Capacity Results on Multiple-Input Single-Output Wireless Optical Channels

Stefan M Moser, Ligong Wang, Michèle Wigger

Capacity Results on Multiple-Input Single-Output Wireless Optical Channels

Stefan M. Moser, *Senior Member, IEEE*, Ligong Wang, *Member, IEEE*, and Michèle Wigger, *Senior Member, IEEE*

***Abstract*— This paper derives upper and lower bounds on the capacity of the multiple-input single-output free-space optical intensity channel with signal-independent additive Gaussian noise subject to both an average-intensity and a peak-intensity constraint. In the limit where the signal-to-noise ratio (SNR) tends to infinity, the asymptotic capacity is specified, while in the limit where the SNR tends to zero, the exact slope of the capacity is given.**

***Index Terms*— Average- and peak-power constraint, channel capacity, direct detection, Gaussian noise, infrared communication, multiple-input single-output (MISO) channel, optical communication.**

I. Introduction

Optical wireless communication is a form of communication in which visible, infrared, or ultraviolet light is transmitted in free space (air or vacuum) to carry a message to its destination. Recent works suggest that it is a promising solution to replacing some of the existing radio-frequency (RF) wireless communication systems in order to prevent future rate bottlenecks [1]–[3]. Particularly attractive are simple *intensity-modulation–direct-detection (IM-DD)* systems. In such a system, the transmitter modulates the intensity of optical signals coming from light emitting diodes (LEDs) or laser diodes (LDs), and the receiver measures incoming optical intensities by means of photodetectors. The electrical output signals of the photodetectors are essentially proportional to the incoming optical intensities, but are corrupted by thermal noise of the photodetectors, relative-intensity noise of random intensity fluctuations inherent to low-cost LEDs and LDs, and shot noise caused by ambient light. In a first approximation, noise coming from these sources is usually modeled as being additive Gaussian and independent of the transmitted light signal; see [1], [2].

The free-space optical intensity channel has been extensively studied in the literature. In the *single-input single-output (SISO)* scenario, where the transmitter employs a single transmit LED or LD, and the receiver a single photodetector, the works [4], [5] established upper and lower bounds on the capacity of this channel that are asymptotically tight in both high-signal-to-noise-ratio (SNR) and low-SNR limits. Improved bounds at finite SNR have subsequently been presented in [6]–[9]. For the *multiple-input and multiple-output (MIMO)* optical intensity channel, where the transmitter is equipped with multiple LEDs or LDs, and the receiver with multiple photodetectors, the recent work [10] determined the asymptotic capacity in the high-SNR limit when *the channel matrix is of full column rank*. A special case of this MIMO result was independently solved in [11].

Previous to [10], [11], various code constructions for this setup have been proposed in [12]–[15]. When there is no crosstalk so the MIMO channel can be modeled through a diagonal channel matrix, bounds on capacity were presented in [9] and [16].

The current work is concerned with the *multiple-input and single-output (MISO)* channel. (Clearly, the channel matrix of the MISO channel cannot have full column rank.) Our main results, some of which presented in part in [10] and [17], include

- several upper and lower bounds on the capacity;
- high-SNR asymptotic capacity for all parameter ranges; and
- low-SNR capacity slope in terms of a maximum variance on the input.

The high-SNR asymptotic capacity is proven based on two capacity lower bounds (Propositions 5 and 6) derived using the Entropy Power Inequality (EPI), and an upper bound (Proposition 10) derived using the duality technique [18]. These proof techniques are similar to those in [5], [9], and [10], but also involve some nontrivial optimization. In particular, the optimal input distribution at high SNR involves LED cooperation (compared to independent signaling in the MIMO full-column-rank case [10]), and, with certain probabilities, assigns to each LED a truncated exponential distribution, whose parameters must be carefully chosen.

The low-SNR capacity slope is proven using a simple upper bound (Proposition 9) and a classic asymptotic lower bound by

Manuscript received September 26, 2017; revised January 23, 2018; accepted March 26, 2018. Date of publication April 12, 2018; date of current version October 18, 2018. The work of M. Wigger was supported by the ERC under Grant 715111. This paper was presented in part at the 2017 IEEE International Symposium on Information Theory, Aachen, Germany, and at the 2017 IEEE Information Theory Workshop, Kaohsiung, Taiwan.

S. M. Moser is with the Signal and Information Processing Laboratory, ETH Zürich, 8092 Zürich, Switzerland, and also with the Institute of Communications Engineering, National Chiao Tung University, Hsinchu 30010, Taiwan (e-mail: stefan.moser@ieee.org).

L. Wang is with ETIS—Université Paris Seine, Université de Cergy-Pontoise, ENSEA, CNRS, 95000 Cergy-Pontoise, France (e-mail: ligong.wang@ensea.fr).

M. Wigger is with LTCI, Telecom ParisTech, Université Paris-Saclay, 75013 Paris, France (e-mail: michele.wigger@telecom-paristech.fr).

Communicated by N. Devroye, Associate Editor for Communications.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2018.2825994

Prelov and Van der Meulen [19]. The expression of this slope involves a maximization of variance, which can be computed numerically (Lemma 8).

All the above-mentioned bounds (except the asymptotic one from [19]) hold at all SNR, and not only in the high- or low-SNR limits. They are compared numerically, together with another indirect upper bound (Proposition 7), which utilizes upper bounds on the capacity of the SISO channel from [5], [6], and [9].

The remainder of this paper is structured as follows. After a few remarks on notation, Section II lays out the specific details of the investigated channel model. Section III reviews the known capacity results from [10] and proves a fundamental proposition giving the optimal structure of an input to the MISO channel with both active average- and peak-power constraints. Section IV then presents all new upper and lower bounds on the capacity and gives the correct high-SNR and low-SNR asymptotics. The detailed proofs for the lower bounds can be found in Section V and the proof for one of the upper bounds in Section VI. Section VII shows the analysis of the high-SNR capacity. The paper is concluded in Section VIII.

We meticulously distinguish between random and deterministic quantities. A random variable is denoted by a capital Roman letter, e.g., Z, while its realization is denoted by the corresponding small Roman letter, e.g., z. Vectors are boldfaced, e.g., $\mathbf{X}$ denotes a random vector and $\mathbf{x}$ its realization. Constants are typeset either in small Romans, in Greek letters or in a special font, e.g., $\mathcal{E}$ or $\mathcal{A}$. Entropy is typeset as $\mathsf{H}(\cdot)$, differential entropy as $\mathsf{h}(\cdot)$, and $\mathcal{I}(\cdot;\cdot)$ denotes the mutual information [20]. The relative entropy (or Kullback-Leibler divergence) [21, Sec. 2.3] is denoted by $\mathcal{D}(\mathbf{p}\|\mathbf{q})$ for some probability vectors $\mathbf{p}$ and $\mathbf{q}$. The logarithm function $\log(\cdot)$ denotes the natural logarithm.

II. Channel Model

Consider a communication link where the transmitter is equipped with n_T LEDs (or LDs), $n_\mathrm{T} \geq 2$, and the receiver with a single photodetector. The photodetector receives a superposition of the signals sent by the LEDs, and we assume that the crosstalk between LEDs is constant. Hence, the channel output is given by

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

where the n_T-vector $\mathbf{x} = (x_1, \ldots, x_{n_\mathrm{T}})^\mathsf{T}$ denotes the channel input, whose entries are proportional to the optical intensities of the corresponding LEDs, and are therefore nonnegative:

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

where the length-n_T row vector $\mathbf{h}^\mathsf{T} = (h_1, \ldots, h_{n_\mathrm{T}})$ is the constant channel state vector with nonnegative entries, which, without loss of generality, we assume to be ordered:

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

and where $Z \sim \mathcal{N}\big(0, \sigma^2\big)$ is additive Gaussian noise. Note that, in contrast to the input $\mathbf{x}$, the output Y can be negative.

Inputs are subject to a peak-power (peak-intensity) and an average-power (average-intensity) constraint:

$$\Pr\big[X_k > \mathcal{A}\big] = 0, \quad \forall\, k \in \{1, \ldots, n_\mathrm{T}\} \tag{4}$$

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

for some fixed parameters $\mathcal{A}, \mathcal{E} > 0$. Note that the average-power constraint is on the expectation of the channel input and not on its square. Also note that $\mathcal{A}$ describes the maximum power of each single LED, while $\mathcal{E}$ describes the allowed average total power of all LEDs together.

We denote the ratio between the allowed average power and the allowed peak power by α:

$$\alpha \triangleq \frac{\mathcal{E}}{\mathcal{A}} \tag{6}$$

where $0 < \alpha \leq n_\mathrm{T}$. For $\alpha = n_\mathrm{T}$ the average-power constraint is inactive in the sense that it is automatically satisfied whenever the peak-power constraint is satisfied. Thus, $\alpha = n_\mathrm{T}$ corresponds to the case with only a peak-power constraint.

We denote the capacity of the channel (1) with allowed peak power $\mathcal{A}$ and allowed average power $\mathcal{E}$ by $\mathsf{C}_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathcal{A}, \mathcal{E})$. The capacity is given by [20]

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

where the supremum is over all laws $Q_\mathbf{X}$ on $\mathbf{X}$ satisfying (2), (4), and (5). When only an average-power constraint is imposed, capacity is denoted by $\mathsf{C}_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathcal{E})$. It is given as in (7) except that the supremum is taken over all laws $Q_\mathbf{X}$ on $\mathbf{X}$ satisfying (2) and (5).

III. Equivalent Capacity Formulas

Denote

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

$$s_k \triangleq \sum_{k'=1}^{k} h_{k'}, \quad k \in \{1, \ldots, n_\mathrm{T}\} \tag{8b}$$

and

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

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

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

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

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

Now notice that because $\mathbf{X} \multimap\!\!- \bar{X} \multimap\!\!- Y$ form a Markov chain and because $\bar{X}$ is a function of $\mathbf{X}$, we have

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

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 $\bar{X}$. So,

$$\mathcal{C}_{\mathbf{h}^{\mathsf{T}},\sigma^2}(\mathcal{A},\mathcal{E}) = \max_{Q_{\bar{X}}} \mathcal{I}(\bar{X};Y) \tag{12}$$

where $Q_{\bar{X}}$ is restricted to the set of admissible distributions. Characterizing this set of admissible distributions is relatively straightforward when there is only an average- or only a peak-power constraint, but is more involved in general. This is the subject of the following three propositions.

If $\mathbf{X}$ is only subject to an average-power constraint $\mathcal{E}$ (and no peak-power constraint), then $\bar{X}$ is only subject to an average-power constraint $h_1\mathcal{E}$.

Proposition 1 (Only Average-Power Constraint): Without a peak-power constraint,

$$\mathcal{C}_{\mathbf{h}^{\mathsf{T}},\sigma^2}(\mathcal{E}) = \max_{\substack{Q_{\bar{X}}:\, \bar{X}\in[0,\infty),\\ \mathsf{E}[\bar{X}]\le h_1\mathcal{E}}} \mathcal{I}(\bar{X};Y) = \mathcal{C}_{1,\sigma^2}(h_1\mathcal{E}) \tag{13}$$

where $\mathcal{C}_{1,\sigma^2}(h_1\mathcal{E})$ denotes the capacity of a SISO channel with unit channel gain under average-power constraint $h_1\mathcal{E}$.

Proof: When $\mathbf{X}$ satisfies (5), we have

$$\mathsf{E}\left[\bar{X}\right] = \sum_{k=1}^{n_{\mathrm{T}}} h_k\,\mathsf{E}\left[X_k\right] \le h_1\mathcal{E} \tag{14}$$

so $\mathcal{C}_{\mathbf{h}^{\mathsf{T}},\sigma^2}(\mathcal{E}) \le \mathcal{C}_{1,\sigma^2}(h_1\mathcal{E})$. For the reverse direction, to achieve any target distribution on $\bar{X}$ satisfying $\mathsf{E}\left[\bar{X}\right] \le h_1\mathcal{E}$, the transmitter can let the LED corresponding to h_1 send $\bar{X}/h_1$ and all the other LEDs send zero. ■

When $\mathbf{X}$ is only subject to a peak-power constraint $\mathcal{A}$ (and no average-power constraint), then $\bar{X}$ is only subject to a peak-power constraint $s_{n_{\mathrm{T}}}\mathcal{A}$. Moreover, for $\alpha \ge \frac{n_{\mathrm{T}}}{2}$, the average-power constraint is inactive because the capacity-achieving input distribution on $\bar{X}$ in the absence of a peak-power constraint can be shown to be symmetric around $\frac{s_{n_{\mathrm{T}}}\mathcal{A}}{2}$ [10, Prop. 1].

