Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Spring Semester 2023

Semester Thesis

Capacity Bounds for MISO Channels Under 1st- and 2nd-Moment Constraints

Victor Kawasaki-Borruat

Advisor: Prof. Stefan M. Moser
Co-Advisors: Dr. Ligong Wang, Prof. Amos Lapidoth

Abstract

In visible light communication (VLC), safety reasons impose peak- and average-power constraints on the transmitted signal. Moreover, hardware considerations have been seen to subject the input to an additional second-moment constraint. This semester thesis will study the lower bounds of the channel capacity of multiple-input single-output (MISO) optical free-space communication systems with peak-limited antennas subjected to first- and second-moment constraints on the input vector. A tight lower bound is given in the case where moment constraints may be alleviated, and an open-form lower bound is offered when all constraints are active.

Acknowledgements

I would like to thoroughly thank Prof. Dr. Stefan Moser, as well as Dr. Ligong Wang for supervising this semester thesis. It has been a privilege to study in such a welcoming environment. I would also like to thank my friends and family for supporting me in my academic journey.

Contents

Chapter 1

Introduction

Optical wireless communication (OWC) is a communication method that uses light waves (visible, infrared or ultraviolet) to transmit information wirelessly through free space. Compared to traditional radio frequency (RF) wireless communication, information is encoded in OWC signals by modulating the signal's real-valued intensity, rather than the amplitude of a complex-valued signal. The widespread use of IM/DD (*intensity-modulation-direct-detection*) systems is a good motivation to further explore such channels' capacities. Due to safety and hardware reasons, the transmitted signal must satisfy peak-power, as well as first- and second-moment constraints. We will use the nomenclature of *peak-power, average-power and electrical-power* to denote all three constraints. A. Lapidoth *et al.* [1] have already offered channel bounds for single-input and single-output (SISO) optical free-space channels under peak-power and average-power constraints. Moreover, M. Wigger *et al.* [2] have extended capacity results to SISO channels under all three constraints.

Among various types of OWC systems, the Multiple-Input Single-Output (MISO) system is a popular configuration that utilizes multiple transmitters and a single receiver. With the same logic, constraints must be applied to the input vector. First- and second-moment constraints thus become one- and two-norm constraints, and the peak-power constraint is still applied to every element of the input.

This paper will combine multiple previous results on (constrained) optical free-space channels to derive capacity bounds on optical MISO channels with a yet-to-be explored second-moment constraint. The derivations in this paper will principally be based off of the aforementioned literature. We will start off by detailing the channel model in Section 1.1. Section 1.2 will present the optimal structure to channel inputs, which will allow us to derive alternative channel capacity formulations in Chapter 2. Chapter 3 will offer lower bounds to channel capacity, and Chapter 4 will discuss numerical simulation results of a toy example.

1.1 Channel Model

The channel model considered in this paper is the same as the one studied by Ma, Moser, Wigger and Wang [3]: a multiple-input and single-output (MISO) antenna system of the form

$$\bar{X} = \mathbf{h}^{\mathsf{T}}\mathbf{X}, \tag{1.1}$$

where $\mathbf{h} = (h_1, ..., h_{n_T})^{\mathsf{T}}$ is a constant channel state vector and $\mathbf{X}$ is an n_T-dimensional channel input vector. A visual representation is offered in Figure 1.1. We call the weighted sum of inputs $\bar{X}$ the *channel image*. To avoid degeneracy of solutions and without loss of generality, we will consider that the channel state vector is ordered:

$$h_1 > h_2 > ... > h_{n_T} > 0. \tag{1.2}$$

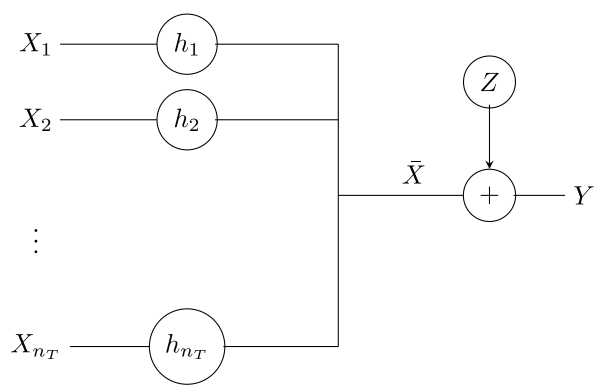

Figure 1.1: Channel Model Diagram

Moreover, each X_i-component of $\mathbf{X}$ is non-negative, and for reasons mentioned in [1], $\bar{X}$ can be considered as corrupted with additive Gaussian noise $Z \sim \mathcal{N}(0, \sigma^2)$ from a combination of thermal, shot, and relative-intensity noise. Hence, our channel output is given by

$$Y = \bar{X} + Z \tag{1.3}$$

Inputs will be subjected to three types of constraints: *peak- , average- and electrical-power constraints*:

$$X_i \in [0, \mathsf{A}] \quad \forall i \in \{1, ..., n_T\}, \tag{1.4}$$

$$P[X_i > \mathsf{A}] = 0 \quad \forall i \in \{1, ..., n_T\}, \tag{1.5}$$

$$\mathsf{E}[\|\mathbf{X}\|_1] \leq \mathcal{E}, \tag{1.6}$$

$$\mathsf{E}\left[\|\mathbf{X}\|_2^2\right] \leq \mathcal{V}, \tag{1.7}$$

for some fixed parameters $\mathcal{E}, \mathcal{V}, \mathsf{A} \geq 0$. We will denote the ratio between the average-power constraint and the allowed peak-power by α_1, and the ratio between the electrical-power and the peak-power by α_2. Thus we have

$$\alpha_1 \triangleq \frac{\mathcal{E}}{\mathsf{A}}, \tag{1.8}$$

$$\alpha_2 \triangleq \frac{\mathcal{V}}{\mathsf{A}^2}. \tag{1.9}$$

Note that [2] treats all three constraints as well, but only on SISO channels. Also note that α_1 and α_2 are not unrelated: by the L1-L2 norm inequality, the expected norms of the channel input must satisfy:

$$\sqrt{\mathsf{E}[\|\mathbf{X}\|_2^2]} \leq \mathsf{E}[\|\mathbf{X}\|_1]. \tag{1.10}$$

The utility of this rather obvious result will come in handy in Chapter 2 when observing whether average- and electrical-power constraints are active or not.

We will denote the capacity of the channel (1.1) with allowed peak-power A and allowed first- (resp. second-) moment $\alpha_1 A$ (resp. $\alpha_2 \mathsf{A}^2$) as $\mathcal{C}_{\mathbf{h}^\mathsf{T},\sigma^2}\big(\mathsf{A}, \alpha_1\mathsf{A}, \alpha_2\mathsf{A}^2\big)$. A general form of the capacity is given in [4]:

$$\mathcal{C}_{\mathbf{h}^\mathsf{T},\sigma^2}\big(\mathsf{A}, \alpha_1\mathsf{A}, \alpha_2\mathsf{A}^2\big) \triangleq \sup_{Q_\mathbf{X}} \mathcal{I}(\mathbf{X}, Y) \tag{1.11}$$

where the supremum is taken over all laws $Q_\mathbf{X}$ satisfying 1.4-1.7.

1.2 Minimum-Energy Signaling

In this section we will briefly present the signaling scheme for our channel model. It is the subject of [3], which covers in detail the entire process.

1.2.1 Optimal Channel Input Vector

Ma *et al.* [3] take a weighted sum approach to the mixed vector-norm minimization and propose an optimal channel input vector $\mathbf{x}$ by inspecting the Karush-Kuhn-Tucker (KKT) conditions of the objective function

$$\mathbf{x} \mapsto \lambda \cdot \frac{\|\mathbf{x}\|_1}{A} + (1-\lambda) \cdot \frac{\|\mathbf{x}\|_2^2}{A^2}. \tag{1.12}$$

The derived solution is presented in [3, Lemma 9], and is defined for a fixed value of $\lambda \in [0,1]$. Since channel capacity bounds for the case of $\lambda = 1$ have already been treated in [5], we will also (as in [3]) only focus on the case of $\lambda \in [0,1)$. The choice of λ will induce a particular partition of the channel image into various sub-intervals, denoted $\mathcal{I}_{\ell,k}$ (see [3, Definitions 5 & 7]), which we recall here for context.

