Received March 20, 2018, accepted May 1, 2018, date of publication May 15, 2018, date of current version June 26, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2837056

Capacity-Achieving Signals for Point-to-Point and Multiple-Access Channels Under Non-Gaussian Noise and Peak Power Constraint

MOHAMMAD RANJBAR[1], NGHI H. TRAN[1], (Senior Member, IEEE), TRUYEN V. NGUYEN[2], MUSTAFA CENK GURSOY[3], (Senior Member, IEEE), AND HUNG NGUYEN-LE[4]

[1]Department of Electrical and Computer Engineering, University of Akron, Akron, OH 44325, USA
[2]Department of Mathematics, University of Akron, Akron, OH 44325, USA
[3]Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244, USA
[4]The University of Danang, University of Science and Technology, Danang 551000, Vietnam

Corresponding author: Nghi H. Tran (nghi.tran@uakron.edu)

This work was supported in part by the Fostering Innovation through Research, Science and Technology (FIRST) Project (governed by Ministry of Science and Technology, Vietnam) under Grant 27/FIRST/1a/DNU and in part by the National Science Foundation, USA, under Grant 1509006.

ABSTRACT This paper generalizes and proves the discrete and finite nature of the capacity-achieving signaling schemes for general classes of non-Gaussian point-to-point and multiple-access channels (MACs) under peak power constraints. Specifically, we first investigate the detailed characteristics of capacity-achieving inputs for a single-user channel that is impaired by two types of noise: a Gaussian mixture (GM) noise Z consisting of Gaussian elements with arbitrary means and the interference U with an arbitrary distribution. The only very mild condition imposed on U is that its second moment is finite. To this end, one of the important results is the establishment of the Kuhn–Tucker condition (KTC) on a capacity-achieving input and the proof of analyticity of the KTC using Fubini–Tonelli's and Morera's theorems. Using the Bolzano–Weierstrass's and Identity's theorems, we then show that a capacity-achieving input is continuous if and only if the KTC function is zero on the entire real line. However, by examining an upper bound on the tail of the output PDF, it is demonstrated that the KTC function must be bounded away from zero. As such, any capacity-achieving input must be discrete with a finite number of mass points. Finally, we exploit U having an arbitrary distribution to show that the optimal input distributions that achieve the sum-capacity of an M-user MAC under GM noise are discrete and finite. We also prove that there exist at least two distinct points that achieve the sum capacity on the rate region.

INDEX TERMS Channel capacity, Gaussian mixture, multiple access channels, non-Gaussian interference, optimal inputs.

I. INTRODUCTION

The presence of multi-tier heterogeneous architectures mixed with limited spectral resources in current and future wired and wireless communication systems leads to increased co-channel interference, which is intermittent and asynchronous with the main communication. In many cases, the aggregate co-channel interference plus noise cannot be treated as Gaussian [2]–[7]. Examples of such channels include impulsive interference channels in digital subscriber line (DSL) [8]–[10], power line communications (PLC) [4], [11]–[14], underwater acoustic channels [15]–[17], and heterogeneous wireless networks involving marcocells, microcells, femtocells, device-to-device (D2D) links, Wi-Fi access points, and cognitive radio (CR) [2], [5]–[7], [18]. As a result, the traditional approach of using Gaussian signaling schemes for system analysis and design no longer holds. It should be noted that non-Gaussian interference is also observed in speech and audio signals [19], as well as in imaging science [20], [21]. For instance, in image processing, corruption of images due to impulsive interference is a significant hurdle. This interference is caused by a noisy sensor or sensor heat. It may also occur during image acquisition, recording, or transmission [20], [21].

Over the years, several interesting results on information-theoretical aspects of non-Gaussian channels have been obtained for single-user point-to-point channels [22]–[28].

For example, in [29], the characteristics of a capacity-achieving input were studied for channels under generalized Gaussian noise, and the discreteness of a capacity-achieving input was shown for some special cases of generalized Gaussian noise. However, a generalized Gaussian distribution is not general enough to represent any arbitrary distribution. In [22] and [23], predefined conditions on noise distributions have been examined to guarantee the discreteness of the optimal input under the input average power constraint. Unfortunately, for many non-Gaussian models, the noise distributions do not satisfy those pre-defined conditions. If a more practical limitation of the input peak constraint is imposed, the capacity-achieving input signal is shown to be finite and discrete [24]–[26]. Recently, [30] has considered a general class of non-Gaussian channels under the assumption of two conditions imposed on the noise distribution. These two conditions rely on a lower bound and an upper bound on the noise distribution. While these conditions are satisfied by several well-known distributions, they do not hold in more general cases. Thus, the characterizations of capacity-achieving inputs for more general non-Gaussian point-to-point channels remain to be elucidated. It is worth mentioning that the discreteness of the optimal input amplitude has also been observed in Gaussian channels under the average power constraint together with other constraints, such as peak constraint [31], [32], non-coherent conditions [33]–[36], rapid phase variations [37], or duty cycle [38]. Due to the difficulty in studying the detailed properties of the optimal input and in establishing the capacity in closed-form for a non-Gaussian channel, a novel approach was proposed in [39] to study the sensitivity of the capacity, i.e., how the capacity changes when the channel parameters vary.

While information-theoretical results for point-to-point links are certainly useful, it is also important to extend the analysis and design to multi-user networks. One of the most fundamental network scenarios is the multiple access channel (MAC) for the uplink, and this scenario has been well-investigated for Gaussian channels under average power constraints with many breakthroughs. Under this line of work, Gaussian signals are usually assumed to be optimal for the design and analysis. In a recent work in [40], a two-user Gaussian MAC under peak power constraints is investigated. It has been shown in [40] that both users can use finite and discrete signaling schemes to achieve any point on the boundary of the capacity region. The proof in [40] is relied on the key result that for a point-to-point link that is affected by a Gaussian noise and an interference having a finite support, the capacity-achieving input is discrete under the peak power constraint. To our knowledge, network information-theoretical results for non-Gaussian noise and interference are lacking. Therefore, considering non-Gaussian interference plus noise in multi-user networks presents new challenges.

In this work, we attempt to generalize and prove the discrete and finite nature of the capacity-achieving signaling schemes for general classes of non-Gaussian point-to-point and multiple-access channels under peak power constraints. Specifically, we first investigate the detailed characteristics of capacity-achieving inputs for a single-user channel that is impaired by two types of noise: a Gaussian mixture (GM) noise Z consisting of Gaussian elements with arbitrary means, and the interference U with an arbitrary distribution. Different from [40] that assumed a finitely supported U, the only very mild condition we impose on U is that its second moment is finite. Note that when U is either Gaussian or GM, the total noise $W = U + Z$ is a GM. We then extend the results to a general M-user Gaussian mixture MAC. Different from the previous literature that only investigated GM noise with zero-mean components (please see [27], [41], and references therein), our considered GM noise with arbitrary-mean Gaussian components can be used to approximate any distribution of engineering interest to arbitrary accuracy [42]–[45]. The significance of our results is two-fold as follows:

- Since the considered GM is general enough to represent all PDFs of engineering interest, the proof of the discreteness and finiteness of the capacity-achieving input partially addresses the conjecture in [24] that the capacity-achieving input for any peak power constrained channel is discrete with finite number of mass points. The results in this work can be used in system analysis for various channels impaired by different sources of interference, such as cognitive radio with imperfect spectrum sensing and 5G and beyond heterogeneous cellular networks [2], [3], [5]–[7], [18], [46].
- The obtained results also shed important light on the characteristics of capacity-achieving signaling schemes for a multiple-access non-Gaussian channel, which has not been explored in the literature so far.

