ satisfies condition (i) in Lemma 3.14 and hence, $d_H(\mathbf{x}_{\bar{c},\cdot}, \mathbf{x}'_{\bar{c}',\cdot}) \geq 2n\eta$. This guarantees that $\min\{d_H(\mathbf{x}_{\bar{c},\cdot}, \mathbf{x}), d_H(\mathbf{x}'_{\bar{c}',\cdot}, \mathbf{x})\} \geq n\eta$, i.e., there exists at least one codeword amongst the two, say $\tilde{\mathbf{x}}$, which is at a Hamming distance $n\eta$ from $\mathbf{x}$, i.e., $d_H(\tilde{\mathbf{x}}, \mathbf{x}) \geq n\eta$. Given that $\mathbf{y} \in \mathcal{T}^{(n)}_{\delta'}(P_{X,Y}|\mathbf{x})$ w.h.p.[6], it follows that for $\eta > \delta'$ appropriately chosen (note that Lemma 3.14 guarantees that such an $\eta > 0$ choice is possible), the probability that $\mathbf{y} \in \mathcal{T}^{(n)}_{\delta'}(P_{X,Y}|\tilde{\mathbf{x}})$ is exponentially decaying. This follows from the Chernoff bound. Thus, for n sufficient large, we can guarantee that our protocol is ϵ-binding.

3.4 Dual Capacity expression for commitment over Non Redundant Channels

Next, we present a dual characterization of the commitment capacity.

Theorem 3.16. *Let $W_{Y|X}$ be a (ρ_X, Γ)-non-redundant discrete memoryless channel. Then, for any $\Gamma \geq \min_{x \in \mathcal{X}} \rho_X(x)$,*

$$\mathbb{C}(\Gamma) = \min_{\gamma \geq 0} \max_{Q_Y} \log \left[\sum_{x \in \mathcal{X}} 2^{-D(W_{Y|X}(\cdot|x)||Q_Y(\cdot)) + \gamma(\Gamma - \rho_X(x))} \right].$$

Furthermore, the maximizing distribution Q_Y is unique and $Q_Y = [P_X W_{Y|X}]_Y$, where P_X is any optimizer of (3.2).

Remark 3.17. The dual capacity characterization offers an alternate method to compute the commitment capacity. Given the channel law and the size of the input and

[6] Here w.h.p. stands for with high probability.

output alphabets, one may prefer either of the two results depending on the computational and/or analytical tractability of the concomitant optimization problems.

To illustrate the utility of our capacity characterizations, we determine the commitment capacity of the binary symmetric channel (BSC) under input (Hamming weight) cost constraints.

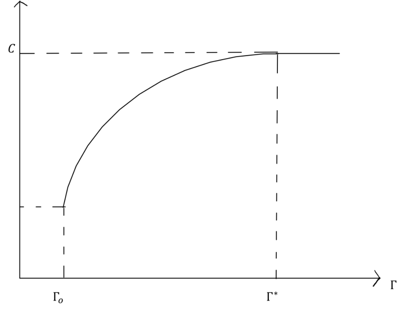
3.4.1 Proof of the Dual Expression

We have already claimed (3.10, 3.11) that $\mathbb{C}(\Gamma)$ is a non decreasing concave function of Γ. Let us define Γ_o as the smallest Γ whose capacity is of interest. Let Γ^* be the largest Γ for which $\mathbb{C}(\Gamma)$ is increasing i.e., beyond Γ^* $\mathbb{C}(\Gamma)$ stays constant. In fact this constant value $\mathbb{C}(\Gamma^*)$ is the capacity of the unconstrained general DMC.

$$\Gamma_o = \arg\min_{\Gamma} \big[\mathbb{C}(\Gamma)\big] \tag{3.16}$$

$$\Gamma^* = \arg\min_{\Gamma} \big[\mathbb{C}(\infty) - \mathbb{C}(\Gamma)\big] \tag{3.17}$$

Now that we have established these characteristics of $\mathbb{C}(\Gamma)$ we can plot it as has been done in Fig (.). Here on we will follow an approach that has been inspired by [(Csiszár and Körner, 2011) , Ch 7] and also has some interesting differences.



Let us pick $\Gamma_1 \in [\Gamma_o, \Gamma^*]$. If we draw a tangent at this point, we see that it has a non negative slope, say γ_1 and a y-intercept $F(\gamma_1)$. It follows from non deceasing concave behaviour of $\mathbb{C}(\Gamma)$ that there is a unique γ for every Γ point. The point slope equation of this tangent is

$$F(\gamma_1) = \mathbb{C}(\Gamma_1) - \gamma_1\Gamma_1$$

We introduce the following notations which will be used hereafter.

$$\begin{aligned}
\rho_X(P_X) &\triangleq \mathbb{E}_{P_X}[\rho_X(X)] && (3.18)\\
P_X W_{Y|X} &= \mathbb{E}_{P_X}[W_{Y|X}] && (3.19)\\
D(Q_Y||W_{Y|X}|P_X) &= \sum_{x\in\mathcal{X}} P_X(x) D(Q_Y(\cdot)||W_{Y|X}(\cdot|X=x)) && (3.20)\\
H(P_X|W_{Y|X}) &\triangleq H(X|Y) && (3.21)\\
&= D(P_X W_{Y|X}||W|P_X) && (3.22)
\end{aligned}$$

We can see from Fig(2) that among a family of parallel lines of slope γ_1 drawn from points on the curve $\mathbb{C}(\Gamma)$, the y-intercept is maximum for $\Gamma = \Gamma_1$.

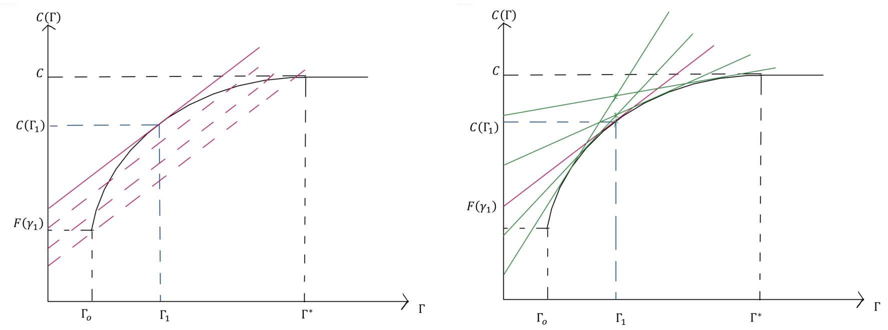
$$\begin{aligned}
\Rightarrow F(\gamma_1) &= \max_{\Gamma} \left[\mathbb{C}(\Gamma) - \gamma_1\Gamma\right] && (3.23)\\
&= H({P_X}^*|W) - \gamma\rho_X({P_X}^*) + \max_{\Gamma} \left[(\mathbb{C}(\Gamma) - H({P_X}^*|W)) - \gamma_1(\Gamma - \rho_X({P_X}^*))\right]
\end{aligned}$$

Where ${P_X}^*$ is some P_X optimising the primal expression i.e. $\mathbb{C}(\Gamma_1) = \max_{P_X:\rho_X(P_X)\leq\Gamma_1} H(P_X|W_{Y|X})$ occurs at ${P_X}^*$ and $\rho_X({P_X}^*) = \Gamma_1$(Note that ${P_X}^*$ need not be unique). So, for the given constraint, both the first and second terms of the maximisation expression of $F(\gamma_1)$ are maximised to zero at ${P_X}^*$.

$$\begin{aligned}
\Rightarrow F(\gamma_1) &= H({P_X}^*|W_{Y|X}) - \gamma_1\rho_X({P_X}^*) && (3.24)\\
&= \max_{P_X} \left[H(P_X|W_{Y|X}) - \gamma_1\rho_X(P_X)\right] && (3.25)\\
&= \max_{P_X} \left[H(P_X) - I(P_X|W_{Y|X}) - \gamma_1\rho_X(P_X)\right] && (3.26)
\end{aligned}$$

As detailed in (Csiszár and Körner, 2011), it can be observed that for a concave function, the lower envelope of the tangents to a curve form the curve itself. This can also be seen in Fig(3). Now, plotting all these tangents with slope-intercepts $\{\gamma, F(\gamma)\}$, and looking at the family of points $F(\gamma)+\gamma\Gamma_1$(the vertical line at Γ_1) we see that the minimum occurs at γ_1. Therefore

$$C(\Gamma_1) = \min_{\gamma \geq 0} \left[F(\gamma) + \gamma\Gamma_1\right] \tag{3.27}$$



To analytically determine $F(\gamma)$ in equation(5), we will follow the following steps. Let's define a function $F(P_X, Q_Y)$ on distributions P_X on inputs $\mathcal{X}$ and Q_Y on outputs $\mathcal{Y}$.

$$\begin{aligned} F(P_X, Q_Y) &\triangleq H(P_X) - I(P_X, W_{Y|X}) - D(P_X W_{Y|X}||Q_Y) - \gamma\rho_X(P_X) && (3.28)\\ &= H(P_X) - D(W_{Y|X}||Q_Y|P_X) - \gamma\rho_X(P_X) && (3.29)\\ &= H(P_X) - \sum_{x\in\mathcal{X}} P_X(x)\left[D(W_{Y|X}(\cdot|x)||Q_Y) + \gamma\rho_X(x)\right] && (3.30)\\ &= H(P_X) - \mathbb{E}_{P_X}\left[D(W_{Y|X}(\cdot|x)||Q_Y) + \gamma\rho_X(x)\right] && (3.31)\end{aligned}$$

Claim 3.18.

$$\max_{P_X} F(P_X, Q_Y) = \log\left(\sum_{x\in\mathcal{X}} \exp(-D(W_{Y|X}(\cdot|x)||Q_Y) - \gamma\rho_X(x))\right) \tag{3.32}$$

Proof: This can be proved from lemma A.1 in Appendix. We substitute $\alpha = 1$ and $f(x) = D(W_{Y|X}(\cdot|x)||Q_Y) + \gamma\rho_X(x)$. The maximum value of the LHS over all P_X equals RHS. This maximum value is
$\log\left(\sum_{x\in\mathcal{X}} \exp(-D(W_{Y|X}(\cdot|x)||Q_Y) - \gamma\rho_X(x))\right)$ ■

Claim 3.19.

$$\max_{Q_Y} F(P_X, Q_Y) = H(P_X) - I(P_X, W_{Y|X}) - \gamma\rho_X(P_X)$$

Proof: From the definition of $F(P_X, Q_Y)$ in equation (3.28)

$$\begin{aligned} F(P_X, Q_Y) &= H(P_X) - I(P_X, W_{Y|X}) - D(P_X W_{Y|X}||Q_Y) - \gamma\rho_X(P_X) \\ \Rightarrow \max_{Q_Y} F(P_X, Q_Y) &= H(P_X) - I(P_X, W_{Y|X}) - \gamma\rho_X(P_X) \end{aligned} \tag{3.33}$$

