

On the Capacity of Block Fading Optical Wireless Channels

Longguang Li, Stefan M Moser, Ligong Wang, Michèle Wigger

▶ **To cite this version:**

Longguang Li, Stefan M Moser, Ligong Wang, Michèle Wigger. On the Capacity of Block Fading Optical Wireless Channels. 2019 IEEE Global Communications Conference (GLOBECOM 2019), Dec 2019, Waikoloa, HI, United States. 10.1109/GLOBECOM38437.2019.9013614 . hal-03009327

HAL Id: hal-03009327
https://hal.science/hal-03009327

Submitted on 31 Aug 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On the Capacity of Block Fading Optical Wireless Channels

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*—This paper investigates the capacity of block fading optical intensity channels with more transmit than receive antennas under different assumptions on the transmitter's channel state information (CSI). Lower and upper bounds on the capacities are derived using the entropy power inequality (EPI) and a dual expression for capacity. Our lower bounds for perfect and partial CSI utilize a transmit-antenna cooperation strategy based on minimum-energy signaling, which we proposed recently. For perfect CSI, this lower bound matches the upper bound asymptotically in the high signal-to-noise ratio (SNR) regime. For imperfect CSI, our lower bound is close to its perfect-CSI counterpart.**

I. Introduction

We consider a block fading optical wireless communication system where the transmitter modulates the intensity of optical signals that are emitted by light emitting diodes (LEDs), and the receiver measures incoming optical intensities by means of photodetectors which produce output electrical currents. As a first approximation, the noise in such systems can be assumed to be Gaussian and independent of the transmitted signal (thermal noise, background light sources). Inputs are nonnegative because they represent optical intensities, while outputs can be negative because they represent output electrical currents. Moreover, inputs are typically subject to both peak- and average-power constraints, where the peak-power constraint is mainly due to technical limitations of the used components, whereas the average-power constraint is imposed by battery limitations and safety considerations.

Recent years have seen extensive studies of *intensity-modulation-direct-detection (IM-DD)* systems [1]–[6], where transmitters are equipped with $n_{\mathrm{T}} \geq 1$ LEDs and receivers with $n_{\mathrm{R}} \geq 1$ photodetectors. In the high signal-to-noise ratio (SNR) regime, [5] determined the asymptotic capacity when the channel matrix is invertible or of full column rank. For general MIMO channels with average-power constraint only, the asymptotic high-SNR capacity was determined in [7], [8]. In the case when the channel input is subject to both peak- and average-power constraints, the achievable high-SNR prelog (degrees of freedom) is shown in [7], [8], and recently, the exact high-SNR asymptotic capacity is characterized in [9]. Furthermore, the asymptotic capacity in the low-SNR regime is also derived in [6], [9].

In this paper, we consider a block fading IM-DD channel where there are more transmit than receive antennas:

$$n_{\mathrm{T}} > n_{\mathrm{R}} \geq 1. \tag{1}$$

The capacity of block fading channels heavily depends on the channel modeling (such as the models for turbulence, fading, and antenna correlation) and the availability of channel state information (CSI) at the transmitter and receiver. Inspired by the previous works [10]–[12] in the classic block fading Gaussian channels, we assume that the receiver has perfect CSI, and study three different versions of transmitter CSI: no CSI, perfect CSI, and limited CSI. The main contributions of this paper are as follows.

- Lower bounds are presented for the capacities without CSI, with perfect CSI, and with limited CSI.
- The lower bounds for perfect and limited CSI are obtained using the minimum-energy signaling strategy we have derived in [9], which sets at each instant in time a subset of $n_{\mathrm{T}} - n_{\mathrm{R}}$ antennas to zero or to full power, and uses only the remaining n_{R} antennas for signaling.
- For perfect CSI, the lower bound matches a new duality-based upper bound in the high-SNR regime.
- For limited CSI, the lower bound is close to the one with perfect CSI, but requires only $(n_{\mathrm{T}} - n_{\mathrm{R}}) \log_2 \binom{n_{\mathrm{T}}}{n_{\mathrm{R}}}$ bits of feedback in each block.

II. Channel Model

Consider the following $n_{\mathrm{R}} \times n_{\mathrm{T}}$ block fading channel:

$$\mathbf{Y}_t[n] = \mathbb{H}_t \mathbf{x}_t[n] + \mathbf{Z}_t[n], \tag{2}$$

where $t \in \{1, 2, \ldots, \mathsf{B}\}$ denotes the block index, with B being the number of blocks; $n \in \{1, 2, \ldots, \mathsf{N}\}$ denotes the symbol index along a block, with N being the block length; $\mathbf{x}_t[n] = \big(x_t^{(1)}[n], \ldots, x_t^{(n_{\mathrm{T}})}[n]\big)^{\mathsf{T}}$ denotes the n_{T}-dimensional channel input vector; $\mathbf{Z}_t[n]$ denotes the n_{R}-dimensional noise vector with independent standard Gaussian entries,

$$\mathbf{Z}_t[n] \sim \mathcal{N}(0, \mathsf{I}); \tag{3}$$

and

$$\mathbb{H}_t = \big[\mathbf{H}_t^{(1)}, \mathbf{H}_t^{(2)}, \ldots, \mathbf{H}_t^{(n_{\mathrm{T}})}\big] \tag{4}$$

is a random $n_{\mathrm{R}} \times n_{\mathrm{T}}$ channel matrix with nonnegative entries. Thus, $\mathbf{H}_t^{(1)}, \ldots, \mathbf{H}_t^{(n_{\mathrm{T}})}$ are n_{R}-dimensional random column

vectors. The channel noises $\{\mathbf{Z}_t[n]\}$ are independent and identically distributed (IID) inside and across blocks. The channel matrix $\mathbb{H}_t$ remains constant within each block and is IID across blocks. We assume it has finite density $f(\mathbb{H}_t)$ over the set of nonnegative real numbers (e.g., Rayleigh, Weibull, Pareto, or gamma distributions).