Definition 1. Set $\kappa_0 \triangleq 0$. For each index $\ell \in \{1, ..., n_T\}$, define the integer

$$\kappa_\ell \triangleq \max\left\{ j \in \{\ell, ..., n_T\} : \frac{\nu}{2}\left(\frac{h_\ell}{h_j} - 1\right) < 1 \right\}; \tag{1.13}$$

define the point

$$t_\ell \triangleq \mathsf{A}\sum_{i=1}^{\ell} h_i + \mathsf{A}\sum_{i=\ell+1}^{\kappa_\ell}\left(\frac{h_i^2}{h_\ell} + \frac{\nu}{2}\left(\frac{h_i^2}{h_\ell} - h_i\right)\right); \tag{1.14}$$

and finally, for every $k \in \{\kappa_{\ell-1}+1, ..., \kappa_\ell\}$, define

$$s_k \triangleq \mathsf{A}\sum_{i=1}^{\ell-1} h_i + \mathsf{A}\sum_{i=\ell}^{k-1}\frac{\nu}{2}\left(\frac{h_i^2}{h_k} - h_i\right). \tag{1.15}$$

[3, Remark 6] explains that these points represent thresholds on $\bar{x}$ corresponding to when the ℓ^{th} antenna should be set to full , and when the k^{th} antenna should be switched on. Lastly, κ_k indicates how many antennas should be switched on before the k^{th} antenna is fixed to maximum value A.

Definition 2. For any $\ell \in \{1, ..., n_T\}$ for which $\kappa_\ell < \kappa_{\ell-1}$, define the sub-intervals

$$\mathcal{I}_{\ell,k} \triangleq \begin{cases} [t_{\ell-1}, s_{k+1}] & \text{if } k = \kappa_{\ell-1} \\ [s_k, s_{k+1}] & \text{if } \kappa_{\ell-1} < k < \kappa_\ell \\ [s_k, t_\ell] & \text{if } k = \kappa_\ell \end{cases} \tag{1.16}$$

For any $\ell \in \{2, ..., n_T\}$ for which $\kappa_{\ell-1} = \kappa_\ell$, define the sub-interval

$$\mathcal{I}_{\ell,\emptyset} \triangleq [t_{\ell-1}, t_\ell] \tag{1.17}$$

From these sub-intervals we can construct the optimal input vector, denoted $\mathbf{x}_{\text{min},\lambda}(\bar{x})$. It is given in [3, Eq. (26), Lemma 9]. The signaling scheme is as follows: conditionally on $\bar{x} \in \mathcal{I}_{\ell,k}$, the $(\ell-1)$ strongest antennas are set to full power, while the $(n_T - k)$ weakest are switched off. The remaining $(k-\ell+1)$ antennas transmit in a shifted beamforming manner. The double-indexing with (ℓ, k) thus denotes how many antennas should be set to full power and how many antennas should be switched on respectively.

1.2.2 Characterization of Input Distribution

Now recall Theorem 12 of [3], which states that the desired channel image $\bar{X}$ can be generated with a random input vector $\mathbf{X}$ if, and only if, there exists a $\lambda \in [0, 1]$ such that

$$\mathsf{E}_{\bar{X}}[m(\bar{X}, \lambda)] \le \alpha_1 \mathsf{A}, \tag{1.18}$$
$$\mathsf{E}_{\bar{X}}[v(\bar{X}, \lambda)] \le \alpha_2 \mathsf{A}^2, \tag{1.19}$$

where $m(\cdot,\cdot)$ and $v(\cdot,\cdot)$ are defined in [3, Eq. (27)-(28)]. They correspond to the one- and two-norm of $\mathbf{x}_{\text{min},\lambda}(\bar{x})$ respectively.

With this result in hand, we will consider a *fixed* value of $\lambda \in [0, 1)$ and combine the conditions (32a)-(32b) of [3] into one. Let

$$\mathsf{E}\big[\rho(\bar{X}, \lambda)\big] \triangleq \lambda \cdot \frac{\mathsf{E}\big[m(\bar{X}, \lambda)\big]}{\mathsf{A}} + (1-\lambda)\cdot\frac{\mathsf{E}\big[v(\bar{X}, \lambda)\big]}{\mathsf{A}^2}, \tag{1.20}$$

then we have that

$$\mathsf{E}\left[\rho(\bar{X}, \lambda)\right] \leq \underbrace{\lambda \cdot \alpha_1 + (1 - \lambda) \cdot \alpha_2}_{\triangleq \alpha}. \tag{1.21}$$

We now have a single condition which must be satisfied in order for $\bar{X}$ to be properly generated. Note that choosing to solve for a fixed λ does not create impossible cases, as there will always exist a corresponding (α_1, α_2)-pair to each fixed λ, α.

Chapter 2

Capacity Formulations

2.1 Preliminaries

We will introduce a two-dimensional random variable $V = (V_1, V_2)^\mathsf{T}$ over the admissible values of (ℓ, k) pairs such that

$$\Big(V = (\ell, k)\Big) \iff \Big(\bar{X} \in \mathcal{I}_{\ell,k}\Big). \tag{2.1}$$

The channel state vector $\mathbf{h}$'s 1- and 2-norms will be thoroughly used, as well as truncated norms. To this goal, we will use the following to denote the truncated norms:

$$\|\mathbf{h}_\ell^k\|_1 \triangleq \sum_{i=\ell}^{k} h_i, \tag{2.2}$$

$$\|\mathbf{h}_\ell^k\|_2^2 \triangleq \sum_{i=\ell}^{k} h_i^2. \tag{2.3}$$

Moreover, we will denote $\Pr[V = (\ell, k)] \triangleq p_{\ell,k}$ for brevity.

Remark 3. We denote *admissible* values of (ℓ, k) the pairs those recalled in Definition 1. △

2.2 Capacity: General Case and Less Constraints

We start with a general result.

Lemma 4 (Capacity in Terms of Channel Image). *The MISO channel capacity can be expressed in terms of a SISO channel capacity, where the input is the channel image $\bar{X}$. Namely,*

$$\mathcal{C}_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathsf{A}, \alpha_1, \alpha_2, \lambda) \triangleq \max_{Q_{\bar{X}}} \mathcal{I}(\bar{X}; Y) \tag{2.4}$$

where $Q_{\bar{X}}$ is the set of admissible distributions. The conditions on these distributions will be explored and derived in subsequent propositions.

Proof. [5, Section III] proves the same result by noticing that $\mathbf{X} \multimap \bar{X} \multimap Y$ forms a Markov chain, and since $\bar{X}$ is a function of $\mathbf{X}$, then

$$\mathcal{I}(\bar{X};Y) = \mathcal{I}(\mathbf{X};Y). \tag{2.5}$$

□

We recall [5, Propositions 1 & 2], which derive results on MISO channels only under average- and peak-power constraints respectively. We translate these results to our new case.

If $\mathbf{X}$ is only subjected to first- and second-moment constraints without a peak-power constraint, then it is reduced to a SISO case.

Proposition 5 (No Peak-Power Constraint). *Without a peak-power constraint,*

$$\mathcal{C}_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathrm{A},\alpha_1,\alpha_2,\lambda) = \max_{\substack{Q_{\bar{X}}:\bar{X}\in[0,\infty] \\ \mathsf{E}[\bar{X}]\le\tilde{h}\mathcal{E} \\ \mathsf{E}[\bar{X}^2]\le\tilde{h}^2\mathcal{V}}} \mathcal{I}(\bar{X};Y) = \mathcal{C}_{\mathbf{1},\sigma^2}(\tilde{h}\mathcal{E},\tilde{h}^2\mathcal{V}) \tag{2.6}$$

where $\tilde{h} = \sqrt{\|\mathbf{h}_1^{n_T}\|_1 h_1}$ and $\mathcal{C}_{\mathbf{1},\sigma^2}(\tilde{h}\mathcal{E},\tilde{h}^2\mathcal{V})$ denotes the capacity of a SISO channel with unit channel gain and first- and second-moment constraints, as studied in [2].

