MISO Wireless Optical Communication under a First-Moment and a Peak-Power Constraint: Theoretical and Practical Aspects

Jordan Hong* and Stefan M. Moser*†
*Signal and Information Processing Lab, ETH Zurich, Switzerland
†Institute of Communications Engineering, National Yang Ming Chiao Tung University, Hsinchu, Taiwan
Email: johong@student.ethz.ch, moser@isi.ee.ethz.ch

***Abstract*—This paper studies the multiple-input single-output free-space optical intensity channel with signal-independent additive Gaussian noise and subject to both an average- and a peak-intensity constraint. Closed-form expressions for the asymptotic high-signal-to-noise-ratio (high-SNR) capacity and for the corresponding capacity-achieving input distribution are presented. Moreover, several practical modulation schemes are proposed that approach the capacity over a wide range of practical SNR values.**

I. Introduction

Optical communication systems based on simple intensity-modulation and direct-detection (*IM-DD systems*) have received much attention in both theoretical and practical research [1]–[15] due to their ability to achieve high transmission rates with simple implementations.

This work studies the multiple-input single-output (MISO) optical intensity channel model with signal-independent additive Gaussian noise, subject to both an average- and a peak-intensity constraint (Section II). To date, the capacity of this channel is not fully characterized. In [14], this channel was investigated and the asymptotic capacity in the limit when the signal-to-noise ratio (SNR) tends to infinity was derived. However, the presented capacity expression in [14, Prop. 5] contains a numerical optimization over one parameter. In Section III, we review the equivalent SISO formulation introduced in [14]. The first contribution of this work is the derivation of a closed-form expression for this asymptotic capacity and its corresponding capacity-achieving input distribution (Section IV).

The second contribution (Sections V–VI) of this paper addresses practical aspects of this optical communication channel by introducing a modulation scheme that fixes the input distribution and optimizes the constellation size. The scheme is designed to achieve near-capacity rates across a wide range of practical SNR values. In Section V, we present the proposed scheme, which fixes the input distribution to be entropy-maximizing while adjusting the number of constellation points for different SNR values to maximize mutual information. Section VI further simplifies the selection of constellation size by using the Ozarow–Wyner lower bound [16]. In its simplest form, our approach reduces to a dichotomous system that consists of only two predefined constellations and input distributions. The choice between the two input distributions is governed by a single SNR threshold derived from the Ozarow–Wyner lower bound. Numerical results presented in Section VII demonstrate that the proposed modulation scheme achieve rates near the channel capacity across a wide range of practical SNR values.

II. Channel Model

We consider a MISO channel of the form

$$Y = \mathbf{h}^{\mathsf{T}}\mathbf{x} + Z, \tag{1}$$

where Y denotes the channel output, $\mathbf{x} = (x_1, \ldots, x_{n_\mathrm{T}})^{\mathsf{T}}$ denotes the channel input vector, $\mathbf{h}$ is the fixed channel state vector, and $Z \sim \mathcal{N}(0, \sigma^2)$ is additive Gaussian noise. The n_T input random variables are proportional to the optical intensities of LEDs and are therefore nonnegative:

$$X_k \in \mathbb{R}_0^+, \quad k = 1, \ldots, n_\mathrm{T}. \tag{2}$$

Similarly, $\mathbf{h} = (h_1, h_2, \ldots, h_{n_\mathrm{T}})$ has positive entries and, without loss of generality, is assumed to be ordered:

$$h_1 \geq h_2 \geq \cdots \geq h_{n_\mathrm{T}} > 0. \tag{3}$$

The input $\mathbf{X}$ is subject to both a per-LED peak-power and an average-power constraint:

$$\mathbb{P}[X_k > \mathsf{A}] = 0, \quad \forall\, k \in \{1, \ldots, n_\mathrm{T}\}, \tag{4}$$

$$\sum_{k=1}^{n_\mathrm{T}} \mathbb{E}[X_k] \leq \mathsf{E}, \tag{5}$$

for some fixed parameters $\mathsf{A}, \mathsf{E} > 0$, where we define

$$\alpha \triangleq \frac{\mathsf{E}}{\mathsf{A}} \in (0, n_\mathrm{T}]. \tag{6}$$

Here, $\mathbb{E}[\cdot]$ denotes the expectation operator. Note that contrary to conventional systems (where an average-power constraint is on the expectation of the squared channel inputs), here the average-power constraint is on the expectation of the channel inputs because the inputs are proportional to optical intensity, i.e. the energy induced by each input vector $\mathbf{X}$ is given by: $\sum_{k=1}^{n_\mathrm{T}} X_k$. We define the SNR as A/σ^2.

