# Asymptotic Capacity Results for MIMO Wireless Optical Communication

Stefan M. Moser[*†], Michail Mylonakis[*], Ligong Wang[‡], and Michèle Wigger[§]

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

*Abstract*—**This paper provides several asymptotic capacity results for the multiple-input multiple-output free-space optical intensity channel in the regime of high signal-to-noise ratio (SNR). For the case where the channel matrix has full column rank, the asymptotic capacity is derived assuming a peak-power constraint on each transmit antenna, or an average-power constraint on the total power across all transmit antennas, or both. For multiple-input and single-output channels, the asymptotic high-SNR capacity is derived when either only the total average power is constrained, or only the per-antenna peak power is constrained, or both but with the average-power constraint being sufficiently loose.**

## 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. For example, the work [4] 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 [5]–[8].

Of increasing practical interest are multiple-input and multiple-output (MIMO) free-space channels, where the transmitter is equipped with multiple LEDs or LDs, and the receiver with multiple photodetectors. Various code constructions for this setup have been proposed in [9]–[12]. First bounds on the capacity were presented in [8], [13], focusing on scenarios without crosstalk so the MIMO channel can be modeled through a diagonal channel matrix. While [8] constrained the *peak optical intensity* at each transmitting LED, [13] constrained the *average total optical intensity* over all LEDs.

In this paper we consider more general MIMO systems accounting for crosstalk between the different LEDs and the different photodetectors (implying that the channel matrix is not diagonal), and we impose both a peak-power constraint on each LED and an average-power constraint on all LEDs. We note that, for safety reasons and implementation considerations, practical systems are usually indeed subject to both constraints. We present the asymptotic high-SNR capacity of the MIMO optical intensity channel for the following cases:

- The channel matrix is of full column rank, i.e., its rank equals the number of transmit LEDs, and the inputs are subject to any peak- and average-power constraints;
- The channel is multiple-input and single-output (MISO), and the inputs are subject to only a peak-power constraint, or only an average-power constraint, or both constraints but with the average-power constraint being sufficiently loose.

We prove these asymptotic capacity results by providing upper and lower bounds on capacity. The upper bounds are obtained by means of the *duality capacity upper-bounding technique*; see, e.g., [4], [14]. The lower bounds are based on the Entropy Power Inequality. Some details in the derivations are omitted due to space limitations.

The rest of this paper is arranged as follows: Section II describes the channel model; Section III presents the results in the case where the channel matrix has full column rank; Section IV presents the results for MISO channels; and Section V concludes the paper with a few remarks.

## II. Channel Model

We consider a communication scenario where the transmitter is equipped with $n_T$ LEDs (or LDs) and the receiver with $n_R$ photodetectors. Each 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 an $n_R$-vector given by

$$\mathbf{Y} = \mathsf{H}\mathbf{x} + \mathbf{Z}, \tag{1}$$

where the $n_T$-vector $\mathbf{x}$ denotes the channel input, whose entries are proportional to the optical intensities of the corresponding LEDs, and are therefore nonnegative:

$$x^{(i)} \in \mathbb{R}_0^+, \quad i = 1, \ldots, n_T; \tag{2}$$

where the $n_R \times n_T$ matrix $\mathsf{H}$ is the constant channel mixing matrix with nonnegative entries:

$$h^{(i,j)} \in \mathbb{R}_0^+, \quad \forall (i,j) \in \{1, \ldots, n_R\} \times \{1, \ldots, n_T\}; \tag{3}$$

and where $\mathbf{Z}$ is a centered Gaussian noise vector with a given positive definite $n_R \times n_R$ covariance matrix $\mathsf{K}$:

$$\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathsf{K}). \tag{4}$$

Note that, in contrast to the input $\mathbf{x}$, the output $\mathbf{Y}$ can have negative components.

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

$$\Pr[X^{(i)} > \mathsf{A}] = 0, \quad \forall i \in \{1, \ldots, n_T\}, \tag{5}$$

$$\sum_{i=1}^{n_T} \mathsf{E}[X^{(i)}] \leq \mathsf{E}, \tag{6}$$

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 $\mathsf{A}$ describes the maximum power of each single LED, while $\mathsf{E}$ 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 $\alpha$:

$$\alpha \triangleq \frac{\mathsf{E}}{\mathsf{A}}, \tag{7}$$

where $0 < \alpha \leq n_T$. For $\alpha = n_T$ the average-power constraint is inactive in the sense that it is automatically satisfied whenever the peak-power constraint is satisfied. Thus, $\alpha = n_T$ corresponds to the case with only a peak-power constraint. On the other hand, $\alpha \ll 1$ corresponds to having a dominant average-power constraint and only a very weak peak-power constraint.