Proof. When $\mathbf{X}$ satisfies (1.6)-(1.7), then we have

$$\begin{aligned}
\mathsf{E}\big[\bar{X}\big] &= \sum_{k=1}^{n_T} h_k\,\mathsf{E}[X_k] \\
&\le h_1\mathcal{E} \\
&\le \sqrt{\|\mathbf{h}_1^{n_T}\|_1 h_1}\mathcal{E} \\
&= \tilde{h}\mathcal{E}
\end{aligned} \tag{2.7}$$

and

$$\begin{aligned}
\mathsf{E}\big[\bar{X}^2\big] &= \|\mathbf{h}_1^{n_T}\|_1^2 \cdot \mathsf{E}\left[\frac{1}{\|\mathbf{h}_1^{n_T}\|_1^2}\left(\sum_{k=1}^{n_T} h_k X_k\right)^2\right] \\
&\overset{i)}{\le} \|\mathbf{h}_1^{n_T}\|_1^2 \cdot \mathsf{E}\left[\sum_{k=1}^{n_T} \frac{h_k}{\|\mathbf{h}_1^{n_T}\|_1} X_k^2\right] \\
&\overset{ii)}{\le} \|\mathbf{h}_1^{n_T}\|_1 h_1\,\mathsf{E}\left[\sum_{k=1}^{n_T} X_k^2\right] \\
&\le \tilde{h}^2\mathcal{V}.
\end{aligned} \tag{2.8}$$

where $i)$ is by Jensen's inequality, and $ii)$ is done as in the inequality of $\mathsf{E}\big[\bar{X}\big]$.

Conversely, any distribution satisfying $\mathsf{E}\big[\bar{X}\big] \le \tilde{h}\mathcal{E}$ and $\mathsf{E}\big[\bar{X}^2\big] \le \tilde{h}^2\mathcal{V}$ can be generated by sending $\frac{\bar{X}}{\tilde{h}}$ on the transmitter corresponding to h_1 and setting all others to zero. □

We recall [6, Proposition 1], which states that in the case of peak- and average-power constraints, when $\alpha \geq \frac{n_T}{2}$, the average-power constraint is dropped. This is due to the capacity-achieving distribution being symmetric around $\frac{A}{2}$. We can now derive the following condition on α_1, α_2.

Lemma 6. *When $\alpha_1 \geq \frac{n_T}{2}$, the average-power constraint is dropped, and when $\alpha_2 \geq \frac{n_T^2}{4}$, the electrical-power constraint is dropped.*

Proof. Proving first part is the same as proving [6, Proposition 1]. The second part is proven by invoking the L1-L2 inequality on the input vector. □

We however note that the case where one moment-constraint is active and not the other is not the subject of this paper. It is of more interest when either *both* moment-constraints are active, with or without a peak-power constraint. The case of no electrical-power constraint is the subject of [5], and the case of no average-power constraint would be superfluous, as the L1-L2 inequality on the input vector norm tells us that electrical-power constraints can be directly imposed by setting average-power constraints. We thus skip these cases and go on to present the result for a channel where $\alpha_1 \geq \frac{n_T}{2}, \alpha_2 \geq \frac{n_T^2}{4}$.

Proposition 7 (Only Peak-Power Constraint). *When $\alpha_1 \geq \frac{n_T}{2}$ and $\alpha_2 \geq \frac{n_T^2}{4}$, then:*

$$\mathcal{C}_{\mathbf{h}^{\mathsf{T}},\sigma^2}(\mathrm{A}, \alpha_1, \alpha_2, \lambda) = \max_{Q_{\bar{X}}:\bar{X}\in[0,\|\mathbf{h}_1^{n_T}\|_1\mathrm{A}]} \mathcal{I}(\bar{X}; Y) \tag{2.9}$$

$$= \mathcal{C}_{1,\sigma^2}(\|\mathbf{h}_1^{n_T}\|_1\mathrm{A}, \frac{\|\mathbf{h}_1^{n_T}\|_1\mathrm{A}}{2}) \tag{2.10}$$

where $\mathcal{C}_{1,\sigma^2}(\|\mathbf{h}_1^{n_T}\|_1\mathrm{A}, \frac{\|\mathbf{h}_1^{n_T}\|_1\mathrm{A}}{2})$ denotes the capacity of a SISO channel with unit gain with allowed peak-power $\|\mathbf{h}_1^{n_T}\|_1\mathrm{A}$ and allowed average-power $\frac{\|\mathbf{h}_1^{n_T}\|_1\mathrm{A}}{2}$.

Proof. Since both conditions are inactive, we can safely reduce the situation to [5, Proposition 2], giving us the desired result. □

We will now turn to the case where both constraints are active.

2.3 Capacity Under all Constraints

Similarly to [5], the set of admissible distributions gets complicated when both $\alpha_1 < \frac{n_T}{2}$ and $\alpha_2 < \frac{n_T^2}{4}$. To clearly define these distributions, we will use one of the main results of [3]. By the new mixed constraint (1.21), we will now denote the channel capacity by $\mathcal{C}_{\mathbf{h}^{\mathsf{T}},\sigma^2}(\mathrm{A}, \alpha)$.

Proposition 8 (Channel Capacity with Active Peak-Power, First- and Second-Moment Constraints). *When $\alpha < \lambda \cdot \frac{n_T}{2} + (1-\lambda) \cdot \frac{n_T^2}{4}$, the channel capacity is given by*

$$\mathcal{C}_{\mathbf{h}^{\mathsf{T}},\sigma^2}(\mathrm{A}, \alpha) \triangleq \max_{Q_{\bar{X}}} \mathcal{I}(\bar{X}; Y) \tag{2.11}$$

where the maximization is over all laws on $\bar{X} \in \mathbb{R}_0^+$ *satisfying*

$$\mathsf{Pr}[\bar{X} > \|\mathbf{h}_1^{n_T}\|_1 \mathsf{A}] = 0, \tag{2.12}$$

and

$$\begin{aligned}\mathsf{E}\left[\rho(\bar{X},\lambda)\right] &= \sum_{\ell,k} p_{\ell,k}\left[\mu(\lambda,\ell,k) + (1-\lambda)\cdot \mathsf{E}\left[\frac{(\bar{X}-s_{\ell,k})^2}{\mathsf{A}^2\|\mathbf{h}_\ell^k\|_2^2}|V=(\ell,k)\right]\right] \\ &\leq \alpha,\end{aligned} \tag{2.13}$$

where $\mu(\lambda,\ell,k) \triangleq (\ell-1) + \frac{(1-\lambda)\nu^2 - 2\lambda\nu}{4}(k-\ell+1)$, $\nu \triangleq \frac{\lambda}{1-\lambda}$, *and*
$s_{\ell,k} \triangleq \mathsf{A}\sum_{i=1}^{\ell-1} h_i + \frac{\nu\mathsf{A}}{2}\sum_{i=\ell}^{k} h_i$.

Proof. Recall Section 1.2. We may now derive our desired distributions by only taking into account minimum-energy input vectors, thus simplifying the mixed moment constraint for $\bar{X} \in \mathcal{I}_{\ell,k}$. We will denote $\bar{X} \in \mathcal{I}_{\ell,k}$ as $\bar{X}_{\ell,k}$ for convenience:

$$\mathsf{E}\left[\rho(\bar{X}_{\ell,k},\lambda)\right] \triangleq \frac{\lambda}{\mathsf{A}}\cdot\mathsf{E}\left[m(\bar{X}_{\ell,k},\lambda)\right] + \frac{1-\lambda}{\mathsf{A}^2}\cdot\mathsf{E}\left[v(\bar{X}_{\ell,k},\lambda)\right] \tag{2.14}$$

$$\begin{aligned}&= \lambda\cdot\left((\ell-1) - (k-\ell+1)\frac{\nu}{2} + \left(\frac{\mathsf{E}\left[\bar{X}_{\ell,k}\right] - s_{\ell,k}}{\mathsf{A}\|h_{\ell,k}\|_2^2}\right)\|h_{\ell,k}\|_1\right) \\ &\quad +(1-\lambda)\cdot\left((\ell-1) + (k-\ell+1)\frac{\nu^2}{4} + \frac{\mathsf{E}\left[(\bar{X}_{\ell,k}-s_{\ell,k})^2\right]}{\mathsf{A}^2\|h_{\ell,k}\|_2^2}\right) \\ &\quad -\lambda\cdot\left(\frac{\mathsf{E}\left[\bar{X}_{\ell,k}\right] - s_{\ell,k}}{\mathsf{A}\|h_{\ell,k}\|_2^2}\right)\|h_{\ell,k}\|_1 \\ &= \underbrace{(\ell-1) + (k-\ell+1)\cdot\left(\frac{(1-\lambda)\nu^2}{4} - \frac{\lambda\nu}{2}\right)}_{\triangleq\mu(\lambda,\ell,k)} \\ &\quad + (1-\lambda)\cdot\mathsf{E}\left[\frac{(\bar{X}_{\ell,k}-s_{\ell,k})^2}{\mathsf{A}^2\|h_{\ell,k}\|_2^2}\right].\end{aligned} \tag{2.15}$$

Hence over all $\mathcal{I}_{\ell,k}$, by law of total probability the overall constraint becomes Eq. (2.13). □

Thus, we are able to see that our mixed moment constraint actually implies only a second moment constraint on $(\bar{X} - s_{\ell,k})$ conditionally on $V = (\ell,k)$. This will be exploited to derive channel capacity bounds.