The capacity of the channel (1) under the given constraints is given by [17]:

$$\mathsf{C}_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathsf{A},\mathsf{E}) = \sup_{Q_\mathbf{X}} \mathrm{I}(\mathbf{X};Y), \tag{7}$$

where the supremum is over all channel input laws $Q_\mathbf{X}$ on $\mathbf{X}$ satisfying (2), (4), and (5).

III. Energy-Optimal Signaling and Equivalent Capacity

We introduce the shorthands

$$s_0 \triangleq 0 \tag{8a}$$

$$s_k \triangleq \sum_{j=1}^{k} h_j, \qquad k \in \{1,\dots,n_\mathrm{T}\}, \tag{8b}$$

and define

$$\bar{X} \triangleq \mathbf{h}^\mathsf{T}\mathbf{X} = \sum_{k=1}^{n_\mathrm{T}} h_k X_k. \tag{9}$$

Then we define the random variable U over the alphabet $\{1,\dots,n_\mathrm{T}\}$ to indicate the interval in which $\bar{X}$ lies:

$$\Big(U=1\Big) \iff \Big(\bar{X} \in [\mathsf{A}s_0, \mathsf{A}s_1]\Big), \tag{10a}$$

and for $k \in \{2,\dots,n_\mathrm{T}\}$:

$$\Big(U=k\Big) \iff \Big(\bar{X} \in (\mathsf{A}s_{k-1}, \mathsf{A}s_k]\Big). \tag{10b}$$

Note that $\bar{X}$ can be viewed as the scalar constellation being sent through the AWGN channel. Indeed, $\mathbf{X} \multimap\!\!- \bar{X} \multimap\!\!- Y$ form a Markov chain and $\bar{X}$ is a deterministic function of $\mathbf{X}$. Thus, we have

$$\mathrm{I}(\mathbf{X};Y) = \mathrm{I}(\bar{X};Y). \tag{11}$$

Hence, the MISO channel (1) is equivalent to a SISO channel with input $\bar{X}$ and output $Y = \bar{X} + Z$ with the power-constraints (4) and (5) on $\mathbf{X}$ transformed to a set of admissible distributions for $\bar{X}$. For any given $\bar{X}$, there exists a unique energy-optimal construction that is summarized in the following lemma [14].

Lemma 1 (Energy-Optimal Construction of $\bar{X}$ [14, Lem. 4]): Let $\bar{X}$ and U be defined as in (9) and (10). The energy-optimal construction of $\bar{X}$ is:

$$X_1 = \cdots = X_{k-1} = \mathsf{A} \tag{12a}$$

$$X_k = \left.\frac{\bar{X} - \mathsf{A}s_{k-1}}{h_k}\right|_{U=k} \tag{12b}$$

$$X_{k+1} = \cdots = X_{n_\mathrm{T}} = 0. \tag{12c}$$

We emphasize that using this construction, there is a unique correspondence between $\mathbf{X}$ and $\bar{X}$. Consequently, both the peak-power constraint and the average-power constraint on $\mathbf{X}$ can be expressed solely in terms of $\bar{X}$. In particular, the average power induced by a specific value of $\bar{X}$ is given by the *energy-generating function*

$$g(\bar{X}) \triangleq \sum_{k=1}^{n_\mathrm{T}} \left(\frac{\bar{X} - s_{k-1}\mathsf{A}}{h_k} + (k-1)\mathsf{A} \right) \mathbb{1}[U=k], \tag{13}$$

where $\mathbb{1}[\cdot]$ denotes the indicator function. Note that $g(\bar{X})$ is defined on $[0, n_\mathrm{T}\mathsf{A}]$ and is continuous, convex, and piece-wise linear with the slope in each segment being $1/h_k$.

It follows that we can express (7) using the following equivalent SISO capacity on $\bar{X}$.

Lemma 2 (Equivalent SISO Capacity [14, Prop. 3]):

$$\mathsf{C}_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathsf{A},\alpha\mathsf{A}) = \max_{Q_{\bar{X}}} \mathrm{I}\big(\bar{X};Y\big) \tag{14}$$

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

$$\mathbb{P}\big[\bar{X} > s_{n_\mathrm{T}}\mathsf{A}\big] = 0 \tag{15a}$$

and

$$\mathbb{E}\big[g(\bar{X})\big] \le \alpha\mathsf{A}. \tag{15b}$$

Note that there exists a threshold value α_th such that if $\alpha > \alpha_\text{th}$, the constraint (15b) becomes inactive. From [14, Sec. IV], we know that

$$\alpha_\text{th} \triangleq \frac{1}{2} + \frac{1}{s_{n_\mathrm{T}}} \sum_{k=1}^{n_\mathrm{T}} h_k(k-1). \tag{16}$$

IV. Maximum Entropy Distribution

In this section, we will focus on the situation when both peak- and average-power constraints are active, i.e., $\alpha < \alpha_\text{th}$.

