On the Capacity of Free-Space Optical Intensity Channels

Amos Lapidoth, *Fellow, IEEE,* Stefan M. Moser, *Member, IEEE,* and Michèle A. Wigger, *Member, IEEE*

***Abstract*—Upper and lower bounds are derived on the capacity of the free-space optical intensity channel. This channel has a nonnegative input (representing the transmitted optical intensity), which is corrupted by additive white Gaussian noise. To preserve the battery and for safety reasons, the input is constrained in both its average and its peak power. For a fixed ratio of the allowed average power to the allowed peak power, the difference between the upper and the lower bound tends to zero as the average power tends to infinity and their ratio tends to one as the average power tends to zero.**

When only an average power constraint is imposed on the input, the difference between the bounds tends to zero as the allowed average power tends to infinity, and their ratio tends to a constant as the allowed average power tends to zero.

***Index Terms*—Channel capacity, direct detection, Gaussian noise, infrared communication, optical communication, PAM.**

I. INTRODUCTION

We consider a channel model for short-range optical communication in free space such as the infrared communication between electronic handheld devices. We assume a channel model based on *intensity modulation*, where the input signal modulates the optical intensity of the emitted light. Thus, the input signal is proportional to the light intensity and is therefore nonnegative. We further assume that at the receiver a front-end photodetector measures the incident optical intensity of the incoming light and produces an output signal which is proportional to the detected intensity, corrupted by white Gaussian noise. To preserve the battery and for safety reasons, we restrict both the average and the peak intensity of the input signal.[1]

This channel model, also known as the *free-space optical intensity channel* or *optical direct-detection channel with Gaussian post-detection noise*, was previously studied in [1], [2], [3], [4], and [5]. In [4], [5] upper and lower bounds on this channel's capacity were derived. A related channel model for optical communication is the Poisson channel: see [6], [7], [8], [2] for the discrete-time channel and [9], [10], [11], [12], [13], [14], [15] for the continuous-time channel. A variation of the free-space optical intensity channel where the noise depends on the input is investigated in [2, Chapter 4], [16].

In this paper we present new upper and lower bounds on the capacity of the free-space optical intensity channel and study the capacity's asymptotic behavior at high and low powers. The maximum gap between the upper and lower bounds never exceeds 1 nat when the ratio of the allowed average power to the allowed peak power is larger than 0.03 or when only an average-power constraint is imposed. If both average- and peak-power constraints are imposed, then the bounds are asymptotically tight both at high and low input powers in the following sense: if the ratio of the allowed average power to the allowed peak power is held fixed, then the difference between the bounds tends to zero as the allowed average power tends to infinity, and the ratio of the bounds tends to one as the allowed average power tends to zero. If only an average-power constraint is imposed, then the difference between the bounds tends to zero as the allowed average power tends to infinity and the ratio tends to $2\sqrt{2}$ as the allowed average power tends to zero.

The derivation of the upper bounds is based on a general technique introduced in [17], [2], [18] using a dual expression for channel capacity. We will not state it in its full generality but only in the form needed in this paper. For more details and for a proof see [17, Sec. V], [2, Ch. 2].

Proposition 1: Consider a memoryless channel with input alphabet $\mathcal{X} = \mathbb{R}_0^+$ and output alphabet $\mathcal{Y} = \mathbb{R}$, where, conditional on the input $x \in \mathcal{X}$, the distribution on the output Y is denoted by the probability measure $W(\cdot|x)$.[2] Then, for any distribution $R(\cdot)$ on $\mathcal{Y}$, the channel capacity under a peak-power constraint $\mathcal{A}$ and an average-power constraint $\mathcal{E}$ is upper-bounded by

$$\mathcal{C}(\mathcal{A},\mathcal{E}) \leq \sup_{Q} \mathsf{E}_Q\big[D\big(W(\cdot|X)\big\|R(\cdot)\big)\big] \tag{1}$$

where the supremum is over all probability laws Q on the input X satisfying $Q(X > \mathcal{A}) = 0$ and $\mathsf{E}_Q[X] \leq \mathcal{E}$. Here, $D(\cdot\|\cdot)$ denotes relative entropy [19, Ch. 2].

Manuscript received June 19, 2008; revised June 12, 2009. Current version published September 23, 2009. The work of S. M. Moser was supported in part by ETH under Grant TH-23 02-2 and in part by the National Science Council, Taiwan, under Grant NSC 96-2221-E-009-012-MY3. The material in this paper was presented in part at the 2005 Winter School on Coding and Information Theory, Bratislava, Slovakia, and at the IEEE International Symposium on Information Theory (ISIT), Toronto, ON, Canada, July 2008.

A. Lapidoth is with the Department of Information Technology and Electrical Engineering, ETH Zurich, 8092 Zurich, Switzerland (e-mail: amos.lapidoth@isi.ee.ethz.ch).

S. M. Moser is with the Department of Electrical Engineering, National Chiao Tung University (NCTU), Hsinchu, Taiwan (e-mail: stefan.moser@ieee.org).

M. A. Wigger was with the Department of Information Technology and Electrical Engineering, ETH Zurich, 8092 Zurich, Switzerland. She is now with the Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla 92093 USA (e-mail: wigger@ece.ucsd.edu).

Communicated by I. Kontoyiannis, Associate Editor for Shannon Theory.

Digital Object Identifier 10.1109/TIT.2009.2027522

[1]For more details on the channel model see Section II.

[2]The proposition requires certain measurability assumptions on the law $W(\cdot|\cdot)$ which we omit for simplicity. However, the channel law under consideration satisfies these assumptions.

There are two challenges in using (1). The first is in finding a clever choice of the law R that will lead to a good upper bound. The second is in upper-bounding the supremum on the right-hand side of (1). To handle this second challenge we shall resort to some further bounding, *e.g.*, Jensen's inequality.

To derive the lower bounds we apply two different techniques: in the high-power regime we use the Entropy Power Inequality (EPI) [19, Theorem 17.7.3] and the theory of entropy maximizing distributions [19, Ch. 12]. The so-derived bounds are asymptotically tight. At low powers we lower-bound capacity by the mutual information corresponding to binary signaling (a choice which was inspired by [20] and [3]). When a peak-power constraint is present, the analysis of the mutual information is based on techniques from [21] and the resulting bounds are tight. When only an average-power constraint is imposed, we further lower-bound the asymptotic expression for the mutual information. In this regime our analysis does not yield the asymptotic capacity expansion, because the ratio of our upper bound to our lower bound does not tend to one as the allowed average power tends to zero; it tends to $2\sqrt{2}$.

The results of this paper are partially based on the results in [22] and [2, Ch. 3].

The rest of the paper is structured as follows. After some remarks about notation at the end of this section, we describe the channel model in detail in Section II. In Section III we state our main results, *i.e.*, the upper and lower bounds on channel capacity and the asymptotic results. The detailed derivations can be found in the appendix.

For random quantities we use uppercase letters and for their realizations lowercase letters. Scalars are typically denoted using Greek letters or lowercase Roman letters. A few exceptions are the following symbols: $\mathcal{C}$ stands for capacity, $D(\cdot\|\cdot)$ denotes the relative entropy between two probability measures, and $I(\cdot;\cdot)$ stands for the mutual information. Moreover, the capitals Q, W, and R denote probability measures:

- $Q(\cdot)$ denotes a generic probability measure on the channel input;
- for any input $x \in \mathcal{X}$, $W(\cdot|x)$ represents a probability measure on the channel output when the channel input is x;
- $R(\cdot)$ denotes a generic probability measure on the channel output.

The expression $I(Q,W)$ stands for the mutual information between input X and output Y of a channel with transition probability measure W when the input has distribution Q, *i.e.*, $I(Q,W) \triangleq I(X;Y)$. The symbol $\mathcal{E}$ denotes average power and $\mathcal{A}$ stands for peak power. All rates specified in this paper are in nats per channel-use, and all logarithms are natural logarithms.

II. Channel Model

In free-space optical communication the input signal is usually transmitted by means of light emitting diodes (LED) or laser diodes (LD). Conventional and most inexpensive diodes emit infrared light of wavelength between 850 and 950 nanometers. For such high frequencies, practical systems often apply intensity modulation where the transmitter modulates the optical intensity of the emitted light, and hence the input signal is proportional to the optical intensity. The receiver first measures the incident optical intensity of the incoming light by means of a front-end photodetector and produces an output signal which is proportional to the detected intensity. Based on this output signal the receiver decodes the transmitted data.

For our model we assume three main sources of noise: thermal noise in the receiver which is well-modeled by a Gaussian distribution; Relative-Intensity Noise (RIN) that models random intensity fluctuations inherent to low-cost laser sources and that also can be assumed to be Gaussian; and shot noise caused by the ambient light. Note that the shot noise only has an impact at large intensities where its distribution will tend to be Gaussian and in a first approximation can be assumed to be independent of the signal itself. At low intensity the thermal noise is the limiting factor. The sum of these three noise sources can be well-modeled by independent and memoryless additive Gaussian noise, *i.e.*, the channel output Y is given by

$$Y = x + Z \tag{2}$$

where x denotes the channel input that is proportional to the optical intensity and therefore cannot be negative,

$$x \in \mathbb{R}_0^+ \tag{3}$$

and where the additive noise is zero-mean Gaussian with variance σ^2. It is important to note that, unlike the input, the output Y may be negative since the noise introduced at the receiver can be negative. See [23], [24] for more on this model.

Note that we neglect random components of the ambient light. Such random components also could be modeled to be Gaussian, however, as they are added to the signal before the detection, their resulting distribution is non-Gaussian. This model—sometimes called *optical direct-detection channel with Gaussian pre-detection noise* [23] and described by $Y = |\sqrt{x} + Z|^2$—is not part of this investigation. But see [25] and [26]. Similarly, we assume that there is no optical pre-amplifier at the receiver which would have a similar effect. However note that at high power the pre-detection noise model will tend to our post-detection noise model (2).

For safety reasons and practical implementation considerations, the optical average power and maximum power must be constrained:

$$\Pr[X > \mathcal{A}] = 0 \tag{4}$$

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

for some fixed parameters $\mathcal{E}$, $\mathcal{A} > 0$. We refer to $\mathcal{E}$ as the allowed average power and to $\mathcal{A}$ as the allowed peak power. Note that the average-power constraint is on the expectation of the channel input and not on its square.

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

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

where $0 < \alpha \leq 1$. Note that for $\alpha = 1$ the average-power constraint is inactive in the sense that it has no influence on the capacity and is automatically satisfied whenever the peak power constraint is satisfied. Thus, $\alpha = 1$ 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 allowed peak power $\mathcal{A}$ and allowed average power $\mathcal{E}$ by $\mathcal{C}(\mathcal{A}, \mathcal{E})$. The capacity is given by [27]

$$\mathcal{C}(\mathcal{A}, \mathcal{E}) = \sup_{Q} I(Q, W) \tag{7}$$

where the supremum is over all laws Q on $X \geq 0$ satisfying $Q(X > \mathcal{A}) = 0$ and $\mathsf{E}_Q[X] \leq \mathcal{E}$.

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

Note that, as shown in [3] using the techniques of [20], the capacity-achieving input distribution to our channel (2) is discrete. But this is not critical to our analysis.