Proposition 2 (Only Peak-Power Constraint is Active): When $\alpha \ge \frac{n_{\mathrm{T}}}{2}$,

$$\begin{aligned}\mathcal{C}_{\mathbf{h}^{\mathsf{T}},\sigma^2}(\mathcal{A},\alpha\mathcal{A}) &= \max_{Q_{\bar{X}}:\, \bar{X}\in[0,s_{n_{\mathrm{T}}}\mathcal{A}]} \mathcal{I}(\bar{X};Y) &(15)\\ &= \mathcal{C}_{1,\sigma^2}\left(s_{n_{\mathrm{T}}}\mathcal{A}, \frac{s_{n_{\mathrm{T}}}\mathcal{A}}{2}\right) &(16)\end{aligned}$$

where $\mathcal{C}_{1,\sigma^2}\left(s_{n_{\mathrm{T}}}\mathcal{A}, \frac{s_{n_{\mathrm{T}}}\mathcal{A}}{2}\right)$ denotes the capacity of a SISO channel with unit channel gain under peak-power constraint $s_{n_{\mathrm{T}}}\mathcal{A}$ and average-power constraint $\frac{s_{n_{\mathrm{T}}}\mathcal{A}}{2}$.

Proof: When $\mathbf{X}$ satisfies the peak-power constraint (4), $\bar{X}$ must satisfy $\bar{X} \le s_{n_{\mathrm{T}}}\mathcal{A}$ with probability one. Hence $\mathcal{C}_{\mathbf{h}^{\mathsf{T}},\sigma^2}(\mathcal{A},\alpha\mathcal{A})$ cannot exceed the capacity of the SISO channel with allowed peak power $s_{n_{\mathrm{T}}}\mathcal{A}$. By [5, Prop. 9], for a SISO channel with allowed peak power $s_{n_{\mathrm{T}}}\mathcal{A}$, adding an average-power constraint of $\frac{s_{n_{\mathrm{T}}}\mathcal{A}}{2}$ does not affect its capacity. We hence know that the left-hand side (LHS) of (16) is upper-bounded by its right-hand side (RHS).

For the reverse direction, consider any target distribution on $\bar{X}$ satisfying peak-power constraint $s_{n_{\mathrm{T}}}\mathcal{A}$ and average-power constraint $\frac{1}{2}s_{n_{\mathrm{T}}}\mathcal{A}$. We need only to show that such an $\bar{X}$ can be generated by some distribution for $\mathbf{X}$ satisfying peak- and average-power constraints $\mathcal{A}$ and $\alpha\mathcal{A}$, respectively. To this end, we let the transmitter send the same signal on all LEDs:

$$X_k = \frac{\bar{X}}{s_{n_{\mathrm{T}}}}, \quad k \in \{1,\ldots,n_{\mathrm{T}}\}. \tag{17}$$

One can easily check that both constraints are indeed satisfied by this choice. ■

As already mentioned, describing the set of admissible distributions is more complicated when $\alpha < \frac{n_{\mathrm{T}}}{2}$. Recall the definition of U in (10) and let $p_k \triangleq \Pr[U=k]$ for $k = 1,\ldots,n_{\mathrm{T}}$.

Proposition 3 (Active Average- and Peak-Power Constraints): When $\alpha < \frac{n_{\mathrm{T}}}{2}$,

$$\mathcal{C}_{\mathbf{h}^{\mathsf{T}},\sigma^2}(\mathcal{A},\alpha\mathcal{A}) = \max_{Q_{\bar{X}}} \mathcal{I}(\bar{X};Y) \tag{18}$$

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

$$\Pr\left[\bar{X} > s_{n_{\mathrm{T}}}\mathcal{A}\right] = 0 \tag{19a}$$

and

$$\sum_{k=1}^{n_{\mathrm{T}}} p_k \left(\frac{\mathsf{E}[\bar{X}\,|\,U=k] - \mathcal{A}s_{k-1}}{h_k} + (k-1)\mathcal{A}\right) \le \alpha\mathcal{A}. \tag{19b}$$

The proof of Proposition 3 is based on the following lemma, which will be of further interest in this paper.

Lemma 4: Without loss in optimality, the maximization in (7) can be restricted to distributions $Q_{\mathbf{X}}$ of the input vector $\mathbf{X}$ satisfying for all $k \in \{1,\ldots,n_{\mathrm{T}}\}$, with probability one,

$$\Big(X_k > 0\Big) \implies \Big(X_1 = \cdots = X_{k-1} = \mathcal{A}\Big). \tag{20}$$

Proof: Fix $X_1,\ldots,X_{n_{\mathrm{T}}}$ satisfying the peak- and average-power constraints (4) and (5) and let $\bar{X}$ and U be defined as in (9) and (10). Define also a set of new inputs $X_1^*,\ldots,X_{n_{\mathrm{T}}}^*$ that with probability $p_k = \Pr[U=k]$ take on values

$$\begin{aligned}X_1^* = \cdots = X_{k-1}^* &= \mathcal{A} &(21a)\\ X_k^* &= \left.\frac{\bar{X} - \mathcal{A}s_{k-1}}{h_k}\right|_{U=k} &(21b)\\ X_{k+1}^* = \cdots = X_{n_{\mathrm{T}}}^* &= 0. &(21c)\end{aligned}$$

Notice that

$$\begin{aligned}\bar{X}^* &\triangleq \sum_{k=1}^{n_{\mathrm{T}}} h_k X_k^* &(22)\\ &= \sum_{k'=1}^{n_{\mathrm{T}}} p_{k'} \sum_{k=1}^{n_{\mathrm{T}}} h_k X_k^*\Big|_{U=k'} &(23)\\ &= \sum_{k'=1}^{n_{\mathrm{T}}} p_{k'} \left(\sum_{k=1}^{k'-1} h_k\mathcal{A} + h_{k'}\frac{\bar{X}\big|_{U=k'} - \mathcal{A}s_{k'-1}}{h_{k'}}\right) &(24)\\ &= \sum_{k'=1}^{n_{\mathrm{T}}} p_{k'}\bar{X}\big|_{U=k'} &(25)\\ &= \bar{X} &(26)\end{aligned}$$

and hence by (11)

$$\mathcal{I}(\mathbf{X}^*; Y) = \mathcal{I}(\bar{X}^*; Y) = \mathcal{I}(\bar{X}; Y) = \mathcal{I}(\mathbf{X}; Y). \tag{27}$$

Moreover, the new inputs $X_1^*, \ldots, X_{n_\mathrm{T}}^*$ are admissible because they trivially satisfy the peak-power constraint (4) and their average power does not exceed the average power of the original inputs $X_1, \ldots, X_{n_\mathrm{T}}$:

$$\sum_{k=1}^{n_\mathrm{T}} \mathsf{E}\left[X_k^*\right] \le \sum_{k=1}^{n_\mathrm{T}} \mathsf{E}\left[X_k\right]. \tag{28}$$

In fact, it is not hard to see that among all input assignments generating a sum-input $\bar{X} \in (s_{k-1}\mathcal{A}, s_k\mathcal{A}]$, the choice in (21) consumes least input energy $\sum_{k=1}^{n_\mathrm{T}} \mathsf{E}\left[X_k^*\right]$. ■

Proof of Proposition 3: By Lemma 4, we can restrict attention to distributions on $\mathbf{X}$ satisfying the implication in (20). Notice that for these distributions there is a one-to-one correspondence between $\mathbf{X}$ and $\bar{X}$. The average input power $\mathsf{E}\left[\sum_{k=1}^{n_\mathrm{T}} X_k\right]$ can thus entirely be expressed in terms of $\bar{X}$. In fact, since the input power consumed for $\bar{X} \in (s_{k-1}\mathcal{A}, s_k\mathcal{A}]$ is

$$\sum_{k'=1}^{n_\mathrm{T}} X'_k = \frac{\bar{X} - \mathcal{A}s_{k-1}}{h_k} + (k-1)\mathcal{A} \qquad \text{(conditional on } \bar{X} \in (s_{k-1}\mathcal{A}, s_k\mathcal{A}]) \tag{29}$$

the average input power is

$$\begin{aligned} \mathsf{E}\left[\sum_{k'=1}^{n_\mathrm{T}} X_{k'}\right] &= \sum_{k=1}^{n_\mathrm{T}} p_k \, \mathsf{E}\left[\sum_{k'=1}^{n_\mathrm{T}} X_{k'} \,\middle|\, U = k\right] &(30)\\ &= \sum_{k=1}^{n_\mathrm{T}} p_k \left(\frac{\mathsf{E}[\bar{X} \mid U = k] - \mathcal{A}s_{k-1}}{h_k} + (k-1)\mathcal{A}\right). &(31) \end{aligned}$$

The proposition now follows from (12) and (31) and because $\bar{X} \in [0, s_{n_\mathrm{T}}\mathcal{A}]$. ■

Bounds on the capacities $\mathcal{C}_{1,\sigma^2}(h_1\mathcal{E})$ and $\mathcal{C}_{1,\sigma^2}\bigl(s_{n_\mathrm{T}}\mathcal{A}, \frac{s_{n_\mathrm{T}}\mathcal{A}}{2}\bigr)$ were presented in [5], [6], [9], and [22]. Moreover, [5] also characterizes exactly the high-SNR asymptotic behavior of these two capacities and the low-SNR asymptotic behavior of $\mathcal{C}_{1,\sigma^2}\bigl(s_{n_\mathrm{T}}\mathcal{A}, \frac{s_{n_\mathrm{T}}\mathcal{A}}{2}\bigr)$. In the rest of this paper we focus on the case with active average- and peak-power constraints, so $\alpha < \frac{n_\mathrm{T}}{2}$, and we bound the RHS of (18) under constraints (19).

IV. Bounds on Capacity When $\alpha < \frac{n_\mathrm{T}}{2}$

Throughout this section we assume $\alpha < \frac{n_\mathrm{T}}{2}$. Some of our bounds depend on whether α is larger or smaller than the threshold

$$\alpha_\mathrm{th} \triangleq \frac{1}{2} + \frac{1}{s_{n_\mathrm{T}}} \sum_{k=1}^{n_\mathrm{T}} h_k (k-1). \tag{32}$$

This threshold represents the smallest α such that $\bar{X}$ can be made uniformly distributed over $[0, s_{n_\mathrm{T}}\mathcal{A}]$; we discuss this toward the end of Section IV-A.

A. Lower Bounds

We first present the lower bounds.