Chapter 3

Lower Bounds with all Constraints Active

Throughout this section it will be assumed that both moment constraints are active. A useful threshold value will be the following:

$$\alpha_{\mathsf{th}} \triangleq \frac{1}{\mathsf{A}\|\mathbf{h}_1^{n_T}\|_1} \sum_{\ell,k} (b_{\ell,k} - a_{\ell,k})\mu(\lambda,\ell,k) \\ + \frac{(1-\lambda)}{\mathsf{A}^3\|\mathbf{h}_1^{n_T}\|_1} \sum_{\ell,k} \left(\frac{(b_{\ell,k} - s_{\ell,k})^3 - (a_{\ell,k} - s_{\ell,k})^3}{3\|\mathbf{h}_\ell^k\|_2^2} \right). \tag{3.1}$$

We will see in the following propositions that the bounds differ, depending on whether $\alpha \gtreqless \alpha_{\mathsf{th}}$. It turns out that this threshold is the smallest value α can take where $\bar{X}$ can be uniformly distributed over $[0, \|\mathbf{h}_1^{n_T}\|_1 \mathsf{A}]$. We will show this in the proof of Proposition 10.

We start off by lower-bounding the channel capacity with the Entropy Power Inequality (EPI), proposed by Shannon in 1949, stated in [7, Theorem 17.7.3].

Lemma 9 (Capacity Lower Bound). *The channel capacity is lower-bounded by*

$$\mathcal{C}_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathsf{A},\alpha) \geq \frac{1}{2}\log\left(1 + \frac{e^{2\mathsf{h}(\bar{X})}}{2\pi e \sigma^2}\right) \tag{3.2}$$

where $\mathsf{h}(\cdot)$ *denotes the differential entropy.*

Proof. We first expand the channel capacity and lower-bounding by the mutual information:

$$\begin{aligned} \mathcal{C}_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathsf{A},\alpha) &\geq \mathcal{I}(\bar{X};Y) && (3.3)\\ &= \mathcal{I}(\bar{X};\bar{X}+Z) && (3.4)\\ &= \mathsf{h}(\bar{X}+Z) - \mathsf{h}(Z) && (3.5) \end{aligned}$$

$$\overset{i)}{\geq} \frac{1}{2}\log\Big(e^{2\mathsf{h}(\bar{X})} + e^{2\mathsf{h}(Z)}\Big) - \mathsf{h}(Z) \tag{3.6}$$

$$= \frac{1}{2}\log\left(1 + \frac{e^{2\mathsf{h}(\bar{X})}}{2\pi e\sigma^2}\right) \tag{3.7}$$

where $i)$ is due to the EPI. We may apply it since the additive Gaussian noise Z is considered to be independent of the channel image. □

From the expression in (3.2), notice that a tight lower bound can be found by saturating both inequalities. This is possible by finding a valid input distribution to $\bar{X}$ which successfully maximizes $\mathsf{h}(\bar{X})$ under the constraint (2.13), which leads us to our first capacity lower bound.

3.1 Lower Bound for $\alpha \geq \alpha_{\mathsf{th}}$

Proposition 10 (Lower Bound when $\alpha \geq \alpha_{\mathsf{th}}$). *When $\alpha \geq \alpha_{\mathsf{th}}$, the capacity is lower-bounded as:*

$$\mathcal{C}_{\mathbf{h}^{\mathsf{T}},\sigma^2}(\mathsf{A},\alpha) \geq \frac{1}{2}\log(1 + \frac{\|\mathbf{h}_1^{n_T}\|_1^2\mathsf{A}^2}{2\pi e\sigma^2}) \tag{3.8}$$

Proof. As mentioned in the beginning of Section 3, we will choose $\bar{X}$ to be uniformly distributed over $[0, \|\mathbf{h}_1^{n_T}\|_1\mathsf{A}]$. With such a distribution, notice that $\Pr[V = (\ell,k)] = p_{\ell,k}$ must be

$$\begin{aligned} p_{\ell,k} &\triangleq \frac{|\mathcal{I}_{\ell,k}|}{\|\mathbf{h}_1^{n_T}\|_1\mathsf{A}} \\ &= \frac{b_{\ell,k} - a_{\ell,k}}{\|\mathbf{h}_1^{n_T}\|_1\mathsf{A}}. \end{aligned} \tag{3.9}$$

This distribution of $\bar{X}$ obviously satisfies (2.12), as well as (2.13) because

$$\begin{aligned} \mathsf{E}_{\mathcal{U}}[\rho(\bar{X},\lambda)] &= \sum_{\ell,k} \frac{b_{\ell,k} - a_{\ell,k}}{\mathsf{A}\|\mathbf{h}_1^{n_T}\|_1}\Bigg(\mu(\lambda,\ell,k) - \frac{2(1-\lambda)s_{\ell,k}}{\mathsf{A}^2\|\mathbf{h}_\ell^k\|_2^2}\frac{b_{\ell,k}+a_{\ell,k}}{2} \\ &\quad + \frac{1-\lambda}{3\mathsf{A}^2\|\mathbf{h}_\ell^k\|_2^2}(b_{\ell,k}^2 + a_{\ell,k}b_{\ell,k} + a_{\ell,k}^2) + \frac{(1-\lambda)s_{\ell,k}^2}{\mathsf{A}^2\|\mathbf{h}_\ell^k\|_2^2}\Bigg) \\ &= \frac{1}{\mathsf{A}\|\mathbf{h}_1^{n_T}\|_1}\sum_{\ell,k}(b_{\ell,k} - a_{\ell,k})\mu(\lambda,\ell,k) \\ &\quad + \frac{1-\lambda}{\mathsf{A}^3\|\mathbf{h}_1^{n_T}\|_1}\sum_{\ell,k}\left(-s_{\ell,k}\frac{b_{\ell,k}^2 - a_{\ell,k}^2}{\|\mathbf{h}_\ell^k\|_2^2} + \frac{b_{\ell,k}^3 - a_{\ell,k}^3}{3\|\mathbf{h}_\ell^k\|_2^2} + s_{\ell,k}^2\frac{b_{\ell,k} - a_{\ell,k}}{\|\mathbf{h}_\ell^k\|_2^2}\right) \\ &= \frac{1}{\mathsf{A}\|\mathbf{h}_1^{n_T}\|_1}\sum_{\ell,k}(b_{\ell,k} - a_{\ell,k})\mu(\lambda,\ell,k) \\ &\quad + \frac{(1-\lambda)}{\mathsf{A}^3\|\mathbf{h}_1^{n_T}\|_1}\sum_{\ell,k}\left(\frac{(b_{\ell,k} - s_{\ell,k})^3 - (a_{\ell,k} - s_{\ell,k})^3}{3\|\mathbf{h}_\ell^k\|_2^2}\right) \\ &= \alpha_{\mathsf{th}}. \end{aligned} \tag{3.10}$$

Thus, we have that the differential entropy of $\bar{X}$ given by

$$\mathsf{h}(\bar{X}) = \log(\mathrm{A}\|\mathbf{h}_1^{n_T}\|_1), \tag{3.11}$$

which we can directly plug into the RHS of (3.2), yielding our desired result. Our capacity is thus lower-bounded by (3.8) when $\alpha \geq \alpha_{\mathsf{th}}$. □