Since the channel inputs correspond to optical intensities, they are nonnegative:

$$x_t^{(k)}[n] \in \mathbb{R}_0^+, \quad k = 1, \ldots, n_{\mathsf{T}}, \tag{5}$$

for all $t \in \{1, \ldots, \mathsf{B}\}$ and $n \in \{1, \ldots, \mathsf{N}\}$. Moreover, we assume that the inputs are subject to a peak-power (peak-intensity) and a *per-block* average-power (average-intensity) constraint:

$$\Pr\big[X_t^{(k)}[n] > \mathsf{A}\big] = 0, \quad \forall\, k \in \{1, \ldots, n_{\mathsf{T}}\}, \tag{6a}$$

$$\frac{1}{\mathsf{N}}\,\mathsf{E}\left[\sum_{n=1}^{\mathsf{N}} \big\|\mathbf{X}_t[n]\big\|_1\right] \le \mathsf{E}, \quad \forall\, t \in \{1, 2, \ldots, \mathsf{B}\}, \tag{6b}$$

for some fixed parameters $\mathsf{A}, \mathsf{E} > 0$. The ratio between the allowed average power and the allowed peak power is denoted by α:

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

Note that A describes the maximum power of each single LED, while E describes the allowed total average power of all LEDs together. Moreover, power allocation is permitted only inside each block. This restriction on power allocation is to prevent large visible fluctuations in the light.

The goal of the communication is to convey a random message $M \in \{1, 2, \ldots, \mathsf{M}\}$ from the transmitter to the receiver over a fixed number of B blocks. Encoding is described separately for each kind of CSI. Decoding is as follows. Based on its observed NB outputs $\mathbf{Y}_1[1], \ldots, \mathbf{Y}_{\mathsf{B}}[\mathsf{N}]$ and its knowledge of the B channel state matrices $\mathbb{H}_1, \ldots, \mathbb{H}_{\mathsf{B}}$, it produces an estimate

$$\hat{M} \triangleq \psi(\mathbf{Y}_1[1], \ldots, \mathbf{Y}_{\mathsf{B}}[\mathsf{N}], \mathbb{H}_1, \ldots, \mathbb{H}_{\mathsf{B}}). \tag{8}$$

The probability of error is defined as

$$P_{\mathrm{e}}^{(\mathsf{NB})} = \Pr\big[\hat{M} \ne M\big], \tag{9}$$

and the communication rate is given as

$$\mathsf{R} = \frac{\log_2 \mathsf{M}}{\mathsf{NB}}. \tag{10}$$

A rate R is said to be achievable if there exists a sequence of codes such that $P_{\mathrm{e}}^{(\mathsf{NB})} \to 0$ as $\mathsf{B} \to \infty$ (N remains fixed). The ergodic capacity $\mathsf{C}_{\mathbb{H}}$ of the channel is defined as the supremum of all achievable rates.

III. No CSI at the Transmitter

In this section we assume that the transmitter has no CSI.[1] In this case, the channel input $\mathbf{X}_t[n]$ is just a function of the message M:

$$\mathbf{X}_t[n] = \phi_t(M), \tag{11}$$

[1]Note that the channel *statistics* are always assumed to be known everywhere.

and the ergodic capacity is given by the following proposition.

Proposition 1: If the transmitter has no CSI, then

$$\mathsf{C}_{\mathbb{H}} = \max_{P_{\mathbf{X}} \text{ satisfying (6)}} \mathsf{E}_{\mathbb{H}}\big[\mathrm{I}(\mathbf{X}; \bar{\mathbf{X}} + \mathbf{Z} | \mathbb{H} = \mathsf{H})\big], \tag{12}$$

where

$$\bar{\mathbf{X}} \triangleq \mathbb{H}\mathbf{X}. \tag{13}$$

To obtain a lower bound on $\mathsf{C}_{\mathbb{H}}$, we lower-bound the mutual information $\mathrm{I}(\mathbf{X}; \mathbf{Y} | \mathbb{H} = \mathsf{H})$ for a given channel realization H by using the independence between $\mathbb{H}$ and $\mathbf{Z}$ and invoking the EPI:

$$\begin{aligned}
&\mathrm{I}(\mathbf{X}; \mathbb{H}\mathbf{X} + \mathbf{Z} | \mathbb{H} = \mathsf{H}) \\
&\quad = \mathsf{h}(\mathbb{H}\mathbf{X} + \mathbf{Z} | \mathbb{H} = \mathsf{H}) - \mathsf{h}(\mathbf{Z}) && (14)\\
&\quad \ge \frac{1}{2}\log_2\Big(e^{2\,\mathsf{h}(\mathbb{H}\mathbf{X}|\mathbb{H}=\mathsf{H})} + e^{2\,\mathsf{h}(\mathbf{Z})}\Big) - \mathsf{h}(\mathbf{Z}) && (15)\\
&\quad = \frac{1}{2}\log_2\left(1 + \frac{e^{2\,\mathsf{h}(\mathbb{H}\mathbf{X}|\mathbb{H}=\mathsf{H})}}{(2\pi e)^{n_{\mathsf{R}}}}\right) && (16)\\
&\quad = \frac{1}{2}\log_2\left(1 + \frac{e^{2\,\mathsf{h}(\mathsf{H}\mathbf{X})}}{(2\pi e)^{n_{\mathsf{R}}}}\right), && (17)
\end{aligned}$$

where the last equality holds because without CSI at the transmitter, the input vector $\mathbf{X}$ is independent of the channel matrix $\mathbb{H}$.

We now present choices of the distribution $P_{\mathbf{X}}$. Fix some $\lambda \in (0, \frac{n_{\mathsf{R}}}{2})$ and let the n_{R}-dimensional random vector $\mathbf{V}$ be exponentially distributed with density

$$f(\mathbf{v}) = \frac{1}{\mathsf{A}^{n_{\mathsf{R}}}}\left(\frac{\mu}{1 - e^{-\mu}}\right)^{n_{\mathsf{R}}} e^{-\frac{\mu\|\mathbf{v}\|_1}{\mathsf{A}}}, \quad \mathbf{v} \in [0, \mathsf{A}]^{n_{\mathsf{R}}}, \tag{18}$$

where μ denotes the unique solution to

$$\frac{1}{\mu} - \frac{e^{-\mu}}{1 - e^{-\mu}} = \frac{\lambda}{n_{\mathsf{R}}}. \tag{19}$$

Choose then some matrix G with nonnegative entries, with $\mathrm{rank}(\mathsf{G}) = n_{\mathsf{R}}$, with $\|\mathsf{G}\|_1 \le \alpha/\lambda$, and with row vectors satisfying $\|\mathbf{g}_i\|_1 \le 1$, $\forall\, i \in \{1, \ldots, n_{\mathsf{T}}\}$, and define $P_{\mathbf{X}}$ as the distribution of the vector

$$\mathbf{X} = \mathsf{G}\mathbf{V}. \tag{20}$$

It can be verified that this choice of $P_{\mathbf{X}}$ satisfies the power constraints (6) and puts nonzero density only on nonnegative inputs.