This follows from the fact that $D(P_X W_{Y|X}||Q_Y) \geq 0$ with equality iff $P_X W_{Y|X}$ and Q are identical. ■

From Claim 2,

$$\max_{P_X} \max_{Q_Y} F(P_X, Q_Y) = \max_{P_X} \left[H(P_X) - I(P_X, W_{Y|X}) - \gamma\rho_X(P_X)\right] \tag{3.34}$$

$$= F(\gamma) \tag{3.35}$$

This follows from definition in (El Gamal and Kim, 2011). Now, we interchange the order of maximisation. The optimising value so obtained shouldn't be different since we are not restricting the domains of P_X and Q_Y.

$$\begin{aligned}
F(\gamma) &= \max_{P_X,Q_Y} F(P_X,Q_Y) && (3.36)\\
&= \max_{Q_Y,P_X} F(P_X,Q_Y) && (3.37)\\
&= \max_{Q_Y}\max_{P_X} F(P_X,Q_Y) && (3.38)\\
&= \max_{Q_Y} \big[\log\left(\sum_{x\in\mathcal{X}} \exp(-D(W_{Y|X}(\cdot|x)||Q_Y) - \gamma\rho_X(x))\right)\big] && (3.39)
\end{aligned}$$

We now determine $\mathbb{C}(\Gamma)$ from equations (3.27), (3.39)

$$\begin{aligned}
\mathbb{C}(\Gamma) &= \min_{\gamma\geq 0}\big[F(\gamma)+\gamma\Gamma\big] && (3.40)\\
\mathbb{C}(\Gamma) &= \min_{\gamma\geq 0}\big[\max_{Q_Y}\log\Big(\sum_{x\in\mathcal{X}}\exp[-D(W_{Y|X}(\cdot|x)||Q_Y)-\gamma\rho_X(x)]\Big)+\gamma\Gamma\big] && (3.41)\\
&= \min_{\gamma\geq 0}\max_{Q_Y}\big[\log\Big(\sum_{x\in\mathcal{X}}\exp[-D(W_{Y|X}(\cdot|x)||Q_Y)+\gamma(\Gamma-\rho_X(x)]\Big)\big] && (3.42)
\end{aligned}$$

This completes the proof of Theorem (3.16).

3.4.2 Example - Input Constrained BSC capacity

Consider the binary symmetric channel BSC(p), where $p \leq 1/2$. Let $\Gamma \in [0,1/2)$ and let $\rho_X(0) = 0$ and $\rho_X(1) = 1$. Thus, $\mathcal{S}(\Gamma) := \{\mathbf{x} \in \{0,1\}^n : wt_H(\mathbf{x}) \leq n\Gamma\}$, where $wt_H(\mathbf{x}) = d_H(\mathbf{0},\mathbf{x})$ is the Hamming weight of $\mathbf{x}$.

The capacity of this (ρ_X,Γ)-input constrained BSC(p) is given by

$$\mathbb{C}_{BSC}(\Gamma) = H_2(\Gamma) + H_2(p) - H_2(p\otimes\Gamma), \tag{3.43}$$

where $H_2(\cdot)$ is the binary entropy function and $a\otimes b := a(1-b)+(1-a)b, \forall a,b \in [0,1]$.

The capacity expression in (3.43) follows from both of our capacity characterizations in Theorems 3.8 and 3.16 and so verify each other. The proof details of this result can be found in the B. Note that when $\Gamma = 1/2$, we have $H_2(\Gamma) = H_2(p\otimes\Gamma) =$

$H_2(1/2) = 1$. Hence, commitment capacity in this case $\mathbb{C}_{BSC}(\Gamma = 1/2) = H_2(p)$ which is the capacity of the unconstrained BSC(p).

This example illustrates the point made earlier that the commitment capacity of a DMC (in the above case, the DMC is a BSC(p)) under cost constraints can differ significantly from its unconstrained commitment capacity.

Chapter 4

Commitment over Compound Channels

A classic BSC is a simple DMC channel. Sometimes however such a channel is not fully characterised. We model such inconsistencies in the form of a compound BSC channel in this chapter.

4.1 Compound Binary Symmetric Channels

Definition 4.1 (Compound-binary symmetric channel (compound-BSC)). A compound binary symmetric channel with parameters[1] $0 < \gamma < \delta < 1/2$ is a binary symmetric channel with transition probability $s \in \mathcal{S}$, where $\mathcal{S} = [\gamma, \delta]$. The channel, also denoted by compound-BSC$[\gamma, \delta]$, has input $X \in \mathcal{X} = \{0, 1\}$, output

[1] As stated earlier, we rule out $s = 0$ and $s = 1/2$ as it can be trivially shown that even a single-bit commitment is impossible for these values.

$Y \in \mathcal{Y} = \{0, 1\}$, and its channel law is specified by:

$$W_{Y|X,S}(y|x,s) = \begin{cases} 1-s & \text{if } x = y \\ s & \text{if } x \neq y \end{cases}$$

The state $s \in [\gamma, \delta]$ is chosen arbitrarily from the set $\mathcal{S}$ and remains fixed throughout; furthermore, the state is assumed to be *unknown* to either Alice or Bob, whether they are honest or dishonest.

Remark 4.2. A compound-BSC$[\gamma, \delta]$ belongs to the class of compound DMCs studied widely in literature (cf. (El Gamal and Kim, 2011, Ch. 7)). Compound DMCs are specified by the channel law given by the conditional distribution $W_{Y|X,S}$, where $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$ are the channel input and output, while $s \in \mathcal{S}$ is the compound channel state. For the purposes of this work, $\mathcal{S}$ is a closed interval, viz., $\mathcal{S} = [\gamma, \delta]$; however, this restriction can be easily removed to allow arbitrary sets $\mathcal{S}$. We avoid the details (corresponding to arbitrary $\mathcal{S}$) in this thesis.

Although we already discussed the commitment problem over general UNCs in the previous chapter, let us now also look at the same problem again over the compound channels.

4.2 Commitment Protocol Problem setup over Binary Compound Channels

The setup for commitment of a binary bit string over a specialised Binary Compound Channel is similar to the one of DMCs in the previous chapter. It is a specialisation of the the setup where there is no cost constraint (i.e., Γ is large) and the Memoryless channel used is the Binary Compound Channel which is as has been described above. Alice uses the compound-BSC for n rounds of one-way communication to

Bob. Alice's transmission to Bob is denoted by $\mathbf{X}$; we also call it Alice's *codeword.* A noisy version $\mathbf{Y}$ of Alice's codeword $\mathbf{X}$ is observed by Bob. We allow Alice and Bob to access to private randomness strings $K_A \in \mathcal{K}_A$ and $K_B \in \mathcal{K}_B$ respectively. Both Alice and Bob also have access to a bi-directional and authenticated noiseless public link. At any time, any message transmitted by either Alice or Bob can depend causally on the information available to them at that time.

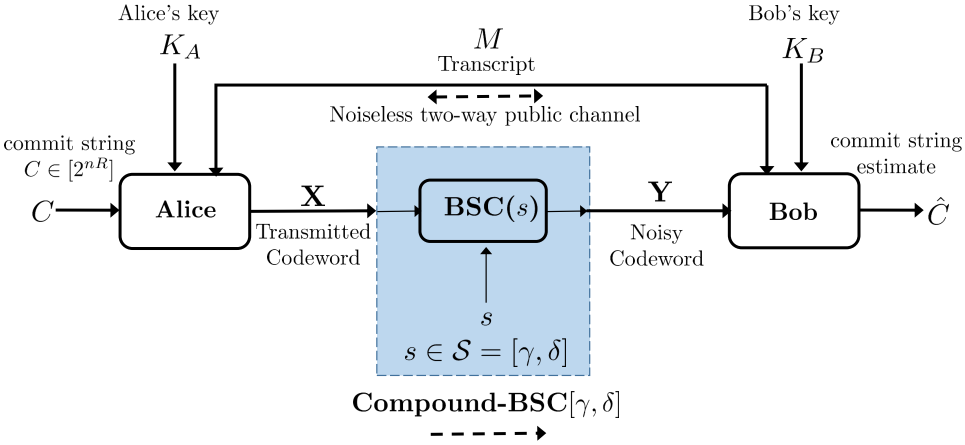


Figure 4.1: Commitment over a compound-BSC$[\gamma, \delta]$

This problem setup will be used to present the results in the second part of the thesis (Chapter 5).

4.2.1 Key Features of a Commitment Protocol

We now define three key parameters of an (n, R) protocol. Let $\epsilon > 0$ be any arbitrary constant.

Definition 4.3 (ϵ-sound). An (n, R) protocol is said to be ϵ-sound[2]: if when *both* parties Alice and Bob are *honest* and execute the protocol,

$$\max_{s \in \mathcal{S}} \mathbb{P}\left(T(C, \mathbf{X}, V_B) = 0 | S = s\right) \leq \epsilon \tag{4.1}$$

[2] In our definition of an ϵ-sound protocol, we average over the (uniformly) random commit strings C. As such, this is an *average* soundness criterion. If instead, we drop the averaging over C and instead demand that $\mathbb{P}\left(T(c, \mathbf{X}, \mathbf{Y}, M) = 0\right) \leq \epsilon$, $\forall c \in [2^{nR}]$, then the resulting criterion is a *maximal* soundness criterion.

Definition 4.4 (ϵ-concealing). An (n, R) protocol is said to be ϵ-concealing if under *any* strategy of Bob

$$\max_{s \in \mathcal{S}} I(C; V_B | S = s) \leq \epsilon \tag{4.2}$$

Definition 4.5 (ϵ-binding). An (n, R) protocol is said to be ϵ-binding if under *any* strategy of Alice with an accompanying choice of $\mathbf{X} \in \{0,1\}^n$ during the concealment phase and for any two pairs $(\bar{c}, \bar{\mathbf{X}})$, $(\hat{c}, \hat{\mathbf{X}})$, where $\bar{c} \neq \hat{c}$, and $\bar{\mathbf{X}}, \hat{\mathbf{X}} \in \{0,1\}^n$,

$$\max_{s \in \mathcal{S}} \mathbb{P}\left(T(\bar{c}, \bar{\mathbf{X}}, V_B) = 1 \quad \& \quad T(\hat{c}, \hat{\mathbf{X}}, V_B) = 1 \middle| S = s \right) \leq \epsilon \tag{4.3}$$

A rate $R \in [0,1]$ is said to be *achievable* if for every $\epsilon > 0$ and sufficiently large n, there exists an (n, R)-commitment protocol which is ϵ*-sound*, ϵ*-binding* and ϵ*-concealing*. We define the *commitment capacity* or *capacity* as the supremum of all achievable rates.

The main result of this Chapter is to find the commitment capacity of the compound-BSC$[\gamma, \delta]$. To establish our result, we first prove a converse for commitment capacity; note that we prove a converse over general compound DMCs, which can be specialized to the compound-BSC$[\gamma, \delta]$. Then we prove the achievability result for compound-BSC$[\gamma, \delta]$.