Remark 11. A expected, α_{th} does *not* depend on A. Notice that $a_{\ell,k}, b_{\ell,k}$ and $s_{\ell,k}$ can have A factored out, thus cancelling out those in the denominator of (3.10). △

3.2 Lower Bound for $\alpha < \alpha_{\mathsf{th}}$

In this section we will consider the case where $\alpha < \alpha_{\mathsf{th}}$, i.e. when a uniformly distributed channel image is not possible. The optimization of $\mathsf{h}(\bar{X})$ this time is much less straightforward. The following will cover the derivations leading to an open-form optimization, unfortunately without an analytical solution. We start by expanding the differential entropy of $\bar{X}$:

$$\begin{aligned}
\mathsf{h}(\bar{X}) &= \mathsf{h}(\bar{X}) - \mathsf{h}(\bar{X}|V) + \mathsf{h}(\bar{X}|V) && (3.12)\\
&= \mathcal{I}(\bar{X}; V) + \mathsf{h}(\bar{X}|V) && (3.13)\\
&= \mathsf{H}(V) - \mathsf{H}(V|\bar{X}) + \mathsf{h}(\bar{X}|V) && (3.14)\\
&\overset{i)}{=} \mathsf{H}(\mathbf{p}) + \sum_{\ell,k} p_{\ell,k} \mathsf{h}(\bar{X}|V = (\ell,k)), && (3.15)
\end{aligned}$$

where $i)$ is due to the non-overlapping property of the $\mathcal{I}_{\ell,k}$ partition of $\bar{\mathcal{X}}$ (explained in [3, Lemma 8]).

Looking at (2.13), we notice that conditionally on $\bar{x} \in \mathcal{I}_{\ell,k}$, we may consider a partition of α into per-interval energy allocations. We now define a vector $\boldsymbol{\alpha} \triangleq (\alpha_{1,1}, ...\alpha_{n_T,\kappa_{n_T}})^{\mathsf{T}}$ such that

$$\mathsf{E}\left[\rho(\bar{X}, \lambda)|\bar{X} \in \mathcal{I}_{\ell,k}\right] = \alpha_{\ell,k}, \tag{3.16a}$$

$$\sum_{\ell,k} p_{\ell,k} \alpha_{\ell,k} = \alpha. \tag{3.16b}$$

Note that any distribution of $\bar{X}$ satisfying (3.16) automatically satisfies (2.13). The maximization of $\mathsf{h}(\bar{X})$ in (3.15) is now twofold: we must optimize over $\mathbf{p}$ as well as find the optimal energy-allocation vector $\boldsymbol{\alpha}$ under the constraints (3.16). Notice how the two vectors are related: $\mathbf{p}$ denotes the probability of the channel image being in a particular interval, and $\boldsymbol{\alpha}$ denotes the energy allocation in said interval. The probability mass vector will thus scale each component of $\boldsymbol{\alpha}$, while $\boldsymbol{\alpha}$ itself sets the distribution parameters.

3.2.1 Conditional Max-Entropy Distribution

We will now derive max-entropy distributions of $\bar{X}$, conditional on $\bar{X} \in \mathcal{I}_{\ell,k}$.

Lemma 12 (Interval-wise Max-Entropy Distribution). *Conditionally on $\bar{X} \in \mathcal{I}_{\ell,k}$, the differential entropy maximizing distribution of $\bar{X}_{\ell,k}$ subjected to (3.16) is given by*

$$f^*_{\bar{X}_{\ell,k}}(\bar{x}) \triangleq \frac{1}{\mathsf{A}\theta_{\ell,k}\|\mathbf{h}^k_\ell\|_2} \cdot \frac{\phi\Big(-\frac{(\bar{x}-s_{\ell,k})^2}{2\mathsf{A}^2\theta^2_{\ell,k}\|\mathbf{h}^k_\ell\|_2^2}\Big)}{Z_{\ell,k}}, \quad \theta_{\ell,k} > 0, \tag{3.17}$$

where

$$Z_{\ell,k} \triangleq \Phi(\mathfrak{b}_{\ell,k}(\theta_{\ell,k})) - \Phi(\mathfrak{a}_{\ell,k}(\theta_{\ell,k})), \tag{3.18}$$

and $\phi(\cdot), \Phi(\cdot)$ are the normal probability distribution and cumulative distribution functions, respectively.

Proof. Let $\tilde{g}$ be the following:

$$\tilde{g} : \bar{x} \overset{\triangle}{\mapsto} \frac{(\bar{x}-s_{\ell,k})^2}{\mathsf{A}^2\|\mathbf{h}^k_\ell\|_2^2} \cdot \mathbf{1}\{\bar{x} \in \mathcal{I}_{\ell,k}\} \tag{3.19}$$

The constraint (3.16) can thus be re-written as

$$\mathsf{E}\big[\tilde{g}(\bar{X}_{\ell,k})\big] = \frac{\alpha_{\ell,k} - \mu(\lambda,\ell,k)}{1-\lambda}. \tag{3.20}$$

Applying [7, Theorem 12.1.1], we get that the distribution maximizing $\mathsf{h}(\bar{X}|\bar{X} \in \mathcal{I}_{\ell,k})$ is of the form:

$$f^*_{\bar{X}_{\ell,k}}(\bar{x}) = e^{-\gamma_{0,\ell,k} - \gamma_{1,\ell,k}\tilde{g}(\bar{x})}. \tag{3.21}$$

By substituting $\frac{1}{2\theta^2_{\ell,k}} \triangleq \gamma_{1,\ell,k}$ and then normalizing over $\mathcal{I}_{\ell,k}$ with the appropriate value of $\gamma_{0,\ell,k}$, we get exactly the distribution in (3.17). □

3.2.2 Solving Distribution Parameters

We thus have that $\bar{X}$ is a concatenation of truncated Gaussian random variables, each restricted to $\mathcal{I}_{\ell,k} = [a_{\ell,k}, b_{\ell,k}]$. Over each of these non-overlapping intervals, (3.16) must be satisfied, which we can now compute[1] with our newfound distribution $f^*_{\bar{X}_{\ell,k}}$ from (3.17).

$$\begin{aligned} \mathsf{E}_{f^*_{\bar{X}_{\ell,k}}}\big[\rho(\lambda,\bar{X})|\bar{X} \in \mathcal{I}_{\ell,k}\big] &\triangleq \mu(\lambda,\ell,k) + (1-\lambda)\cdot\mathsf{E}\left[\frac{(\bar{X}_{\ell,k}-s_{\ell,k})^2}{\mathsf{A}^2\|\mathbf{h}^k_\ell\|_2^2}\right] \\ &= \mu(\lambda,\ell,k) \\ &\quad + (1-\lambda)\cdot\theta^2_{\ell,k}(1+\eta_{\ell,k}(\theta_{\ell,k})) \\ &\overset{!}{=} \alpha_{\ell,k}, \end{aligned} \tag{3.22}$$

[1] The full calculation is deferred to Appendix A.1.

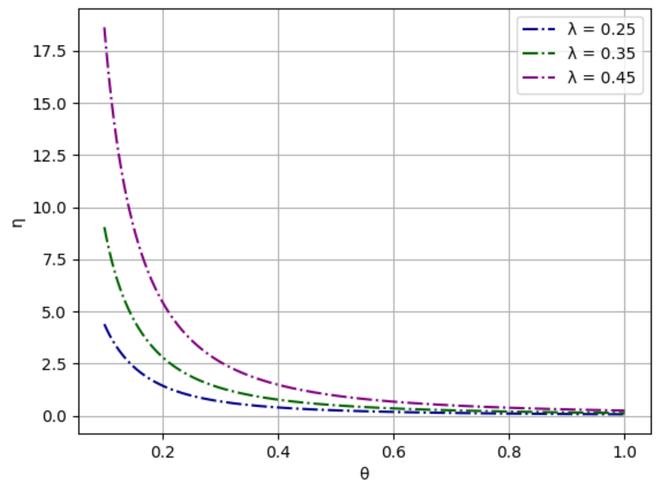

Figure 3.1: The $\eta_{\ell,k}(\cdot)$-function over various θ-values for different choices of λ. The dark blue and dark green lines are only plotted over values for which they are well-defined.