We denote the capacity of the channel (1) with allowed peak power $\mathsf{A}$ and allowed average power $\mathsf{E}$ by $\mathsf{C}_{\mathsf{H},\mathsf{K}}(\mathsf{A}, \mathsf{E})$. The capacity is given by [15]

$$\mathsf{C}_{\mathsf{H},\mathsf{K}}(\mathsf{A}, \mathsf{E}) = \sup_Q \mathrm{I}(Q, W) \tag{8}$$

where the supremum is over all laws $Q$ on $\mathbf{X}$ satisfying (2), (5), and (6). When only an average-power constraint is imposed, capacity is denoted by $\mathsf{C}_{\mathsf{H},\mathsf{K}}(\mathsf{E})$. It is given as in (8) except that the supremum is taken over all laws $Q$ on $\mathbf{X}$ satisfying (2) and (6).

Due to the symmetry of the channel law, we have the following basic property.

*Proposition 1:* If $\alpha \geq \frac{n_T}{2}$, then there exists an optimal input distribution $Q^*$ such that

$$\mathsf{E}_{Q^*}[X^{(i)}] = \frac{\mathsf{A}}{2}, \quad \forall i \in \{1, \ldots, n_T\}, \tag{9}$$

and

$$\mathsf{C}_{\mathsf{H},\mathsf{K}}(\mathsf{A}, \alpha\mathsf{A}) = \mathsf{C}_{\mathsf{H},\mathsf{K}}\left(\mathsf{A}, \frac{n_T}{2}\mathsf{A}\right). \tag{10}$$

Thus, for $\alpha \geq \frac{n_T}{2}$, the average-power constraint is inactive.

*Proof:* The proof of this proposition follows the proof of [4, Prop. 9] with all variables replaced by corresponding vectors. The details are omitted. ∎

## III. THE MIMO CHANNEL WITH FULL-COLUMN-RANK $\mathsf{H}$

In this section we derive the asymptotic high-SNR capacity of the channel (1) when the channel matrix $\mathsf{H}$ has full column rank.

### A. Asymptotic High-SNR Capacity

We start with the case where $\mathsf{H}$ is square.

*Theorem 2:* Let $\mathsf{H}$ be square (so $n_T = n_R = n$) and invertible. If $0 < \alpha < \frac{n}{2}$, then

$$\lim_{\mathsf{A} \to \infty} \{\mathsf{C}_{\mathsf{H},\mathsf{K}}(\mathsf{A}, \alpha\mathsf{A}) - n \log \mathsf{A}\}$$
$$= \frac{1}{2} \log \det(\mathsf{H}^\mathsf{T}\mathsf{K}^{-1}\mathsf{H}) - \frac{n}{2} \log 2\pi e$$
$$- n\mu^*\left(1 - \frac{\alpha}{n}\right) - n \log\left(1 - \frac{\alpha\mu^*}{n}\right) \tag{11}$$

where $\mu^*$ is the unique solution to

$$\frac{\alpha}{n} = \frac{1}{\mu^*} - \frac{e^{-\mu^*}}{1 - e^{-\mu^*}}. \tag{12}$$

If $\alpha \geq \frac{n}{2}$, then

$$\lim_{\mathsf{A} \to \infty} \{\mathsf{C}_{\mathsf{H},\mathsf{K}}(\mathsf{A}, \alpha\mathsf{A}) - n \log \mathsf{A}\}$$
$$= \frac{1}{2} \log \det(\mathsf{H}^\mathsf{T}\mathsf{K}^{-1}\mathsf{H}) - \frac{n}{2} \log 2\pi e. \tag{13}$$

If there is only an average-power constraint (6), then

$$\lim_{\mathsf{E} \to \infty} \{\mathsf{C}_{\mathsf{H},\mathsf{K}}(\mathsf{E}) - n \log \mathsf{E}\}$$
$$= \frac{1}{2} \log \det(\mathsf{H}^\mathsf{T}\mathsf{K}^{-1}\mathsf{H}) - \frac{n}{2} \log \frac{2\pi n^2}{e}. \tag{14}$$

*Proof:* See Section III-B. ∎

We next generalize Theorem 2 to the case where $\mathsf{H}$ is not necessarily square but has full column rank.