III. Results

Subject to (4) and (5), our channel (2) has a unique capacity-achieving input distribution, which we denote by Q^* [3]. Using this observation, the symmetry of the channel law, and the concavity of mutual information in the input distribution, we have:

Lemma 2: If the allowed average power $\mathcal{E}$ is larger than half the allowed peak power $\mathcal{A}$, then the optimal input distribution Q^* in (7) satisfies

$$\mathsf{E}_{Q^*}[X] = \frac{1}{2}\mathcal{A}. \tag{8}$$

Thus,

$$\mathcal{C}(\mathcal{A}, \alpha\mathcal{A}) = \mathcal{C}\left(\mathcal{A}, \frac{\mathcal{A}}{2}\right), \quad \frac{1}{2} < \alpha \leq 1. \tag{9}$$

Proof: See Appendix D. ■

To state our results we distinguish between three cases:

- Case I: both an average- and a peak-power constraint are imposed, with $\alpha \in \left(0, \frac{1}{2}\right)$;
- Case II: both an average- and a peak-power constraint are imposed, with $\alpha \in \left[\frac{1}{2}, 1\right]$;
- Case III: only an average-power constraint is imposed.

We present firm upper and lower bounds on the channel capacity in all three cases. In all three cases their difference tends to zero as the allowed average power tends to infinity, thus revealing the asymptotic capacity at high power. We also present results on the asymptotic capacity at low power. In cases I and II our results are precise: we present asymptotic upper and lower bounds whose ratio tends to 1 as the power tends to 0. For case III we present asymptotic upper and lower bounds whose ratio tends to $2\sqrt{2}$ as the power tends to 0.

A. Bounds on Channel Capacity for Case I

Theorem 3 (Bounds): If $0 < \alpha < \frac{1}{2}$, then $\mathcal{C}(\mathcal{A}, \alpha\mathcal{A})$ is lower-bounded by

$$\mathcal{C}(\mathcal{A}, \alpha\mathcal{A}) \geq \frac{1}{2}\log\left(1 + \mathcal{A}^2 \frac{e^{2\alpha\mu^*}}{2\pi e\sigma^2}\left(\frac{1 - e^{-\mu^*}}{\mu^*}\right)^2\right) \tag{10}$$

and upper-bounded by each of the two bounds

$$\mathcal{C}(\mathcal{A}, \alpha\mathcal{A}) \leq \frac{1}{2}\log\left(1 + \alpha(1-\alpha)\frac{\mathcal{A}^2}{\sigma^2}\right) \tag{11}$$

$$\begin{aligned}\mathcal{C}(\mathcal{A}, \alpha\mathcal{A}) \leq{}& \left(1 - \mathcal{Q}\left(\frac{\delta + \alpha\mathcal{A}}{\sigma}\right) - \mathcal{Q}\left(\frac{\delta + (1-\alpha)\mathcal{A}}{\sigma}\right)\right) \\ &\cdot \log\left(\frac{\mathcal{A}}{\sigma} \cdot \frac{e^{\frac{\mu\delta}{\mathcal{A}}} - e^{-\mu\left(1 + \frac{\delta}{\mathcal{A}}\right)}}{\sqrt{2\pi}\mu\left(1 - 2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)}\right) \\ &- \frac{1}{2} + \mathcal{Q}\left(\frac{\delta}{\sigma}\right) + \frac{\delta}{\sqrt{2\pi}\sigma}e^{-\frac{\delta^2}{2\sigma^2}} \\ &+ \frac{\sigma}{\mathcal{A}}\frac{\mu}{\sqrt{2\pi}}\left(e^{-\frac{\delta^2}{2\sigma^2}} - e^{-\frac{(\mathcal{A}+\delta)^2}{2\sigma^2}}\right) \\ &+ \mu\alpha\left(1 - 2\mathcal{Q}\left(\frac{\delta + \frac{\mathcal{A}}{2}}{\sigma}\right)\right).\end{aligned} \tag{12}$$

Here, $\mu > 0$ and $\delta > 0$ are free parameters, and μ^* is the unique solution to

$$\alpha = \frac{1}{\mu^*} - \frac{e^{-\mu^*}}{1 - e^{-\mu^*}}. \tag{13}$$

It is straightforward to show that the solution to (13) always exists and is unique.

A suboptimal but useful choice of the free parameters in (12) is

$$\delta = \sigma\log\left(1 + \frac{\mathcal{A}}{\sigma}\right) \tag{14}$$

$$\mu = \mu^*\left(1 - e^{-\alpha\frac{\delta^2}{2\sigma^2}}\right) \tag{15}$$

where μ^* is the solution to (13).

Figures 1 and 2 depict the bounds of Theorem 3 for $\alpha = 0.1$ and 0.4, where (12) is numerically minimized over $\delta, \mu > 0$.

Corollary 4 (Asymptotics): If $0 < \alpha < \frac{1}{2}$, then

$$\begin{aligned}\lim_{\mathcal{A}\uparrow\infty}&\left\{\mathcal{C}(\mathcal{A}, \alpha\mathcal{A}) - \log\frac{\mathcal{A}}{\sigma}\right\} \\ &= -\frac{1}{2}\log 2\pi e - (1-\alpha)\mu^* - \log(1 - \alpha\mu^*)\end{aligned} \tag{16}$$

and

$$\lim_{\mathcal{A}\downarrow 0}\frac{\mathcal{C}(\mathcal{A}, \alpha\mathcal{A})}{\mathcal{A}^2/\sigma^2} = \frac{\alpha(1-\alpha)}{2}. \tag{17}$$

B. Bounds on Channel Capacity for Case II

Theorem 5 (Bounds): If $\frac{1}{2} \leq \alpha \leq 1$, then $\mathcal{C}(\mathcal{A}, \alpha\mathcal{A})$ is lower-bounded by

$$\mathcal{C}(\mathcal{A}, \alpha\mathcal{A}) \geq \frac{1}{2}\log\left(1 + \frac{\mathcal{A}^2}{2\pi e\sigma^2}\right) \tag{18}$$

and is upper-bounded by each of the two bounds

$$\mathcal{C}(\mathcal{A}, \alpha\mathcal{A}) \leq \frac{1}{2}\log\left(1 + \frac{\mathcal{A}^2}{4\sigma^2}\right) \tag{19}$$

$$\begin{aligned}\mathcal{C}(\mathcal{A}, \alpha\mathcal{A}) \leq{}& \left(1 - 2\mathcal{Q}\left(\frac{\delta + \frac{\mathcal{A}}{2}}{\sigma}\right)\right)\log\frac{\mathcal{A} + 2\delta}{\sigma\sqrt{2\pi}\left(1 - 2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)} \\ &- \frac{1}{2} + \mathcal{Q}\left(\frac{\delta}{\sigma}\right) + \frac{\delta}{\sqrt{2\pi}\sigma}e^{-\frac{\delta^2}{2\sigma^2}}\end{aligned} \tag{20}$$

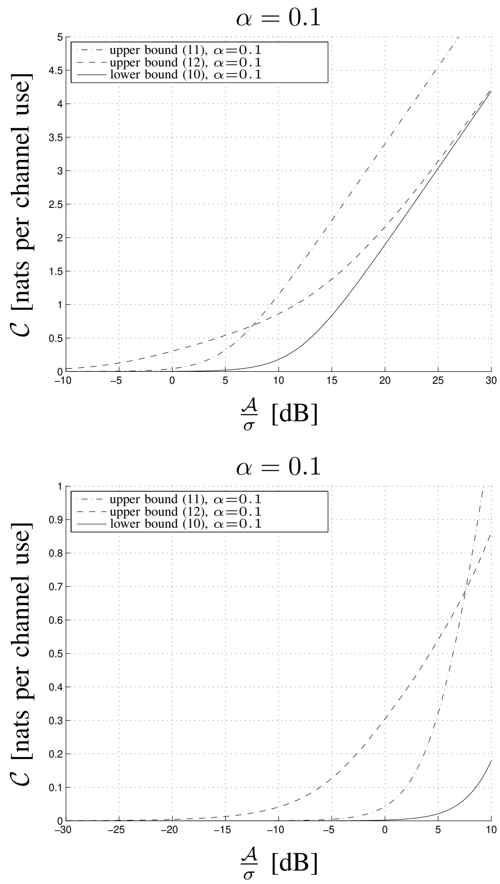

Fig. 1. Bounds of Theorem 3 for $\alpha = 0.1$ when upper bound (12) is numerically minimized over $\delta, \mu > 0$. The maximum gap between upper and lower bound is 0.68 nats (for $\frac{\mathcal{A}}{\sigma} \approx 10.5$ dB).

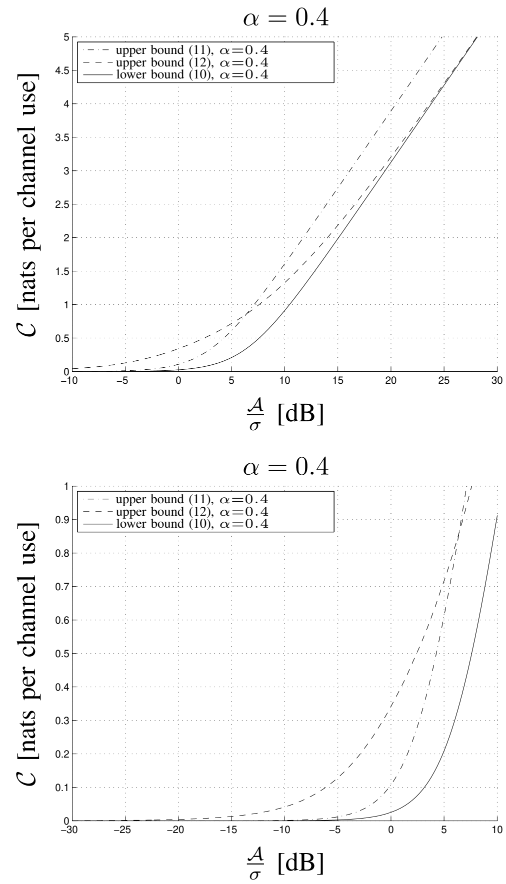

Fig. 2. Bounds of Theorem 3 for $\alpha = 0.4$ with numerically optimized upper bound (12). The maximum gap between upper and lower bound is 0.52 nats (for $\frac{\mathcal{A}}{\sigma} \approx 6.4$ dB).

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

A useful but suboptimal choice for δ is

$$\delta = \sigma \log\left(1 + \frac{\mathcal{A}}{\sigma}\right). \tag{21}$$

Figure 3 depicts the bounds of Theorem 5, where upper bound (20) is numerically minimized over $\delta > 0$.

Corollary 6 (Asymptotics): If $\frac{1}{2} \leq \alpha \leq 1$, then

$$\lim_{\mathcal{A}\uparrow\infty} \left\{ \mathcal{C}(\mathcal{A}, \alpha\mathcal{A}) - \log\frac{\mathcal{A}}{\sigma} \right\} = -\frac{1}{2}\log 2\pi e \tag{22}$$

and

$$\lim_{\mathcal{A}\downarrow 0} \frac{\mathcal{C}(\mathcal{A}, \alpha\mathcal{A})}{\mathcal{A}^2/\sigma^2} = \frac{1}{8}. \tag{23}$$

Note that (22) and (23) exhibit the well-known asymptotic behavior of the capacity of a Gaussian channel under a peak-power constraint only [27].

Based on the right-hand sides of (16) and (22) we define

$$\chi(\alpha) \triangleq \begin{cases} -\frac{1}{2}\log 2\pi e - (1-\alpha)\mu^* - \log(1 - \alpha\mu^*) & \\ & 0 < \alpha < \frac{1}{2}, \\ -\frac{1}{2}\log 2\pi e & \frac{1}{2} \leq \alpha \leq 1. \end{cases} \tag{24}$$

Thus for $\alpha \in (0,1)$, $\chi(\alpha)$ represents the second term in the high-SNR asymptotic expansion of the channel capacity $\mathcal{C}(\mathcal{A}, \alpha\mathcal{A})$. It is depicted in Figure 4. Note that when α tends to 0, then $\chi(\alpha)$ tends to $-\infty$. This can be seen by rewriting $\chi(\alpha)$ for $\alpha \in \left(0, \frac{1}{2}\right)$ using (13) as

$$\chi(\alpha) = -\frac{1}{2}\log 2\pi e - \alpha\mu^* - \log\frac{\mu^*}{1 - e^{-\mu^*}} \tag{25}$$

and then noting that $\mu^* \uparrow \infty$ and $\alpha\mu^* \uparrow 1$ when $\alpha \downarrow 0$.