with

$$\eta_{\ell,k}(\theta_{\ell,k}) \triangleq \frac{\mathfrak{a}_{\ell,k}(\theta_{\ell,k})\phi(\mathfrak{a}_{\ell,k}(\theta_{\ell,k})) - \mathfrak{b}_{\ell,k}(\theta_{\ell,k})\phi(\mathfrak{b}_{\ell,k}(\theta_{\ell,k}))}{Z_{\ell,k}}, \tag{3.23}$$

and

$$\mathfrak{a}_{\ell,k}(\theta_{\ell,k}) \triangleq \frac{a_{\ell,k} - s_{\ell,k}}{\mathrm{A}\theta_{\ell,k}\|\mathbf{h}_\ell^k\|_2}, \tag{3.24}$$

$$\mathfrak{b}_{\ell,k}(\theta_{\ell,k}) \triangleq \frac{b_{\ell,k} - s_{\ell,k}}{\mathrm{A}\theta_{\ell,k}\|\mathbf{h}_\ell^k\|_2}. \tag{3.25}$$

Lemma 13 (Distribution Parameter Solution). *For $\bar{X} \in \mathcal{I}_{\ell,k}$ distributed according to (3.17) with parameter $\theta_{\ell,k}$, then $\theta_{\ell,k}$ is the unique solution to the equation*

$$\eta_{\ell,k}(\theta_{\ell,k}) = \frac{\mu(\lambda, \ell, k) - \alpha_{\ell,k}}{\theta_{\ell,k}^2 \cdot (1-\lambda)} - 1, \tag{3.26}$$

Proof. The form of (3.26) is a simple reformulation of (3.16) under max-entropy distribution. Uniqueness is given by monotonicity of $\eta_{\ell,k}(\theta_{\ell,k})$ in $\theta_{\ell,k}$, which can be observed in Figure 3.1. □

3.2.3 Evaluating Differential Entropy

The differential entropy of $\bar{X} \in \mathcal{I}_{\ell,k}$ is that of a normal distribution of mean $s_{\ell,k}$ and variance $\mathrm{A}^2\theta_{\ell,k}^2\|\mathbf{h}_\ell^k\|_2^2$, truncated over $\mathcal{I}_{\ell,k}$. It is given by

$$\begin{aligned}
\mathsf{h}(\bar{X} \in \mathcal{I}_{\ell,k}) &= \log(\sqrt{2\pi e}\mathrm{A}\theta_{\ell,k}\|\mathbf{h}_\ell^k\|_2 Z_{\ell,k}) \\
&\quad + \underbrace{\frac{\mathfrak{a}_{\ell,k}(\theta_{\ell,k})\phi(\mathfrak{a}_{\ell,k}(\theta_{\ell,k})) - \mathfrak{b}_{\ell,k}(\theta_{\ell,k})\phi(\mathfrak{b}_{\ell,k}(\theta_{\ell,k}))}{2Z_{\ell,k}}}_{\triangleq \delta_{\ell,k}(\theta_{\ell,k})} \\
&= \log(\sqrt{2\pi e}\mathrm{A}\theta_{\ell,k}\|\mathbf{h}_\ell^k\|_2 Z_{\ell,k}) + \delta_{\ell,k}(\theta_{\ell,k}),
\end{aligned} \tag{3.27}$$

with $\theta_{\ell,k}$ being the solution to (3.26) for each (ℓ, k). Thus

$$\begin{aligned}
\mathsf{h}(\bar{X}) &= \mathsf{H}(\mathbf{p}) + \log(\sqrt{2\pi e}\mathrm{A}) + \sum_{\ell,k} p_{\ell,k} \log\left(\underbrace{\theta_{\ell,k}\|\mathbf{h}_\ell^k\|_2 Z_{\ell,k} e^{\delta_{\ell,k}(\theta_{\ell,k})}}_{\triangleq Q_{\ell,k}(\theta_{\ell,k})}\right) \\
&= -\mathscr{D}\left(\mathbf{p}\| \frac{Q_{\ell,k}(\theta_{\ell,k})}{\sum_{\ell,k} Q_{\ell,k}(\theta_{\ell,k})}\right) + \log(\sqrt{2\pi e}\mathrm{A}) + \sum_{\ell,k} \log Q_{\ell,k}(\theta_{\ell,k})
\end{aligned} \tag{3.28}$$

In order to maximize (3.15), we must now optimize over both the probability mass vector $\mathbf{p}$ and the per-interval energy allocation $\boldsymbol{\alpha}$, while satisfying $\boldsymbol{\alpha}^\mathsf{T}\mathbf{p} = \alpha$. This culminates to the following capacity lower bound.

Proposition 14 (Lower Bound when $\alpha < \alpha_{\mathsf{th}}$). *When $\alpha < \alpha_{\mathsf{th}}$, the channel capacity is lower-bounded by*

$$\mathcal{C}_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathrm{A}, \alpha) \geq \log\left(1 + \frac{\mathrm{A}^2 e^{2\tau}}{\sigma^2}\right) \tag{3.29}$$

with

$$\tau \triangleq \sup_{\substack{\mathbf{p}:\sum_{\ell,k} p_{\ell,k}=1, \\ \boldsymbol{\alpha}:\sum_{\ell,k} p_{\ell,k}\alpha_{\ell,k}=\alpha}} \left\{-\mathscr{D}\left(\mathbf{p}\| \frac{Q_{\ell,k}(\theta_{\ell,k})}{\sum_{\ell,k} Q_{\ell,k}(\theta_{\ell,k})}\right) + \sum_{\ell,k} \log Q_{\ell,k}(\theta_{\ell,k})\right\} \tag{3.30}$$

and where each $\theta_{\ell,k}$ is the solution to (3.26) for each (ℓ, k).

Unfortunately, an analytical method to perform the maximization of (3.30) has yet to be found.

Chapter 4

Numerical Simulation

Due to the lack of analytical optimization methods so far, we have attempted to compute the maximization of $\mathsf{h}(\bar{X})$ numerically. This section will cover the implementation and results obtained by simulating a toy example. The code was written in the Python programming language.

4.1 Example Setup

We will consider a two-antenna system, with the following channel state vector

$$\mathbf{h} \triangleq [2.0, 1.0]^{\mathsf{T}}, \tag{4.1}$$

and unit channel gain (i.e. $\mathsf{A} = 1$). The goal was to implement a simulation with a partition of $\bar{\mathcal{X}} = [0, 3.0]$ into three sub-intervals. The first one corresponding to a transmission with only antenna 1, the second interval corresponding to shifted beamforming mode, and the third corresponding to antenna 1 set to full, while modulating only antenna 2. To this goal, we will fix

$$\lambda \triangleq 0.25. \tag{4.2}$$

Such a choice of λ sets the following sub-intervals:

- $\mathcal{I}_{1,1} \triangleq [0, \frac{1}{3}]$
- $\mathcal{I}_{1,2} \triangleq [\frac{1}{3}, 2.41\bar{6}]$
- $\mathcal{I}_{2,2} \triangleq [2.41\bar{6}, 3.0]$

Remark 15. Since the double-indexing is quite cumbersome for such a small example, we will not employ it, and rather denote the intervals with single indices: $1, 2$ and 3. △

Remark 16. λ was chosen arbitrarily, and multiple tests with other λ-values which still generated a partition into three sub-intervals have been performed. We discuss this later. △

4.2 Numerical Solver

Algorithm 1 is a rather straightforward way to solve the optimization problem in Eq. (3.30). It is a simple template we have used as a stepping stone to hopefully inspire better solvers.