*Corollary 3:* If $\mathsf{H}$ has full column rank, i.e., if $n_R \geq n_T$ and $\mathrm{rank}(\mathsf{H}) = n_T$, all results of Theorem 2 still apply with $n$ replaced by $n_T$.

*Proof:* See Section III-C. ∎

*Remark 4:* For comparison, consider the SISO channel $Y = x + Z$ where $Z \sim \mathcal{N}(0, 1)$. Denote its capacity under constraints (5) and (6) by $\mathsf{C}_{1,1}(\mathsf{A}, \mathsf{E})$. Comparing Corollary 3 and the results in [4] we observe the following. For any $\alpha \geq 0$

$$\lim_{\mathsf{A} \to \infty} \left\{\mathsf{C}_{\mathsf{H},\mathsf{K}}(\mathsf{A}, \alpha\mathsf{A}) - n_T\mathsf{C}_{1,1}\left(\mathsf{A}, \frac{\alpha}{n_T}\mathsf{A}\right)\right\}$$
$$= \frac{1}{2} \log \det(\mathsf{H}^\mathsf{T}\mathsf{K}^{-1}\mathsf{H}). \tag{15}$$

Similarly, in the absence of a peak-power constraint,

$$\lim_{\mathsf{E}\to\infty}\left\{\mathsf{C}_{\mathsf{H},\mathsf{K}}(\mathsf{E}) - n_{\mathsf{T}}\mathsf{C}_{1,1}\left(\frac{\mathsf{E}}{n_{\mathsf{T}}}\right)\right\} = \frac{1}{2}\log\det(\mathsf{H}^{\mathsf{T}}\mathsf{K}^{-1}\mathsf{H}). \quad (16)$$

### B. Proof of Theorem 2

Since H is square and invertible,

$$\mathrm{I}(\mathbf{X};\mathsf{H}\mathbf{X}+\mathbf{Z}) = \mathrm{I}(\mathbf{X};\mathbf{X}+\mathsf{H}^{-1}\mathbf{Z}) = \mathrm{I}(\mathbf{X};\mathbf{X}+\mathbf{Z}') \quad (17)$$

where $\mathbf{Z}' \sim \mathcal{N}(\mathbf{0},\mathsf{H}^{-1}\mathsf{K}\mathsf{H}^{-\mathsf{T}})$.

We start with the case where $0 < \alpha < \frac{n}{2}$. We first derive an upper bound on capacity based on the following duality upper bound (see [4, Eq. (1)]). For any output distribution $R(\cdot)$ on the real line $\mathbb{R}$,

$$\mathsf{C}_{\mathsf{H},\mathsf{K}}(A, \alpha A) \le \sup_{Q} \mathsf{E}\left[\mathscr{D}\big(W(\cdot|\mathbf{X})\big\|R(\cdot)\big)\right], \quad (18)$$

where the supremum is over all input distributions $Q$ satisfying both constraints (5) and (6). We pick the output distribution

$$R(\mathbf{y}) = \prod_{i=1}^{n} R_i\big(y^{(i)}\big) \quad (19)$$

with each $R_i(\cdot)$ chosen as in [4, Eq. (47)]:

$$R_i(y) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{y^2}{2\sigma_i^2}} & y < -\delta_i, \\ \frac{1}{A} \cdot \frac{\mu^*\left(1 - 2\,\mathcal{Q}\left(\frac{\delta_i}{\sigma_i}\right)\right)}{e^{\frac{\mu^*\delta_i}{A}} - e^{-\mu^*\left(1+\frac{\delta_i}{A}\right)}} e^{-\frac{\mu^* y}{A}} & -\delta_i \le y \le A+\delta_i, \\ \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(y-A)^2}{2\sigma_i^2}} & y > A+\delta_i, \end{cases} \quad (20)$$

where

$$\sigma_i \triangleq \sqrt{\mathsf{Var}\big[Z'^{(i)}\big]}, \quad (21)$$

which is given by the square root of the $i$th diagonal entry in $\mathsf{H}^{-1}\mathsf{K}\mathsf{H}^{-\mathsf{T}}$; and

$$\delta_i \triangleq \sigma_i \cdot \delta, \quad i \in \{1,\dots,n\}, \quad (22)$$

with

$$\delta \triangleq \log(1+A). \quad (23)$$

Here, $\mathcal{Q}(\cdot)$ denotes the complementary cumulative distribution function of the standard Gaussian.

Note that for each $i$, $R_i(\cdot)$ consists of a truncated exponential distribution with Gaussian tails on both sides. The truncated exponential distribution is well-known to maximize differential entropy under a peak and a first-moment constraint [16, Ch. 12].

