On the Capacity of Free-Space Optical Intensity Channels

Amos Lapidoth
ETH Zurich
Zurich, Switzerland
Email: lapidoth@isi.ee.ethz.ch

Stefan M. Moser
National Chiao Tung University (NCTU)
Hsinchu, Taiwan
Email: stefan.moser@ieee.org

Michèle A. Wigger
ETH Zurich
Zurich, Switzerland
Email: wigger@isi.ee.ethz.ch

***Abstract*— New upper and lower bounds are presented on the capacity of the free-space optical intensity channel. This channel is characterized by inputs that are nonnegative (representing the transmitted optical intensity) and by outputs that are corrupted by additive white Gaussian noise (because in free space the disturbances arise from many independent sources). Due to battery and safety reasons the inputs are simultaneously constrained in both their average and peak power. For a fixed ratio of the average power to the peak power the difference between the upper and the lower bounds tends to zero as the average power tends to infinity, and the ratio of the upper and lower bounds tends to one as the average power tends to zero.**

The case where only an average-power constraint is imposed on the input is treated separately. In this case, the difference of the upper and lower bound tends to 0 as the average power tends to infinity, and their ratio tends to a constant as the power tends to zero.

I. Introduction

We consider a channel model for short optical communications in free space such as the communication between a remote control and a TV. We assume a channel model based on *intensity modulation* where the signal is modulated onto the optical intensity of the emitted light. Thus, the channel input is proportional to the light intensity and is therefore nonnegative. We further assume that the receiver directly measures the incident optical intensity of the incoming signal, *i.e.*, it produces an electrical current at the output which is proportional to the detected intensity. Since in ambient light conditions the received signal is disturbed by a high number of independent sources, we model the noise as Gaussian. Moreover, we assume that the line-of-sight component is dominant and ignore any effects due to multiple-path propagation like fading or inter-symbol interference.

Optical communication is restricted not only by battery power, but also for safety reasons by the maximum allowed peak power. We therefore assume simultaneously two constraints: an average-power constraint $\mathcal{E}$ and a maximum allowed peak power A. The situation where only a peak-power constraint is imposed corresponds to $\mathcal{E} = \mathsf{A}$. The case of only an average-power constraint is treated separately.

In this work we study the channel capacity of such an optical communication channel and present new upper and lower bounds. The maximum gap between upper and lower bound never exceeds 1 nat when the ratio of the average-power constraint to the peak-power constraint is larger than 0.03 or when only the average power is constrained but not the peak power. Asymptotically when the available average and peak power tend to infinity with their ratio held fixed, the upper and lower bounds coincide, *i.e.*, their difference tends to 0. We also present the asymptotic behavior of the channel capacity in the limit when the power tends to 0.

The channel model has been studied before, *e.g.*, in [1] and is described in detail in the following. The received signal is corrupted by additive noise due to strong ambient light that causes high-intensity shot noise in the electrical output signal. In a first approximation this shot noise can be assumed to be independent of the signal itself, and since the noise is caused by many independent sources it is reasonable to model it as an independent and identically distributed (IID) Gaussian process. Also, without loss of generality we assume the noise to be zero-mean, since the receiver can always subtract or add any constant signal.

Hence, the channel output Y_k at time k, modeling a sample of the electrical output signal, is given by

$$Y_k = x_k + Z_k, \tag{1}$$

where $x_k \in \mathbb{R}_0^+$ denotes the time-k channel input and represents a sample of the electrical input current that is proportional to the optical intensity and therefore nonnegative, and where the random process $\{Z_k\}$ modeling the additive noise is given by

$$\{Z_k\} \sim \text{IID } \mathcal{N}_{\mathbb{R}}\big(0, \sigma^2\big)\,. \tag{2}$$

It is important to note that, unlike the input x_k, the output Y_k may be negative since the noise introduced at the receiver can be negative.