C. Bounds on Channel Capacity for Case III

Theorem 7 (Bounds): In the absence of a peak-power constraint the channel capacity $\mathcal{C}(\mathcal{E})$ is lower-bounded by

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

and is upper-bounded by each of the bounds

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

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

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

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

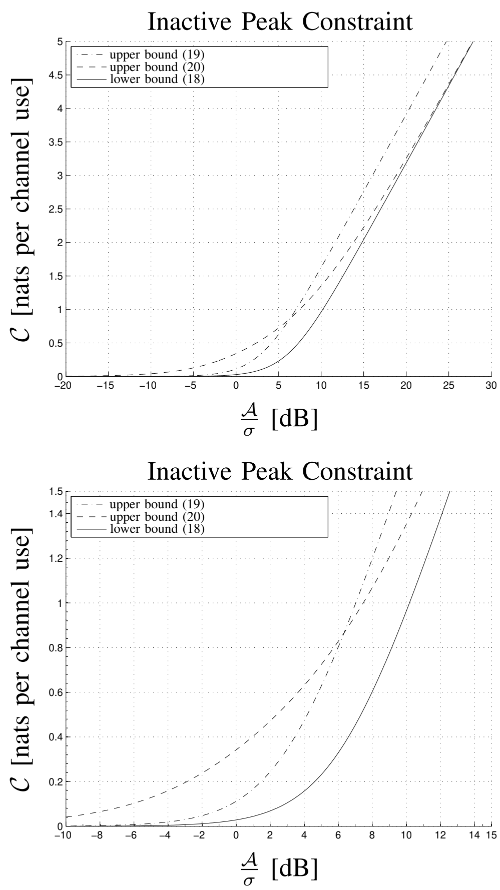

Fig. 3. Bounds on capacity for $\alpha \in \left[\frac{1}{2}, 1\right]$ according to Theorem 5, where upper bound (20) is numerically minimized over $\delta > 0$. The maximum gap between upper and lower bound is 0.50 nats (for $\frac{\mathcal{A}}{\sigma} \approx 6.4$ dB).

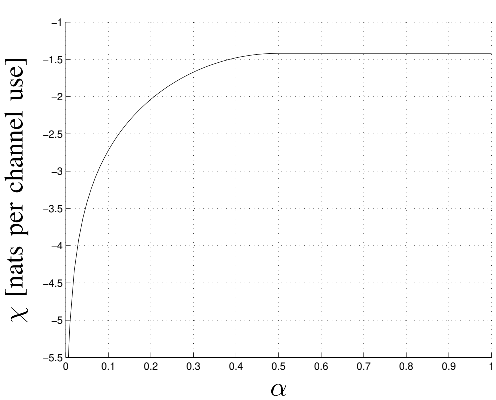

Fig. 4. The second term $\chi(\alpha)$ of the hign-SNR expansion of capacity for $\alpha \in (0, 1]$.

$$+ \frac{1}{\beta}\left(\delta + \mathcal{E} + \frac{\sigma}{\sqrt{2\pi}}e^{-\frac{\delta^2}{2\sigma^2}}\right) - \frac{1}{2}\log 2\pi e\sigma^2, \quad \delta \geq 0 \tag{28}$$

where $\beta > 0$ and δ are free parameters. 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

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

and β given in (30) on the top of this page; and for the free parameters in bound (28) is

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

and β given in (32) on the top of this page. Figure 5 depicts the bounds of Theorem 7 when the upper bounds (27) and (28) are numerically minimized over the allowed values of β and δ.

Proposition 8 (Asymptotics): In the absence of a peak-power constraint,

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

and

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

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

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.

D. Basic Ideas of the Derivations

One can always find a lower bound on capacity by dropping the maximization and choosing an arbitrary input distribution Q in (7). To get a tight bound, this choice of Q should yield a mutual information that is reasonably close to capacity. Such a choice is difficult to find and might make the evaluation of $I(Q, W)$ intractable. We circumvent these problems by using the EPI. For any probability distribution Q with a probability density function (PDF) we have

$$\mathcal{C} \geq I(Q, W) \tag{36}$$
$$= h(Y) - h(Y|X) \tag{37}$$
$$= h(X + Z) - h(Z) \tag{38}$$
$$\geq \frac{1}{2}\log\left(e^{2h(X)} + e^{2h(Z)}\right) - h(Z) \tag{39}$$

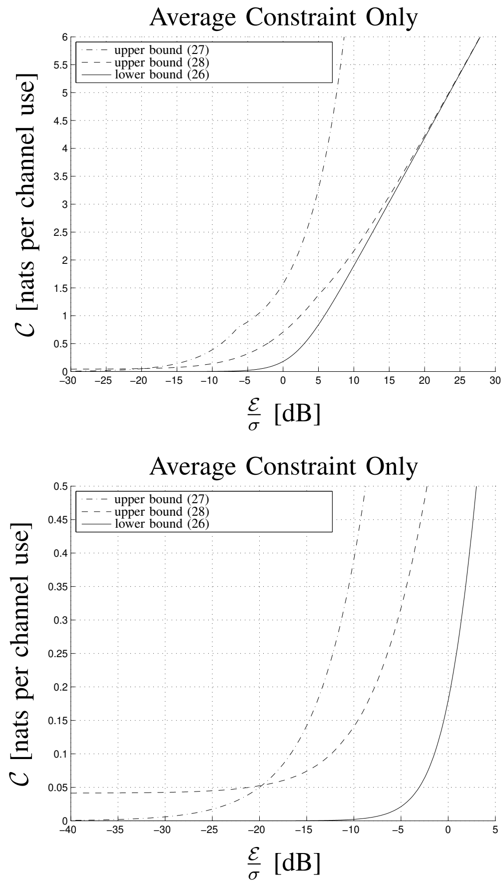

Fig. 5. Bounds on capacity according to Theorem 7 when upper bounds (27) and (28) are numerically minimized over allowed values of β, δ. The maximum gap between upper and lower bound is 0.57 nats (for $\frac{\mathcal{E}}{\sigma} \approx 2.8$ dB).

$$= \frac{1}{2}\log\left(1 + \frac{e^{2h(X)}}{2\pi e\sigma^2}\right) \tag{40}$$

where (39) follows from the EPI. To make this lower bound as tight as possible we will choose a distribution Q that maximizes differential entropy under the given constraints [19, Ch. 12]. The details can be found in Appendix A.[3]

The derivation of the upper bounds is based on Proposition 1. Choosing a PDF f on the output alphabet we get

$$\mathcal{C} \le \sup_Q \mathsf{E}_Q\left[-\int_{-\infty}^{\infty} \log f(y)\,\mathrm{d}W(y|X)\right] - \frac{1}{2}\log 2\pi e\sigma^2. \tag{41}$$

The details can be found in Appendix B.

The derivations of the asymptotic results are shown in Appendix C.

Appendix A
Derivation of the Firm Lower Bounds

The input distribution Q_1 that maximizes differential entropy under a nonnegativity constraint, a peak constraint, and an average constraint has PDF [19, Ch. 12]

$$\frac{1}{\mathcal{A}} \cdot \frac{\mu^*}{1 - e^{-\mu^*}} e^{-\frac{\mu^* x}{\mathcal{A}}}, \qquad 0 \le x \le \mathcal{A} \tag{42}$$

where μ^* has to be chosen such that the average-power constraint is satisfied, *i.e.*, μ^* is given as the solution to (13). The bound (10) now follows from (40) by computing $h(X)$ under the probability law Q_1.

Similarly, the lower bound (18) follows from (40) under the uniform distribution Q_2 over $[0, \mathcal{A}]$, and (26) follows from (40) under the exponential distribution Q_3. Note that Q_2 represents the pointwise limit of Q_1 when $\alpha \uparrow \frac{1}{2}$ and that Q_3 represents the pointwise limit of Q_1 when $\alpha \downarrow 0$.

Appendix B
Derivation of the Firm Upper Bounds

A. *Upper Bound* (11) *of Theorem 3*

To derive the first upper bound (11) we choose a PDF f_1 corresponding to a Gaussian random variable of mean $\mathcal{E}$ and of variance $(\sigma^2 + (\mathcal{A} - \mathcal{E})\mathcal{E})$. For arbitrary law Q satisfying $\mathsf{E}_Q[X] \le \mathcal{E}$ and $Q(X > A) = 0$ this yields

$$\begin{aligned} &\mathsf{E}_Q\left[-\int_{-\infty}^{\infty} \log f_1(y)\,\mathrm{d}W(y|X)\right] \\ &\quad= \log\sqrt{2\pi\big(\sigma^2 + \mathcal{E}(\mathcal{A} - \mathcal{E})\big)} \\ &\qquad + \mathsf{E}_Q\left[\frac{X^2 + \sigma^2 - 2\mathcal{E}X + \mathcal{E}^2}{2\sigma^2 + 2\mathcal{E}(\mathcal{A} - \mathcal{E})}\right] && (43) \\ &\quad\le \log\sqrt{2\pi\big(\sigma^2 + \mathcal{E}(\mathcal{A} - \mathcal{E})\big)} \\ &\qquad + \mathsf{E}_Q\left[\frac{(A - 2\mathcal{E})X + \sigma^2 + \mathcal{E}^2}{2\sigma^2 + 2\mathcal{E}(\mathcal{A} - \mathcal{E})}\right] && (44) \\ &\quad\le \log\sqrt{2\pi\big(\sigma^2 + \mathcal{E}(\mathcal{A} - \mathcal{E})\big)} + \frac{(A - 2\mathcal{E})\mathcal{E} + \sigma^2 + \mathcal{E}^2}{2\sigma^2 + 2\mathcal{E}(\mathcal{A} - \mathcal{E})} && (45) \\ &\quad= \frac{1}{2}\log 2\pi e\left(\sigma^2 + \mathcal{E}(\mathcal{A} - \mathcal{E})\right) && (46) \end{aligned}$$

where the first inequality follows from $X^2 \le \mathcal{A}X$ due to the peak-power constraint, and where the second inequality follows from the average-power constraint using that $\frac{\mathcal{E}}{\mathcal{A}} = \alpha \le \frac{1}{2}$, *i.e.*, $\mathcal{A} - 2\mathcal{E} \ge 0$.