Evaluating the lower bound in (17) for the proposed $P_{\mathbf{X}}$ and averaging over the channel matrix $\mathbb{H}$, yields the following result.

Theorem 2: If the transmitter has no CSI, then

$$\mathsf{C}_{\mathbb{H}} \ge \frac{1}{2}\sup_{\lambda \in (0, \frac{n_{\mathsf{R}}}{2})}\ \sup_{\mathsf{G}}\, \mathsf{E}_{\mathbb{H}}\left[\log_2\left(1 + \frac{\mathsf{A}^{2n_{\mathsf{R}}}(\det \mathbb{H}\mathsf{G})^2\, e^{2\nu}}{(2\pi e)^{n_{\mathsf{R}}}}\right)\right]. \tag{21}$$

Here, the supremum is over all $n_{\mathsf{T}} \times n_{\mathsf{R}}$ matrices G with nonnegative entries, with $\mathrm{rank}(\mathsf{G}) = n_{\mathsf{R}}$, with $\|\mathsf{G}\|_1 \le \alpha/\lambda$,

and with row vectors satisfying $\|\mathbf{g}_i\|_1 \le 1, \forall\, i \in \{1,\dots,n_{\mathsf{T}}\}$; and ν is defined as

$$\nu \triangleq n_{\mathsf{R}}\left(1 - \log_2 \frac{\mu}{1-e^{-\mu}} - \frac{\mu\, e^{-\mu}}{1-e^{-\mu}}\right) \tag{22}$$

with μ being the unique solution to (19).

Proof: We use the choice of $P_{\mathbf{X}}$ specified before Theorem 2. For a given realization of the channel matrix $\mathbb{H} = \mathsf{H}$, it implies the following density on the channel image vector $\bar{\mathbf{X}} = \mathsf{H}\mathbf{X}$:

$$f_{\mathsf{H}\mathbf{X}}(\bar{\mathbf{x}}) = \frac{1}{\mathsf{A}^{n_{\mathsf{R}}}|\det \mathsf{H}\mathsf{G}|} \cdot \left(\frac{\mu}{1-e^{-\mu}}\right)^{n_{\mathsf{R}}} e^{\frac{-\mu\|(\mathsf{H}\mathsf{G})^{-1}\bar{\mathbf{x}}\|_1}{\mathsf{A}}}. \tag{23}$$

The differential entropy $\mathsf{h}(\mathsf{H}\mathbf{X})$ thus evaluates to:

$$\begin{aligned}\mathsf{h}(\mathsf{H}\mathbf{X}) &= n_{\mathsf{R}} \log_2(\mathsf{A}\cdot|\det \mathsf{H}\mathsf{G}|) \\ &\quad + n_{\mathsf{R}}\left(1 - \log_2 \frac{\mu}{1-e^{-\mu}} - \frac{\mu\, e^{-\mu}}{1-e^{-\mu}}\right).\end{aligned} \tag{24}$$

Combining this with (17), taking expectation over $\mathbb{H}$, and maximizing over the free parameters conclude the proof. ∎

IV. Perfect CSI at Transmitter

In this section, we assume that the transmitter has perfect CSI, i.e., the channel input $\mathbf{X}_t[n]$ is a function of the message M and the channel matrix $\mathbb{H}_t$:[2]

$$\mathbf{X}_t[n] = \phi_t(M, \mathbb{H}_t). \tag{25}$$

The ergodic capacity in this scenario is characterized as follows.

Proposition 3: If the transmitter has perfect CSI, then

$$\mathsf{C}_{\mathbb{H}} = \mathsf{E}_{\mathbb{H}}\Big[\max_{P_{\mathbf{X}|\mathbb{H}=\mathsf{H}} \text{ satisfying (6)}} \mathrm{I}(\mathbf{X}; \bar{\mathbf{X}} + \mathbf{Z}|\mathbb{H} = \mathsf{H})\Big]. \tag{26}$$

As before, we present a choice of the input distribution $P_{\mathbf{X}|\mathbb{H}=\mathsf{H}}$ and evaluate the lower bound on capacity obtained with this distribution and the EPI. Our choice of $P_{\mathbf{X}|\mathbb{H}=\mathsf{H}}$ is based on the minimum-energy signaling strategy proposed in [9], [13], which we recall in the following.

A. Minimum-Energy Signaling

We notice that for a fixed channel matrix H, when there are more transmit than receive antennas, i.e., $n_{\mathsf{T}} > n_{\mathsf{R}}$, then for a certain image vector $\bar{\mathbf{x}}$ there are different input vectors $\mathbf{x}$ *inducing* $\bar{\mathbf{x}}$, i.e., satisfying $\mathsf{H}\mathbf{x} = \bar{\mathbf{x}}$. In [9], [13], an algorithm is presented which for each $\bar{\mathbf{x}}$ characterizes the vector $\mathbf{x}$ that has minimum energy $\|\mathbf{x}\|_1$ among all input vectors inducing $\bar{\mathbf{x}}$.

Fix an $n_{\mathsf{R}} \times n_{\mathsf{T}}$ channel matrix H for which every tuple of n_{R} columns is linearly independent. (The probability of the random channel matrix $\mathbb{H}$ not being of this kind is 0 for any bounded density $f(\mathbb{H})$.) Note that since the input vector $\mathbf{x}$ lies in $[0,\mathsf{A}]^{n_{\mathsf{T}}}$, the image vector $\bar{\mathbf{x}} = \mathsf{H}\mathbf{x}$ takes value in the zonotope

$$\mathcal{R}(\mathsf{H}) \triangleq \big\{\bar{\mathbf{x}}\colon \bar{\mathbf{x}} = \mathsf{H}\mathbf{x},\ \forall \mathbf{x} \in [0,\mathsf{A}]^{n_{\mathsf{T}}}\big\}. \tag{27}$$

Figure 1 shows the zonotope $\mathcal{R}(\mathsf{H})$ for two different 2×4 channel matrices H.

