On the Discreteness of Capacity-Achieving Distributions for Fading and Signal-Dependent Noise Channels With Amplitude-Limited Inputs

Ahmad ElMoslimany, *Student Member, IEEE*, and Tolga M. Duman, *Fellow, IEEE*

Abstract— We address the problem of finding the capacity of two classes of channels with amplitude-limited inputs. The first class is frequency flat fading channels with an arbitrary (but finite support) channel gain with the channel state information available only at the receiver side; while the second one we consider is the class of additive noise channels with signal-dependent Gaussian noise. We show that for both channel models and under some regularity conditions, the capacity-achieving distribution is discrete with a finite number of mass points. Furthermore, finding the capacity-achieving distribution turns out to be a finite-dimensional optimization problem, and efficient numerical algorithms can be developed using standard optimization techniques to compute the channel capacity. We demonstrate our findings via several examples. In particular, we present an example for a block fading channel where the channel gain follows a truncated Rayleigh distribution, and two instances of signal-dependent noise that are used in the literature of magnetic recording and optical communication channels.

Index Terms— Fading channels, signal-dependent noise, amplitude-limited inputs, peak power constraints, channel capacity.

I. INTRODUCTION

SINCE the formulation of the relevant information theoretic problem about fifty years ago, there has been much progress on the characterization of channel capacity with the practical constraint of amplitude limited inputs. Smith [2] studied the capacity of a scalar Gaussian channel under peak and average power constraints, and showed that there is a unique optimal distribution that maximizes the mutual information and this distribution has a finite number of mass points. Tchamkerten [3] generalized Smith's results to channels with additive noise, which is not necessarily Gaussian, and proved that the capacity-achieving distribution under certain conditions on the noise distribution is also discrete.

Discrete distributions appear as optimal inputs in other cases as well. As an example, quadrature Gaussian channels are studied in [4], where the authors show that the capacity-achieving distribution has a uniformly distributed phase and discrete amplitude. The authors in [5] consider noncoherent additive white Gaussian noise (AWGN) channels and prove that the optimal input distribution is discrete. They also compute tight lower bounds on the capacity of the channel based on a close examination of some suboptimal input distributions. The authors in [6]–[8] study the capacity of the Gaussian Multiple Access Channels (MACs) with amplitude-constrained inputs. They show that the sum-capacity achieving distribution is discrete, and that discrete distributions achieve rates at any of the corner points of the capacity region. In other related work, transmissions over peak-power constrained channels are considered in [9]–[11], and multiple-input multiple-output (MIMO) channels with amplitude-limited inputs are studied in [12]. In [13], the authors study conditionally Gaussian channels with amplitude-limited inputs for which the distribution of the channel output conditioned on the channel input is Gaussian. They show that the channel capacity is achieved by a discrete input distribution with a finite number of mass points.

In this paper, we consider two channel models. The first one covers fading channels with amplitude-limited inputs where the channel coefficients have a finite support and the channel state information is available only at the receiver side; while the second one covers signal-dependent additive Gaussian noise channels such as those encountered in magnetic recording systems and optical communication channels.

The capacity of fading channels with different constraints on the inputs has been studied in the previous literature for certain fading distributions. For instance, in [14] the authors consider transmission over Rayleigh fading channels with average power constrained inputs where neither the transmitter nor the receiver has the channel state information. In [15], the authors investigate the capacity of Rician fading channels with inputs having constraints on the second and the fourth moments. The capacity-achieving distribution is shown to be discrete with a finite number of mass points for both cases.

Certain aspects of the problem of finding the capacity of signal-dependent Gaussian noise channels with amplitude-limited inputs have also been studied previously. For example, upper and lower bounds on the capacity of optical intensity channels with input-dependent Gaussian noise where the

variance of the noise varies linearly with the input are derived in [16]. The upper bound relies on a dual expression for the channel capacity and a new notion called the capacity-achieving input distributions that escape to infinity. The lower bound is based on a lower bound on the differential entropy of the channel output in terms of the differential entropy of the input.

We draw particular attention to [13] where the general framework adopted covers certain classes of fading channels and signal-dependent Gaussian channels encompassing a wide range of scenarios with amplitude-limited inputs. We note, however, that neither of our models falls within its framework. For the case of fading channels, [13] covers the case of complex Gaussian channel coefficients while our model considers arbitrary channel coefficients with finite support. For the case of signal-dependent noise channels, the results of [13] apply for a class of noise variance functions, which converge to zero for some limit points while we impose other technical conditions. We will provide a specific example of a channel model adopted in the literature of magnetic recording systems, which falls within our framework but not that of [13] in the numerical examples section.

We show that the capacity of fading channels with amplitude-limited inputs is achieved by a unique input distribution and when the channel coefficients have a finite support, the capacity-achieving distribution is discrete with a finite number of mass points. For signal-dependent Gaussian channels with noise variance that depends on the transmitted signal, we prove that, under some technical conditions, the capacity is achieved by a discrete input distribution with a finite number of mass points as well. In solving both problems, we extend techniques utilized in [17] and [6]. For both channel models considered, we prove that the mutual information is a concave, continuous and weakly differentiable function of the input distribution and the space of cumulative distribution functions is compact and convex. Hence, the problem is a convex optimization problem, which facilitates deriving conditions on the optimal input distribution. The proof of the discreteness of the optimal input distribution is achieved by adopting some techniques from complex analysis.

The paper is organized as follows: in Sections II and III, we describe the two channel models under consideration and provide the required preliminaries for the rest of the paper. In Section IV, we prove that the capacity of fading channels with an arbitrary but finite support distribution on the channel coefficients and amplitude-limited inputs is achieved by discrete inputs with a finite number of mass points. In Section V, we show that under some technical conditions, the capacity of general signal-dependent Gaussian noise channels is achieved by a discrete input distribution as well. In Section VI, we present some numerical examples that illustrate our results. Finally, we conclude the paper in Section VII.

A. Notation

Unless mentioned otherwise, the notation is as follows. The set of real numbers, complex numbers and natural numbers are denoted by $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{N}$, respectively. We denote other sets by calligraphic letters such as $\mathcal{F}$. For a complex number z, the real part is denoted by $\text{Re}\{z\}$ while the imaginary part is denoted by $\text{Im}\{z\}$. Random variables are written in bold lowercase letters, e.g., $\mathbf{x}$, and their realizations are shown using standard lowercase letters, e.g., x. Constants are shown using capital letters. The probability density function (PDF) of a random variable $\mathbf{x}$ is denoted as $f_{\mathbf{x}}(x)$, and the corresponding cumulative distribution function (CDF) is referred to as $F_{\mathbf{x}}(x)$. The PDF of a channel output $\mathbf{y}$ for an input $\mathbf{x}$ with CDF $F_{\mathbf{x}}$ is denoted by $f_{\mathbf{y}}(y; F_{\mathbf{x}})$ and the corresponding CDF is denoted by $F_{\mathbf{y}}(y; F_{\mathbf{x}})$. We refer to the PDF of a random variable $\mathbf{x}$ whose density depends on a parameter z as $f_{\mathbf{x}}(x, z)$. We reserve the notation $\mathbf{E}$ for the expectation operator, i.e., the expectation of a function of a random variable $\mathbf{x}$ with a CDF $F_{\mathbf{x}}$ is written as $\mathbf{E}_{F_{\mathbf{x}}}[g(\mathbf{x})]$. We denote the differential entropy of the random variable $\mathbf{x}$ by $H_{F_{\mathbf{x}}}(\mathbf{x})$ and its entropy density by $h_{F_{\mathbf{x}}}(x)$. The average mutual information functional between the channel input $\mathbf{x}$ with distribution $F_{\mathbf{x}}$ and the corresponding output $\mathbf{y}$ is denoted by $I_{F_{\mathbf{x}}}(\mathbf{x}; \mathbf{y})$, and the information density is denoted by $i_{F_{\mathbf{x}}}(x)$. The conditional entropy of the channel output $\mathbf{y}$ conditioned on channel state $\mathbf{u}$ for the corresponding input distribution $F_{\mathbf{x}}$ is denoted by $H_{F_{\mathbf{x}}}(\mathbf{y}|\mathbf{u})$ while the conditional entropy given that the channel state takes on a specific value u is shown as $H_{F_{\mathbf{x}}}(\mathbf{y}|\mathbf{u} = u)$. The conditional mutual information between the channel input $\mathbf{x}$ and the output $\mathbf{y}$ conditioned on the channel state $\mathbf{u}$ is denoted by $I_{F_{\mathbf{x}}}(\mathbf{x}; \mathbf{y}|\mathbf{u})$ while the conditional mutual information for a specific channel state value u is denoted by $I_{F_{\mathbf{x}}}(\mathbf{x}; \mathbf{y}|\mathbf{u} = u)$. Finally, the conditional mutual information density between the channel input $\mathbf{x}$ and the output $\mathbf{y}$ conditioned on the channel state $\mathbf{u}$ is denoted by $i_{F_{\mathbf{x}}}(x|\mathbf{u})$ while the condition mutual information density for a specific channel state u is denoted by $i_{F_{\mathbf{x}}}(x|\mathbf{u} = u)$.

II. CHANNEL MODELS

In this section, we describe the two channel models under consideration, namely, fading channels with amplitude-limited inputs and signal dependent additive Gaussian noise channels.

A. Fading Channels With Amplitude-Limited Inputs

The received signal $\mathbf{y}$ is given by

$$\mathbf{y} = \mathbf{ux} + \mathbf{z}, \qquad (1)$$

where $\mathbf{x}$ is the channel input that is amplitude constrained to $[-A, A]$, i.e., it has a CDF $F_{\mathbf{x}}$ that belongs to the class of CDFs $\mathcal{F}_{\mathbf{x}}$ such that for any $F_{\mathbf{x}} \in \mathcal{F}_{\mathbf{x}}$, $F_{\mathbf{x}}(x) = 0$ for $x < -A$ and $F_{\mathbf{x}}(x) = 1$ for $x \geq A$. The random variable $\mathbf{u}$ is the fading channel coefficient with a CDF $F_{\mathbf{u}}$. We assume that $\mathbf{u}$ has a finite support, i.e., for a fixed $\epsilon_0 > 0$, $\mathbf{u} \in [\epsilon_0, U]$ for some $U < \infty$.[1] Note that most of the results throughout the paper apply for the case $\mathbf{u} \in [0, U]$, however,

---

[1]We note that, this restriction will be used in the discreteness proof to be able to make some technical arguments. It is clear that this model can be used as an arbitrarily tight approximation to other fading channel models with coefficients on the interval $[0, \infty)$ (e.g., Rayleigh fading) by properly selecting ϵ_0 and U.

the specific argument on the discreteness of the capacity-achieving distribution requires $u > 0$. We also note that the channel state information is available only at the receiver side. The noise $\mathbf{z}$ is additive white Gaussian with zero mean and variance $\sigma_{\mathbf{z}}^2$, i.e., $\mathbf{z} \sim \mathcal{N}(0, \sigma_{\mathbf{z}}^2)$. We assume that the input $\mathbf{x}$ and the fading coefficient $\mathbf{u}$ are independent. We also assume an ergodic block fading channel model where coding is performed over multiple fading blocks.

We emphasize that this model differs from the previous models studied in the literature in the context of amplitude limited inputs. The most closely related one is in [13] where the authors study the conditionally Gaussian channels. When the fading gain is zero mean complex Gaussian (i.e., for Rayleigh fading), the channel becomes conditionally Gaussian, and the results of [13] apply. However, here we consider fading channels with an arbitrary (but finite support) distribution, hence our model does not fall within this framework.