In this paper, we follow a standard procedure to prove the optimality of discrete input by first showing the existence of the optimal solution before eliminating the possibility of having an optimal continuous input. However, the novelty in technical derivations and analysis makes our work stand out from the current state-of-the-art. Specifically, for the point-to-point channel, by using lower and upper bounds on the probability density functions (PDFs) of the total noise W and the output, we first show that the mutual information (MI) is continuous and concave. Therefore, the capacity-achieving input exists. We then establish a necessary and sufficient Kuhn-Tucker condition (KTC) on an optimal input, and show that the KTC function is analytic on the complex plane. While this analyticity is obvious for Gaussian channels, the consideration of the general noise model $W = U + Z$ is much more challenging, and it requires new techniques. In particular, our proof relies on Fubini-Tonelli's and Morera's theorems, which is very different from the standard approach employed in [31]. Then using the Bolzano-Weierstrass and Identity Theorems, we show that a capacity-achieving input is continuous if only if the KTC is zero on the entire real line. However, by establishing and examining an upper bound on the tail of the output PDF, the KTC is proved to be bounded away from zero. As such, any capacity-achieving input must be discrete with a finite number of mass points. Note that the traditional approach of using the Fourier transform of the noise distribution as in [31] cannot be used, since there exist zeros in the transformation. Finally, building upon the results of the point-to-point channel, we characterize the sum-capacity achieving signaling schemes of a multiple access channel under GM noise. Specifically, we exploit the arbitrary nature of U to show that the optimal input distributions that achieve the sum-capacity of an M-user MAC under GM noise are discrete and finite. In addition, we prove that there exist at least two different points that achieve the sum capacity on the rate region.

The results and tools developed in this work will be important to major communication systems, such as power line communications, digital subscriber lines, urban indoor and outdoor wireless communications, cognitive radio, and underwater acoustic communications. More importantly, insights from the developed information-theoretic results can be exploited to develop powerful and efficient signaling and coding techniques for practical purposes. In addition, other disciplines such as imaging science can also benefit from the proposed results and methodologies to advance the current state-of-the-art.

The rest of this paper is organized as follows. Sec. II introduces the considered single-user and multiple-access non-Gaussian channels. Detailed characteristics of a capacity-achieving input for the single-user channel are studied in Sec. III. Numerical examples are also provided in this section to confirm the analysis. In Sec. IV, we address the characterization of optimal inputs for a multiple-access channel. Finally, conclusions are drawn in Sec. V.

II. POINT-TO-POINT AND MULTI-USER NON-GAUSSIAN CHANNELS: CHANNEL MODELS

A. SINGLE-USER CHANNEL

We consider a non-Gaussian point-to-point channel with the following input-output model

$$V = X + U + Z, \tag{1}$$

where X and V are the channel input and output, respectively. Here, it is assumed that the input is subject to a peak power constraint $|X| \leq A$, which is a practical constraint for any communication system. In (1), U and Z are two sources of additive noise/interference. We assume that U follows a fixed yet arbitrary distribution with only a mild condition that its second moment is finite. This assumption in fact holds true for almost all practical interference sources. In addition, $Z \in \mathbb{R}$ is a Gaussian Mixture (GM) with its PDF being a weighted sum of N Gaussian component densities with arbitrary means. The PDF of Z can be expressed as

$$f_Z(z) = \sum_{n=1}^{N} \varepsilon_n \mathcal{N}(z, a_n, \sigma_n^2), \tag{2}$$

where $\mathcal{N}(z, \mu, \sigma^2)$ denotes a real Gaussian distribution with mean μ and variance σ^2, and $\varepsilon_n > 0$ is the mixing probability satisfying $\sum_{n=1}^{N} \varepsilon_n = 1$. For convenience, let $W = U + Z$ be the total noise. It is not difficult to realize that when U is selected as either Gaussian, GM or a discrete distribution, W becomes a GM.

For a given input cumulative distribution function (CDF) $F_X(x)$, the mutual information between the input and output of the channel can be calculated as

$$I(X; V) = h(V; F_X(x)) - h(V \mid X) \triangleq I(F_X(x)), \tag{3}$$

where $h(V; F_X(x))$ is the output entropy given as

$$h(V; F_X(x)) = -\int_{\mathbb{R}} f_V(v; F_X(x)) \ln f_V(v; F_X(x))\, dv. \tag{4}$$

Note that $f_V(v; F_X(x))$ is the PDF of the output V for a given input $F_X(x)$. In addition, $h(V \mid X) = h(W)$ is the total noise entropy. Under the peak input power constraint $|X| \leq A$, the Shannon capacity of the channel is the supremum of $I(X; V)$ over the feasible set of $F_X(x)$, which is expressed as

$$C = \sup_{F_X(x) \in \mathcal{F}} I(F_X(x)), \tag{5}$$

where $\mathcal{F}$ is the set of all CDFs having all mass points in the interval $[-A, A]$, i.e., $\mathcal{F} = \left\{F : \int_{-A}^{A} dF(x) = 1\right\}$. The optimal input distribution, which is denoted as $F_X^{\star}(x)$, is referred to as the capacity-achieving input. Since $h(W)$ is independent of $F_X(x)$, the capacity-achieving input maximizes the output entropy. For convenience, hereafter, the use of $F_X(x)$ and F_X to indicate the distribution of the input is interchangeable.

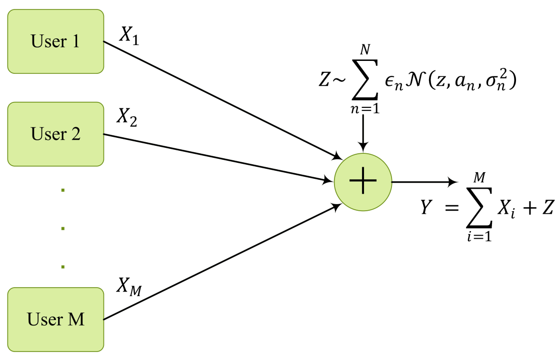

FIGURE 1. An M-user multiple-access channel (MAC) under GM noise Z.

B. MULTIPLE-ACCESS CHANNEL (MAC) UNDER GM NOISE

In this paper, we are also interested in a MAC under a GM noise Z in (2) in which M users communicate with a common receiver as shown in Fig. 1. Let's assume that X_i is the signal of user i, $i = 1 \cdots M$, and it is imposed by the peak constraint $|X_i| \leq A_i$. The received signal Y is then written as:

$$Y = \sum_{i=1}^{M} X_i + Z. \tag{6}$$

Note that with a suitable choice of $\{N, \varepsilon_n, a_n, \sigma_n\}$, a GM Z can be used to represent any PDF of engineering interest [42]–[45]. Furthermore, with arbitrary mean Gaussian components, the GM model considered in this work is more general than other GM channels studied earlier using all zero-mean Gaussian elements.