Proposition 5 (Lower Bound for $\alpha < \alpha_\mathrm{th}$): If $\alpha < \alpha_\mathrm{th}$, the capacity is lower-bounded as

$$\mathcal{C}_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathcal{A}, \alpha\mathcal{A}) \ge \frac{1}{2}\log\left(1 + \frac{\mathcal{A}^2 s_{n_\mathrm{T}}^2}{2\pi e\sigma^2} e^{2\nu}\right) \tag{33}$$

with

$$\nu \triangleq \sup_{\lambda \in \left(\max\{0, \frac{1}{2}+\alpha-\alpha_\mathrm{th}\}, \min\{\frac{1}{2}, \alpha\}\right)} \left\{1 - \log\frac{\mu(\lambda)}{1 - e^{-\mu(\lambda)}} - \frac{\mu(\lambda)\, e^{-\mu(\lambda)}}{1 - e^{-\mu(\lambda)}} - \mathcal{D}\left(\mathbf{p} \,\middle\|\, \frac{\mathbf{h}}{s_{n_\mathrm{T}}}\right)\right\} \tag{34}$$

where $\mu(\lambda)$ is the unique positive solution to the following equation in μ:

$$\frac{1}{\mu} - \frac{e^{-\mu}}{1 - e^{-\mu}} = \lambda; \tag{35}$$

and where

$$p_k = \frac{h_k a^k}{\sum_{j=1}^{n_\mathrm{T}} h_j a^j}, \qquad k \in \{1, \ldots, n_\mathrm{T}\} \tag{36a}$$

with a being the unique positive solution to

$$\frac{\sum_{k=1}^{n_\mathrm{T}} h_k k a^k}{\sum_{j=1}^{n_\mathrm{T}} h_j a^j} = \alpha - \lambda + 1. \tag{36b}$$

Proof: See Section V-A. ■

We remark that ν in (34) is the maximum differential entropy of $\frac{\bar{X}}{s_{n_\mathrm{T}}\mathcal{A}}$ under the power constraints; see discussion after the next proposition. We shall plot ν after stating our asymptotic high-SNR result; see Figure 2.

Proposition 6 (Lower Bound for $\alpha \ge \alpha_\mathrm{th}$): If $\alpha \ge \alpha_\mathrm{th}$, the capacity is lower-bounded as

$$\mathcal{C}_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathcal{A}, \alpha\mathcal{A}) \ge \frac{1}{2}\log\left(1 + \frac{s_{n_\mathrm{T}}^2 \mathcal{A}^2}{2\pi e\sigma^2}\right). \tag{37}$$

Proof: See Section V-B. ■

The lower bounds are obtained by choosing the inputs so as to maximize the differential entropy $\mathsf{h}(\bar{X})$ under the constraints (19). To find these entropy-maximizing inputs in the general situation under both a peak- and an (active) average-power constraint, we need two insights. First, relying on Lemma 4, we note that in order to reach a certain range of amplitude levels $\bar{X} \in (s_{k-1}\mathcal{A}, s_k\mathcal{A}]$, it is most energy-efficient to set all strong LEDs (i.e., those with large channel gains) to the maximum level, $X_j = \mathcal{A}$, $j = 1, \ldots, k-1$; to switch the weaker LEDs off, $X_j = 0$, $j = k+1, \ldots, n_\mathrm{T}$; and to choose $X_k = \bar{X} - s_{k-1}\mathcal{A}$. Second, conditional on a given range $(s_{k-1}\mathcal{A}, s_k\mathcal{A}]$, the entropy-maximizing input X_k has a truncated exponential distribution. It then only remains to optimize over the probability masses assigned to each of the different amplitude ranges and over the parameters of the truncated exponentials.

Thus, for $\alpha < \alpha_{\mathrm{th}}$, we let with probability p_k, $k \in \{1, \ldots, n_{\mathrm{T}}\}$,

$$X_j = \mathcal{A}, \qquad j \in \{1, \ldots, k-1\} \tag{38a}$$

$$X_k \sim \text{truncated exponential of parameter } \mu(\lambda) \tag{38b}$$

$$X_j = 0, \quad j \in \{k+1, \ldots, n_{\mathrm{T}}\} \tag{38c}$$

where $\mu(\lambda)$ is the unique positive solution to (35) and $\mathbf{p}$ is given in (36). This choice results in a concatenation of n_{T} truncated exponentials for $\bar{X}$, with $\mathsf{h}(\bar{X}) = \log(s_{n_{\mathrm{T}}}\mathcal{A}) + \nu$, where ν is given in (34).

For $\alpha > \alpha_{\mathrm{th}}$, the average-power constraint on $\bar{X}$ becomes inactive, and we can replace the truncated exponential distribution in (38b) by a uniform distribution and choose the probability vector $\mathbf{p}$ as $p_k = h_k/s_{n_{\mathrm{T}}}$, for $k \in \{1, \ldots, n_{\mathrm{T}}\}$. This choice yields a uniform distribution on $[0, s_{n_{\mathrm{T}}}\mathcal{A}]$ for $\bar{X}$, with $\mathsf{h}(\bar{X}) = \log(s_{n_{\mathrm{T}}}\mathcal{A})$.

That the described choices maximize $\mathsf{h}(\bar{X})$ under constraints (19) can be proved, e.g., using [21, Th. 12.1.1].

B. Upper Bounds

We next present our upper bounds. First we present a simple upper bound by the SISO capacity.

Proposition 7 (Upper Bound by SISO Capacity): The capacity is upper-bounded as

$$\mathcal{C}_{\mathbf{h}^{\mathsf{T}},\sigma^2}(\mathcal{A}, \alpha\mathcal{A}) \leq \mathcal{C}_{1,\sigma^2}\left(s_{n_{\mathrm{T}}}\mathcal{A}, \frac{s_{n_{\mathrm{T}}}\mathcal{A}}{2}\right). \tag{39}$$

Proof: The bound follows by the equivalence in (16), and because the capacity is nondecreasing in the parameter α (as the transmitter can always choose not to use all of its available power). ∎

Note that the SISO capacity $\mathcal{C}_{1,\sigma^2}\left(s_{n_{\mathrm{T}}}\mathcal{A}, \frac{s_{n_{\mathrm{T}}}\mathcal{A}}{2}\right)$ is itself unknown to date. Upper bounds on it were given in [5], [6], [9], and [22]. In fact, under a peak-power constraint $s_{n_{\mathrm{T}}}\mathcal{A}$, the average-power constraint $\frac{s_{n_{\mathrm{T}}}\mathcal{A}}{2}$ is not active.

Our next upper bound, like Proposition 7, is valid for all values of $\alpha < \frac{n_{\mathrm{T}}}{2}$. It depends on the maximum variance of $\bar{X}$ under constraints (19):

$$\mathcal{V}_{\max}(\mathcal{A}, \alpha\mathcal{A}) \triangleq \max_{Q_{\bar{X}}} \mathsf{E}\left[\left(\bar{X} - \mathsf{E}[\bar{X}]\right)^2\right] \tag{40}$$

where the maximization is over all distributions on $\bar{X} \geq 0$ that satisfy constraints (19). This maximum variance is calculated numerically using the following lemma.

Lemma 8: Consider the maximum variance $\mathcal{V}_{\max}(\mathcal{A}, \alpha\mathcal{A})$ as defined in (40).

1) The maximum variance can be achieved by restricting $Q_{\bar{X}}$ to the support set

$$\{0, s_1\mathcal{A}, s_2\mathcal{A}, \ldots, s_{n_{\mathrm{T}}}\mathcal{A}\}. \tag{41}$$

2) The maximum variance satisfies

$$\mathcal{V}_{\max}(\mathcal{A}, \alpha\mathcal{A}) = \mathcal{A}^2\gamma \tag{42}$$

where

$$\gamma \triangleq \max_{\substack{q_1,\ldots,\,q_{n_{\mathrm{T}}}\geq 0:\\ \sum_{k=1}^{n_{\mathrm{T}}} q_k \leq 1 \\ \sum_{k=1}^{n_{\mathrm{T}}} k\cdot q_k \leq \alpha}} \left(\sum_{k=1}^{n_{\mathrm{T}}} s_k^2 q_k - \left(\sum_{k=1}^{n_{\mathrm{T}}} s_k q_k\right)^2\right). \tag{43}$$

Proof: See Appendix. ∎

TABLE I

MAXIMUM VARIANCE FOR DIFFERENT CHANNEL COEFFICIENTS

channel gains	α	$\mathcal{V}_{\max}$	$Q_{\bar{X}}$ achieving $\mathcal{V}_{\max}$
$\mathbf{h} = (3, 2.2, 0.1)$	0.9	$6.6924\mathcal{A}^2$	$Q_{\bar{X}}(0) = 0.55,\ Q_{\bar{X}}(s_2\mathcal{A}) = 0.45$
$\mathbf{h} = (3, 2.2, 1.1)$	0.7	$7.1001\mathcal{A}^2$	$Q_{\bar{X}}(0) = 0.7667,\ Q_{\bar{X}}(s_3\mathcal{A}) = 0.2333$
$\mathbf{h} = (3, 1.5, 0.3)$	0.95	$5.1158\mathcal{A}^2$	$Q_{\bar{X}}(0) = 0.5907,\ Q_{\bar{X}}(s_2\mathcal{A}) = 0.2780,$ $Q_{\bar{X}}(s_3\mathcal{A}) = 0.1313$

For a three-LED MISO channel, some examples of variances $\mathcal{V}_{\max}$ are shown in Table I. The last column of the table indicates the probability mass function that achieves $\mathcal{V}_{\max}$.

Proposition 9 (Upper Bound in Terms of $\mathcal{V}_{\max}$): The capacity is upper-bounded as

$$\mathcal{C}_{\mathbf{h}^{\mathsf{T}},\sigma^2}(\mathcal{A}, \alpha\mathcal{A}) \leq \frac{1}{2}\log\left(1 + \frac{\mathcal{V}_{\max}(\mathcal{A}, \alpha\mathcal{A})}{\sigma^2}\right). \tag{44}$$

Proof: Since $\bar{X}$ and Z are independent, we know that the variance of Y cannot exceed $\mathcal{V}_{\max}(\mathcal{A}, \alpha\mathcal{A}) + \sigma^2$, and therefore

$$\mathsf{h}(Y) \leq \frac{1}{2}\log 2\pi e\left(\mathcal{V}_{\max}(\mathcal{A}, \alpha\mathcal{A}) + \sigma^2\right). \tag{45}$$

The bound follows by subtracting $\mathsf{h}(Z) = \frac{1}{2}\log 2\pi e\sigma^2$ from the above. ∎

Our last upper bound is more involved than the previous two. It holds only when $\alpha < \alpha_{\mathrm{th}}$.

Proposition 10 (Upper Bound for $\alpha < \alpha_{\mathrm{th}}$): If $\alpha < \alpha_{\mathrm{th}}$, then the capacity is upper-bounded as

$$\begin{aligned}
&\mathcal{C}_{\mathbf{h}^{\mathsf{T}},\sigma^2}(\mathcal{A}, \alpha\mathcal{A}) \\
&\leq \sup_{\mathbf{p}} \inf_{\delta,\mu>0} \Bigg\{ \frac{1}{2}\log\frac{\mathcal{A}^2 s_{n_{\mathrm{T}}}^2}{2\pi e\sigma^2} - \log\mu - \log\left(1 - 2\,\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right) \\
&\qquad + \mathcal{Q}\left(\frac{\delta}{\sigma}\right) + \frac{\delta}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{\delta^2}{2\sigma^2}} - \mathcal{D}\left(\mathbf{p}\,\middle\|\,\frac{\mathbf{h}}{s_{n_{\mathrm{T}}}}\right) \\
&\qquad + \sum_{k=1}^{n_{\mathrm{T}}} p_k \log\left(e^{\frac{\mu\delta}{\mathcal{A}h_k}} - e^{-\mu\left(1+\frac{\delta}{\mathcal{A}h_k}\right)}\right) \\
&\qquad + \frac{\mu\sigma}{\mathcal{A}\sqrt{2\pi}} \sum_{k=1}^{n_{\mathrm{T}}} \frac{p_k}{h_k}\left(e^{-\frac{\delta^2}{2\sigma^2}} - e^{-\frac{(\mathcal{A}h_k+\delta)^2}{2\sigma^2}}\right) \\
&\qquad + \mu\left(\alpha - \sum_{k=1}^{n_{\mathrm{T}}} p_k(k-1)\right)\Bigg\}
\end{aligned} \tag{46}$$

where the supremum is over all probability vectors $\mathbf{p} = (p_1, \ldots, p_{n_{\mathrm{T}}})$ satisfying

$$\sum_{k=1}^{n_{\mathrm{T}}} p_k(k-1) \leq \alpha. \tag{47}$$

Proof: See Section VI. ∎

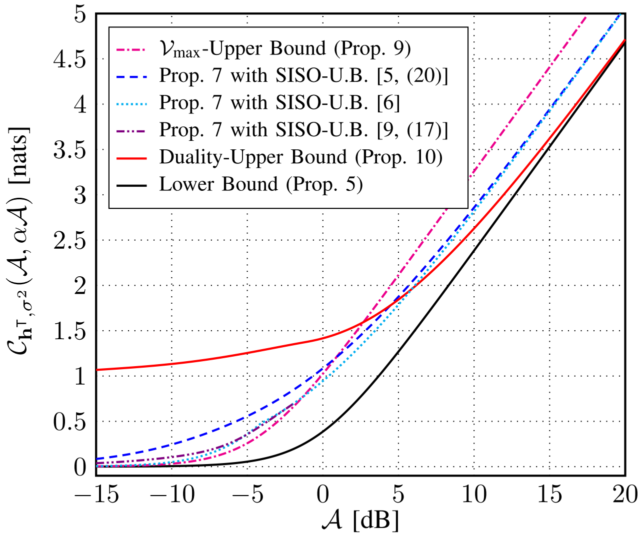

Fig. 1. Bounds on the capacity of the MISO channel with three LEDs and channel gains $\mathbf{h} = (3, 2, 1.5)$ for the case $\alpha = 0.6$. Note that the threshold of this channel is $\alpha_{\text{th}} = 1.2692$ (and $\frac{n_\text{T}}{2} = 1.5$). The upper bound from Proposition 7 is plotted in combination with three known SISO capacity bounds taken from [5], [6], and [9]. Note that the bound from [9] is only valid for $\mathcal{A} \leq -1.97$ dB.