The PDF of the channel output $\mathbf{y}$ is given by

$$f_{\mathbf{y}}(y; F_{\mathbf{x}}) = \int_0^U \int_{-A}^A f_{\mathbf{z}}(y - ux)\, dF_{\mathbf{x}}(x)\, dF_{\mathbf{u}}(u),$$

where $f_{\mathbf{z}}(y - ux) = f_{\mathbf{y}|\mathbf{x},\mathbf{u}}(y|x, u)$ is the PDF of the channel output $\mathbf{y}$ conditioned on specific values of $\mathbf{x}$ and $\mathbf{u}$, and $f_{\mathbf{y}}(y; F_{\mathbf{x}})$ is the PDF of the channel output $\mathbf{y}$ when the input has a CDF $F_{\mathbf{x}}$. The existence of $f_{\mathbf{y}}(y; F_{\mathbf{x}})$ can be verified by computing the CDF of the channel output

$$\begin{aligned}
F_{\mathbf{y}}(y; F_{\mathbf{x}}) &= \int_0^U \int_{-A}^A F_{\mathbf{z}}(y - ux)\, dF_{\mathbf{x}}(x)\, dF_{\mathbf{u}}(u), \\
&= \int_0^U \int_{-A}^A \int_{-\infty}^y f_{\mathbf{z}}(y' - ux)\, dy'\, dF_{\mathbf{x}}(x)\, dF_{\mathbf{u}}(u), \\
&= \int_{-\infty}^y \int_0^U \int_{-A}^A f_{\mathbf{z}}(y' - ux)\, dF_{\mathbf{x}}(x)\, dF_{\mathbf{u}}(u)\, dy', \\
&= \int_{-\infty}^y f_{\mathbf{y}}(y'; F_{\mathbf{x}})\, dy',
\end{aligned}$$

where $F_{\mathbf{z}}(\cdot)$ is the CDF of the noise. The interchange of the order of the integrals is justified since the PDF of the noise $f_{\mathbf{z}}(z)$ is nonnegative and integrable. Note that with the same reasoning we can also write

$$\begin{aligned}
f_{\mathbf{y}}(y; F_{\mathbf{x}}) &= \int_0^U \int_{-A}^A f_{\mathbf{z}}(y - ux)\, dF_{\mathbf{x}}(x)\, dF_{\mathbf{u}}(u), \\
&= \int_{-A}^A \int_0^U f_{\mathbf{z}}(y - ux)\, dF_{\mathbf{u}}(u)\, dF_{\mathbf{x}}(x).
\end{aligned}$$

In the following, we provide bounds on the PDF of the noise $f_{\mathbf{z}}(y - ux)$ and the conditional PDF $f_{\mathbf{y}|\mathbf{u}}(y|u)$ for later use. Following similar arguments as in Smith [2], it is straightforward to show that, for $u > 0$, the noise PDF evaluated at $y - ux$ is bounded as follows

$$\gamma(y, u) \leq f_{\mathbf{z}}(y - ux) \leq \Gamma(y, u), \qquad (2)$$

where

$$\gamma(y, u) = \begin{cases} k_1 \exp(-k_2(y - uA)^2) & \text{if } y \leq 0, \\ k_1 \exp(-k_2(y + uA)^2) & \text{if } y > 0, \end{cases} \qquad (3)$$

and

$$\Gamma(y, u) = \begin{cases} k_3 \exp(-k_4(y + uA)^2) & \text{if } y \leq -uA, \\ k_3 & \text{if } y \in [-uA, uA], \\ k_3 \exp(-k_4(y - uA)^2) & \text{if } y > uA, \end{cases} \qquad (4)$$

for some finite and positive k_1, k_2, k_3, and k_4. Let us also define

$$q(m) \triangleq k_1 \exp(-k_2 m^2). \qquad (5)$$

Clearly, $q(y - ux) \leq f_{\mathbf{z}}(y - ux)$.

We have the conditional PDF $f_{\mathbf{y}|\mathbf{u}}(y|u; F_{\mathbf{x}})$ as

$$f_{\mathbf{y}|\mathbf{u}}(y|u; F_{\mathbf{x}}) = \int_{-A}^A f_{\mathbf{z}}(y - ux)\, dF_{\mathbf{x}}(x). \qquad (6)$$

Hence, using (2), we can also write

$$\gamma(y, u) \leq f_{\mathbf{y}|\mathbf{u}}(y|u; F_{\mathbf{x}}) \leq \Gamma(y, u). \qquad (7)$$

B. Signal-Dependent Additive Noise Channels

For the case of additive noise channels with signal dependent Gaussian noise, the received signal $\mathbf{y}$ is given by

$$\mathbf{y} = \mathbf{x} + \mathbf{n}, \qquad (8)$$

where $\mathbf{x}$ is the channel input, $\mathbf{n}$ is the noise term, and $\mathbf{y}$ is the channel output. As in Section II-A, we assume that the random variable $\mathbf{x}$ is subject to an amplitude-constraint such that $|\mathbf{x}| \leq A$ for some $A > 0$. Let $\mathcal{F}_{\mathbf{x}}$ denote the corresponding class of input CDFs, i.e., $F_{\mathbf{x}}$ in $\mathcal{F}_{\mathbf{x}}$ implies $F_{\mathbf{x}} = 0$ for all $x < -A$ and $F_{\mathbf{x}} = 1$ for all $x \geq A$. The additive Gaussian noise $\mathbf{n}$ has a zero mean, and its variance has two components: one is independent of the input signal and the other depends on the input, i.e., $\sigma_{\mathbf{n}}^2(x) = \sigma_0^2 + \sigma_1^2(x)$, if $\mathbf{x} = x$ with $|x| \leq A$. In other words, the noise term conditioned on the channel input $f_{\mathbf{n}|\mathbf{x}}(n|x)$ is a zero mean Gaussian random variable with variance $\sigma_{\mathbf{n}}^2(x)$.

We assume that $\sigma_1^2(x)$ is a bounded, continuous and differentiable function of the input x. We also note that we can consider an arbitrary extension of the noise variance $\sigma_{\mathbf{n}}^2(x)$ for $|x| > A$ as this will not change the capacity of the channel (under the specific input constraint). However, we assume that an extension for $|x| > A$, which satisfies the technical conditions stated in IV.D, is utilized to make sure that the arguments in that section follow.

The existence of $f_{\mathbf{y}}(y; F_{\mathbf{x}})$ is guaranteed since the noise PDF $f_{\mathbf{n}}(n)$ is a nonnegative and integrable function, i.e., we have

$$f_{\mathbf{y}}(y; F_{\mathbf{x}}) = \int_{-A}^A f_{\mathbf{n}}(y - x, x)\, dF_{\mathbf{x}}(x),$$

where

$$f_{\mathbf{n}}(y - x, x) = \frac{1}{\sqrt{2\pi(\sigma_0^2 + \sigma_1^2(x))}} \exp\left[-\frac{(y - x)^2}{2(\sigma_0^2 + \sigma_1^2(x))}\right],$$

which can be bounded as

$$k_1' \exp\left[-k_2'(y-x)^2\right] \le f_{\mathbf{n}}(y-x, x) \le k_3' \exp\left[-k_4'(y-x)^2\right],$$

for some positive finite values k_1', k_2', k_3', k_4' where

$$k_1' \le \min_{|x| \le A} \frac{1}{\sqrt{2\pi(\sigma_0^2 + \sigma_1^2(x))}}, \quad k_2' \ge \max_{|x| \le A} \frac{1}{2(\sigma_0^2 + \sigma_1^2(x))},$$

$$k_3' \ge \max_{|x| \le A} \frac{1}{\sqrt{2\pi(\sigma_0^2 + \sigma_1^2(x))}}, \quad k_4' \le \min_{|x| \le A} \frac{1}{2(\sigma_0^2 + \sigma_1^2(x))}.$$

Define the functions $q'(m)$, $\gamma'(y)$, and $\Gamma'(y)$ as

$$q'(m) = k_1' \exp(-k_2' m^2), \tag{9}$$

$$\gamma'(y) = \begin{cases} k_1' \exp(-k_2'(y - A)^2) & \text{if } y \le 0, \\ k_1' \exp(-k_2'(y + A)^2) & \text{if } y > 0, \end{cases} \tag{10}$$

and

$$\Gamma'(y) = \begin{cases} k_3' \exp(-k_4'(y + A)^2) & \text{if } y \le -A, \\ k_3' & \text{if } y \in [-A, A], \\ k_3' \exp(-k_4'(y - A)^2) & \text{if } y > A, \end{cases} \tag{11}$$

respectively. It is easy to see that the output PDF $f_{\mathbf{y}}(y; F_{\mathbf{x}})$ is bounded as

$$\gamma'(y) \le f_{\mathbf{y}}(y; F_{\mathbf{x}}) \le \Gamma'(y), \tag{12}$$

and

$$q'(y - x) \le f_{\mathbf{n}}(y - x, x). \tag{13}$$

III. DEFINITIONS AND PRELIMINARIES

In this section, we provide some definitions and the preliminaries required for the rest of the paper.

A. Preliminaries

In this subsection, we provide statements of several theorems and lemmas that will be used throughout the paper.

- *Helly-Bray Theorem [18]:* Let $F_{\mathbf{x}}$ and $F_{\mathbf{x}}^{(1)}, F_{\mathbf{x}}^{(2)}, \ldots, F_{\mathbf{x}}^{(n)}$ be CDFs on $\mathbb{R}$. The Helly-Bray theorem states that if $F_{\mathbf{x}}^{(n)} \to F_{\mathbf{x}}$ (weak convergence), then

$$\int_{\mathbb{R}} g(x) \, dF_{\mathbf{x}}^{(n)}(x) \xrightarrow[n \to \infty]{} \int_{\mathbb{R}} g(x) \, dF_{\mathbf{x}}(x),$$

  for each bounded and continuous function $g(x)$.

- *The Identity Theorem [19, p. 78]:* Suppose $f : \Omega \to \mathbb{C}$ is holomorphic (complex analytic) and $Z_f = \{z \in \Omega : f(z) = 0\}$. Then either $Z_f = \Omega$, or Z_f has no limit points in Ω.

- *Corollary to the Identity Theorem [19, p. 79]:* Suppose f and g are holomorphic in a region Ω and $f(z) = g(z)$ for all z in some nonempty open subset of Ω (or, more generally for z in some sequence of distinct points with a limit point in Ω). Then $f(z) = g(z)$ throughout Ω.

- *Ash's Lemma [20]:* Let $f(x)$ and $g(x)$ be arbitrary PDFs, if

$$-\int f(x) \log g(x) \, dx < \infty,$$

then

$$-\int f(x) \log f(x) \, dx \le -\int f(x) \log g(x) \, dx < \infty. \tag{14}$$

- *Bounding the Logarithm of a Bounded PDF:* Let $f : \mathbb{R} \to \mathbb{R}$ be a positive-valued and bounded function, i.e., $0 \le f(x) \le c < \infty$ for any $x \in \mathbb{R}$ and some positive constant c. We have

$$|\log(f(x))| + \log(f(x)) = \begin{cases} 0 & \log(f(x)) \le 0, \\ 2\log(f(x)) & \log(f(x)) > 0, \end{cases}$$

  and

$$\log(f(x)) \le \log c, \quad \forall x \in \mathbb{R}.$$

  Thus,

$$|\log(f(x))| \le -\log(f(x)) + 2|\log c|. \tag{15}$$

  We note that this bounding approach was used in [6] as well.

B. Fading Channels

The average mutual information between the input and the output conditioned on the channel gain $\mathbf{u}$ is defined as [21], [22]

$$I_{F_{\mathbf{x}}}(\mathbf{x}; \mathbf{y}|\mathbf{u}) \triangleq \int_0^U I_{F_{\mathbf{x}}}(\mathbf{x}; \mathbf{y}|\mathbf{u} = u) \, dF_{\mathbf{u}}(u), \tag{16}$$

where

$$I_{F_{\mathbf{x}}}(\mathbf{x}; \mathbf{y}|\mathbf{u} = u)$$
$$\triangleq \int_{-\infty}^{\infty} \int_{-A}^{A} f_{\mathbf{z}}(y - ux) \log\left(\frac{f_{\mathbf{z}}(y - ux)}{f_{\mathbf{y}|\mathbf{u}}(y|u; F_{\mathbf{x}})}\right) dF_{\mathbf{x}}(x) \, dy. \tag{17}$$

We define the conditional entropy $H_{F_{\mathbf{x}}}(\mathbf{y}|\mathbf{u})$ as

$$H_{F_{\mathbf{x}}}(\mathbf{y}|\mathbf{u}) \triangleq \int_0^U H_{F_{\mathbf{x}}}(\mathbf{y}|\mathbf{u} = u) \, dF_{\mathbf{u}}(u), \tag{18}$$

where

$$H_{F_{\mathbf{x}}}(\mathbf{y}|\mathbf{u} = u) \triangleq -\int_{-\infty}^{\infty} f_{\mathbf{y}|\mathbf{u}}(y|u; F_{\mathbf{x}}) \log f_{\mathbf{y}|\mathbf{u}}(y|u; F_{\mathbf{x}}) \, dy. \tag{19}$$

For noise with finite variance and bounded PDF, the conditional mutual information can be written as,

$$I_{F_{\mathbf{x}}}(\mathbf{x}; \mathbf{y}|\mathbf{u}) = H_{F_{\mathbf{x}}}(\mathbf{y}|\mathbf{u}) - D_{\mathbf{z}}, \tag{20}$$

where $D_{\mathbf{z}}$ is the noise entropy which is defined as

$$D_{\mathbf{z}} \triangleq -\int_{-\infty}^{\infty} f_{\mathbf{z}}(z) \log f_{\mathbf{z}}(z) \, dz. \tag{21}$$

For a Gaussian noise density with zero mean and variance $\sigma_{\mathbf{z}}^2$, as considered in this section,

$$D_{\mathbf{z}} = \frac{1}{2} \log\left(2\pi e \sigma_{\mathbf{z}}^2\right).$$

The channel capacity is defined as[2]

$$C_1 \triangleq \sup_{F_{\mathbf{x}} \in \mathcal{F}_{\mathbf{x}}} I_{F_{\mathbf{x}}}(\mathbf{x}; \mathbf{y} | \mathbf{u}). \tag{22}$$

We define the conditional mutual information density and the conditional entropy density conditioned on a specific value of $\mathbf{u}$ as

$$i_{F_{\mathbf{x}}}(x | \mathbf{u} = u) \triangleq \int_{-\infty}^{\infty} f_{\mathbf{z}}(y - ux) \log \frac{f_{\mathbf{z}}(y - ux)}{f_{\mathbf{y}|\mathbf{u}}(y | u; F_{\mathbf{x}})} \, dy,$$

$$h_{F_{\mathbf{x}}}(x | \mathbf{u} = u) \triangleq - \int_{-\infty}^{\infty} f_{\mathbf{z}}(y - ux) \log f_{\mathbf{y}|\mathbf{u}}(y | u; F_{\mathbf{x}}) \, dy,$$

respectively. Clearly, the following equality holds

$$i_{F_{\mathbf{x}}}(x | \mathbf{u} = u) = h_{F_{\mathbf{x}}}(x | \mathbf{u} = u) - D_{\mathbf{z}}.$$