For the considered MAC, the sum-capacity is the supremum of the joint MI between the inputs and output over the feasible set of the inputs, and it is given as:

$$C_{\text{sum}} = \sup_{F_{X_i}\in\mathcal{F}_i, i=1\cdots M} I\left(X_1, X_2, \ldots, X_M; Y\right) \triangleq \sup_{F_{X_i}\in\mathcal{F}_i, i=1\cdots M} I\left(\{F_{X_i}\}_{i=1}^{M}\right). \quad (7)$$

In (7), each set $\mathcal{F}_i$, $1 \le i \le M$, consists of all CDFs having mass points in the interval $[-A_i, A_i]$. A given set of the input distributions $\{F^{\star}_{X_i}\}$ that achieves C_{sum} is called the optimal set.

III. DETAILED CHARACTERISTICS OF OPTIMAL $F_X^{\star}$ IN POINT-TO-POINT CHANNELS

In this section, the focus is on the point-to-point channel in (1). In the following, we will first show the existence of an optimal input distribution $F_X^{\star}$ before investigating its detailed characteristics.

A. EXISTENCE OF $F_X^{\star}$

It has been well-known that the set $\mathcal{F}$ of all distributions F_X satisfying the peak power constraint is convex and weakly compact with respect to weak* topology [47]. In addition, the output entropy and the mutual information are concave with respect to F_X. As such, to prove the existence of the optimal $F_X^{\star}$, we only need to show the weak continuity of the output entropy on the feasible set of input distributions. Therefore, the key steps are to show the weak continuity of $f_V(v; F_X)$ on F_X and obtain an integrable upper bound on $|f_V(v; F_X)\ln f_V(v; F_X)|$. Toward this end, we first have the following proposition regarding the continuity of the PDF $f_W(\cdot)$.

Proposition 1: The function $f_W(w) : \mathbb{R} \to \mathbb{R}$ is continuous.

Proof: We first have:

$$f_W(w) = \int_{-\infty}^{\infty} f_Z(w-u)dF_U(u). \quad (8)$$

Now, let $\{w_n\} \in \mathbb{R}$ be a sequence such that $\lim_{n\to\infty} w_n = w_0$ for some $w_0 \in \mathbb{R}$. Since the GM PDF $f_Z(z)$ is continuous, we have

$$\lim_{n\to\infty} f_Z(w_n - u) = f_Z(w_0 - u), \quad \forall u \in \mathbb{R}.$$

Without loss of generality, it is assumed that $\sigma_1 \ge \cdots \ge \sigma_N$. Therefore, $f_Z(z) \le \frac{1}{\sqrt{2\pi\sigma_N^2}}$. It then follows that

$$0 \le f_Z\left(w_n - u\right) \le \frac{1}{\sqrt{2\pi\sigma_N^2}}. \quad (9)$$

Then by applying Dominated Convergence Theorem [48], we can conclude that

$$\lim_{n\to\infty} f_W(w_n) = f_W(w_0),$$

which implies the continuity of $f_W(w)$. ■

Note that in a similar manner, we can show the continuity of $f_V(v; F_X)$ in v for a given F_X. The next proposition states the weak continuity of $f_V(v; F_X)$ with respect to the input distribution.

Proposition 2: $f_V(v; F_X)$ is weakly continuous in F_X.

Proof: Assume we have a sequence $\{F_X^{(n)}\}$ that weakly converges to some specific $F_X^{(0)}$, i.e., $\{F_X^{(n)}\} \overset{w*}{\to} F_X^{(0)}$. Besides its continuity, $f_W(\cdot)$ is also upper bounded by the constant $d = \sum_{n=1}^{N} \frac{\varepsilon_n}{\sqrt{2\pi\sigma_n^2}}$ as shown in Appendix A. In addition, we have $f_V(v; F) = \int_{-\infty}^{\infty} f_W(v-x)\mathrm{d}F(x)$. Therefore,

$$\lim_{n\to\infty} f_V(v; F_X^{(n)}) = \lim_{n\to\infty} \int_{-\infty}^{\infty} f_W(v-x)\mathrm{d}F_X^{(n)}(x)$$
$$= \int_{-\infty}^{\infty} f_W(v-x)\mathrm{d}F_X^{(0)}(x) \quad (10)$$
$$= f_V(v; F_X^{(0)}) \quad (11)$$

It should be noted that (10) comes from Helly-Bray theorem [49]. ■

We then have the following theorem regarding the weak continuity of the output entropy $h\left(V; F_X(x)\right)$.

Theorem 1: The entropy $h\left(V; F_X\right)$ is weakly continuous in F_X.

Proof: First, assume that $F_X^{(n)} \overset{w*}{\to} F_X^{(0)}$. From Proposition 2, we have $\lim_{n\to\infty} f_V(v; F_X^{(n)}) = f_V(v; F_X^{(0)})$ for all $v \in \mathbb{R}$. Moreover, since $f(\xi) = \xi \ln \xi$ is continuous for $\xi > 0$, we obtain

$$\lim_{n\to\infty} f_V(v; F_X^{(n)}) \ln f_V(v; F_X^{(n)}) = f_V(v; F_X^{(0)}) \ln f_V(v; F_X^{(0)}).$$

Therefore, by Lebesgue Dominated Convergence Theorem, to prove the weak continuity of output entropy, it is enough to show that $|f_V(v; F_X^{(n)}) \ln f_V(v; F_X^{(n)})|$ is upper bounded by an integrable function of v uniformly in n. First, as shown in Appendix A, there always exist positive k_0 and k_1 such that

$$f_W\left(w\right) \le k_1 w^{-2} \text{ when } |w| > k_0. \quad (12)$$

Therefore, for all $|v| > A + k_0$ we have:

$$f_V\left(v; F_X^{(n)}\right) = \int_{-A}^{A} f_W(v-x)dF_X^{(n)}(x)$$
$$\le \int_{-A}^{A} k_1\left(v-x\right)^{-2} dF_X^{(n)}(x)$$
$$\le k_1\left(|v| - A\right)^{-2}. \quad (13)$$

In addition, as we demonstrate in Appendix B, $f_V\left(v; F_X^{(n)}\right)$ is always upper bounded by the constant d. Therefore, we have

$$f_V\left(v; F_X^{(n)}\right) \le \begin{cases} k_1\left(|v| - A\right)^{-2}, & |v| > A + k_0 \\ d, & |v| \le A + k_0. \end{cases} \quad (14)$$

It is well known that $\ln x^{\gamma-1} \le x^{\gamma-1}$ for $0 < x, \gamma < 1$, and $\ln x \le \frac{x^{\gamma}}{\gamma}$ for any positive x and γ. As such, for any $0 < \gamma < 1$, we have $|x \ln x| \le \max\{\frac{1}{1-\gamma}, \frac{\beta}{\gamma}\}x^{\gamma}$ when $0 < x < \beta$ for any positive β. By combining this with (14), we have

$$|f_V(v; F_X^{(n)}) \ln f_V(v; F_X^{(n)})| \le g(v), \quad (15)$$

where $g(v)$ is given as:

$$g(v) = \begin{cases} \max\{\frac{1}{1-\gamma}, \frac{d}{\gamma}\} \left(k_1 \left(|v| - A\right)^{-2}\right)^{\gamma}, & |v| > A + k_0 \\ \max\{\frac{1}{1-\gamma}, \frac{d}{\gamma}\} d^{\gamma}, & |v| \le A + k_0. \end{cases} \tag{16}$$

It can be verified that $g(v)$ is an integrable function for $\frac{1}{2} < \gamma < 1$. Therefore, $h(V; F_X)$ is weakly continuous. ■

Given the result in Theorem 1, it can be concluded that the output entropy, and as a consequence, the mutual information is weakly continuous and concave. Therefore, the optimal input distribution exists on $\mathcal{F}$.

