The MISO Free-Space Optical Channel at Low and Moderate SNR

Longguang Li*, Stefan M. Moser†‡, Ligong Wang§, and Michèle Wigger*

*LTCI, Telecom ParisTech, Université Paris-Saclay, 75013 Paris, France
†Signal and Information Processing Lab, ETH Zürich, Switzerland
‡Institute of Communications Engineering, National Chiao Tung University, Hsinchu, Taiwan
§ETIS—Université Paris Seine, Université de Cergy-Pontoise, ENSEA, CNRS, Cergy-Pontoise, France

Abstract—The capacity of the multiple-input single-output (MISO) free-space optical channel with a per-antenna peak-power constraint and a sum (over all antennas) average-power constraint is studied. The asymptotic low-signal-to-noise-ratio (low-SNR) capacity is determined exactly and close upper and lower bounds are presented in the low- and moderate-SNR regimes. The asymptotic low-SNR limit is achieved by having each transmit antenna signal either with zero or with the maximally allowed peak power, and the latter only if all stronger antennas also send at maximum peak power. In particular, for almost all channel gains, the input to achieve the asymptotic low-SNR capacity is such that its projection on the channel-gain vector has only two or three positive probability point masses, one of them being at zero. The lower bounds at finite SNR are numerical and are obtained using input distributions whose projection on the channel-gain vector has either two, three, or four positive probability masses. Finally, the paper presents two analytic upper bounds on the capacity of the MISO channel: the first one closely follows the proposed numerical lower bounds in the low-SNR regime, and the second one can improve on previous bounds in the moderate-SNR regime.

I. INTRODUCTION AND CHANNEL MODEL

In optical wireless communication, visible, infrared, or ultraviolet light is transmitted in free space to carry a message to its destination [1]–[3]. Particularly attractive are simple *intensity-modulation–direct-detection (IM-DD)* systems where 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. As a first approximation, the noise in such systems can be assumed Gaussian, and the inputs are typically subject to both peak- and average-power constraints. The peak power is limited mainly due to technical limitations of the used components. The average-power constraint is imposed by battery limitations and by safety considerations. Notice that, unlike in radio communication, the average-power constraint applies directly to the transmit signal and not to its second moment, the reason being that the power of the transmit signal is proportional to the optical intensity, which directly relates to the transmit signal.

The main interest of this article is the *capacity* of such free-space optical intensity channels. For the *single-input and single-output (SISO)* scenario, where the transmitter employs a single transmit LED or LD, and the receiver a single photodetector, bounds on the capacity were presented in [4]–

[10]. They coincide asymptotically at high and low signal-to-noise ratio (SNR) [6]. 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, similar capacity bounds were presented in [11], [12] for the special case where the channel matrix has full column rank.

The current work considers the *multiple-input and single-output (MISO)* channel. The asymptotic high-SNR capacity of this channel was derived in [13]. The current work has the following contributions.

- The capacity of the MISO channel is expressed as the capacity of a SISO channel subject to complicated power constraints (Lemma 1).
- The asymptotic low-SNR capacity slope is derived for the MISO channel (Theorem 3).
- For almost all channel gains, we show that the asymptotic low-SNR slope is achieved by an input distribution that assigns either two or three point masses on the projection of the inputs on the channel-gain vector (Lemma 2). One of these point masses corresponds to setting all input antennas to 0. Each of the other (one or two) point masses corresponds to setting, for some $k \in \{1, \ldots, n_T\}$, the k strongest antennas to the maximum allowed peak power A and the remaining $n_T - k$ weaker antennas to 0.
- New upper bounds on the MISO capacity are presented for finite SNRs (Proposition 4 and Theorem 7). In the low-SNR regime, the first upper bound numerically matches the lower bounds that are obtained by choosing inputs as described above. The second upper bound improves on both the first upper bound and the upper bound in [13] in the moderate-SNR regime.

A. Channel Model

We consider a communication link where the transmitter is equipped with n_T LEDs (or LDs), $n_T \geq 2$, 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 $Z \sim \mathcal{N}(0, \sigma^2)$ 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:

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

$$\sum_{k=1}^{n_\mathrm{T}} \mathsf{E}[X_k] \leq \mathsf{E}, \tag{4b}$$

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

$$\alpha \triangleq \frac{\mathsf{E}}{\mathsf{A}}. \tag{5}$$

Throughout this paper we assume $\alpha < \frac{n_\mathrm{T}}{2}$. For $\alpha \geq \frac{n_\mathrm{T}}{2}$ the average-power constraint is not active and the capacity is the same as with only a peak-power constraint [14].

