First- and Second-Moment Constrained Gaussian Channels

Shuai Ma*†‡ and Michèle Wigger*
*LTCI, Telecom Paris, IP Paris, 91120 Palaiseau, France
†National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China
‡School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China
mashuai001@cumt.edu.cn, michele.wigger@telecom-paris.fr

***Abstract*—This paper studies the channel capacity of intensity-modulation direct-detection (IM/DD) visible light communication (VLC) systems under both optical and electrical power constraints. Specifically, it derives the asymptotic capacities in the high and low signal-to-noise ratio (SNR) regimes under peak, first-moment, and second-moment constraints. The results show that first- and second-moment constraints are never simultaneously active in the asymptotic low-SNR regime, and only in few cases in the asymptotic high-SNR regime. Moreover, the second-moment constraint is more stringent in the asymptotic low-SNR regime than in the high-SNR regime.**

I. Introduction

The ever-increasing number of wireless devices and high-speed communication requirements cause a spectrum scarcity of conventional radio-frequencies (RF). A promising solution is visible light communication (VLC) with its abundant unlicensed spectrum [1], [2]. In particular, when utilizing the simple and practical intensity modulation–direct detection (IM/DD) technology, transmitters directly modulate information onto the real, non-negative optical intensity of the VLC signals (in contrast to RF signals which modulate the complex field) and receivers apply photodetectors to measure incoming optical intensities. For eye safety reasons and hardware limitations, both the maximum and average optical intensities of VLC transmit signals typically have to be restricted. Since these apply directly to the intensities, they impose both peak and *first-moment constraints* on the transmit signals. Additional *second-moment constraints* are imposed by limitations of the electronic circuits that control the transmit signal, such as the boundedness of the linear amplification regime and electric power consumption [3]–[8].

A close-form expression for the capacity of such IM/DD systems is still unknown, even when some of the first or second-moment constraints are relaxed. However, bounds and asymptotic results in the high and low signal-to-noise ratio (SNR) regimes are known under certain relaxations. For example, various upper and lower bounds on the capacity, as well as its exact high- and low-SNR asymptotics, have been derived under only a first-moment constraint without a second-moment constraint [9]–[15]. An interesting related model with optical amplifiers was also studied in [16].

In this work, we derive the exact expressions for the asymptotic high- and low-SNR capacities under peak, first-moment, *and* second-moment constraints. Our results show that in the asymptotic low-SNR regime, only one of the two moment constraints is stringent. Specifically, the second-moment constraint is active if the peak-constraint A times the first-moment-constraint $\alpha_1\mathsf{A}$ exceeds this second-moment constraint $\alpha_2\mathsf{A}^2$, and otherwise the first-moment constraint is active. This can be seen as a consequence of the optimality of on-off keying in the asymptotic low-SNR regime. Our results further show that for most constraint-parameters (α_1, α_2) also in the high-SNR regime, only one of the moment-constraints is active. Interestingly, the second-moment constraint is inactive over a larger region of (α_1, α_2)-pairs in the high-SNR regime than in the low-SNR regime, and the first-moment constraint over a smaller region. An additional second-moment constraint is thus more restrictive in the low-SNR regime than in the high-SNR regime. In the asymptotic high-SNR regime, we further observe a small region of (α_1, α_2)-pairs where both moment-constraints are simultaneously active and limit the asymptotic capacity.

II. Channel Model

Consider a typical VLC communication link, where the transmitter is equipped with a single LED or laser and the receiver with a single photodetector. The photodetector measures the incoming light intensity, which can be modeled as [9]–[11], [13]–[15].

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

where x denotes the input signal produced by the transmitter's LED or laser, and Z is standard additive white Gaussian noise independent of x. The noise Z includes both optical noise and thermal noise. Note that, in contrast to the input x, the output Y can be negative.

Inputs x are subject to both a peak and an average optical power (average-intensity) constraints:

$$X \in [0, \mathsf{A}], \tag{2}$$

$$\mathsf{E}[X] \le \alpha_1 \mathsf{A}, \tag{3}$$

for some fixed parameters $\mathsf{A} > 0$ and $\alpha_1 \in (0, 1)$. These constraints come from (eye- and skin-) safety reasons, and from limitations (caused by non-linearities) on the optical operating regimes of LEDs or lasers.

Due to battery limitations on the attached RF circuit and power amplifier limitations, the *second moment* of the transmit signal also has to be restricted:

$$\mathsf{E}\big[X^2\big] \leq \alpha_2 \mathsf{A}^2. \tag{4}$$

where $\alpha_2 \in (0, 5/4)$.

We denote the capacity of the channel (1) with allowed peak power A, maximum average power $\alpha_1 \mathsf{A}$, and maximum second moment $\alpha_2 \mathsf{A}^2$ by $\mathsf{C}(\alpha_1, \alpha_2, \mathsf{A})$. It is given by [17]

$$\mathsf{C}(\alpha_1, \alpha_2, \mathsf{A}) = \sup_{P_X} \mathrm{I}(X; Y), \tag{5}$$

where the supremum is over input laws P_X satisfying (2)–(4).