B. DISCRETENESS OF $F_X^\star$

Given the existence of $F_X^\star$, this subsection focuses on the characterization of $F_X^\star$. In particular, we prove that the optimal input support set $E^\star$ of any optimal $F_X^\star$ has a finite number of mass points. To this end, we first establish the Kuhn-Tucker condition (KTC), which is a necessary and sufficient condition for an input to be optimal, and show that the KTC is analytic on the complex plane. By exploiting this analyticity of the KTC, we then demonstrate that it is not possible to have a continuous $F_X^\star$.

1) THE KTC AND ITS ANALYTICITY

For convenience, some preliminaries regarding the concepts of directional derivative and optimization theory are given in Appendix D. We then have the following theorem regarding the KTC:

Theorem 2 (KTC): $F_X^\star$ is an optimal input distribution if and only if

i) $\Phi(x) = \int_{-\infty}^{\infty} f_W(v-x) \ln f_V(v; F_X^\star) dv + H^\star \ge 0, \quad \forall x \in [-A, A]$
where $H^\star = C + h(W)$.

ii) The equality is achieved when $x \in E^\star$, with $E^\star$ being the set of points of increase for $F_X^\star$ defined as

$$E^\star = \left\{z \in \mathbb{R} : F_X^\star(z+r) > F_X^\star(z-r), \quad \forall r > 0\right\}. \tag{17}$$

The proof of this theorem is similar to the arguments in [31]. It is because $I(F_X)$ is concave and continuous, and we have the existence and finiteness of the directional derivative for $I(F_X)$ (See Lemma 7 in Appendix D).

In the following, we will show that the KTC function $\Phi(x)$ is analytic on the complex plane. To this end, we have the following lemmas regarding the analyticity of $f_W(v-s)$ and $\rho(s) = \int_{-\infty}^{\infty} f_W(v-s) \ln f_V(v; F_X) dv$ as functions of s.

Lemma 1: For any $v \in \mathbb{R}$, the function $f_W(v-s) : \mathbb{C} \to \mathbb{C}$ is analytic.

Proof: We will first demonstrate that $f_W(v-s)$ is continuous. Specifically, let $\{s_n\}_{n\in\mathbb{N}}$, be a sequence in $\mathbb{C}$ such that $\lim_{n\to\infty} s_n = s_*$ for $s_* \in \mathbb{C}$. From (8), and the fact that $f_Z(.)$ is continuous and analytic over $\mathbb{C}$, which results in $\lim_{n\to\infty} f_Z(v-s_n-u) = f_Z(v-s_*-u)$ for all $u, v \in \mathbb{R}$, the continuity of $f_W(v-s)$ is reduced to finding an integrable upper bound for $f_Z(v-s_n-u)$. Since s_n converges, there exists $b > 0$ such that $|s_n| \le b$ for any $n \in \mathbb{N}$. We then have:

$$\begin{aligned} &|f_Z(v-s_n-u)| \\ &\le \sum_{n=1}^{N} \frac{\varepsilon_n}{\sqrt{2\pi\sigma_n^2}} \left| \exp\left(-\frac{(v-s_n-u-a_n)^2}{2\sigma_n^2}\right)\right| \\ &= \sum_{n=1}^{N} \frac{\varepsilon_n \exp\left(\frac{[(\Im(s_n)]^2}{2\sigma_n^2}\right)}{\sqrt{2\pi\sigma_n^2}} \exp\left(-\frac{(v-\Re(s_n)-u-a_n)^2}{2\sigma_n^2}\right) \\ &\le e^{\frac{b^2}{2\sigma_N^2}} f_Z(v-\Re(s_n)-u) \quad (18) \\ &\le \frac{1}{\sqrt{2\pi\sigma_N^2}} e^{\frac{b^2}{2\sigma_N^2}}, \quad (19) \end{aligned}$$

where $\Re$ and $\Im$ denote the real and imaginary parts, respectively. Note that the inequality in (18) is obtained by using $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_N$ and the fact that $|s_n| \le b$. Then by using the Dominated Convergence Theorem, we have $\lim_{n\to\infty} f_W(v-s_n) = f_W(v-s_*)$. Equivalently, $s \to f_W(v-s)$ is continuous in $\mathbb{C}$. Now, let $\triangle$ be an arbitrary triangular path in $\mathbb{C}$. Then,

$$\left|\oint_\triangle f_W(v-s)\, ds\right| \le len(\triangle) \max_{s\in\triangle} |f_W(v-s)| < \infty, \tag{20}$$

where $len(\triangle)$ denotes the length of the path $\triangle$ and we have used the fact that $\max_{s\in\triangle} |f_W(v-s)|$ exists and is finite due to continuity of $f_W(v-s)$ on the compact set $\triangle$. As a result, we can use Fubini-Tonelli theorem [50] and exchange the integrals $\oint_\triangle$ and $\int_{-\infty}^{\infty}$ to obtain the following:

$$\oint_\triangle f_W(v-s)\, ds = \int_{-\infty}^{\infty} \oint_\triangle f_Z(v-s-u) ds dF_U(u). \tag{21}$$

As $f_Z(\cdot)$ is analytic everywhere, we have $\oint_\triangle f_Z(v-s-u)ds = 0$ from Cauchy's Integral theorem [51]. It then follows that $\oint_\triangle f_W(v-s)\, ds = 0$ for any triangular path $\triangle$. Then by Morera's Theorem [51], we can conclude that $f_W(v-s)$ is analytic everywhere. ■

Lemma 2: The function $\rho(s) : \mathbb{C} \to \mathbb{C}$ is analytic for any $F_X \in \mathcal{F}$.

Proof: We will first show that $\rho(s)$ is continuous. As before, let $\{s_n\}_{n\in\mathbb{N}}$, be a sequence in $\mathbb{C}$ such that $\lim_{n\to\infty} s_n = s_*$ for $s_* \in \mathbb{C}$. As we know that $f_W(v-s)$ is continuous, then by Generalized Lebesgue Dominated Convergence Theorem, it is enough to show that $\exists\{g_n(v)\}, g(v)$ such that

i) $|f_W(v-s_n) \ln f_V(v; F_X)| \le g_n(v)$.
ii) $g_n(v) \to g(v)$ almost everywhere.
iii) $\lim_{n\to\infty} \int_{-\infty}^{\infty} g_n(v)\, dv = \int_{-\infty}^{\infty} g(v)\, dv < \infty$.