We need some notation to describe the minimum-energy signaling in Lemma 4 ahead. Let $\mathcal{U}$ be the set

$$\mathcal{U} \triangleq \big\{\mathcal{I} \subseteq \{1,2,\dots,n_{\mathsf{T}}\}\colon |\mathcal{I}| = n_{\mathsf{R}}\big\}, \tag{28}$$

and for each $\mathcal{I} \in \mathcal{U}$, define

- the complement

$$\mathcal{I}^{\mathrm{c}} \triangleq \{1,\dots,n_{\mathsf{T}}\} \setminus \mathcal{I}; \tag{29}$$

- the n_{R}-dimensional parallelepiped

$$\mathcal{D}_{\mathcal{I}} \triangleq \left\{\sum_{i\in\mathcal{I}} \lambda_i \mathbf{h}^{(i)}\colon \lambda_i \in [0,\mathsf{A}],\ \forall\, i \in \mathcal{I}\right\}; \tag{30}$$

- the $n_{\mathsf{R}} \times n_{\mathsf{R}}$ submatrix

$$\mathsf{H}_{\mathcal{I}} \triangleq \big[\mathbf{h}^{(i)}\colon i \in \mathcal{I}\big]; \tag{31}$$

- the nonnegative number

$$s_{\mathcal{I}} \triangleq \sum_{j\in\mathcal{I}^{\mathrm{c}}} \mathbb{1}\big\{\mathbf{1}_{n_{\mathsf{R}}}^{\mathsf{T}} \mathsf{H}_{\mathcal{I}}^{-1}\mathbf{h}^{(j)} > 1\big\}, \quad \mathcal{I} \in \mathcal{U}, \tag{32}$$

  where $\mathbb{1}\{\cdot\}$ denotes the indicator function; and
- the n_{R}-dimensional vector

$$\mathbf{v}_{\mathcal{I}} \triangleq \mathsf{A} \sum_{j\in\mathcal{I}^{\mathrm{c}}} \mathbb{1}\big\{\mathbf{1}_{n_{\mathsf{R}}}^{\mathsf{T}} \mathsf{H}_{\mathcal{I}}^{-1}\mathbf{h}^{(j)} > 1\big\} \cdot \mathbf{h}^{(j)}. \tag{33}$$

Lemma 4 ([9], [13]): Given an $n_{\mathsf{R}} \times n_{\mathsf{T}}$ channel matrix H, the following holds:

1) The zonotope $\mathcal{R}(\mathsf{H})$ is covered by the parallelepipeds $\{\mathbf{v}_{\mathcal{I}} + \mathcal{D}_{\mathcal{I}}\}_{\mathcal{I}\in\mathcal{U}}$, which overlap only on sets of measure zero.
2) For any $\mathcal{I} \in \mathcal{U}$ and $\bar{\mathbf{x}} \in (\mathbf{v}_{\mathcal{I}} + \mathcal{D}_{\mathcal{I}})$, the vector $\mathbf{x} = (x_1,\dots,x_{n_{\mathsf{T}}})^{\mathsf{T}}$ that induces $\bar{\mathbf{x}}$ with minimum energy is:

$$x_i = \begin{cases} a_{\mathcal{I},i}(\mathsf{H}) & \text{if } i \in \mathcal{I}^{\mathrm{c}}, \\ \beta_{\mathcal{I},i}(\mathsf{H}, \bar{\mathbf{x}}) & \text{if } i \in \mathcal{I}, \end{cases} \tag{34}$$

where

$$a_{\mathcal{I},i}(\mathsf{H}) \triangleq \mathsf{A}\cdot\mathbb{1}\big\{\mathbf{1}_{n_{\mathsf{R}}}^{\mathsf{T}} \mathsf{H}_{\mathcal{I}}^{-1}\mathbf{h}^{(i)} > 1\big\}, \quad i \in \mathcal{I}^{\mathrm{c}}, \tag{35}$$

and where the vector

$$\boldsymbol{\beta}_{\mathcal{I}}(\mathsf{H}, \bar{\mathbf{x}}) = \big(\beta_{\mathcal{I},i}(\mathsf{H}, \bar{\mathbf{x}})\colon i \in \mathcal{I}\big)^{\mathsf{T}} \tag{36a}$$

is given by

$$\boldsymbol{\beta}_{\mathcal{I}}(\mathsf{H}, \bar{\mathbf{x}}) \triangleq \mathsf{H}_{\mathcal{I}}^{-1}(\bar{\mathbf{x}} - \mathbf{v}_{\mathcal{I}}). \tag{36b}$$

Figure 1 shows the partition of $\mathcal{R}(\mathsf{H})$ into 6 parallelepipeds for two 2×4 examples.

[2] More generally, one could allow the channel input $\mathbf{X}_t[n]$ to depend on all previous channel matrices $\mathbb{H}_1,\dots,\mathbb{H}_t$. But since we do not allow for power allocation across blocks and since the channel matrices are independent across blocks, the more general definition is not helpful.

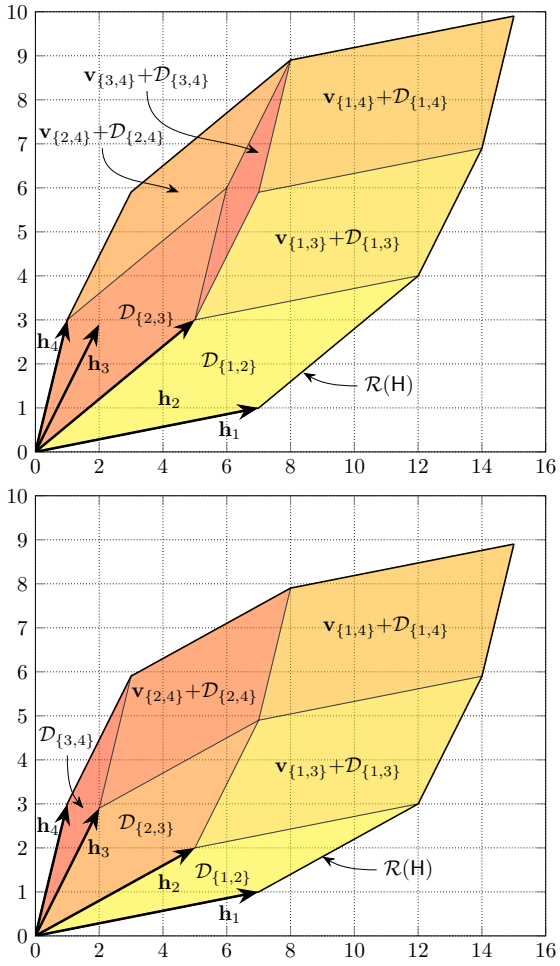


Fig. 1. Partition of $\mathcal{R}(\mathsf{H})$ into 6 parallelepipeds for two 2×4 MIMO examples. The example on the top is for $\mathsf{H} = [7, 5, 2, 1; 1, 2, 2.9, 3]$ and the example on the bottom for $\mathsf{H} = [7, 5, 2, 1; 1, 3, 2.9, 3]$.