Since the chosen $R(\cdot)$ has independent components, the relative entropy decouples:

$$\mathscr{D}\big(W(\cdot|\mathbf{x})\big\|R(\cdot)\big) = \sum_{i=1}^{n}\mathscr{D}\big(W^{(i)}(\cdot|x^{(i)})\big\|R_i(\cdot)\big), \quad (24)$$

and

$$\mathsf{C}_{\mathsf{H},\mathsf{K}}(A, \alpha A) \le \sum_{i=1}^{n}\sup_{Q}\mathsf{E}\left[\mathscr{D}\big(W^{(i)}(\cdot|X^{(i)})\big\|R_i(\cdot)\big)\right]. \quad (25)$$

By applying the same derivations as in [4, App. B.B] to each summand on the right-hand side (RHS) of (25), the following upper bound can be established:

$$\mathsf{C}_{\mathsf{H},\mathsf{K}}(A, \alpha A)$$
$$\le n\mathcal{Q}(\delta) + \frac{n\delta}{\sqrt{2\pi}}e^{-\frac{\delta^2}{2}} - \frac{1}{2}\sum_{i=1}^{n}\log\sigma_i^2 - \frac{n}{2}$$
$$+ \frac{1}{2}\log\det(\mathsf{H}^{\mathsf{T}}\mathsf{K}^{-1}\mathsf{H}) + \sum_{i=1}^{n}\left(1 - 2\mathcal{Q}\left(\delta + \frac{A}{2\sigma_i}\right)\right)$$
$$\cdot\log\frac{A\big(e^{\mu^*\delta\frac{\sigma_i}{A}} - e^{-\mu^*(1+\delta\frac{\sigma_i}{A})}\big)}{\sigma_i\sqrt{2\pi}\mu^*(1 - 2\,\mathcal{Q}(\delta))}$$
$$+ \mu^*\alpha\left(1 - 2\,\mathcal{Q}\left(\delta + \frac{A}{2\sigma_{\min}}\right)\right)$$
$$+ \sum_{i=1}^{n}\frac{\mu^*}{\sqrt{2\pi}}\frac{\sigma_i}{A}\left(e^{-\frac{\delta^2}{2}} - e^{-\frac{\left(\frac{A}{\sigma_i}+\delta\right)^2}{2}}\right), \quad (26)$$

where $\sigma_{\min} \triangleq \min\{\sigma_1,\dots,\sigma_n\}$. Taking $A \to \infty$ in the above proves that the left-hand side (LHS) of (11) is upper-bounded by its RHS.

To show that the LHS of (11) is also lower-bounded by its RHS, we pick the following distribution for $\mathbf{X}$

$$Q(\mathbf{x}) = \prod_{i=1}^{n}\frac{1}{A}\frac{\mu^*}{1 - e^{-\mu^*}}e^{-\frac{\mu^*}{A}x^{(i)}}\cdot\mathbf{I}\big\{0 \le x^{(i)} \le A\big\}, \quad (27)$$

with $\mathbf{I}\{\cdot\}$ denoting the indicator function. Using the Entropy Power Inequality (EPI) [16, Thm. 17.7.3] we have

$$\mathsf{C}_{\mathsf{H},\mathsf{K}}(A, \alpha A) \ge \mathrm{I}(\mathbf{X};\mathbf{X}+\mathbf{Z}') \quad (28)$$
$$= \mathrm{h}(\mathbf{X}+\mathbf{Z}') - \mathrm{h}(\mathbf{Z}') \quad (29)$$
$$\ge \frac{n}{2}\log\left(e^{\frac{2}{n}\mathrm{h}(\mathbf{X})} + e^{\frac{2}{n}\mathrm{h}(\mathbf{Z}')}\right) - \mathrm{h}(\mathbf{Z}'). \quad (30)$$

Evaluating the differential entropy terms $\mathrm{h}(\mathbf{X})$ and $\mathrm{h}(\mathbf{Z}')$ and letting $A$ tend to infinity establish the desired asymptotic lower bound.

The statements for the case $\alpha \ge \frac{n}{2}$ and for the case with only an average-power constraint can be shown following similar lines as above, if one adapts the choices of $R_i(\cdot)$ in (20) and $Q(\cdot)$ in (27).