4.3 Converse

Theorem 4.6 (Converse). *Consider a general compound channel specified by the conditional law $W_{Y|X,S}$, where $s \in \mathcal{S}$ and $\mathcal{S}$ is an arbitrary set. Then, every sequence of commitment protocols $(\mathscr{P}_n)_{n \in \mathbb{N}}$, where $\forall n$, $\mathscr{P}_n$ is ϵ_n-sound, ϵ_n-concealing and ϵ_n-binding and $\epsilon_n \to 0$ as $n \to \infty$, has rate $R \leq \max_{P_X} \min_{s \in \mathcal{S}} H(X|Y)$.*

Remark 4.7. Note that the max (over P_X) and min (over $s \in \mathcal{S}$) cannot be interchanged in general. In fact, the alternate expression $\min_{s\in\mathcal{S}} \max_{P_X} H(X|Y)$ is generally larger (and hence, is a weaker upper bound). We can show that the latter expression is a tight upper bound for a 'state-aware' scenario where either Alice or both parties know (but cannot control) the compound channel state $s \in \mathcal{S}$ a priori.

Even though the converse is proved for bounded sets $\mathcal{S}$ in this work (recall that $\mathcal{S} = [\gamma, \delta]$ for the compound-BSC$[\gamma, \delta]$), the proof can be extended to arbitrary compound states, i.e., $\mathcal{S}$ can be any arbitrary set. Here we use the standard approach where the converse is first proved for finite $\mathcal{S}$ followed by a discretization procedure (over the arbitrary set $\mathcal{S}$) to extend the converse[3] for general sets $\mathcal{S}$.

The Proof

Consider a sequence of protocols $(\mathscr{P}_n)_{n\geq 1}$. Here protocol $\mathscr{P}_n$, $\forall n$, is ϵ_n-sound, ϵ_n-concealing and ϵ_n-binding for every state $s \in \mathcal{S}$, where $\epsilon_n \geq 0$ and $\epsilon_n \to 0$ as $n \to \infty$.

We now state and prove the following lemma which will be used in our converse.

Lemma 4.8. *For every* $\mathscr{P}_n$, *we have* $H(C|\mathbf{X}, V_B) \leq n\epsilon'_n$, $\forall s \in \mathcal{S}$, *where* $\epsilon'_n \to 0$ *as* $n \to \infty$, *.*

Proof of lemma: Recall that V_B denotes the view of Bob at the end of commit phase.

[3] For arbitrary sets $\mathcal{S}$, the upper bound may be stated as $\max_{P_X} \inf_{s\in\mathcal{S}} H(X|Y)$ if the min is not meaningfully defined.

Let us define[4] $\tilde{c} := \arg\max_{c\in[2^{nR}]} T(\tilde{c}, \mathbf{X}, V_B)$. We now bound $\mathbb{P}(\hat{C} \neq C)$, where $\hat{C} = \hat{C}(V_B, \mathbf{X}) = \tilde{c}$. As the code is ϵ_n-binding for every $s \in \mathcal{S}$, we know that $\forall s \in \mathcal{S}$

$$\mathbb{P}\left(T(\bar{c}, \bar{\mathbf{X}}, V_B) = 1 \quad \& \quad T(\hat{c}, \hat{\mathbf{X}}, V_B) = 1 \middle| s\right) \leq \epsilon_n \tag{4.4}$$

for any two distinct $(\bar{c}, \bar{\mathbf{X}})$ and $(\hat{c}, \hat{\mathbf{X}})$. For the given decoder, we have

$$\mathbb{P}(\hat{C} \neq C) = \mathbb{P}(\hat{C} = 0) + \mathbb{P}(\hat{C} \neq C | C \neq 0) \tag{4.5}$$

$$\leq \epsilon_n + \epsilon_n \tag{4.6}$$

$$= 2\epsilon_n. \tag{4.7}$$

where in the penultimate inequality, the first part follows from noting that $\mathscr{P}_n$ is ϵ_n-binding, and the second part follows from the fact that conditioned on $\mathscr{P}_n$ being ϵ_n-binding, the probability that $\hat{C}$ is different from C is at most ϵ_n due to $\mathscr{P}_n$ being ϵ_n-sound.

We now use Fano's inequality (cf. (El Gamal and Kim, 2011)) to bound the conditional entropy.

$$H(C|\mathbf{X}, V_B) \leq 1 + \mathbb{P}(\hat{C} \neq C)nR \tag{4.8}$$

$$\leq n\left(\frac{1}{n} + 2\epsilon_n R\right) \tag{4.9}$$

$$\leq n\epsilon'_n \tag{4.10}$$

where $\epsilon'_n \to 0$ as $n \to \infty$. This completes the proof of the lemma. ■

[4] Although Bob's test T is a randomized test, it can be shown that one can construct from T a deterministic test with essentially the same soundness and bindingness performance (cf. ()). Hence, for the rest of the converse, we consider that Bob's test is a deterministic function; as such, $\tilde{c}$ is well defined for such a determinstic test.

Let us now bound the rate R. Consider the following:

$$\begin{aligned} nR &= H(C) \\ &= H(C|V_B) + I(C;V_B) \\ &\overset{(a)}{\leq} H(C|V_B) + \epsilon_n \\ &\overset{(b)}{\leq} H(C|\mathbf{Y}_s, K_B, M) + \epsilon_n \\ &\overset{(c)}{=} H(C, \mathbf{X}|\mathbf{Y}_s, K_B, M) - H(\mathbf{X}|\mathbf{Y}_s, K_B, M, C) + \epsilon_n \\ &\overset{(d)}{\leq} H(C, \mathbf{X}|\mathbf{Y}_s, K_B, M) + \epsilon_n \\ &\overset{(e)}{\leq} H(\mathbf{X}|\mathbf{Y}_s, K_B, M) + H(C|\mathbf{X}, \mathbf{Y}_s, K_B, M) + \epsilon_n \\ &= H(\mathbf{X}|\mathbf{Y}_s, K_B, M) + H(C|\mathbf{X}, V_B) + \epsilon_n \\ &\overset{(f)}{\leq} H(\mathbf{X}|\mathbf{Y}_s) + n\epsilon'_n + \epsilon_n \\ &\overset{(g)}{\leq} \sum_{i=1}^{n} H(X_i|Y_{s,i}) + n\epsilon'_n + \epsilon_n \end{aligned} \tag{4.11}$$

Here

(a) follows from the fact that each code $\mathcal{C}_n$ is ϵ_n-concealing.

(b) here we denote by $\mathbf{Y}_s$, the channel output under the state $s \in \mathcal{S}$. We then get (b) by noting that $V_B = (M, K_B, \mathbf{Y}_s)$.

(c) follows from the chain rule of joint entropy

(d) Note that conditional entropy is a positive quantity

(e) follows from the chain rule of joint entropy

(f) follows from the fact that conditioning reduces entropy and Lemma 4.8

(g) follows from the chain rule for conditional entropy

We now introduce a time-sharing random variable $W \sim \text{Unif}(\{1, n\})$ which follows a uniform distribution over the set $\{1, 2, \cdots, n\}$. Then, it follows from (4.11)

$$
\begin{aligned}
nR &\leq n \left(\sum_{i=1}^{n} \frac{1}{n} H(X_i|Y_{s,i}) \right) + n\epsilon'_n + \epsilon_n && (4.12)\\
&\overset{(a)}{=} n \left(\sum_{i=1}^{n} P(W = i) H(X_i|Y_{s,i}) \right) + n\epsilon'_n + \epsilon_n && (4.13)\\
&= nH(X_W|Y_{s,W}, W) + n\epsilon'_n + \epsilon_n && (4.14)\\
&\overset{(b)}{\leq} nH(X_W|Y_{s,W}) + n\epsilon'_n + \epsilon_n && (4.15)\\
&\overset{(c)}{\leq} nH(X|Y_s) + n\epsilon'_n + \epsilon_n && (4.16)
\end{aligned}
$$

(a) follows from the definition of W

(b) follows from noting that conditioning reduces entropy

(c) by defining $X := X_W$, $Y_s = Y_{s,W}$ (noting that $X_W \in \mathcal{X}$ and $Y_{s,W} \in \mathcal{Y}$)

Note that (4.16) holds for every $s \in \mathcal{S}$. Furthermore, we know that $\epsilon_n, \epsilon'_n \to 0$ as $n \to \infty$. Hence, it follows that

$$R \leq \min_{s \in \mathcal{S}} H(X|Y_s)$$

for some appropriate distribution P_X on X. Now optimizing the distribution on $\mathcal{X}$, we have the following bound on R:

$$R \leq \max_{P_X} \min_{s \in \mathcal{S}} H(X|Y) \tag{4.17}$$

Noting that $s \in [\gamma, \delta]$, and solving (4.17), we have

$$R \leq H(\gamma)$$

where we observe that the optimizing $X \sim \textit{Bernoulli}(1/2)$. This completes our converse.

4.4 Achievability

Next, we state the achievability result for the compound-BSC.

Theorem 4.9 (Achievability). *Let $\epsilon > 0$. Then, for every $R < H(\gamma)$ and for blocklength n sufficiently large, there exists a computationally-efficient commitment protocol of rate R which is ϵ-sound, ϵ-concealing and ϵ-binding over the compound-BSC$[\gamma, \delta]$.*

The Proof

Outline: Our achievability uses Damgård et al. (Damgård et al., 1999) classic commitment scheme involving two rounds of random hash excahnge and a strong randomness extractor. The commit phase and the reveal phase are described below:
Commit Phase: Alice wishes to commit to a binary string $c \in [2^{nR}]$ with Bob, and proceeds in the following manner:

- Given $c \in [2^{nR}]$, Alice chooses $\mathbf{X} \in \{0,1\}^n$ uniformly at random and sends it over the compound-BSC.
- Bob recieves the noisy version $\mathbf{Y}$ of the transmitted bit string $\mathbf{X}$. For the recieved binary string $\mathbf{Y} = \mathbf{y}$, Bob calculates the list of candidate binary vectors[5]:

$$\mathcal{L}(\mathbf{y}) := \{\mathbf{x} \in \{0,1\}^n : n(\gamma - \alpha_1) \leq d_H(\mathbf{x}, \mathbf{y}) \leq n(\delta + \alpha_1)\}.$$

- Bob picks a hash function G_1, uniformly at random, from a $4n$-universal hash family $\mathcal{G}_1 := \{g_1 : \{0,1\}^n \to \{0,1\}^{n\beta_1}\}$, and $\beta_1 > 0$ is a small enough constant, and sends a description of G_1 to Alice over the noiseless public channel.

[5] Here the parameter $\alpha_1 > 0$ is chosen appropriately small.

- Alice computes $G_1(\mathbf{X})$ and sends it back to Bob over the noiseless public channel.

- Bob picks another hash function G_2, uniformly at random, from a 2-universal hash family $\mathcal{G}_2 := \{g_2 : \{0,1\}^n \to \{0,1\}^{n\beta_2}\}$, where $\beta_2 > 0$ is a small enough constant, and sends a description of G_2 to Alice over the noiseless public channel.