Proposition 3 (Maximum SISO Entropy): For a nonnegative continuous random variable X with peak-constraint A and first-moment constraint $\alpha\mathsf{A}$, the entropy-maximizing distribution is given by

$$Q(x) = \frac{1}{\kappa_\mathsf{A}(\mu)}\, e^{-\frac{\mu x}{\mathsf{A}}}, \qquad 0 \le x \le \mathsf{A}, \tag{17}$$

where $\kappa_\mathsf{A}(\mu)$ is the normalizing factor

$$\kappa_\mathsf{A}(\mu) = \int_0^\mathsf{A} e^{-\frac{\mu x}{\mathsf{A}}}\,\mathrm{d}x = \frac{\mathsf{A}(1-e^{-\mu})}{\mu}, \tag{18}$$

and μ is chosen such that the first-moment constraint $\alpha\mathsf{A}$ is satisfied, i.e., μ is the unique solution to

$$\alpha = \frac{1}{\mu} - \frac{e^{-\mu}}{1-e^{-\mu}}. \tag{19}$$

The corresponding maximum differential entropy is

$$\mathrm{h}(Q) = \log\mathsf{A} + \log\left(\frac{1-e^{-\mu}}{\mu}\right) + \mu\alpha. \tag{20}$$

Proof: Omitted for space reasons. ∎

In [14, Sec. IV.A], a lower bound on the MISO capacity is presented and shown to be tight at high SNR, achieving the exact capacity in the limit when $\mathsf{A} \to \infty$. The derivation of this bound is based on the maximum SISO entropy results presented in Proposition 3. However, the given expression is not in closed form, but contains a numerical optimization. We will next provide a closed-form expression for this distribution and the corresponding value of the asymptotic high-SNR capacity.

Proposition 4 (Optimal Distribution of $\bar{X}$): Under the constraints defined in (15), the entropy-maximizing distribution of $\bar{X}$ is

$$Q(\bar{x}) = \frac{1}{\mathsf{K}_{\mathsf{A}}(\mu)}\, e^{-\frac{\mu g(\bar{x})}{\mathsf{A}}}, \qquad 0 \le \bar{x} \le \mathsf{A} s_{n_{\mathrm{T}}}, \tag{21}$$

with the normalizing factor

$$\mathsf{K}_{\mathsf{A}}(\mu) = \kappa_{\mathsf{A}}(\mu) \sum_{k=1}^{n_{\mathrm{T}}} h_k\, e^{-\mu(k-1)} \tag{22}$$

and where μ is chosen such that the first-moment constraint is satisfied:

$$\int_0^{s_{n_{\mathrm{T}}}} g(\bar{x})\, Q(\bar{x})\, \mathrm{d}\bar{x} = \alpha \mathsf{A}, \tag{23}$$

i.e., μ satisfies

$$\frac{1}{\mu} - \frac{1}{e^{\mu}-1} + \sum_{k=1}^{n_{\mathrm{T}}} p_k (k-1) = \alpha, \tag{24}$$

where

$$p_k \triangleq \mathbb{P}[U=k] = \frac{h_k\, e^{-\mu k}}{\sum_{j=1}^{n_{\mathrm{T}}} h_j\, e^{-\mu j}}, \quad k \in \{1, \dots, n_{\mathrm{T}}\}. \tag{25}$$

The corresponding maximum differential entropy is

$$\mathsf{h}(Q) = \log \mathsf{K}_{\mathsf{A}}(\mu) + \mu\alpha \tag{26}$$

$$= \log \kappa_{\mathsf{A}}(\mu) + \log\left(\sum_{k=1}^{n_{\mathrm{T}}} h_k\, e^{-\mu(k-1)} \right) + \mu\alpha. \tag{27}$$