Since the optical intensity is proportional to the optical power, in such a system the instantaneous optical power is proportional to the electrical input current [2]. This is in contrast to radio communication where usually the power is proportional to the square of the input current. Therefore, in addition to the implicit nonnegativity constraint on the input,

$$X_k \geq 0, \tag{3}$$

we assume constraints both on the peak and the average power, *i.e.*,

$$\Pr[X_k > \mathsf{A}] = 0, \tag{4}$$

$$\mathsf{E}[X_k] \leq \mathcal{E}. \tag{5}$$

We shall denote the ratio between the average power and the peak power by α,

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

where we assume $0 < \alpha \leq 1$. Note that $\alpha = 1$ corresponds to the case with only a peak-power constraint. Similarly, $\alpha \ll 1$ corresponds to a dominant average-power constraint and only a very weak peak-power constraint.

We denote the capacity of the described channel with peak-power constraint A and average-power constraint $\mathcal{E}$ by $\mathsf{C}(\mathsf{A}, \mathcal{E})$. The capacity is given by [3]

$$\mathsf{C}(\mathsf{A}, \mathcal{E}) = \sup I(X; Y) \tag{7}$$

where $I(X;Y)$ stands for the mutual information between the channel input X and the channel output Y, where conditional on the input x the output Y is Gaussian $\sim \mathcal{N}(x, \sigma^2)$; and where the supremum is over all laws on $X \geq 0$ satisfying $\Pr[X \geq \mathsf{A}] = 0$ and $\mathsf{E}[X] \leq \mathcal{E}$.

In the case of only an average-power constraint the capacity is denoted by $\mathsf{C}(\mathcal{E})$. It is given as in (7) except that the supremum is taken over all laws on $X \geq 0$ satisfying $\mathsf{E}[X] \leq \mathcal{E}$.

The derivation of the upper bounds is based on a technique introduced in [4]. There a dual expression of mutual information is used to show that for any channel law $W(\cdot|\cdot)$ and for an arbitrary distribution $R(\cdot)$ over the channel output alphabet, the channel capacity is upper-bounded by

$$\mathsf{C} \leq \mathsf{E}_{Q^*}\left[D\big(W(\cdot|X)\big\|R(\cdot)\big)\right]. \tag{8}$$

Here, $D(\cdot\|\cdot)$ stands for the relative entropy [5, Ch. 2], and $Q^*(\cdot)$ denotes the capacity-achieving input distribution. For more details about this technique and for a proof of (8), see [4, Sec. V], [6, Ch. 2]. The challenge of using (8) lies in a clever choice of the arbitrary law $R(\cdot)$ that will lead to a good upper bound. Moreover, note that the bound (8) still contains an expectation over the (unknown) capacity-achieving input distribution $Q^*(\cdot)$. To handle this expectation we will need to resort to some further bounding like, *e.g.*, Jensen's inequality [5, Ch. 2.6].

The derivation of the firm lower bounds relies on the entropy power inequality [5, Th. 17.7.3].

The asymptotic results at high power follow directly by evaluation of the firm upper and lower bounds. For the low power regime we introduce an additional lower bound which does not rely on the entropy power inequality. This lower bound is obtained by choosing a binary input distribution, a choice which was inspired by [7], and by then evaluating the corresponding mutual information. For the cases involving a peak-power constraint we further resort to the results on the asymptotic expression of mutual information for weak signals in [8].

The results of this paper are partially based on the results in [9] and [6, Ch. 3].

The remainder is structured as follows. In the subsequent section we state our results and in Section III we give a brief outline of some of the derivations.

II. Results

We start with an auxiliary lemma which is based on the symmetry of the channel law and the concavity of channel capacity in the input distribution.

Lemma 1: Consider a peak-power constraint A and an average-power constraint $\mathcal{E}$ such that $\alpha = \frac{\mathsf{A}}{\mathcal{E}} > \frac{1}{2}$. Then the optimal input distribution[1] in (7) has an average power equal to half the peak power

$$\frac{\mathsf{E}_{Q^*}[X]}{\mathsf{A}} = \frac{1}{2}, \tag{9}$$

irrespective of α. *I.e.*, the average-power constraint is inactive for all $\alpha \in \left(\frac{1}{2}, 1\right]$, and in particular

$$\mathsf{C}(\mathsf{A}, \alpha\mathsf{A}) = \mathsf{C}\left(\mathsf{A}, \frac{\mathsf{A}}{2}\right), \quad \frac{1}{2} < \alpha \leq 1. \tag{10}$$

We are now ready to state our results. We will distinguish between three different cases: in the first two cases we impose on the input both an average- and a peak-power constraint: in the first case the average-to-peak power ratio α is restricted to lie in $\left(0, \frac{1}{2}\right)$, and in the second case it is restricted to lie in $\left[\frac{1}{2}, 1\right]$. (Note that by Lemma 1, $\frac{1}{2} < \alpha \leq 1$ represents the situation with an inactive average-power constraint.) In the third case we impose on the input an average power constraint only.