B. *Upper Bound* (12) *of Theorem 3*

To derive (12) we choose the PDF

$$f_2(y) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}}, & y < -\delta \\ \frac{1}{\mathcal{A}} \cdot \frac{\mu\left(1 - 2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)}{e^{\frac{\mu\delta}{\mathcal{A}}} - e^{-\mu\left(1 + \frac{\delta}{\mathcal{A}}\right)}} e^{-\frac{\mu y}{\mathcal{A}}}, & -\delta \le y \le \mathcal{A} + \delta \\ \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y - \mathcal{A})^2}{2\sigma^2}}, & y > \mathcal{A} + \delta \end{cases} \tag{47}$$

where $\delta > 0$ and $\mu > 0$ are free parameters. This leads to the expression given in (48) on the top of the next page. We investigate each term individually. We start with c_1:

$$\begin{aligned} c_1 = \mathsf{E}_Q\Bigg[&\log\left(\sqrt{2\pi}\sigma\right) \cdot \mathcal{Q}\left(\frac{\delta + X}{\sigma}\right) \\ &+ \frac{1}{2\sigma^2}\int_{-\infty}^{-\delta} y^2 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y - X)^2}{2\sigma^2}}\,\mathrm{d}y\Bigg] \end{aligned} \tag{49}$$

[3] Of course, if under the law Q one can compute the differential entropy of $X + Z$ precisely, then one need not employ the EPI. But then the choice of Q might not be so clear.

$$\mathsf{E}_Q\left[-\int_{-\infty}^{\infty}\log f_2(y)\,\mathrm{d}W(y|X)\right] = \underbrace{\mathsf{E}_Q\left[\int_{-\infty}^{-\delta}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(y-X)^2}{2\sigma^2}}\log\left(\sqrt{2\pi}\sigma e^{\frac{y^2}{2\sigma^2}}\right)\mathrm{d}y\right]}_{c_1}$$
$$\underbrace{-\,\mathsf{E}_Q\left[\int_{-\delta}^{\mathcal{A}+\delta}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(y-X)^2}{2\sigma^2}}\log\left(\frac{e^{-\frac{\mu y}{\mathcal{A}}}}{\mathcal{A}}\frac{\mu\left(1-2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)}{e^{\frac{\mu\delta}{\mathcal{A}}}-e^{-\mu\left(1+\frac{\delta}{\mathcal{A}}\right)}}\right)\mathrm{d}y\right]}_{c_2}$$
$$\underbrace{+\,\mathsf{E}_Q\left[\int_{\mathcal{A}+\delta}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(y-X)^2}{2\sigma^2}}\log\left(\sqrt{2\pi}\sigma e^{\frac{(y-\mathcal{A})^2}{2\sigma^2}}\right)\mathrm{d}y\right]}_{c_3}. \tag{48}$$

$$\leq \mathsf{E}_Q\left[\log\left(\sqrt{2\pi}\sigma\right)\cdot\mathcal{Q}\left(\frac{\delta+X}{\sigma}\right)\right] + \frac{1}{2\sigma^2}\int_{-\infty}^{-\delta}y^2\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{y^2}{2\sigma^2}}\,\mathrm{d}y \tag{50}$$
$$= \mathsf{E}_Q\left[\log\left(\sqrt{2\pi}\sigma\right)\cdot\mathcal{Q}\left(\frac{\delta+X}{\sigma}\right)\right] + \frac{1}{2}\mathcal{Q}\left(\frac{\delta}{\sigma}\right) + \frac{\delta}{2\sqrt{2\pi}\sigma}e^{-\frac{\delta^2}{2\sigma^2}} \tag{51}$$

where the inequality follows from the assumption $\delta > 0$ that ensures that $(y-x)^2 \geq y^2$ for all $x \geq 0$ and $y \leq -\delta$. Similarly we get for c_3:

$$c_3 \leq \mathsf{E}_Q\left[\log\left(\sqrt{2\pi}\sigma\right)\cdot\mathcal{Q}\left(\frac{\delta+\mathcal{A}-X}{\sigma}\right)\right] + \frac{1}{2\sigma^2}\int_{\mathcal{A}+\delta}^{\infty}(y-\mathcal{A})^2\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(y-\mathcal{A})^2}{2\sigma^2}}\,\mathrm{d}y \tag{52}$$
$$= \mathsf{E}_Q\left[\log\left(\sqrt{2\pi}\sigma\right)\cdot\mathcal{Q}\left(\frac{\delta+\mathcal{A}-X}{\sigma}\right)\right] + \frac{1}{2}\mathcal{Q}\left(\frac{\delta}{\sigma}\right) + \frac{\delta}{2\sqrt{2\pi}\sigma}e^{-\frac{\delta^2}{2\sigma^2}}. \tag{53}$$

Here, the inequality follows because $(y-x)^2 \geq (y-\mathcal{A})^2$ for all $x \leq \mathcal{A}$ and $y \geq \mathcal{A}+\delta$. Finally, for c_2 we have

$$c_2 = \mathsf{E}_Q\left[\left(1-\mathcal{Q}\left(\frac{\delta+X}{\sigma}\right)-\mathcal{Q}\left(\frac{\delta+\mathcal{A}-X}{\sigma}\right)\right)\cdot\log\left(\frac{e^{\frac{\mu\delta}{\mathcal{A}}}-e^{-\mu\left(1+\frac{\delta}{\mathcal{A}}\right)}}{1-2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)}\frac{\mathcal{A}}{\mu}\right)\right]$$
$$+\,\mathsf{E}_Q\left[\left(1-\mathcal{Q}\left(\frac{\delta+X}{\sigma}\right)-\mathcal{Q}\left(\frac{\delta+\mathcal{A}-X}{\sigma}\right)\right)\frac{\mu}{\mathcal{A}}X\right] + \mathsf{E}_Q\left[\frac{\mu\sigma}{\mathcal{A}\sqrt{2\pi}}\left(e^{-\frac{(\delta+X)^2}{2\sigma^2}}-e^{-\frac{(\mathcal{A}+\delta-X)^2}{2\sigma^2}}\right)\right]. \tag{54}$$

Plugging c_1, c_2, and c_3 into (48) and combining this with (41) we get the following bound:

$$\mathcal{C} \leq \mathcal{Q}\left(\frac{\delta}{\sigma}\right) + \frac{\delta}{\sqrt{2\pi}\sigma}e^{-\frac{\delta^2}{2\sigma^2}} - \frac{1}{2} + \mathsf{E}_Q\left[\left(1-\mathcal{Q}\left(\frac{\delta+X}{\sigma}\right)-\mathcal{Q}\left(\frac{\delta+\mathcal{A}-X}{\sigma}\right)\right)\cdot\log\left(\frac{\mathcal{A}\left(e^{\frac{\mu\delta}{\mathcal{A}}}-e^{-\mu\left(1+\frac{\delta}{\mathcal{A}}\right)}\right)}{\sqrt{2\pi}\sigma\mu\left(1-2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)}\right)\right]$$
$$+\,\mathsf{E}_Q\left[\left(1-\mathcal{Q}\left(\frac{\delta+X}{\sigma}\right)-\mathcal{Q}\left(\frac{\delta+\mathcal{A}-X}{\sigma}\right)\right)\frac{\mu}{\mathcal{A}}X\right] + \mathsf{E}_Q\left[\frac{\mu\sigma}{\mathcal{A}\sqrt{2\pi}}\left(e^{-\frac{(\delta+X)^2}{2\sigma^2}}-e^{-\frac{(\mathcal{A}+\delta-X)^2}{2\sigma^2}}\right)\right]. \tag{55}$$

It can be shown that

$$\log\left(\frac{\mathcal{A}\left(e^{\frac{\mu\delta}{\mathcal{A}}}-e^{-\mu\left(1+\frac{\delta}{\mathcal{A}}\right)}\right)}{\sqrt{2\pi}\sigma\mu\left(1-2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)}\right) \geq 0 \tag{56}$$

for any values of $\mathcal{A}$, σ, δ, $\mu > 0$. Moreover, $\xi \mapsto 1 - \mathcal{Q}\left(\xi_0+\xi\right) - \mathcal{Q}\left(\xi_0+\gamma-\xi\right)$ for all $\xi \in [0,\gamma]$ is monotonically increasing and concave for all $\xi \in [0,\gamma]$. Therefore, using Jensen's inequality, we conclude that

$$\mathsf{E}_Q\left[\left(1-\mathcal{Q}\left(\frac{\delta+X}{\sigma}\right)-\mathcal{Q}\left(\frac{\delta+\mathcal{A}-X}{\sigma}\right)\right)\cdot\log\left(\frac{\mathcal{A}\left(e^{\frac{\mu\delta}{\mathcal{A}}}-e^{-\mu\left(1+\frac{\delta}{\mathcal{A}}\right)}\right)}{\sqrt{2\pi}\sigma\mu\left(1-2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)}\right)\right]$$
$$\leq \left(1-\mathcal{Q}\left(\frac{\delta+\alpha\mathcal{A}}{\sigma}\right)-\mathcal{Q}\left(\frac{\delta+(1-\alpha)\mathcal{A}}{\sigma}\right)\right)\cdot\log\left(\frac{\mathcal{A}\left(e^{\frac{\mu\delta}{\mathcal{A}}}-e^{-\mu\left(1+\frac{\delta}{\mathcal{A}}\right)}\right)}{\sqrt{2\pi}\sigma\mu\left(1-2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)}\right). \tag{57}$$

Next, we upper-bound $\left(1-\mathcal{Q}\left(\frac{\delta+x}{\sigma}\right)-\mathcal{Q}\left(\frac{\delta+\mathcal{A}-x}{\sigma}\right)\right)$ by its maximum value that is taken on for $x = \frac{\mathcal{A}}{2}$:

$$\mathsf{E}_Q\left[\left(1-\mathcal{Q}\left(\frac{\delta+X}{\sigma}\right)-\mathcal{Q}\left(\frac{\delta+\mathcal{A}-X}{\sigma}\right)\right)\frac{\mu}{\mathcal{A}}X\right]$$
$$\leq \mathsf{E}_Q\left[\left(1-\mathcal{Q}\left(\frac{\delta+\frac{\mathcal{A}}{2}}{\sigma}\right)-\mathcal{Q}\left(\frac{\delta+\mathcal{A}-\frac{\mathcal{A}}{2}}{\sigma}\right)\right)\frac{\mu}{\mathcal{A}}X\right] \tag{58}$$
$$\leq \mu\alpha\left(1-2\mathcal{Q}\left(\frac{\delta+\frac{\mathcal{A}}{2}}{\sigma}\right)\right). \tag{59}$$

And finally we use the monotonicity of the exponential function and the fact that $X \in [0,\mathcal{A}]$ to show the following:

$$\mathsf{E}_Q\left[\frac{\mu\sigma}{\mathcal{A}\sqrt{2\pi}}\left(e^{-\frac{(\delta+X)^2}{2\sigma^2}}-e^{-\frac{(\mathcal{A}+\delta-X)^2}{2\sigma^2}}\right)\right] \le \frac{\mu\sigma}{\mathcal{A}\sqrt{2\pi}}\left(e^{-\frac{\delta^2}{2\sigma^2}}-e^{-\frac{(\mathcal{A}+\delta)^2}{2\sigma^2}}\right). \quad (60)$$

Combining (55) with (57), (59), and (60) yields the bound on channel capacity given in (12).

C. Upper Bound (19) of Theorem 5

To derive bound (19) we choose a Gaussian PDF f_3 with mean $\frac{\mathcal{A}}{2}$ and variance $\sigma^2+\frac{\mathcal{A}^2}{4}$. This yields

$$\mathsf{E}_Q\left[-\int_{-\infty}^{\infty}\log f_3(y)\,\mathrm{d}W(y|X)\right] = \log\sqrt{2\pi\left(\sigma^2+\frac{\mathcal{A}^2}{4}\right)}+\mathsf{E}_Q\left[\frac{X^2+\sigma^2-\mathcal{A}X+\frac{\mathcal{A}^2}{4}}{2\sigma^2+\frac{\mathcal{A}^2}{2}}\right] \quad (61)$$

$$\le \log\sqrt{2\pi\left(\sigma^2+\frac{\mathcal{A}^2}{4}\right)}+\frac{\sigma^2+\frac{\mathcal{A}^2}{4}}{2\sigma^2+\frac{\mathcal{A}^2}{2}} \quad (62)$$

$$= \frac{1}{2}\log 2\pi e\sigma^2\left(1+\frac{\mathcal{A}^2}{4\sigma^2}\right) \quad (63)$$

where the inequality follows because $X^2 \le \mathcal{A}X$ due to the peak-power constraint. Combined with (41) this yields the claimed result. Note that the relation $\mathsf{E}_Q[X]\le\alpha\mathcal{A}$ has not been used. Therefore this bound is valid for all $\alpha\in[0,1]$ and especially for all $\alpha\in\left[\frac{1}{2},1\right]$.