Define the conditional mutual information density and the conditional entropy density as

$$i_{F_{\mathbf{x}}}(x | \mathbf{u}) \triangleq \int_0^U i_{F_{\mathbf{x}}}(x | \mathbf{u} = u) \, dF_{\mathbf{u}}(u), \tag{23}$$

$$h_{F_{\mathbf{x}}}(x | \mathbf{u}) \triangleq \int_0^U h_{F_{\mathbf{x}}}(x | \mathbf{u} = u) \, dF_{\mathbf{u}}(u), \tag{24}$$

respectively. We can write

$$i_{F_{\mathbf{x}}}(x | \mathbf{u}) = h_{F_{\mathbf{x}}}(x | \mathbf{u}) - D_{\mathbf{z}},$$

$$I_{F_{\mathbf{x}}}(\mathbf{x}; \mathbf{y} | \mathbf{u}) = \int_{-A}^A i_{F_{\mathbf{x}}}(x | \mathbf{u}) \, dF_{\mathbf{x}}(x),$$

$$H_{F_{\mathbf{x}}}(\mathbf{y} | \mathbf{u}) = \int_{-A}^A h_{F_{\mathbf{x}}}(x | \mathbf{u}) \, dF_{\mathbf{x}}(x).$$

We note that the change of the integration orders above is justified by using Fubini's theorem, which applies since the mutual information density $i_{F_{\mathbf{x}}}(x | \mathbf{u})$ and the entropy density $h_{F_{\mathbf{x}}}(x | \mathbf{u})$ are finite as proved next.

Lemma 1: The conditional entropy density $h_{F_{\mathbf{x}}}(x | \mathbf{u})$, the conditional entropy $H_{F_{\mathbf{x}}}(\mathbf{y} | \mathbf{u})$, the conditional mutual information density $i_{F_{\mathbf{x}}}(x | \mathbf{u})$, and the conditional mutual information $I_{F_{\mathbf{x}}}(\mathbf{x}; \mathbf{y} | \mathbf{u})$ are finite.

Proof: To show the finiteness of $i_{F_{\mathbf{x}}}(x | \mathbf{u})$, it is sufficient to show the finiteness of $h_{F_{\mathbf{x}}}(x | \mathbf{u})$ as the difference between them is constant. That is,

$$|h_{F_{\mathbf{x}}}(x | \mathbf{u})|$$

$$= \left| \int_0^U h_{F_{\mathbf{x}}}(x | \mathbf{u} = u) \, dF_{\mathbf{u}}(u) \right|,$$

$$\leq \int_0^U \int_{-\infty}^{\infty} \left| f_{\mathbf{z}}(y - ux) \log(f_{\mathbf{y}|\mathbf{u}}(y | u; F_{\mathbf{x}})) \right| \, dy \, dF_{\mathbf{u}}(u),$$

$$\leq \int_0^U \int_{-\infty}^{\infty} f_{\mathbf{z}}(y - ux) [- \log(f_{\mathbf{y}|\mathbf{u}}(y | u; F_{\mathbf{x}}))$$
$$+ 2 | \log(k_3) |] \, dy \, dF_{\mathbf{u}}(u),$$

$$\leq \int_0^U \int_{-\infty}^{\infty} \Gamma(y, u) [- \log(\gamma(y, u)) + 2 | \log(k_3) |] \, dy \, dF_{\mathbf{u}}(u),$$

[2]We reserve C_1 to denote the capacity of fading channels with amplitude-limited inputs while C_2 is used to denote the capacity of signal-dependent Gaussian channels.

which is integrable. Hence, we can conclude the finiteness of $h_{F_{\mathbf{x}}}(x | \mathbf{u})$ and $i_{F_{\mathbf{x}}}(x | \mathbf{u})$, i.e., $|h_{F_{\mathbf{x}}}(x | \mathbf{u})| \leq L < \infty$. On the other hand, for the conditional entropy $H_{F_{\mathbf{x}}}(\mathbf{y} | \mathbf{u})$, we have

$$|H_{F_{\mathbf{x}}}(\mathbf{y} | \mathbf{u})| = \left| \int_{-A}^A h_{F_{\mathbf{x}}}(x | \mathbf{u}) \, dF_{\mathbf{x}}(x) \right|,$$

$$\leq \int_{-A}^A \left| h_{F_{\mathbf{x}}}(x | \mathbf{u}) \right| \, dF_{\mathbf{x}}(x),$$

$$= L < \infty. \tag{25}$$

Thus, we can conclude that $H_{F_{\mathbf{x}}}(\mathbf{y} | \mathbf{u})$ and $I_{F_{\mathbf{x}}}(\mathbf{x}; \mathbf{y} | \mathbf{u})$ are both finite as well. $\square$

C. Signal-Dependent Additive Noise Channels

The conditional entropy $H_{F_{\mathbf{x}}}(\mathbf{y} | \mathbf{x})$ (for a given input distribution $F_{\mathbf{x}}$) is given by

$$H_{F_{\mathbf{x}}}(\mathbf{y} | \mathbf{x}) = \int_{-A}^A H(\mathbf{y} | \mathbf{x} = x) \, dF_{\mathbf{x}}(x),$$

$$= \int_{-A}^A \frac{1}{2} \log \left(2\pi e \, (\sigma_0^2 + \sigma_1^2(x)) \right) dF_{\mathbf{x}}(x),$$

$$= \int_{-A}^A \frac{1}{2} \log(2\pi e \sigma_0^2) \, dF_{\mathbf{x}}(x)$$

$$+ \frac{1}{2} \int_{-A}^A \log \left(1 + \frac{\sigma_1^2(x)}{\sigma_0^2} \right) dF_{\mathbf{x}}(x),$$

$$= \frac{1}{2} \log(2\pi e \sigma_0^2) + \frac{1}{2} \mathbf{E}_{F_{\mathbf{x}}} \left[\log(\sigma^2(\mathbf{x})) \right], \tag{26}$$

where $\sigma^2(\mathbf{x}) = 1 + \frac{\sigma_1^2(\mathbf{x})}{\sigma_0^2}$. We note that the function $\sigma^2(x)$ is bounded (by assumption), and it is greater than or equal to one, hence the expectation $\mathbf{E}_{F_{\mathbf{x}}} \left[\log(\sigma^2(\mathbf{x})) \right]$ exists. Thus, the average mutual information between the random variables $\mathbf{x}$ and $\mathbf{y}$ becomes

$$I_{F_{\mathbf{x}}}(\mathbf{x}; \mathbf{y}) = H_{F_{\mathbf{x}}}(\mathbf{y}) - D_0 - \frac{1}{2} \mathbf{E}_{F_{\mathbf{x}}} \left[\log(\sigma^2(\mathbf{x})) \right], \tag{27}$$

where $D_0 = \frac{1}{2} \log(2\pi e \sigma_0^2)$.

We define the mutual information density $i_{F_{\mathbf{x}}}(x)$ and the entropy density $h_{F_{\mathbf{x}}}(x)$ as

$$i_{F_{\mathbf{x}}}(x) \triangleq \int_{-\infty}^{\infty} f_{\mathbf{n}}(y - x, x) \log \frac{f_{\mathbf{n}}(y - x, x)}{f_{\mathbf{y}}(y; F_{\mathbf{x}})} \, dy, \tag{28}$$

$$h_{F_{\mathbf{x}}}(x) \triangleq - \int_{-\infty}^{\infty} f_{\mathbf{n}}(y - x, x) \log f_{\mathbf{y}}(y; F_{\mathbf{x}}) \, dy. \tag{29}$$

Clearly, we can write

$$i_{F_{\mathbf{x}}}(x) = h_{F_{\mathbf{x}}}(x) - \frac{1}{2} \log(\sigma^2(x)) - D_0. \tag{30}$$

The channel capacity is defined as the supremum of the mutual information over the space of the CDFs in $\mathcal{F}_{\mathbf{x}}$ given by

$$C_2 \triangleq \sup_{F_{\mathbf{x}} \text{ in } \mathcal{F}_{\mathbf{x}}} I_{F_{\mathbf{x}}}(\mathbf{x}; \mathbf{y}). \tag{31}$$

IV. Capacity of Fading Channels With Amplitude-Limited Inputs

In this section, we show that the mutual information is a strictly concave, weakly differentiable and continuous function of the input distribution, and by utilizing the Identity Theorem, we prove that the capacity of fading channels with amplitude-limited inputs is achieved by a unique input distribution and this distribution is discrete. We note that the results in the subsections IV-A and IV-B have also been derived in [23], however, we keep a brief description of these results with a highlight of their proofs as our proof approach is different (and to make the paper self-contained).

A. The Mutual Information Is a Continuous Function of the Input Distribution

In order to show the continuity of the term $H_{F_\mathbf{x}}(\mathbf{y}|\mathbf{u})$ in (20), we first show that for any sequence of input distributions, $H_{F_\mathbf{x}^{(n)}}(\mathbf{y}|\mathbf{u} = u)$ is bounded by an integrable function. That is, let us fix a sequence $\{F_\mathbf{x}^{(n)}(x)\}_{n\geq 1}$ in $\mathcal{F}_\mathbf{x}$ such that $F_\mathbf{x}^{(n)}(x) \to F_\mathbf{x}(x)$ (convergence in the Levy metric [18]) for some $F_\mathbf{x} \in \mathcal{F}_\mathbf{x}$. Thus,

$$\lim_{n\to\infty} f_{\mathbf{y}|\mathbf{u}}(y|u; F_\mathbf{x}^{(n)}) = \lim_{n\to\infty} \int_{-A}^{A} f_\mathbf{z}(y - ux)\, dF_\mathbf{x}^{(n)}(x),$$
$$\overset{(a)}{=} \int_{-A}^{A} f_\mathbf{z}(y - ux)\, dF_\mathbf{x}(x),$$
$$= f_{\mathbf{y}|\mathbf{u}}(y|u; F_\mathbf{x}),$$

where (a) follows from the Helly-Bray Theorem [18]. That is,

$$\lim_{n\to\infty} -f_{\mathbf{y}|\mathbf{u}}\left(y|u; F_\mathbf{x}^{(n)}\right) \log\left(f_{\mathbf{y}|\mathbf{u}}(y|u; F_\mathbf{x}^{(n)})\right)$$
$$= -f_{\mathbf{y}|\mathbf{u}}(y|u; F_\mathbf{x}) \log\left(f_{\mathbf{y}|\mathbf{u}}(y|u; F_\mathbf{x})\right)$$

From (7) and (15), we have

$$\left| -f_{\mathbf{y}|\mathbf{u}}(y|u; F_\mathbf{x}^{(n)}) \log\left(f_{\mathbf{y}|\mathbf{u}}(y|u; F_\mathbf{x}^{(n)})\right) \right|$$
$$\leq \Gamma(y, u)\left[-\log(\gamma(y, u)) + 2|\log(k_3)|\right] < \infty,$$

hence, by following similar steps as in [7], with the use of the Dominated Convergence Theorem, we can argue that the conditional entropy and the conditional mutual information are continuous.

B. The Mutual Information Is a Strictly Concave and Weakly Differentiable Function of the Input Distribution

From (20), it is enough to show that the conditional entropy $H_{F_\mathbf{x}}(\mathbf{y}|\mathbf{u})$ is a strictly concave function of the input distribution to conclude the strict concavity of the mutual information. This in turn can be verified using $H_{F_\mathbf{x}}(\mathbf{y}|\mathbf{u} = u)$ from (18).

We define a new random variable $\mathbf{y}' = \frac{\mathbf{y}}{u}$ for a fixed $u > 0$. Thus, an equivalent channel model for a given channel gain becomes

$$\mathbf{y}' = \mathbf{x} + \frac{\mathbf{z}}{u}. \tag{32}$$

We assume that $F_\mathbf{u}(0) \neq 1$, i.e., the measure of the set of the nonzero values of the channel coefficients is not zero. Thus, the equivalent model in (32) is the same as the scalar Gaussian channel model studied by Smith [2], which leads to the strict concavity of the conditional entropy for any $u \in [\epsilon_0, U]$, i.e., $H_{F_\mathbf{x}}(\mathbf{y}|\mathbf{u} = u)$ is a strictly concave function. Using this result, we can conclude the strict concavity of the conditional output entropy since the integration of strictly concave functions is strictly concave [24, p. 79].

Lemma 2: The mutual information $I_{F_\mathbf{x}}(\mathbf{x}; \mathbf{y}|\mathbf{u})$ is a weakly differentiable function at every point in $\mathcal{F}_\mathbf{x}$, and its weak derivative denoted by $I'_{F_1, F_2}(\mathbf{x}; \mathbf{y}|\mathbf{u})$ is given by

$$I'_{F_1, F_2}(\mathbf{x}; \mathbf{y}|\mathbf{u}) = \int_{-A}^{A} i_{F_1}(x|\mathbf{u})\, dF_2(x) - I_{F_1}(\mathbf{x}; \mathbf{y}|\mathbf{u}). \tag{33}$$

The proof of the lemma follows from similar line of arguments as in [7], and it is omitted.