Notice that, for any random variable $X \in [0, \mathsf{A}]$, we have $\mathsf{E}\big[X^2\big] \leq \mathsf{E}[X]\mathsf{A}$ and of course $\mathsf{E}[X] \leq \sqrt{\mathsf{E}[X^2]}$. Therefore, whenever $\alpha_1 < \alpha_2$, the second moment constraint (4) is inactive in view of the first moment-constraint (3), and whenever $\sqrt{\alpha_2} < \alpha_1$, the first moment constraint (3) is inactive in view of the second moment-constraint (4). Observe further that by the symmetry of the Gaussian density, for any input X, we have $\mathrm{I}(X;Y) = \mathrm{I}(X';Y)$ for the derived input $X' = \mathsf{A} - X$, which has smaller first and second moments than X if $\mathsf{E}[X] \geq \mathsf{A}/2$:

$$\mathsf{E}[X'] = \mathsf{A} - \mathsf{E}[X] \leq 1/2\mathsf{A} \leq \mathsf{E}[X], \tag{6}$$

and

$$\mathsf{E}\big[X'^2\big] = \mathsf{A}^2 - 2\,\mathsf{E}[X]\mathsf{A} + \mathsf{E}\big[X^2\big] \leq \mathsf{E}\big[X^2\big]. \tag{7}$$

Therefore, for any $\alpha_1 \geq 1/2$, the first-moment constraint (3) is not active, and for $\alpha_2 \geq 1/2$, the second-moment constraint (4) is not active.

As a consequence:

$$\mathsf{C}(\alpha_1, \alpha_2, \mathsf{A}) = \mathsf{C}(1, \alpha_2, \mathsf{A}), \qquad \forall \alpha_1 \geq \min\{\sqrt{\alpha_2}, 1/2\}, \tag{8}$$

and

$$\mathsf{C}(\alpha_1, \alpha_2, \mathsf{A}) = \mathsf{C}(\alpha_1, 1, \mathsf{A}), \qquad \forall \alpha_2 \geq \min\{\alpha_1, 1/2\}. \tag{9}$$

In the remainder of the paper, we present bounds on the capacities, and establish the exact asymptotic results in the high and low SNR regimes, respectively.

The following functions will be used throughout the paper. For $i = 0, 1, 2, 3, 4$, define:

$$\zeta_i(\lambda_1, \lambda_2) := \int_0^1 y^i e^{-\lambda_1 y - \lambda_2 y^2} \mathrm{d}y. \tag{10}$$

III. The Asymptotic High-SNR Capacity

Consider first the asymptotic high-SNR regime, where α_1, α_2 are fixed and A grows without bound.

Theorem 1: Depending on the parameters $\alpha_1, \alpha_2 > 0$, the asymptotic high-SNR capacity satisfies one of the following limiting behaviours.

1) If $\alpha_1 \geq \frac{1}{2}$ and $\alpha_2 \geq \frac{1}{3}$, then both the first- and second-moment constraints are inactive and

$$\varlimsup_{\mathsf{A}\to\infty} \left(\mathsf{C}(\alpha_1, \alpha_2, \mathsf{A}) - \log \frac{\mathsf{A}}{\sqrt{2\pi e \sigma^2}} \right) = 0. \tag{11}$$

2) If $0 < \alpha_1 < 1/2$ is such that the unique solution λ_1^* to the equation (in λ_1)

$$\frac{1}{\lambda_1} - \frac{e^{-\lambda_1}}{1 - e^{-\lambda_1}} = \alpha_1, \tag{12a}$$

satisfies

$$\frac{2}{(\lambda_1^*)^2} - \frac{e^{-\lambda_1^*}\left(1 + \frac{2}{\lambda_1^*}\right)}{1 - e^{-\lambda_1^*}} < \alpha_2, \tag{12b}$$

then only the first moment constraint is active and

$$\varlimsup_{\mathsf{A}\to\infty} \left(\mathsf{C}(\alpha_1, \alpha_2, \mathsf{A}) - \log \frac{\mathsf{A}}{\sqrt{2\pi e \sigma^2}} \right) = \log \zeta_0(\lambda_1^*, 0) + \lambda_1^* \alpha_1. \tag{12c}$$

3) If $0 < \alpha_2 < 1/3$ is such that the unique solution λ_2^* to the equation (in λ_2)

$$2\sqrt{\pi\lambda_2}\big((2\lambda_2)^{-1} - \alpha_2\big)\left[\frac{1}{2} - \mathcal{Q}\Big(\sqrt{2\lambda_2}\Big)\right] = e^{-\lambda_2}, \tag{13a}$$

satisfies

$$2\sqrt{\pi\lambda_2^*}\alpha_1\left[\frac{1}{2} - \mathcal{Q}\Big(\sqrt{2\lambda_2^*}\Big)\right] > 1 - e^{-\lambda_2^*}, \tag{13b}$$

then only the second moment constraint is active and

$$\varlimsup_{\mathsf{A}\to\infty} \left(\mathsf{C}(\alpha_1, \alpha_2, \mathsf{A}) - \log \frac{\mathsf{A}}{\sqrt{2\pi e \sigma^2}} \right) = \log \zeta_0(0, \lambda_2^*) + \lambda_2^* \alpha_2. \tag{13c}$$