- Alice computes $G_2(\mathbf{X})$ and sends it back to Bob over the noiseless public channel.

- Alice chooses, uniformly at random, an extractor function Ext from the $2-$universal hash family $\{e : \{0,1\}^n \to \{0,1\}^{n(H(\gamma)-\beta_3)}\}$, where $\beta_3 > 0$ is a constant chosen appropriately. Alice sends $Z = c \oplus \text{Ext}(\mathbf{X})$ (where $\oplus$ denotes component-wise modulo-2 addition) and a description of Ext to Bob over the noiseless public channel.

Reveal phase: In the reveal phase, all announcements are over the public noiseless channel. The parties proceed in the following manner:

- Alice announces the pair $(\tilde{c}, \tilde{\mathbf{x}})$.

- Bob accepts $\tilde{c}$ if the following three conditions hold simultaneously and uniquely for $\tilde{c}$ and $\tilde{\mathbf{x}}$: *(i)* $\tilde{\mathbf{x}} \in \mathcal{L}(\mathbf{y})$, *(ii)* $g_1(\tilde{\mathbf{x}}) = g_1(\mathbf{x})$, $g_2(\tilde{\mathbf{x}}) = g_2(\mathbf{x})$ and *(iii)* $\tilde{c} = z \ \oplus \text{Ext}(\tilde{\mathbf{x}})$. Otherwise, it outputs $\tilde{c} = 0$ to declare error.

Analysis: Intuitively, it is virtually impossible for Alice to reveal a bit string $\tilde{\mathbf{x}}$, different from the one she transmitted over the compound-BSC, without being caught in the first two tests run by Bob. This is guaranteed by the very low probability of the hash values of a new $\tilde{\mathbf{x}}$ matching both $G_1(\mathbf{X})$ and $G_2(\mathbf{X})$. The first hash challenge reduces the number of *confusing* strings that Alice can use to reveal to Bob from exponentially many to polynomially many (in n), while the second hash challenge further reduces such strings to a singleton set. In essence, these hash challenges bind Alice to her original transmitted bit string. To see why Bob's third test is useful, suppose that Alice revealed an arbitrary $\tilde{c}$, $\tilde{c} \neq c$, and $\tilde{\mathbf{x}}$ same as the transmitted string in the commit phase. Since Bob requires $Z \ \oplus \text{Ext}(\tilde{\mathbf{x}})$ to be equal to c, Alice will be caught in the third test run by Bob (as $\tilde{c} \neq c$). Note that $\text{Ext}(\mathbf{X})$ is a nearly random string for Bob. As a result, Bob learns nothing about c from the one-time pad of c and $\text{Ext}(\mathbf{X})$, i.e., Z. Together, these tests ensure that soundness, concealment and bindingness are realised in the protocol. A more detailed analysis now follows. we now analyse the security guarantees for the above defined (n, R) commitment protocol:

1. $\epsilon-$*sound:* For a honest Alice and Bob, the protocol is sound if $\mathbf{X} \in \mathcal{L}(\mathbf{Y})$ with high probability (w.h.p.); to proceed, we analyse the event $\{\mathbf{X} \notin \mathcal{L}(\mathbf{Y})\}$.

$$\begin{aligned}\mathbb{P}\left(\mathbf{X} \notin \mathcal{L}(\mathbf{y})\right) \leq \mathbb{P}&\left(\mathbf{X} : wt_H(\mathbf{X}) \notin \left[n(\frac{1}{2} - \delta_1'), n(\frac{1}{2} + \delta_1')\right]\right) \\ &+ \mathbb{P}\left(\mathbf{Y} : d_H(\mathbf{X}, \mathbf{Y}) \notin [n(\gamma - \delta_1), n(\delta + \delta_1)] \middle| wt_H(\mathbf{X}) \geq \frac{n}{2}(1 - \delta_1')\right)\end{aligned} \tag{4.18}$$

Since $\mathbf{X} \sim$ Bernouulli(1/2) i.i.d., it follows from Chernoff bound that for n sufficiently large

$$\mathbb{P}\left(\mathbf{X} : wt_H(\mathbf{X}) \notin \left[n(\frac{1}{2} - \delta_1'), n(\frac{1}{2} + \delta_1')\right]\right) \leq 2^{-\frac{n\delta_1'^2}{3}} \tag{4.19}$$

As compound-BSC$[\gamma, \delta]$ instantiates a BSC with a transition probability $s \in [\gamma, \delta]$, we have

$$\mathbb{P}\left(\mathbf{Y} : d_H(\mathbf{X}, \mathbf{Y}) \geq n(\delta + \delta_1) \middle| wt_H(\mathbf{X}) \geq \frac{n}{2}(1 - \delta_1')\right) \leq 2^{-\frac{n\delta\delta_1^2}{3}} \tag{4.20}$$

$$\mathbb{P}\left(\mathbf{Y} : d_H(\mathbf{X}, \mathbf{Y}) \leq n(\gamma - \delta_1) \middle| wt_H(\mathbf{X}) \geq \frac{n}{2}(1 - \delta_1')\right) \leq 2^{-\frac{n\gamma\delta_1^2}{2}}. \tag{4.21}$$

Using (4.19), (4.20) and (4.21) in (4.18), we have for n sufficiently large

$$\mathbb{P}\left(\mathbf{X} \notin \mathcal{L}(\mathbf{y})\right) \leq 2^{-\frac{n\gamma\delta_1^2}{4}} \tag{4.22}$$

Hence, for n sufficiently large, it follows that $\mathbb{P}\left(\mathbf{X} \notin \mathcal{L}(\mathbf{y})\right) \leq \epsilon$; which is negligible in n, this shows that the protocol is ϵ-sound.

2. *ϵ-concealing:* Note that any ϵ-concealing commitment protocol with positive rate, where $\epsilon > 0$ is *exponentially decreasing* in blocklength n, satisfies the so-called *capacity-based secrecy* (cf. (Damgard et al., 1998, Def. 3.2)) and vice versa. In our proof, we leverage a well known equivalence between capacity-based secrecy and another notion of secrecy called *biased-based secrecy* (cf. (Damgard et al., 1998, Def. 3.1)). In particular, we first show that our commitment protocol has biased-secrecy; here we lower bind the term $H_\infty(\mathbf{X}|\mathbf{Y}, G_1(\mathbf{X}), G_1, G_2(\mathbf{X}), G_2)$ and then crucially use the generalized left-over hash lemma (cf. (Dodis et al., 2004)). Then, we use (Damgard et al., 1998, Th. 4.1) to conclude that our protocol satisfies capacity-based secrecy which implies that our protocol is ϵ-concealing for any given ϵ for n chosen sufficiently large.

To begin, consider the following smooth-min-entropy:

$$
\begin{aligned}
&H_\infty^{\epsilon_1}(\mathbf{X}|\mathbf{Y}, G_1(\mathbf{X}), G_1, G_2(\mathbf{X}), G_2)\\
&\overset{(a)}{\geq} H_\infty(\mathbf{X}, G_1(\mathbf{X}), G_2(\mathbf{X})|\mathbf{Y}, G_1, G_2)\\
&\qquad - H_0(G_1(\mathbf{X}), G_2(\mathbf{X})|\mathbf{Y}, G_1, G_2) - \log(\epsilon_1^{-1})\\
&\overset{(b)}{=} H_\infty(\mathbf{X}|\mathbf{Y}, G_1, G_2) + H_\infty(G_1(\mathbf{X}), G_2(\mathbf{X})|\mathbf{Y}, G_1, G_2, \mathbf{X})\\
&\qquad - H_0(G_1(\mathbf{X}), G_2(\mathbf{X})|\mathbf{Y}, G_1, G_2) - \log(\epsilon_1^{-1})\\
&\overset{(c)}{=} H_\infty(\mathbf{X}|\mathbf{Y}, G_1, G_2)\\
&\qquad - H_0(G_1(\mathbf{X}), G_2(\mathbf{X})|\mathbf{Y}, G_1, G_2) - \log(\epsilon_1^{-1})\\
&\overset{(d)}{=} H_\infty(\mathbf{X}|\mathbf{Y}) - H_0(G_1(X), G_2(X)|\mathbf{Y}, G_1, G_2) - \log(\epsilon_1^{-1})\\
&\overset{(e)}{\geq} n(H(\gamma) - \zeta) - H_0(G_1(\mathbf{X})|G_2(\mathbf{X}), \mathbf{Y}, G_1, G_2)\\
&\qquad - H_0(G_2(\mathbf{X})|\mathbf{Y}, G_1, G_2) - \log(\epsilon_1^{-1})\\
&\overset{(f)}{\geq} n(H(\gamma) - \zeta) - n(\beta_1 + \beta_2) - \log(\epsilon_1^{-1})\\
&= n(H(\gamma) - \zeta - \beta_1 - \beta_2) - \log(\epsilon_1^{-1}) \qquad (4.23)
\end{aligned}
$$

Here,

(a) follows from the chain rule of smooth min-entropy (2.1).

(b) follows from the chain rule of min-entropy (2.1).

(c) follows from the fact that $G_1(\mathbf{X})$ and $G_2(\mathbf{X})$ are deterministic functions of G_1, G_2 and $\mathbf{X}$.

(d) follows from the Markov chain $\mathbf{X} \to \mathbf{Y} \to (G_1, G_2)$.

(e) follows from the chain rule for max-entropy (2.2) and the fact that the minimum crossover probability of the (memoryless) Compound-BSC channel is γ (this results in $H_\infty(\mathbf{X}|\mathbf{Y}) \geq nH(\gamma - \zeta)$).

(f) follows from the definition of max-entropy (and noting that the range of G_1 and G_2 is $\{0,1\}^{n\beta_1}$ and $\{0,1\}^{n\beta_2}$ respectively).

Then, from (4.23), the definitions of $H_\infty^{\epsilon_1}(\cdot)$ and $H_\infty(\cdot)$, and Lemma A.2 (in Appendix A) we get

$$H_\infty(\mathbf{X}|\mathbf{Y}, H_1(\mathbf{X}), H_1, H_2(\mathbf{X}), H_2) \geq H_\infty^{\epsilon_1}(\mathbf{X}|\mathbf{Y}, H_1(\mathbf{X}), H_1, H_2(\mathbf{X}), H_2) - \epsilon_1' \tag{4.24}$$

$$\geq n(H(\gamma) - \zeta - \gamma_1 - \gamma_2 - \gamma' - \gamma'') \tag{4.25}$$

for some $\epsilon_1' := 2^{-n\gamma_1'}$, $\gamma_1' > 0$, where $\epsilon_1' \to 0$ as $\epsilon_1 \to 0$.