It is clear that for a function $g(\alpha)$ such that $0 < g(\alpha) \le \beta < \infty$ for all $\alpha \in \mathbb{R}$ and some $\beta > 0$, we have $|\ln g(\alpha)| \le -\ln g(\alpha) + 2|\ln \beta|$. As $f_V(v; F_X^{(n)})$ is upper bounded by d, we have:

$$|\ln f_V(v; F_X^{(n)})| \le -\ln f_V(v; F_X^{(n)}) + 2|\ln d|. \tag{22}$$

Then using the upper bound on $-\ln f_V(v; F_X^{(n)})$ given in Appendix B, we end up with the following inequality:

$$|\ln f_V(v; F_X^{(n)})| \le av^2 - 2b_2 v + c_2 + 2|\ln d|. \quad (23)$$

As a result, using (18) and (23), we obtain the following bound:

$$\begin{aligned} &|f_W(v-s_n)\ln f_V(v, F_X)| \\ &\le e^{\frac{b^2}{2\sigma_N^2}} f_W\left(v - \Re\left(s_n\right)\right)\left(av^2 - 2b_2 v + c_2 + 2|\ln d|\right) \\ &= g_n\left(v\right). \end{aligned} \quad (24)$$

Furthermore, we have:

$$\begin{aligned} &\int_{\mathbb{R}} g_n\left(v\right) dv \\ &= e^{\frac{b^2}{2\sigma_N^2}} \int_{\mathbb{R}} f_W\left(v - \Re\left(s_n\right)\right)\left(av^2 - 2b_2 v + c_2 + 2|\ln d|\right) dv \\ &= e^{\frac{b^2}{2\sigma_N^2}} \left[a \int_{\mathbb{R}} v^2 f_W\left(v - \Re\left(s_n\right)\right) dv \right. \\ &\quad \left. - 2b_2 \int_{\mathbb{R}} v f_W\left(v - \Re\left(s_n\right)\right) dv + c_2 + 2\left|\ln d\right|\right]. \end{aligned} \quad (25)$$

Now by changing the variables as $t = v - \Re\left(s_n\right)$ we have

$$\begin{aligned} &\int_{\mathbb{R}} g_n\left(v\right) dv \\ &= e^{\frac{b^2}{2\sigma_N^2}} \left[a \int_{\mathbb{R}} t^2 f_W(t)dt + 2a\Re(s_n) \int_{\mathbb{R}} t f_W(t)dt + a\left(\Re(s_n)\right)^2 \right. \\ &\quad \left. - 2b_2 \int_{\mathbb{R}} t f_W\left(t\right) dt - 2b_2\Re\left(s_n\right) + c_2 + 2\left|\ln d\right|\right] \\ &= e^{\frac{b^2}{2\sigma_N^2}} \left[aE\left[W^2\right] + \left(2a\Re\left(s_n\right) - 2b_2\right) E\left[W\right] + a\left(\Re\left(s_n\right)\right)^2 \right. \\ &\quad \left. - 2b_2\Re\left(s_n\right) + c_2 + 2\left|\ln d\right|\right]. \end{aligned} \quad (26)$$

Since $E\left[W^2\right]$ and $E\left[W\right]$ are finite, and the fact that $|\Re\left(s_n\right)| \le b$, we have $\int_{\mathbb{R}} g_n\left(v\right) dv < \infty$. Moreover, as $\Re\left(s_n\right) \to \Re\left(s_*\right)$, it is straightforward to verify that $\lim_{n\to\infty} \int_{-\infty}^{\infty} g_n\left(v\right) dv = \int_{-\infty}^{\infty} g\left(v\right) dv$ based on equation (26), where $g\left(v\right)$ is defined as $g_n\left(v\right)$ with $s_\star$ replacing s_n. Then using the Generalized Dominated Convergence Theorem, we can conclude that $\rho(s)$ is continuous. Considering the analyticity of $f_W\left(v-s\right)$ and following a similar procedure as in the proof of Lemma 1 using Fubini-Tonelli, Cauchy's Integral, and Morera's theorems, it can be shown that $\rho\left(s\right)$ is analytic. ■

Combining Lemmas 1 and 2, it can be concluded that the KTC function $\Phi\left(s\right) : \mathbb{C} \to \mathbb{C}$ is analytic everywhere.

2) DISCRETENESS OF OPTIMAL INPUT

Given the analyticity of the KTC function $\Phi\left(s\right) : \mathbb{C} \to \mathbb{C}$, we then now shed light on the characterization of an optimal input $F_X^\star$. To this end, assume that the set $E^\star$ of $F_X^\star$ includes an infinite number of mass points on a bounded interval. This includes a continuous $F_X^\star$ as a special case. From Bolzano-Weierstrass Theorem [50], this set of mass points admits a limit point. Furthermore, following the Identity Theorem [52], we know that if two analytic functions are identical on an infinite set of points in a region along with their limit points, these two functions must be identical in the entire region. As a result, we obtain the extended KTC as follows:

$$\Phi(x) = \int_{-\infty}^{\infty} f_W(v-x) \ln f_V\left(v; F_X^\star\right) dv + H^\star = 0 \quad \forall x \in \mathbb{R}. \quad (27)$$

In the following, we will show that it is not possible to have (27). Consider the two constants k_0 and k_1 defined in the proof of Theorem 1. We can then select a constant m such that $m > A + k_0 + E\left[|W|\right]$ and $\ln\left(\frac{e^{H^*}k_1}{(m-A)^2}\right) < 0$. It then follows that for all $|v| > m$

$$\begin{aligned} \ln\left(e^{H^*} f_V\left(v; F_X^\star\right)\right) &\le \ln\left(\frac{e^{H^*}k_1}{\left(|v| - A\right)^2}\right) \\ &\le \ln\left(\frac{e^{H^*}k_1}{\left(m - A\right)^2}\right) < 0, \end{aligned} \quad (28)$$

where the first inequality is obtained from (13). Now, we rewrite the KTC in (27) as

$$\begin{aligned} &\int_{\Omega^+} f_W(v-x) \ln\left(e^{H^*} f_V\left(v; F_X^\star\right)\right) dv \\ &\quad = -\int_{\Omega^-} f_W(v-x) \ln\left(e^{H^*} f_V\left(v; F_X^\star\right)\right) dv, \end{aligned} \quad (29)$$

where

$$\Omega^+ = \left\{v \in \mathbb{R} : \ln\left(e^{H^*} f_V\left(v; F_X^\star\right)\right) \ge 0\right\}, \quad (30)$$

$$\Omega^- = \left\{v \in \mathbb{R} : \ln\left(e^{H^*} f_V\left(v; F_X^\star\right)\right) \le 0\right\}. \quad (31)$$

Now, let first examine the left-hand side of (29). From (28), we have $\Omega^+ \subset [-m, m]$. Thus,

$$\begin{aligned} 0 &\le \int_{\Omega^+} f_W(v-x) \ln\left(e^{H^*} f_V\left(v; F_X^\star\right)\right) dv \\ &\le \left[\max_{v\in\Omega^+} \ln\left(e^{H^*} f_V\left(v; F_X^\star\right)\right)\right] \int_{-m}^{m} f_W(v-x) dv. \end{aligned} \quad (32)$$

Because of the continuity of $\ln\left(e^{H^*} f_V\left(v; F_X^\star\right)\right)$, $\zeta = \max_{v\in\Omega^+} \ln\left(e^{H^*} f_V\left(v; F_X^\star\right)\right)$ exists and is finite. Hence, for $x > m + k_0$, we have

$$\begin{aligned} 0 &\le \int_{\Omega^+} f_W(v-x) \ln\left(e^{H^*} f_V\left(v; F_X^\star\right)\right) dv \\ &\le \zeta \int_{-m}^{m} f_W(v-x) dv \\ &\le 2m\zeta k_1 \left(x - m\right)^{-2}, \end{aligned} \quad (33)$$

where the last inequality follows from (12). Therefore, as $x \to \infty$, the left hand side of (29) goes to zero.