4) Else, both moment constraints are active and

$$\varlimsup_{\mathsf{A}\to\infty} \left(\mathsf{C}(\alpha_1, \alpha_2, \mathsf{A}) - \log \frac{\mathsf{A}}{\sqrt{2\pi e \sigma^2}} \right) = \log \zeta_0(\lambda_1^*, \lambda_2^*) + \lambda_1^* \alpha_1 + \lambda_2^* \alpha_2, \tag{14}$$

for $\lambda_1^*, \lambda_2^* > 0$ the unique solution to the equations

$$\sqrt{\pi\lambda_2} e^{\frac{\lambda_1^2}{4\lambda_2}} \left(2\alpha_1 + \frac{\lambda_1}{\lambda_2}\right)\left[\mathcal{Q}\left(\frac{\lambda_1}{\sqrt{2\lambda_2}}\right) - \mathcal{Q}\left(\frac{\lambda_1 + 2\lambda_2}{\sqrt{2\lambda_2}}\right)\right] = 1 - e^{-(\lambda_1 + \lambda_2)} \tag{15a}$$

and

$$\sqrt{\frac{\pi}{\lambda_2}} e^{\frac{\lambda_1^2}{4\lambda_2}} \left(\alpha_2 - \frac{\lambda_2 - \lambda_1^2}{2\lambda_2^2}\right)\left[\mathcal{Q}\left(\frac{\lambda_1}{\sqrt{2\lambda_2}}\right) - \mathcal{Q}\left(\frac{\lambda_1 + 2\lambda_2}{\sqrt{2\lambda_2}}\right)\right] = \frac{1}{2\lambda_2} e^{-(\lambda_1 + \lambda_2)}\left(\frac{\lambda_1}{\lambda_2} - 1\right) - \frac{\lambda_1}{2\lambda_2^2}. \tag{15b}$$

IV. The Asymptotic Low-SNR Capacity

Consider now the asymptotic low-SNR regime, where α_1, α_2 are again kept fixed and $\mathsf{A} \to 0$.

Proposition 2: Given parameters $\alpha_1, \alpha_2 > 0$,

$$\lim_{\mathsf{A}\downarrow 0} \frac{\mathsf{C}(\alpha_1, \alpha_2, \mathsf{A})}{\mathsf{A}^2} = \max_{\substack{T\in[0,1]:\ \mathsf{E}[T]\leq\alpha_1 \\ \frac{1}{2}\mathsf{E}[T^2]\leq\alpha_2}} \mathsf{Var}[T]. \tag{16}$$

Proof: The achievability follows directly from Prelov's and Verdú's classical result on the mutual information of peak-constrained channels [18, Corollary 2]. The converse follows by the well-known Gaussian max-entropy bound:

$$\mathsf{C}(\alpha_1, \alpha_2, \mathsf{A}) \le \max \frac{1}{2} \log\left(1 + \frac{\mathsf{Var}[X]}{\sigma^2}\right),$$

where the maximization is over random variables $X \in [0, 1]$ satisfying (2)–(4). Defining $T := X/\mathsf{A}$ and using that $\lim_{t\downarrow 0} \frac{\log(1+bt)}{t} = b$, for any constant $b > 0$, establishes the desired asymptotic converse bound. ∎

Lemma 3: The maximization in Proposition 2 is attained by a binary random variable $T \in \{0, \mathsf{A}\}$:

$$\max_{\substack{T\in[0,1]:\\ \mathsf{E}[T]\le\alpha_1\\ \mathsf{E}[T^2]\le\alpha_2}} \mathrm{Var}(T) = \max_{\substack{T\in\{0,\mathsf{A}\}:\\ \mathsf{E}[T]\le\alpha_1\\ \mathsf{E}[T^2]\le\alpha_2}} \mathrm{Var}(T) \tag{17}$$

Proof: Fix T satisfying the conditions in the minimization and construct a new random variable $T' \in \{0, \mathsf{A}\}$ with $p_\mathsf{A} := \Pr[T' = \mathsf{A}] = \frac{\mathsf{E}[T^2]}{\mathsf{A}^2}$ and $\Pr[T' = 0] = 1 - p_\mathsf{A}$. Notice that $\mathsf{E}\left[(T')^2\right] = p_\mathsf{A}\mathsf{A}^2 = \mathsf{E}\left[T^2\right]$ and

$$\mathsf{E}[T'] = p_\mathsf{A}\mathsf{A} = \frac{\mathsf{E}\left[T^2\right]}{\mathsf{A}} \le \frac{\mathsf{E}[T]\cdot\mathsf{A}}{\mathsf{A}} = \mathsf{E}[T]. \tag{18}$$

The new random variable T' thus also satisfies the conditions in the maximization, and moreover it has larger objective function (variance) than T because $\mathsf{Var}[T'] = \mathsf{E}\left[(T')^2\right] - (\mathsf{E}[T'])^2 \ge \mathsf{E}\left[(T)^2\right] - (\mathsf{E}[T])^2 = \mathsf{Var}[T]$. ∎

Combining Proposition 2 with Lemma 3 establishes the desired low-SNR asymptotics.

Theorem 4: For any parameters $\alpha_1, \alpha_2 > 0$:

$$\lim_{\mathsf{A}\downarrow 0} \frac{\mathsf{C}(\alpha_1, \alpha_2, \mathsf{A})}{\mathsf{A}^2} = p^*(1 - p^*), \tag{19}$$

where $p^* := \min\{\alpha_1, \alpha_2, 1/2\}$.