For $\alpha \ge \frac{n}{2}$, we choose each $R_i(\cdot)$ as [4, Eq. (64)]

$$R_i(y) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{y^2}{2\sigma_i^2}} & y < -\delta_i, \\ \frac{1 - 2\,\mathcal{Q}\left(\frac{\delta_i}{\sigma_i}\right)}{A + 2\delta_i} & -\delta_i \le y \le A+\delta_i, \\ \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(y-A)^2}{2\sigma_i^2}} & y > A+\delta_i \end{cases} \quad (31)$$

with $\sigma_i$ and $\delta_i$ given by (21)–(23) ($R_i(\cdot)$ consists of a uniform distribution with Gaussian tails); and choose $Q(\cdot)$ as

$$Q(\mathbf{x}) = \prod_{i=1}^{n}\frac{1}{A}\cdot\mathbf{I}\big\{0 \le x^{(i)} \le A\big\}. \quad (32)$$

For the case with only an average-power constraint, we choose $R_i(\cdot)$ as [4, Eq. (70)]

$$R_i(y) = \begin{cases} \dfrac{1}{\beta\, e^{-\frac{\delta_i^2}{2\sigma_i^2}} + \sqrt{2\pi}\sigma_i\, \mathcal{Q}\left(\frac{\delta_i}{\sigma_i}\right)}\, e^{-\frac{y^2}{2\sigma^2}} & y < -\delta_i, \\[3ex] \dfrac{1}{\beta\, e^{-\frac{\delta_i^2}{2\sigma_i^2}} + \sqrt{2\pi}\sigma_i\, \mathcal{Q}\left(\frac{\delta_i}{\sigma_i}\right)}\, e^{-\frac{\delta_i^2}{2\sigma_i^2}}\, e^{-\frac{y+\delta_i}{\beta}} & y \geq -\delta_i \end{cases}$$
$$(33)$$

with $\sigma_i$ given by (21), $\delta_i \triangleq \sigma_i\sqrt{\log \mathsf{E}}$, and $\beta \triangleq \frac{\mathsf{E}}{n}$ ($R_i(\cdot)$ consists of an exponential distribution with a Gaussian tail); and choose the distribution of $\mathbf{X}$ as

$$Q(\mathbf{x}) = \prod_{i=1}^{n} \frac{n}{\mathsf{E}}\, e^{-\frac{n}{\mathsf{E}}x^{(i)}} \cdot \mathbf{I}\{x^{(i)} \geq 0\}. \qquad (34)$$

∎

### C. Proof of Corollary 3

Since $\mathsf{K}$ is positive definite, it can be written as $\mathsf{K} = \mathsf{S}^\mathsf{T}\mathsf{S}$ and

$$\mathrm{I}(\mathbf{X}; \mathsf{H}\mathbf{X} + \mathbf{Z}) = \mathrm{I}(\mathbf{X}; \mathsf{S}^{-\mathsf{T}}\mathsf{H}\mathbf{X} + \mathsf{S}^{-\mathsf{T}}\mathbf{Z}) \qquad (35)$$
$$= \mathrm{I}(\mathbf{X}; \tilde{\mathsf{H}}\mathbf{X} + \tilde{\mathbf{Z}}) \qquad (36)$$

with $\tilde{\mathbf{Z}} \sim \mathcal{N}(\mathbf{0}, \mathsf{I}_{n_\mathrm{R}})$. Let the singular value decomposition (SVD) of $\tilde{\mathsf{H}} = \mathsf{S}^{-\mathsf{T}}\mathsf{H}$ be

$$\tilde{\mathsf{H}} = \mathsf{U}\mathsf{\Sigma}\mathsf{V}, \qquad (37)$$

where $\mathsf{U}$ is an $n_\mathrm{R} \times n_\mathrm{R}$ unitary matrix, $\mathsf{V}$ is an $n_\mathrm{T} \times n_\mathrm{T}$ unitary matrix, and $\mathsf{\Sigma}$ is an $n_\mathrm{R} \times n_\mathrm{T}$ diagonal matrix of the form

$$\mathsf{\Sigma} = \begin{pmatrix} \mathsf{\Sigma}_{n_\mathrm{T}} \\ 0 \end{pmatrix} \qquad (38)$$

with $\mathsf{\Sigma}_{n_\mathrm{T}}$ being an $n_\mathrm{T} \times n_\mathrm{T}$ diagonal matrix. Because $\mathsf{H}$ has full column rank and $\mathsf{S}$ is invertible, $\tilde{\mathsf{H}}$ also has full column rank and the diagonal elements of $\mathsf{\Sigma}_{n_\mathrm{T}}$ are all positive. Thus,