In all three cases we present firm upper and lower bounds on the channel capacity. The difference of the upper and lower bounds tends to 0 when the available average and peak power tend to infinity with their ratio held constant at α. Thus we can derive the asymptotic capacity at high power exactly. We also present the asymptotics of capacity at low power: for the cases involving a peak-power constraint we are able to state them exactly, and for the case of only an average-power constraint we give the asymptotics up to a constant factor.

A. Bounds on Channel Capacity with both an Average- and a Peak-Power Constraint ($0 < \alpha < \frac{1}{2}$)

Theorem 2: If $0 < \alpha < \frac{1}{2}$, then $\mathsf{C}(\mathsf{A}, \alpha\mathsf{A})$ is lower-bounded by

$$\mathsf{C}(\mathsf{A}, \alpha\mathsf{A}) \geq \frac{1}{2}\log\left(1 + \mathsf{A}^2 \frac{e^{2\alpha\mu^*}}{2\pi e\sigma^2}\left(\frac{1-e^{-\mu^*}}{\mu^*}\right)^2\right), \tag{11}$$

and upper-bounded by each of the two bounds

$$\mathsf{C}(\mathsf{A}, \alpha\mathsf{A}) \leq \frac{1}{2}\log\left(1 + \alpha(1-\alpha)\frac{\mathsf{A}^2}{\sigma^2}\right), \tag{12}$$

$$\mathsf{C}(\mathsf{A}, \alpha\mathsf{A}) \leq \left(1 - \mathcal{Q}\left(\frac{\delta + \alpha\mathsf{A}}{\sigma}\right) - \mathcal{Q}\left(\frac{\delta + (1-\alpha)\mathsf{A}}{\sigma}\right)\right) \cdot \cdot \log\left(\frac{\mathsf{A}}{\sigma} \cdot \frac{e^{\frac{\mu\delta}{\mathsf{A}}} - e^{-\mu\left(1+\frac{\delta}{\mathsf{A}}\right)}}{\sqrt{2\pi}\mu\left(1 - 2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)}\right)$$

[1] It was shown in [1] that the optimal input distribution in (7) is unique.

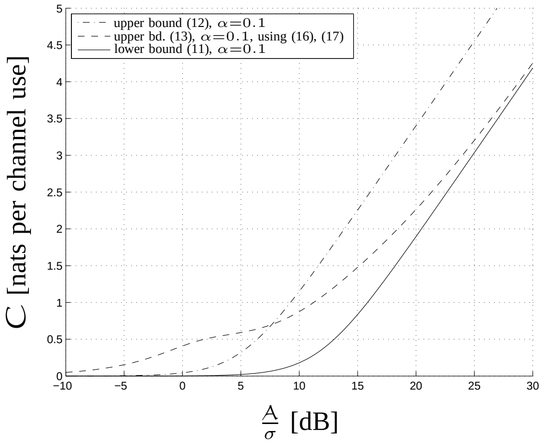

Fig. 1. Bounds of Theorem 2 for a choice of the average-to-peak power ratio $\alpha = 0.1$. The free parameters have been chosen as suggested in (16) and (17). The maximum gap between upper and lower bound is 0.72 nats (for $\mathsf{A}/\sigma \approx 11.8$ dB).