The capacity of the channel (1) is [15]

$$C_{\mathbf{h}^\mathsf{T}, \sigma^2}(\mathsf{A}, \mathsf{E}) = \sup_{Q_{\mathbf{X}}} I(\mathbf{X}; Y) \tag{6}$$

where the supremum is over all laws $Q_{\mathbf{X}}$ on $\mathbf{X} = (X_1, \ldots, X_{n_\mathrm{T}})$ satisfying (2), (4a), and (4b). It can be represented in an alternative form using the following definitions. Let

$$s_0 \triangleq 0 \tag{7a}$$

$$s_k \triangleq \sum_{k'=1}^{k} h_{k'}, \quad k \in \{1, \ldots, n_\mathrm{T}\}, \tag{7b}$$

and

$$\bar{X} \triangleq \mathbf{h}^\mathsf{T}\mathbf{X} = \sum_{k=1}^{n_\mathrm{T}} h_k X_k. \tag{8}$$

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

$$\bar{X} \in [\mathsf{A}s_{k-1}, \mathsf{A}s_k) \implies U = k, \tag{9}$$

and $U = n_\mathrm{T}$ if $\bar{X} = \mathsf{A}s_{n_\mathrm{T}}$. Let $\mathbf{p} = (p_1, \ldots, p_{n_\mathrm{T}})$ denote the probability vector of U:

$$p_k \triangleq \Pr[U = k], \quad k \in \{1, \ldots, n_\mathrm{T}\}. \tag{10}$$

Lemma 1: The MISO capacity satisfies

$$C_{\mathbf{h}^\mathsf{T}, \sigma^2}(\mathsf{A}, \alpha\mathsf{A}) = \max_{Q_{\bar{X}}} I(\bar{X}; Y), \tag{11}$$

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

$$\Pr[\bar{X} > s_{n_\mathrm{T}}\mathsf{A}] = 0 \tag{12a}$$

and

$$\sum_{k=1}^{n_\mathrm{T}} p_k \left(\frac{\mathsf{E}[\bar{X} \,|\, U = k] - \mathsf{A}s_{k-1}}{h_k} + (k-1)\mathsf{A} \right) \leq \alpha\mathsf{A}. \tag{12b}$$

Proof: The proof follows mainly by the Markov chain $\mathbf{X} \multimap \bar{X} \multimap Y$ and by noting that the ordering in (3) implies that the most energy-efficient way of signaling is to set $X_{k+1} = \cdots = X_{n_\mathrm{T}} = 0$ whenever $X_k = 0$ and to set $X_1 = \cdots = X_{k-1} = \mathsf{A}$ whenever $X_k > 0$. Under such signaling, constraints (12) are equivalent to constraints (4). ∎

II. MAIN RESULTS

A. Asymptotic Low-SNR Capacity

The asymptotic low-SNR capacity of the MISO channel is characterized by the maximum variance of $\bar{X}$ under constraints (12), i.e., by

$$V_{\max}(\mathsf{A}, \alpha\mathsf{A}) \triangleq \max_{Q_{\bar{X}}} \mathsf{E}\left[(\bar{X} - \mathsf{E}[\bar{X}])^2 \right], \tag{13}$$

where the maximization is over all distributions on $\bar{X} \in \mathbb{R}_0^+$ satisfying (12).

Lemma 2:

1) The maximum variance $V_{\max}(\mathsf{A}, \alpha\mathsf{A})$ can be achieved by restricting $Q_{\bar{X}}$ to the support set

$$\{0, s_1\mathsf{A}, s_2\mathsf{A}, \ldots, s_{n_\mathrm{T}}\mathsf{A}\}. \tag{14}$$

2) The maximum variance $V_{\max}(\mathsf{A}, \alpha\mathsf{A})$ satisfies

$$V_{\max}(\mathsf{A}, \alpha\mathsf{A}) = \mathsf{A}^2\gamma \tag{15}$$

where

$$\gamma \triangleq \max_{\substack{q_1, \ldots, q_{n_\mathrm{T}} \geq 0: \\ \sum_{k=1}^{n_\mathrm{T}} q_k \leq 1 \\ \sum_{k=1}^{n_\mathrm{T}} k \cdot q_k \leq \alpha}} \left\{ \sum_{k=1}^{n_\mathrm{T}} s_k^2 q_k - \left(\sum_{k=1}^{n_\mathrm{T}} s_k q_k \right)^2 \right\}. \tag{16}$$

3) A solution $\mathbf{q}^* = (q_1^*, \ldots, q_{n_\mathrm{T}}^*)$ to the optimization problem (16) satisfies $\sum_{k=1}^{n_\mathrm{T}} q_k < 1$ with strict inequality and $\sum_{k=1}^{n_\mathrm{T}} k \cdot q_k = \alpha$ with equality. (It is assumed that $\alpha \leq \frac{n_\mathrm{T}}{2}$.)