Proof: By Lemma 3:

$$\max_{\substack{T\in[0,1]:\\ \mathsf{E}[T]\le\alpha_1\\ \mathsf{E}[T^2]\le\alpha_2}} \mathrm{Var}(T) = \max_{\substack{p_\mathsf{A}\in[0,1]:\\ p_\mathsf{A}\le\alpha_1\\ p_\mathsf{A}\le\alpha_2}} p_\mathsf{A}(1 - p_\mathsf{A}). \tag{20}$$

Since the function $t \mapsto t(1 - t)$ is continuous and monotonically increasing over $[0, 1/2]$ but monotonically decreasing over $[1/2, 1]$, the maximum value is obtained for $p_\mathsf{A} = \min\{\alpha_1, \alpha_2, 1/2\}$. Plugging this into Proposition 2 establishes the desired result. ∎

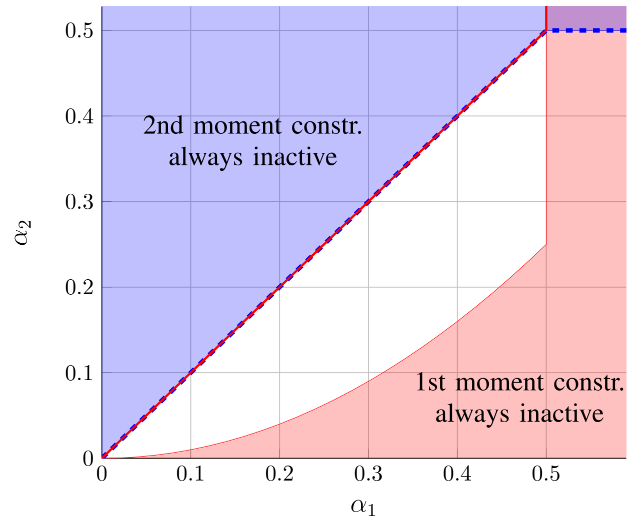

Fig. 1: The shaded regions illustrate the parameter values (α_1, α_2) where the two moment constraints (3) and (4) are inactive at any SNR value. In the *asymptotic low-SNR regime* the first-moment constraint is inactive on the right of the thin red solid line and the second-moment constraint is inactive above the thick blue dashed line.

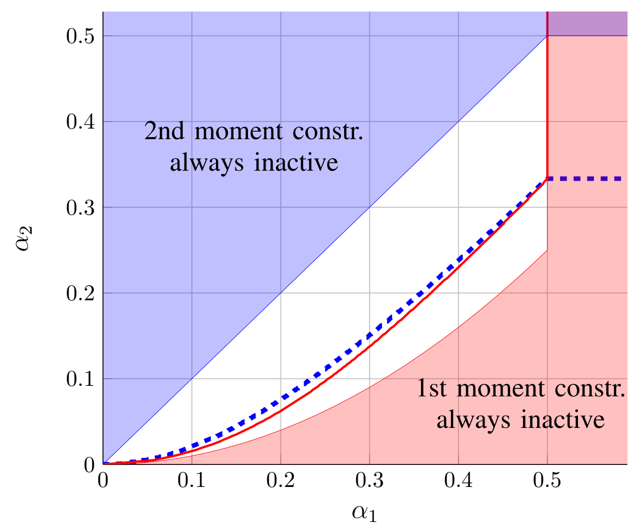

Fig. 2: The shaded regions indicate the parameter values where the two constraints are inactive for all SNR values. In the *in the asymptotic high-SNR regime* the first-moment constraint is inactive on the right of the thin red solid line and the second-moment constraint is inactive above the thick blue dashed line.

V. Discussion of Asymptotic Results

Figure 1 and Figure 2 both illustrate the regions of (α_1, α_2)-pairs where both the first- and the second-moment constraints, i.e., (3) and (4), are inactive *at all SNR values*: the first-moment constraint (3) is not active in the red shaded region and the second-moment constraint (4) is not active in the blue shaded region.

Figure 1 further shows the (α_1, α_2)-regions where the constraints are inactive in the *asymptotic low-SNR regime*. Specifically, in the low-SNR regime the first-moment constraint is not active on the right of the thin red solid line, i.e., whenever $\alpha_1 \ge \min\{\alpha_2, 1/2\}$, and the second-moment constraint is not active above the thick blue dashed line, i.e., whenever $\alpha_2 \ge \min\{\alpha_1, 1/2\}$. We see that the first-moment constraint is far less stringent at low SNR than for general SNR values.

Figure 2 shows the corresponding regions for the *asymptotic high-SNR regime*. In the high-SNR regime, the first-moment constraint is not active on the right of the thin red solid line , and the second-moment constraint is not active above the thick blue dashed line. Both constraints are simultaneously active only in the small region between these two lines. We observe that both constraints are less stringent in the high-SNR regime than for general SNR values, and the second-moment constraint is also less stringent than in the low-SNR regime. The first-moment constraint however is more stringent in the high-SNR regime than in the low-SNR regime.