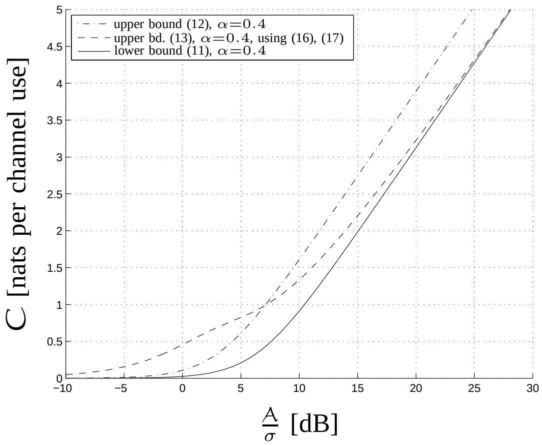

Fig. 2. Bounds of Theorem 2 for a choice of the average-to-peak power ratio $\alpha = 0.4$. The free parameters have been chosen as suggested in (16) and (17). The maximum gap between upper and lower bound is 0.56 nats (for $\mathsf{A}/\sigma \approx 7.1$ dB).

$$
\begin{aligned}
&-\frac{1}{2} + \mathcal{Q}\left(\frac{\delta}{\sigma}\right) + \frac{\delta}{\sqrt{2\pi}\sigma} e^{-\frac{\delta^2}{2\sigma^2}} \\
&+ \frac{\sigma}{\mathsf{A}} \frac{\mu}{\sqrt{2\pi}} \left(e^{-\frac{\delta^2}{2\sigma^2}} - e^{-\frac{(\mathsf{A}+\delta)^2}{2\sigma^2}} \right) \\
&+ \mu\alpha \left(1 - 2\mathcal{Q}\left(\frac{\delta + \frac{\mathsf{A}}{2}}{\sigma} \right) \right). \qquad (13)
\end{aligned}
$$

Here $\mathcal{Q}(\cdot)$ denotes the $\mathcal{Q}$-function defined by

$$
\mathcal{Q}(\xi) \triangleq \int_{\xi}^{\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}} \, \mathrm{d}t, \qquad \forall\, \xi \in \mathbb{R}; \qquad (14)
$$

$\mu > 0$ and $\delta > 0$ are free parameters; and μ^* is the unique solution of

$$
\alpha = \frac{1}{\mu^*} - \frac{e^{-\mu^*}}{1 - e^{-\mu^*}}. \qquad (15)
$$

Note that μ^* is well-defined as the function $\mu^* \mapsto \frac{1}{\mu^*} - \frac{e^{-\mu^*}}{1-e^{-\mu^*}}$ is strictly monotonically decreasing over $(0, \infty)$ and tends to $\frac{1}{2}$ for $\mu^* \downarrow 0$ and to 0 for $\mu^* \uparrow \infty$.

A suboptimal but useful choice for the free parameters in the upper bound (13) is

$$
\delta = \delta(\mathsf{A}) \triangleq \sigma \log\left(1 + \frac{\mathsf{A}}{\sigma}\right), \qquad (16)
$$

$$
\mu = \mu(\mathsf{A}, \alpha) \triangleq \mu^* \left(1 - e^{-\alpha \frac{\delta^2}{2\sigma^2}}\right), \qquad (17)
$$

where μ^* is the solution to (15). For this choice and for $\alpha = 0.1$ and 0.4 the bounds of Theorem 2 are depicted in Figures 1 and 2.

Theorem 3: If α lies in $\left(0, \frac{1}{2}\right)$, then

$$
\begin{aligned}
\chi(\alpha) &\triangleq \lim_{\mathsf{A}\uparrow\infty} \left\{ \mathsf{C}(\mathsf{A}, \alpha\mathsf{A}) - \log\frac{\mathsf{A}}{\sigma} \right\} \\
&= -\frac{1}{2}\log 2\pi e - (1-\alpha)\mu^* - \log(1 - \alpha\mu^*) \qquad (18)
\end{aligned}
$$

and

$$
\lim_{\mathsf{A}\downarrow 0} \frac{\mathsf{C}(\mathsf{A}, \alpha\mathsf{A})}{\mathsf{A}^2/\sigma^2} = \frac{\alpha(1-\alpha)}{2}. \qquad (19)
$$

B. Bounds on Channel Capacity with a Strong Peak-Power and Inactive Average-Power Constraint ($\frac{1}{2} \le \alpha \le 1$)

By Lemma 1 we have that for $\frac{1}{2} < \alpha \le 1$ the average-power constraint is inactive and $\mathcal{C}(\mathsf{A}, \alpha\mathsf{A}) = \mathcal{C}(\mathsf{A}, \frac{1}{2}\mathsf{A})$. Thus we can obtain the results in this section by simply deriving bounds for the case $\alpha = \frac{1}{2}$.