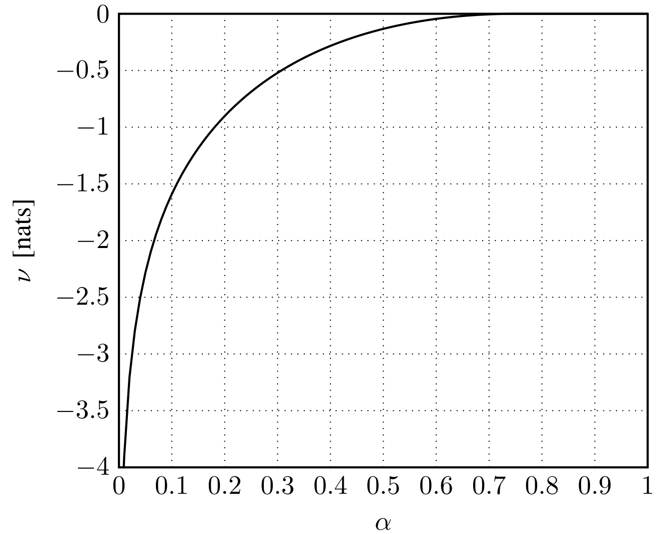

Fig. 2. The parameter ν, see (34) and (49), is depicted as a function of $\alpha \in (0, 1)$ for the two-LED MISO channel with gains $h_1 = 3$ and $h_2 = 1$. This represents the asymptotic capacity gap to the case with no average-power constraint, i.e., to $\mathcal{C}_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathcal{A},\mathcal{A})$.

Figure 1 shows our lower and upper bounds for the MISO channel with gains $\mathbf{h} = (3, 2, 1.5)$. It suggests that the upper bound in Proposition 10 and the lower bound in Proposition 5 are asymptotically tight as $\mathcal{A} \to \infty$. This is indeed the case, as we show in Proposition 11 below. In the low-SNR regime, all our upper and lower bounds, except the upper bound in Proposition 10, tend to zero, but not with the same slope. To characterize the capacity slope at low SNR, we shall use an asymptotic lower bound from [19]; see Proposition 13.

C. Asymptotic Results

Proposition 11 (High-SNR Asymptotics): If $\alpha \geq \alpha_{\text{th}}$, then

$$\lim_{\mathcal{A}\to\infty} \left\{\mathcal{C}_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathcal{A},\alpha\mathcal{A}) - \log\mathcal{A}\right\} = \frac{1}{2}\log\frac{s_{n_\text{T}}^2}{2\pi e\sigma^2}. \tag{48}$$

If $\alpha < \alpha_{\text{th}}$, then

$$\lim_{\mathcal{A}\to\infty} \left\{\mathcal{C}_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathcal{A},\alpha\mathcal{A}) - \log\mathcal{A}\right\} = \frac{1}{2}\log\frac{s_{n_\text{T}}^2}{2\pi e\sigma^2} + \nu \tag{49}$$

where $\nu \leq 0$ is defined in Proposition 5 (see (34)).

Proof: Achievability of (48) and (49) follows immediately from the lower bounds in Propositions 6 and 5, respectively. The converse to (48) is based on the upper bound (39) in Proposition 7 and the high-SNR analysis of the SISO capacity in [5, Cor. 6]. The converse to (49) is based on Proposition 10 and is given in detail in Section VII. ■

Example 12: Consider a MISO channel with two LEDs and with channel parameters $h_1 = 3$ and $h_2 = 1$. We plot ν against α in Figure 2. Note that ν characterizes the capacity gap to the case with no average-power constraint in the high-SNR limit. As expected, the gap becomes zero when α reaches $\alpha_{\text{th}} = 0.75$, and approaches infinity when α tends to zero. ♢

We now turn to the low-SNR asymptotic regime. In this regime the capacity is determined by $\mathcal{V}_{\max}(\mathcal{A},\alpha\mathcal{A})$.

Proposition 13 (Low-SNR Asymptotics): The low-SNR asymptotic capacity is

$$\lim_{\mathcal{A}\downarrow 0} \frac{\mathcal{C}_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathcal{A},\alpha\mathcal{A})}{\mathcal{A}^2/\sigma^2} = \frac{\gamma}{2} \tag{50}$$

where γ is defined in (43).

Proof: The converse follows immediately from the upper bound (44) in Proposition 9. Achievability follows from [19, Th. 2], which states that

$$\mathcal{C}_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathcal{A},\alpha\mathcal{A}) \geq \frac{\mathcal{V}_{\max}(\mathcal{A},\alpha\mathcal{A})}{2\sigma^2} + o(\mathcal{A}^2) \tag{51}$$

where $o(\mathcal{A}^2)$ decreases to 0 faster than $\mathcal{A}^2$, i.e.,

$$\lim_{\mathcal{A}\downarrow 0} \frac{o(\mathcal{A}^2)}{\mathcal{A}^2} = 0. \tag{52}$$

Note that the MISO channel under consideration in this paper satisfies the technical conditions A–F in [19]. ■

Example 14: We return to the example from before and reconsider the two-LED MISO channel of Figure 2 with channel gains $h_1 = 3$ and $h_2 = 1$. Figure 3 plots the low-SNR slope of its capacity $\gamma/2$ as a function of the parameter α. We notice that the low-SNR slope $\gamma/2$ does not reach a constant value when $\alpha \geq \alpha_{\text{th}}$, but it is strictly increasing for all values of $\alpha < \frac{n_\text{T}}{2}$. ♢

V. Achievability Proofs (Proofs of Lower Bounds)

A. Proof of Proposition 5

Fix

$$\lambda \in \left(\max\left\{0, \frac{1}{2} + \alpha - \alpha_{\text{th}}\right\}, \min\left\{\frac{1}{2}, \alpha\right\}\right) \tag{53}$$

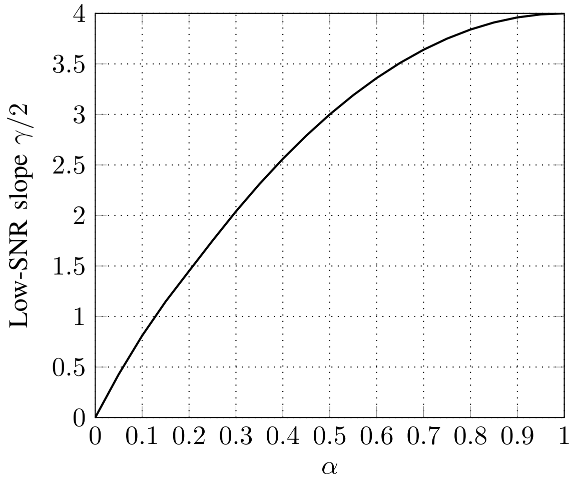

Fig. 3. The low-SNR asymptotic slope $\gamma/2$, see (50), is depicted as a function of $\alpha \in (0,1)$ for the two-LED MISO channel with gains $h_1 = 3$ and $h_2 = 1$ (the same channel as considered in Figure 2).

and choose a probability vector $\mathbf{p} = (p_1, \ldots, p_{n_\mathrm{T}})$ satisfying

$$\sum_{k=1}^{n_\mathrm{T}} p_k (k-1) = \alpha - \lambda. \tag{54}$$

Such a choice always exists. To see this, first note $0 < \alpha - \lambda < n_\mathrm{T}/2$. When we choose $\mathbf{p} = (1, 0, \ldots, 0)$, the LHS of (54) equals 0; when we choose $\mathbf{p} = (0, \ldots, 0, 1)$, it equals $n_\mathrm{T} - 1$, which is larger than or equal to $n_\mathrm{T}/2$ for all $n_\mathrm{T} \geq 2$. The existence of a $\mathbf{p}$ satisfying (54) then follows by the continuity of the LHS of (54) in $\mathbf{p}$.

Now let $\bar{X}$ have the following probability density function (PDF): for every $k \in \{1, \ldots, n_\mathrm{T}\}$, for $\bar{x} \in (s_{k-1}\mathcal{A}, s_k\mathcal{A}]$,

$$f_{\bar{X}}(\bar{x}) = \frac{p_k}{h_k \mathcal{A}} \frac{\mu(\lambda)}{1 - e^{-\mu(\lambda)}} e^{-\frac{\mu(\lambda)(\bar{x} - s_{k-1}\mathcal{A})}{h_k \mathcal{A}}} \tag{55}$$

where $\mu(\lambda)$ is the unique positive solution to (35); and let U be the random variable (RV) defined in (10) corresponding to this choice of $\bar{X}$. It is easy to check that the choice (55) indeed constitutes a PDF. Further, it satisfies (19) because it has positive probability only on the interval $[0, s_{n_\mathrm{T}}\mathcal{A}]$, and because

$$\begin{aligned}
&\sum_{k=1}^{n_\mathrm{T}} p_k \left(\frac{\mathsf{E}\left[\bar{X} | U = k\right] - \mathcal{A}s_{k-1}}{h_k} + (k-1)\mathcal{A} \right) \\
&\quad = \sum_{k=1}^{n_\mathrm{T}} p_k \left(\left(\frac{1}{\mu(\lambda)} - \frac{e^{-\mu(\lambda)}}{1 - e^{-\mu(\lambda)}} \right) \mathcal{A} + (k-1)\mathcal{A} \right) && (56) \\
&\quad = \sum_{k=1}^{n_\mathrm{T}} p_k \left(\lambda\mathcal{A} + (k-1)\mathcal{A} \right) && (57) \\
&\quad = \lambda\mathcal{A} + \sum_{k=1}^{n_\mathrm{T}} p_k (k-1)\mathcal{A} && (58) \\
&\quad = \alpha\mathcal{A}. && (59)
\end{aligned}$$

Here, (57) follows from (35), and (59) from (54). We then evaluate the mutual information in (18) for this $\bar{X}$ and use the Entropy Power Inequality (EPI) [21, Th. 17.7.3] to obtain

$$\begin{aligned}
&\mathcal{C}_{\mathbf{h}^\mathsf{T}, \sigma^2}(\mathcal{A}, \alpha\mathcal{A}) \\
&\quad \geq \mathcal{I}(\bar{X}; \bar{X} + Z) && (60) \\
&\quad = \mathsf{h}(\bar{X} + Z) - \mathsf{h}(Z) && (61) \\
&\quad \geq \frac{1}{2} \log\left(e^{2\,\mathsf{h}(\bar{X})} + e^{2\,\mathsf{h}(Z)}\right) - \mathsf{h}(Z) && (62) \\
&\quad = \frac{1}{2} \log\left(e^{2\,\mathsf{h}(\bar{X})} + 2\pi e\sigma^2\right) - \frac{1}{2}\log 2\pi e\sigma^2 && (63) \\
&\quad = \frac{1}{2} \log\left(1 + \frac{e^{2\,\mathsf{h}(\bar{X})}}{2\pi e\sigma^2}\right). && (64)
\end{aligned}$$