Theorem 1: The capacity of a fading channel with an amplitude-limited input and a fading coefficient with finite support is achieved by a unique CDF F^* in $\mathcal{F}_\mathbf{x}$, and

$$\int_{-A}^{A} i_{F^*}(x|\mathbf{u})\, dF_\mathbf{x}(x) - I_{F^*}(\mathbf{x}; \mathbf{y}|\mathbf{u}) \leq 0, \quad \forall F^* \in \mathcal{F}_\mathbf{x}. \tag{34}$$

Proof: The space $\mathcal{F}_\mathbf{x}$ is convex and compact in the Levy metric topology (topology of weak convergence) [17]. We showed that the function $I: \mathcal{F}_\mathbf{x} \to \mathbb{R}$ is strictly concave, continuous and weakly differentiable in $\mathcal{F}_\mathbf{x}$. Thus, there is a unique optimal input distribution that maximizes the capacity. The weak derivative of the mutual information (as shown in Lemma 2) is

$$I'_{F_1, F_2}(\mathbf{x}; \mathbf{y}|\mathbf{u}) = \int_{-A}^{A} i_{F_1}(x|\mathbf{u})\, dF_2(x) - I_{F_1}(\mathbf{x}; \mathbf{y}|\mathbf{u}). \tag{35}$$

As a result, F^* is the optimal input distribution if and only if

$$\int_{-A}^{A} i_{F^*}(x|\mathbf{u})\, dF_\mathbf{x}(x) - I_{F^*}(\mathbf{x}; \mathbf{y}|\mathbf{u}) \leq 0, \quad \forall F^* \in \mathcal{F}_\mathbf{x}. \tag{36}$$

which concludes the proof. □

Corollary 1: The following conditions are necessary and sufficient for an optimal input CDF

$$i_{F^*}(x|\mathbf{u}) \leq I_{F^*}(\mathbf{x}; \mathbf{y}|\mathbf{u}), \quad \forall x \in [-A, A], \tag{37}$$
$$i_{F^*}(x|\mathbf{u}) = I_{F^*}(\mathbf{x}; \mathbf{y}|\mathbf{u}), \quad \forall x \in E^*, \tag{38}$$

where E^* is the set of points of increase of the CDF F^*.

Proof: The proof follows from similar lines of reasoning as in [2]. Assume that F^* is optimal but the first inequality is not valid. Then, there exists an $x_1 \in [-A, A]$ such that $i_{F^*}(x_1|\mathbf{u}) > I_{F^*}(\mathbf{x}; \mathbf{y}|\mathbf{u})$. Let $F_{\mathbf{x}_1}(x) \triangleq \mathbf{U}(x - x_1)$, where $\mathbf{U}(\cdot)$ is the unit step function. Then,

$$\int_{-A}^{A} i_{F^*}(x|\mathbf{u})\, dF_{\mathbf{x}_1}(x) = i_{F^*}(x_1|\mathbf{u}) > I_{F^*}(\mathbf{x}; \mathbf{y}|\mathbf{u}).$$

This contradicts the result of Theorem 1. Thus, the first assertion is valid, i.e.,

$$i_{F^*}(x|\mathbf{u}) \leq I_{F^*}(\mathbf{x}; \mathbf{y}|\mathbf{u}), \quad \forall x \in [-A, A]. \tag{39}$$

Define E' as a subset of E^* with positive measure, i.e.,

$$\int_{E'} dF^*(x) = \delta > 0, \tag{40}$$

for which

$$i_{F^*}(x|\mathbf{u}) < I_{F^*}(\mathbf{x}; \mathbf{y}|\mathbf{u}), \quad \forall x \in E'. \tag{41}$$

Since

$$\int_{E^* - E'} \mathrm{d}F^*(x) = 1 - \delta, \tag{42}$$

and

$$i_{F^*}(x|\mathbf{u}) = I_{F^*}(\mathbf{x}; \mathbf{y}|\mathbf{u}) \quad \text{on } E^* - E',$$

we have

$$
\begin{aligned}
I_{F^*}&(\mathbf{x}; \mathbf{y}|\mathbf{u}) \\
&= \int_{E^*} i_{F^*}(x|\mathbf{u}) \, \mathrm{d}F^*(x) \\
&= \int_{E'} i_{F^*}(x|\mathbf{u}) \, \mathrm{d}F^*(x) + \int_{E^*-E'} i_{F^*}(x|\mathbf{u}) \, \mathrm{d}F^*(x) \\
&\overset{(a)}{<} \delta I_{F^*}(\mathbf{x}; \mathbf{y}|\mathbf{u}) + (1-\delta)I_{F^*}(\mathbf{x}; \mathbf{y}|\mathbf{u}) \\
&= I_{F^*}(\mathbf{x}; \mathbf{y}|\mathbf{u})
\end{aligned}
$$

where (a) follows from (40), (41), and (42). Thus, we have a contradiction and the second statement also holds. $\square$

C. Discreteness of the Optimal Distribution

In this subsection, we prove that the optimal distribution that maximizes the mutual information in (22) is discrete with a finite number of mass points. In a nutshell, we show that the extension of the conditional entropy density to the complex plane is well-defined and this extension is analytic. Then, we assume that the set of points of increase of the input CDF E^* contains an infinite number of elements. Finally, we demonstrate via Bolzano-Weierstrass and Identity Theorems that this assumption leads to a contradiction, hence proving that the optimal input distribution has a finite number of mass points.

The conditional entropy density corresponding to the optimal input distribution F^* is given by

$$h_{F^*}(x|\mathbf{u}) = \int_0^U h_{F^*}(x|\mathbf{u} = u) \, \mathrm{d}F_{\mathbf{u}}(u). \tag{43}$$

We first extend $h_{F^*}(x|\mathbf{u} = u)$ to the complex plane. For any $z = \eta + j\zeta \in \mathbb{C}$ and $u \in [0, U]$,

$$
\begin{aligned}
|h&_{F_{\mathbf{x}}}(z|\mathbf{u} = u)| \\
&\leq \int_{-\infty}^{\infty} |f_{\mathbf{z}}(y - uz)| |\log f_{\mathbf{y}|\mathbf{u}}(y|u; F_{\mathbf{x}})| \, \mathrm{d}y, \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi \sigma_{\mathbf{z}}^2}} \left|\exp\left(\frac{-(y - uz)^2}{2\sigma_{\mathbf{z}}^2}\right)\right| \\
&\qquad\qquad |\log f_{\mathbf{y}|\mathbf{u}}(y|u; F_{\mathbf{x}})| \, \mathrm{d}y, \\
&\leq \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi \sigma_{\mathbf{z}}^2}} \left|\exp\left(\frac{-(y - u\eta - ju\zeta)^2}{2\sigma_{\mathbf{z}}^2}\right)\right| \\
&\qquad\qquad [-\log(\gamma(y, u)) + 2|\log(k_3)|] \, \mathrm{d}y,
\end{aligned}
$$

$$
\begin{aligned}
&\leq \frac{1}{\sqrt{2\pi \sigma_{\mathbf{z}}^2}} \exp\left(\frac{u^2 \zeta^2}{2\sigma_{\mathbf{z}}^2}\right) \int_{-\infty}^{\infty} \exp\left(\frac{-(y - u\eta)^2}{2\sigma_{\mathbf{z}}^2}\right) \\
&\qquad\qquad \left[-\log(k_1) + k_2(|y| + uA)^2 + 2|\log(k_3)|\right] \mathrm{d}y, \\
&= \exp\left(\frac{u^2 \zeta^2}{2\sigma_{\mathbf{z}}^2}\right) \int_{-\infty}^{\infty} f_{\mathbf{z}}(y - u\eta) \\
&\qquad\qquad \left[-\log(k_1) + k_2(|y| + uA)^2 + 2|\log(k_3)|\right] \mathrm{d}y, \\
&< \infty.
\end{aligned}
$$

Hence $|h(z|\mathbf{u} = u)|$ is finite for any $|z| < \infty$. Thus, the extension of $h_{F_{\mathbf{x}}}(z|\mathbf{u} = u)$ is well defined.

Since there exists a finite $B > 0$ such that $\forall u \in [0, U]$, we have $|h_{F_{\mathbf{x}}}(z|\mathbf{u} = u)| \leq B$, i.e.,

$$
\begin{aligned}
|h_{F^*}(z|\mathbf{u})| &= \left|\int_0^U h_{F_{\mathbf{x}}}(z|\mathbf{u} = u) \, \mathrm{d}F_{\mathbf{u}}(u)\right|, \\
&\leq \int_0^U \left|h_{F_{\mathbf{x}}}(z|\mathbf{u} = u)\right| \mathrm{d}F_{\mathbf{u}}(u), \\
&\leq B \int_0^U \mathrm{d}F_{\mathbf{u}}(u) = B < \infty,
\end{aligned}
$$

namely, $h_{F^*}(z|\mathbf{u})$ has an extension to the complex plane as well.

Since the Gaussian PDF extended to the complex plane is an analytic function, using the Cauchy Integral Theorem [25], we have

$$\oint_{\omega} f_{\mathbf{z}}(z) \, \mathrm{d}z = 0, \tag{44}$$

where ω is any simple closed contour. To show the analyticity of the conditional entropy density, we use Morera's Theorem, i.e., by showing that the integration of the conditional entropy over any simple closed contour is zero, we conclude that the function is analytic. That is,

$$
\begin{aligned}
\oint_{\omega} &h_{F_{\mathbf{x}}}(z|\mathbf{u}) \, \mathrm{d}z \\
&= -\oint_{\omega} \int_0^U \int_{-\infty}^{\infty} f_{\mathbf{z}}(y - uz) \log(f_{\mathbf{y}|\mathbf{u}}(y|u; F_{\mathbf{x}})) \, \mathrm{d}y \, \mathrm{d}F_{\mathbf{u}}(u) \, \mathrm{d}z, \\
&\overset{(a)}{=} -\int_0^U \int_{-\infty}^{\infty} \log(f_{\mathbf{y}|\mathbf{u}}(y|u; F_{\mathbf{x}})) \oint_{\omega} f_{\mathbf{z}}(y - uz) \, \mathrm{d}z \, \mathrm{d}y \, \mathrm{d}F_{\mathbf{u}}(u), \\
&= 0,
\end{aligned}
$$

where the order of integrals in (a) is changed by invoking the Fubini's Theorem that requires the finiteness of $\oint_{\omega} |h_{F_{\mathbf{x}}}(z|\mathbf{u})| \, \mathrm{d}z$. This can be justified as follows. Define M_{ω} as

$$M_{\omega} = \max_{z \in \omega} \left|h_{F_{\mathbf{x}}}(z|\mathbf{u})\right|. \tag{45}$$

M_{ω} exists since the conditional entropy $|h_{F_{\mathbf{x}}}(z|\mathbf{u})|$ is bounded, continuous (the continuity is shown in Appendix B) in z, and the contour ω is closed. Hence,

$$
\begin{aligned}
\left|\oint_{\omega} h_{F_{\mathbf{x}}}(z|\mathbf{u}) \, \mathrm{d}z\right| &\leq \oint_{\omega} \left|h_{F_{\mathbf{x}}}(z|\mathbf{u})\right| \mathrm{d}z, \\
&\leq M_{\omega} l_{\omega} < \infty,
\end{aligned}
$$

where l_ω is the length of ω, which is finite as ω is a closed contour.

It is now easy to argue that the extension of the conditional mutual information density $i_{F^*}(z|\mathbf{u})$ to the complex plane is well defined (since its difference with the entropy density is a constant), and it is analytic.