$$\mathrm{I}(\mathbf{X}; \mathsf{H}\mathbf{X} + \mathbf{Z})$$
$$= \mathrm{I}(\mathbf{X}; \mathsf{\Sigma}\mathsf{V}\mathbf{X} + \mathsf{U}^\mathsf{T}\tilde{\mathbf{Z}}) \qquad (39)$$
$$= \mathrm{I}(\mathbf{X}; \mathsf{\Sigma}\mathsf{V}\mathbf{X} + \tilde{\mathbf{Z}}) \qquad (40)$$
$$= \mathrm{I}\big(\mathbf{X}; \mathsf{\Sigma}_{n_\mathrm{T}}\mathsf{V}\mathbf{X} + \tilde{\mathbf{Z}}_{n_\mathrm{T}}, \tilde{Z}^{(n_\mathrm{T}+1)}, \dots, \tilde{Z}^{(n_\mathrm{R})}\big) \qquad (41)$$
$$= \mathrm{I}(\mathbf{X}; \mathsf{\Sigma}_{n_\mathrm{T}}\mathsf{V}\mathbf{X} + \tilde{\mathbf{Z}}_{n_\mathrm{T}}) \qquad (42)$$
$$= \mathrm{I}(\mathbf{X}; \mathbf{X} + \mathsf{V}^\mathsf{T}\mathsf{\Sigma}_{n_\mathrm{T}}^{-1}\tilde{\mathbf{Z}}_{n_\mathrm{T}}) \qquad (43)$$
$$= \mathrm{I}(\mathbf{X}; \mathbf{X} + \mathbf{Z}'_{n_\mathrm{T}}). \qquad (44)$$

Here, (40) holds because both $\tilde{\mathbf{Z}}$ and $\mathsf{U}^\mathsf{T}\tilde{\mathbf{Z}}$ have the same distribution $\mathcal{N}(\mathbf{0}, \mathsf{I}_{n_\mathrm{R}})$; (41) follows from (38); and (42) follows by dropping the noise terms $\tilde{Z}^{(n_\mathrm{T}+1)}, \dots, \tilde{Z}^{(n_\mathrm{R})}$, which are independent of $\tilde{\mathbf{Z}}_{n_\mathrm{T}}$ and $\mathbf{X}$. In (44) we have

$$\mathbf{Z}'_{n_\mathrm{T}} \sim \mathcal{N}\big(\mathbf{0}, \mathsf{H}^{-1}\mathsf{K}\mathsf{H}^{-\mathsf{T}}\big). \qquad (45)$$

Hence, (44) is identical to (17) with $n = n_\mathrm{T}$. The remainder of the proof is identical to that of Theorem 2. ∎

## IV. THE MISO CHANNEL

In this section we consider the MISO channel where $n_\mathrm{R} = 1$ and solve its asymptotic high-SNR capacity in some special cases. For the MISO channel, the channel matrix $\mathsf{H}$ becomes a row vector $\mathbf{h}^\mathsf{T}$ with nonnegative entries, while the covariance matrix $\mathsf{K}$ for the additive noise is reduced to a scalar $\sigma^2$. Thus (1) becomes

$$Y = \mathbf{h}^\mathsf{T}\mathbf{x} + Z, \qquad (46)$$

where $Z \sim \mathcal{N}(0, \sigma^2)$. The input $\mathbf{X}$ is again subject to the constraints (5) and (6). Denote $\bar{X} \triangleq \mathbf{h}^\mathsf{T}\mathbf{X}$, then, because $\mathbf{X} \multimap \bar{X} \multimap Y$ form a Markov chain, and because $\bar{X}$ is a function of $\mathbf{X}$, we have

$$\mathrm{I}(\mathbf{X}; Y) = \mathrm{I}(\bar{X}; Y). \qquad (47)$$

Hence the MISO channel (46) is equivalent to the SISO channel whose input is $\bar{X}$, while the constraints (5) and (6) on $\mathbf{X}$ are transformed to a set of admissible distributions for $\bar{X}$. For example:

- When $\mathbf{X}$ is only subject to a peak-power constraint $\mathsf{A}$ (and no average-power constraint), then $\bar{X}$ is only subject to a peak-power constraint $h_\mathrm{sum}\mathsf{A}$, where

$$h_\mathrm{sum} \triangleq \sum_{i=1}^{n_\mathrm{T}} h^{(i)}. \qquad (48)$$

- When $\mathbf{X}$ is only subject to an average-power constraint $\mathsf{E}$ (and no peak-power constraint), then $\bar{X}$ is only subject to an average-power constraint $h_\mathrm{max}\mathsf{E}$, where

$$h_\mathrm{max} \triangleq \max_{i \in \{1,\dots,n_\mathrm{T}\}} h^{(i)}. \qquad (49)$$

Formally, we have the following two propositions.