VI. Proof of Theorem 1

A. Lower Bound

We first lower-bound the capacity with some simple entropy-manipulations and by using the entropy-maximizing input-density $f^*_X(x)$ over $[0, \mathsf{A}]$. Under constraints (2)–(4), $f^*_X(x)$ has the form:

$$f^*_X(x) = (\mathsf{A}\zeta_0(\lambda_1,\lambda_2))^{-1} \cdot e^{-\frac{\lambda_1}{\mathsf{A}}x - \frac{\lambda_2}{\mathsf{A}^2}x^2}, \quad x \in [0,\mathsf{A}], \tag{21}$$

where the parameters λ_1, λ_2 have to be chosen to satisfy

$$\int_0^{\mathsf{A}} f^*_X(x) \cdot x \, \mathrm{d}x \le \alpha_1 \mathsf{A}, \tag{22a}$$

$$\int_0^{\mathsf{A}} f^*_X(x) \cdot x^2 \, \mathrm{d}x \le \alpha_2 \mathsf{A}^2. \tag{22b}$$

Given the form in (21), through a simple variable substitution $y = \frac{x}{\mathsf{A}}$, one can prove that (22) are equivalent to

$$\frac{\zeta_1(\lambda_1,\lambda_2)}{\zeta_0(\lambda_1,\lambda_2)} \le \alpha_1, \tag{23a}$$

$$\frac{\zeta_2(\lambda_1,\lambda_2)}{\zeta_0(\lambda_1,\lambda_2)} \le \alpha_2, \tag{23b}$$

where recall that the functions ζ_i, for $i = 0, 1, \ldots, 4$, are defined in (10). Then,

$$\begin{aligned}
&\mathsf{C}(\alpha_1,\alpha_2,\mathsf{A}) \\
&\quad \ge I_{f^*_X}(X;Y) && (24)\\
&\quad = h_{f^*_X}(Y) - h(Z) \ge h_{f^*_X}(Y|Z) - h(Z) && (25)\\
&\quad = h_{f^*_X}(X) - h(Z) && (26)\\
&\quad = \mathsf{E}_{f^*_X}[-\log f^*_X(X)] - \frac{1}{2}\log(2\pi e\sigma^2) && (27)\\
&\quad = \log(\mathsf{A}\cdot\zeta_0(\lambda_1,\lambda_2)) + \frac{\lambda_1}{\mathsf{A}}\mathsf{E}_{f^*_X}[X] + \frac{\lambda_2}{\mathsf{A}^2}\mathsf{E}_{f^*_X}\left[X^2\right] \\
&\qquad -\frac{1}{2}\log(2\pi e\sigma^2) && (28)\\
&\quad = \log\left(\frac{\mathsf{A}\cdot\zeta_0(\lambda_1,\lambda_2)}{\sqrt{2\pi e\sigma^2}}\right) + \lambda_1\alpha_1 + \lambda_2\alpha_2, && (29)
\end{aligned}$$

where all (λ_1, λ_2) satisfying (23) yield valid lower bounds.

B. Upper bound

We turn to the duality-based upper bound with the choice of output density

$$f_Y(y) = \tau \cdot f^{(1)}_Y(y) + (1-\tau)\cdot f^{(2)}_Y(y), \tag{30}$$

where $\tau \in (0,1)$ is a parameter that we specify later on; $f^{(1)}_Y(y)$ is a probability density function over the interval $\mathcal{I} := [0, \mathsf{A}]$ of the form

$$f^{(1)}_Y(y) = \frac{1}{\mathsf{A}\cdot\zeta_0(\lambda_1,\lambda_2)} e^{-\frac{\lambda_1}{\mathsf{A}}y - \frac{\lambda_2}{\mathsf{A}}y^2} \cdot \mathbb{1}\{y \in \mathcal{I}\}, \tag{31}$$

where $\lambda_1, \lambda_2 \ge 0$ are free parameters, over which we will optimize in a latter stage; and $f^{(2)}_Y(y)$ is a probability density function over the rest of the real line $\mathcal{I}^c := \mathbb{R}\backslash\mathcal{I}$:

$$f^{(2)}_Y(y) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y^2}{2\sigma^2}} & \text{if } y < 0, \\ \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mathsf{A})^2}{2\sigma^2}} & \text{if } y > \mathsf{A}. \end{cases} \tag{32}$$

For the choice in (30), the duality-based upper bound yields

$$\begin{aligned}
&\mathsf{C}(\alpha_1,\alpha_2,\mathsf{A}) && (33)\\
&\le \mathsf{E}_{f^*_Y}[-\log f_Y(Y)] - \frac{1}{2}\log(2\pi e\sigma^2) && (34)\\
&\le \mathsf{E}_{f^*_Y}\left[-\log\Big(\tau f^{(1)}_Y(Y)\Big)\Big| Y \in \mathcal{I}\right]\cdot P_{f^*_Y}(\mathcal{I}) \\
&\quad + \mathsf{E}_{f^*_Y}\left[-\log\Big((1-\tau) f^{(2)}_Y(Y)\Big)\Big| Y \in \mathcal{I}^c\right]\cdot P_{f^*_Y}(\mathcal{I}^c) \\
&\quad - \frac{1}{2}\log(2\pi e\sigma^2) && (35)\\
&= \log\frac{\mathsf{A}\cdot\zeta_0(\lambda_1,\lambda_2)}{\tau}\cdot P_{f^*_Y}(\mathcal{I}) \\
&\quad + \mathsf{E}_{f^*_Y}\left[-\log\Big(\tau\cdot f^{(2)}_Y(Y)\Big)\Big| Y \in \mathcal{I}^c\right]\cdot P_{f^*_Y}(\mathcal{I}^c) \\
&\quad + \left(\frac{\lambda_1}{\mathsf{A}}\mathsf{E}_{f^*_Y}[Y|Y\in\mathcal{I}] + \frac{\lambda_2}{\mathsf{A}^2}\mathsf{E}_{f^*_Y}\left[Y^2|Y\in\mathcal{I}\right]\right)\cdot P_{f^*_Y}(\mathcal{I}) \\
&\quad - \frac{1}{2}\log(2\pi e\sigma^2). && (36)
\end{aligned}$$