Theorem 4: If $\alpha \in \left[\frac{1}{2}, 1\right]$, then $\mathsf{C}(\mathsf{A}, \alpha\mathsf{A})$ is lower-bounded by

$$
\mathsf{C}(\mathsf{A}, \alpha\mathsf{A}) \ge \frac{1}{2}\log\left(1 + \frac{\mathsf{A}^2}{2\pi e \sigma^2}\right), \qquad (20)
$$

and is upper-bounded by each of the two bounds

$$
\mathsf{C}(\mathsf{A}, \alpha\mathsf{A}) \le \frac{1}{2}\log\left(1 + \frac{\mathsf{A}^2}{4\sigma^2}\right), \qquad (21)
$$

$$
\begin{aligned}
\mathsf{C}(\mathsf{A}, \alpha\mathsf{A}) \le &\left(1 - 2\mathcal{Q}\left(\frac{\delta + \frac{\mathsf{A}}{2}}{\sigma}\right)\right) \log \frac{\mathsf{A} + 2\delta}{\sigma\sqrt{2\pi}\left(1 - 2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)} \\
&- \frac{1}{2} + \mathcal{Q}\left(\frac{\delta}{\sigma}\right) + \frac{\delta}{\sqrt{2\pi}\sigma} e^{-\frac{\delta^2}{2\sigma^2}}, \qquad (22)
\end{aligned}
$$

where $\delta > 0$ is a free parameter.

We suboptimally choose

$$
\delta = \delta(\mathsf{A}) \triangleq \sigma \log\left(1 + \frac{\mathsf{A}}{\sigma}\right). \qquad (23)
$$

For this choice the bounds of Theorem 4 are depicted in Figure 3.

Theorem 5: If α lies in $\left[\frac{1}{2}, 1\right]$, then

$$
\chi(\alpha) \triangleq \lim_{\mathsf{A}\uparrow\infty} \left\{ \mathsf{C}(\mathsf{A}, \alpha\mathsf{A}) - \log\frac{\mathsf{A}}{\sigma} \right\} = -\frac{1}{2}\log 2\pi e \qquad (24)
$$

and

$$
\lim_{\mathsf{A}\downarrow 0} \frac{\mathsf{C}(\mathsf{A}, \alpha\mathsf{A})}{\mathsf{A}^2/\sigma^2} = \frac{1}{8}. \qquad (25)
$$

Note that (24) and (25) exhibit the well-known asymptotic behavior of the capacity of a Gaussian channel under a peak-power constraint only [3].

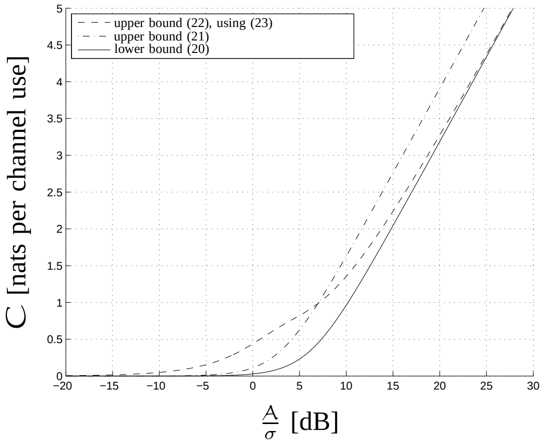

Fig. 3. Bounds on the capacity of the free-space optical intensity channel with average- and peak-power constraints for $\alpha \geq \frac{1}{2}$ according to Theorem 4. This includes the case of only a peak-power constraint $\alpha = 1$. The free parameter has been chosen as suggested in (23). The maximum gap between upper and lower bound is 0.54 nats (for $\mathsf{A}/\sigma \approx 7$ dB).