*Proposition 5:* When $\alpha \geq \frac{n_\mathrm{T}}{2}$,

$$\mathsf{C}_{\mathbf{h}^\mathsf{T}, \sigma^2}(\mathsf{A}, \alpha\mathsf{A}) = \mathsf{C}_{1, \sigma^2}\left(h_\mathrm{sum}\mathsf{A}, \frac{h_\mathrm{sum}\mathsf{A}}{2}\right), \qquad (50)$$

where $h_\mathrm{sum}$ is defined in (48).

*Proof:* When $\mathbf{X}$ satisfies the peak-power constraint (5), $\bar{X}$ must satisfy $\bar{X} \leq h_\mathrm{sum}\mathsf{A}$ with probability one. Hence $\mathsf{C}_{\mathbf{h}^\mathsf{T}, \sigma^2}(\mathsf{A}, \alpha\mathsf{A})$ cannot exceed the capacity of the SISO channel with allowed peak power $h_\mathrm{sum}\mathsf{A}$. By [4, Prop. 9], for a SISO channel with allowed peak power $h_\mathrm{sum}\mathsf{A}$, adding an average-power constraint of $\frac{h_\mathrm{sum}\mathsf{A}}{2}$ does not affect its capacity. We hence know that the LHS of (50) is upper-bounded by the RHS.

For the reverse direction, consider any target distribution on $\bar{X}$ satisfying peak-power constraint $h_\mathrm{sum}\mathsf{A}$ and average-power constraint $\frac{1}{2}h_\mathrm{sum}\mathsf{A}$. 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 $\mathsf{A}$ and $\alpha\mathsf{A}$, respectively. To this end, we let the transmitter send the same signal on all LEDs:

$$X^{(i)} = \frac{\bar{X}}{h_\mathrm{sum}}, \quad i \in \{1, \dots, n_\mathrm{T}\}. \qquad (51)$$

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

*Proposition 6:* Without a peak-power constraint,

$$C_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathsf{E}) = C_{1,\sigma^2}(h_{\max}\mathsf{E}), \tag{52}$$

where $h_{\max}$ is defined in (49).

*Proof:* When $\mathbf{X}$ satisfies (6), we have

$$\mathsf{E}\big[\bar{X}\big] = \sum_{i=1}^{n_\mathsf{T}} h^{(i)}\,\mathsf{E}\big[X^{(i)}\big] \le h_{\max}\mathsf{E}, \tag{53}$$

so $C_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathsf{E}) \le C_{1,\sigma^2}(h_{\max}\mathsf{E})$. For the reverse direction, to achieve any target distribution on $\bar{X}$ satisfying $\mathsf{E}\big[\bar{X}\big] \le h_{\max}\mathsf{E}$, the transmitter can let the LED corresponding to $h_{\max}$ send $\bar{X}/h_{\max}$, and let all the other LEDs send zero. ∎

Using Propositions 5 and 6 together with the asymptotic capacity results for SISO channels [4, Cor. 6 and Prop. 8], we immediately obtain the following.

*Corollary 7:* When $\alpha \ge \frac{n_\mathsf{T}}{2}$,

$$\lim_{A \to \infty} \big\{ C_{\mathbf{h}^\mathsf{T},\sigma^2}(A, \alpha A) - \log A \big\} = \frac{1}{2}\log\frac{h_{\mathrm{sum}}^2}{2\pi e\sigma^2}. \tag{54}$$

Without a peak-power constraint,

$$\lim_{\mathsf{E} \to \infty} \big\{ C_{\mathbf{h}^\mathsf{T},\sigma^2}(\mathsf{E}) - \log \mathsf{E} \big\} = \frac{1}{2}\log\frac{eh_{\max}^2}{2\pi\sigma^2}. \tag{55}$$

The last result of this section shows that (54) holds under a less restrictive condition than the one in Corollary 7.