Following similar steps as, e.g., in [19, Eq. (209)–(226)], we obtain the following lemmas proved in the extended version of this paper [20].

Lemma 5: For the Gaussian-tail distribution defined in (32):

$$\mathsf{E}_{f^*_Y}\left[-\log\Big(f^{(2)}_Y(Y)\Big)\Big| Y \in \mathcal{I}^c\right] \le \log\sqrt{2\pi e\sigma^2}. \tag{37}$$

Lemma 6: For the distribution in (31):

$$\begin{aligned}
\mathsf{E}_{f^*_Y}[Y|Y\in\mathcal{I}]\cdot P_{f^*_Y}(\mathcal{I}) &\le \mathsf{E}_{f^*_Y}[Y] + \frac{1}{\sqrt{2\pi}}\left(1 - e^{-\frac{\mathsf{A}^2}{2}}\right) && (38)\\
&= \mathsf{E}_{f^*_X}[X] + \left(1 - \frac{1}{\sqrt{2\pi}}e^{-\frac{\mathsf{A}^2}{2}}\right) && (39)
\end{aligned}$$

and

$$\mathsf{E}_{f^*_Y}\left[Y^2|Y\in\mathcal{I}\right]\cdot P_{f^*_Y}(\mathcal{I}) \le \mathsf{E}_{f^*_Y}\left[Y^2\right] = \mathsf{E}_{f^*_X}\left[X^2\right] + \sigma^2. \tag{40}$$

We continue with our upper bound. By plugging these lemmas into (36) and choosing

$$\tau = \frac{\mathsf{A}\cdot\zeta_0(\lambda_1,\lambda_2)}{\mathsf{A}\cdot\zeta_0(\lambda_1,\lambda_2) + \sqrt{2\pi e\sigma^2}}, \tag{41}$$

we obtain:

$$\begin{aligned}
&\mathsf{C}(\alpha_1,\alpha_2,\mathsf{A}) \\
&\quad \le \log\frac{\mathsf{A}\cdot\zeta_0(\lambda_1,\lambda_2)}{\tau}\cdot P_{f^*_Y}(\mathcal{I}) + \log\frac{\sqrt{2\pi e\sigma^2}}{1-\tau}\cdot P_{f^*_Y}(\mathcal{I}^c) \\
&\qquad + \frac{\lambda_1}{\mathsf{A}}\mathsf{E}_{f^*_X}[X] + \frac{\lambda_2}{\mathsf{A}^2}\left(\mathsf{E}_{f^*_X}\left[X^2\right] + \sigma^2\right) \\
&\qquad + \frac{\lambda_1}{\mathsf{A}}\left(1 - \frac{1}{\sqrt{2\pi}}e^{-\frac{\mathsf{A}^2}{2}}\right) - \frac{1}{2}\log(2\pi e\sigma^2) && (42)\\
&\quad = \log\left(\mathsf{A}\cdot\zeta_0(\lambda_1,\lambda_2) + \sqrt{2\pi e\sigma^2}\right)\cdot P_{f^*_Y}(\mathcal{I}) \\
&\qquad + \log\left(\mathsf{A}\cdot\zeta_0(\lambda_1,\lambda_2) + \sqrt{2\pi e\sigma^2}\right)\cdot P_{f^*_Y}(\mathcal{I}^c) \\
&\qquad + \frac{\lambda_1}{\mathsf{A}}\mathsf{E}_{f^*_X}[X] + \frac{\lambda_2}{\mathsf{A}^2}\left(\mathsf{E}_{f^*_X}\left[X^2\right] + \sigma^2\right) \\
&\qquad + \frac{\lambda_1}{\mathsf{A}}\left(1 - \frac{1}{\sqrt{2\pi}}e^{-\frac{\mathsf{A}^2}{2}}\right) - \frac{1}{2}\log(2\pi e\sigma^2) && (43)\\
&\quad \le \log\left(\mathsf{A}\cdot\zeta_0(\lambda_1,\lambda_2) + \sqrt{2\pi e\sigma^2}\right) + \lambda_1\alpha_1 + \lambda_2\alpha_2 + \lambda_2\frac{\sigma^2}{\mathsf{A}^2}
\end{aligned}$$

$$+\frac{\lambda_1}{\mathsf{A}}\left(1-\frac{1}{\sqrt{2\pi}}e^{-\frac{\mathsf{A}^2}{2}}\right)-\frac{1}{2}\log(2\pi e\sigma^2) \quad (44)$$

$$=\log\left(1+\frac{\mathsf{A}\cdot\zeta_0(\lambda_1,\lambda_2)}{2\pi e\sigma^2}\right)+\lambda_1\alpha_1+\lambda_2\alpha_2$$

$$+\lambda_2\frac{\sigma^2}{\mathsf{A}^2}+\frac{\lambda_1}{\mathsf{A}}\left(1-\frac{1}{\sqrt{2\pi}}e^{-\frac{\mathsf{A}^2}{2}}\right). \quad (45)$$