4) Whenever

$$\mathrm{rank} \begin{pmatrix} 1 & \frac{1}{s_1} & s_1 \\ 1 & \frac{2}{s_2} & s_2 \\ \vdots & \vdots & \vdots \\ 1 & \frac{n_\mathrm{T}}{s_{n_\mathrm{T}}} & s_{n_\mathrm{T}} \end{pmatrix} = 3, \tag{17}$$

the solution $\mathbf{q}^*$ to (16) has at most two nonzero elements, i.e., under condition (17), the maximum variance

V_{max} is achieved by an $\bar{X}^*$ with positive probability masses at 0 and at most two points from the set $\{s_1 A, \ldots, s_{n_T} A\}$.

Proof: See Appendix A. ∎

For many examples, the optimizing $\mathbf{q}^*$ has only a single positive entry, and thus V_{max} is achieved by an $\bar{X}^*$ that has only two point masses (one of them at 0). Table I (on top of the next page) presents some examples of maximum variances V_{max} and the probability mass functions of $\bar{X}^*$ achieving V_{max}.

In the low-SNR asymptotic regime, the capacity is determined by $V_{max}(A, \alpha A)$. According to the above lemma and the achievability proof of the following theorem, it is achieved by an input $\mathbf{X}$ that puts only 2 or 3 mass points on $\bar{X}$. One of these mass points being 0 and the other one or two mass points are elements of $\{s_1 A, s_2 A, \ldots, s_{n_T} A\}$.

Theorem 3 (Low-SNR Asymptotics): The low-SNR asymptotic capacity slope is

$$\lim_{A \downarrow 0} \frac{C_{\mathbf{h}^\mathsf{T}, \sigma^2}(A, \alpha A)}{A^2/\sigma^2} = \frac{\gamma}{2}, \tag{18}$$

where γ is defined in (16).

Proof: The converse follows immediately from a finite-SNR upper bound that we prove later; see Proposition 4. Achievability follows from [16, Thm. 2], which states that

$$C_{\mathbf{h}^\mathsf{T}, \sigma^2}(A, \alpha A) \geq \frac{V_{max}(A, \alpha A)}{2\sigma^2} + o(A^2), \tag{19}$$

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

$$\lim_{A \downarrow 0} \frac{o(A^2)}{A^2} = 0. \tag{20}$$

Note that the MISO channel under consideration in this paper satisfies the technical conditions A–F in [16]. ∎

Example 1: Consider a two-LED MISO channel with channel gains $h_1 = 3$ and $h_2 = 1$. Figure 1 plots the low-SNR slope of its capacity $\gamma/2$ as a function of the parameter α. We notice that the low-SNR slope $\gamma/2$ is strictly increasing for all values of $\alpha < \frac{n_T}{2}$. Interestingly, this strict monotonicity does not hold for the high-SNR asymptotic capacity, which saturates at a threshold level below $\frac{n_T}{2}$; see [14]. ◇

B. Finite-SNR Bounds

We present two analytic upper bounds and compare them to numerical lower bounds. As we will see, our first bound is very close to the actual capacity at low SNR. The second bound can improve on previous bounds in the regime of moderate SNR.

Proposition 4 (Upper Bound in Terms of V_{max}): The capacity is upper-bounded as

$$C_{\mathbf{h}^\mathsf{T}, \sigma^2}(A, \alpha A) \leq \frac{1}{2} \log\left(1 + \frac{V_{max}(A, \alpha A)}{\sigma^2}\right). \tag{21}$$

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

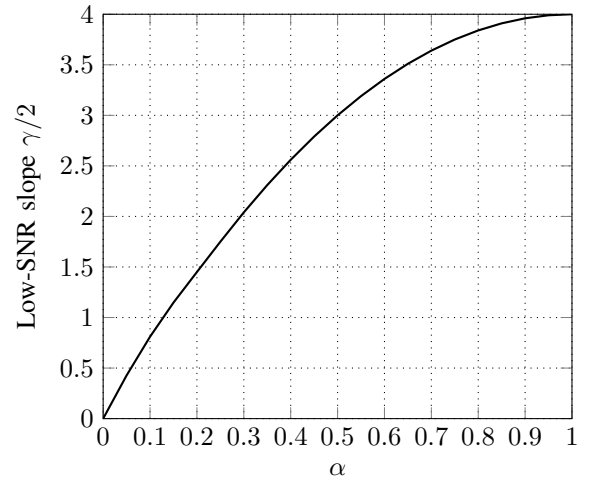$$h(Y) \leq \frac{1}{2} \log 2\pi e \big(V_{max}(A, \alpha A) + \sigma^2\big). \tag{22}$$