Lemma 4.10 (Generalized leftover hash lemma). *Let* $\{G_x : \{0,1\}^n \to \{0,1\}^l\}_{x\in\mathcal{X}}$ *be a family of universal hash functions. Then, for any joint distribution* (W, I)*, we have*

$$\|(P_{G_X(W),X,I} - P_{U_l,X,I})\| \leq \frac{1}{2}\sqrt{2^{-H_\infty(W|I)2^l}} \tag{4.26}$$

where $U_l \sim$ *Unif*$\big(\{0,1\}^l\big)$.

Now, we use the generalized leftover hash lemma (4.26) () to prove the security of the key, Ext($\mathbf{X}$) against Bob, by showing that it is statistically close to a uniform distribution and therefore achieves the bias-based secrecy. Let us fix $\epsilon_1 := 2^{-n\gamma'}$, where $\gamma' > 0$ is an arbitrary small constant. we make the following correspondence: $X \leftrightarrow$ Ext, $W \leftrightarrow \mathbf{X}$, $I \leftrightarrow (\mathbf{Y}, G_1(\mathbf{X}), G_1, G_2(\mathbf{X}), G_2)$. we now have:

$$\begin{aligned}
&\|(P_{\mathrm{Ext}(\mathbf{X}),Ext,\mathbf{Y},G_1(\mathbf{X}),G_1,G_2(\mathbf{X}),G_2} - P_{U_l,\mathrm{EXT},\mathbf{Y},G_1(\mathbf{X}),G_1,G_2(\mathbf{X}),G_2})\| \\
&\quad\leq \frac{1}{2}\sqrt{2^{-H_\infty(\mathbf{X}|\mathbf{Y},G_1(\mathbf{X}),G_1,G_2(\mathbf{X}),G_2)2^l}} \\
&\quad\overset{(a)}{\leq} \frac{1}{2}\sqrt{2^{-n(H(\gamma)-\zeta-\beta_1-\beta_2-\gamma')}2^{n(H(\gamma)-\beta_3))}} \\
&\quad= \frac{1}{2}\sqrt{2^{n(\zeta+\beta_1+\beta_2+\gamma'-\beta_3))}} \\
&\quad\overset{(b)}{\leq} 2^{-n\delta'}
\end{aligned} \tag{4.27}$$

Here,

(a) follows from (4.23) and noting the choice of 2-universal hash function Ext, chosen uniformly from the class $e : \{0,1\}^n \to \{0,1\}^{n(H(\gamma)-\beta_3)}$. .

(b) follows from noting that β_3 is choosen such that $\beta_3 > \beta_1 + \beta_2$ and ζ and η can be chosen arbitrarily small for sufficiently large n.

As the extractor Ext is chosen uniformly at random from the class e of 2-univesal hash family, the generalized leftover hash lemma (cf. 4.10) guarantees that we can extract $n(H(\gamma) - \beta_3)$ almost nearly uniform random bits. Thus, from (4.27), it follows that our commitment protocol satisfies biased-secrecy (cf. (Damgard et al., 1998, Def. 3.1)). We now use (Damgard et al., 1998, Th. 4.1) to conclude that our commitment scheme satisfies capacity-based secrecy (cf. (Damgard et al., 1998, Def. 3.2)). As capacity-based secrecy implies that $I(c; V_B)$ is decreasing exponentially w.r.t. n, it follows that for n sufficiently large, our protocol is ϵ-concealing.

3. ϵ-binding: Here, we need to guarantee that under *any* behaviour of Alice, Bob is able to verify (with high probability) if Alice's reveal choice $(\tilde{c}, \tilde{\mathbf{x}})$ correspond to it's choices in the commit phase or are different. We show that from Bob's perspective, a dishonest Alice 'appears' as capable as one over the UNC$[\gamma, \delta]$ (recall that in a UNC, unlike in the compound-BSC$[\gamma, \delta]$, a dishonest Alice knows and can control the channel transition probability). Once this correspondence is established, the analysis for ϵ-binding for our commitment protocol follows exactly along the lines of that in (Crépeau et al., 2020).

   Let a potentially dishonest Alice transmit $\mathbf{X} = \mathbf{x}$ in the commit phase and let Bob receive $\mathbf{Y} = \mathbf{y}$. Note that $n(\gamma - \alpha_1) \leq d_H(\mathbf{x}, \mathbf{y}) \leq n(\delta + \alpha_1)$. Alice can successfully cheat (to confuse Bob) in the reveal phase if she finds two different strings $\mathbf{x}'$, $\bar{\mathbf{x}}$ such that the following two conditions hold simultaneously: (i) Hamming-distance condition: $n(\gamma - \alpha_1) \leq d_H(\mathbf{x}', \mathbf{y}), d_H(\bar{\mathbf{x}}, \mathbf{y}) \leq n(\delta + \alpha_1)$ and

(ii) Hash-challenge condition: $\mathbf{x}'$, $\bar{\mathbf{x}}$ satisfy the hash function conditions (Bob knows the hash functions and hash values corresponding to $\mathbf{x}$ from Alice in the commit phase). From Bob's perspective, the 'worst' scenario is one where the set of Alice's candidate codewords, say $\{\mathbf{x}''\}$, which satisfy the Hamming distance condition is the largest; it is not too hard to see that the compound-BSC$[\gamma, \delta]$ state $s = \gamma$ instantiates this scenario (on the other hand, $s = \delta$ would have resulted in the 'smallest' such set of candidate codewords). But, such a situation at Bob is exactly the one involving a dishonest Alice over the UNC$[\gamma, \delta]$, where a dishonest Alice may 'actively' fix the channel state of the UNC to γ.

The number of such strings that Alice can use to confuse Bob and that can pass the Bob's first test in the reveal phase (viz. $\mathbf{X} \in \mathcal{L}(\mathbf{Y})$) are exponentially many in n, upper bounded by $2^{n\eta)}$, where $\eta > 0$ for sufficiently large n. The first round of random hash exchange reduces the number of such confusing strings from exponentially many to a polynomially many in n. Fix a $G_1(X) \in \{0,1\}^{n(H(\kappa)+\beta_1)}$ and for the i^{th} confusable bit string, let's define an indicator random variable M_i and $M = \sum_i M_i$, such that M_i equals 1 if the i^{th} confusable string maps to $G_1(X)$, transmitted in the commit phase and $M_i = 0$ otherwise. Noting that $\beta_1 \geq \eta$, we have $\mathbf{E}[M] < 1$.Now, we use the following result by Rompel ():

Lemma 4.11. *Let* $X_1, X_2, X_3....X_m \in \{0,1\}$ *be t-wise independent random variable, where t is an even and positive integer. Let* $X := \sum_i^m X_i$, $\mu := \mathbf{E}[X]$, *and* $A > 0$ *be a constant. Then,*

$$\mathbb{P}\left(|X - \mu| > A\right) < O\left(\left(\frac{t\mu + t^2}{A^2}\right)^{t/2}\right) \tag{4.28}$$

we make the following correspondence: $t \leftrightarrow 4n$, $A \leftrightarrow 2t = 8n$. we now have:

$$P[M > 8n+1] < O\Big(\Big(\frac{t\mu + t^2}{(2t)^2}\Big)^{t/2}\Big) \tag{4.29}$$

$$< O\Big(\Big(\frac{1+t}{4t}\Big)^{t/2}\Big) \tag{4.30}$$

$$< O((2)^{-t/2}) \tag{4.31}$$

4.5 Result - Commitment Capacity of Compound BSC

Using the above two results, we state the commitment capacity of the compound-BSC$[\gamma, \delta]$ in the following theorem.

Theorem 4.12 (Compound-BSC$[\gamma, \delta]$ commitment capacity). *The commitment capacity of the compound binary symmetric channel (compound-BSC), is given by*

$$\mathbb{C} = H(\gamma) \tag{4.32}$$

Chapter 5

Conclusion

5.1 Summary

The reliability of currently popular cryptographic protocols such as the RSA depend largely on computational constraints (no poly time algorithm to factor large prime numbers). An information theoretically secure protocol like the bit commitment doesn't depend on such constraints. Studying them is thus of useful. In Chapter 4 we studied commitment capacity over general discrete memoryless channels and arrived at a dual expression for it. It is interesting to observe that there can exist multiple input symbol probability distributions all achieving the same maximum commitment rate. In Chapter 5, we model the compound BSCs and derive a commitment rate expression for it. We also present a a rate achieving commitment protocol. We see that this capacity ($h(\gamma)$) is strictly higher than that of a closely similar Unfair Noisy Channel ($UNC[\gamma, \delta]$ has capacity $h(\gamma) - h(\frac{\delta-\gamma}{1-2\gamma})$).

5.2 Future Prospects

While compound BSC is an interesting channel to look at, it looks like it belongs to a more general class of channels that includes simple BSCs, UNCs and elastic channels. We have recently been able to expand the scope to this general class via a model of generalised UNCs. We are currently studying the commitment problem over such class of channels.

Appendix A

Useful Entropy Relations

Lemma A.1.

$$H(P) - \alpha \mathbb{E}_P\{f(X)\} \leq \log\left(\sum_x \exp(-\alpha f(x))\right) \tag{A.1}$$

where $f : \mathcal{X} \to \mathbb{R}$ and $\alpha \in \mathbb{R}$. The maximising distribution P for the equality is

$$P(x) = \frac{1}{A}\exp(-\alpha f(x)) \tag{A.2}$$

$$A = \sum_x \exp(-\alpha f(x)) \tag{A.3}$$

The proof uses the log sum inequality as has been done in Csiszár and Körner (2011). Refer Ch 2 Csiszár and Körner (2011) for details.