Proof: Omitted for space reasons. ∎

We remark that in [14], $\mathsf{h}(\bar{X})$ is maximized by taking the supremum over μ, but the explicit form of the entropy-maximizing distribution was not determined. In Proposition 4, we show that the entropy-maximizing distribution is exponential in $g(\bar{x})$ with the exponential parameter μ. Thus, [14, (34)] can be shown to be precisely:

$$v = \mathsf{h}\left(\frac{\bar{X}}{s_{n_{\mathrm{T}}} \mathsf{A}} \right) \tag{28}$$

$$= 1 - \log \frac{\mu}{1-e^{-\mu}} - \frac{\mu\, e^{-\mu}}{1-e^{-\mu}} - \mathscr{D}\left(\mathbf{p} \,\middle\|\, \frac{\mathbf{h}}{s_{n_{\mathrm{T}}}} \right), \tag{29}$$

where $\mathscr{D}(\cdot\|\cdot)$ denotes relative entropy, μ satisfies (24), and $\mathbf{p}$ is the probability vector with components p_k. By [14, Prop. 5], the capacity can be lower-bounded by

$$\mathsf{C}_{\mathbf{h}^{\mathsf{T}},\sigma^2}(\mathsf{A}, \alpha\mathsf{A}) \ge \frac{1}{2} \log\left(1 + \frac{\mathsf{A}^2 s_{n_{\mathrm{T}}}^2}{2\pi e \sigma^2}\, e^{2v} \right). \tag{30}$$

It is interesting to note that the optimal μ is also used for the optimal distribution of X_k conditional on $U = k$.

Proposition 5 (Optimal Distribution of X_k): Conditional on $U = k$, the optimal[1] X_k has an exponential distribution of the form (17), with μ chosen identically as for $Q_{\bar{X}}$:

$$Q_{X_k|U=k}(x_k) = \frac{1}{\kappa_{\mathsf{A}}(\mu)}\, e^{-\frac{\mu}{\mathsf{A}} x_k}. \tag{31}$$

Proof: Omitted for space reasons. ∎

[1]Recall that according to (12), the other random variables X_j, $j \ne k$, are deterministically equal to 0 or A.

V. Practical Discrete Modulation: Optimization of Constellation Size for Fixed Distribution

The maximum-entropy analysis in the previous section implicitly assumes a continuous $Q_{\bar{X}}$. However, in practical digital communication systems, $\bar{X}$ must be chosen to be discrete, taking value from a finite constellation $\bar{\mathcal{X}}$. We denote the discrete distribution on $\bar{\mathcal{X}}$ by $q_{\bar{X}}$.

In this section, we will discuss some practical design considerations. As it is our objective of maximizing $\mathrm{I}(\bar{X};Y)$, two design parameters arise: 1) the choice of constellation $\bar{\mathcal{X}}$, and 2) the choice of the discrete distribution $q_{\bar{X}}$.

A possible approach is to fix some number N of evenly spaced constellation points, and then numerically optimize the probability distribution over these points (e.g., via gradient search) by leveraging the concavity of mutual information in the input distribution. The main drawback of this method lies in its computational complexity. Since the optimal distribution depends on the SNR, the optimization must be repeated for each value of A. Moreover, selecting N poses additional challenges because it also depends on A: a value of N that is too small results in a suboptimal rate, while excessively large values add unnecessary complexity without improving the rate.

We propose a simpler approach that focuses on optimizing the constellation size N, while maintaining a fixed input distribution based on the entropy-maximizing distribution from Section IV. This choice of distribution is known to be effective in the high-SNR regime as it achieves the asymptotic high-SNR capacity [14]. In the low-SNR region, the optimization over N will lead to a choice $\mathsf{N} = 2$, i.e., binary signaling on 0 and $s_{n_{\mathrm{T}}}\mathsf{A}$, which approximates the optimal variance-maximizing distribution closely.[2]

In detail, the constellation $\bar{\mathcal{X}}$ is chosen to consist of N evenly spaced points over $[0, s_{n_{\mathrm{T}}}\mathsf{A}]$:

$$\bar{\mathcal{X}} \triangleq \left\{ 2i\Delta \colon \Delta = \frac{s_{n_{\mathrm{T}}}\mathsf{A}}{2(\mathsf{N}-1)},\ i = 0, \dots, \mathsf{N}-1 \right\}, \tag{32}$$

where Δ corresponds to the half-spacing between two neighboring constellation points. The discrete distribution $q_{\bar{X}}$ is chosen to be discrete-entropy-maximizing, i.e.,

$$q_{\bar{X}}(\bar{x}) = \frac{1}{\mathsf{K}(\mu)}\, e^{-\frac{\mu}{\mathsf{A}} g(\bar{x})}, \qquad \bar{x} \in \bar{\mathcal{X}}, \tag{33}$$

where $\mathsf{K}(\mu)$ is the normalizing factor[3] and μ is selected in accordance with the average-power constraint (15b):

$$\sum_{\bar{x} \in \bar{\mathcal{X}}} \frac{g(\bar{x})}{\mathsf{A}}\, q_{\bar{X}}(\bar{x}) = \min\{\alpha, \alpha_{\mathrm{th},\mathsf{N}}\}. \tag{34}$$

Note that the entropy-maximizing distribution in (33) degenerates to a uniform distribution if α is large enough such that the average-power constraint becomes inactive. The corresponding

[2]The optimal variance-maximizing distribution is exactly $(1/2, 1/2)$ on the points $(0, s_{n_{\mathrm{T}}}\mathsf{A})$ when the average-power constraint is inactive.