Fig. 1. The low-SNR asymptotic slope $\gamma/2$, see (18), 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 bound follows by subtracting $h(Z) = \frac{1}{2} \log 2\pi e \sigma^2$ from the above. ∎

The second upper bound is based on the following two propositions:

Proposition 5 (Upper Bound by SISO Capacity): Let $\mathbf{X}^*$ be a capacity-achieving input distribution for the MISO channel with gain vector $\mathbf{h}$. Define for all $k \in \{1, \ldots, n_T\}$:

$$p_k^* \triangleq \mathrm{Pr}_{\mathbf{X}^*}[U = k], \tag{23a}$$

$$\alpha_k^* \triangleq \mathsf{E}_{\mathbf{X}^*}\left[\frac{\bar{X} - s_{k-1} A}{h_k A} \,\bigg|\, U = k\right]. \tag{23b}$$

The capacity of the MISO channel is upper-bounded as

$$C_{\mathbf{h}^\mathsf{T}, \sigma^2}(A, \alpha A) \leq \mathrm{H}(\mathbf{p}^*) + \sum_{k=1}^{n_T} p_k^* C_{1, \sigma^2}(h_k A, \alpha_k^* h_k A), \tag{24}$$

and it holds that

$$\sum_{k=1}^{n_T} p_k^* \big(\alpha_k^* + (k-1)\big) \leq \alpha. \tag{25}$$

Proof: Omitted for brevity. See [14, Sec. 6, Eq. (88)]. ∎

Note that the SISO capacity C_{1, σ^2} under peak- and average-power constraints is itself unknown to date. Upper bounds on it were given, for example, in [6, Eq. (12)]. We present here a new upper bound.

Proposition 6 (Upper Bound on SISO Capacity): For any $\mu > 0$, the SISO capacity $C_{1, \sigma^2}(A, \alpha A)$ under peak-power constraint A and average-power constraint αA is upper-bounded as:

$$\begin{aligned} C_{1, \sigma^2}&(A, \alpha A) \\ &\leq \log\left(1 + \frac{A}{\sqrt{2\pi e}} \frac{1 - e^{-\mu}}{\mu}\right) \\ &\quad + \frac{1}{\sqrt{2\pi}} \frac{\mu}{A}\left(1 - e^{-\frac{A^2}{2}}\right) + \mu \alpha\left(1 - 2\mathcal{Q}\left(\frac{A}{2}\right)\right). \end{aligned} \tag{26}$$

TABLE I
MAXIMUM VARIANCE FOR DIFFERENT CHANNEL COEFFICIENTS

channel gains	α	V_{max}	$Q_{\bar{X}}$ achieving V_{max}
$\mathbf{h} = (3, 2.2, 0.1)$	0.9	$6.6924A^2$	$Q_{\bar{X}}(0) = 0.55$, $Q_{\bar{X}}(s_2A) = 0.45$
$\mathbf{h} = (3, 2.2, 1.1)$	0.7	$7.1001A^2$	$Q_{\bar{X}}(0) = 0.7667$, $Q_{\bar{X}}(s_3A) = 0.2333$
$\mathbf{h} = (3, 1.5, 0.3)$	0.95	$5.1158A^2$	$Q_{\bar{X}}(0) = 0.5907$, $Q_{\bar{X}}(s_2A) = 0.2780$, $Q_{\bar{X}}(s_3A) = 0.1313$

Proof: The proof is based on the duality-based upper bound [17] and on a choice of the output distribution inspired by [7] and [10]. See Appendix B for details. ∎

We can now state our new upper bound on the MISO capacity.