For the right-hand side of (29), we have

$$
\begin{aligned}
-\int_{\Omega^-} & f_W(v-x)\ln\left(e^{H^*} f_V\left(v; F_X^\star\right)\right) dv \\
&\geq -\int_m^{\infty} f_W(v-x)\ln\left(e^{H^*} f_V\left(v; F_X^\star\right)\right) dv \\
&\geq -\ln\left(\frac{e^{H^*}k_1}{(m-A)^2}\right)\int_m^{\infty} f_W(v-x)dv \\
&= -\ln\left(\frac{e^{H^*}k_1}{(m-A)^2}\right)\Pr\left(W > m-x\right). \qquad (34)
\end{aligned}
$$

Moreover, for all $x > 2m$, by using Markov's inequality, we have

$$
\Pr\left(W > m-x\right) \geq 1-\Pr\left(|W| > m\right) \geq 1-\frac{E\left[|W|\right]}{m} > 0. \qquad (35)
$$

By combining (35) and (34) we then have

$$
-\int_{\Omega^-} f_W(v-x)\ln\left(e^{H^*} f_V\left(v; F_X^\star\right)\right) dv > l, \quad \forall x > 2m, \qquad (36)
$$

where $l = -\ln\left(\frac{e^{H^*}k_1}{(m-A)^2}\right)\left(1-\frac{E[|W|]}{m}\right) > 0$. The inequalities in (33) and (36) therefore result in a contradiction. As a consequence, $E^\star$ consists of only a finite number of elements and, hence, any optimal $F_X^\star$ is discrete with a finite number of mass points.

C. NUMERICAL EXAMPLES

In this section, several numerical examples are provided to confirm the discreteness of the capacity-achieving input for the considered non-Gaussian channels. To find the optimal input numerically, we apply the well-known gradient descent-based method [31], [33]. In particular, starting with a single mass point distribution, we will identify the number of mass points, the locations of the mass points and their corresponding probabilities in an iterative manner. To guarantee the global optimality of the solution, the obtained distribution is verified with the necessary and sufficient condition given in Theorem 2.

We first consider a non-Gaussian channel impaired by a two-term GM noise Z with $f_Z(z) = 0.2\mathcal{N}(z, -4, 10) + 0.8\mathcal{N}(z, 1, 1)$ and a BPSK-like interference U with $f_U(u) = \frac{1}{2}\delta(u-2) + \frac{1}{2}\delta(u+2)$. It is not hard to verify that the total interference W is a four-term GM, and its PDF is shown in Fig. 2. The optimal input distributions at different signal-to-noise ratios (SNRs), (with SNR defined as SNR $= \frac{A^2}{E[W^2]}$), are plotted in Fig. 3. Observe from Fig. 3 that at sufficiently low SNRs, the optimal input has only two mass points. As the SNR increases, the number of mass points also increases. To verify the optimality of the solutions, we have also plotted the left hand side of KTC function at SNR = 1dB in Fig. 4. Clearly, the KTC is equal to zero at three mass points. It is also interesting to note that in this case, we always have transmission at the peak power level, i.e., there are two mass points at A and $-A$.

In the second example, we consider another non-Gaussian channel impaired by a GM noise Z with $f_Z(z) = 0.1\mathcal{N}(z, 0, 2) + 0.9\mathcal{N}(z, 3, 1)$, and an interference U that follows a Laplace distribution with $f_U(u) = \frac{1}{2}\exp(-|u|)$.

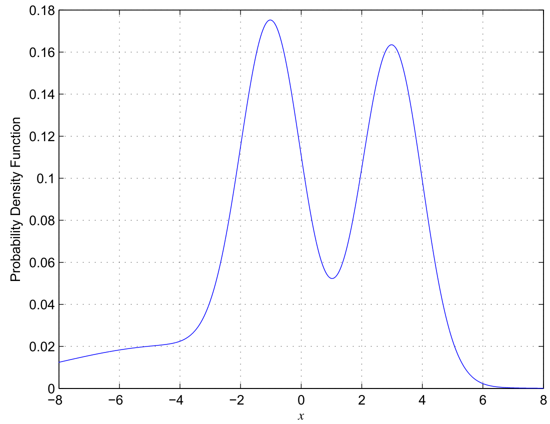

FIGURE 2. The probability distribution function of a four-term GM W.

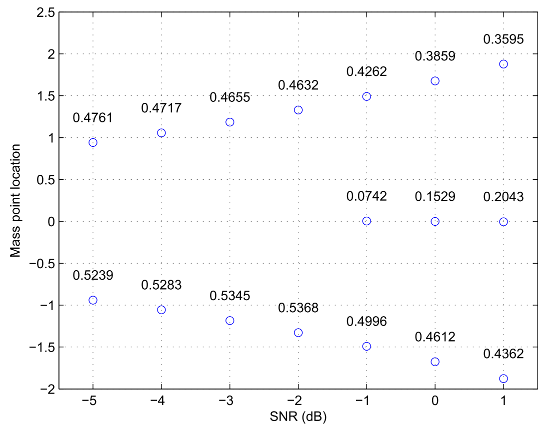

FIGURE 3. The location of mass points of the optimal input distribution and the corresponding probabilities with $f_Z(z) = 0.2\mathcal{N}(z, -4, 10) + 0.8\mathcal{N}(z, 1, 1)$ and $f_U(u) = \frac{1}{2}\delta(u-2) + \frac{1}{2}\delta(u+2)$.

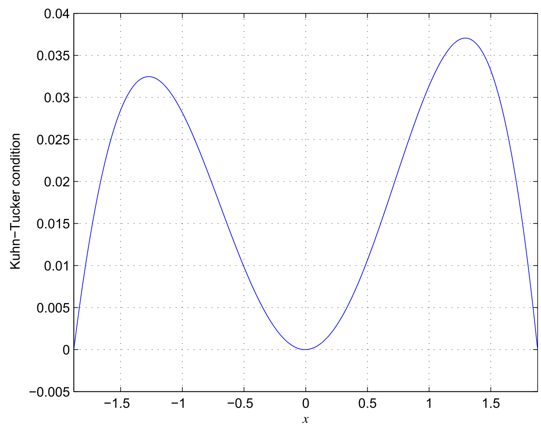

FIGURE 4. The KTC values for the input distribution $f_X(x) = 0.4362\,\delta(x+1.8775) + 0.2043\,\delta(x-0.004) + 0.3595\,\delta(x-1.8775)$ at SNR=1dB.