[3]Note that contrary to the normalizing factor $\mathsf{K}_{\mathsf{A}}(\mu)$ in the probability density (21), $\mathsf{K}(\mu)$ does not depend on A.

threshold value $\alpha_{\text{th,N}}$ equals the ratio of average power to A that is induced by a uniform distribution over the N points, and thus depends on N. It is given as

$$\alpha_{\text{th,N}} = \frac{1}{\mathsf{AN}} \sum_{j=0}^{\mathsf{N}-1} g\left(j \frac{s_{n_\text{T}}\mathsf{A}}{\mathsf{N}-1} \right). \tag{35}$$

Proposition 6 (Discrete Threshold $\alpha_{\text{th,N}}$): The threshold value $\alpha_{\text{th,N}}$ for a discrete, evenly spaced constellation on $\bar{\mathcal{X}}$ is monotonically decreasing in N and is lower-bounded by the threshold α_{th} in (16).

Proof: Omitted for space reasons. ■

We see that when $\alpha < \alpha_{\text{th}}$, one can be sure that the average-power constraint is always active for all N. When $\alpha \geq \alpha_{\text{th},2}$, the average-power constraint is never active for any N (and thus a uniform distribution can always be used). For $\alpha \in [\alpha_{\text{th}}, \alpha_{\text{th},2})$ and once an N is chosen, one needs to check the exact value of $\alpha_{\text{th,N}}$ to ensure the compliance of the average-power constraint. We remark that for $\mathsf{N} = 2$, the equiprobable distribution has mass points on $\bar{X} = 0$ and $\bar{X} = s_{n_\text{T}}\mathsf{A}$, therefore $\alpha_{\text{th},2} = n_\text{T}/2$ holds for any channel.

VI. Constellation Size based on Ozarow–Wyner Mutual Information Lower Bound

Selecting the appropriate constellation size N is crucial for good information rates. It is generally observed that for small A, a small N is optimal, while for large A, a larger N is required. To address this, we derive a lower bound on the mutual information as a function of N and A. This bound will serve as a metric to determine a suitable N for fixed A. The bound is largely based on the Ozarow–Wyner bound [16, Th. b], which we adapt to our setting in the following lemma.

Lemma 7 (Ozarow–Wyner Mutual Information Lower Bound [16]): Let A and N be fixed, and consequently $\bar{\mathcal{X}}$ and Δ are also fixed according to (32). Let $\bar{X}$ be a discrete random variable with arbitrary distribution $q_{\bar{X}}$ over $\bar{\mathcal{X}}$. Then

$$\mathrm{I}(\bar{X};Y) \geq \mathrm{H}(\bar{X}) + \log(2\Delta) - \frac{1}{2}\log\left(\frac{\Delta^2}{3} + \frac{\mathbb{E}\left[\bar{X}^2\right]}{\mathbb{E}\left[\bar{X}^2\right]+1} \right) - \frac{1}{2}\log(2\pi e). \tag{36}$$

Proof: See [16, (15)–(19)]. ■

The lower bound in Lemma 7 is a function of the peak-power constraint A, the number of constellation points N, and the distribution $q_{\bar{X}}$. Let A be fixed, we now propose to maximize (36) over $\mathsf{N} \in \mathcal{N}$ for a predefined set $\mathcal{N}$. Maximizing the lower bound provides a practical heuristic method to maximize the mutual information.

$$\mathsf{N}^* = \operatorname*{argmax}_{\mathsf{N}\in\mathcal{N}} \Bigg\{ \mathrm{H}(\bar{X}) + \log(2\Delta) - \frac{1}{2}\log\left(\frac{\Delta^2}{3} + \frac{\mathbb{E}\left[\bar{X}^2\right]}{\mathbb{E}\left[\bar{X}^2\right]+1} \right) \Bigg\}. \tag{37}$$