We next evaluate the differential entropy for the chosen $\bar{X}$:

$$\begin{aligned}
\mathsf{h}(\bar{X}) &= \mathsf{H}(U) - \underbrace{\mathsf{H}(U|\bar{X})}_{=0} + \mathsf{h}(\bar{X}|U) && (65) \\
&= \mathsf{H}(\mathbf{p}) + \sum_{k=1}^{n_\mathrm{T}} p_k\, \mathsf{h}(\bar{X}|U = k) && (66) \\
&= \mathsf{H}(\mathbf{p}) + \sum_{k=1}^{n_\mathrm{T}} p_k \log h_k + \log\mathcal{A} - \log\frac{\mu(\lambda)}{1 - e^{-\mu(\lambda)}} \\
&\quad + 1 - \frac{\mu(\lambda)\, e^{-\mu(\lambda)}}{1 - e^{-\mu(\lambda)}} && (67) \\
&= -\mathcal{D}\left(\mathbf{p} \middle\| \frac{\mathbf{h}}{s_{n_\mathrm{T}}}\right) + \log s_{n_\mathrm{T}}\mathcal{A} - \log\frac{\mu(\lambda)}{1 - e^{-\mu(\lambda)}} \\
&\quad + 1 - \frac{\mu(\lambda)\, e^{-\mu(\lambda)}}{1 - e^{-\mu(\lambda)}} && (68)
\end{aligned}$$

where (67) follows from the differential entropy expression of a truncated exponential RV.

The proposition then follows by plugging (68) into (64) and by maximizing the lower bound over the choice of the probability vector $\mathbf{p} = (p_1, \ldots, p_{n_\mathrm{T}})$ subject to constraint (54) and then maximizing over the choice of λ.

It only remains to show that the optimal choice of $\mathbf{p}$ in (34) indeed has the form given in (36). To that goal note that the first three terms inside the supremum in (34):

$$1 - \log\frac{\mu(\lambda)}{1 - e^{-\mu(\lambda)}} - \frac{\mu(\lambda)\, e^{-\mu(\lambda)}}{1 - e^{-\mu(\lambda)}} \tag{69}$$

correspond to the differential entropy of a truncated exponential RV $V \in [0,1]$ with mean $\mathsf{E}[V] = \lambda$ for $\lambda \in \left(0, \frac{1}{2}\right)$. This expression is monotonically strictly increasing in λ and approaches its maximum for $\lambda \to \frac{1}{2}$ (in which case V approaches a uniformly distributed random variable on [0, 1]).

The last term inside the supremum in (34):

$$-\mathcal{D}\left(\mathbf{p} \middle\| \frac{\mathbf{h}}{s_{n_\mathrm{T}}}\right) \tag{70}$$

is nonpositive and equals zero if, and only if, $\mathbf{p} = \frac{\mathbf{h}}{s_{n_\mathrm{T}}}$, in which case

$$\sum_{k=1}^{n_\mathrm{T}} p_k (k-1) = \alpha_\mathrm{th} - \frac{1}{2}. \tag{71}$$

Thus, as long as

$$\alpha - \lambda < \alpha_\mathrm{th} - \frac{1}{2} \tag{72}$$

the constraint (54) is active and [21, Probl. 12.2] tells us that the unique optimal $\mathbf{p}$ that maximizes (70) and simultaneously satisfies (54) has the form given in (36).

B. Proof of Proposition 6

We choose $\bar{X}$ to be uniform on $[0, s_{n_\mathrm{T}}\mathcal{A}]$. Let U be the RV defined in (10) for this choice of $\bar{X}$, then $\Pr[U = k] = p_k$ where

$$p_k \triangleq \frac{h_k}{s_{n_\mathrm{T}}}, \quad k = 1, \ldots, n_\mathrm{T} \tag{73}$$

and, conditional on $U = k$, $\bar{X}$ is uniform over $(s_{k-1}\mathcal{A}, s_k\mathcal{A}]$. The chosen $\bar{X}$ satisfies (19) because it has positive probability only on the interval $[0, s_{n_\mathrm{T}}\mathcal{A}]$ and because

$$\begin{aligned}
&\sum_{k=1}^{n_\mathrm{T}} p_k \left(\frac{\mathsf{E}\left[\bar{X}|U = k\right] - \mathcal{A}s_{k-1}}{h_k} + (k-1)\mathcal{A} \right) \\
&\quad = \sum_{k=1}^{n_\mathrm{T}} p_k \left(\frac{\mathcal{A}}{2} + (k-1)\mathcal{A} \right) && (74)\\
&\quad = \frac{\mathcal{A}}{2} + \mathcal{A}\sum_{k=1}^{n_\mathrm{T}} \frac{h_k}{s_{n_\mathrm{T}}}(k-1) && (75)\\
&\quad = \left(\frac{1}{2} + \frac{1}{s_{n_\mathrm{T}}}\sum_{k=1}^{n_\mathrm{T}} h_k(k-1) \right)\mathcal{A} && (76)\\
&\quad = \alpha_\mathrm{th}\mathcal{A} && (77)\\
&\quad \le \alpha\mathcal{A} && (78)
\end{aligned}$$

where (77) follows from the definition of α_th in (32). Like in (60)–(64), we obtain

$$\mathcal{C}_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathcal{A}, \alpha\mathcal{A}) \ge \frac{1}{2}\log\left(1 + \frac{e^{2\,\mathsf{h}(\bar{X})}}{2\pi e\sigma^2} \right) \tag{79}$$

for the above choice of $\bar{X}$. Since $\bar{X}$ is uniform, we have

$$\mathsf{h}(\bar{X}) = \log s_{n_\mathrm{T}}\mathcal{A}. \tag{80}$$

Plugging this into (79) yields the desired bound.

VI. Proof of Proposition 10

Choose $\bar{X} \sim Q^*_{\bar{X}}$ to be a maximizer of the capacity expression in (18). That means in particular that $Q^*_{\bar{X}}$ satisfies (19), and thus that the conditional expectations

$$\alpha_k^* \triangleq \mathsf{E}_{Q^*_{\bar{X}}}\left[\frac{\bar{X} - s_{k-1}\mathcal{A}}{h_k\mathcal{A}} \middle| U = k \right] \tag{81}$$

satisfy

$$\sum_{k=1}^{n_\mathrm{T}} p_k^*\big(\alpha_k^* + (k-1)\big) \le \alpha \tag{82}$$

where U is the RV defined in (10) and $p_k^* = \Pr[U = k]$.

The capacity is upper-bounded as follows:

$$\begin{aligned}
&\mathcal{C}_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathcal{A}, \alpha\mathcal{A}) \\
&\quad = \mathcal{I}(\bar{X}; \bar{X} + Z) && (83)\\
&\quad \le \mathcal{I}(\bar{X}; \bar{X} + Z, U) && (84)\\
&\quad = \mathcal{I}(\bar{X}; U) + \mathcal{I}(\bar{X}; \bar{X} + Z|U) && (85)\\
&\quad = \mathsf{H}(U) - \underbrace{\mathsf{H}(U|\bar{X})}_{=0} + \sum_{k=1}^{n_\mathrm{T}} p_k^*\,\mathcal{I}(\bar{X}; \bar{X} + Z|U = k) && (86)\\
&\quad = \mathsf{H}(\mathbf{p}^*) + \sum_{k=1}^{n_\mathrm{T}} p_k^*\,\mathcal{I}(\bar{X}; \bar{X} + Z|U = k) && (87)\\
&\quad = \mathsf{H}(\mathbf{p}^*) + \sum_{k=1}^{n_\mathrm{T}} p_k^*\,\mathcal{I}(X_k; h_kX_k + Z|U = k) && (88)\\
&\quad \le \mathsf{H}(\mathbf{p}^*) + \sum_{k=1}^{n_\mathrm{T}} p_k^*\mathcal{C}_{1,\sigma^2}(h_k\mathcal{A}, \alpha_k^* h_k\mathcal{A}) && (89)
\end{aligned}$$

where the inequality in (89) holds because, by Lemma 4, given $U = k$ we have $\bar{X} = s_{k-1}\mathcal{A} + h_kX_k$, where X_k lies on the interval $[0, \mathcal{A}]$ and is of average power $\alpha_k^*\mathcal{A}$.

The SISO capacity $\mathcal{C}_{1,\sigma^2}(h_k\mathcal{A}_k, \alpha_k^* h_k\mathcal{A})$ has been upper-bounded in [5, Eq. (12)]. Plugging this bound into (89) and performing some simple bounding steps prove that for every choice of positive parameters δ and μ:

$$\begin{aligned}
&\mathcal{C}_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathcal{A}, \alpha\mathcal{A}) \\
&\quad \le \mathsf{H}(\mathbf{p}^*) + \sum_{k=1}^{n_\mathrm{T}} p_k^* \log\left(\frac{\mathcal{A}h_k}{\sigma} \cdot \frac{e^{\frac{\mu\delta}{\mathcal{A}h_k}} - e^{-\mu\left(1+\frac{\delta}{\mathcal{A}h_k}\right)}}{\sqrt{2\pi}\,\mu\left(1 - 2\,\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)} \right) \\
&\qquad - \frac{1}{2} + \mathcal{Q}\left(\frac{\delta}{\sigma}\right) + \frac{\delta}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{\delta^2}{2\sigma^2}} + \mu\sum_{k=1}^{n_\mathrm{T}} p_k^*\alpha_k^* \\
&\qquad + \frac{\mu\sigma}{\mathcal{A}\sqrt{2\pi}} \sum_{k=1}^{n_\mathrm{T}} \frac{p_k^*}{h_k}\left(e^{-\frac{\delta^2}{2\sigma^2}} - e^{-\frac{(\mathcal{A}h_k+\delta)^2}{2\sigma^2}} \right) && (90)\\
&\quad \le \sum_{k=1}^{n_\mathrm{T}} p_k^* \log\frac{h_k}{p_k^*} + \log\mathcal{A} - \frac{1}{2}\log(2\pi e\sigma^2) - \log\mu \\
&\qquad - \log\left(1 - 2\,\mathcal{Q}\left(\frac{\delta}{\sigma}\right) \right) + \frac{\delta}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{\delta^2}{2\sigma^2}} \\
&\qquad + \sum_{k=1}^{n_\mathrm{T}} p_k^* \log\left(e^{\frac{\mu\delta}{\mathcal{A}h_k}} - e^{-\mu\left(1+\frac{\delta}{\mathcal{A}h_k}\right)} \right) + \mathcal{Q}\left(\frac{\delta}{\sigma}\right) \\
&\qquad + \frac{\mu\sigma}{\mathcal{A}\sqrt{2\pi}} \sum_{k=1}^{n_\mathrm{T}} \frac{p_k^*}{h_k}\left(e^{-\frac{\delta^2}{2\sigma^2}} - e^{-\frac{(\mathcal{A}h_k+\delta)^2}{2\sigma^2}} \right) \\
&\qquad + \mu\left(\alpha - \sum_{k=1}^{n_\mathrm{T}} p_k^*(k-1) \right) && (91)
\end{aligned}$$

where (91) follows from (82) and by rearranging terms. Since (91) holds for all $\delta, \mu > 0$, it must also hold when we take the infimum of its RHS over $\delta, \mu > 0$. Then we relax this bound by further taking a supremum[1] over $\mathbf{p}$, which establishes (46).

[1]Here we also need to make sure that $\mathbf{p}$ is chosen such that $\alpha - \sum_{k=1}^{n_\mathrm{T}} p_k^*(k-1) \ge 0$.