Algorithm 1 Optimization Algorithm for a partition of $\overline{\mathcal{X}}$ into 3 sub-intervals

---

Input: $\alpha, \lambda, \mathbf{h}$
  $\mathbf{\Theta} \leftarrow [0, 0, 0]$
  $\mathbf{p} \leftarrow [0, 0, 0]$
  $\mathsf{h}_{\mathsf{max}} \leftarrow -\infty$
  $\boldsymbol{\alpha} \leftarrow [0, 0, 0]$
  for $p_1 \in [0, 1]$ **do** ▷ Iterating over probabilities
    for $p_2 \in [0, 1 - p_1]$ **do**
      $p_3 \leftarrow 1 - p_1 - p_2$
      for $\alpha_1 \in [0, \frac{\alpha}{p_1}]$ **do** ▷ Iterating over energy allocation
        for $\alpha_2 \in [0, \frac{\alpha - \alpha_1 \cdot p_1}{p_2}]$ **do**
          $\alpha_3 \leftarrow (\alpha - p_1 \cdot \alpha_1 - p_2 \cdot \alpha_2)/p_3$
          $\theta_i^2 \leftarrow$ result of solving (3.26) for every i
          $\mathsf{h} \leftarrow \mathsf{h}(\bar{X})$
          if $\mathsf{h} > \mathsf{h}_{\mathsf{max}}$ **then** ▷ Storing optimal result so-far
            $\mathsf{h}_{\mathsf{max}} \leftarrow \mathsf{h}$
            $\mathbf{p} \leftarrow [p_1, p_2, p_3]$
            $\boldsymbol{\alpha} \leftarrow [\alpha_1, \alpha_2, \alpha_3]$
            $\mathbf{\Theta} \leftarrow [\theta_1^2, \theta_2^2, \theta_3^2]$
          end if
        end for
      end for
    end for
  end for
Output: $\mathbf{p}, \boldsymbol{\alpha}, \mathbf{\Theta}, \mathsf{h}_{\mathsf{max}}$

Remark 17. For the sake of brevity, some fine-tuning in Algorithm 1 has been omitted. Namely that α_2 (resp. α_3) must be zero if p_2 (resp. p_3) equals zero; in which case we solve Eq. (3.26) with both equal to zero. Also, we have not given implementations of computing $\mathsf{h}(\bar{X})$ or finding the roots of (3.26). The former was computer in a straightforward code translation of (3.15). The latter was numerically performed using the built-in root-finder `scipy.optimize.fsolve`. △

4.3 Results and Discussion

Tables 4.1 and 4.2 will present the obtained results, and Figure 4.1 visualizes Table 4.1. Since the most error-prone part of the optimization lies in the finding the $\theta_{\ell,k}$-parameters (distribution parameters over every $\mathcal{I}_{\ell,k}$), we have quantified the

Fixed $\lambda = 0.25$			
Values	**Allocated Energy** α		
	0.4	0.42	0.47
θ_1^2	0.0395821	0.1037130991	0.133155856415
θ_2^2	0.15206302	0.1038154926	0.116576025792
θ_3^2	0.26747476	0.3278110617	0.731420239186
α_1	0.02641053	0.0272526315	0.027368421052
α_2	0.4446739	0.4062559225	0.419050833020
α_3	1.46545951	1.4830114177	1.531166889124
p_1	0.21052711	0.21428649	$0.142857\bar{9}$
p_2	0.74697437	0.6975232	$0.761015\bar{2}$
p_3	0.04249852	0.08819027	$0.096126\bar{7}$
$\mathsf{h}_{\mathsf{max}}$	0.93777075	0.95930191	1.00825641
$\varepsilon(\theta_1)$	$-1.5265 \cdot 10^{-16}$	$-1.94289 \cdot 10^{-16}$	$-1.3877 \cdot 10^{-16}$
$\varepsilon(\theta_2)$	$-4.7184 \cdot 10^{-15}$	$-4.291012 \cdot 10^{-14}$	$-1.2379 \cdot 10^{-12}$
$\varepsilon(\theta_3)$	$-5.5511 \cdot 10^{-15}$	$-1.830979 \cdot 10^{-12}$	$-4.1633 \cdot 10^{-14}$

Table 4.1: Simulation Results using Algorithm 1 for a fixed λ and varying α

accuracy of our computed solution by simply evaluating Eq. (3.26) with the obtained $\theta_{\ell,k}$, as to make sure that the solver has actually converged. We denote this accuracy by $\varepsilon(\theta_{\ell,k})$, which corresponds to the difference between the computed and the expected result.

First off, we notice that the maximal differential entropy increases as we allow more energy, which is consistent with monotonic increase of transmission capabilities in energy. Secondly, the accuracy of the solver seems quite good, as the error lies in the order of machine-precision errors.

Lastly, recalling Section 1.2; we had briefly mentioned that both α and λ are fixed. We note that it is also of interest how the capacity changes with a *fixed* energy allocation and varying beamforming intervals $\mathcal{I}_{\ell,k}$ induced by λ. Due to this, we have also performed numerical optimizations over a fixed α, and varying λ. Results are given in Table 4.2 and displayed in Figure 4.2. Note that he different choices of λ yielded different partitions of $\bar{\mathcal{X}}$, which are given in Table 4.3.

Remark 18 (Reducing the Search Space)**.** A clear way to yield better results is to reduce the range of the iterations, especially for the α-components. As no analytical bounds were of any meaning, we have simply proceeded by manually reducing the search space. Another option would have been to allow the process to run over more iterations over the same bounds. We have opted for the former, since the latter yielded exponential runtime increases. $\triangle$

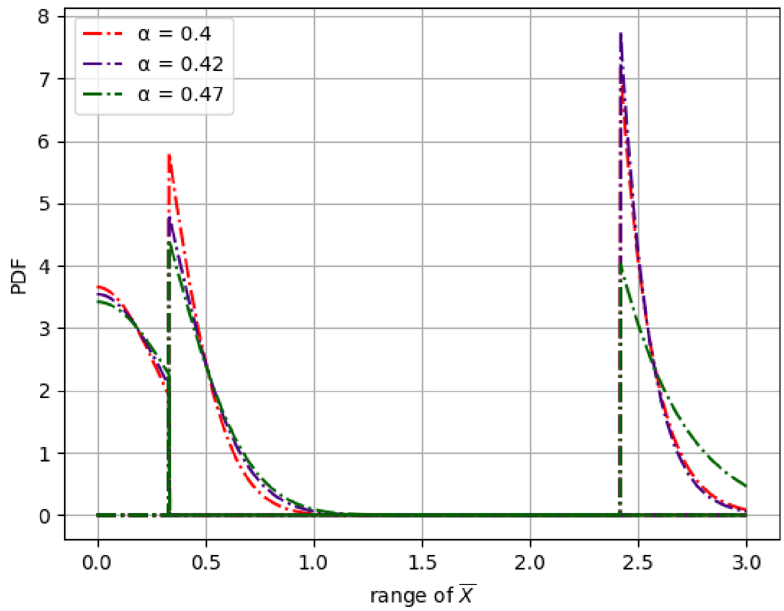

Figure 4.1: The simulated differential entropy maximizing distribution of $\bar{X}$ using results from Table (4.1). ($\lambda = \frac{1}{4}$)

Fixed $\alpha = 0.4$			
Values	**Trade-off Parameter λ**		
	0.3	0.34	0.35
θ_1^2	0.0766106	0.05737169	0.11334305
θ_2^2	0.14184885	0.09866077	0.12270888
θ_3^2	1.53114158	0.16041785	0.4193806
α_1	0.04105263	0.05368421	0.06
α_2	0.47014571	0.46171853	0.49352713
α_3	1.54680101	1.38242224	1.47597066
p_1	0.26315863	0.31579016	0.36842168
p_2	0.69711293	0.61129312	0.56420176
p_3	0.03972844	0.07291672	0.06737656
$\mathsf{h}_{\mathsf{max}}$	0.92822555	0.91764985	0.920063045
$\varepsilon(\theta_1)$	$-5.140332 \cdot 10^{-14}$	$-1.290634 \cdot 10^{-15}$	$-2.63123 \cdot 10^{-14}$
$\varepsilon(\theta_2)$	$-1.11022 \cdot 10^{-16}$	$-6.328271 \cdot 10^{-15}$	$-3.79341 \cdot 10^{-12}$
$\varepsilon(\theta_3)$	$-3.33067 \cdot 10^{-16}$	$-1.88088 \cdot 10^{-12}$	$-1.69864 \cdot 10^{-14}$

Table 4.2: Simulation Results using Algorithm 1 for a fixed α and varying λ.