We can conclude that for any choice of $\lambda_1, \lambda_2 \geq 0$:

$$\overline{\lim_{\mathsf{A}\to\infty}}\left(\mathsf{C}(\alpha_1,\alpha_2,\mathsf{A})-\log\frac{\mathsf{A}}{\sqrt{2\pi e\sigma^2}}\right) \leq \log\zeta_0(\lambda_1,\lambda_2)+\lambda_1\alpha_1+\lambda_2\alpha_2. \quad (46)$$

C. Distinction of the Four Cases

We now show that the case distinction proposed in the theorem partitions the set of all (α_1, α_2)-parameters and that the described choice of λ_1^*, λ_2^*-parameters exists in each subset. More specifically, we show that the proposed case distinction coincides with the case distinction that arises when minimizing the right-hand side of (46), i.e., the function

$$\Gamma(\lambda_1,\lambda_2):=\log\zeta_0(\lambda_1,\lambda_2)+\lambda_1\alpha_1+\lambda_2\alpha_2, \quad (47)$$

over the choices $\lambda_1, \lambda_2 > 0$, and we show that the λ_1^*, λ_2^* values given in the theorem are the minimizers of this function.

Consider the partial derivatives of this function:

$$\frac{\partial\Gamma}{\partial\lambda_1}=-\frac{\zeta_1(\lambda_1,\lambda_2)}{\zeta_0(\lambda_1,\lambda_2)}+\alpha_1 \quad (48a)$$

and

$$\frac{\partial\Gamma}{\partial\lambda_2}=-\frac{\zeta_2(\lambda_1,\lambda_2)}{\zeta_0(\lambda_1,\lambda_2)}+\alpha_2, \quad (48b)$$

as well as its Hessian matrix

$$\mathbb{H}\Gamma(\lambda_1,\lambda_2):=\begin{pmatrix}\frac{\partial^2\Gamma(\lambda_1,\lambda_2)}{\partial\lambda_1^2} & \frac{\partial^2\Gamma(\lambda_1,\lambda_2)}{\partial\lambda_1\partial\lambda_2}\\ \frac{\partial^2\Gamma(\lambda_1,\lambda_2)}{\partial\lambda_1\partial\lambda_2} & \frac{\partial^2\Gamma(\lambda_1,\lambda_2)}{\partial\lambda_2^2}\end{pmatrix} \quad (49)$$

$$=\begin{pmatrix}\zeta_2(\lambda_1,\lambda_2)-\zeta_1^2(\lambda_1,\lambda_2) & c(\lambda_1,\lambda_2)\\ c(\lambda_1,\lambda_2) & \zeta_4(\lambda_1,\lambda_2)-\zeta_2^2(\lambda_1,\lambda_2)\end{pmatrix}, \quad (50)$$

where

$$c(\lambda_1,\lambda_2):=\zeta_3(\lambda_1,\lambda_2)-\zeta_1(\lambda_1,\lambda_2)\cdot\zeta_2(\lambda_1,\lambda_2). \quad (51)$$

Since for any pair (λ_1, λ_2) the Hessian $\mathbb{H}\Gamma(\lambda_1, \lambda_2)$ is a two-by-two matrix with positive trace and determinant, all its eigenvalues are positive, and the Hessian itself is positive definite for all (λ_1, λ_2). As a consequence, the function $\Gamma(\lambda_1, \lambda_2)$ is jointly strictly convex in both arguments and the minimizer $(\lambda_1^*, \lambda_2^*)$ of $\Gamma(\lambda_1, \lambda_2)$, for $\lambda_1, \lambda_2 \geq 0$ is accordingly obtained as follows, depending on the values of α_1 and α_2:

1) If both partial derivatives of Γ at the origin are strictly positive, i.e.,

$$-\frac{\zeta_1(0,0)}{\zeta_0(0,0)}+\alpha_1=-\frac{1}{2}+\alpha_1>0, \quad (52)$$

$$-\frac{\zeta_2(0,0)}{\zeta_0(0,0)}+\alpha_2=-\frac{1}{3}+\alpha_2>0, \quad (53)$$

then $\lambda_1^* = \lambda_2^* = 0$.

2) If for some $\lambda_1' > 0$ the partial derivatives of Γ satisfy

$$-\frac{\zeta_1(\lambda_1',0)}{\zeta_0(\lambda_1',0)}+\alpha_1=0, \quad (54)$$

$$-\frac{\zeta_2(\lambda_1',0)}{\zeta_0(\lambda_1',0)}+\alpha_2>0, \quad (55)$$

then $\lambda_1^* = \lambda_1'$ and $\lambda_2^* = 0$.