D. Upper Bound (20) of Theorem 5

The derivation of this bound is similar to the derivation of (12). We choose

$$f_4(y)=\begin{cases}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{y^2}{2\sigma^2}}, & y<-\delta\\ \frac{1-2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)}{\mathcal{A}+2\delta}, & -\delta\le y\le\mathcal{A}+\delta\\ \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(y-\mathcal{A})^2}{2\sigma^2}}, & y>\mathcal{A}+\delta\end{cases} \quad (64)$$

where $\delta>0$ is a free parameter. This leads to the following expression:

$$\mathsf{E}_Q\left[-\int_{-\infty}^{\infty}\log f_4(y)\,\mathrm{d}W(y|X)\right]=c_1+c_3 + \underbrace{\mathsf{E}_Q\left[\int_{-\delta}^{\mathcal{A}+\delta}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(y-X)^2}{2\sigma^2}}\log\frac{\mathcal{A}+2\delta}{1-2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)}\,\mathrm{d}y\right]}_{\tilde{c}_2} \quad (65)$$

where c_1 and c_3 are defined in (48) and are upper-bounded in (51) and (53) (assuming that $\delta>0$), respectively. Similarly to c_2, we compute $\tilde{c}_2$ as follows:

$$\tilde{c}_2=\mathsf{E}_Q\left[\left(1-\mathcal{Q}\left(\frac{\delta+X}{\sigma}\right)-\mathcal{Q}\left(\frac{\delta+\mathcal{A}-X}{\sigma}\right)\right)\cdot\log\frac{\mathcal{A}+2\delta}{1-2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)}\right]. \quad (66)$$

Plugging c_1, $\tilde{c}_2$, and c_3 into (65) and combining this with (41) we get

$$\mathcal{C}\le\mathsf{E}_Q\left[\left(1-\mathcal{Q}\left(\frac{\delta+X}{\sigma}\right)-\mathcal{Q}\left(\frac{\delta+\mathcal{A}-X}{\sigma}\right)\right)\cdot\log\frac{\mathcal{A}+2\delta}{\sqrt{2\pi}\sigma\left(1-2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)}\right]-\frac{1}{2}+\mathcal{Q}\left(\frac{\delta}{\sigma}\right) + \frac{\delta}{\sqrt{2\pi}\sigma}e^{-\frac{\delta^2}{2\sigma^2}}. \quad (67)$$

It can be shown that for any $\mathcal{A}>0$ and $\delta>0$

$$\log\frac{\mathcal{A}+2\delta}{\sqrt{2\pi}\sigma\left(1-2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)}\ge 0 \quad (68)$$

and hence analogously to (58)–(59) we get

$$\mathsf{E}_Q\left[\left(1-\mathcal{Q}\left(\frac{\delta+X}{\sigma}\right)-\mathcal{Q}\left(\frac{\delta+\mathcal{A}-X}{\sigma}\right)\right)\cdot\log\frac{\mathcal{A}+2\delta}{\sqrt{2\pi}\sigma\left(1-2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)}\right] \le\left(1-2\mathcal{Q}\left(\frac{\delta+\frac{\mathcal{A}}{2}}{\sigma}\right)\right)\log\frac{\mathcal{A}+2\delta}{\sqrt{2\pi}\sigma\left(1-2\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)}. \quad (69)$$

Again we have not used the relation $\mathsf{E}_Q[X]\le\alpha\mathcal{A}$, and hence the bound is valid for arbitrary $\alpha\in[0,1]$.

E. Upper Bound (27) of Theorem 7

One of the main challenges of deriving the upper bounds of Theorem 7 using duality is that without a peak-power constraint the input can be arbitrarily large (albeit with small probability). This makes it much harder to find bounds on expressions like $\mathsf{E}_Q[X^2]$. We choose

$$f_5(y)=\begin{cases}\frac{1}{\beta e^{-\frac{\delta^2}{2\sigma^2}}+\sqrt{2\pi}\sigma\mathcal{Q}\left(\frac{\delta}{\sigma}\right)}e^{-\frac{y^2}{2\sigma^2}}, & y<-\delta\\ \frac{1}{\beta e^{-\frac{\delta^2}{2\sigma^2}}+\sqrt{2\pi}\sigma\mathcal{Q}\left(\frac{\delta}{\sigma}\right)}e^{-\frac{\delta^2}{2\sigma^2}}e^{-\frac{y+\delta}{\beta}}, & y\ge-\delta\end{cases} \quad (70)$$

where $\delta\in\mathbb{R}$ and $\beta>0$ are free parameters. This leads to the following expression:

$$\mathsf{E}_Q\left[-\int_{-\infty}^{\infty}\log f_5(y)\,\mathrm{d}W(y|X)\right] =\log\left(\beta e^{-\frac{\delta^2}{2\sigma^2}}+\sqrt{2\pi}\sigma\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right) +\underbrace{\mathsf{E}_Q\left[\left(\frac{1}{2}+\frac{X^2}{2\sigma^2}\right)\mathcal{Q}\left(\frac{\delta+X}{\sigma}\right)+\frac{\delta-X}{2\sqrt{2\pi}\sigma}e^{-\frac{(\delta+X)^2}{2\sigma^2}}\right]}_{c_4} +\underbrace{\mathsf{E}_Q\left[\frac{\delta^2}{2\sigma^2}\left(1-\mathcal{Q}\left(\frac{\delta+X}{\sigma}\right)\right)\right]}_{c_5} +\underbrace{\mathsf{E}_Q\left[\frac{\delta+X}{\beta}\left(1-\mathcal{Q}\left(\frac{\delta+X}{\sigma}\right)\right)+\frac{\sigma}{\sqrt{2\pi}\beta}e^{-\frac{(\delta+X)^2}{2\sigma^2}}\right]}_{c_6}. \quad (71)$$

We now restrict the free parameter δ to satisfy

$$\delta \leq -\frac{\sigma}{\sqrt{e}} \tag{72}$$

and note that for arbitrary input law Q such that $\mathsf{E}_Q[X] \leq \mathcal{E}$:

$$\mathsf{E}_Q\left[\frac{1}{2}\mathcal{Q}\left(\frac{\delta+X}{\sigma}\right)\right] \leq \frac{1}{2}; \tag{73}$$

$$\begin{aligned}&\mathsf{E}_Q\left[\frac{X^2}{2\sigma^2}\mathcal{Q}\left(\frac{\delta+X}{\sigma}\right)\right]\\ &\quad= \mathsf{E}_Q\left[\frac{X}{2\sigma}\cdot\frac{X}{\sigma}\mathcal{Q}\left(\frac{X}{\sigma}-\frac{-\delta}{\sigma}\right)\right] &(74)\\ &\quad\leq \mathsf{E}_Q\left[\frac{X}{2\sigma}\cdot\frac{-\delta}{\sigma}\right] \leq -\frac{\delta\mathcal{E}}{2\sigma^2}; &(75)\end{aligned}$$

$$\mathsf{E}_Q\left[\frac{\delta-X}{2\sqrt{2\pi}\sigma}e^{-\frac{(\delta+X)^2}{2\sigma^2}}\right] \leq 0; \tag{76}$$

$$\begin{aligned}&\mathsf{E}_Q\left[\frac{\delta^2}{2\sigma^2}\left(1-\mathcal{Q}\left(\frac{\delta+X}{\sigma}\right)\right)\right]\\ &\quad\leq \mathsf{E}_Q\left[\frac{\delta^2}{2\sigma^2}\left(1-\mathcal{Q}\left(\frac{\delta}{\sigma}\right)+\frac{X}{-\delta}\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)\right] &(77)\\ &\quad\leq \frac{\delta^2}{2\sigma^2}\left(1-\mathcal{Q}\left(\frac{\delta}{\sigma}\right)-\frac{\mathcal{E}}{\delta}\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right); &(78)\end{aligned}$$

$$\mathsf{E}_Q\left[\frac{\delta}{\beta}\left(1-\mathcal{Q}\left(\frac{\delta+X}{\sigma}\right)\right)\right] \leq 0; \tag{79}$$

$$\mathsf{E}_Q\left[\frac{X}{\beta}\left(1-\mathcal{Q}\left(\frac{\delta+X}{\sigma}\right)\right)\right] \leq \mathsf{E}_Q\left[\frac{X}{\beta}\right] \leq \frac{\mathcal{E}}{\beta}; \tag{80}$$

$$\mathsf{E}_Q\left[\frac{\sigma}{\sqrt{2\pi}\beta}e^{-\frac{(\delta+X)^2}{2\sigma^2}}\right] \leq \frac{\sigma}{\sqrt{2\pi}\beta}. \tag{81}$$

Here (75) follows because $\xi\mathcal{Q}(\xi-\mu) \leq \mu$ for $\mu \geq \frac{1}{\sqrt{e}}$ and $\xi \geq 0$ and because of the assumption (72); and in (77) we use $1-\mathcal{Q}(\xi-\mu) \leq 1-\mathcal{Q}(-\mu)+\frac{\xi}{\mu}\mathcal{Q}(-\mu)$ for $\mu,\xi \geq 0$. Combining (73)–(81) with (71) and (41) yields the claimed result.

F. Upper Bound (28) *of Theorem 7*

The bound (28) follows from the same choice (70) as we have used for the bound (27). However, here we will restrict the free parameter δ to be nonnegative:

$$\delta \geq 0. \tag{82}$$

We can then bound c_4 as

$$\begin{aligned}c_4 &\leq \mathsf{E}_Q\left[\frac{1}{2\sigma^2}\int_{-\infty}^{-\delta} y^2\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{y^2}{2\sigma^2}}\,\mathrm{d}y\right] &(83)\\ &= \frac{1}{2}\mathcal{Q}\left(\frac{\delta}{\sigma}\right)+\frac{\delta}{2\sqrt{2\pi}\sigma}e^{-\frac{\delta^2}{2\sigma^2}} &(84)\end{aligned}$$

where the inequality follows from the assumption $\delta \geq 0$ and the nonnegativity of X that ensure that $(\delta+X)^2 \geq \delta^2$.

Moreover, using the concavity and monotonicity of $\xi \mapsto (1-\mathcal{Q}(\xi))$ for $\xi \geq 0$ we bound

$$c_5 \leq \frac{\delta^2}{2\sigma^2}\left(1-\mathcal{Q}\left(\frac{\delta+\mathcal{E}}{\sigma}\right)\right) \tag{85}$$

and, using the nonnegativity of $\mathcal{Q}(\cdot)$ and of X,

$$\begin{aligned}c_6 &\leq \mathsf{E}_Q\left[\frac{\delta+X}{\beta}+\frac{\sigma}{\sqrt{2\pi}\beta}e^{-\frac{\delta^2}{2\sigma^2}}\right] &(86)\\ &\leq \frac{\delta+\mathcal{E}}{\beta}+\frac{\sigma}{\sqrt{2\pi}\beta}e^{-\frac{\delta^2}{2\sigma^2}}. &(87)\end{aligned}$$

Combining (71), (84), (85), and (87) with (41) yields the claimed result.

APPENDIX C
DERIVATION OF ASYMPTOTIC RESULTS

A. High-SNR Asymptotic Expressions

The analysis of the asymptotic behavior (16) is based on the lower bound (10) and on the upper bound (12) with δ as in (14) and μ equal to μ^*, the solution to (13).

To derive the asymptotic behavior (22) we compare the lower bound (18) with the upper bound (20) using δ as in (21).

To derive (33) we compare the lower bound (26) with the upper bound (28) using the following choice of the free parameters β and δ:

$$\beta \triangleq \mathcal{E} \tag{88}$$

$$\delta \triangleq \sigma\sqrt{\log\frac{\mathcal{E}}{\sigma}} \tag{89}$$

for $\mathcal{E} \geq \sigma$.