A.1 Continuity of Smooth Entropies

Lemma A.2. *The min entropy $\mathbb{H}_\infty(X)$ and the smooth min entropy $\mathbb{H}_\infty^\varepsilon(X)$ of a disctrete random variable X with some probability distribution P_X converge with* *converging* ε

$$\mathbb{H}_{\infty}^{\varepsilon}(X) - \frac{\varepsilon}{\ln(2)(\max_{x\in\mathcal{X}}\ P(x) - \varepsilon)} \leq \mathbb{H}_{\infty}(X) \leq \mathbb{H}_{\infty}^{\varepsilon}(X) \tag{A.4}$$

Proof.

$$\mathbb{H}_{\infty}(X) \triangleq \min_{x\in\mathcal{X}}\ \log\frac{1}{P(x)} \tag{A.5}$$

$$\mathbb{H}_{\infty}^{\varepsilon}(X) \triangleq \max_{X'\sim P':SD(P,P')\leq\varepsilon}\ \mathbb{H}_{\infty}(X') \tag{A.6}$$

Now, let us define

$$P^* \triangleq \operatorname*{arg\,max}_{X'\sim P':SD(P,P')\ \leq\varepsilon}\ \mathbb{H}_{\infty}(X') \tag{A.7}$$

$$|P - P^*| \leq \varepsilon \quad \forall x \in \mathcal{X} \tag{A.8}$$

So that

$$\mathbb{H}_{\infty}^{\varepsilon}(X) = \min_{x\in\mathcal{X}}\ \log\frac{1}{P^*(x)} \tag{A.9}$$

Also let us define

$$x_1 \triangleq \operatorname*{arg\,min}_{x\in\mathcal{X}}\ \log\frac{1}{P(x)} \tag{A.10}$$

$$= \operatorname*{arg\,max}_{x\in\mathcal{X}}\ P(x) \tag{A.11}$$

$$x_2 \triangleq \operatorname*{arg\,min}_{x\in\mathcal{X}}\ \log\frac{1}{P^*(x)} \tag{A.12}$$

$$= \operatorname*{arg\,max}_{x\in\mathcal{X}}\ P^*(x) \tag{A.13}$$

So that

$$\mathbb{H}_{\infty}(X) = \log\frac{1}{P(x_1)} \tag{A.14}$$

$$\mathbb{H}_{\infty}^{\varepsilon}(X) = \log\frac{1}{P^*(x_2)} \tag{A.15}$$

From definition of $\mathbb{H}_\infty^\varepsilon(X)$

$$\begin{aligned} \mathbb{H}_\infty(X) &\leq \mathbb{H}_\infty^\varepsilon(X) && \text{(A.16)} \\ \Rightarrow \quad \log \frac{1}{P(x_1)} &\leq \log \frac{1}{P^*(x_2)} && \text{(A.17)} \\ \Rightarrow \quad P(x_1) &\geq P^*(x_2) && \text{(A.18)} \end{aligned}$$

From A.8, A.11, A.13 and A.18

$$\begin{aligned} P(x_1) &\leq P^*(x_1) + \varepsilon && \text{(A.19)} \\ &\leq P^*(x_2) + \varepsilon && \text{(A.20)} \end{aligned}$$

Now looking at $\mathbb{H}_\infty(X)$ from A.14

$$\begin{aligned} \mathbb{H}_\infty(X) &= \log \frac{1}{P(x_1)} && \text{(A.21)} \\ &\geq \log \left(\frac{1}{P^*(x_2) + \varepsilon} \right) && \text{(A.22)} \\ &= \log \left(\frac{1}{P^*(x_2)} \right) - \log \left(\frac{P^*(x_2) + \varepsilon}{P^*(x_2)} \right) && \text{(A.23)} \\ &= \mathbb{H}_\infty^\varepsilon(X) - \log \left(1 + \frac{\varepsilon}{P^*(x_2)} \right) && \text{(A.24)} \\ &\geq \mathbb{H}_\infty^\varepsilon(X) - (\log e) \left(\frac{\varepsilon}{P^*(x_2)} \right) && \text{(A.25)} \\ &= \mathbb{H}_\infty^\varepsilon(X) - \frac{1}{\ln(2)} \frac{\varepsilon}{P^*(x_2)} && \text{(A.26)} \\ &\geq \mathbb{H}_\infty^\varepsilon(X) - \frac{\varepsilon}{\ln(2)(P(x_1) - \varepsilon)} && \text{(A.27)} \\ &= \mathbb{H}_\infty^\varepsilon(X) - \frac{\varepsilon}{\ln(2)(\max_{x \in \mathcal{X}} \ P(x) - \varepsilon)} && \text{(A.28)} \\ & && \text{(A.29)} \end{aligned}$$

From A.16 and A.28

$$\mathbb{H}_\infty^\varepsilon(X) - \frac{\varepsilon}{\ln(2)(\max_{x \in \mathcal{X}} \ P(x) - \varepsilon)} \leq \mathbb{H}_\infty(X) \leq \mathbb{H}_\infty^\varepsilon(X) \qquad \text{(A.30)}$$

For finite set $\mathcal{X}$, $\max_{x\in\mathcal{X}} P(x)$ has to be non negligible and thus the bounds converge.

$\square$

Appendix B

Verification of BSC Primal and Dual Capacities

B.1 Primal BSC capacity

For the Binary case we will set the cost of input symbols $\rho_X(0) = 0$ and $\rho_X(1) = 1$. We will also parametrise the input distribution P_X and the channel which is a BSC with p probability of flipping.

$$\begin{aligned} P_X(0) &= 1 - t \\ P_X(1) &= t \end{aligned} \tag{B.1}$$

$$\begin{aligned} W_{Y|X}(y|x) &= 1 - p \quad \text{if } y = x \\ &= p \qquad\;\; \text{if } y = x \oplus 1 \end{aligned} \tag{B.2}$$

with,

$$0 < t < \frac{1}{2} \tag{B.3}$$
$$0 < p < \frac{1}{2} \tag{B.4}$$
$$0 < \Gamma < \frac{1}{2} \tag{B.5}$$

Let us solve $\mathbb{C}(\Gamma)$ using the primal expression.

$$\begin{aligned} \mathbb{C}(\Gamma) &= \max_{P_X:\rho_X(P_X)\leq\Gamma} \big[H(X|Y)\big] \qquad \text{(B.6)} \\ &= \max_{P_X:\mathbb{E}[\rho_X(P_X)]\leq\Gamma} \big[H(Y|X) + H(X) - H(Y)\big] \qquad \text{(B.7)} \end{aligned}$$

For the binary case, the output follows a Bernoulli($t \circledast p$) distribution. The optimisation constraint can also be simplified as

$$\begin{aligned} &\Gamma \geq \mathbb{E}[\rho_X(P_X)] = 0(1-t) + 1t \\ &\Rightarrow \Gamma \geq t \qquad \text{(B.8)} \end{aligned}$$

Therefore, it follows that

$$\begin{aligned} \mathbb{C}(\Gamma) &= \max_t \big[H_2(p) + H_2(t) - H_2(t \circledast p)\big] \\ &\text{s.t.} \qquad t \leq \Gamma \qquad \text{(B.9)} \\ \Leftrightarrow \mathbb{C}(\Gamma) &= -\min_t \big[-H_2(p) - H_2(t) + H_2(t \circledast p\big]) \\ &\text{s.t.} \qquad t \leq \Gamma \qquad \text{(B.10)} \end{aligned}$$

The optimisation expression $-H(X|Y)$ is convex and the constraint $t \leq \Gamma$ is linear w.r.t parameter t. So, we can use a Lagrange optimiser to solve this problem

$$\mathcal{L} = -H_2(p) - H_2(t) + H_2(p \circledast t) + \lambda(t - \Gamma) \tag{B.11}$$

According to the KKT conditions at the minimising value t',

$$\left.\frac{\partial \mathcal{L}}{\partial t}\right|_{t=t'} = 0 \quad \& \quad t' - \Gamma = 0 \quad \& \quad \lambda > 0 \tag{B.12}$$

$$(or)$$

$$\left.\frac{\partial \mathcal{L}}{\partial t}\right|_{t=t'} = 0 \quad \& \quad t' - \Gamma < 0 \quad \& \quad \lambda = 0 \tag{B.13}$$

$$\tag{B.14}$$

Applying these KKT conditions on the common condition $\left.\frac{\partial \mathcal{L}}{\partial t}\right|_{t=t'} = 0$

$$\begin{aligned} \left.\frac{\partial}{\partial t}\left[H_2(p) + H_2(t) - H_2(p \circledast t) - \lambda(t - \Gamma)\right]\right|_{t=t'} &= 0 \\ \Rightarrow \quad 0 - \log\left(\frac{t'}{1-t'}\right) + (1-2p)\log\left(\frac{p \circledast t'}{1 - p \circledast t'}\right) - \lambda &= 0 \end{aligned}$$

$$\begin{aligned} \Rightarrow \quad \lambda &= -\log\left(\frac{t'}{1-t'}\right) + (1-2p)\log\left(\frac{p \circledast t'}{1 - p \circledast t'}\right) \\ \Rightarrow \quad \lambda &= -(1-2p)\left[\log\left(\frac{t'}{1-t'}\right) - \log\left(\frac{p \circledast t'}{1 - p \circledast t'}\right)\right] - 2p\log\left(\frac{t'}{1-t'}\right) \end{aligned} \tag{B.15}$$

Here let us look at both the terms seperately. From constraints we imposed on the input we know that $\Gamma < \frac{1}{2}$, $t < \frac{1}{2}$. Hence,

$$\begin{aligned} t' &< \frac{1}{2} \\ \frac{t'}{1-t'} &< 1 \\ \log_2\left(\frac{t'}{1-t'}\right) &< 0 \end{aligned}$$

That means term II in (B.15) is always positive. Now let us look at term I. Here we will use the property that $p > 0$.

$$\begin{aligned} p \circledast t' &= p + t' - 2pt' && \text{(B.16)} \\ &= t' + p(1-2t') \quad > \quad t' && \text{(B.17)} \\ p \circledast t' &> t' && \text{(B.18)} \\ \overset{(a)}{\Rightarrow} \qquad \frac{p \circledast t'}{1 - p \circledast t'} &\geq \frac{t'}{1-t'} && \text{(B.19)} \\ \overset{(b)}{\Rightarrow} \quad \log\left(\frac{p \circledast t'}{1 - p \circledast t'}\right) &\geq \log\left(\frac{t'}{1-t'}\right) && \text{(B.20)} \end{aligned}$$

(a) Follows from the fact that $\frac{x}{1-x}$ is an increasing function. (b) Follows from the increasing nature of the log function. That result is that term I of equation (B.15) is always non negative. It can be deduced from this and the previous result that $\lambda > 0$. That given, it follows from the KKT condition that $t' = \Gamma$ is the only solution for the optimisation problem. Substituting this back into the capactiy expression we get.

$$\mathbb{C}(\Gamma) = H_2(p) + H_2(\Gamma) - H_2(p \circledast \Gamma) \tag{B.21}$$

This is the capacity for the binary input BSC channel, we estimated using the primal expression.