Theorem 7: The MISO capacity is upper-bounded as:

$$C_{\mathbf{h}^\intercal, \sigma^2}(A, \alpha A)$$

$$\leq \sup_{\mathbf{p}} \inf_{\mu > 0} \left\{ H(\mathbf{p}) + \sum_{k=1}^{n_T} p_k \log\left(1 + \frac{Ah_k}{\sqrt{2\pi e}} \cdot \frac{1 - e^{-\mu}}{\mu}\right) \right.$$

$$+ \frac{\mu}{\sqrt{2\pi}A} \sum_{k=1}^{n_T} \frac{p_k}{h_k}\left(1 - e^{-\frac{A^2 h_k^2}{2}}\right)$$

$$\left. + \mu\left(\alpha - \sum_{k=1}^{n_T} p_k(k-1)\right) \right\} \quad (27)$$

where the supremum is over vectors $\mathbf{p} = (p_1, \ldots, p_{n_T})$ satisfying

$$\sum_{k=1}^{n_T} p_k(k-1) \leq \alpha. \quad (28)$$

Proof: Combine Propositions 5 and 6, and use the bound $1 - 2Q\left(\frac{A}{2}\right) < 1$ and (25). ∎

Example 2: Consider the 3-LED MISO channel with gains $\mathbf{h} = (3, 2, 1.5)$. The asymptotic low-SNR capacity slope is $\gamma/2 = 5.07$ and is attained by choosing $\bar{X}$ equal to 0 with probability $q_0 = 0.6$ and equal to s_3A with probability $q_3 = 0.4$. Figure 2 shows lower and upper bounds on the channel capacity at different SNR values. The blue lower bound is obtained by numerically evaluating $I(\bar{X}; Y)$ for the choice of $\bar{X}$ that achieves the asymptotic low-SNR capacity. The magenta lower bound follows by numerically optimizing $I(\bar{X}; Y)$ over all choices of $\bar{X}$ that have positive probability on $\bar{X} = 0$ and on at most two point masses from $\{s_1A, \ldots, s_{n_T}A\}$. In the low SNR regime, these numerical lower bounds improve over the previous analytic lower bounds in [13] and are very close to the maximum-variance upper bound in Proposition 4. The gap between the best upper and lower bounds is larger in the moderate SNR regime. In this regime, the best upper bound (see the black line) is given in Theorem 7. ◇

Example 3: Consider now a SISO channel with channel gain $h = 1$. Figure 3 shows lower and upper bounds for this channel. It can be noted that the curve obtained by our new upper bound is very close to the three-point and four-point numerical lower bounds at moderate SNR. This indicates that
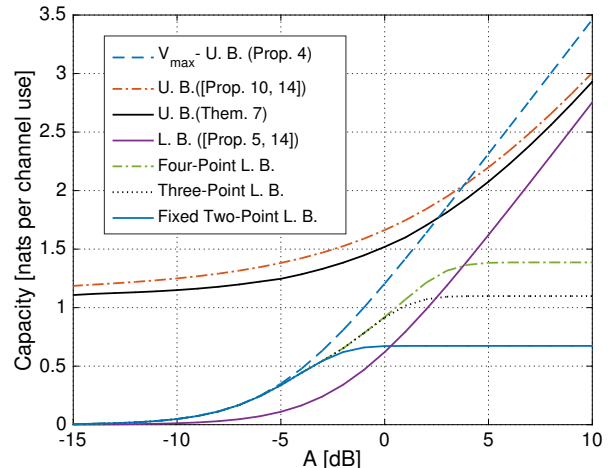

Fig. 2. Bounds on capacity of MISO channel with gains $h = (3, 2, 1.5)$ and average-to-peak power ratio $\alpha = 1.2$.

it gives a good approximation to the capacity and dominates other existing upper bounds in the moderate SNR regime. ◇
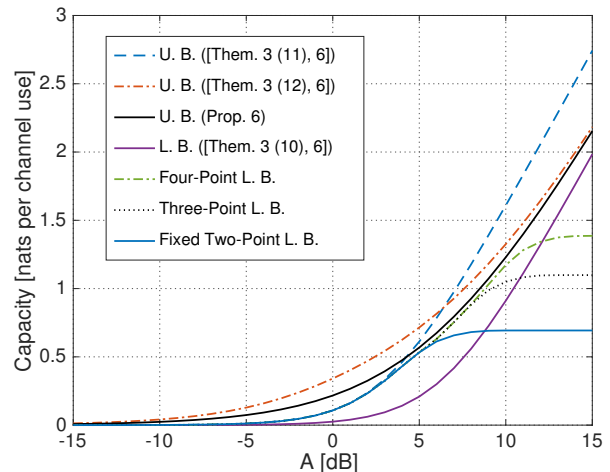

Fig. 3. Bounds on capacity of SISO channel with $h = 1$ and $\alpha = 0.4$.