B. Low-SNR Asymptotic Expressions

In order to prove the low-SNR asymptotic expression (17) in Corollary 4, we derive an asymptotic lower bound that combined with upper bound (11) yields the desired result. The lower bound we propose is based on Theorem 2 in [21]. Note that for the channel (2) under consideration, the technical conditions A–F in [21] are satisfied. Theorem 2 in [21] states that for peak-constrained inputs $|X| < \mathcal{A}$ the mutual information satisfies

$$\mathcal{C}(\mathcal{A},\mathcal{E}) \geq I(X;Y) = \frac{\mathrm{Var}(X)}{2\sigma^2}+o(\mathcal{A}^2) \tag{90}$$

where $o(\mathcal{A}^2)$ decreases faster to 0 than $\mathcal{A}^2$, *i.e.*,

$$\lim_{\mathcal{A}\downarrow 0}\frac{o(\mathcal{A}^2)}{\mathcal{A}^2} = 0. \tag{91}$$

We restrict attention to settings where $0 < \mathcal{A} < 1$ and choose a binary input

$$X = \begin{cases}0, & \text{with probability } 1-\alpha\\ \mathcal{A}(1-\mathcal{A}), & \text{with probability } \alpha.\end{cases} \tag{92}$$

This yields the correct asymptotics.

The low-SNR asymptotic expression (23) is derived analogously by choosing X to equiprobably take on the values 0 and $\mathcal{A}(1-\mathcal{A})$, for $0 < \mathcal{A} < 1$ and comparing the corresponding lower bound (90) with the upper bound (19).

The asymptotic upper bound (34) follows from the upper bound (27) with a choice of δ as in (29) and with $\beta \triangleq \frac{1}{\mathcal{E}}$:

$$\frac{\mathcal{C}(\mathcal{E})}{\frac{\mathcal{E}}{\sigma}\sqrt{\log\frac{\sigma}{\mathcal{E}}}} \leq \frac{\log\left(\frac{\mathcal{E}}{\sqrt{2\pi}\sigma^3}+\mathcal{Q}\left(\frac{\delta}{\sigma}\right)\right)}{\frac{\mathcal{E}}{\sigma}\sqrt{\log\frac{\sigma}{\mathcal{E}}}}+\frac{2\sqrt{\log\frac{\sigma}{\mathcal{E}}}}{\frac{\mathcal{E}}{\sigma}}\mathcal{Q}\left(-\frac{\delta}{\sigma}\right)$$

$$+ 1 + \mathcal{Q}\left(\frac{\delta}{\sigma}\right) + \frac{\sigma\left(\mathcal{E} + \frac{\sigma}{\sqrt{2\pi}}\right)}{\sqrt{\log\frac{\sigma}{\mathcal{E}}}} \tag{93}$$

which leads to the claimed asymptotic behavior.

The derivation of (35) is more involved and requires a new lower bound on capacity. The lower bound is obtained by lower-bounding the mutual information $I(Q,W)$ for a binary input

$$Q(x) = \begin{cases} 1 - \frac{\mathcal{E}}{x_1}, & \text{if } x = 0 \\ \frac{\mathcal{E}}{x_1}, & \text{if } x = x_1 \end{cases} \tag{94}$$

where for sufficiently small $\mathcal{E}$ we choose

$$x_1 \triangleq \sigma\sqrt{c \log\frac{\sigma}{\mathcal{E}}} \tag{95}$$

for some constant $c > 2$. Note that $x_1 \uparrow \infty$ as $\mathcal{E} \downarrow 0$. In the remainder we assume $\frac{\mathcal{E}}{\sigma} \leq \frac{1}{2}$ so that (94) is well-defined. The PDF of the channel output Y corresponding to the input (94) is given by

$$f_Y(y) = \left(1 - \frac{\mathcal{E}}{x_1}\right)\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{y^2}{2\sigma^2}} + \frac{\mathcal{E}}{x_1}\cdot\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(y-x_1)^2}{2\sigma^2}}. \tag{96}$$

In order to evaluate the mutual information $I(Q,W)$ for the chosen binary input distribution we write it as (see [28])

$$I(Q,W) = \int D\big(W(\cdot|x)\big\|W(\cdot|0)\big)\,\mathrm{d}Q(x) - D\big(R(\cdot)\big\|W(\cdot|0)\big) \tag{97}$$

and evaluate the first term on the right-hand side as

$$\int D\big(W(\cdot|x)\big\|W(\cdot|0)\big)\,\mathrm{d}Q(x) = \frac{\sqrt{c}}{2}\cdot\frac{\mathcal{E}}{\sigma}\sqrt{\log\frac{\sigma}{\mathcal{E}}}. \tag{98}$$

Evaluating the second term is more difficult, and in fact we only derive an upper bound on it which exhibits the desired asymptotic behavior at low SNR. We shall show that

$$\varlimsup_{\mathcal{E}\downarrow 0}\frac{D\big(R(\cdot)\big\|W(\cdot|0)\big)}{\frac{\mathcal{E}}{\sigma}\sqrt{\log\frac{\sigma}{\mathcal{E}}}} \leq \frac{\sqrt{c}}{2} - \frac{1}{\sqrt{c}} \tag{99}$$

from which follows by (97) and (98)

$$\varliminf_{\mathcal{E}\downarrow 0}\frac{I(Q,W)}{\frac{\mathcal{E}}{\sigma}\sqrt{\log\frac{\sigma}{\mathcal{E}}}} \geq \frac{1}{\sqrt{c}}. \tag{100}$$

The desired asymptotic lower bound in (35) then follows because (100) holds for any $c > 2$.

Thus, in order to prove (99), we write

$$\begin{aligned} D\big(R(\cdot)\big\|W(\cdot|0)\big) &= \underbrace{\int_{-\infty}^{\frac{x_1}{2}} f_Y(y)\log\left(1 - \frac{\mathcal{E}}{x_1} + \frac{\mathcal{E}}{x_1}e^{\frac{yx_1}{\sigma^2} - \frac{x_1^2}{2\sigma^2}}\right)\mathrm{d}y}_{c_7} \\ &\quad + \underbrace{\int_{\frac{x_1}{2}}^{\frac{x_1}{2}+\frac{x_1}{c}} f_Y(y)\log\left(1 - \frac{\mathcal{E}}{x_1} + \frac{\mathcal{E}}{x_1}e^{\frac{yx_1}{\sigma^2} - \frac{x_1^2}{2\sigma^2}}\right)\mathrm{d}y}_{c_8} \\ &\quad + \underbrace{\int_{\frac{x_1}{2}+\frac{x_1}{c}}^{\infty} f_Y(y)\log\left(1 - \frac{\mathcal{E}}{x_1} + \frac{\mathcal{E}}{x_1}e^{\frac{yx_1}{\sigma^2} - \frac{x_1^2}{2\sigma^2}}\right)\mathrm{d}y}_{c_9} \end{aligned} \tag{101}$$

and upper-bound c_7, c_8, and c_9. For $y \geq \frac{x_1}{2}$:

$$c_7 \leq \int_{-\infty}^{\frac{x_1}{2}} f_Y(y)\log\left(1 - \frac{\mathcal{E}}{x_1} + \frac{\mathcal{E}}{x_1}e^{\frac{x_1}{\sigma^2}\frac{x_1}{2} - \frac{x_1^2}{2\sigma^2}}\right)\mathrm{d}y \tag{102}$$

$$= \int_{-\infty}^{\xi_1} f_Y(y)\log 1\,\mathrm{d}y = 0. \tag{103}$$

Next we examine c_8. Using $\frac{\mathcal{E}}{x_1} \geq 0$ and $\log(1+\xi) \leq \xi$ for all $\xi \geq 0$ we get

$$\begin{aligned} \log\left(1 - \frac{\mathcal{E}}{x_1} + \frac{\mathcal{E}}{x_1}e^{\frac{yx_1}{\sigma^2} - \frac{x_1^2}{2\sigma^2}}\right) &\leq \log\left(1 + \frac{\mathcal{E}}{x_1}e^{\frac{yx_1}{\sigma^2} - \frac{x_1^2}{2\sigma^2}}\right) \\ &\leq \frac{\mathcal{E}}{x_1}e^{\frac{yx_1}{\sigma^2} - \frac{x_1^2}{2\sigma^2}} \end{aligned} \tag{104}$$

and hence

$$c_8 \leq \int_{\frac{x_1}{2}}^{\frac{x_1}{2}+\frac{x_1}{c}} f_Y(y)\frac{\mathcal{E}}{x_1}e^{\frac{yx_1}{\sigma^2} - \frac{x_1^2}{2\sigma^2}}\,\mathrm{d}y \tag{105}$$

$$\begin{aligned} &= \left(1 - \frac{\frac{\mathcal{E}}{\sigma}}{\sqrt{c}\sqrt{\log\frac{\sigma}{\mathcal{E}}}}\right)\frac{\frac{\mathcal{E}}{\sigma}}{\sqrt{c}\sqrt{\log\frac{\sigma}{\mathcal{E}}}} \\ &\quad\cdot\left(\mathcal{Q}\left(\left(\frac{\sqrt{c}}{2} - \frac{1}{\sqrt{c}}\right)\sqrt{\log\frac{\sigma}{\mathcal{E}}}\right) - \mathcal{Q}\left(\frac{\sqrt{c}}{2}\sqrt{\log\frac{\sigma}{\mathcal{E}}}\right)\right) \\ &\quad + \frac{\left(\frac{\mathcal{E}}{\sigma}\right)^{2-c}}{c\log\frac{\sigma}{\mathcal{E}}}\left(\mathcal{Q}\left(\left(\frac{3\sqrt{c}}{2} - \frac{1}{\sqrt{c}}\right)\sqrt{\log\frac{\sigma}{\mathcal{E}}}\right)\right. \\ &\quad\left. - \mathcal{Q}\left(\frac{3\sqrt{c}}{2}\sqrt{\log\frac{\sigma}{\mathcal{E}}}\right)\right). \end{aligned} \tag{106}$$

Note that

$$\begin{aligned} \lim_{\mathcal{E}\downarrow 0}&\frac{\left(1 - \frac{\frac{\mathcal{E}}{\sigma}}{\sqrt{c}\sqrt{\log\frac{\sigma}{\mathcal{E}}}}\right)\frac{\frac{\mathcal{E}}{\sigma}}{\sqrt{c}\sqrt{\log\frac{\sigma}{\mathcal{E}}}}}{\frac{\mathcal{E}}{\sigma}\sqrt{\log\frac{\sigma}{\mathcal{E}}}} \\ &\cdot\left(\mathcal{Q}\left(\left(\frac{\sqrt{c}}{2} - \frac{1}{\sqrt{c}}\right)\sqrt{\log\frac{\sigma}{\mathcal{E}}}\right) - \mathcal{Q}\left(\frac{\sqrt{c}}{2}\sqrt{\log\frac{\sigma}{\mathcal{E}}}\right)\right) \\ &= 0. \end{aligned} \tag{107}$$