B. The Choice of $P_{\mathbf{X}|\mathbb{H}=\mathsf{H}}$

Fix an $n_{\mathrm{R}} \times n_{\mathrm{T}}$ channel matrix H for which every tuple of n_{R} columns is linearly independent.

The following parameters are important in this section. Let $\mathbf{q}$ be a probability vector on $\mathcal{U}$ with entries

$$q_{\mathcal{I}} \triangleq \frac{|\det \mathsf{H}_{\mathcal{I}}|}{\sum_{\mathcal{I}' \in \mathcal{U}} |\det \mathsf{H}_{\mathcal{I}'}|}, \quad \mathcal{I} \in \mathcal{U}, \tag{37}$$

and define

$$\alpha_{\mathrm{th}}(\mathsf{H}) \triangleq \frac{n_{\mathrm{R}}}{2} + \sum_{\mathcal{I} \in \mathcal{U}} s_{\mathcal{I}}\, q_{\mathcal{I}}. \tag{38}$$

If $\alpha \geq \alpha_{\mathrm{th}}(\mathsf{H})$, we choose $P_{\mathbf{X}}$ so that $\bar{\mathbf{X}}$ is uniform over $\mathcal{R}(\mathsf{H})$. This is obtained by defining a random variable over $\mathcal{U}$ with probability mass function (PMF) $\mathbf{q}$ and conditional on $\tilde{U} = \mathcal{I}$, choose $\mathbf{X}$ according to the minimum-energy signaling in (34) and so that $\bar{\mathbf{X}}$ is uniform over the shifted parallelepiped $\mathbf{v}_{\mathcal{I}} + \mathcal{D}_{\mathcal{I}}$.

If $\alpha < \alpha_{\mathrm{th}}(\mathsf{H})$, we fix a parameter

$$\lambda \in \left(\max\left\{0, \frac{n_{\mathrm{R}}}{2} + \alpha - \alpha_{\mathrm{th}}(\mathsf{H})\right\}, \min\left\{\frac{n_{\mathrm{R}}}{2}, \alpha\right\}\right) \tag{39}$$

and a PMF $\mathbf{p} = (p_{\mathcal{I}} \colon \mathcal{I} \in \mathcal{U})$ over the set $\mathcal{U}$ so that

$$\sum_{\mathcal{I} \in \mathcal{U}} p_{\mathcal{I}} s_{\mathcal{I}} = \alpha - \lambda. \tag{40}$$

Let then $\tilde{U}$ be a random variable over $\mathcal{U}$ with PMF $\mathbf{p}$, and conditional on $\tilde{U} = \mathcal{I}$, choose $\mathbf{X}$ according to the minimum-energy signaling in (34) and so that $\bar{\mathbf{X}}$ follows an n_{R}-dimensional truncated exponential distribution over the shifted parallelepiped $\mathbf{v}_{\mathcal{I}} + \mathcal{D}_{\mathcal{I}}$. Specifically, given $\mathbb{H} = \mathsf{H}$ and $\tilde{U} = \mathcal{I}$, the inputs $\{X_i \colon i \in \mathcal{I}^{\mathrm{c}}\}$ are deterministically set to

$$X_i = a_{\mathcal{I},i}(\mathsf{H}), \quad i \in \mathcal{I}^{\mathrm{c}}, \tag{41}$$

where $a_{\mathcal{I},i}(\mathsf{H})$ is defined in (35), and the remaining inputs $\{X_i \colon i \in \mathcal{I}\}$ are chosen according to the truncated exponential distribution

$$f_{X_i|\tilde{U}=\mathcal{I}}(x_i) = \frac{\mu}{1 - e^{-\mu}}\, e^{-\frac{\mu x_i}{\mathsf{A}}}, \quad \forall i \in \mathcal{I}. \tag{42}$$

Then at the receiver side, the image vector $\bar{\mathbf{X}} = \mathsf{H}\mathbf{X}$ is of conditional density

$$f_{\bar{\mathbf{X}}|\tilde{U}=\mathcal{I}}(\bar{\mathbf{x}}) = \frac{1}{\mathsf{A}^{n_{\mathrm{R}}} |\det \mathsf{H}_{\mathcal{I}}|} \cdot \left(\frac{\mu}{1 - e^{-\mu}}\right)^{n_{\mathrm{R}}} e^{\frac{-\mu \|\mathsf{H}_{\mathcal{I}}^{-1}(\bar{\mathbf{x}} - \mathbf{v}_{\mathcal{I}})\|_1}{\mathsf{A}}}. \tag{43}$$

C. Capacity Lower Bound

Theorem 5: If the transmitter has perfect CSI, then

$$\mathsf{C}_{\mathbb{H}} \geq \frac{1}{2}\, \mathsf{E}_{\mathbb{H}}\left[\log_2\left(1 + \frac{\mathsf{A}^{2n_{\mathrm{R}}} \mathsf{V}_{\mathbb{H}}^2}{(2\pi e)^{n_{\mathrm{R}}}}\, e^{2\nu(\mathbb{H})}\right)\right], \tag{44}$$

where for each realization of the channel matrix H:

$$\mathsf{V}_{\mathsf{H}} \triangleq \sum_{\mathcal{I} \in \mathcal{U}} |\det \mathsf{H}_{\mathcal{I}}|, \tag{45}$$

and if $\alpha \geq \alpha_{\mathrm{th}}(\mathsf{H})$, then

$$\nu(\mathsf{H}) \triangleq 0, \tag{46}$$

whereas if $\alpha < \alpha_{\mathrm{th}}(\mathsf{H})$, then

$$\nu(\mathsf{H}) \triangleq \sup_{\lambda \in \left(\max\left\{0, \frac{n_{\mathrm{R}}}{2} + \alpha - \alpha_{\mathrm{th}}\right\}, \min\left\{\frac{n_{\mathrm{R}}}{2}, \alpha\right\}\right)} \left\{ n_{\mathrm{R}}\left(1 - \log_2 \frac{\mu}{1 - e^{-\mu}} - \frac{\mu\, e^{-\mu}}{1 - e^{-\mu}}\right) - \inf_{\mathbf{p}} \mathsf{D}(\mathbf{p} \| \mathbf{q}) \right\}. \tag{47}$$

Thus, $\nu(\mathsf{H})$ is nonpositive and corresponds to the penalty due to α.