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

$$\mathsf{E}\left[\left(\bar{X} - \mathsf{E}[\bar{X}]\right)^2\right]$$
$$= \sum_{k=1}^{n_\mathrm{T}} h_k^2 \, \mathsf{E}\left[(X_k - \mathsf{E}[X_k])^2\right]$$
$$+ \sum_{\substack{i,j=1 \\ i \neq j}}^{n_\mathrm{T}} h_i h_j \big(\mathsf{E}[X_i X_j] - \mathsf{E}[X_i]\,\mathsf{E}[X_j]\big). \qquad (29)$$

Let us fix the joint distribution on $(X_1, \ldots, X_{n_\mathrm{T}-1})$, and fix with probability one the conditional mean $\mathsf{E}[X_{n_\mathrm{T}} | X_1, \ldots, X_{n_\mathrm{T}-1}]$. These determine the consumed average input power, as well as every summand on the RHS of (29) except $\mathsf{E}\left[(X_{n_\mathrm{T}} - \mathsf{E}[X_{n_\mathrm{T}}])^2\right]$. For any choice above, the value of $\mathsf{E}\left[(X_{n_\mathrm{T}} - \mathsf{E}[X_{n_\mathrm{T}}])^2\right]$ is maximized by X_{n_T} taking value only in the set $\{0, \mathsf{A}\}$. We hence conclude that, to maximize the variance (29) subject to a constraint on average input power, it is optimal to restrict X_{n_T} to taking value only in $\{0, \mathsf{A}\}$. Repeating this argument for $X_{n_\mathrm{T}-1}, X_{n_\mathrm{T}-2}$, etc., we conclude that every X_k, $k = 1, \ldots, n_\mathrm{T}$, should take value only in $\{0, \mathsf{A}\}$.

Next, using the same arguments as the ones leading to [14, Lem. 4], it can be shown that it is optimal to consider the following joint distributions: for each $k \in \{0, \ldots, n_\mathrm{T}\}$, set with probability q_k

$$X_1 = \cdots = X_k = \mathsf{A} \quad \text{and} \quad X_{k+1} = \cdots = X_{n_\mathrm{T}} = 0. \quad (30)$$

Such a choice produces an $\bar{X}$ that takes value only in (14). 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

$$\sum_{k=0}^{n_\mathrm{T}} q_k k \leq \alpha. \qquad (31)$$

Part 2 of the lemma is then proven by noting that with the choice in (30), the variance of $\bar{X}$ is

$$\mathsf{E}\left[\left(\bar{X} - \mathsf{E}[\bar{X}]\right)^2\right] = \mathsf{E}\left[\bar{X}^2\right] - \left(\mathsf{E}[\bar{X}]\right)^2 \qquad (32)$$
$$= \sum_{k=1}^{n_\mathrm{T}} q_k \mathsf{A}^2 s_k^2 - \left(\sum_{k=1}^{n_\mathrm{T}} q_k \mathsf{A} s_k\right)^2. \qquad (33)$$

To prove Part 3, we notice that a solution $\mathbf{q}^* = (q_1^*, \ldots, q_{n_\mathrm{T}}^*)$ to (16) has to satisfy $\sum_{k=1}^{n_\mathrm{T}} q_k^* < 1$ because any $\bar{X}$ that achieves maximum variance $\mathsf{V}_{\max}$ puts nonzero probability on $\bar{X} = 0$. We further notice that the optimization problem (16) is convex because $s_1, \ldots, s_{n_\mathrm{T}} > 0$. Consider the Karush-Kuhn-Tucker (KKT) conditions for this optimization problem, and account for the fact that the linear constraint $\sum_{k=1}^{n_\mathrm{T}} q_k^* \leq 1$ is not active. The KKT conditions are then given by the following six (in)equalities:

$$-s_k(s_k - 2\mathbf{s}^\mathsf{T}\mathbf{q}) + \mu_0 \cdot k - \mu_k = 0, \quad k \in \{1, \ldots, n_\mathrm{T}\}, \quad (34a)$$

$$\mu_k q_k = 0, \quad k \in \{1, \ldots, n_\mathrm{T}\}, \quad (34b)$$

$$\mu_0 \left(\sum_{k=1}^{n_\mathrm{T}} k \cdot q_k - \alpha\right) = 0 \qquad (34c)$$

$$\mu_i \geq 0, \quad i \in \{0, 1, \ldots, n_\mathrm{T}\}, \qquad (34d)$$

$$\sum_{k=1}^{n_\mathrm{T}} k \cdot q_k \leq \alpha, \qquad (34e)$$

$$q_k \geq 0, \quad k \in \{1, \ldots, n_\mathrm{T}\}. \quad (34f)$$

A solution to these conditions always exists. Assume now that for such a solution the inequality constraint $\sum_{k=1}^{n_\mathrm{T}} k \cdot q_k^* \leq \alpha$ holds with strict inequality. The corresponding Lagrange multiplier μ_0 must then equal 0. Now, since for any $q_i^* > 0$, $\mu_i = 0$, (34a) then implies