We prove the discreteness of the capacity-achieving distribution by a contradiction. We first assume that the set of points of increase E^* has an infinite cardinality. From the optimality condition in (38) we have[3]

$$\int_{\epsilon_0}^{U} (i_{F^*}(x|\mathbf{u}=u) - I_{F^*}(\mathbf{x}; \mathbf{y}|\mathbf{u}=u))\, dF_{\mathbf{u}}(u) = 0, \quad \forall x \in E^*.$$

The set E^* is bounded, and hence it has a limit point (Bolzano-Weierstrass Theorem). The conditional mutual information density $i_{F^*}(z|\mathbf{u})$ is analytic on the entire complex plane. That is, we can invoke the Identity Theorem to argue that the optimality condition implies

$$\int_{\epsilon_0}^{U} (i_{F^*}(x|\mathbf{u}=u) - I_{F^*}(\mathbf{x}; \mathbf{y}|\mathbf{u}=u))\, dF_{\mathbf{u}}(u) = 0, \quad \forall x \in \mathbb{R},$$

and hence

$$\int_{\epsilon_0}^{U} \left(-\int_{-\infty}^{\infty} f_{\mathbf{z}}(y - ux) \log f_{\mathbf{y}|\mathbf{u}}(y|u)\, dy - D_{\mathbf{z}} \right.$$
$$\left. - I_{F^*}(\mathbf{x}; \mathbf{y}|\mathbf{u}=u) \right) dF_{\mathbf{u}}(u) = 0, \quad \forall x \in \mathbb{R}. \quad (46)$$

We now extend the approach in [3] and [6] to the case of fading channels. For a fixed $\epsilon_0 < u \le U$, let us define $L(u) \triangleq I_{F^*}(\mathbf{x}; \mathbf{y}|\mathbf{u}=u) + \frac{1}{2}\log(2\pi\sigma_{\mathbf{z}}^2)$, $\rho(y, u) \triangleq \log\left(f_{\mathbf{y}|\mathbf{u}}(y|u)\right) + L(u)$, and $x' \triangleq ux$. Also define the sets

$$\Omega_u^+ = \{y : \rho(y, u) \ge 0\}, \quad \text{and} \quad \Omega_u^- = \{y : \rho(y, u) < 0\}.$$

We can then write,

$$\int_{\epsilon_0}^{U} \left[\int_{\Omega_u^+} f_{\mathbf{z}}(y - ux)\rho(y, u)\, dy \right. \quad (47)$$
$$\left. + \int_{\Omega_u^-} f_{\mathbf{z}}(y - ux)\rho(y, u)\, dy \right] dF_{\mathbf{u}}(u) = 0. \quad (48)$$

For any $y \in \Omega_u^+$, we have $\rho(y, u) \le \log(\Gamma(y, u)) + L(u) \le \log(k_3) + L(u)$. Defining

$$l \triangleq \max_{u \in [\epsilon_0, U]} uA + \sqrt{\frac{\log(k_3) + L(u)}{k_4 \log(e)}} < \infty, \quad (49)$$

for any $y \in \Omega_u^+$, we have $y \in [-l, l]$. The last step in (49) follows since $L(u) < \infty$ for any $u \in [\epsilon_0, U]$ (Lemma 1). In other words, $\Omega_u^+ \subseteq [-l, l]$. Therefore, for any $u \in [\epsilon_0, U]$

$$\int_{\Omega_u^+} f_{\mathbf{z}}(y - ux)\rho(y, u)\, dy$$
$$\le \int_{\Omega_u^+} f_{\mathbf{z}}(y - ux)\left(\log(k_3) + \max_{u' \in [\epsilon_0, U]} L(u')\right) dy, \quad (50)$$
$$\le \left(\log(k_3) + \max_{u' \in [\epsilon_0, U]} L(u')\right) \int_{-l}^{l} \Gamma(y - x', u)\, dy, \quad (51)$$

[3] So far, we have written the integrals over the channel fading coefficient from 0 to U, however, at this point we write the integration explicitly over the interval where the coefficient belongs to. The exclusion of some neighborhood of 0 is needed in the technical arguments about the discreteness of the optimal input distribution.

which can be made arbitrarily small by choosing large values for x'.

On the other hand, for any given $u > \epsilon_0$ and $x' > l + A$,

$$\int_{\Omega_u^-} f_{\mathbf{z}}(y - ux)\rho(y, u)\, dy$$
$$\overset{(a)}{\le} \int_{l}^{\infty} f_{\mathbf{z}}(y - x')\rho(y, u)\, dy,$$
$$\le \int_{x'-A}^{x'+A} q(y - x')\left[\log(\Gamma(y, u)) + L(u)\right] dy$$
$$+ \int_{x'+A}^{\infty} q(y - x')\left[\log(\Gamma(y, u)) + L(u)\right] dy, \quad (52)$$
$$\overset{(b)}{<} \int_{x'-A}^{x'+A} q(A)\left[\log(\Gamma(x' - A, u)) + L(u)\right] dy,$$
$$< 2Aq(A)\left[\log(\Gamma(x' - A, u)) + L(u)\right] \overset{(c)}{<} 0, \quad (53)$$

where (a) follows since $[l, \infty) \subset \Omega_u^-$ and the definition of l guarantees that both $\rho(y, u)$ and $\log(\Gamma(y, u)) + L(u)$ are negative on $[l, \infty)$. The inequality (b) follows since $q(y - x') \le f_{\mathbf{z}}(y - x')$ and it is nonzero on its support by the definition in (7). That is, the second integral in (52) is strictly negative. Also notice that $q(A) \le q(y - x')$ over the region of integration for the first term in (52). Finally, (c) follows since $q(A) > 0$, and the function $\log(\Gamma(y, u)) + L(u)$ is negative and monotonically decreasing in y for $y > l$, i.e., $\log(\Gamma(y, u)) + L(u) \le \log(\Gamma(x' - A, u)) + L(u)$.

Therefore, from (51) and (53), one can argue that $\forall u \in [\epsilon_0, U]$, there exists an $x \in \mathbb{R}$ (same value of x for any $u \in [\epsilon_0, U]$) such that the integration in (46) is strictly less than zero, which is a contradiction, hence we conclude that the set E^* does not have an infinite number of mass points concluding the proof.

Remark: The channel capacity in (22) can be computed by finding the optimal input distribution and then evaluating the mutual information corresponding to this distribution. As we have shown, the capacity optimization problem is convex since the space of input distribution functions is convex and the mutual information is strictly concave. We also have shown that the capacity is achieved by a discrete distribution with a finite number of mass points. Thus, the problem of finding the optimal input distribution boils down to a finite-dimensional optimization problem that aims to find the location of mass points and the associated probabilities corresponding to this distribution. To do this, an efficient numerical optimization algorithm can be developed which iterates over the number of mass points and its associated probabilities until the optimality conditions are satisfied and hence the optimal input distribution is found.

V. CAPACITY OF SIGNAL-DEPENDENT NOISE CHANNELS WITH AMPLITUDE-LIMITED INPUTS

In this section, we present our results on the capacity of signal-dependent Gaussian channels with amplitude-limited inputs. We basically show that there is an optimal distribution that maximizes the mutual information, and this distribution is discrete under certain technical conditions. To accomplish

this, we first show that the mutual information is a concave, continuous and weakly differentiable function of the input distribution, i.e., the capacity optimization problem is a convex optimization problem for which we can derive conditions on the optimality of the input distribution. We investigate the properties of the optimal distribution using techniques from analysis and show that under some technical conditions on the function that relates the noise variance to the input signal the capacity-achieving distribution is discrete with a finite number of mass points.

We note that even though the noise is assumed to be Gaussian (with parameters depending on the input) throughout this section, the techniques developed here can potentially be generalized to other types of noise similar to the generalization in [3] (extending the work in [2]) for the case with no input dependency. We also note that, as stated previously, the mutual information in (27) is different from the one in the original work of Smith [17] due the existence of the expectation term $\frac{1}{2}\mathbf{E}_{F_\mathbf{x}}\left[\log\left(\sigma^2(\mathbf{x})\right)\right]$ and difference in the computation of the output entropy $H_{F_\mathbf{x}}(\mathbf{y})$.

A. The Mutual Information Is a Continuous Function of the Input Distribution

Let us fix a sequence $\{F_\mathbf{x}^{(n)}(x)\}_{n\geq 1}$ in $\mathcal{F}_\mathbf{x}$ such that $F_\mathbf{x}^{(n)}(x) \to F_\mathbf{x}$ for some $F_\mathbf{x} \in \mathcal{F}_\mathbf{x}$ at points of continuity of $F_\mathbf{x}$. Since the noise variance function $\sigma^2(\mathbf{x})$ is continuous by assumption, $f_\mathbf{n}(y - x, x)$ is bounded and continuous in x, y, we have

$$\lim_{n\to\infty} f_\mathbf{y}(y; F_\mathbf{x}^{(n)}) = \lim_{n\to\infty} \int_{-A}^{A} f_\mathbf{n}(y - x, x)\,\mathrm{d}F_\mathbf{x}^{(n)}(x),$$
$$\overset{(a)}{=} \int_{-A}^{A} f_\mathbf{n}(y - x, x)\,\mathrm{d}F_\mathbf{x}(x),$$
$$= f_\mathbf{y}(y; F_\mathbf{x}),$$

where (a) follows by the Helly-Bray Theorem [18]. Then, we can write

$$\lim_{n\to\infty} f_\mathbf{y}(y; F_\mathbf{x}^{(n)}) \log\left(f_\mathbf{y}(y; F_\mathbf{x}^{(n)})\right)$$
$$= f_\mathbf{y}(y; F_\mathbf{x}) \log\left(f_\mathbf{y}(y; F_\mathbf{x})\right).$$

Thus, from (11), and by applying the Dominated Convergence Theorem, we conclude that $H_{F_\mathbf{x}}(\mathbf{y})$ is a continuous function of the input distribution. The continuity of $\mathbf{E}_{F_\mathbf{x}}\left[\log\left(\sigma^2(\mathbf{x})\right)\right]$ in the input distribution function $F_\mathbf{x}(x)$ follows from the Helly-Bray Theorem as well since $\sigma^2(x)$ is bounded.

B. Concavity of the Mutual Information

For the mutual information term in (27), since $1 \leq \sigma^2(x) < \infty$, we have $0 \leq \log(\sigma^2(x)) \leq M_0 < \infty$. Thus,

$$\mathbf{E}_{F_\mathbf{x}}\left[\log\left(\sigma^2(\mathbf{x})\right)\right] = \int_{-A}^{A} \log(\sigma^2(x))\,\mathrm{d}F_\mathbf{x}(x),$$
$$\leq M_0 \int_{-A}^{A} \mathrm{d}F_\mathbf{x}(x),$$
$$= M_0 < \infty.$$

The concavity of the output entropy $H_{F_\mathbf{x}}(\mathbf{y})$ can be shown using a similar line of arguments as in [2] by considering an input distribution of the form $F_\theta = \theta F_1 + (1-\theta)F_2$, where θ is a scalar in [0, 1], and F_1, F_2 are two arbitrary CDFs satisfying the input constraints, and by invoking Ash's Lemma given in (14). We omit the details as they are also available in [13].

C. The Mutual Information Function $I_{F_\mathbf{x}}(\mathbf{x}; \mathbf{y})$ Is Weakly Differentiable

For arbitrary distribution functions F_1, F_2, and $\theta \in [0, 1]$, defining $F_\theta = (1 - \theta)F_1 + \theta F_2$, we have

$$I(F_\theta) = H(F_\theta) - D_0 - \frac{1}{2}\mathbf{E}_{F_\theta}\left[\log(\sigma^2(\mathbf{x}))\right].$$

Defining $J(\theta, F_1, F_2) = \frac{I(F_\theta) - I(F_1)}{\theta}$, we can write

$$J(\theta, F_1, F_2)$$
$$= \frac{H(F_\theta) - H(F_1)}{\theta}$$
$$\quad - \frac{1}{2}\frac{\mathbf{E}_{F_\theta}\left[\log(\sigma^2(\mathbf{x}))\right] - \mathbf{E}_{F_1}\left[\log(\sigma^2(\mathbf{x}))\right]}{\theta},$$
$$= \frac{H(F_\theta) - H(F_1)}{\theta}$$
$$\quad - \frac{1}{2}\left(\mathbf{E}_{F_2}\left[(\log(\sigma^2(\mathbf{x}))\right] - \mathbf{E}_{F_1}\left[\log(\sigma^2(\mathbf{x}))\right]\right).$$

We then have

$$I'_{F_1}(F_2) = \lim_{\theta\to 0} J(\theta, F_1, F_2),$$
$$= \lim_{\theta\to 0}\left[\frac{H(F_\theta) - H(F_1)}{\theta}\right]$$
$$\quad - \frac{1}{2}\left(\mathbf{E}_{F_2}\left[\log(\sigma^2(\mathbf{x}))\right] - \mathbf{E}_{F_1}\left[\log(\sigma^2(\mathbf{x}))\right]\right).$$

Following similar arguments as in Smith [2], we can show the weak differentiability of the mutual information.

Theorem 2: The capacity of an additive signal-dependent Gaussian noise channel is achieved by a random variable $\mathbf{x}^*$ with CDF $F^* \in \mathcal{F}_\mathbf{x}$, i.e., $C_2 = I_{F^*}(\mathbf{x}; \mathbf{y})$ for some $F^* \in \mathcal{F}_\mathbf{x}$. A necessary and sufficient condition for F^* to achieve the channel capacity is

$$\int_{-A}^{A} i_{F^*}(x)\,\mathrm{d}F_\mathbf{x}(x) - I_{F^*}(\mathbf{x}; \mathbf{y}) \leq 0. \tag{54}$$

Proof: The space $\mathcal{F}_\mathbf{x}$ is convex and compact in the Levy metric topology [17]. We showed that the function $I : \mathcal{F}_\mathbf{x} \to \mathbb{R}$ is concave, continuous and weakly differentiable in $\mathcal{F}_\mathbf{x}$. The weak derivative of the mutual information is

$$I'_{F_1}(F_2) = \int_{-A}^{A} i(x; F_1)\,\mathrm{d}F_2 - I(F_1),$$

Hence, a distribution $F^* \in \mathcal{F}_\mathbf{x}$ is optimal if and only if

$$\int_{-A}^{A} i(x; F_1)\,\mathrm{d}F_2 - I(F_1) \leq 0. \qquad \square$$

As a result of this theorem and similar to what is done in [2] and Section IV, we can show that the following conditions are necessary and sufficient for F^* to be optimal:

$$i_{F^*}(x) - I_{F^*}(\mathbf{x}; \mathbf{y}) \leq 0, \quad \forall x \in [-A, A], \quad (55)$$

$$i_{F^*}(x) - I_{F^*}(\mathbf{x}; \mathbf{y}) = 0, \quad \forall x \in E^*, \quad (56)$$

where E^* is the set of points of increase of F^*.