To deal with the second summand in (106) we note that

$$\mathcal{Q}(\xi) < \frac{1}{\sqrt{2\pi}\xi}\,e^{-\frac{\xi^2}{2}}, \quad \xi > 0 \tag{108}$$

and get

$$\begin{aligned} &\frac{\left(\frac{\mathcal{E}}{\sigma}\right)^{2-c}}{c\log\frac{\sigma}{\mathcal{E}}}\left(\mathcal{Q}\left(\left(\frac{3\sqrt{c}}{2} - \frac{1}{\sqrt{c}}\right)\sqrt{\log\frac{\sigma}{\mathcal{E}}}\right)\right. \\ &\quad\left.\underbrace{- \mathcal{Q}\left(\frac{3\sqrt{c}}{2}\sqrt{\log\frac{\sigma}{\mathcal{E}}}\right)}_{\geq 0}\right) \\ &< \frac{\left(\frac{\mathcal{E}}{\sigma}\right)^{2-c}}{c\log\frac{\sigma}{\mathcal{E}}}\cdot\frac{1}{\sqrt{2\pi}\sqrt{\log\frac{\sigma}{\mathcal{E}}}\left(\frac{3\sqrt{c}}{2} - \frac{1}{\sqrt{c}}\right)}\left(\frac{\mathcal{E}}{\sigma}\right)^{\frac{c}{2}\left(\frac{3}{2} - \frac{1}{c}\right)^2} \end{aligned} \tag{109}$$

$$= \frac{\left(\frac{\mathcal{E}}{\sigma}\right)^{\frac{1}{2} + \frac{1}{2}\left(\frac{c}{4} + \frac{1}{c}\right)}}{\left(\log\frac{\sigma}{\mathcal{E}}\right)^{3/2}}\cdot\frac{1}{\sqrt{2\pi}\left(\frac{3c^{3/2}}{2} - \sqrt{c}\right)}. \tag{110}$$

Note that whenever $c > 2$ then

$$\left(\frac{c}{4}+\frac{1}{c}\right) > 1 \tag{111}$$

and

$$\left(\frac{3c^{3/2}}{2}-\sqrt{c}\right) \neq 0 \tag{112}$$

and therefore

$$\overline{\lim_{\mathcal{E}\downarrow 0}} \frac{c_8}{\frac{\mathcal{E}}{\sigma}\sqrt{\log\frac{\sigma}{\mathcal{E}}}} \leq 0. \tag{113}$$

Finally, we examine the limiting behavior of c_9. To this end we rewrite c_9 as

$$\begin{aligned} c_9 =& \underbrace{\int_{\frac{x_1}{2}+\frac{x_1}{c}}^{\infty} f_Y(y)\log\left(e^{\frac{yx_1}{\sigma^2}-\frac{x_1^2}{2\sigma^2}}\right)\mathrm{d}y}_{c_{9,1}} \\ &+ \underbrace{\int_{\frac{x_1}{2}+\frac{x_1}{c}}^{\infty} f_Y(y)\log\left(\frac{\mathcal{E}}{x_1}\right)\mathrm{d}y}_{c_{9,2}} \\ &+ \underbrace{\int_{\frac{x_1}{2}+\frac{x_1}{c}}^{\infty} f_Y(y)\log\left(1+\frac{\left(1-\frac{\mathcal{E}}{x_1}\right)}{\frac{\mathcal{E}}{x_1}}e^{-\frac{yx_1}{\sigma^2}+\frac{x_1^2}{2\sigma^2}}\right)\mathrm{d}y}_{c_{9,3}} \end{aligned} \tag{114}$$

and note that

$$\begin{aligned} c_{9,1} =& \underbrace{\left(1-\frac{\mathcal{E}}{x_1}\right)}_{\leq 1}\frac{x_1}{\sigma^2}\cdot\frac{\sigma}{\sqrt{2\pi}}e^{-\frac{x_1^2\left(\frac{1}{2}+\frac{1}{c}\right)^2}{2\sigma^2}} \\ &- \underbrace{\left(1-\frac{\mathcal{E}}{x_1}\right)}_{\geq 0 \text{ if } \frac{\mathcal{E}}{\sigma}\leq\frac{1}{2}}\underbrace{\frac{x_1^2}{2\sigma^2}\mathcal{Q}\left(\frac{x_1}{2\sigma}+\frac{x_1}{c\sigma}\right)}_{\geq 0} \\ &+ \frac{\mathcal{E}}{x_1}\frac{x_1}{\sigma^2}\cdot\frac{\sigma}{\sqrt{2\pi}}e^{-\frac{x_1^2}{2\sigma^2}\left(\frac{1}{c}-\frac{1}{2}\right)^2} + \frac{\mathcal{E}x_1}{2\sigma^2}\underbrace{\mathcal{Q}\left(\frac{x_1}{c\sigma}-\frac{x_1}{2\sigma}\right)}_{\leq 1} \end{aligned} \tag{115}$$

$$\begin{aligned} \leq& \frac{\sqrt{c}\sqrt{\log\frac{\sigma}{\mathcal{E}}}}{\sqrt{2\pi}}\left(\frac{\mathcal{E}}{\sigma}\right)^{\frac{c}{2}\left(\frac{1}{2}+\frac{1}{c}\right)^2} + \frac{1}{\sqrt{2\pi}}\left(\frac{\mathcal{E}}{\sigma}\right)^{\frac{c}{2}\left(\frac{1}{c}-\frac{1}{2}\right)^2+1} \\ &+ \frac{\mathcal{E}}{\sigma}\cdot\frac{\sqrt{c}}{2}\sqrt{\log\frac{\sigma}{\mathcal{E}}}. \end{aligned} \tag{116}$$

Again, since $c > 2$ we have $\frac{c}{4}+\frac{1}{c} > 1$, and therefore

$$\lim_{\mathcal{E}\downarrow 0}\frac{\frac{\sqrt{c}\sqrt{\log\frac{\sigma}{\mathcal{E}}}}{\sqrt{2\pi}}\left(\frac{\mathcal{E}}{\sigma}\right)^{\frac{c}{2}\left(\frac{1}{2}+\frac{1}{c}\right)^2}}{\frac{\mathcal{E}}{\sigma}\sqrt{\log\frac{\sigma}{\mathcal{E}}}} = 0. \tag{117}$$

Moreover,

$$\lim_{\mathcal{E}\downarrow 0}\frac{\frac{1}{\sqrt{2\pi}}\left(\frac{\mathcal{E}}{\sigma}\right)^{\frac{c}{2}\left(\frac{1}{c}-\frac{1}{2}\right)^2+1}}{\frac{\mathcal{E}}{\sigma}\sqrt{\log\frac{\sigma}{\mathcal{E}}}} = 0 \tag{118}$$

and hence

$$\overline{\lim_{\mathcal{E}\downarrow 0}}\frac{c_{9,1}}{\frac{\mathcal{E}}{\sigma}\sqrt{\log\frac{\sigma}{\mathcal{E}}}} \leq \frac{\sqrt{c}}{2}. \tag{119}$$

Next we analyze $c_{9,2}$. Note that $\frac{\mathcal{E}}{x_1} \leq 1$, for $\frac{\mathcal{E}}{\sigma} \leq \frac{1}{2}$, and hence $\log\frac{\mathcal{E}}{x_1} \leq 0$. Therefore,

$$\begin{aligned} c_{9,2} =& \left(\underbrace{\left(1-\frac{\mathcal{E}}{x_1}\right)}_{\geq 0 \text{ if } \frac{\mathcal{E}}{\sigma}\leq\frac{1}{2}}\underbrace{\mathcal{Q}\left(\frac{x_1}{2\sigma}+\frac{x_1}{c\sigma}\right)}_{\geq 0}+\frac{\mathcal{E}}{x_1}\mathcal{Q}\left(\frac{x_1}{c\sigma}-\frac{x_1}{2\sigma}\right)\right) \\ &\cdot\log\left(\frac{\mathcal{E}}{x_1}\right) \end{aligned} \tag{120}$$

$$\leq \frac{\mathcal{E}}{x_1}\mathcal{Q}\left(\frac{x_1}{c\sigma}-\frac{x_1}{2\sigma}\right)\log\left(\frac{\mathcal{E}}{x_1}\right) \tag{121}$$

$$\begin{aligned} =& \frac{\frac{\mathcal{E}}{\sigma}}{\sqrt{c}\sqrt{\log\frac{\sigma}{\mathcal{E}}}}\mathcal{Q}\left(\left(\frac{1}{\sqrt{c}}-\frac{\sqrt{c}}{2}\right)\sqrt{\log\frac{\sigma}{\mathcal{E}}}\right) \\ &\cdot\log\left(\frac{\frac{\mathcal{E}}{\sigma}}{\sqrt{c}\sqrt{\log\frac{\sigma}{\mathcal{E}}}}\right). \end{aligned} \tag{122}$$

Since $c > 2$ the term $\mathcal{Q}\left(\left(\frac{1}{\sqrt{c}}-\frac{\sqrt{c}}{2}\right)\sqrt{\log\frac{\sigma}{\mathcal{E}}}\right)$ tends to 1 when $\mathcal{E}$ tends to 0, and therefore

$$\overline{\lim_{\mathcal{E}\downarrow 0}}\frac{c_{9,2}}{\frac{\mathcal{E}}{\sigma}\sqrt{\log\frac{\sigma}{\mathcal{E}}}} = -\frac{1}{\sqrt{c}}. \tag{123}$$

Finally, we analyze $c_{9,3}$. Using that $\frac{x_1}{\mathcal{E}}-1 \geq 0$ if $\frac{\mathcal{E}}{\sigma} \leq \frac{1}{2}$ we lower-bound $y \geq \frac{x_1}{2}+\frac{x_1}{c}$ to get

$$c_{9,3} \leq \int_{\frac{x_1}{2}+\frac{x_1}{c}}^{\infty} f_Y(y)\log\left(1+\left(\frac{x_1}{\mathcal{E}}-1\right)e^{-\frac{x_1^2}{c\sigma^2}}\right)\mathrm{d}y \tag{124}$$

$$\begin{aligned} =& \left(\underbrace{\left(1-\frac{\mathcal{E}}{x_1}\right)}_{\leq 1}\mathcal{Q}\left(\frac{x_1}{2\sigma}+\frac{x_1}{c\sigma}\right)+\frac{\mathcal{E}}{x_1}\mathcal{Q}\left(\frac{x_1}{c\sigma}-\frac{x_1}{2\sigma}\right)\right) \\ &\cdot\log\left(1+\underbrace{\left(\frac{x_1}{\mathcal{E}}-1\right)}_{\leq\frac{x_1}{\mathcal{E}}}e^{-\frac{x_1^2}{c\sigma^2}}\right) \end{aligned} \tag{125}$$

$$\begin{aligned} \leq& \left(\mathcal{Q}\left(\frac{x_1}{2\sigma}+\frac{x_1}{c\sigma}\right)+\frac{\mathcal{E}}{x_1}\mathcal{Q}\left(\frac{x_1}{c\sigma}-\frac{x_1}{2\sigma}\right)\right) \\ &\cdot\log\left(1+\frac{x_1}{\mathcal{E}}e^{-\frac{x_1^2}{c\sigma^2}}\right) \end{aligned} \tag{126}$$

$$\begin{aligned} \leq& \left(\frac{1}{2}e^{-\frac{x_1^2}{2\sigma^2}\left(\frac{1}{2}+\frac{1}{c}\right)^2}+\frac{\mathcal{E}}{x_1}\cdot\frac{1}{2}e^{-\frac{x_1^2}{2\sigma^2}\left(\frac{1}{c}-\frac{1}{2}\right)^2}\right) \\ &\cdot\left(\frac{x_1}{\mathcal{E}}e^{-\frac{x_1^2}{c\sigma^2}}\right) \end{aligned} \tag{127}$$

$$\begin{aligned} =& \left(\frac{1}{2}\left(\frac{\mathcal{E}}{\sigma}\right)^{\frac{c}{2}\left(\frac{1}{2}+\frac{1}{c}\right)^2}+\frac{1}{2\sqrt{c\log\frac{\sigma}{\mathcal{E}}}}\left(\frac{\mathcal{E}}{\sigma}\right)^{\frac{c}{2}\left(\frac{1}{c}-\frac{1}{2}\right)^2+1}\right) \\ &\cdot\sqrt{c}\sqrt{\log\frac{\sigma}{\mathcal{E}}}. \end{aligned} \tag{128}$$

Here, (127) follows from $\mathcal{Q}(\xi) \leq \frac{1}{2}e^{-\frac{\xi^2}{2}}$ and from $\log(1+\xi) \leq \xi$.