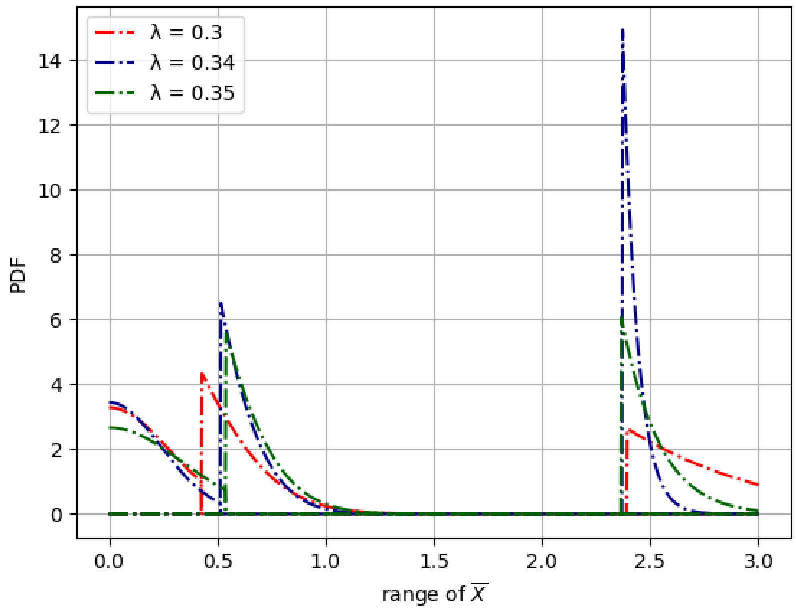

Figure 4.2: The simulated differential entropy maximizing distribution of $\bar{X}$ using a fixed energy allocation $\alpha = \frac{2}{5}$ and varying trade-off parameter λ using results from Table 4.2. Conversely to Table 4.1, the distributions are not defined over the same sub-intervals. These are given in Table 4.3.

Partition	$\lambda = 0.3$	$\lambda = 0.34$	$\lambda = 0.35$
$\mathcal{I}_1$	[0, 0.42857]	[0, 0.51515]	[0, 0.53846]
$\mathcal{I}_2$	[0.42857, 2.39285]	[0.51515, 2.3712]	[0.53846, 2.36538]
$\mathcal{I}_3$	[2.3928571, 3.0]	[2.3712, 3.0]	[2.36538, 3.0]

Table 4.3: Induced partitions of the channel image $[0, 3]$ for various choices of λ. The same desired structure of a partition into 3 sub-intervals is still held.

Chapter 5

Conclusions and Outlooks

This semester thesis was a first dive into the derivations of upper- and lower-bounds on the capacity of MISO optical free-space channels with additive Gaussian noise, where the channel input vector is subjected to peak-power, first- and second-moment constraints. We have found that capacity bounds are derivable by maximizing the differential entropy $\mathsf{h}(\mathbf{h}^\mathsf{T}\mathbf{X})$. The optimal input is found by considering minimum-energy input vectors (given by [3]) and that conditional on a given range $\mathcal{I}_{\ell,k}$, the channel image should be distributed as a truncated Gaussian. The optimization is thus performed over a partition of the energy constraint $\boldsymbol{\alpha}$ and the probability mass vector $\mathbf{p}$ such that $\boldsymbol{\alpha}^\mathsf{T}\mathbf{p} = \alpha$. The distribution parameters are found by solving a per-interval energy constraint, given by each component of $\boldsymbol{\alpha}$.

Looking at Eq. (3.30), we notice that various optimization methods are possible. Due to the non-linear nature of the constraints, no linear program seems to be a good solution. However, the problem could potentially be solved using a non-linear optimization technique, such as the Karush-Kuhn-Tucker conditions. This is just considerations, and has not been tested by any means.

We have opted for a numerical simulation of a simple example, as to hopefully give us some insights on analytical results. The expected hypothesis of differential entropy increasing under more allowed energy with a fixed λ seems to hold, as seen in Table (4.1). We note however that the nature of the problem is more as seen in Table (4.2), where we can see that certain values of λ will yields higher differential entropy. We suggest that this approach be taken in further tests. By fixing α and iterating over λ, we could successfully find the optimal trade-off parameter for a certain energy constraint.

A clear extension aside from finding an optimization routine to solve Eq. (3.15) would be to find upper bounds to the channel capacity. An idea developed but not concluded during this work was to mimic the upper bound derivations in [5, Proposition 10] with the SISO capacity upper-bound offered in [2]. An optimal distribution of $\bar{X}$ is however still required. With both results in hand, asymptotic results for high-SNR could be derived.

Appendix A

Calculations

A.1 Developing Interval-wise Constraints

Here we will offer the full calculations leading to Eq. (3.26). We begin by computing

$$\begin{aligned}
\mathsf{E}_{f^*_{\bar{X}_{\ell,k}}}[\tilde{g}(\bar{X})|V=(\ell,k)] &= \int_{a_{\ell,k}}^{b_{\ell,k}} \frac{(\bar{x}-s_{\ell,k})^2}{\mathsf{A}^2\|\mathbf{h}_\ell^k\|_2^2} \cdot f^*_{\bar{X}_{\ell,k}}(\bar{x})\,\mathrm{d}\bar{x} \\
&= \frac{1}{\mathsf{A}^3\|\mathbf{h}_\ell^k\|_2^3\theta_{\ell,k}Z_{\ell,k}} \int_{a_{\ell,k}}^{b_{\ell,k}} (\bar{x}-s_{\ell,k})^2 \cdot e^{-\frac{(\bar{x}-s_{\ell,k})^2}{2\mathsf{A}^2\theta_{\ell,k}^2\|\mathbf{h}_\ell^k\|_2^2}}\,\mathrm{d}\bar{x} \\
&= \theta_{\ell,k}\frac{(b_{\ell,k}-s_{\ell,k})e^{-\frac{(b_{\ell,k}-s_{\ell,k})^2}{2\mathsf{A}^2\theta_{\ell,k}^2\|\mathbf{h}_\ell^k\|_2^2}}}{\sqrt{2\pi}\mathsf{A}\|\mathbf{h}_\ell^k\|_2 Z_{\ell,k}} \\
&\quad - \theta_{\ell,k}\frac{(a_{\ell,k}-s_{\ell,k})e^{-\frac{(a_{\ell,k}-s_{\ell,k})^2}{2\mathsf{A}^2\theta_{\ell,k}^2\|\mathbf{h}_\ell^k\|_2^2}}}{\sqrt{2\pi}\mathsf{A}\|\mathbf{h}_\ell^k\|_2 Z_{\ell,k}} + \theta_{\ell,k}^2 && \text{(A.1)} \\
&= \theta_{\ell,k}^2\Bigg(\frac{\mathfrak{a}_{\ell,k}(\theta_{\ell,k})\phi(\mathfrak{a}_{\ell,k}(\theta_{\ell,k}))}{Z_{\ell,k}} \\
&\quad - \frac{\mathfrak{b}_{\ell,k}(\theta_{\ell,k})\phi(\mathfrak{b}_{\ell,k}(\theta_{\ell,k}))}{Z_{\ell,k}}\Bigg) + \theta_{\ell,k}^2 && \text{(A.2)} \\
&= \theta_{\ell,k}^2\Big(\eta_{\ell,k}(\theta_{\ell,k})+1\Big) && \text{(A.3)}
\end{aligned}$$

Bibliography

[1] A. Lapidoth, S. M. Moser, and M. A. Wigger, "On the Capacity of Free-Space Optical Intensity Channels," *IEEE Transactions on Information Theory*, vol. 55, no. 10, pp. 4449–4461, 2009.

[2] S. Ma and M. Wigger, "First- and Second-Moment Constrained Gaussian Channels," in *2021 IEEE International Symposium on Information Theory (ISIT)*, 2021, pp. 432–437.

[3] S. Ma, S. M. Moser, L. Wang, and M. Wigger, "Signaling for MISO Channels Under First- and Second-Moment Constraints," in *2022 IEEE International Symposium on Information Theory (ISIT)*, 2022, pp. 2648–2653.

[4] C. E. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 1948. [Online]. Available: http://plan9.bell-labs.com/cm/ms/what/shannonday/shannon1948.pdf

[5] S. M. Moser, L. Wang, and M. Wigger, "Capacity Results on Multiple-Input Single-Output Wireless Optical Channels," *IEEE Transactions on Information Theory*, vol. 64, no. 11, pp. 6954–6966, 2018.

[6] S. M. Moser, M. Mylonakis, L. Wang, and M. Wigger, "Asymptotic capacity results for mimo wireless optical communication," in *2017 IEEE International Symposium on Information Theory (ISIT)*, 2017, pp. 536–540.

[7] T. M. Cover and J. A. Thomas, *Elements of Information Theory 2nd Edition (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, July 2006.