VII. Asymptotic High-SNR Analysis—Converse Proof to (49)

Recall that, throughout this section, we are only concerned with the case where

$$\alpha < \alpha_{\text{th}}. \tag{92}$$

Consider the upper bound in Proposition 10. We relax this bound by choosing specific values for δ and μ depending on $\mathbf{p} = (p_1, \ldots, p_{n_\text{T}})$ and $\mathcal{A}$. The relaxed upper bound will establish the converse to (49).

Fix $\mathcal{A} \geq 1$. For any $\mathbf{p} = (p_1, \ldots, p_{n_\text{T}})$ satisfying (47), define

$$\lambda = \lambda(\mathbf{p}) \triangleq \alpha - \sum_{k=1}^{n_\text{T}} p_k(k-1) \tag{93}$$

and fix some $0 < \zeta < 1$. We choose

$$\delta = \log(1+\mathcal{A}) \tag{94}$$

and

$$\mu = \begin{cases} \mu^*(\mathbf{p}) & \text{if } \frac{1}{\mathcal{A}^{1-\zeta}} < \lambda(\mathbf{p}) < \frac{1}{2} \\ \mathcal{A}^{1-\zeta} & \text{if } \lambda(\mathbf{p}) \leq \frac{1}{\mathcal{A}^{1-\zeta}} \\ \dfrac{1}{\mathcal{A}} & \text{if } \lambda(\mathbf{p}) \geq \frac{1}{2} \end{cases} \tag{95}$$

where $\mu^*(\mathbf{p})$ is the unique positive solution to

$$\frac{1}{\mu^*} - \frac{e^{-\mu^*}}{1-e^{-\mu^*}} = \lambda(\mathbf{p}). \tag{96}$$

Note that in the first case of (95),

$$\frac{1}{\mathcal{A}^{1-\zeta}} \leq \lambda(\mathbf{p}) = \frac{1}{\mu^*(\mathbf{p})} - \frac{e^{-\mu^*(\mathbf{p})}}{1-e^{-\mu^*(\mathbf{p})}} \leq \frac{1}{\mu^*(\mathbf{p})} \tag{97}$$

and thus

$$\mu^*(\mathbf{p}) \leq \mathcal{A}^{1-\zeta}. \tag{98}$$

Our choice (95) thus guarantees in all three cases that

$$\mu \leq \mathcal{A}^{1-\zeta} \quad \text{for } \mathcal{A} \geq 1 \tag{99}$$

and we can bound

$$\begin{aligned} &\frac{\mu\sigma}{\mathcal{A}\sqrt{2\pi}} \sum_{k=1}^{n_\text{T}} \frac{p_k}{h_k} \left(e^{-\frac{\delta^2}{2\sigma^2}} - e^{-\frac{(\mathcal{A}h_k+\delta)^2}{2\sigma^2}} \right) \\ &\quad \leq \frac{\sigma}{\mathcal{A}^\zeta \sqrt{2\pi}} \sum_{k=1}^{n_\text{T}} \frac{1}{h_k} \left(e^{-\frac{\delta^2}{2\sigma^2}} - e^{-\frac{(\mathcal{A}h_k+\delta)^2}{2\sigma^2}} \right) \end{aligned} \tag{100}$$

and

$$\begin{aligned} &\sum_{k=1}^{n_\text{T}} p_k \log\left(e^{\frac{\mu\delta}{\mathcal{A}h_k}} - e^{-\mu-\frac{\mu\delta}{\mathcal{A}h_k}} \right) \\ &\quad \leq \sum_{k=1}^{n_\text{T}} p_k \log\left(e^{\frac{\delta\mathcal{A}^{-\zeta}}{h_{n_\text{T}}}} - e^{-\mu-\frac{\delta\mathcal{A}^{-\zeta}}{h_{n_\text{T}}}} \right) \qquad (101) \\ &\quad = \log\left(e^{\frac{\delta\mathcal{A}^{-\zeta}}{h_{n_\text{T}}}} - e^{-\mu-\frac{\delta\mathcal{A}^{-\zeta}}{h_{n_\text{T}}}} \right) \qquad (102) \end{aligned}$$

where we have also used $p_k \leq 1$ and $h_k \geq h_{n_\text{T}}$.

With the described choices of μ, δ and the proposed relaxations, upper bound (46) becomes

$$\mathcal{C}_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathcal{A}, \alpha\mathcal{A}) \leq \frac{1}{2}\log\frac{\mathcal{A}^2 s_{n_\text{T}}^2}{2\pi e\sigma^2} + f(\mathcal{A}) + \sup_{\mathbf{p}} g(\mathcal{A}, \mathbf{p}, \mu) \tag{103}$$

with

$$\begin{aligned} f(\mathcal{A}) \triangleq\; &\mathcal{Q}\left(\frac{\log(1+\mathcal{A})}{\sigma}\right) - \log\left(1 - 2\,\mathcal{Q}\left(\frac{\log(1+\mathcal{A})}{\sigma}\right)\right) \\ &+ \frac{\sigma}{\mathcal{A}^\zeta\sqrt{2\pi}} \sum_{k=1}^{n_\text{T}} \frac{1}{h_k}\left(e^{-\frac{\log^2(1+\mathcal{A})}{2\sigma^2}} - e^{-\frac{(\mathcal{A}h_k+\log(1+\mathcal{A}))^2}{2\sigma^2}} \right) \\ &+ \frac{\log(1+\mathcal{A})}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{\log^2(1+\mathcal{A})}{2\sigma^2}} \end{aligned} \tag{104}$$

and

$$\begin{aligned} g(\mathcal{A}, \mathbf{p}, \mu) \triangleq\; &-\mathcal{D}\left(\mathbf{p} \middle\| \frac{\mathbf{h}}{s_{n_\text{T}}}\right) - \log\mu + \mu\left(\alpha - \sum_{k=1}^{n_\text{T}} p_k(k-1)\right) \\ &+ \log\left(e^{\frac{\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\text{T}}}} - e^{-\mu-\frac{\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\text{T}}}} \right). \end{aligned} \tag{105}$$

We note that

$$\lim_{\mathcal{A}\to\infty} f(\mathcal{A}) = 0. \tag{106}$$

Next, we bound $g(\mathcal{A}, \mathbf{p}, \mu)$ by bounding the function individually for the three different cases defined in (95), and then taking the maximum over the three obtained bounds. Notice first that when $\lambda(\mathbf{p}) = \alpha - \sum_{k=1}^{n_\text{T}} p_k(k-1)$ lies in the open interval $(\mathcal{A}^{1-\zeta}, 1/2)$,

$$\begin{aligned} &g(\mathcal{A}, \mathbf{p}, \mu) \\ &= -\mathcal{D}\left(\mathbf{p} \middle\| \frac{\mathbf{h}}{s_{n_\text{T}}}\right) - \log\mu^*(\mathbf{p}) + \mu^*(\mathbf{p})\lambda(\mathbf{p}) \\ &\quad + \log\left(e^{\frac{\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\text{T}}}} - e^{-\mu^*(\mathbf{p})-\frac{\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\text{T}}}} \right) \qquad (107) \\ &= -\mathcal{D}\left(\mathbf{p} \middle\| \frac{\mathbf{h}}{s_{n_\text{T}}}\right) - \log\mu^*(\mathbf{p}) \\ &\quad + \mu^*(\mathbf{p})\left(\frac{1}{\mu^*(\mathbf{p})} - \frac{e^{-\mu^*(\mathbf{p})}}{1-e^{-\mu^*(\mathbf{p})}} \right) \\ &\quad + \log\left(e^{\frac{\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\text{T}}}} - e^{-\mu^*(\mathbf{p})-\frac{\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\text{T}}}} \right) \qquad (108) \\ &\leq \sup_{\mathbf{p}\colon \lambda(\mathbf{p}) \in \left(\frac{1}{\mathcal{A}^{1-\zeta}}, \frac{1}{2}\right)} \Bigg\{ -\mathcal{D}\left(\mathbf{p} \middle\| \frac{\mathbf{h}}{s_{n_\text{T}}}\right) - \log\mu^*(\mathbf{p}) \\ &\qquad + \mu^*(\mathbf{p})\left(\frac{1}{\mu^*(\mathbf{p})} - \frac{e^{-\mu^*(\mathbf{p})}}{1-e^{-\mu^*(\mathbf{p})}} \right) \\ &\qquad + \log\left(e^{\mathcal{A}^{-\zeta}\frac{\log(1+\mathcal{A})}{h_{n_\text{T}}}} - e^{-\mu^*(\mathbf{p})-\frac{\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\text{T}}}} \right) \Bigg\} \qquad (109) \end{aligned}$$

$$= \sup_{\mathbf{p}:\,\lambda(\mathbf{p})\in\left(\frac{1}{\mathcal{A}^{1-\zeta}},\frac{1}{2}\right)} \Biggl\{ -\mathcal{D}\left(\mathbf{p}\middle\|\frac{\mathbf{h}}{s_{n_\mathrm{T}}}\right) - \log\frac{\mu^*(\mathbf{p})}{1-e^{-\mu^*(\mathbf{p})}} + 1 - \frac{\mu^*(\mathbf{p})\,e^{-\mu^*(\mathbf{p})}}{1-e^{-\mu^*(\mathbf{p})}} + \log\left(\frac{e^{\frac{2\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\mathrm{T}}}} - e^{-\mu^*(\mathbf{p})}}{1-e^{-\mu^*(\mathbf{p})}}\right)\Biggr\} - \frac{\log(1+\mathcal{A})}{\mathcal{A}^\zeta h_{n_\mathrm{T}}} \tag{110}$$

$$\triangleq g_1(\mathcal{A}). \tag{111}$$

Here, (108) follows from (96); the inequality in (109) holds because $\lambda(\mathbf{p}) \in \left(\mathcal{A}^{\zeta-1}, 1/2\right)$; and (110) follows by rearranging terms.

When $\lambda(\mathbf{p}) \le \mathcal{A}^{\zeta-1}$,

$$\begin{aligned} &g\left(\mathcal{A},\mathbf{p},\mu\right)\\ &= -\mathcal{D}\left(\mathbf{p}\middle\|\frac{\mathbf{h}}{s_{n_\mathrm{T}}}\right) - \log\frac{\mathcal{A}^{1-\zeta}}{e^{\frac{\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\mathrm{T}}}} - e^{-\mathcal{A}^{1-\zeta}-\frac{\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\mathrm{T}}}}}\\ &\quad + \mathcal{A}^{1-\zeta}\lambda(\mathbf{p}) \qquad (112)\\ &\le -\mathcal{D}\left(\mathbf{p}\middle\|\frac{\mathbf{h}}{s_{n_\mathrm{T}}}\right) - \log\frac{\mathcal{A}^{1-\zeta}}{e^{\frac{\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\mathrm{T}}}} - e^{-\mathcal{A}^{1-\zeta}-\frac{\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\mathrm{T}}}}}\\ &\quad + 1 \qquad (113)\\ &= \underbrace{\sum_{k=1}^{n_\mathrm{T}} p_k\log h_k}_{\le \log h_1} + \underbrace{\sum_{k=1}^{n_\mathrm{T}} p_k\log\frac{1}{p_k}}_{\le \log n_\mathrm{T}} + \log s_{n_\mathrm{T}}\\ &\quad - \log\frac{\mathcal{A}^{1-\zeta}}{e^{\frac{\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\mathrm{T}}}} - \underbrace{e^{-\mathcal{A}^{1-\zeta}-\frac{\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\mathrm{T}}}}}_{\ge 0}} + 1 \qquad (114)\\ &\le -(1-\zeta)\log\mathcal{A} + \log h_1 + \log n_\mathrm{T} + \log s_{n_\mathrm{T}}\\ &\quad + \frac{\log(1+\mathcal{A})}{\mathcal{A}^\zeta h_{n_\mathrm{T}}} + 1 \qquad (115)\\ &\triangleq g_2(\mathcal{A}) \qquad (116) \end{aligned}$$

where the inequality in (113) follows because $\lambda(\mathbf{p}) \le \mathcal{A}^{\zeta-1}$.