The set $\mathcal{N}$ is a design parameter. An obvious choice is to optimize over all natural numbers $\mathcal{N} = \mathbb{N}$. However, empirical observations show that this often results in excessively large N^* values. Thus, it is practical to impose an upper limit on $\mathcal{N}$. We heed the heuristic recommendation in [16] and define the upper limit as $2^{\lceil \mathsf{C}^*+1 \rceil}$, where C^* can be obtained by evaluating known capacity upper bounds [14] for some chosen maximum value of A. We propose two practical choices of $\mathcal{N}$ with varying levels of complexity:

1) *Bounded Powers of Two*: $\mathcal{N} = \left\{2^1, 2^2, \ldots, 2^{\lceil \mathsf{C}^*+1 \rceil}\right\}$.
2) *Dichotomous*: $\mathcal{N} = \left\{2, 2^{\lceil \mathsf{C}^*+1 \rceil}\right\}$.

In particular, the dichotomous system represents the simplest approach. The maximization of (37) amounts to defining an SNR-threshold A_th, which is precisely when the lower bound in (36) with $\mathsf{N} = 2^{\lceil \mathsf{C}^*+1 \rceil}$ achieves higher value than that of $\mathsf{N} = 2$. For $\mathsf{A} < \mathsf{A}_\text{th}$, we use a binary constellation $\bar{\mathcal{X}} = \{0, s_{n_\text{T}}\mathsf{A}\}$ and the corresponding entropy-maximizing (binary) distribution (33); for $\mathsf{A} \geq \mathsf{A}_\text{th}$, we use constellation (32) and distribution (33) with $\mathsf{N} = 2^{\lceil \mathsf{C}^*+1 \rceil}$.

Recall from the previous section that with a selected N, our proposed choice of $q_{\bar{X}}$ is the N-point entropy-maximizing distribution (33), which can be either equiprobable or energy-exponential depending on α, i.e., $\mathbb{E}[\bar{X}^2]$ and $\mathrm{H}(\bar{X})$ depend on α. The corresponding expressions are omitted for space reasons.

VII. Numerical Results

Figures 1 and 2 illustrate the performance of the two proposed modulation schemes, respectively, in comparison with known capacity bounds. In this example, the MISO channel has gains $\mathbf{h} = [3, 2, 1.5]$ and $\alpha = 0.6 < \alpha_\text{th}$. Hence, the maximum-entropy distribution is exponential in $g(\bar{X})$. All plots use $\sigma^2 = 1$, and thus the SNR corresponds to A.

In Figure 1, we set $\mathcal{N} = \left\{2^1, 2^2, \ldots, 2^{\lceil \mathsf{C}^*+1 \rceil}\right\}$, where C^* is evaluated using the tightest upper bound known (SISO Duality-Upper Bound [14, Prop. 10]). In this case, $2^{\lceil \mathsf{C}^*+1 \rceil} = 256$. We show the following:

1) (Solid blue) The Ozarow–Wyner lower bound (36), maximized over $\mathcal{N} = \left\{2^1, 2^2, \ldots, 2^{\lceil \mathsf{C}^*+1 \rceil}\right\}$.
2) (Circled yellow) Information rates achieved by the proposed maximum-entropy distribution (33) with the corresponding N^*.
3) (Purple) Information rates achieved by a gradient descent optimization on a constellation grid with 256 points.

First, we notice that the Ozarow–Wyner lower bound closely approximates the known entropy-lower bound [14, Prop. 5] from 0 dB onward, which is consistent with the findings in [16]; the bounds are close for large SNR. Second, the rates achieved by gradient descent is of particular interest because it represents the maximum achievable rate given a fixed constellation grid with $2^{\lceil \mathsf{C}^*+1 \rceil}$ points. We remark that this baseline is of particular value in the mid-SNR region because the capacity of this channel is unknown (except in the asymptotic low- and high-SNR cases [14]). By construction, the maximum constellation size in our set $\mathcal{N}$ is also $2^{\lceil \mathsf{C}^*+1 \rceil}$. Therefore, any distribution selected with $\mathsf{N} \in \mathcal{N}$ is also a feasible solution in the gradient descent search. We observe that our proposed