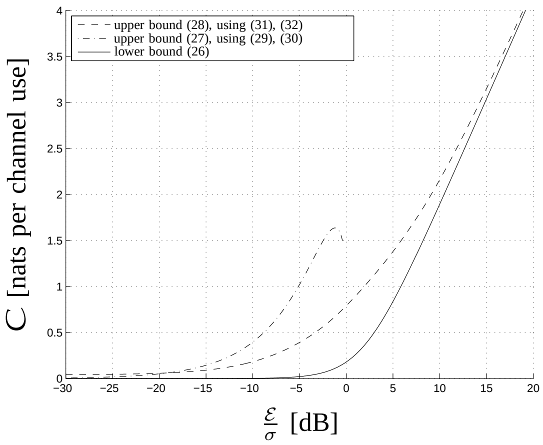

Fig. 4. Bounds on the capacity of the free-space optical intensity channel with only an average-power constraint according to Theorem 6. The free parameters have been chosen as suggested in (29)–(32). The maximum gap between upper and lower bound is 0.64 nats (for $\mathcal{E}/\sigma \approx 1.8$ dB).

C. Bounds on Channel Capacity with an Average-Power Constraint

Finally, we consider the case with an average-power constraint only.

Theorem 6: In the absence of a peak-power constraint the channel capacity $\mathsf{C}(\mathcal{E})$ is lower-bounded by

$$\mathsf{C}(\mathcal{E}) \geq \frac{1}{2}\log\left(1 + \frac{\mathcal{E}^2 e}{2\pi\sigma^2}\right), \tag{26}$$

and is upper-bounded by each of the bounds

$$\begin{aligned}\mathsf{C}(\mathcal{E}) \leq{}& \log\left(\beta e^{-\frac{\delta^2}{2\sigma^2}} + \sqrt{2\pi}\sigma\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right) - \log\left(\sqrt{2\pi}\sigma\right) \\ & - \frac{\delta\mathcal{E}}{2\sigma^2} + \frac{\delta^2}{2\sigma^2}\left(1 - \mathcal{Q}\left(\frac{\delta}{\sigma}\right) - \frac{\mathcal{E}}{\delta}\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right) \\ & + \frac{1}{\beta}\left(\mathcal{E} + \frac{\sigma}{\sqrt{2\pi}}\right), \qquad \delta \leq -\frac{\sigma}{\sqrt{e}},\end{aligned} \tag{27}$$

$$\begin{aligned}\mathsf{C}(\mathcal{E}) \leq{}& \log\left(\beta e^{-\frac{\delta^2}{2\sigma^2}} + \sqrt{2\pi}\sigma\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right) + \frac{1}{2}\mathcal{Q}\left(\frac{\delta}{\sigma}\right) \\ & + \frac{\delta}{2\sqrt{2\pi}\sigma}e^{-\frac{\delta^2}{2\sigma^2}} + \frac{\delta^2}{2\sigma^2}\left(1 - \mathcal{Q}\left(\frac{\delta + \mathcal{E}}{\sigma}\right)\right) \\ & + \frac{1}{\beta}\left(\delta + \mathcal{E} + \frac{\sigma}{\sqrt{2\pi}}e^{-\frac{\delta^2}{2\sigma^2}}\right) - \frac{1}{2}\log 2\pi e\sigma^2, \\ & \delta \geq 0,\end{aligned} \tag{28}$$

where $\beta > 0$ and δ are free parameters. Note that bound (27) only holds for $\delta \leq -\sigma e^{-\frac{1}{2}}$, while bound (28) only holds for $\delta \geq 0$.

A suboptimal but useful choice for the free parameters in bound (27) is shown in (29) and (30) and for the free parameters in bound (28) is shown in (31) and (32) at the top of the next page. For these choices, the bounds of Theorem 6 are depicted in Figure 4.

Theorem 7: In the case of only an average-power constraint,

$$\chi_{\mathcal{E}} \triangleq \lim_{\mathcal{E}\uparrow\infty}\left\{\mathsf{C}(\mathcal{E}) - \log\frac{\mathcal{E}}{\sigma}\right\} = \frac{1}{2}\log\frac{e}{2\pi} \tag{33}$$

and

$$\overline{\lim_{\mathcal{E}\downarrow 0}}\, \frac{\mathsf{C}(\mathcal{E})}{\frac{\mathcal{E}}{\sigma}\sqrt{\log\frac{\sigma}{\mathcal{E}}}} \leq 2, \tag{34}$$

$$\underline{\lim_{\mathcal{E}\downarrow 0}}\, \frac{\mathsf{C}(\mathcal{E})}{\frac{\mathcal{E}}{\sigma}\sqrt{\log\frac{\sigma}{\mathcal{E}}}} \geq \frac{1}{\sqrt{2}}. \tag{35}$$

Note that the asymptotic upper and lower bound at low SNR do not coincide in the sense that their ratio equals $2\sqrt{2}$ instead of 1. However, they exhibit similar behavior.