$$s_i = 2\mathbf{s}^\mathsf{T}\mathbf{q}^*. \qquad (35)$$

But this can hold at most for a single $i \in \{1, \ldots, n_\mathrm{T}\}$ because all values of s_i are different ($h_{n_\mathrm{T}} > 0$). Moreover, it can hold only for $k = n_\mathrm{T}$. Indeed, if (35) holds for some $i < n_\mathrm{T}$, then the KKT condition (34a) cannot be satisfied for all $k > i$ because $\mu_k \geq 0$. To conclude, if the inequality constraint $\sum_{k=1}^{n_\mathrm{T}} k \cdot q_k^* \leq \alpha$ holds with strict inequality, then the optimal $\mathbf{q}^*$ satisfies $q_1^* = q_2^* = \ldots = q_{n_\mathrm{T}-1}^*$ and $q_{n_\mathrm{T}} = 1/2$. But this choice is only feasible for $\alpha = \frac{n_\mathrm{T}}{2}$, in which case the inequality constraint $\sum_{k=1}^{n_\mathrm{T}} k \cdot q_k^* \leq \alpha$ holds with equality. We have thus reached the desired contradiction, irrespective of the value of $\alpha \leq \frac{n_\mathrm{T}}{2}$.

We now prove Part 4 of the lemma by contradiction. Assume that for positive integers $k > i > j$ the optimal solution $\mathbf{q}^*$ satisfies $q_k^*, q_i^*, q_j^* > 0$. Then, by (34b), $\mu_k = \mu_j = \mu_i = 0$, and (34a) implies

$$\begin{pmatrix} 1 & \frac{j}{s_j} \\ 1 & \frac{i}{s_i} \\ 1 & \frac{k}{s_k} \end{pmatrix} \begin{pmatrix} 2\mathbf{s}^\mathsf{T}\mathbf{q} \\ \mu_0 \end{pmatrix} = \begin{pmatrix} s_j \\ s_i \\ s_k \end{pmatrix}. \qquad (36)$$

This is an overdetermined system of linear equations in the two "variables" $(2\mathbf{s}^\mathsf{T}\mathbf{q})$ and $\mu_0 \geq 0$, and it has a solution if, and only if,

$$\mathrm{rank}\begin{pmatrix} 1 & \frac{j}{s_j} \\ 1 & \frac{i}{s_i} \\ 1 & \frac{k}{s_k} \end{pmatrix} = \mathrm{rank}\begin{pmatrix} 1 & \frac{j}{s_j} & s_j \\ 1 & \frac{i}{s_i} & s_i \\ 1 & \frac{k}{s_k} & s_k \end{pmatrix} \qquad (37)$$

The proof follows then by noticing that this is only possible if the rank of both matrices in (37) is 2. In particular, the left-most matrix cannot have rank 1. And neither can the matrix in (17).

We evaluate the upper bound [17]

$$\mathsf{C} \leq \sup_Q \mathsf{E}_Q\left[\mathsf{D}\big(W(\cdot|X)\big\|f(\cdot)\big)\right] \qquad (38)$$

(where $D(\cdot\|\cdot)$ denotes the relative entropy or Kullback-Leibler divergence) for the test density

$$f(y) = \begin{cases} (1-\beta) \cdot f_1(y) & \text{if } y \in [0, A], \\ \beta \cdot f_2(y) & \text{otherwise,} \end{cases} \quad (39)$$

where $f_1(y)$ is a density over $[0, A]$,

$$f_1(y) = \frac{1}{A} \cdot \frac{\mu}{1 - e^{-\mu}} \cdot e^{-\frac{\mu y}{A}}, \quad y \in [0, A], \quad (40)$$

for some $\mu > 0$; $f_2(y)$ is a density over $\mathbb{R}\backslash[0, A]$,

$$f_2(y) = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} & \text{if } y < 0, \\ \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-A)^2}{2}} & y > A; \end{cases} \quad (41)$$

and $\beta \in (0, 1)$ will be specified later.