D. The Capacity-Achieving Distribution Is Discrete

In this subsection, we show that the capacity-achieving distribution of the signal-dependent Gaussian noise channels is discrete under certain assumptions on the noise variance function. This is shown by assuming that the set of points of increase of the capacity-achieving input distribution has an infinite number of mass points, and then by proving via a series of arguments that this assumption leads to a contradiction. Before going into detail, we state the specific technical conditions as follows:

- The noise variance $\sigma^2(x)$ and its logarithm $\log(\sigma^2(x))$ can be extended to an open connected set in the complex plane containing the real line $\mathbb{R}$ (possibly excluding a finite number of branch points denoted by the set $\mathcal{S}$. We assume that the set of branch points is a finite set of points which guarantees that the extension of the $\log(\sigma^2(x))$ will be defined over a connected set in the complex plane and the Identity Theorem applies).
- The extension of the noise variance denoted by $\sigma^2(z)$ is analytic on an open connected set $\mathcal{D} \in \mathbb{C}$ including the real line. From this assumption, it is easy to show that the function $\log(\sigma^2(z))$ is analytic as well.[4]

The aforementioned technical conditions on the noise variance are imposed to apply the Identity Theorem to prove the discreteness of the optimal input distribution. We note that there are examples for which these technical conditions are not satisfied, nevertheless for many important cases used in the existing literature (e.g., those used in magnetic recording and optical communications channels), they are satisfied as will be illustrated later in Section VI. We also note that these technical conditions are different from those in [13]. In the numerical results section, we will provide an example of a magnetic recording channel for which the adopted noise model is included in our set-up while it does not fall within the framework of [13].

First, we extend the function $i_{F_{\mathbf{x}}}(x)$ to the complex plane as detailed in Appendix A where we show that the extension of the mutual information density to the complex plane is well defined. The discreteness of the optimal distribution can be established through the following contradiction arguments. Assuming that the set E^* has an infinite number of points, since it is a bounded set, any infinite sequence in E^* has a limit point (Bolzano-Weierstrass Theorem). We show in Appendix B that the function $i_{F^*}(z)$ is analytic on some open connected set $\mathcal{D}$ in the complex plane $\mathbb{C}$ that includes the real line $\mathbb{R}$ except the set of branch points $\mathcal{S}$. Using the Identity

Theorem, we conclude that $i_{F^*}(z) = I_{F^*}(\mathbf{x}; \mathbf{y})$ for all $\mathcal{D} \setminus \mathcal{S}$. We can then write

$$h_{F^*}(z) = I_{F^*}(\mathbf{x}; \mathbf{y}) + \frac{1}{2}\log(\sigma^2(z)) + D_0 = 0, \quad \forall z \in \mathcal{D} \setminus \mathcal{S},$$

which implies that $h_{F^*}(x) = I_{F^*}(\mathbf{x}; \mathbf{y}) + \frac{1}{2}\log(\sigma^2(x)) + D_0$ on the entire real line except the set $\mathcal{S}$. We note that the branch points of the function $\log(\sigma^2(z))$ are only located in the interval between $[-A, A]$ because outside this region the extension of the noise variance is arbitrary and can be defined in such a way that it does not have any branch points (constant for example).

Clearly, $\sigma^2(x) \geq 1$, and hence $\log(\sigma^2(x)) \geq 0$. Let us define $L \triangleq I_{F^*}(\mathbf{x}; \mathbf{y}) + D_0 + \frac{1}{2}\log(\hat{\sigma}^2)$ and $\rho(y) \triangleq \log(f_{\mathbf{y}}(y; F^*)) + L$, where $\hat{\sigma}^2$ is a constant.[5] For sufficiently large values of x, we have

$$h_{F^*}(x) - I_{F^*}(\mathbf{x}; \mathbf{y}) - \frac{1}{2}\log(\hat{\sigma}^2) - D_0 = 0,$$

namely,

$$\int_{-\infty}^{\infty} f_{\mathbf{n}}(y - x, x)\left[\log(f_{\mathbf{y}}(y; F^*)) + I_{F^*}(\mathbf{x}; \mathbf{y})\right.$$
$$\left. + \frac{1}{2}\log(\hat{\sigma}^2) + D_0\right]\mathrm{d}y = 0. \quad (57)$$

We define

$$f_{\mathbf{n}}^*(y - x) \triangleq \frac{1}{\sqrt{2\pi\left(\sigma_0^2 + \min_{|x'| \leq A} \sigma_1^2(x')\right)}}$$
$$\times \exp\left(-\frac{(y - x)^2}{2\left(\sigma_0^2 + \max_{|x'| \leq A} \sigma_1^2(x')\right)}\right), \quad (58)$$

and

$$\Omega^+ \triangleq \{y : \rho(y) \geq 0\}, \quad \text{and } \Omega^- \triangleq \{y : \rho(y) < 0\}.$$

Then,

$$\int_{\Omega^+} f_{\mathbf{n}}(y - x, x)\rho(y)\,\mathrm{d}y + \int_{\Omega^-} f_{\mathbf{n}}(y - x, x)\rho(y)\,\mathrm{d}y = 0. \quad (59)$$

By (11), we get $\rho(y) \leq \log(\Gamma'(y)) + L \leq \log(k_3') + L$, for any $y \in \mathbb{R}$, and $x > A$. Hence, $k_3' > 2^{-L}$. Choose a constant l such that $l > A + \sqrt{\frac{\log(k_3') + L}{k_4' \log(e)}}$. Using (11), one has $\Omega^+ \subseteq [-l, l]$. Therefore,

$$\int_{\Omega^+} f_{\mathbf{n}}(y - x, x)\rho(y)\,\mathrm{d}y$$
$$\leq \int_{\Omega^+} f_{\mathbf{n}}^*(y - x)\left(\log(k_3') + L\right)\mathrm{d}y,$$
$$\leq \left(\log(k_3') + L\right)\int_{-l}^{l} f_{\mathbf{n}}^*(y - x)\,\mathrm{d}y, \quad (60)$$

---

[4] The extension of the noise variance function $log(\sigma^2(x))$ is defined over an open connected set on the complex plane except branch cuts that can be defined such that they do not include the entire real line.

[5] We choose the constant $\sigma^2(x) = \hat{\sigma}^2$ to be an arbitrary extension of the noise variance function for sufficiently large values of x. This can be done in such a way that differentiability and boundedness of $\sigma^2(x)$ are still guaranteed.

which can be made arbitrarily small by choosing large values for x.

On the other hand, for $x > A + l$, we have

$$\int_{\Omega^-} f_{\mathbf{n}}(y - x, x)\rho(y)\,dy$$

$$\overset{(a)}{\leq} \int_l^\infty f_{\mathbf{n}}(y - x, x)\rho(y)\,dy,$$

$$\leq \int_l^\infty f_{\mathbf{n}}(y - x, x)\left[\log(\Gamma'(y)) + L\right]dy,$$

$$\leq \int_{x-A}^{x+A} q'(y - x)\left[\log(\Gamma'(x - A)) + L\right]dy$$

$$+ \int_{x+A}^\infty q'(y - x)\left[\log(\Gamma'(x - A)) + L\right]dy \quad (61)$$

$$\overset{(b)}{<} 2Aq'(A)\left[\log(\Gamma'(x - A)) + L\right] \overset{(c)}{<} 0 \quad (62)$$

where (a) follows since $[l, \infty) \subset \Omega^-$, and (b) follows since the function $q'(y - x) \leq f_{\mathbf{n}}(y - x, x)$ and it is nonzero on its support by definition in (10). Namely, the second integral in (61) is strictly negative, i.e.,

$$\int_{x+A}^\infty q'(y - x)\left[\log(\Gamma'(x - A)) + L\right]dy < 0. \quad (63)$$

Also note that $q'(A) \leq q'(y-x)$ over the region of integration for the first term in (61). Finally, (c) follows since $q'(A) > 0$ and the function $\log(\Gamma'(x - A)) + L$ is negative and monotonically decreasing in y for $y > l$. Therefore, we establish that (57) does not hold for very large (but finite) values of x and hence there is a contradiction, i.e., the set E^* cannot have an infinite number of points, completing the proof.

VI. Numerical Examples

In this section, we present some numerical examples to illustrate our findings. For the case of fading channels, we exemplify our results by finding the optimal input distribution for a truncated Rayleigh channel, and we compare the capacity of peak and average power constrained cases. In the case of signal-dependent Gaussian noise channels, we compare our results with the results in [16] and [26].

A. Fading Channels With Amplitude-Limited Inputs

We consider a fading channel for which the channel coefficient $\mathbf{u}$ follows a truncated Rayleigh distribution, i.e., the PDF of the channel coefficient is given by

$$f_{\mathbf{u}}(u) = \frac{4u}{1 - \exp(-32)}\exp(-2u^2), \quad u \in [0, 4]. \quad (64)$$

We take a noise variance of 1.5, and an amplitude constraint of $A = 3$. We compute the capacity-achieving distribution by following an iterative algorithm similar to the one in [2] that starts by assuming that the input distribution has only two points, and increases the number of mass points until the optimality conditions are satisfied. Fig. 1 shows the resulting optimal input distribution for our example.

We also compare the capacity of the truncated Rayleigh fading channel, with the same fading distribution, for two different input constraints: peak-power constrained inputs and
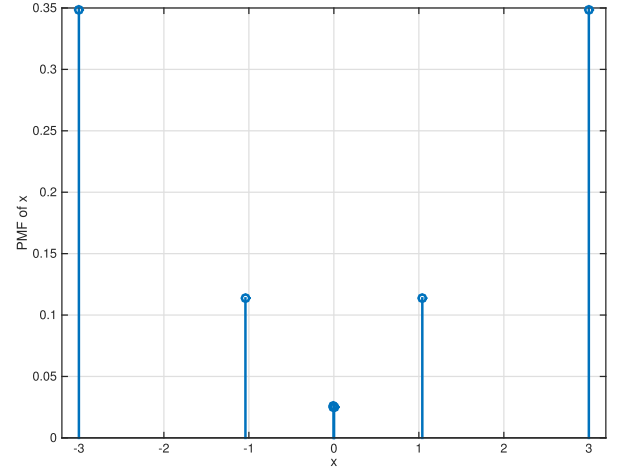

Fig. 1. Probability mass function (PMF) of the optimal input for the fading channel example for $A = 3$.

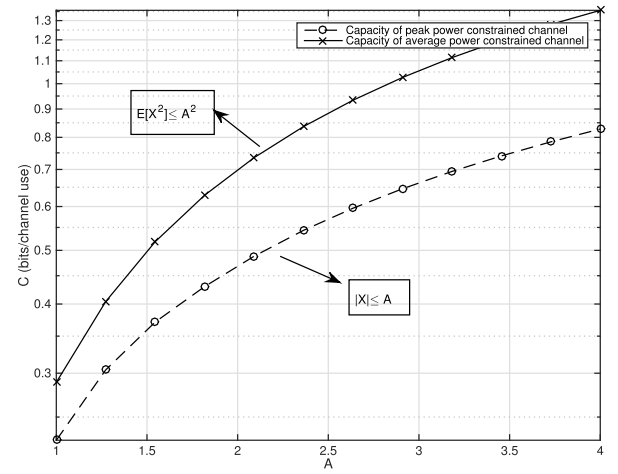
Fig. 2. The capacity of the Rayleigh fading channel versus the amplitude constraint A.

average power constrained inputs. Both capacities are plotted in Fig. 2, which shows that constraining the peak power reduces the channel capacity compared to the same constraint on the average power, as expected.