B.2 Dual BSC capacity

Recall that $\rho_X(0) = 0$ and $\rho_X(1) = 1$. Here we parametrise the output distribution Q_Y and the Binary Symmetric Channel.

$$\begin{aligned} Q_Y(0) &= 1 - q \\ Q_Y(1) &= q \end{aligned} \tag{B.22}$$

$$\begin{aligned} W_{Y|X}(y|x) &= 1 - p \quad \text{if } y = x \\ &= p \qquad\;\;\, \text{if } y = x \oplus 1 \end{aligned} \tag{B.23}$$

with,

$$0 < q < \frac{1}{2} \tag{B.24}$$

$$0 < p < \frac{1}{2} \tag{B.25}$$

$$0 < \Gamma < \frac{1}{2} \tag{B.26}$$

Let us solve $\mathbb{C}(\Gamma)$ using the dual expression. Recall from (3.27)

$$\mathbb{C}(\Gamma) = \min_{\gamma \geq 0} \; \left[F(\gamma) + \gamma\Gamma \right] \tag{B.27}$$

$$F(\gamma) = \max_{Q_Y} \left[\log \left(\sum_{x \in \mathcal{X}} \exp(-D(W_{Y|X}(\cdot|x)||Q_Y) - \gamma\rho_X(x)) \right) \right] \tag{B.28}$$

Consider the following,

$$\begin{aligned}
D(W_{Y|X}(\cdot|0)||Q) &= (1-p)\log\frac{1-p}{1-q} + p\log\frac{p}{q} \\
&= \log\frac{(1-p)^{1-p}}{(1-q)^{1-p}}\frac{p^p}{q^q} \\
D(W_{Y|X}(\cdot|1)||Q) &= p\log\frac{p}{1-q} + (1-p)\log\frac{1-p}{q} \\
\Rightarrow \quad \exp(-D(W_{Y|X}(\cdot|x)||Q) - \gamma\rho_X(x)) &= \exp(-D(W_{Y|X}(\cdot||x)||Q)\exp(-\gamma\rho_X(x)) \\
\Rightarrow \quad \sum_{x\in\mathcal{X}}\exp(-D(W_{Y|X}(\cdot|x)||Q) - \gamma\rho_X(x)) &= \sum_{x\in\mathcal{X}}\exp(-D(W_{Y|X}(\cdot||x)||Q)\exp(-\gamma x(x)) \\
&= \frac{(1-q)^{1-p}}{(1-p)^{1-p}}\frac{q^q}{p^p} + \frac{q^{1-p}}{(1-p)^{1-p}}\frac{(1-q)^q}{p^p}\exp(-\gamma)
\end{aligned} \tag{B.29}$$

Let us define parameters α and z which are functions of γ and q respectively.

$$\alpha \triangleq exp(-\gamma) \tag{B.30}$$

$$z \triangleq \frac{1-q}{q} \tag{B.31}$$

$$\text{Note that} \quad \frac{dz}{dq} = -\frac{1}{q^2} \tag{B.32}$$

Substituting (B.30, B.29) in (B.28)

$$F(\gamma) = \max_q \sum_{x\in\mathcal{X}}\log\frac{1}{(1-p)^{1-p}p^p}\left[(1-q)^{1-p}q^p + \alpha(1-q)^p q^{1-p}\right] \tag{B.33}$$

$$= \max_q \left[\sum_{x\in\mathcal{X}}[\log\frac{1}{(1-p)^{1-p}}] + \sum_{x\in\mathcal{X}}\log\left[(1-q)^{1-p}q^p + \alpha(1-q)^p q^{1-p}\right]\right] \tag{B.34}$$

$$= H_2(p) + \log\max_q\left[(1-q)^{1-p}q^p + \alpha(1-q)^p q^{1-p}\right] \tag{B.35}$$

$$= H_2(p) + \log G(\gamma) \tag{B.36}$$

where

$$G(\gamma) \quad \triangleq \max_q G(\gamma, q) \tag{B.37}$$

$$G(\gamma, q) \triangleq (1-q)^{1-p}q^p + \alpha(1-q)^p q^{1-p} \tag{B.38}$$

$G(\gamma, q)$ can be simplified as

$$G(\gamma, q) = (1-q)^{\frac{1}{2}} q^{\frac{1}{2}} \big[\big(\frac{1-q}{q}\big)^{\frac{1}{2}-p} + \alpha \big(\frac{1-q}{q}\big)^{p-\frac{1}{2}} \big] \tag{B.39}$$

$$= (1-q)^{\frac{1}{2}} q^{\frac{1}{2}} \big[z^{\frac{1}{2}-p} + \alpha z^{p-\frac{1}{2}} \big] \tag{B.40}$$

To determine $\max_q G(\gamma, q)$ we equate $\frac{\partial G(\gamma,q)}{\partial q}$ to zero and check. For now let's assume the derivative becomes zero at $\tilde{q}$. Let $\tilde{z}$ be the corresponding z. $\tilde{z} = z(\tilde{q})$

$$\left. \frac{\partial G(\gamma, q)}{\partial q} \right|_{q=\tilde{q}} = 0 \tag{B.41}$$

$$\tag{B.42}$$

$$\Rightarrow\ 0 = \frac{1}{2} \frac{1-2\tilde{q}}{(1-\tilde{q})^{\frac{1}{2}} \tilde{q}^{\frac{1}{2}}} \left[\tilde{z}^{\frac{1}{2}-p} + \alpha \tilde{z}^{p-\frac{1}{2}} \right] + (1-\tilde{q})^{\frac{1}{2}} \tilde{q}^{\frac{1}{2}} \left[\big(\frac{1}{2} - p\big) \tilde{z}^{\frac{1}{2}-p} + \alpha \tilde{z}^{p-\frac{1}{2}} \right] \frac{-1}{\tilde{q}^2}$$

$$= \frac{1}{2} \frac{1-2\tilde{q}}{(1-\tilde{q})^{\frac{1}{2}} \tilde{q}^{\frac{1}{2}}} \left[\tilde{z}^{\frac{1}{2}-p} + \alpha \tilde{z}^{p-\frac{1}{2}} \right] + \frac{((1-\tilde{q})\tilde{q})^{\frac{1}{2}}}{\tilde{q}^2} \frac{\frac{1}{2} - p}{\tilde{z}} \left[\tilde{z}^{\frac{1}{2}-p} - \alpha \tilde{z}^{p-\frac{1}{2}} \right] \tag{B.43}$$

$$= \frac{1-2\tilde{q}}{2(1-\tilde{q})^{\frac{1}{2}} \tilde{q}^{\frac{1}{2}}} \left[\tilde{z}^{\frac{1}{2}-p} + \alpha \tilde{z}^{p-\frac{1}{2}} \right] + \frac{1-2p}{2(1-\tilde{q})^{\frac{1}{2}} \tilde{q}^{\frac{1}{2}}} \left[\tilde{z}^{\frac{1}{2}-p} - \alpha \tilde{z}^{p-\frac{1}{2}} \right] \tag{B.44}$$

$$\Rightarrow 0 = (1-2\tilde{q}) \big[\tilde{z}^{\frac{1}{2}-p} + \alpha \tilde{z}^{p-\frac{1}{2}} \big] + (1-2p) \big[\tilde{z}^{\frac{1}{2}-p} - \alpha \tilde{z}^{p-\frac{1}{2}} \big] \tag{B.45}$$

$$= \alpha \tilde{z}^{p-\frac{1}{2}} \big[2 - 2\tilde{q} - 2p \big] - \tilde{z}^{\frac{1}{2}-p} \big[2\tilde{q} - 2p \big] \tag{B.46}$$

$$\Rightarrow \alpha = \tilde{z}^{1-2p} \frac{\tilde{q} - p}{1 - \tilde{q} - p} \tag{B.47}$$

Note that from our defintion of α in (B.30) and the conditions (B.25, B.24), its only feasible range $\alpha \in (0, 1]$. Also note that

$$\tilde{z} \in (1, \infty) \tag{B.48}$$

$$\tilde{q} \in (0, \frac{1}{2}) \tag{B.49}$$

Now,

$$
\begin{aligned}
G(\gamma) &= \max_q G(\gamma, q) && \text{(B.50)}\\
&= G(\gamma, \tilde{q}) && \text{(B.51)}\\
&= (\tilde{q}(1-\tilde{q}))^{\frac{1}{2}} \big[\tilde{z}^{\frac{1}{2}-p} + \alpha \tilde{z}^{p-\frac{1}{2}}\big] && \text{(B.52)}\\
&= (\tilde{q}(1-\tilde{q}))^{\frac{1}{2}} \big[\tilde{z}^{\frac{1}{2}-p} + \tilde{z}^{1-2p}\frac{\tilde{q}-p}{1-\tilde{q}-p}\tilde{z}^{p-\frac{1}{2}}\big] && \text{(B.53)}\\
&= (\tilde{q}(1-\tilde{q}))^{\frac{1}{2}} \big[\tilde{z}^{\frac{1}{2}-p} + \tilde{z}^{\frac{1}{2}-p}\frac{\tilde{q}-p}{1-\tilde{q}-p}\big] && \text{(B.54)}\\
&= (\tilde{q}(1-\tilde{q}))^{\frac{1}{2}} \tilde{z}^{\frac{1}{2}-p}\frac{1-2p}{1-\tilde{q}-p} && \text{(B.55)}\\
&= (\tilde{q}(1-\tilde{q}))^{\frac{1}{2}} \big(\frac{1-\tilde{q}}{\tilde{q}}\big)^{\frac{1}{2}-p}\frac{1-2p}{1-\tilde{q}-p} && \text{(B.56)}\\
&= \tilde{q}^p(1-\tilde{q})^{1-p}\frac{1-2p}{1-\tilde{q}-p} && \text{(B.57)}
\end{aligned}
$$

We have from the original definitions

$$
\begin{aligned}
F(\gamma) &= H_2(p) + logG(\gamma) && \text{(B.58)}\\
\mathbb{C}(\Gamma) &= \min_{\gamma\geq 0}\ F(\gamma) + \gamma\Gamma && \text{(B.59)}\\
&= \min_{\gamma\geq 0}\ \big[H_2(p) + \log G(\gamma) + \gamma\Gamma\big] && \text{(B.60)}\\
&= H_2(p) + \min_{\tilde{q}}\ \big[\log\big[\tilde{q}^p(1-\tilde{q})^{1-p}\frac{1-2p}{1-\tilde{q}-p}\big] + \gamma\Gamma\big] && \text{(B.61)}\\
&= H_2(p) + E && \text{(B.62)}
\end{aligned}
$$

Where E is defined as follows.

$$
E \triangleq \min_{\tilde{q}\in(0,\frac{1}{2})} E(\tilde{q}) \qquad \text{(B.63)}
$$

Note that we have replaced the minimisation over variable γ with over variable $\tilde{q}$. This is valid for the given input ranges because $\tilde{q}$ is solely a function of γ.