Finally, when $\lambda(\mathbf{p}) \ge 1/2$,

$$\begin{aligned} &g(\mathcal{A},\mathbf{p},\mu)\\ &= -\mathcal{D}\left(\mathbf{p}\middle\|\frac{\mathbf{h}}{s_{n_\mathrm{T}}}\right) - \log\frac{\mathcal{A}^{-1}}{e^{\frac{\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\mathrm{T}}}} - e^{-\mathcal{A}^{-1}-\frac{\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\mathrm{T}}}}}\\ &\quad + \frac{1}{\mathcal{A}}\underbrace{\left(\alpha - \sum_{k=1}^{n_\mathrm{T}} p_k(k-1)\right)}_{\le\alpha} \qquad (117)\\ &\le -\mathcal{D}\left(\mathbf{p}\middle\|\frac{\mathbf{h}}{s_{n_\mathrm{T}}}\right) - \log\frac{\mathcal{A}^{-1}}{e^{\frac{\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\mathrm{T}}}} - e^{-\mathcal{A}^{-1}-\frac{\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\mathrm{T}}}}}\\ &\quad + \frac{\alpha}{\mathcal{A}} \qquad (118)\\ &\le -\inf_{\mathbf{p}:\,\lambda(\mathbf{p})\ge\frac{1}{2}} \mathcal{D}\left(\mathbf{p}\middle\|\frac{\mathbf{h}}{s_{n_\mathrm{T}}}\right)\\ &\quad - \log\frac{\mathcal{A}^{-1}}{e^{\frac{\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\mathrm{T}}}} - e^{-\mathcal{A}^{-1}-\frac{\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\mathrm{T}}}}} + \frac{\alpha}{\mathcal{A}} \qquad (119)\\ &\triangleq g_3(\mathcal{A}). \qquad (120) \end{aligned}$$

We note that depending on the value of $\lambda(\mathbf{p})$, the function $g(\mathcal{A},\mathbf{p},\mu)$ is upper-bounded by one of the three functions $g_1(\mathcal{A})$, $g_2(\mathcal{A})$, or $g_3(\mathcal{A})$. Thus, it is also upper-bounded by their maximum:

$$g(\mathcal{A},\mathbf{p},\mu) \le \max\bigl\{g_1(\mathcal{A}), g_2(\mathcal{A}), g_3(\mathcal{A})\bigr\}. \tag{121}$$

We now analyze this maximum when $\mathcal{A}\to\infty$. Since $g_2(\mathcal{A})$ tends to $-\infty$ as $\mathcal{A}\to\infty$ and since $g_1(\mathcal{A})$ and $g_3(\mathcal{A})$ are both bounded for $\mathcal{A} > 1$, $g_2(\mathcal{A})$ is strictly smaller than $\max\{g_1(\mathcal{A}), g_3(\mathcal{A})\}$ for $\mathcal{A}$ large enough. Moreover,

$$\begin{aligned} \lim_{\mathcal{A}\to\infty} g_3(\mathcal{A}) &= -\inf_{\mathbf{p}:\,\alpha-\sum_{k=1}^{n_\mathrm{T}} p_k(k-1)\ge\frac{1}{2}} \mathcal{D}\left(\mathbf{p}\middle\|\frac{\mathbf{h}}{s_{n_\mathrm{T}}}\right) \qquad (122)\\ &= -\inf_{\mathbf{p}:\,\alpha-\sum_{k=1}^{n_\mathrm{T}} p_k(k-1)=\frac{1}{2}} \mathcal{D}\left(\mathbf{p}\middle\|\frac{\mathbf{h}}{s_{n_\mathrm{T}}}\right) \qquad (123) \end{aligned}$$

where the second equality follows because, given $\alpha < \alpha_\mathrm{th}$, an optimal choice of $\mathbf{p}$ will make full use of the available constraint

$$\sum_{k=1}^{n_\mathrm{T}} p_k(k-1) \le \alpha - \frac{1}{2}. \tag{124}$$

It remains to investigate the behavior of $g_1(\mathcal{A})$ for $\mathcal{A}\to\infty$. To that goal define

$$\tilde{g}_1(\mathcal{A},\mathbf{p}) \triangleq -\mathcal{D}\left(\mathbf{p}\middle\|\frac{\mathbf{h}}{s_{n_\mathrm{T}}}\right) - \log\frac{\mu^*(\mathbf{p})}{1-e^{-\mu^*(\mathbf{p})}} - \frac{\mu^*(\mathbf{p})\,e^{-\mu^*(\mathbf{p})}}{1-e^{-\mu^*(\mathbf{p})}} + 1 + \log\left(\frac{e^{\frac{2\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\mathrm{T}}}} - e^{-\mu^*(\mathbf{p})}}{1-e^{-\mu^*(\mathbf{p})}}\right) \tag{125}$$

where $\mu^*(\mathbf{p})$ is the unique positive solution to (96) (with $\lambda(\mathbf{p})$ defined in (93)).

Notice that

$$g_1(\mathcal{A}) = \sup_{\mathbf{p}:\,\alpha-\sum_{k=1}^{n_\mathrm{T}} p_k(k-1)\in\left(\frac{1}{\mathcal{A}^{1-\zeta}},\frac{1}{2}\right)} \tilde{g}_1(\mathcal{A},\mathbf{p}) - \frac{\log(1+\mathcal{A})}{\mathcal{A}^\zeta h_{n_\mathrm{T}}}. \tag{126}$$

Note further that, for a fixed $\mathbf{p}$,

$$\begin{aligned} \Delta(\mathcal{A},\mathbf{p}) &\triangleq \tilde{g}_1(\mathcal{A},\mathbf{p}) - \lim_{\mathcal{A}\to\infty}\tilde{g}_1(\mathcal{A},\mathbf{p}) \qquad (127)\\ &= \log\left(\frac{e^{\frac{2\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_\mathrm{T}}}} - e^{-\mu^*(\mathbf{p})}}{1-e^{-\mu^*(\mathbf{p})}}\right). \qquad (128) \end{aligned}$$

Since $\frac{b-\xi}{1-\xi}$ is increasing in ξ for $b > 1$ and because of (98),

$$0 \geq \Delta(\mathcal{A}, \mathbf{p}) \geq \log\left(\frac{e^{\frac{2\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_{\mathrm{T}}}}} - e^{-\mathcal{A}^{1-\zeta}}}{1 - e^{-\mathcal{A}^{1-\zeta}}}\right) \tag{129}$$

and therefore for any $\mathbf{p}$:

$$\left|\Delta(\mathcal{A}, \mathbf{p})\right| \leq \left|\log\left(\frac{e^{\frac{2\mathcal{A}^{-\zeta}\log(1+\mathcal{A})}{h_{n_{\mathrm{T}}}}} - e^{-\mathcal{A}^{1-\zeta}}}{1 - e^{-\mathcal{A}^{1-\zeta}}}\right)\right| \tag{130}$$

$$\overset{\mathcal{A}\to\infty}{\to} \log(1) = 0. \tag{131}$$

This proves that $\tilde{g}_1(\mathcal{A}, \mathbf{p})$ converges uniformly as $\mathcal{A} \to \infty$ and we are allowed to swap limit and supremum to obtain

$$\lim_{\mathcal{A}\to\infty} g_1(\mathcal{A})$$

$$= \lim_{\mathcal{A}\to\infty}\left\{\sup_{\mathbf{p}:\, \alpha-\sum_{k=1}^{n_{\mathrm{T}}} p_k(k-1)\in\left(\frac{1}{\mathcal{A}^{1-\zeta}}, \frac{1}{2}\right)} \tilde{g}_1(\mathcal{A}, \mathbf{p}) - \frac{\log(1+\mathcal{A})}{\mathcal{A}^{\zeta} h_{n_{\mathrm{T}}}}\right\} \tag{132}$$

$$= \sup_{\mathbf{p}:\, \alpha-\sum_{k=1}^{n_{\mathrm{T}}} p_k(k-1)\in\left(0, \frac{1}{2}\right)} \lim_{\mathcal{A}\to\infty} \tilde{g}_1(\mathcal{A}, \mathbf{p}) \tag{133}$$

$$= \sup_{\mathbf{p}:\, \alpha-\sum_{k=1}^{n_{\mathrm{T}}} p_k(k-1)\in\left(0, \frac{1}{2}\right)} \left\{1 - \log\frac{\mu^*(\mathbf{p})}{1-e^{-\mu^*(\mathbf{p})}} - \frac{\mu^*(\mathbf{p})\, e^{-\mu^*(\mathbf{p})}}{1-e^{-\mu^*(\mathbf{p})}} - \mathcal{D}\left(\mathbf{p}\middle\|\frac{\mathbf{h}}{s_{n_{\mathrm{T}}}}\right)\right\}. \tag{134}$$

Since for any $\lambda \in \left(0, \frac{1}{2}\right)$,

$$1 - \log\frac{\mu^*}{1-e^{-\mu^*}} - \frac{\mu^* e^{-\mu^*}}{1-e^{-\mu^*}} \geq 0 \tag{135}$$

and since

$$\lim_{\lambda\to 0}\left\{1 - \log\frac{\mu^*}{1-e^{-\mu^*}} - \frac{\mu^* e^{-\mu^*}}{1-e^{-\mu^*}}\right\} = 0 \tag{136}$$

we conclude by comparing (123) and (134) that for sufficiently large values of $\mathcal{A}$, $g_1(\mathcal{A}) \geq g_3(\mathcal{A})$, and thus

$$\lim_{\mathcal{A}\to\infty} \sup_{\mathbf{p}} g(\mathcal{A}, \mathbf{p}, \mu)$$

$$\leq \lim_{\mathcal{A}\to\infty} g_1(\mathcal{A}) \tag{137}$$

$$= \sup_{\mathbf{p}:\, \alpha-\sum_{k=1}^{n_{\mathrm{T}}} p_k(k-1)\in\left(0, \frac{1}{2}\right)} \left\{1 - \log\frac{\mu^*(\mathbf{p})}{1-e^{-\mu^*(\mathbf{p})}} - \frac{\mu^*(\mathbf{p})\, e^{-\mu^*(\mathbf{p})}}{1-e^{-\mu^*(\mathbf{p})}} - \mathcal{D}\left(\mathbf{p}\middle\|\frac{\mathbf{h}}{s_{n_{\mathrm{T}}}}\right)\right\} \tag{138}$$

$$= \sup_{\lambda\in\left(\max\left\{0, \frac{1}{2}+\alpha-\alpha_{\mathrm{th}}\right\}, \min\left\{\frac{1}{2}, \alpha\right\}\right)} \left\{1 - \log\frac{\mu(\lambda)}{1-e^{-\mu(\lambda)}} - \frac{\mu(\lambda)\, e^{-\mu(\lambda)}}{1-e^{-\mu(\lambda)}} - \inf_{\mathbf{p}:\, \alpha-\sum_{k=1}^{n_{\mathrm{T}}} p_k(k-1)=\lambda} \mathcal{D}\left(\mathbf{p}\middle\|\frac{\mathbf{h}}{s_{n_{\mathrm{T}}}}\right)\right\} \tag{139}$$

where in (139) $\mu(\lambda)$ is the unique positive solution to (35). Note that for the re-parametrizing in (139) we have defined

$$\lambda = \alpha - \sum_{k=1}^{n_{\mathrm{T}}} p_k(k-1) \tag{140}$$

and then used the same argumentation as given at the end of Section V-A to restrict the required range of λ. By [21, Probl. 12.2] it then follows that the infimum is achieved for the $\mathbf{p}$ given in (36).

Combining (139) with (103) and (106) then proves the proposition.

VIII. Concluding Remarks

In this paper we present upper and lower bounds on the capacity of a multiple-input and single-output (MISO) free-space optical intensity channel with signal-independent additive Gaussian noise and with both a peak- and an average-power constraint on the input. Asymptotically, when both peak and average power tend either to infinity or to zero (with their ratio held fixed), we succeed in specifying the capacity exactly.