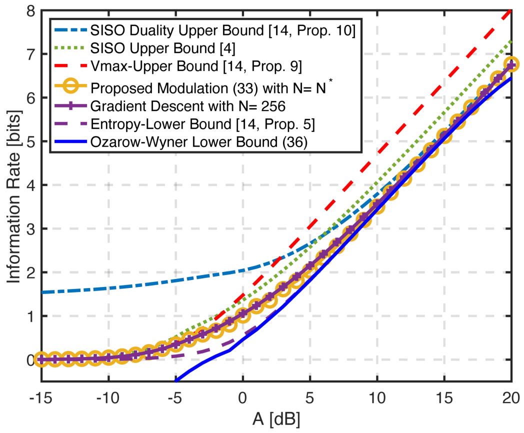

Figure 1. Capacity bounds and mutual information rates of the MISO channel with three LEDs and channel gains $\mathbf{h} = [3, 2, 1.5]$ ($\alpha_{\text{th}} = 1.2692$, $\alpha_{\text{th},2} = 1.5$), $\alpha = 0.6$, and noise variance $\sigma^2 = 1$. The Ozarow–Wyner lower bound (blue) is plotted in combination with known capacity bounds from [4], [14]. The mutual information rates achieved by the proposed modulation (33) with N selected according to (37) with $\mathcal{N} = \{2^1, 2^2, \ldots, 2^{\lceil \mathsf{C}^*+1 \rceil}\}$ is shown (yellow circle) to be close to the maximum achievable rate on the same maximum constellation of 256 points (purple).

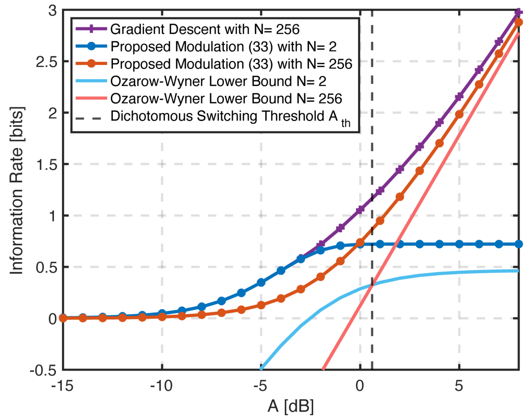

Figure 2. Bounds and mutual information rates with N set to 2 (blue) and 256 (red). The respective Ozarow–Wyner lower bounds with the same N are also shown in the solid lines. This demonstrates the dichotomous scheme with $\mathcal{N} = \{2, 256\}$. The intersection of the Ozarow–Wyner bounds mark A_{th}, where we switch from $\mathsf{N} = 2$ to $\mathsf{N} = 256$. In this case, $\mathsf{A}_{\text{th}} = 1\,\text{dB}$. As a baseline, the maximum achievable rate via gradient search on a 256-point constellation is plotted. The channel is the same channel as in Figure 1.

maximum-entropy modulation achieves rates (yellow) that are very close to the maximum achievable rate (purple). This suggests that i) using a maximum-entropy distribution on an N-point constellation grid is effective, and ii) the Ozarow–Wyner bound is an effective optimization objective to select the constellation size N, despite that the Ozarow–Wyner bound is loose in the low-SNR regime.

Figure 2 demonstrates the performance of the dichotomous modulation scheme, i.e., when $\mathcal{N} = \{2, 2^{\lceil \mathsf{C}^*+1 \rceil}\}$. We plot:

1) (Purple) Maximum information rates achieved by a gradient descent optimization on a constellation grid with 256 points.
2) (Light blue) The Ozarow–Wyner lower bound (36) and (dark blue) information rates of the proposed maximum-entropy distribution (33), both with $\mathsf{N} = 2$.
3) (Light red) The Ozarow–Wyner lower bound (36) and (dark red) information rates of the proposed maximum-entropy distribution (33), both with $\mathsf{N} = 256$.

The dichotomous modulation scheme switches from $\mathsf{N} = 2$ to $\mathsf{N} = 256$ according to A_{th} where the Ozarow–Wyner lower bounds intersect. We observe that A_{th} obtained from the Ozarow–Wyner is remarkably close to the true threshold when $\mathsf{N} = 256$ achieves a higher rate than $\mathsf{N} = 2$. The Ozarow–Wyner bound also shows to be an effective method to maximize the mutual information when $\mathcal{N} = \{2, 2^{\lceil \mathsf{C}^*+1 \rceil}\}$. Empirical observations show that the dichotomous scheme is not too far away from the maximum achieved by the gradient descent. In this example, the dichotomous scheme is at most 0.34 bits lower than the gradient descent search.

VIII. Conclusions

This paper presents theoretical and practical contributions to the study of the multiple-input and single-output (MISO) freespace optical intensity channel with signal-independent additive Gaussian noise and with both a peak-power and first-moment constraint on the input vector $\mathbf{X}$.