*Theorem 8:* Let the transmit LEDs be ordered such that

$$h^{(1)} \ge h^{(2)} \ge \cdots \ge h^{(n_\mathsf{T})}. \tag{56}$$

Limit (54) holds whenever

$$\alpha \ge \alpha_{\mathrm{th}} \triangleq \frac{1}{h_{\mathrm{sum}}} \sum_{i=1}^{n_\mathsf{T}} \left(i - \frac{1}{2}\right) h^{(i)}. \tag{57}$$

Note that the RHS of (57) is less than or equal to $\frac{n_\mathsf{T}}{2}$, where equality holds if and only if all $h^{(i)}$ are the same. For example, when $n_\mathsf{T} = 2$, it is given by

$$\alpha_{\mathrm{th}} = \frac{1}{2} + \frac{h^{(2)}}{h^{(1)} + h^{(2)}}. \tag{58}$$

*Proof of Theorem 8:* A tight upper bound on capacity follows immediately from Corollary 7. Hence we only need to derive a matching lower bound. To this end, we use [4, Eq. (40)], which says

$$\mathsf{I}(\bar{X}; Y) \ge \frac{1}{2}\log\left(1 + \frac{e^{2\,\mathsf{h}(\bar{X})}}{2\pi e\sigma^2}\right). \tag{59}$$

We show that, under (57), one can make $\bar{X}$ uniformly distributed on the interval $[0, h_{\mathrm{sum}}A]$, which gives $\mathsf{h}(\bar{X}) = \log(h_{\mathrm{sum}}A)$, yielding the desired lower bound. The uniform distribution is achieved as follows. For every $i \in \{1, \ldots, n_\mathsf{T}\}$, choose with probability $p_i \triangleq h^{(i)}/h_{\mathrm{sum}}$ the input signals $X^{(1)}, \ldots, X^{(n_\mathsf{T})}$ jointly as

$$X^{(j)} = A \quad \text{w.p. } 1, \qquad j \in \{1, \ldots, i-1\}, \tag{60a}$$

$$X^{(i)} \text{ uniform over } [0, A], \tag{60b}$$

$$X^{(j)} = 0 \quad \text{w.p. } 1, \qquad j \in \{i+1, \ldots, n_\mathsf{T}\}. \tag{60c}$$

The average input power of this strategy is equal to the RHS of (57) times $A$. ∎

## V. Concluding Remarks

MIMO wireless optical channels are different from their RF counterparts due to different constraints on the input power. The presence of peak and first-moment input constraints does not seem to allow a closed-form capacity expression at finite SNR. However, in the high-SNR regime, when the channel matrix $\mathsf{H}$ has full column rank, the MIMO optical channel considered here behaves similarly to the MIMO Gaussian channel under a second-moment constraint only. The capacity prelog is given by the column rank $n_\mathsf{T}$, and the asymptotic high-SNR capacity depends on $\mathsf{H}$ and the noise covariance $\mathsf{K}$ only via the constant term $\frac{1}{2}\log\det(\mathsf{H}^\mathsf{T}\mathsf{K}^{-1}\mathsf{H})$. This asymptotic capacity can be achieved by an input vector with IID components.

When $\mathsf{H}$ does not have full column rank, the wireless optical channel is more difficult to analyze even at high SNR. For the special case of MISO channels, we have identified the optimal input strategies at high SNR for a certain range of power constraints. Remarkably, it is possible to achieve a uniform distribution at the single receiving antenna even with a nontrivial average-power constraint that is smaller than $\frac{n_\mathsf{T}}{2}$. A generalization of these results is part of ongoing research.

## 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 Comm. Surveys & Tutorials*, vol. 16, no. 4, pp. 2231–2258, fourth quarter 2014.

[4] 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.

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

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

[7] ——, "Capacity bounds for wireless optical intensity channels with Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6066–6077, Dec. 2010.

[8] A. Thangaraj, G. Kramer, and G. Böcherer, "Capacity bounds for discrete-time, amplitude-constrained, additive white Gaussian noise channels," *arXiv*, Nov. 2015, arXiv:1511.08742.

[9] S. M. Haas, J. H. Shapiro, and V. Tarokh, "Space-time codes for wireless optical communications," *EURASIP J. App. Sig. Proc.*, no. 3, pp. 211–220, 2002.

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

[11] X. Song and J. Cheng, "Subcarrier intensity modulated MIMO optical communications in atmospheric turbulence," *IEEE/OSA J. of Optical Comm. & Networking*, vol. 5, no. 9, pp. 1001–1009, Sept. 2013.

[12] L. Mroueh and J.-C. Belfiore, "Quadratic extension field codes for free space optical intensity communications," 2017, to app. in *IEEE Trans. Commun.*

[13] 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.

[14] 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.

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

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