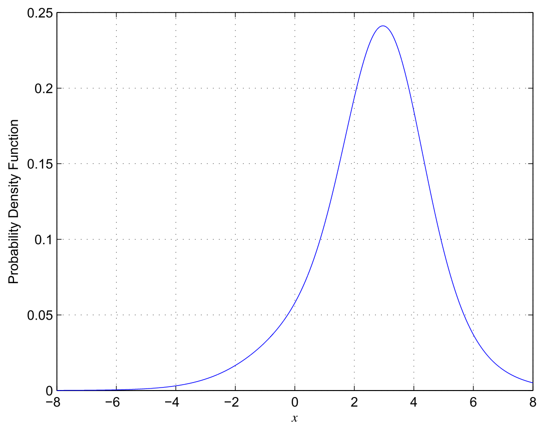

FIGURE 5. The probability distribution function of W resulting from Laplace distributed U.

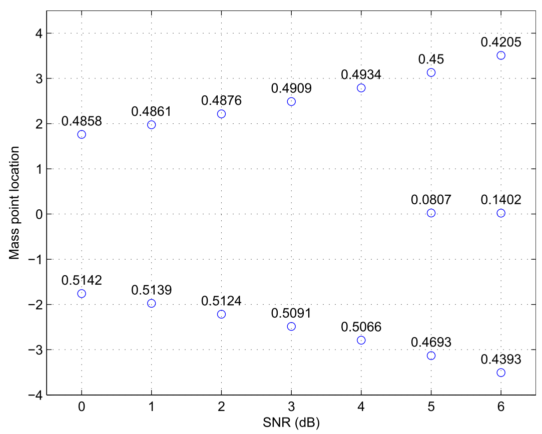

FIGURE 6. The location of mass points of the optimal input distribution and the corresponding probabilities with $f_Z(z) = 0.1\mathcal{N}(z,0,2) + 0.9\mathcal{N}(z,3,1)$ and $f_U(u) = \frac{1}{2}\exp(-|u|)$.

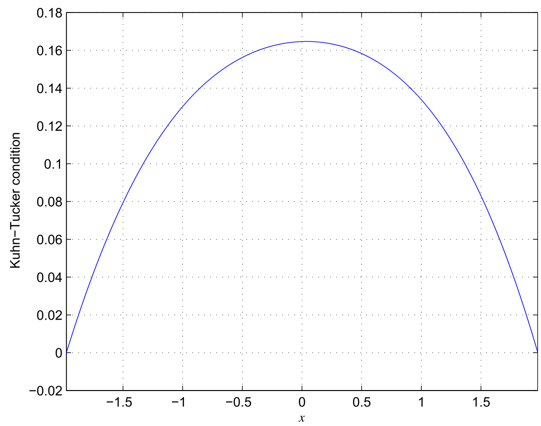

FIGURE 7. The KTC value for the input distribution $f_X(x) = 0.5139\,\delta(x+1.9755) + 0.4861\,\delta(x-1.9755)$ at SNR=1dB.

In this case, W is no longer a GM, and its PDF is given in Fig. 5. The optimal inputs for this channel are shown in Fig. 6 at different SNRs. We also plot the corresponding KTC at SNR $=$ 1dB. As similar to the previous example, at a given SNR, it is clear that the optimal input is discrete.

IV. EXTENSION TO MULTI-USER CHANNELS UNDER GM NOISE

In this section, we address the detailed characterization of optimal inputs for an M-user MAC with GM noise under the peak-power constraints in (6). While we follow closely the arguments in [40] for a Gaussian MAC, the results in this section mainly rely on the arbitrary property of U and the results established in Section III. Specifically, we first show that the sum-capacity of the channels can only be achieved if each user uses a discrete distribution with a finite number of mass points. An interesting important property of the rate region is also given.

Now, for the considered MAC, by following a similar analysis as in Section III, it can be shown that a set of optimal input distributions $\{F^{\star}_{X_i}\}$ that achieves C_{sum} always exists. However, such a set is not necessarily unique. For a given optimal set $\{F^{\star}_{X_i}\}$, it is clear that:

$$\{F^{\star}_{X_i}\} = \arg \sup_{F_{X_i}\in\mathcal{F}_i, i=1\cdots M} h\left(\sum_{i=1}^{M} X_i + Z\right). \tag{37}$$

To shed light on the characteristics of $\{F^{\star}_{X_i}\}$, $1 \le i \le M$, consider a point-to-point link with an input X_i subject to the peak constraint in the feasible set $\mathcal{F}_i$ and the output V impaired by two types of noise, the GM noise Z and the interference $U = \sum_{j=1, j\neq i}^{M} X^{\star}_j$. Here, each $X^{\star}_j, j \in \{1,\ldots,M\}\backslash\{i\}$, follows the distribution $F^{\star}_{X_j}$, which is unknown yet fixed. Let F'_{X_i} be an optimal input distribution of this point-to-point channel, i.e.,

$$F'_{X_i} = \arg \sup_{F_{X_i}\in\mathcal{F}_i} I\left(F_{X_i}\right) = \arg \sup_{F_{X_i}\in\mathcal{F}_i} I\left(X_i; X_i + U + Z\right). \tag{38}$$

We then have the following lemma regarding the characteristics of F'_{X_i}.

Lemma 3: F'_{X_i} exists, and it is discrete and unique.

Proof: Since each $F^{\star}_{X_i}$, $1 \le i \le M$, is peak constrained, the distribution of U is fixed but unknown with a finite second-order moment. Therefore, it follows from the results for the point-to-point channel that F'_{X_i} exists, and it is discrete.

Next, it can be verified that the moment-generating function (MGF) of $U = \sum_{j=1, j\neq i}^{M} X^{\star}_j$ exists on an interval around 0. As a result, the MGF of $W = U + Z$ denoted as $M_W(t)$ exists in the range $|t| < t_0$ for a positive constant t_0. As we show in Appendix C, when extended to the complex plane, the function $M_W(s) : \mathcal{D} \to \mathbb{C}$ is analytic on $\mathcal{D}$, where $\mathcal{D} = \{s \in \mathbb{C} : |\Re(s)| < t_0\}$. As a result, $M_W(s)$ has isolated zeros on $\mathcal{D}$. Since the characteristic function (CF) $\phi_W(t)$ of W is its MGF evaluated along the imaginary axis, i.e., $\phi_W(t) = M_W(jt)$, $\phi_W(t)$ is analytic on the real line and has isolated zeros. Now, besides F'_{X_i}, assume that F''_{X_i} is another optimal distribution. It means both F'_{X_i} and F''_{X_i} maximize the output entropy $h(V;\cdot)$. Since the output entropy

$h(\cdot)$ is strictly concave in $f_V(\cdot)$, we then have $f_V\left(v, F'_{X_i}\right) = f_V\left(v, F''_{X_i}\right)$. Therefore,

$$\phi_W(t)\left(\phi'_{X_i}(t) - \phi''_{X_i}(t)\right) = 0, \quad \forall t \in \mathbb{R}, \tag{39}$$

where $\phi(.)$ denotes the CFs of the corresponding random variables. Because zeros of $\phi_W(.)$ are isolated, and a CF is uniformly continuous on the entire real line [53], it is clear that $\phi'_{X_i}(t) - \phi''_{X_i}(t) = 0$, $\forall\, t \in \mathbb{R}$. Therefore, $F'_{X_i} = F''_{X_i}$. ■

Given the result in Lemma 3, a similar argument as in [40] can be used to addresses the characteristics of the optimal input distributions $\{F^\star_{X_i}\}$. The result is stated in the following proposition.