At low SNR, a good input vector $\mathbf{X}$ maximizes the variance of $\mathbf{h}^{\mathsf{T}}\mathbf{X}$ under the given power constraints. This is achieved by $\mathbf{X}$ having only entries of $\mathcal{A}$ and 0, i.e., by each LED sending either full or no power.

At high SNR, a good input vector $\mathbf{X}$ maximizes the differential entropy $\mathsf{h}(\mathbf{h}^{\mathsf{T}}\mathbf{X})$. For the case of only an average-power constraint or only a peak-power constraint (or both a peak- and an average-power constraint but with the latter being sufficiently loose), this is relatively straightforward. For the general situation of both a peak- and an average-power constraint, maximizing $\mathsf{h}(\mathbf{h}^{\mathsf{T}}\mathbf{X})$ is more involved. The optimal input can be found based on two insights. First, in order to reach a certain range of amplitude levels $\mathbf{h}^{\mathsf{T}}\mathbf{X} \in (s_{k-1}\mathcal{A}, s_k\mathcal{A}]$, it is most energy-efficient to set all LEDs with strong channel gains to the maximum level, $X_j = \mathcal{A}$, $j = 1, \ldots, k-1$; to switch the weaker LEDs off, $X_j = 0$, $j = k+1, \ldots, n_{\mathrm{T}}$; and to exclusively use X_k to signal. Second, conditional on a given range $(s_{k-1}\mathcal{A}, s_k\mathcal{A}]$, X_k should have a truncated exponential distribution in order to maximize the conditional differential entropy under the given power constraints. It then only remains to optimize over the probability masses assigned to each of the different amplitude ranges and the parameters of the truncated exponentials. Note that this optimization characterizes an implicit trade-off: higher probabilities on the higher amplitude ranges will increase the effectively used total range of $\mathbf{h}^{\mathsf{T}}\mathbf{X}$, but at the cost of using more power for the LEDs that are set deterministically to $\mathcal{A}$.

Our lower bound in Proposition 5 is based on such a choice of input distribution. Our upper bound in Proposition 10 and its asymptotic analysis in Section VII are based on the same intuition, but borrow from known upper bounds on the SISO capacity. Alternatively, one can also derive a new upper bound using the duality-based bounding technique [18], as we outlined in [17]. Such a bound, however, is much harder to prove.

A close look at the results in [10] confirms that also for the MIMO optical intensity channel when the channel matrix H

has full column rank, the high-SNR asymptotic capacity is given by the maximum differential entropy of $\mathsf{H}\mathbf{X}$ minus that of the noise vector. With the current work and [10], the only MIMO optical intensity channels whose high-SNR asymptotic capacities are not yet known are those with more than one receive antennas (photodetectors), and with channel matrices that do not have full column rank. It is natural to conjecture that, for those channels, the high-SNR asymptotic capacity is again given by the maximum of $\mathsf{h}(\mathsf{H}\mathbf{X})$ minus the differential entropy of the noise.

APPENDIX
PROOF OF LEMMA 8

The variance of $\bar{X}$ can be decomposed as

$$\begin{aligned}\mathsf{E}\left[\left(\bar{X}-\mathsf{E}[\bar{X}]\right)^2\right] &= \sum_{k=1}^{n_\mathrm{T}} h_k^2\,\mathsf{E}\left[(X_k-\mathsf{E}\left[X_k\right])^2\right] \\ &\quad+\sum_{\substack{i,j=1\\ i\neq j}}^{n_\mathrm{T}} h_i h_j\left(\mathsf{E}\left[X_i X_j\right]-\mathsf{E}\left[X_i\right]\mathsf{E}\left[X_j\right]\right). \qquad (141)\end{aligned}$$

Let us fix the joint distribution on $(X_1,\ldots,X_{n_\mathrm{T}-1})$, and fix with probability one the conditional mean $\mathsf{E}\left[X_{n_\mathrm{T}}\middle|X_1,\ldots,X_{n_\mathrm{T}-1}\right]$. These determine the consumed average input power, as well as every summand on the RHS of (141) except $\mathsf{E}\left[(X_{n_\mathrm{T}}-[X_{n_\mathrm{T}}])^2\right]$. For any choice above, the value of $\mathsf{E}\left[(X_{n_\mathrm{T}}-[X_{n_\mathrm{T}}])^2\right]$ is maximized by X_{n_T} taking value only in the set $\{0,\mathcal{A}\}$. We hence conclude that, to maximize the variance (141) subject to a constraint on average input power, it is optimal to restrict X_{n_T} to taking value only in $\{0,\mathcal{A}\}$. Repeating this argument, we conclude that every X_k, $k=1,\ldots,n_\mathrm{T}$, should take value only in $\{0,\mathcal{A}\}$.

Next, using the same argument as in Lemma 4, we know that it is optimal to consider joint distributions as follows: for each $k\in\{0,\ldots,n_\mathrm{T}\}$, with probability q_k

$$X_1=\cdots=X_k=\mathcal{A} \quad \text{and} \quad X_{k+1}=\cdots=X_{n_\mathrm{T}}=0. \qquad (142)$$

Such a choice produces an $\bar{X}$ that takes value only in (41). This proves Part 1 of the lemma. Further, this choice of inputs consumes an average power of $\sum_{k=0}^{n_\mathrm{T}} q_k k$. The condition for a probability vector $q_0,\ldots,q_{n_\mathrm{T}}$ to be valid is thus

$$\sum_{k=0}^{n_\mathrm{T}} q_k k \le \alpha. \qquad (143)$$

Part 2 of the lemma is then proven by noting that with the choice in (142), the variance of $\bar{X}$ is

$$\begin{aligned}\mathsf{E}\left[\left(\bar{X}-\mathsf{E}[\bar{X}]\right)^2\right] &= \mathsf{E}\left[\bar{X}^2\right]-\left(\mathsf{E}[\bar{X}]\right)^2 \qquad (144)\\ &= \sum_{k=1}^{n_\mathrm{T}} q_k\mathcal{A}^2\ s_k^2-\left(\sum_{k=1}^{n_\mathrm{T}} q_k\mathcal{A}s_k\right)^2. \qquad (145)\end{aligned}$$

ACKNOWLEDGMENTS

We gratefully acknowledge helpful discussions with Tobias Koch and Christoph Pfister. In particular, Tobi's suggestion helped us simplify Proposition 10 and its proof.

REFERENCES

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

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

[3] M. A. Khalighi and M. Uysal, "Survey on free space optical communication: A communication theory perspective," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 2231–2258, 4th Quart., 2014.

[4] M. A. Wigger, "Bounds on the capacity of free-space optical intensity channels," M.S. thesis, Signal Inf. Process. Lab., ETH Zürich, Zürich, Switzerland, Mar. 2003.

[5] 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.

[6] A. L. McKellips, "Simple tight bounds on capacity for the peak-limited discrete-time channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Chicago, IL, USA, Jun./Jul. 2004, p. 348.

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

[8] 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.

[9] 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.

[10] S. M. Moser, M. Mylonakis, L. Wang, and M. Wigger, "Asymptotic capacity results for MIMO wireless optical communication," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, Jun. 2017, pp. 536–540.

[11] A. Chaaban, Z. Rezki, and M.-S. Alouini, "MIMO intensity-modulation channels: Capacity bounds and high SNR characterization," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, May 2017, pp. 1–6.

[12] S. M. Haas, J. H. Shapiro, and V. Tarokh, "Space-time codes for wireless optical communications," *EURASIP J. Appl. Signal Process.*, vol. 2002, no. 3, pp. 211–220, Dec. 2002.

[13] 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.

[14] X. Song and J. Cheng, "Subcarrier intensity modulated MIMO optical communications in atmospheric turbulence," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 5, no. 9, pp. 1001–1009, Sep. 2013.

[15] L. Mroueh and J.-C. Belfiore, "Quadratic extension field codes for free space optical intensity communications," *IEEE Trans. Commun.*, vol. 65, no. 2, pp. 751–763, Feb. 2017.

[16] A. Chaaban, Z. Rezki, and M. S. Alouini, "Fundamental limits of parallel optical wireless channels: Capacity results and outage formulation," *IEEE Trans. Commun.*, vol. 65, no. 1, pp. 296–311, Jan. 2017.

[17] S. M. Moser, L. Wang, and M. Wigger, "Asymptotic high-SNR capacity of MISO optical intensity channels," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Kaohsiung, Taiwan, Nov. 2017, pp. 86–90.

[18] A. Lapidoth and S. M. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2426–2467, Oct. 2003.

[19] V. V. Prelov and E. C. van der Meulen, "An asymptotic expression for the information and capacity of a multidimensional channel with weak input signals," *IEEE Trans. Inf. Theory*, vol. 39, no. 5, pp. 1728–1735, Sep. 1993.

[20] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423 and 623–656, Jul./Oct. 1948.

[21] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York, NY, USA: Wiley, 2006.

[22] J. G. Smith, "The information capacity of amplitude- and variance-constrained scalar Gaussian channels," *Inf. Control*, vol. 18, no. 3, pp. 203–219, Apr. 1971.

Stefan M. Moser (S'01–M'05–SM'10) received the diploma (M.Sc.) in electrical engineering, with distinction, in 1999, the M.Sc. degree in industrial management (M.B.A.) in 2003, and the Ph.D. degree (Dr. sc. techn.) in the field of information theory in 2004, all from ETH Zurich, Switzerland. From 1999 to 2003, he was a Research and Teaching Assistant, and from 2004 to 2005, he was a Senior Research Assistant with the Signal and Information Processing Laboratory, ETH Zurich. From 2005 to 2013, he was a Professor with the Department of Electrical and Computer Engineering, National Chiao Tung University (NCTU), Hsinchu, Taiwan. Currently he is a Senior Researcher and a Lecturer with the Signal and Information Processing Laboratory, ETH Zurich, and an Adjunct Professor with the Institute of Communications Engineering, NCTU, Hsinchu, Taiwan. Besides he also teaches math at the Kantonsschule Uster, Switzerland. He is an Associate Editor for the IEEE TRANSACTIONS ON MOLECULAR, BIOLOGICAL, AND MULTI-SCALE COMMUNICATIONS.

Dr. Moser was a recipient of the 2012 Wu Ta-You Memorial Award from the National Science Council of Taiwan, the 2009 Best Paper Award for Young Scholars from the IEEE Communications Society Taipei and Tainan Chapters and the IEEE Information Theory Society Taipei Chapter. He received various awards from the National Chiao Tung University, including two awards for outstanding teaching, the Willi Studer Award of ETH, the ETH Medal, and the Sandoz (Novartis) Basler Maturandenpreis. He was named an IEEE Communications Society Exemplary Reviewer.

Ligong Wang (S'08–M'12) received the B.E. degree in electronic engineering from Tsinghua University, Beijing, China, in 2004, and the M.Sc. and Dr.Sc. degrees in electrical engineering from ETH Zurich, Switzerland, in 2006 and 2011, respectively. In the years 2011–2014 he was a Postdoctoral Associate at the Department of Electrical Engineering and Computer Science at the Massachusetts Institute of Technology, Cambridge, MA, USA. He is now a researcher (chargé de recherche) with CNRS, France, and is affiliated with ETIS laboratory in Cergy-Pontoise. His research interests include classical and quantum information theory, physical-layer security, and digital, in particular optical communications.

Michèle Wigger (S'05–M'09–SM'14) received the M.Sc. degree in electrical engineering, with distinction, and the Ph.D. degree in electrical engineering both from ETH Zurich in 2003 and 2008, respectively. In 2009, she was first a post-doctoral fellow at the University of California, San Diego, USA, and then joined Telecom Paris Tech, Paris, France, where she is currently a full professor. Dr. Wigger has held visiting professor appointments at the Technion–Israel Institute of Technology and ETH Zurich. Dr. Wigger has previously served as an Associate Editor of the IEEE COMMUNICATION LETTERS, and is now Associate Editor for Shannon Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY. She is currently also serving on the Board of Governors of the IEEE Information Theory Society. Dr. Wigger's research interests are in multi-terminal information theory, in particular in distributed source coding and in capacities of networks with states, feedback, user cooperation, or caching.