Proof: Consider first the case $\alpha \geq \alpha_{\mathrm{th}}(\mathsf{H})$. With the proposed choice ($\bar{\mathbf{X}}$ uniform over $\mathcal{R}(\mathsf{H})$), for each realization $\mathbb{H} = \mathsf{H}$, the conditional differential entropy is

$$\mathsf{h}(\bar{\mathbf{X}} | \mathbb{H} = \mathsf{H}) = \log_2(\mathsf{A}^{n_{\mathrm{R}}} \cdot \mathsf{V}_{\mathsf{H}}). \tag{48}$$

Combining this with (16) yields the desired result.

Consider now the case $\alpha < \alpha_{\mathrm{th}}(\mathsf{H})$. We define U to be a discrete random variable that is obtained by applying a function on $\bar{\mathbf{X}}$ in a way to satisfy:

$$\big(U = \mathcal{I}\big) \iff \big(\bar{\mathbf{X}} \in (\mathbf{v}_{\mathcal{I}} + \mathcal{D}_{\mathcal{I}})\big). \tag{49}$$

We then start from (16) and decompose the conditional differential entropy $\mathsf{h}(\bar{\mathbf{X}}|\mathbb{H}=\mathsf{H})$ as

$$\mathsf{h}(\bar{\mathbf{X}}|\mathbb{H}=\mathsf{H}) = \mathsf{I}(\bar{\mathbf{X}};U|\mathbb{H}=\mathsf{H}) + \mathsf{h}(\bar{\mathbf{X}}|U,\mathbb{H}=\mathsf{H}) \quad (50)$$

$$= \mathsf{H}(U|\mathbb{H}=\mathsf{H}) + \mathsf{h}(\bar{\mathbf{X}}|U,\mathbb{H}=\mathsf{H}) \quad (51)$$

where we used that U is a function of $\bar{\mathbf{X}}$. By the choice of the PMF $\mathbf{p}$ in (40) and the exponential distribution in (43), we then have:

$$\mathsf{h}(\bar{\mathbf{X}}|\mathbb{H}=\mathsf{H}) = \mathsf{H}(U|\mathbb{H}=\mathsf{H}) + h(\bar{\mathbf{X}}|U,\mathbb{H}=\mathsf{H}) \quad (52)$$

$$= \mathsf{H}(\mathbf{p}) + \log_2|\det \mathsf{H}_{\mathcal{I}}| + n_{\mathsf{R}}\log_2 \mathsf{A} + n_{\mathsf{R}}\left(1 - \log_2\frac{\mu}{1-e^{-\mu}} - \frac{\mu\, e^{-\mu}}{1-e^{-\mu}}\right) \quad (53)$$

$$= -\mathsf{D}(\mathbf{p}\|\mathbf{q}) + \log_2 \mathsf{V}_{\mathsf{H}} + n_{\mathsf{R}}\log_2 \mathsf{A} + n_{\mathsf{R}}\left(1 - \log_2\frac{\mu}{1-e^{-\mu}} - \frac{\mu\, e^{-\mu}}{1-e^{-\mu}}\right). \quad (54)$$

The result now follows from substituting (54) into (16), maximizing over the free parameters λ and $\mathbf{p}$, and taking expectation over $\mathbb{H}$. ∎

Next, we present an upper bound on capacity. Its proof is omitted.

Theorem 6: If the transmitter has perfect CSI, then

$$\mathsf{C}_{\mathbb{H}} \le \mathsf{E}_{\mathbb{H}}\Big[\mathsf{C}_{\mathbb{H},1}\cdot \mathbb{1}\big\{\alpha \ge \alpha_{\text{th}}(\mathbb{H})\big\} + \mathsf{C}_{\mathbb{H},2}\cdot \mathbb{1}\big\{\alpha < \alpha_{\text{th}}(\mathbb{H})\big\}\Big], \quad (55)$$

where for each realization of the channel matrix H:

$$\mathsf{C}_{\mathsf{H},1} \triangleq n_{\mathsf{R}}\log_2\left(\sigma_{\max} + \frac{\mathsf{A}}{\sqrt{2\pi e}}\right) + \log_2 \mathsf{V}_{\mathsf{H}} \quad (56)$$

and

$$\mathsf{C}_{\mathsf{H},2} \triangleq \sup_{\mathbf{p}}\inf_{\mu>0}\Bigg\{ n_{\mathsf{R}}\log_2\left(\sigma_{\max} + \frac{\mathsf{A}}{\sqrt{2\pi e}}\frac{1-e^{-\mu}}{\mu}\right) + \log_2 \mathsf{V}_{\mathsf{H}} + \frac{\mu n_{\mathsf{R}}\sigma_{\max}}{\mathsf{A}\sqrt{2\pi}}\left(1 - e^{-\frac{\mathsf{A}^2}{2\sigma_{\min}}}\right) + \mu\left(\alpha - \sum_{\mathcal{I}\in\mathcal{U}} p_{\mathcal{I}} s_{\mathcal{I}}\right) - \mathsf{D}(\mathbf{p}\|\mathbf{q})\Bigg\} \quad (57)$$

with the supremum being over all $\mathbf{p}$ on $\mathcal{U}$ and with

$$\sigma_{\max} \triangleq \max_{\substack{\mathcal{I}\in\mathcal{U},\\ \ell\in\{1,\ldots,n_{\mathsf{R}}\}}} \sigma_{\mathcal{I},\ell} \quad \text{and} \quad \sigma_{\min} \triangleq \min_{\substack{\mathcal{I}\in\mathcal{U},\\ \ell\in\{1,\ldots,n_{\mathsf{R}}\}}} \sigma_{\mathcal{I},\ell}, \quad (58)$$

where $\sigma_{\mathcal{I},\ell}$ denotes the square root of the ℓth diagonal entry of the matrix $\mathsf{H}_{\mathcal{I}}^{-1}\mathsf{H}_{\mathcal{I}}^{-\mathsf{T}}$.

Finally, we analyze the asymptotic capacity for $\mathsf{A},\mathsf{E}\to\infty$ with α held fixed. The following result follows immediately from Theorems 5 and 6.