III. Derivation

In the following we will outline the derivations of the firm lower and upper bounds given in the previous section.

One easily finds a lower bound on capacity by dropping the maximization and choosing an arbitrary input distribution $Q(\cdot)$ to compute the mutual information between input and output. However, in order to get a tight bound, this choice of $Q(\cdot)$ should yield a mutual information that is reasonably close to capacity. Such a choice is difficult to find and might make the evaluation of $I(X;Y)$ intractable. The reason for this is that even for relatively "easy" distributions $Q(\cdot)$, the corresponding distribution on the channel output Y may be difficult to compute, let alone $h(Y)$. We circumvent these problems by using the entropy power inequality [5, Th. 17.7.3] to lower-bound $h(Y)$ by an expression that depends only on $h(X)$. *I.e.*, we "transfer" the problem of computing (or bounding) $h(Y)$ to the input side of the channel, where it is much easier to choose an appropriate distribution that leads to a tight lower bound on channel capacity:

$$\begin{aligned}\mathsf{C} &= \sup_{Q(\cdot)} I(X;Y) && (36)\\ &\geq I(X;Y)\big|_{\text{for a specific } Q(\cdot)} && (37)\end{aligned}$$

$$\delta = \delta(\mathcal{E}) \triangleq -2\sigma\sqrt{\log\frac{\sigma}{\mathcal{E}}}, \qquad \text{for } \frac{\mathcal{E}}{\sigma} \leq e^{-\frac{1}{4e}} \approx -0.4 \text{ dB}, \tag{29}$$

$$\beta = \beta(\mathcal{E}) \triangleq \frac{1}{2}\left(\mathcal{E} + \frac{\sigma}{\sqrt{2\pi}}\right) + \frac{1}{2}\sqrt{\left(\mathcal{E} + \frac{\sigma}{\sqrt{2\pi}}\right)^2 + 4\left(\mathcal{E} + \frac{\sigma}{\sqrt{2\pi}}\right)\sqrt{2\pi}\sigma e^{\frac{\delta^2}{2\sigma^2}}\mathcal{Q}\left(\frac{\delta}{\sigma}\right)}, \tag{30}$$

$$\delta = \delta(\mathcal{E}) \triangleq \sigma \log\left(1 + \frac{\mathcal{E}}{\sigma}\right), \tag{31}$$

$$\beta = \beta(\mathcal{E}) \triangleq \frac{1}{2}\left(\delta + \mathcal{E} + \frac{\sigma}{\sqrt{2\pi}}e^{-\frac{\delta^2}{2\sigma^2}}\right) + \frac{1}{2}\sqrt{\left(\delta + \mathcal{E} + \frac{\sigma}{\sqrt{2\pi}}e^{-\frac{\delta^2}{2\sigma^2}}\right)^2 + 4\left(\delta + \mathcal{E} + \frac{\sigma}{\sqrt{2\pi}}e^{-\frac{\delta^2}{2\sigma^2}}\right)\sqrt{2\pi}\sigma e^{\frac{\delta^2}{2\sigma^2}}\mathcal{Q}\left(\frac{\delta}{\sigma}\right)}. \tag{32}$$

$$= \big(h(Y) - h(Y|X)\big)\Big|_{\text{for a specific } Q(\cdot)} \tag{38}$$

$$= h(X+Z)\Big|_{\text{for a specific } Q(\cdot)} - h(Z) \tag{39}$$

$$\geq \frac{1}{2}\log\left(e^{2h(X)} + e^{2h(Z)}\right)\Bigg|_{\text{for a specific } Q(\cdot)} - h(Z) \tag{40}$$

$$= \frac{1}{2}\log\left(1 + \frac{e^{2h(X)}}{2\pi e\sigma^2}\right)\Bigg|_{\text{for a specific } Q(\cdot)} \tag{41}$$

where the inequality in (40) follows from the entropy power inequality. To make this lower bound as tight as possible we will choose a distribution $Q(\cdot)$ that maximizes differential entropy under the given constraints [5, Ch. 12].