3) If for some $\lambda_2' > 0$ the partial derivatives of Γ satisfy

$$-\frac{\zeta_1(0,\lambda_2')}{\zeta_0(0,\lambda_2')}+\alpha_1>0 \quad (56)$$

$$-\frac{\zeta_2(0,\lambda_2')}{\zeta_0(0,\lambda_2')}+\alpha_2=0, \quad (57)$$

then $\lambda_1^* = 0$ and $\lambda_2^* = \lambda_2'$.

4) If for some $\lambda_1', \lambda_2' > 0$ the partial derivatives of Γ at (λ_1', λ_2') are both zero, i.e.,

$$-\frac{\zeta_1(\lambda_1',\lambda_2')}{\zeta_0(\lambda_1',\lambda_2')}+\alpha_1=0 \quad (58)$$

$$-\frac{\zeta_2(\lambda_1',\lambda_2')}{\zeta_0(\lambda_1',\lambda_2')}+\alpha_2=0, \quad (59)$$

then $\lambda_1^* = \lambda_1'$ and $\lambda_2^* = \lambda_2'$.

Since the strictly convex function $\Gamma(\lambda_1, \lambda_2)$ has exactly one minimizing pair, combined with continuity considerations, this concludes the proof of the theorem.

VII. Acknowledgement

The authors thank Lina Mroueh for interesting discussions. The work of M. Wigger was funded by the European Research Council (ERC) under the European Union's Horizon 2020 under grant agreement No 715111. The work of S. Ma was funded by the Young Talents of Xuzhou Science and Technology Plan Project under Grant KC19051, and by the open research fund of National Mobile Communications Research Laboratory, Southeast University(No.2021D02).

References

[1] A. Jovicic, J. Li, and T. Richardson, "Visible light communication: opportunities, challenges and the path to market," *IEEE Commun. Mag.*, vol. 51, no. 12, pp. 26–32, Dec. 2013.

[2] P. H. Pathak, X. Feng, P. Hu, and P. Mohapatra, "Visible light communication, networking, and sensing: a survey, potential and challenges," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2047–2077, Sept. 2015.

[3] "IEEE standard for local and metropolitan area networks–part 15.7: Short-range wireless optical communication using visible light," *IEEE Std 802.15.7-2011*, pp. 1–309, 2011.

[4] S. Ma, H. Li, Y. He, R. Yang, S. Lu, W. Cao, and S. Li, "Capacity bounds and interference management for interference channel in visible light communication networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 182–193, Jan. 2019.

[5] X. Ling, J. Wang, X. Liang, Z. Ding, and C. Zhao, "Offset and power optimization for DCO-OFDM in visible light communication systems," *IEEE Trans. Signal Process.*, vol. 64, no. 2, pp. 349–363, Jan. 2016.

[6] X. Huang, J. Shi, J. Li, Y. Wang, and N. Chi, "A Gb/s VLC transmission using hardware preequalization circuit," *IEEE Photon. Technol. Lett.*, vol. 27, no. 18, pp. 1915–1918, Sep. 2015.

[7] F. Che, L. Wu, B. Hussain, X. Li, and C. P. Yue, "A fully integrated IEEE 802.15.7 visible light communication transmitter with on-chip 8-W 85% efficiency boost LED driver," *J. Lightw. Technol.*, vol. 34, no. 10, pp. 2419–2430, 2016.

[8] S. Ma, R. Yang, H. Li, Z.-L. Dong, H. Gu, and S. Li, “Achievable rate with closed-form for SISO channel and broadcast channel in visible light communication networks,” *J. Lightw. Technol.*, vol. 35, no. 14, pp. 2778–2787, Jul. 2017.

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

[10] A. Chaaban, J. Morvan, and M. Alouini, “Free-space optical communications: capacity bounds, approximations, and a new sphere-packing perspective,” *IEEE Trans. Commun.*, vol. 64, no. 3, pp. 1176–1191, Mar. 2016.

[11] J. B. Wang, Q. S. Hu, J. Wang, M. Chen, and J. Y. Wang, “Tight bounds on channel capacity for dimmable visible light communications,” *J. Lightw. Technol.*, vol. 31, no. 23, pp. 3771–3779, Dec. 2013.

[12] T. Chan, S. Hranilovic, and F. Kschischang, “Capacity-achieving probability measure for conditionally Gaussian channels with bounded inputs,” *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2073–2088, Jun. 2005.

[13] R. Jiang, Z. Wang, Q. Wang, and L. Dai, “A tight upper bound on channel capacity for visible light communications,” *IEEE Commun. lett.*, vol. 20, no. 1, pp. 1089–7798, Jan. 2016.

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

[15] A. A. Farid and S. Hranilovic, “Channel capacity and non-uniform signaling for free-space optical intensity channels,” *IEEE J. Sel. Areas Commun.*, vol. 17, no. 9, pp. 1553–1563, Dec. 2009.

[16] K. Keykhosravi, E. Agrell, M. Secondini, and M. Karlsson, “When to use optical amplification in noncoherent transmission: An information-theoretic approach,” *IEEE Transactions on Communications*, vol. 68, no. 4, pp. 2438–2445, 2020.

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

[18] V. V. Prelov and S. Verdú, “Second-order asymptotics of mutual information,” *IEEE Trans. Inf. Theory*, vol. 50, no. 8, pp. 1567–1580, Aug. 2004.

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

[20] S. Ma and M. Wigger, “First- and second-moment constrained Gaussian channels,” *arXiv Preprint*, 2021.