Theorem 7: If the transmitter has perfect CSI, then

$$\lim_{\mathsf{A}\to\infty}\big\{\mathsf{C}_{\mathbb{H}} - n_{\mathsf{R}}\log_2\mathsf{A}\big\} = \frac{1}{2}\,\mathsf{E}_{\mathbb{H}}\left[\log_2\left(\frac{\mathsf{V}_{\mathbb{H}}^2}{(2\pi e)^{n_{\mathsf{R}}}}\right) + \nu(\mathbb{H})\cdot\mathbb{1}\big\{\alpha<\alpha_{\text{th}}(\mathbb{H})\big\}\right], \quad (59)$$

where for each realization of the channel matrix H, $\nu(\mathsf{H})$ is defined by (46) for $\alpha \ge \alpha_{\text{th}}(\mathsf{H})$ and by (47) for $\alpha < \alpha_{\text{th}}(\mathsf{H})$.

V. Limited CSI at Transmitter

In this section we assume an instantaneous rate-limited *channel state feedback link* from the receiver to the transmitter. At the very beginning of each block, before transmission begins, the receiver learns $\mathbb{H}_t$ and sends a function of it,

$$\mathcal{F}_t(\mathbb{H}_t), \quad (60)$$

back to the transmitter.

The transmitter can thus compute its channel inputs $\mathbf{X}_t[n]$ as a function of the received feedback $\mathcal{F}_t(\mathbb{H}_t)$ and the message M:

$$\mathbf{X}_t[n] = \phi\big(\mathcal{F}_t(\mathbb{H}_t), M\big). \quad (61)$$

Of course, the capacity of this channel depends on the functions $\{\mathcal{F}_t\}_{t=1}^{\mathsf{B}}$. We will assume a *stationary* feedback policy where

$$\mathcal{F}_1 = \cdots = \mathcal{F}_{\mathsf{B}} = \mathcal{F}. \quad (62)$$

Proposition 8: The capacity $\mathsf{C}_{\mathbb{H},\mathcal{F}}$ of a channel with limited CSI $\mathcal{F}(\mathbb{H})$ at the transmitter is:

$$\mathsf{C}_{\mathbb{H},\mathcal{F}} = \mathsf{E}_{\mathbb{H}}\left[\max_{P_{\mathbf{X}|\mathcal{F}(\mathsf{H})} \text{ satisfying (6)}} \mathsf{I}(\mathbf{X};\bar{\mathbf{X}}+\mathbf{Z}|\mathbb{H}=\mathsf{H})\right]. \quad (63)$$

We present a choice of the function $\mathcal{F}$ with only $\binom{n_{\mathsf{T}}}{n_{\mathsf{R}}}^{n_{\mathsf{T}}-n_{\mathsf{R}}}$ function values, for which the corresponding capacity with limited CSI is close to the one with perfect CSI. Obviously, to implement this function it suffices that the receiver can feed back

$$\mathsf{R}_{\mathsf{FB}} = (n_{\mathsf{T}} - n_{\mathsf{R}})\log_2\binom{n_{\mathsf{T}}}{n_{\mathsf{R}}} \quad (64)$$

bits in each block.

Before describing our choice of $\mathcal{F}$, we notice that to achieve the rates in Theorem 5, it suffices that the transmitter learn the binary values $\{a_{\mathcal{I},i}(\mathsf{H})\}_{\mathcal{I}\in\mathcal{U}}$ and the PMF $\mathbf{p}$ (and μ when $\alpha<\alpha_{\text{th}}$) for each realization of H. Learning the PMF $\mathbf{p}$ at the transmitter requires the state-feedback to have infinite number of bits. In contrast, the binary values $\{a_{\mathcal{I},i}(\mathsf{H})\}_{\mathcal{I}\in\mathcal{U}}$ can be learned with only $(n_{\mathsf{T}}-n_{\mathsf{R}})\log_2\binom{n_{\mathsf{T}}}{n_{\mathsf{R}}}$ bits of feedback.

We thus propose to set

$$\mathcal{F}(\mathsf{H}) \triangleq \big(a_{\mathcal{I},i}(\mathsf{H})\colon\ \mathcal{I}\in\mathcal{U}, i\in\mathcal{I}^{\mathrm{c}}\big). \quad (65)$$

The transmitter then uses the following conditional distribution $P_{\mathbf{X}|\mathcal{F}}$ to generate its (random) codebook for transmitting the desired message M. For any realization $\mathcal{F}$, pick an arbitrary positive number $\lambda \in (0, \frac{n_{\mathsf{R}}}{2})$ and let μ denote the solution to (19) for the picked value of λ. Then for this λ, choose an arbitrary PMF $\mathbf{p}$ over $\mathcal{U}$ satisfying (40) and let $\tilde{U}\sim\mathbf{p}$. (By (32), (35), and (65), the parameters $\{s_{\mathcal{I}}\}_{\mathcal{I}\in\mathcal{U}}$ in condition (40) depend on H only through $\mathcal{F}(\mathsf{H})$, and thus the proposed $\mathbf{p}$ can be used as a parameter for the distribution $P_{\mathbf{X}|\mathcal{F}(\mathsf{H})}$.) Similarly to before, given any realizations of $\mathcal{F}$ and $\tilde{U}$, we deterministically set

$$X_i = a_{\mathcal{I},i}(\mathsf{H}), \quad i\in\mathcal{I}^{\mathrm{c}}, \quad (66)$$

and choose the remaining inputs $\{X_i\colon i\in\mathcal{I}\}$ according to the distribution in (42).

Theorem 9: For the limited CSI function $\mathcal{F}$ in (65):

$$\mathsf{C}_{\mathbb{H},\mathcal{F}} \geq \mathsf{E}_{\mathcal{F}(\mathbb{H})}\left[\sup_{\lambda\in(0,\frac{n_{\mathrm{R}}}{2})}\sup_{\mathbf{p}}\mathsf{E}_{\mathbb{H}|\mathcal{F}}\left[\frac{1}{2}\log\left(1+\frac{\mathsf{A}^{2n_{\mathrm{R}}}\mathsf{V}_{\mathbb{H}}{}^{2}\,e^{2\nu^\star}}{(2\pi e)^{n_{\mathrm{R}}}}\right)\right]\right] \tag{67}$$

where

$$\nu^\star \triangleq n_{\mathrm{R}}\left(1-\log_2\frac{\mu}{1-e^{-\mu}}-\frac{\mu\,e^{-\mu}}{1-e^{-\mu}}\right)-\mathrm{D}(\mathbf{p}\|\mathbf{q}) \tag{68}$$

with μ satisfying (19), and where the supremum over $\mathbf{p}$ is over all PMFs satisfying (40).