B. Additive Channels With Signal-Dependent Noise

We now present two specific examples of the capacity of some Gaussian channels with signal-dependent noise under peak power constrained inputs. The first model we consider is the optical communication channel based on intensity modulation, which has been studied in detail in [16]. The received signal $\mathbf{y}$ is given by,

$$\mathbf{y} = \mathbf{x} + \sqrt{\mathbf{x}}\mathbf{z}_1 + \mathbf{z}_0, \quad (65)$$

where $\mathbf{x} \geq 0$ denotes the channel input $\mathbf{z}_0 \sim \mathcal{N}(0, \sigma^2)$ is a zero-mean, variance σ^2 Gaussian random variable describing the input-independent noise and $\mathbf{z}_1 \sim \mathcal{N}(0, \varsigma\sigma^2)$ is a zero-mean, variance $\varsigma\sigma^2$ Gaussian random variable describing the input-dependent noise. Here $\mathbf{z}_0$ and $\mathbf{z}_1$ are assumed to be independent. The parameter $\sigma^2 > 0$ describes the strength of the
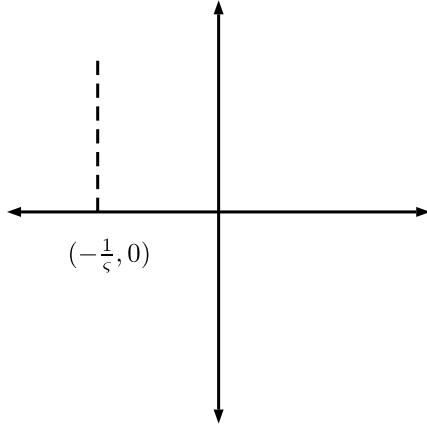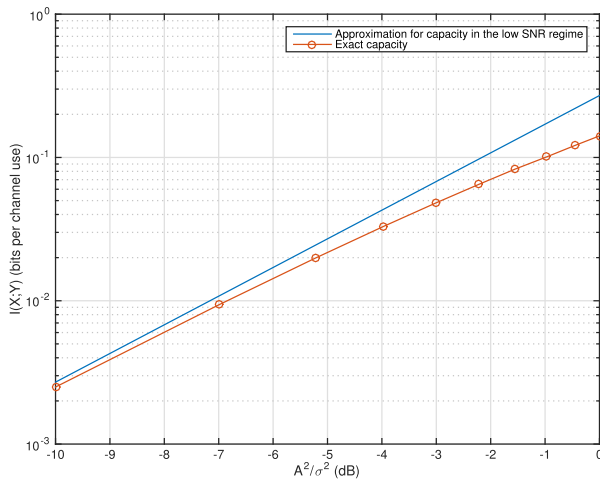
Fig. 3.   The branch cut of the function $\log(1 + \varsigma z)$.

Fig. 5.   The capacity of intensity modulated optical communication channel along with upper and lower bounds.

Fig. 4.   Asymptotic capacity of intensity modulated optical channel for low SNRs and the exact capacity.

Fig. 6.   The capacity of the magnetic recording system modeled as in [26].

input-independent noise, while $\varsigma > 0$ is the ratio of the input-dependent noise variance to the input-independent noise. Thus, $\sigma^2(x) = 1 + \varsigma x$. Moser [16] derives an approximation for the channel capacity for small signal-to-noise ratios (SNRs), a universal lower bound for any amplitude constraints, and an upper bound on the capacity that is valid only at high SNRs.

The function $\log(\sigma^2(z)) = \log(1 + \varsigma z)$ has a branch point at $z = -\frac{1}{\varsigma}$, hence we can define a branch cut as the line connecting the two points $\{(-\frac{1}{\varsigma}, 0), (-\frac{1}{\varsigma}, \tilde{\infty})\}$, where $\tilde{\infty}$ represents the complex infinity. Then, the extension of the function $\sigma^2(x)$ is well defined on the entire complex plane except the line connecting the two points $\{(-\frac{1}{\varsigma}, 0), (-\frac{1}{\varsigma}, \tilde{\infty})\}$, as shown in Fig. 3. Fig. 4 shows the capacity of this channel at low SNRs along with the approximate formula, and Fig. 5 shows the capacity along with the lower and upper bounds in [16].

As a further illustration, we present another example in which the signal-dependent noise term appears as the dominant noise component, i.e., in magnetic recording systems where the media noise is strongly signal-dependent and is modeled as Gaussian noise with variance $\sigma^2(x) = 1 + \sqrt{1 - x^2}$, where the input signal $|x| < 1$ [26]. The extension of the function
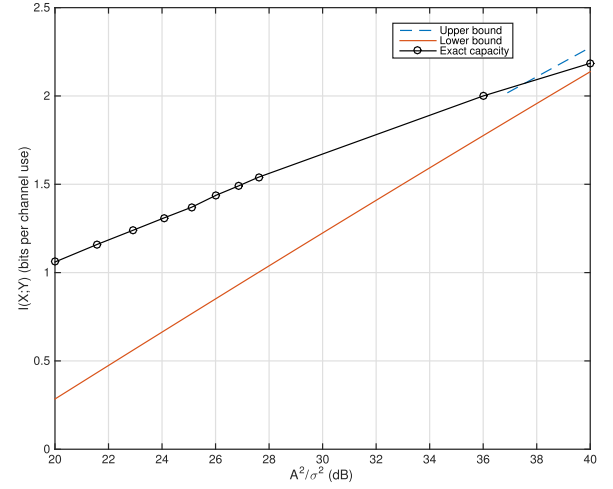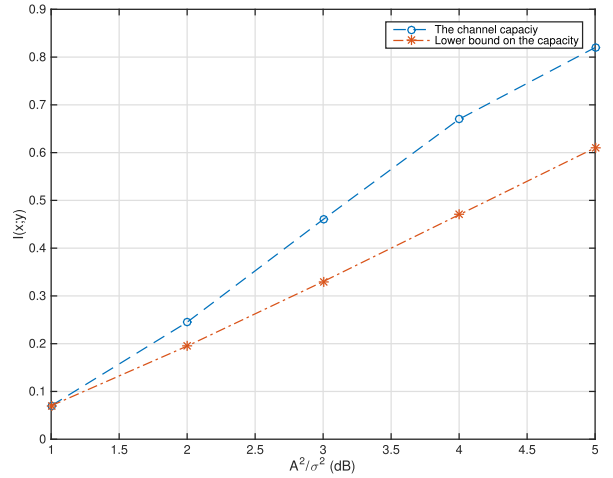
$\log(\sigma^2(x))$ to the complex plane $\log(1 + \sqrt{1 - z^2})$ has two branch points, $(1, 0)$, $(-1, 0)$, the branch cuts can be chosen such that they do not include other parts of the real line. Fig. 6 shows the capacity of this magnetic recording system along with a lower bound. We note that this channel model does not fall within the framework of [13] as the required technical condition (that is, the noise variance converging to zero for some limit point in the set of admissible channel inputs where the extension of the noise variance function behaves well) is not satisfied, however, our results apply indicating that the optimal input distribution is discrete. The lower bound is computed based on an evaluation of the mutual information with a suboptimal input distribution (namely truncate Gaussian distributions) in [26]. We observe that the actual channel capacity is significantly higher than the mutual information evaluated previously.

VII. CONCLUSIONS

We study the capacity of two classes of channels with amplitude limited inputs, namely, fading channels with the

channel state information available only at the receiver, and signal dependent Gaussian noise channels. The former model is useful for many wireless communications settings while the latter is suitable for other applications including optical communication and magnetic recording channels. For the first model, we prove that if the fading coefficients have a finite support (but otherwise arbitrary), the capacity achieving distribution is discrete with a finite number of mass points. For the latter, we prove that under certain technical conditions on the noise variance function, the capacity is achieved by discrete inputs with a finite number of mass points as well. For both cases, the capacity computation is a finite-dimensional optimization problem and hence the optimal input distribution and the channel capacity can be computed efficiently. The findings are illustrated via several examples and through comparisons with existing results in the literature.

APPENDIX A
THE MUTUAL INFORMATION DENSITY IS EXTENDABLE TO THE COMPLEX PLANE

In this appendix, we show that the mutual information density of the signal-dependent Gaussian noise channel can be extended to the complex plane. First, we assume that the function $\log(\sigma^2(x))$ can be extended to the entire complex plane $\mathbb{C}$ excluding the branch cuts of $\log(\sigma^2(z))$. We define

$$\sigma^2(z) \triangleq \sigma_r^2(z) + j\sigma_i^2(z). \tag{A.1}$$

where $\sigma_r^2(z)$ and $\sigma_i^2(z)$ are the real and imaginary parts of $\sigma^2(z)$, respectively. Since we assume that the extension of the noise variance to the complex plane is well defined, we have, $\forall\ z\ \in \mathbb{C}$, i.e., $z = a + jb$, $|\sigma^2(z)| < \infty$, $|\sigma_r^2(z)| < \infty$, and $|\sigma_i^2(z)| < \infty$.

The function $h_{F_{\mathbf{x}}}(x)$ given in (29) can be extended to the entire complex plane $\mathbb{C}$ by showing that $\forall\ z$ s.t. $|z| < \infty$.

$$
\begin{aligned}
|h_{F_{\mathbf{x}}}(z)| &\leq \int_{-\infty}^{\infty} |f_{\mathbf{n}}(y - z, z)|\, \big|\log(f_{\mathbf{y}}(y; F_{\mathbf{x}}))\big|\, \mathrm{d}y, \\
&= \frac{1}{\sqrt{2\pi |\sigma^2(z)|}} \int_{-\infty}^{\infty} \left| \exp\left(-\frac{(y - a - jb)^2}{2\sigma_r^2(z) + j2\sigma_i^2(z)} \right) \right| \\
&\qquad\qquad \big|\log(f_{\mathbf{y}}(y; F_{\mathbf{x}}))\big|\, \mathrm{d}y, \\
&= \frac{\eta}{\sqrt{2\pi |\sigma^2(z)|}} \int_{-\infty}^{\infty} \exp\left(-\frac{(y - \zeta)^2}{\epsilon} \right) \\
&\qquad\qquad \big|\log(f_{\mathbf{y}}(y; F_{\mathbf{x}}))\big|\, \mathrm{d}y, \\
&\leq \frac{\eta}{\sqrt{2\pi |\sigma^2(z)|}} \int_{-\infty}^{\infty} \exp\left(-\frac{(y - \zeta)^2}{\epsilon} \right) \\
&\qquad\qquad \big[-\log(\gamma'(y)) + 2|\log k_3'| \big]\, \mathrm{d}y < \infty,
\end{aligned}
$$

where η, ζ, and ϵ are real values that depend on $a, b, \sigma_r^2, \sigma_i^2$. The last step follows since we can show the finiteness of the integral as in [6]. Thus, $h_{F_{\mathbf{x}}}(x)$ can be extended to the complex plane, and hence, $i_{F_{\mathbf{x}}}(x)$ can also be extended to the complex plane.

APPENDIX B
THE MUTUAL INFORMATION DENSITY IS AN ANALYTIC FUNCTION

In this appendix, we show that the mutual information density of signal-dependent Gaussian noise channels is analytic under some conditions on the noise variance. We assume that the function $\log(\sigma^2(x))$ can be extended to an open connected set in the complex plane containing the real line excluding some branch points. We also note that, since $\sigma^2(z)$ is an analytic function (by assumption) and $\log(\sigma^2(z))$ is well defined, the function $\log(\sigma^2(z))$ is analytic over some open connected set in the complex plane excluding some branch cuts.

First, we show that the function $h_{F_{\mathbf{x}}}(z)$ is continuous on any open connected set $\mathcal{D}_\delta$. If we can show that there is an integrable function $g\ :\ \mathbb{R}\ \rightarrow\ [0, \infty)$ such that $\big|f_{\mathbf{n}}(y - z_n, z_n) \log(f_{\mathbf{y}}(y; F_{\mathbf{x}}))\big|\ \leq\ g(y)$ for any $y \in \mathbb{R}$, and $\int_{-\infty}^{\infty} g(y)\, \mathrm{d}y\ <\ \infty$ then we can invoke the Dominated Convergence Theorem to conclude the continuity of $f_{\mathbf{n}}(y - z_n, z_n)$.

Let $\{z_n\}_{n \geq 1}$ be a sequence of complex numbers in $\mathcal{D}_\delta$ converging to $z \in \mathcal{D}_\delta$. Let $z_n = \eta_n + j\xi_n$ such that $|\xi_n| \leq \delta$. For a fixed $\epsilon > 0$, there exists $t \in \mathbb{N}$ for which for all $n \geq t$ such that $|\eta - \eta_n| < \epsilon$. Thus, the extension of the entropy density for the sequence $\{z_n\}_{n \geq t}$ is

$$h_{F_{\mathbf{x}}}(z_n) = -\int_{-\infty}^{\infty} f_{\mathbf{n}}(y - z_n, z_n) \log(f_{\mathbf{y}}(y; F_{\mathbf{x}}))\, \mathrm{d}y. \tag{B.1}$$

We define the extension of the noise variance as

$$\sigma^2(z_n) \triangleq \sigma_r^2(z_n) + j\sigma_i^2(z_n), \tag{B.2}$$

where $\sigma_r^2(z_n)$ and $\sigma_i^2(z_n)$ are the real and imaginary parts of $\sigma^2(z_n)$, respectively. We have,

$$
\begin{aligned}
&|f_{\mathbf{n}}(y - z_n, z_n)| \\
&= \left| \frac{1}{\sqrt{2\pi(\sigma^2(z_n))}} \exp\left(-\frac{(y - z_n)^2}{2\sigma^2(z_n)} \right) \right|, \\
&= \frac{1}{\sqrt{2\pi\, |\sigma^2(z_n)|}} \left| \exp\left(-\frac{(y - z_n)^2}{2(\sigma^2(z_n))} \right) \right|, \\
&\leq \frac{1}{\sqrt{2\pi\, |\sigma^2(z_n)|}} \exp\left(\frac{\delta^2}{2\left((\sigma_r^2(z_n))^2 + (\sigma_i^2(z_n))^2 \right)} \right) \\
&\quad \times \exp\left(\frac{\delta^2 (\sigma_i^2(z_n))^2}{2\left((\sigma_r^2(z_n))^2 + (\sigma_i^2(z_n))^2 \right) \sigma_r^2(z_n)} \right) \\
&\quad \times \exp\left(-\frac{\sigma_r^2(z_n)\left(y - \eta_n - \delta\frac{\sigma_i^2(z_n)}{\sigma_r^2(z_n)} \right)^2}{2\left((\sigma_r^2(z_n))^2 + (\sigma_i^2(z_n))^2 \right)} \right).
\end{aligned}
$$