On the theoretical side, we extend the asymptotic high-SNR capacity results from [14], which maximized the differential entropy $\mathsf{h}(\mathbf{h}^\mathsf{T}\mathbf{X})$ over an exponential parameter μ. We show that the maximum-entropy distribution of $\mathbf{h}^\mathsf{T}\mathbf{X}$ must be exponential in the energy of $\mathbf{X}$, with the exponential parameter μ uniquely defined by the first-moment constraint. We present closed-form expressions for the asymptotic high-SNR capacity-achieving distribution, as well as the corresponding capacity and differential entropy expressions.

On the practical side, we address the design of modulation schemes for this channel. Specifically, we propose a simple scheme that fixes the input distribution to be entropy-maximizing across all SNR regimes. The mutual information is then maximized by adapting the constellation size: smaller constellations for low SNR and larger constellations for high SNR. The selection of the constellation size is guided by the Ozarow–Wyner lower bound, which provides a practical yet low-complexity heuristic for maximizing the mutual information. Numerical results showcase the effectiveness of the proposed scheme and the value of the Ozarow–Wyner bound in determining the constellation size.

References

[1] J. M. Kahn and J. R. Barry, “Wireless infrared communications,” *Proc. IEEE*, vol. 85, no. 2, pp. 265–298, Feb. 1997.

[2] M. A. Wigger, “Bounds on the capacity of free-space optical intensity channels,” Master’s thesis, Signal and Inf. Proc. Lab., ETH Zürich, Switzerland, Mar. 2003.

[3] S. Hranilovic and F. R. Kschischang, “Capacity bounds for power- and band-limited optical intensity channels corrupted by Gaussian noise,” *IEEE Trans. Inf. Theory*, vol. 50, no. 5, pp. 784–795, May 2004.

[4] A. L. McKellips, “Simple tight bounds on capacity for the peak-limited discrete-time channel,” in *Proc. IEEE Int. Symp. Inf. Theory*, Chicago, IL, USA, Jun. 27 – Jul. 2, 2004, p. 348.

[5] S. M. Moser, *Duality-Based Bounds on Channel Capacity*, ser. ETH Series in Information Theory and its Applications. Konstanz, Germany: Hartung-Gorre Verlag, Jan. 2005, vol. 1, edited by Amos Lapidoth. ISBN 3–89649–956–4. [Online]. Available: https://moser-isi.ethz.ch/publications.html

[6] S. Hranilovic, *Wireless Optical Communication Systems*. New York, NY, USA: Springer Verlag, 2005.

[7] A. A. Farid and S. Hranilovic, “Channel capacity and non-uniform signalling for free-space optical intensity channels,” *IEEE J. Select. Areas Commun.*, vol. 27, no. 9, pp. 1553–1563, Dec. 2009.

[8] A. Lapidoth, S. M. Moser, and M. A. Wigger, “On the capacity of free-space optical intensity channels,” *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4449–4461, Oct. 2009.

[9] A. A. Farid and S. Hranilovic, “Capacity bounds for wireless optical intensity channels with Gaussian noise,” *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6066–6077, Dec. 2010.

[10] E. Bayaki and R. Schober, “On space–time coding for free–space optical systems,” *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 58–62, Jan. 2010.

[11] A. Thangaraj, G. Kramer, and G. Böcherer, “Capacity bounds for discrete-time, amplitude-constrained, additive white Gaussian noise channels,” *IEEE Trans. Inf. Theory*, vol. 63, no. 7, pp. 4172–4182, Jul. 2017.

[12] A. Chaaban, Z. Rezki, and M.-S. Alouini, “Capacity bounds and high-SNR capacity of MIMO intensity-modulation optical channels,” *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 3003–3017, May 2018.

[13] ——, “Low-SNR asymptotic capacity of MIMO optical intensity channels with peak and average constraints,” *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4694–4705, Oct. 2018.

[14] S. M. Moser, L. Wang, and M. Wigger, “Capacity results on multiple-input single-output wireless optical channels,” *IEEE Trans. Inf. Theory*, vol. 64, no. 11, pp. 6954–6966, Nov. 2018.

[15] L. Li, S. M. Moser, L. Wang, and M. Wigger, “On the capacity of MIMO optical wireless channels,” *IEEE Trans. Inf. Theory*, vol. 66, no. 9, pp. 5660–5682, Sept. 2020.

[16] L. H. Ozarow and A. D. Wyner, “On the capacity of the Gaussian channel with a finite number of input levels,” *IEEE Trans. Inf. Theory*, vol. 36, no. 6, pp. 1426–1428, Nov. 1990.

[17] C. E. Shannon, “A mathematical theory of communication,” *Bell Syst. Tech. J.*, vol. 27, pp. 379–423 and 623–656, Jul. and Oct. 1948.