Proof: The proof is similar to the proof of Theorem 5 except that now the choices of λ and $\mathbf{p}$ can depend on $\mathbb{H}$ only through $\mathcal{F}(\mathbb{H})$. The details are omitted. ■

VI. Numerical Example

Figure 2 illustrates the derived upper and lower bounds on the capacities for a 1×3 MISO channel where the entries $\{h_i\}_{i\in\{1,2,3\}}$ in the channel matrix $\mathbb{H}$ are IID and follow a Gamma-Gamma distribution [14]:

$$f(h_i)=\frac{2(\beta\gamma)^{\frac{\beta+\gamma}{2}}}{\Gamma(\beta)\Gamma(\gamma)}\,h_i^{\frac{\beta+\gamma}{2}-1}\,\mathsf{K}_{\beta-\gamma}(2(\beta\gamma h_i)^{\frac{1}{2}})\cdot\mathbb{1}\{h_i\geq 0\},\quad \forall\, i\in\{1,2,3\}, \tag{69}$$

where $\Gamma(\cdot)$ is a Gamma function, and $\mathsf{K}_{\beta-\gamma}(\cdot)$ is a modified Bessel function of the second kind. We have set $\beta\triangleq 2.23$ and $\gamma\triangleq 1.54$. As shown in Figure 2, when the SNR tends to infinity, the lower bound in Theorem 5 coincides with the upper bound in Theorem 6. Furthermore, the gap between the lower bound in Theorem 9 below and the upper bound in Theorem 6 tends to a small constant at high SNR. Thus, at high SNR a relatively small number of feedback bits suffices to approach the perfect-CSI capacity.

VII. Conclusion

In this paper we derived lower bounds on the capacity of optical block fading channels in the three cases where the transmitter has no CSI, perfect CSI, and limited CSI, respectively. For perfect and limited CSI, our lower bounds are based on the minimum-energy signaling strategy in [9]. In particular, for limited CSI, in each block, the receiver feeds back only the binary values $\{a_{\mathcal{I},i}(\mathsf{H})\}_{\mathcal{I}\in\mathcal{U}}$ where H denotes the channel realization in this block. This requires only $(n_{\mathrm{T}}-n_{\mathrm{R}})\log_2\binom{n_{\mathrm{T}}}{n_{\mathrm{R}}}$ bits of feedback per block, but the corresponding lower bound performs close to the lower bound with perfect CSI.

In the case of perfect CSI we also provided an upper bound on capacity and showed that it approaches the proposed lower bound asymptotically in the high-SNR regime. In this regime, the perfect CSI capacity can thus be closely approached with only a small number of feedback bits per block.

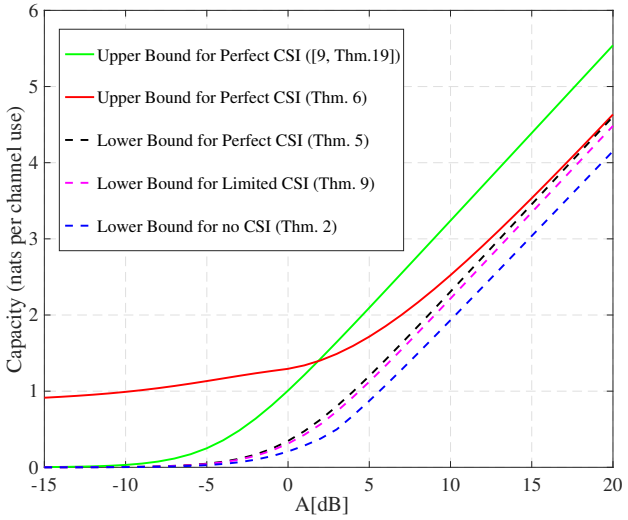


Fig. 2. A 1×3 MISO channel, where the entries in $\mathbb{H}$ follow a Gamma-Gamma distribution, with $\alpha=0.4$.

References

[1] 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.

[2] D. Karunatilaka, F. Zafar, V. Kalavally, and R. Parthiban, "LED based indoor visible light communications: State of the art," *IEEE Comm. Surveys & Tutorials*, vol. 17, no. 3, pp. 1649–1678, third quarter 2015.

[3] M. Uysal and H. Nouri, "Optical wireless communications – an emerging technology," in *Proc. IEEE Int. Conf. Transp. Optical Netw.*, Graz, Austria, Jul. 6–10, 2014, pp. 1–7.

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

[5] 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*, Aachen, Germany, Jun. 25–30, 2017, pp. 536–540.

[6] A. Chaaban, Z. Rezki, and M.-S. Alouini, "Low-SNR asymptotic capacity of MIMO optical intensity channels with peak and average constraints," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4694–4705, Oct. 2018.

[7] ——, "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.

[8] ——, "Capacity bounds and high-SNR capacity of MIMO intensity-modulation optical channels," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 3003–3017, May 2018.

[9] L. Li, S. Moser, L. Wang, and M. Wigger, "On the capacity of MIMO optical wireless channels," Feb. 2019, subm. to *IEEE Trans. Inf. Theory*.

[10] İ. E. Telatar, "Capacity of multi-antenna Gaussian channels," AT&T Bell Laboratories, Tech. Rep., 1995.

[11] G. J. Foschini, "Layered space-time architecture for wireless communication in fading environments when using multi-element antennas," *Bell Syst. Tech. J.*, pp. 41–59, 1996.

[12] S. A. Jafar, S. Vishwanath, and A. Goldsmith, "Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback," in *Proc. IEEE Int. Conf. Commun.*, Helsinki, Finland, Jun. 11–14, 2001, pp. 2266–2270.

[13] L. Li, S. Moser, L. Wang, and M. Wigger, "On the capacity of MIMO optical wireless channels," in *Proc. IEEE Inf. Theory Workshop*, Guangzhou, China, Nov. 25–29, 2018, pp. 1–5.

[14] A. Al-Habash, L. C. Andrews, and R. L. Phillips, "Mathematical model for the irradiance probability density function of a laser beam propagating through turbulent media," *Opt. Eng.*, vol. 40, no. 8, pp. 1554–1563, Aug. 2001.