Let us define the following,

$$m_1 \triangleq \sup_{\substack{\lambda \in \mathbb{C} \\ \eta - \epsilon \leq \mathrm{Re}\{\lambda\} \leq \eta + \epsilon \\ |\mathrm{Im}\{\lambda\}| \leq \delta}} \frac{1}{\sqrt{2\pi\, |\sigma^2(\lambda)|}},$$

$$m_2 \triangleq \sup_{\substack{\lambda \in \mathbb{C} \\ \eta - \epsilon \leq \mathrm{Re}\{\lambda\} \leq \eta + \epsilon \\ |\mathrm{Im}\{\lambda\}| \leq \delta}} \frac{\sigma_r^2(\lambda)}{2\left((\sigma_r^2(\lambda))^2 + (\sigma_i^2(\lambda))^2 \right)},$$

$$m_3 \triangleq \sup_{\substack{\lambda \in \mathbb{C} \\ \eta - \epsilon \leq \mathrm{Re}\{\lambda\} \leq \eta + \epsilon \\ |\mathrm{Im}\{\lambda\}| \leq \delta}} \frac{(\sigma_i^2(\lambda))^2}{2\left((\sigma_r^2(\lambda))^2 + (\sigma_i^2(\lambda))^2\right)\sigma_r^2(\lambda)},$$

$$m_4 \triangleq \sup_{\substack{\lambda \in \mathbb{C} \\ \eta - \epsilon \leq \mathrm{Re}\{\lambda\} \leq \eta + \epsilon \\ |\mathrm{Im}\{\lambda\}| \leq \delta}} \left(\frac{\sigma_i^2(\lambda)}{\sigma_r^2(\lambda)}\right)^2,$$

$$m_5 \triangleq \inf_{\substack{\lambda \in \mathbb{C} \\ \eta - \epsilon \leq \mathrm{Re}\{\lambda\} \leq \eta + \epsilon \\ |\mathrm{Im}\{\lambda\}| \leq \delta}} \left(\frac{\sigma_i^2(\lambda)}{\sigma_r^2(\lambda)}\right)^2,$$

and finally,

$$m^* \triangleq m_1 \exp\left(\delta^2(m_3 + m_2)\right). \tag{B.3}$$

Thus, we can construct an upper bound on the extension of the noise PDF by defining $\tau(y, \epsilon, \delta)$ as,

$$\tau(y,\epsilon,\delta) \triangleq \begin{cases} m^* \exp\left(-m_2(y - \eta + \epsilon - \delta m_5)^2\right) \\ \qquad\qquad \text{if } y < \eta - \epsilon + \delta m_5, \\ m^* \qquad \text{if } \eta - \epsilon + \delta m_5 \leq y \leq \eta + \epsilon + \delta m_4, \\ m^* \exp\left(-m_2(y - \eta - \epsilon - \delta m_4)^2\right) \\ \qquad\qquad \text{if } y > \eta + \epsilon + \delta m_4. \end{cases} \tag{B.4}$$

Namely,

$$|f_{\mathbf{n}}(y - z_n, z_n)| \leq \tau(y, \epsilon, \delta). \tag{B.5}$$

We can write

$$|f_{\mathbf{n}}(y - z_n, z_n)| \, |\log(f_{\mathbf{y}}(y; F_{\mathbf{x}}))| \\ \leq \tau(y,\epsilon,\delta)\left[-2\log(\gamma'(y)) + 2|\log(k_3')|\right].$$

In the case of $\eta > \epsilon - \delta m_5$,

$$|h_{F_{\mathbf{x}}}(z_n)| \leq \left| \int_{-\infty}^{\infty} f_{\mathbf{n}}(y - z_n, z_n) \log(f_{\mathbf{y}}(y; F_{\mathbf{x}})) \, dy \right|,$$
$$\leq \int_{-\infty}^{\infty} \tau(y,\epsilon,\delta)\left[-2\log(\gamma'(y)) + 2|\log(k_3')|\right],$$
$$< \infty.$$

Similarly, we can show the finiteness of $|h_{F_{\mathbf{x}}}(z_n)|$ for the other cases of $\eta < -\epsilon + \delta m_4$ and $-\epsilon + \delta m_4 < \eta < \epsilon - \delta m_5$. Thus, we can invoke the Dominated Convergence Theorem to conclude the continuity of $h_{F_{\mathbf{x}}}(z)$ for any domain $\mathcal{D}_\delta$.

We now show that $\left|\oint_\omega h_{F_{\mathbf{x}}}(z) \, dz\right|$ is finite. Let us define M_ω as

$$M_\omega = \max_{z \in \omega} |h_{F_{\mathbf{x}}}(z)|, \tag{B.6}$$

where ω is any simple closed contour. Hence,

$$\left|\oint_\omega h_{F_{\mathbf{x}}}(z) \, dz\right| = \left|\oint_\omega \int_{-\infty}^{\infty} f_{\mathbf{n}}(y - z, z) \log(f_{\mathbf{y}}(y; F_{\mathbf{x}})) \, dy \, dz\right|,$$
$$\leq \oint_\omega \int_{-\infty}^{\infty} |f_{\mathbf{n}}(y - z, z) \log(f_{\mathbf{y}}(y; F_{\mathbf{x}}))| \, dy \, dz,$$
$$\leq \oint_\omega M_\omega \, dz,$$
$$\leq M_\omega l_\omega < \infty, \tag{B.7}$$

where l_ω is the length of ω which is finite as ω is a closed contour.

We then argue that the extension of the entropy density $h_{F_{\mathbf{x}}}(x)$ is analytic. We note that the extension of the noise PDF $f_{\mathbf{n}}(y - z, z)$ is analytic since $\sigma^2(z)$ is analytic (by assumption). We have,

$$\oint_\omega h_{F_{\mathbf{x}}}(z) \, dz$$
$$= -\oint_\omega \int_{-\infty}^{\infty} f_{\mathbf{n}}(y - z, z) \log(f_{\mathbf{y}}(y; F_{\mathbf{x}})) \, dy \, dz,$$
$$\overset{(a)}{=} -\int_{-\infty}^{\infty} \log(f_{\mathbf{y}}(y; F_{\mathbf{x}})) \oint_\omega f_{\mathbf{n}}(y - z, z) \, dy \, dz = 0,$$

where the order of integration in (a) is changed due to (B.7). Since this is true for any simple contour ω, by invoking Morera's Theorem, we conclude that the extension of the entropy density function $h_{F_{\mathbf{x}}}(z)$ is analytic on any open connected set on the complex plane $\mathbb{C}$.

REFERENCES

[1] A. ElMoslimany, "A new communication scheme implying amplitude limited inputs and signal dependent noise: System design, information theoretic analysis and channel," Ph.D. dissertation, Dept. Elect., Comput. Energy Eng., Arizona State Univ., Tempe, AZ, USA, Dec. 2015.

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

[3] A. Tchamkerten, "On the discreteness of capacity-achieving distributions," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2773–2778, Nov. 2004.

[4] S. Shamai (Shitz) and I. Bar-David, "The capacity of average and peak-power-limited quadrature Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 41, no. 4, pp. 1060–1071, Jul. 1995.

[5] M. Katz and S. Shamai (Shitz), "On the capacity-achieving distribution of the discrete-time noncoherent and partially coherent AWGN channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2257–2270, Oct. 2004.

[6] B. Mamandipoor, K. Moshksar, and A. K. Khandani, "On the sum-capacity of Gaussian MAC with peak constraint," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2012, pp. 26–30.

[7] B. Mamandipoor, K. Moshksar, and A. K. Khandani, "Capacity-achieving distributions in Gaussian multiple access channel with peak power constraints," *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 6080–6092, Oct. 2014.

[8] O. Ozel and S. Ulukus, "On the capacity region of the Gaussian MAC with batteryless energy harvesting transmitters," in *Proc. IEEE Global Commun. Conf.*, Dec. 2012, pp. 2385–2390.

[9] M. Raginsky, "On the information capacity of Gaussian channels under small peak power constraints," in *Proc. Annu. Allerton Conf. Commun., Control, Comput.*, Sep. 2008, pp. 286–293.

[10] J. Cao, J. Chen, and S. Hranilovic, "Discreteness of sum-capacity-achieving distributions for discrete-time poisson multiple access channels with peak constraints," *IEEE Commun. Lett.*, vol. 17, no. 8, pp. 1644–1647, Aug. 2013.

[11] A. Das, "Capacity-achieving distributions for non-Gaussian additive noise channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2000, p. 432.

[12] A. ElMoslimany and T. M. Duman, "On the capacity of multiple-antenna systems and parallel Gaussian channels with amplitude-limited inputs," *IEEE Trans. Commun.*, vol. 64, no. 7, pp. 2888–2899, Jul. 2016.

[13] 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.

[14] I. C. Abou-Faycal, M. D. Trott, and S. Shamai (Shitz), "The capacity of discrete-time memoryless Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1290–1301, May 2001.

[15] M. C. Gursoy, H. V. Poor, and S. Verdú, "The noncoherent Rician fading channel—Part I: Structure of the capacity-achieving input," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2193–2206, Sep. 2005.

[16] S. M. Moser, "Capacity results of an optical intensity channel with input-dependent Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 207–223, Jan. 2012.

[17] J. G. Smith, "On the information capacity of peak and average power constrained Gaussian channels," Ph.D. dissertation, Dept. Elect. Eng., Univ. California, Berkeley, CA, USA, 1969.

[18] R. M. Dudley, *Real Analysis and Probability*, vol. 74. Cambridge, U.K.: Cambridge Univ. Press, 2002.

[19] J. B. Conway, *Functions of One Complex Variable*, vol. 2. New York, NY, USA: Springer, 1973.

[20] R. B. Ash, *Information Theory*. New York, NY, USA: Dover, 1965.

[21] R. M. Gray, *Entropy and Information Theory*. New York, NY, USA: Springer, 1990.

[22] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes*. San Francisco, CA, USA: Holden-Day, 1964.

[23] M. Fozunbal, S. W. McLaughlin, and R. W. Schafer, "Capacity analysis for continuous-alphabet channels with side information, part I: A general framework," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3075–3085, Sep. 2005.

[24] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2009.

[25] J. W. Brown and R. V. Churchill, *Complex Variables and Applications*, 8th ed. New York, NY, USA: McGraw-Hill, 2009.

[26] R. Wood, "Detection and capacity limits in magnetic media noise," *IEEE Trans. Magn.*, vol. 34, no. 4, pp. 1848–1850, Jul. 1998.

Ahmad ElMoslimany (S'09) received the BS degree from Ain Shams University, Cairo, Egypt, in 2009, MS degree from Nile University, Cairo, Egypt in 2011, PhD degree from Arizona State University, Tempe, Arizona, United States in 2015, all in Electrical Engineering. He is currently a Senior Research Engineer in Axxcelera Egypt, Cairo, Egypt.

Dr. Ahmad ElMoslimany was a Research Assistant with the Wireless Intelligent Network Center (WINC) at Nile University from October 2009 to August 2011. From September 2011 to December 2015, he was a Graduate Research Associate at the Arizona State University. From April 2016 to January 2017, he has been a Senior Research Engineer in Polaris Wireless, Mountain View, CA. His research interest includes signal processing, information theory, channel coding theory, communications theory, wireless channel modeling, MIMO communications and channel estimation.

Tolga M. Duman (S'95–M'98–SM'03–F'11) received the B.S. degree from Bilkent University, Ankara, Turkey, in 1993, and the M.S. and Ph.D. degrees from Northeastern University, Boston, MA, USA, in 1995 and 1998, respectively, all in electrical engineering. He is currently a Professor of Electrical and Electronics Engineering Department at Bilkent University, in Turkey, and an Adjunct Professor with the School of Electrical, Computer and Energy Engineering (ECEE) at Arizona State University. Prior to joining Bilkent University in September 2012, he was with the School of ECEE at Arizona State University, as an Assistant Professor from 1998 to 2004, an Associate Professor from 2004 to 2008, and a Professor after 2008. His current research interests are in systems, with particular focus on communication and signal processing, including wireless and mobile communications, coding/modulation, coding for wireless communications, data storage systems, and underwater acoustic communications.

Dr. Duman is a recipient of the National Science Foundation CAREER Award and the IEEE Third Millennium medal. He served as an Editor of the IEEE TRANSACTION ON WIRELESS COMMUNICATIONS from 2003 to 2008, the IEEE COMMUNICATIONS SURVEYS AND TUTORIALS from 2002 to 2007, the IEEE TRANSACTION ON COMMUNICATIONS from 2007 to 2012, and the Physical Communication (Elsevier) from 2010 to 2016. He has been the coding and information theory Area Editor of the IEEE TRANSACTION ON COMMUNICATIONS since 2011, an Editor of the IEEE TRANSACTION ON WIRELESS COMMUNICATION and the Editor-in-Chief of the *Physical Communication* (Elsevier) since 2016.