Since for $c > 2$ we have $\frac{c}{2}\left(\frac{1}{2}+\frac{1}{c}\right)^2 > 1$ and $\left(\frac{1}{c}-\frac{1}{2}\right)^2 > 0$, we obtain the following limiting behavior:

$$\overline{\lim_{\mathcal{E}\downarrow 0}}\frac{c_{9,3}}{\frac{\mathcal{E}}{\sigma}\sqrt{\log\frac{\sigma}{\mathcal{E}}}} \leq 0. \tag{129}$$

By (119), (123), and (129) we conclude that

$$\overline{\lim_{\mathcal{E}\downarrow 0}} \frac{c_9}{\frac{\mathcal{E}}{\sigma}\sqrt{\log\frac{\sigma}{\mathcal{E}}}} \leq \frac{\sqrt{c}}{2} - \frac{1}{\sqrt{c}} \tag{130}$$

which combined with (103) and (113) yields the claimed behavior.

Appendix D
Proof of Lemma 2

Lemma 2 is a direct consequence of the following proposition.

Proposition 9: Let the random variable X take value in the interval $[0,\mathcal{A}]$, and let $Z \sim \mathcal{N}_{\mathbb{R}}(0,\sigma^2)$ be independent of X. Then, there exists a random variable $\tilde{X}$ taking value in $[0,\mathcal{A}]$ and independent of Z that satisfies

$$\mathsf{E}\left[\tilde{X}\right] = \frac{1}{2}\mathcal{A} \tag{131}$$

and

$$I(\tilde{X};\tilde{X}+Z) \geq I(X;X+Z). \tag{132}$$

Proof: Define $\bar{X} = \mathcal{A} - X$ and note that

$$\begin{aligned} I(X;X+Z) &= I(X;\mathcal{A}-X-Z) && (133)\\ &= I(X;\mathcal{A}-X+Z) && (134)\\ &= I(\mathcal{A}-X;\mathcal{A}-X+Z) && (135)\\ &= I(\bar{X};\bar{X}+Z) && (136)\end{aligned}$$

where (133) and (135) follow because $I(U;V) = I(U;g(V))$ whenever g is one-to-one; where (134) follows from the symmetry of the centered Gaussian; and where (136) follows from the definition of $\bar{X}$.

Let B be a binary random variable that takes on the values 0 and 1 equiprobably and independently of the pair (X,Z). Define the random variable $\tilde{X}$ equal to X when $B=1$ and equal to $\bar{X}$ when $B=0$. We show that $\tilde{X}$ (which takes value in $[0,\mathcal{A}]$) satisfies both (131) and (132). Condition (131) follows by the total law of expectation, by the definition of $\tilde{X}$, by the independence of B and $(X,\bar{X})$, and because $\mathsf{E}\left[\bar{X}\right] = \mathcal{A} - \mathsf{E}[X]$:

$$\begin{aligned} \mathsf{E}[\tilde{X}] &= \frac{1}{2}\mathsf{E}[\tilde{X} \mid B=1] + \frac{1}{2}\mathsf{E}[\tilde{X} \mid B=0]\\ &= \frac{1}{2}\mathsf{E}[X] + \frac{1}{2}\mathsf{E}\left[\bar{X}\right]\\ &= \frac{1}{2}\mathcal{A}. && (137)\end{aligned}$$

Condition (132) follows because conditioning reduces differential entropy, because $\tilde{X}$ is independent of $(X,\bar{X},Z)$, and by (136):

$$\begin{aligned} &I(\tilde{X};\tilde{X}+Z)\\ &= h(\tilde{X}+Z) - h(Z) && (138)\\ &\geq h(\tilde{X}+Z|B) - h(Z) && (139)\\ &= \frac{1}{2}h(\tilde{X}+Z|B=1) + \frac{1}{2}h(\tilde{X}+Z|B=0) - h(Z) && (140)\\ &= \frac{1}{2}\big(h(X+Z)-h(Z)\big) + \frac{1}{2}\big(h(\bar{X}+Z)-h(Z)\big) && (141)\\ &= \frac{1}{2}I(X+Z;X) + \frac{1}{2}I(\bar{X}+Z;\bar{X}) && (142)\\ &= I(X;X+Z). && (143)\end{aligned}$$

■

Acknowledgments

The comments of Ding-Jie Lin, the Associate Editor, and the anonymous reviewers are gratefully acknowledged.

References

[1] S. Hranilovic and F. R. Kschischang, "Capacity bounds for power- and band-limited optical intensity channels corrupted by Gaussian noise," *IEEE Transactions on Information Theory*, vol. 50, no. 5, pp. 784–795, May 2004.

[2] S. M. Moser, "Duality-based bounds on channel capacity," Ph.D. dissertation, ETH Zurich, October 2004, Diss. ETH No. 15769. [Online]. Available: http://moser.cm.nctu.edu.tw/

[3] T. H. Chan, S. Hranilovic, and F. R. Kschischang, "Capacity-achieving probability measure for conditionally Gaussian channels with bounded inputs," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 2073–2088, June 2005.

[4] A. A. Farid and S. Hranilovic, "Upper and lower bounds on the capacity of wireless optical intensity channels," in *Proceedings IEEE International Symposium on Information Theory (ISIT)*, Nice, France, June 24–30, 2007, pp. 2416–2420.

[5] ——, "Design of non-uniform capacity-approaching signaling for optical wireless intensity channels," in *Proceedings IEEE International Symposium on Information Theory (ISIT)*, Toronto, Canada, July 6–11, 2008, pp. 2327–2331.

[6] S. Shamai (Shitz), "Capacity of a pulse amplitude modulated direct detection photon channel," in *Proceedings of the IEE*, vol. 137, part I (Communications, Speech and Vision), no. 6, December 1990, pp. 424–430.

[7] D. Brady and S. Verdú, "The asymptotic capacity of the direct detection photon channel with a bandwidth constraint," in *Proceedings Twenty-Eighth Allerton Conference on Communication, Control and Computing*, Allerton House, Monticello, IL, USA, October 3–5, 1990, pp. 691–700.

[8] A. Lapidoth and S. M. Moser, "On the capacity of the discrete-time Poisson channel," *IEEE Transactions on Information Theory*, vol. 55, no. 1, pp. 303–322, January 2009.

[9] Y. Kabanov, "The capacity of a channel of the Poisson type," *Theory of Probability and Its Applications*, vol. 23, pp. 143–147, 1978.

[10] M. H. A. Davis, "Capacity and cutoff rate for Poisson-type channels," *IEEE Transactions on Information Theory*, vol. 26, no. 6, pp. 710–715, November 1980.

[11] A. D. Wyner, "Capacity and error exponent for the direct detection photon channel — part I and II," *IEEE Transactions on Information Theory*, vol. 34, no. 6, pp. 1462–1471, November 1988.

[12] M. R. Frey, "Capacity of the L_p norm-constrained Poisson channel," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 445–450, March 1992.

[13] ——, "Information capacity of the Poisson channel," *IEEE Transactions on Information Theory*, vol. 37, no. 2, pp. 244–256, March 1991.

[14] S. Shamai (Shitz) and A. Lapidoth, "Bounds on the capacity of a spectrally constrained Poisson channel," *IEEE Transactions on Information Theory*, vol. 39, no. 1, pp. 19–29, January 1993.

[15] I. Bar-David and G. Kaplan, "Information rates of photon-limited overlapping pulse position modulation channels," *IEEE Transactions on Information Theory*, vol. 30, no. 3, pp. 455–464, May 1984.

[16] A. Lapidoth and S. M. Moser, "On the capacity of an optical intensity channel with input-dependent noise," 2009, in preparation.

[17] ——, "Capacity bounds via duality with applications to multiple-antenna systems on flat fading channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2426–2467, October 2003.

[18] F. Topsøe, "An information theoretical identity and a problem involving capacity," *Studia Scientarum Mathematicarum Hungarica*, vol. 2, pp. 291–292, 1967.

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

[20] J. G. Smith, "The information capacity of amplitude- and variance-constrained scalar Gaussian channels," *Information and Control*, vol. 18, no. 3, pp. 203–219, February 1971.

[21] 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 Transactions on Information Theory*, vol. 39, no. 5, pp. 1728–1735, September 1993.
[22] M. A. Wigger, "Bounds on the capacity of free-space optical intensity channels," Master's thesis, Signal and Information Processing Laboratory, ETH Zurich, Switzerland, March 2003, supervised by Prof. Dr. Amos Lapidoth.
[23] K.-P. Ho, *Phase-Modulated Optical Communication Systems*. Springer Verlag, 2005.
[24] J. M. Kahn and J. R. Barry, "Wireless infrared communications," *Proceedings of the IEEE*, vol. 85, no. 2, pp. 265–298, February 1997.
[25] A. Lapidoth, "On phase noise channels at high SNR," in *Proceedings IEEE Information Theory Workshop (ITW)*, Bangalore, India, October 20–25, 2002, pp. 1–4.
[26] M. Katz and S. Shamai, "On the capacity-achieving distribution of the discrete-time noncoherent and partially coherent AWGN channels," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2257–2270, October 2004.
[27] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423 and 623–656, July and October 1948.
[28] S. Verdú, "On channel capacity per unit cost," *IEEE Transactions on Information Theory*, vol. 36, no. 5, pp. 1019–1030, September 1990.

Amos Lapidoth (S'89–M'95–SM'00–F'04) received the B.A. degree in mathematics (*summa cum laude*, 1986), the B.Sc. degree in electrical engineering (*summa cum laude*, 1986), and the M.Sc. degree in electrical engineering (1990), all from the Technion–Israel Institute of Technology, Haifa. He received the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1995.

During 1995–1999, he was an Assistant and Associate Professor in the Department of Electrical Engineering and Computer Science at the Massachusetts Institute of Technology, Cambridge, and was the KDD Career Development Associate Professor in Communications and Technology. He is now Professor of Information Theory at the Signal and Information Processing Laboratory, ETH Zurich, Switzerland. His research interests are in digital communications and information theory.

Dr. Lapidoth currently serves and has also served in the years 2003–2004 as Associate Editor for Shannon Theory for the IEEE TRANSACTIONS ON INFORMATION THEORY.

Stefan M. Moser (S'01–M'05) was born in Switzerland. He received the M.Sc. degree in electrical engineering (with distinction) in 1999, the M.Sc. degree in industrial management (M.B.A.) in 2003, and the Ph.D. degree in the field of information theory in 2004, all from ETH Zurich, Switzerland.

During 1999 to 2003, he was a Research and Teaching Assistant, and from 2004 to 2005, he was a Senior Research Assistant with the Signal and Information Processing Laboratory, ETH Zurich. Since August 2005, he has been an Assistant Professor, and since August 2008, an Associate Professor with the Department of Electrical Engineering, National Chiao Tung University (NCTU), Hsinchu, Taiwan. His research interests are in information theory and digital communications.

Dr. Moser received the National Chiao Tung University Outstanding Researchers Award in 2007, 2008, and 2009, the National Chiao Tung University Excellent Teaching Award and the National Chiao Tung University Outstanding Mentoring Award both in 2007, the Willi Studer Award of ETH in 1999, and the ETH Medal for an excellent Master's thesis in 1999.

Michèle A. Wigger (S'05–M'08) received the M.Sc. degree in electrical engineering (with distinction) and the Ph.D. degree in electrical engineering both from ETH Zurich in 2003 and 2008, respectively.

Currently, she is a postdoctoral researcher at the Department of Electrical and Computer Engineering, University of California, San Diego. Her research interests are in the areas of information theory, wireless communication networks, feedback channels, and channels with states.