We notice that

$$-\int_{-\infty}^{0} W(y|x) \log f(y) \, dy$$

$$= -\int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-x)^2}{2}} \left(\log \frac{\beta}{\sqrt{2\pi}} - \frac{y^2}{2} \right) dy \quad (42)$$

$$= -\log \frac{\beta}{\sqrt{2\pi}} \mathcal{Q}(x) + \frac{1}{2} \left(\mathcal{Q}(x) + x^2 \mathcal{Q}(x) - x\phi(x) \right) \quad (43)$$

$$\leq -\left(\log \frac{\beta}{\sqrt{2\pi}} - \frac{1}{2} \right) \mathcal{Q}(x), \quad (44)$$

where $\phi(x) \triangleq \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, and (44) holds because $x > 0$ and $x\mathcal{Q}(x) < \phi(x)$.

Following the same arguments, we also obtain:

$$-\int_{A}^{\infty} W(y|x) \log f(y) \, dy \leq -\left(\log \frac{\beta}{\sqrt{2\pi}} - \frac{1}{2} \right) \mathcal{Q}(A - x). \quad (45)$$

Moreover,

$$-\int_{0}^{A} W(y|x) \log f(y) \, dy$$

$$= -\int_{0}^{A} \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-x)^2}{2}} \left(\log \frac{(1-\beta)}{A} \frac{\mu}{1 - e^{-\mu}} - \frac{\mu}{A} y \right) dy \quad (46)$$

$$= -\log \left(\frac{1-\beta}{A} \frac{\mu}{1 - e^{-\mu}} \right) \left(1 - \mathcal{Q}(x) - \mathcal{Q}(A - x) \right)$$
$$+ \frac{\mu}{A} \left(\phi(x) - \phi(A - x) + x \left(1 - \mathcal{Q}(x) - \mathcal{Q}(A - x) \right) \right) \quad (47)$$

$$\leq -\log \left(\frac{1-\beta}{A} \frac{\mu}{1 - e^{-\mu}} \right) \left(1 - \mathcal{Q}(x) - \mathcal{Q}(A - x) \right)$$
$$+ \frac{\mu}{A} \left(\phi(0) - \phi(A) + x \left(1 - 2\mathcal{Q}\left(\frac{A}{2} \right) \right) \right), \quad (48)$$

where (48) follows from the fact that $1 - \mathcal{Q}(x) - \mathcal{Q}(A - x)$ achieves the maximum value at $x = \frac{A}{2}$ when $x \in [0, A]$.

Combining (44), (45), and (48) with (38), and choosing

$$\beta = \frac{\mu\sqrt{2\pi e}}{A \cdot (1 - e^{-\mu}) + \mu\sqrt{2\pi e}}, \quad (49)$$

now yields the desired upper bound in the theorem.

REFERENCES

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

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

[3] Mohammad Ali Khalighi and Murat Uysal, "Survey on free space optical communication: A communication theory perspective," *IEEE Comm. Surveys & Tutorials*, vol. 16, no. 4, pp. 2231–2258, fourth quarter 2014.

[4] Joel G. Smith, "The information capacity of amplitude- and variance-constrained scalar Gaussian channels," *Inf. Contr.*, vol. 18, no. 3, pp. 203–219, Feb. 1971.

[5] Michèle 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.

[6] Amos Lapidoth, Stefan M. Moser, and Michèle A. Wigger, "On the capacity of free-space optical intensity channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4449–4461, Oct. 2009.

[7] Andrew 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.

[8] Ahmed A. Farid and Steve 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.

[9] Ahmed A. Farid and Steve 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] Andrew Thangaraj, Gerhard Kramer, and Georg 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.

[11] Stefan M. Moser, Michail Mylonakis, Ligong Wang, and Michèle Wigger, "Asymptotic capacity results for MIMO wireless optical communication," in *Proc. IEEE Int. Symp. Inf. Theory*, Aachen, Germany, Jun. 25–30, 2017, pp. 536–540.

[12] Anas Chaaban, Zouheir Rezki, and Mohamed-Slim Alouini, "MIMO intensity-modulation channels: Capacity bounds and high SNR characterization," in *Proc. IEEE Int. Conf. Commun.*, Paris, France, May 21–25, 2017, pp. 1–6.

[13] Stefan M. Moser, Ligong Wang, and Michèle Wigger, "Asymptotic high-SNR capacity of MISO optical intensity channels," in *Proc. IEEE Inf. Theory Workshop*, Kaohsiung, Taiwan, Nov. 6–10, 2017, pp. 86–90.

[14] Stefan M. Moser, Ligong Wang, and Michèle A. Wigger, "Capacity results on multiple-input single-output wireless optical channels," Oct. 2017, arXiv:1710.09177 [cs.IT]. Available: http://arxiv.org/abs/1710.09177

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

[16] Vyacheslav V. Prelov and Edward 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, Sept. 1993.

[17] Amos Lapidoth and Stefan 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.