Consequently, the inequalities $\gamma \geq 0$ and $\tilde{q} \in (0, \frac{1}{2})$ span over the same spaces.

$$\begin{aligned}
E(\tilde{q}) &= \log\big[\tilde{q}^p(1-\tilde{q})^{1-p}\frac{1-2p}{1-\tilde{q}-p}\big] + \gamma\Gamma && \text{(B.64)}\\
&= p\log\tilde{q} + (1-p)\log(1-\tilde{q}) + \log(1-2p) - \log(1-\tilde{q}-p) - \log\alpha^\Gamma && \text{(B.65)}\\
&= p\log\tilde{q} + (1-p)\log(1-\tilde{q}) + \log(1-2p) - \log(1-\tilde{q}-p) - \log\big(\tilde{z}^{1-2p}\frac{\tilde{q}-p}{1-\tilde{q}-p}\big)^\Gamma\\
&= p\log\tilde{q} + (1-p)\log(1-\tilde{q}) + \log(1-2p) - \log(1-\tilde{q}-p)\\
&\quad -(\Gamma-2p\Gamma)\log\tilde{z} - \Gamma\log(\tilde{q}-p) + \Gamma\log(1-\tilde{q}-p) && \text{(B.66)}\\
&= p\log\tilde{q} + (1-p)\log(1-\tilde{q}) + \log(1-2p) - \log(1-\tilde{q}-p)\\
&\quad -(\Gamma-2p\Gamma)(\log(1-\tilde{q}) - \log\tilde{q}) + \Gamma\log(1-\tilde{q}-p) && \text{(B.67)}\\
&= \log(1-2p) + (p+\Gamma-2p\Gamma)\log\tilde{q} + (1-p-\Gamma+2p\Gamma)\log(1-\tilde{q})\\
&\quad -\Gamma\log(\tilde{q}-p) - (1-\Gamma)\log(1-\tilde{q}-p) && \text{(B.68)}
\end{aligned}$$

Let's assume the minima in $E(\tilde{q})$ occurs at q' with derivative becoming zero.

$$\left.\frac{\partial E(\tilde{q})}{\partial\tilde{q}}\right|_{\tilde{q}=q'} = 0 \qquad \text{(B.69)}$$

$$\begin{aligned}
\Rightarrow 0 &= \frac{p+\Gamma-2p\Gamma}{q'} - \frac{1-p-\Gamma+2p\Gamma}{1-q'} - \frac{\Gamma}{q'-p} - \frac{\Gamma}{1-q'-p} && \text{(B.70)}\\
&= \frac{(p+\Gamma-2p\Gamma)(1-q') - (1-p-\Gamma+2p\Gamma)q'}{q'(1-q')} - \frac{\Gamma(1-q'-p) - (1-\Gamma)(q'-p)}{(q'-p)(1-q'-p)} && \text{(B.71)}\\
&= \frac{p+\Gamma-2p\Gamma-q'}{q'(1-q')} - \frac{p+\Gamma-2p\Gamma-q'}{q'(1-q')-p+p^2} && \text{(B.72)}\\
&= (p+\Gamma-2p\Gamma-q')\big[\frac{1}{q'(1-q')} - \frac{1}{q'(1-q')-p+p^2}\big] && \text{(B.73)}\\
\Rightarrow 0 &= (p+\Gamma-2p\Gamma-q')(p^2-p) && \text{(B.74)}\\
\Rightarrow q' &= p+\Gamma-2p\Gamma = p\circledast\Gamma && \text{(B.75)}
\end{aligned}$$

This is because of our assumptions on the range of p that $(p^2 - p)$ cannot be zero.

$$\begin{aligned}
E &= \min_{\tilde{q}} E(\tilde{q}) && \text{(B.76)}\\
&= E(q') && \text{(B.77)}\\
&= \log(1-2p) + (p+\Gamma-2p\Gamma)\log q' + (1-p-\Gamma+2p\Gamma)\log(1-q')\\
&\quad -\Gamma\log(q'-p) - (1-\Gamma)\log(1-q'-p) && \text{(B.78)}\\
&= \log(1-2p) + (p \circledast \Gamma)\log(p \circledast \Gamma) + (1-p\circledast\Gamma)\log(1-p\circledast\Gamma)\\
&\quad -\Gamma\log(p+\Gamma-2p\Gamma-p) - (1-\Gamma)\log(1-p-\Gamma+2p\Gamma-p) && \text{(B.79)}\\
&= \log(1-2p) - H_2(p\circledast\Gamma) - \Gamma\log(\Gamma(1-p)) - (1-\Gamma)\log((1-\Gamma)(1-p))\\
&= \log(1-2p) - H_2(p\circledast\Gamma) - \Gamma\log\Gamma - (1-\Gamma)\log(1-\Gamma) - (\Gamma+1-\Gamma)(1-p)\\
&= -H_2(p\circledast\Gamma) + H_2(\Gamma) && \text{(B.80)}
\end{aligned}$$

From (B.62),

$$\begin{aligned}
\mathbb{C}(\Gamma) &= H_2(p) + E && \text{(B.81)}\\
&= H_2(p) + H_2(\Gamma) - H_2(p\circledast\Gamma) && \text{(B.82)}
\end{aligned}$$

From (B.21) and (B.82) we verify that the capacity of both the primal and dual expressions match.

Bibliography

Ahlswede, R. and Csiszár, I. (1993). Common randomness in information theory and cryptography: I secret sharing. *IEEE Transactions on Information Theory*, 39(4):1121–1132.

Bloch, M. and Barros, J. (2011). *Physical-Layer Security: From Information Theory to Security Engineering.* Cambridge University Press, Cambridge.

Blum, M. (1983). Coin flipping by telephone a protocol for solving impossible problems. *ACM SIGACT News*, 15(1):23–27.

Brassard, G., Chaum, D., and Crépeau, C. (1988). Minimum disclosure proofs of knowledge. *Journal of Computer and System Sciences*, 37(2):156–189.

Carter, J. L. and Wegman, M. N. (1977). Universal classes of hash functions (extended abstract). In *Proceedings of the Ninth Annual ACM Symposium on Theory of Computing*, STOC '77, page 106–112, New York, NY, USA. Association for Computing Machinery.

Carter, J. L. and Wegman, M. N. (1979). Universal classes of hash functions. *Journal of Computer and System Sciences*, 18(2):143–154.

Chaum, D., Crépeau, C., and Damgard, I. (1988). Multiparty unconditionally secure protocols. In *Proceedings of the twentieth annual ACM symposium on Theory of computing*, STOC '88, pages 11–19, New York, NY, USA. Association for Computing Machinery.

Cover, T. M. (1999). *Elements of information theory*. John Wiley & Sons.

Crépeau, C., Dowsley, R., and Nascimento, A. C. (2020). On the commitment capacity of unfair noisy channels. *IEEE Transactions on Information Theory*, 66(6):3745–3752.

Crépeau, C. (1997). Efficient cryptographic protocols based on noisy channels. In *Proceedings of the 16th annual international conference on Theory and application of cryptographic techniques*, EUROCRYPT'97, pages 306–317, Berlin, Heidelberg. Springer-Verlag.

Crépeau, C. and Kilian, J. (1988). Achieving Oblivious Transfer Using Weakened Security Assumptions (Extended Abstract). pages 42–52.

Csiszár, I. and Korner, J. (1978). Broadcast channels with confidential messages. *IEEE transactions on information theory*, 24(3):339–348.

Csiszár, I. and Körner, J. (2011). *Information theory: coding theorems for discrete memoryless systems*. Cambridge University Press.

Csiszár, I. and Körner, J. (2011). *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge University Press, Cambridge, 2 edition.

Cuff, P. (2015). A stronger soft-covering lemma and applications. In *2015 IEEE Conference on Communications and Network Security (CNS)*, pages 40–43. IEEE.

Damgård, I., Kilian, J., and Salvail, L. (1999). On the (im) possibility of basing oblivious transfer and bit commitment on weakened security assumptions. In *International Conference on the Theory and Applications of Cryptographic Techniques*, pages 56–73. Springer.

Damgard, I. B., Pedersen, T. P., and Pfitzmann, B. (1998). Statistical secrecy and multibit commitments. *IEEE Transactions on Information Theory*, 44(3):1143–1151.

Dodis, Y., Ostrovsky, R., Reyzin, L., and Smith, A. (2008). Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. *SIAM journal on computing*, 38(1):97–139.

Dodis, Y., Reyzin, L., and Smith, A. (2004). Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. In *International conference on the theory and applications of cryptographic techniques*, pages 523–540. Springer.

El Gamal, A. and Kim, Y.-H. (2011). *Network Information Theory*. Cambridge University Press.

Even, S., Goldreich, O., and Lempel, A. (1985). A randomized protocol for signing contracts. *Communications of the ACM*, 28(6):637–647.

Goldreich, O., Micali, S., and Wigderson, A. (1987). How to play ANY mental game. In *Proceedings of the nineteenth annual ACM symposium on Theory of computing*, STOC '87, pages 218–229, New York, NY, USA. Association for Computing Machinery.

Goldreich, O., Micali, S., and Wigderson, A. (1991). Proofs that yield nothing but their validity or all languages in NP have zero-knowledge proof systems. *Journal of the ACM*, 38(3):690–728.

Halevi, S. (1999). Efficient Commitment Schemes with Bounded Sender and Unbounded Receiver. *Journal of Cryptology*, 12(2):77–89.

Halevi, S. and Micali, S. (1996). Practical and Provably-Secure Commitment Schemes from Collision-Free Hashing. In Koblitz, N., editor, *Advances in Cryptology — CRYPTO '96*, Lecture Notes in Computer Science, pages 201–215, Berlin, Heidelberg. Springer.

Khurana, D., Maji, H. K., and Sahai, A. (2016). Secure computation from elastic noisy channels. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 184–212. Springer.

Maurer, U. M. (1993). Secret key agreement by public discussion from common information. *IEEE transactions on information theory*, 39(3):733–742.

Mishra, M., Dey, B. K., Prabhakaran, V. M., and Diggavi, S. N. (2017). Wiretapped oblivious transfer. *IEEE Transactions on Information Theory*, 63(4):2560–2595.

Naor, M. (1991a). Bit commitment using pseudorandomness. *Journal of cryptology*, 4(2):151–158.

Naor, M. (1991b). Bit commitment using pseudorandomness. *Journal of Cryptology*, 4(2):151–158.

Nascimento, A. C., Barros, J., Skludarek, S., and Imai, H. (2008). The Commitment Capacity of the Gaussian Channel Is Infinite. *IEEE Transactions on Information Theory*, 54(6):2785–2789.

Nisan, N. and Zuckerman, D. (1996). Randomness is linear in space. *Journal of Computer and System Sciences*, 52(1):43–52.

Nojoumian, M. and Stinson, D. R. (2010). Unconditionally secure first-price auction protocols using a multicomponent commitment scheme. In *International Conference on Information and Communications Security*, pages 266–280. Springer.

Ostrovsky, R., Venkatesan, R., and Yung, M. (1992). Secure commitment against a powerful adversary: A security primitive based on average intractability. In Goos, G., Hartmanis, J., Finkel, A., and Jantzen, M., editors, *STACS 92*, volume 577, pages 437–448. Springer Berlin Heidelberg, Berlin, Heidelberg. Series Title: Lecture Notes in Computer Science.

Winter, A., Nascimento, A. C., and Imai, H. (2003). Commitment capacity of discrete memoryless channels. In *IMA International Conference on Cryptography and Coding*, pages 35–51. Springer.

Wyner, A. D. (1975). The wire-tap channel. *The Bell System Technical Journal*, 54(8):1355–1387. Conference Name: The Bell System Technical Journal.

Publications

- M. Mamindlapally, A.K. Yadav, M. Mishra, A.J. Budkuley, *Commitment Capacity under Cost Constraints*, **IEEE International Symposium on Information Theory (ISIT) 2021** (accepted)

- A.K.Yadav, M. Mamindlapally, A.J. Budkuley, M. Mishra, *Multibit Commitment over Compound Channels*, **National Conference on Communications (NCC) 2021** (under review)

- A.K. Yadav, M. Mamindlapally, P. Joshi[1], A.J. Budkuley, *Commitment Capacity over generalised Unfair Noisy Channels*, **IEEE Information Theory Workshop (ITW) 2021** (under preperation)

[1] The names are written in alphabetical order. However, all authors had equal contribution