The derivation of the upper bounds in Section II are based on the duality approach (8). Hence, we need to specify a distribution $R(\cdot)$ and evaluate the relative entropy in (8).

We have chosen output distributions $R(\cdot)$ with the following densities. For (12) we choose

$$R'(y) \triangleq \frac{1}{\sqrt{2\pi\left(\sigma^2 + \mathcal{E}(\mathsf{A} - \mathcal{E})\right)}} e^{-\frac{(y-\mathcal{E})^2}{2\sigma^2 + 2\mathcal{E}(\mathsf{A}-\mathcal{E})}}; \tag{42}$$

for (13) we choose

$$R'(y) \triangleq \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}}, & y < -\delta, \\ \frac{1}{\mathsf{A}} \cdot \frac{\mu\left(1 - 2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)}{e^{\frac{\mu\delta}{\mathsf{A}}} - e^{-\mu\left(1 + \frac{\delta}{\mathsf{A}}\right)}} e^{-\frac{\mu y}{\mathsf{A}}}, & -\delta \leq y \leq \mathsf{A} + \delta, \\ \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mathsf{A})^2}{2\sigma^2}}, & y > \mathsf{A} + \delta, \end{cases} \tag{43}$$

where $\delta > 0$ and $\mu > 0$ are free parameters; for (21) we choose

$$R'(y) \triangleq \frac{1}{\sqrt{2\pi\left(\sigma^2 + \frac{\mathsf{A}^2}{4}\right)}} e^{-\frac{\left(y - \frac{\mathsf{A}}{2}\right)^2}{2\sigma^2 + \frac{\mathsf{A}^2}{2}}}; \tag{44}$$

for (22) we choose

$$R'(y) \triangleq \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}}, & y < -\delta, \\ \frac{1 - 2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)}{\mathsf{A} + 2\delta}, & -\delta \leq y \leq \mathsf{A} + \delta, \\ \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mathsf{A})^2}{2\sigma^2}}, & y > \mathsf{A} + \delta, \end{cases} \tag{45}$$

where $\delta > 0$ is a free parameter; and for (27) and (28) we choose

$$R'(y) \triangleq \begin{cases} \frac{1}{\beta e^{-\frac{\delta^2}{2\sigma^2}} + \sqrt{2\pi}\sigma\mathcal{Q}\left(\frac{\delta}{\sigma}\right)} e^{-\frac{y^2}{2\sigma^2}}, & y < -\delta, \\ \frac{1}{\beta e^{-\frac{\delta^2}{2\sigma^2}} + \sqrt{2\pi}\sigma\mathcal{Q}\left(\frac{\delta}{\sigma}\right)} e^{-\frac{\delta^2}{2\sigma^2}} e^{-\frac{y+\delta}{\beta}}, & y \geq -\delta, \end{cases} \tag{46}$$

where $\delta \in \mathbb{R}$ and $\beta > 0$ are free parameters. In the derivation of (28) we then restrict δ to be nonnegative, while in the derivation of (27) we restrict $\delta \leq -\sigma e^{-\frac{1}{2}}$.

ACKNOWLEDGMENTS

The authors would like to thank Ding-Jie Lin for fruitful discussions. The work of S. M. Moser was supported in part by the ETH under TH-23 02-2.

REFERENCES

[1] T. H. Chan, S. Hranilovic, and F. R. 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.

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

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

[4] 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.

[5] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, 2006.

[6] S. M. Moser, "Duality-based bounds on channel capacity," Ph.D. dissertation, Swiss Fed. Inst. of Technol. (ETH), Zurich, Oct. 2004, Diss. ETH No. 15769. [Online]. Available: http://moser.cm.nctu.edu.tw/

[7] J. G. Smith, "The information capacity of amplitude and variance-constrained scalar Gaussian channels," Inform. Contr., vol. 18, pp. 203–219, 1971.

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

[9] M. A. Wigger, "Bounds on the capacity of free-space optical intensity channels," Master's thesis, Signal and Inf. Proc. Lab., ETH Zurich, Switzerland, Mar. 2003, supervised by Prof. Dr. Amos Lapidoth.