Proposition 3: For any set of optimal distributions $\{F^\star_{X_i}\}$ *in (37),* $F^\star_{X_i} = F'_{X_i}$ *where* F'_{X_i}, $1 \le i \le M$, *is the optimal input defined in (38). As a result, each* $F^\star_{X_i}$ *is discrete with a finite number of mass points.*

Proof: The proof follows the method used in [40]. In particular, from Lemma 3, we know that F'_{X_i} is unique and discrete, having a finite number of mass points. Furthermore, it is clear from (38) that $F'_{X_i} = \arg \sup_{F_{X_i} \in \mathcal{F}_i} h(X_i + U + Z)$. It then follows that $h\left(X'_i + \sum_{j=1, j\neq i}^{M} X^\star_j + Z\right) \ge h\left(\sum_{j=1}^{M} X^\star_j + Z\right)$, where X'_i follows the distribution F'_{X_i}. On the other hand, since the set $\{F^\star_{X_i}\}$ achieves the sum-rate, we have from (37) that $h\left(X'_i + \sum_{j=1, j\neq i}^{M} X^\star_j + Z\right) \le h\left(\sum_{j=1}^{M} X^\star_j + Z\right)$. Thus, $h\left(X'_i + \sum_{j=1, j\neq i}^{M} X^\star_j + Z\right) = h\left(\sum_{j=1}^{M} X^\star_j + Z\right)$. Equivalently, $F^\star_{X_i}$ is also optimal for the point-to-point channel. Therefore, $F^\star_{X_i} = F'_{X_i}$, and each $F^\star_{X_i}$ is discrete with a finite support set. ■

While the sum-capacity is one of the most important benchmarks, it is also of interest to understand the characteristics of the rate region of the considered MAC. By following a similar analysis as in [40], we can also show that for the considered MAC, there exist at least two distinct points achieving the sum capacity on the rate region. For completeness, the proof of this result is given in Appendix E. Finally, it is worth mentioning that we can use time-sharing arguments to show that there exists a linear segment between these two points in which any point belonging to the segment is also sum-capacity achieving. As a result, there exists an infinite number of sum-capacity achieving points.

Before closing this section, we should emphasize that it is certainly of great interest to obtain some numerical results to confirm the finiteness and discreteness of the optimal inputs for the case of multiple-access channels. Towards this end, an effective numerical method to accurately calculate the optimal mass points and their corresponding probabilities is required. The development of such a method is, however, non-trivial. It is because for the considered multiple-access channel, we need to deal with the optimization of multiple input distributions simultaneously. It should be noted that even for the single-user cases, there exist several drawbacks of the well-known gradient decent-based method we adopted earlier to find the optimal input [33], [36]. It is due to the relatively small sensitivity of the MI to the number of mass points as well as their locations and probabilities used in each iteration. As a consequence, it is difficult to find the optimal input consisting of more mass points, some of which having low probabilities, with high accuracy. Dealing with multiple inputs, and at the same time, paying attention to the convergence behavior of the solution are therefore more challenging. Given that, we believe the investigation on such new numerical methods is beyond the scope of the current work, and the topic deserves further studies.

V. CONCLUSION

This paper has proved the existence and discreteness of the capacity-achieving input signals for general class of point-to-point and multiple access channels with additive non-Gaussian noise under peak-power constraints. In particular, the considered non-Gaussian link consists of a Gaussian mixture noise having Gaussian elements with arbitrary means, and an arbitrary interference U having finite second order moment. The novelty of the work lies in the establishment of the necessary and sufficient condition for an input signal to be optimal and the use of Fubini-Tonelli's and Morera's theorems to show the analyticity of this condition. The continuity of the optimal input is then ruled out by proving that the optimal support set admits no limit point. Taking these into account, it is concluded that the capacity-achieving input is discrete with a finite number of mass points. We also exploited the arbitrary property of U to show that the optimal input distributions that achieve the sum-capacity of an M-user multiple access channel under GM noise are discrete. In addition, there exist at least two distinct points that achieve the sum capacity on the rate region. The point-to-point and multiple-user channels considered in this paper are general enough to represent all non-Gaussian additive channels of engineering interest.

APPENDIX A
TWO UPPER BOUNDS AND A LOWER BOUND ON $f_W(\cdot)$

In this section, upper and lower bounds on the distributions $f_W(\cdot)$, which are useful for the developments in Sections III-A and III-B, are derived. The bounds are given in the following two lemmas.

Lemma 4: The distribution $f_W(\cdot)$ *can be upper-bounded by a constant* d *as* $f_W(w) \le d$. *Furthermore, there exist two positive constants* k_0 *and* k_1 *such that* $f_W(w) \le k_1\, w^{-2}$ *when* $|w| > k_0$.

Proof: Since $W = U + Z$, we have

$$\begin{aligned} f_W(w) &= \int_{-\infty}^{\infty} f_Z(w-u)\,dF_U(u) \\ &= \sum_{n=1}^{N} \frac{\varepsilon_n}{\sqrt{2\pi\sigma_n^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{(w-a_n-u)^2}{2\sigma_n^2}\right) dF_U(u) \\ &\le \sum_{n=1}^{N} \frac{\varepsilon_n}{\sqrt{2\pi\sigma_n^2}} = d \end{aligned} \tag{40}$$

Furthermore, as Z is a GM, it can be verified that there are positive constants α_0, α_1 and α_2 such that $f_Z(z) \leq \alpha_1 e^{-\alpha_2|z|}$ for all $|z| > \alpha_0$. As a consequence, there exist finite constants $\alpha_3 > 0$ and $\alpha_4 \geq \alpha_0$ such that $f_Z(z) \leq \frac{\alpha_3}{z^2}$ for all $|z| > \alpha_4$. Then when $w > 2\alpha_4$ we have:

$$\begin{aligned} f_W(w) &= \int_{-\infty}^{w/2} f_Z(w-u)dF_U(u) + \int_{w/2}^{\infty} f_Z(w-u)dF_U(u) \\ &\leq \int_{-\infty}^{w/2} \frac{\alpha_3}{(w-u)^2} dF_U(u) + c\Pr(U \geq w/2) \quad (41) \\ &\leq \frac{\alpha_3}{(w-w/2)^2} + \frac{4cE\left[U^2\right]}{w^2} \quad (42) \\ &= \frac{4\left(\alpha_3 + cE\left[U^2\right]\right)}{w^2} = T_u(w). \quad (43) \end{aligned}$$

where (41) comes from the fact that the PDF of any GM random variable is upper bounded by any constant c that is great than d, and (42) follows from Markov's inequality. By using the same procedure, we can also obtain the same upper bound when $w < -2\alpha_4$. Therefore, by choosing $k_0 = 2\alpha_4$ and $k_1 = 4\left(\alpha_3 + cE\left[U^2\right]\right)$, the lemma is proved. ■

Now, a lower bound on $f_W(\cdot)$ is stated in the following.

Lemma 5: The distribution $f_W(\cdot)$ can be lower bounded as

$$T_l(w) = \frac{e^{-c_1}}{e^{(aw^2 - 2b_1 w)}} \leq f_W(w) \quad (44)$$

for finite constants a, b_1, c_1.

Proof: We know that the function $-\ln(x)$ is convex for $x \in (0, \infty)$. It then follows that:

$$\begin{aligned} &-\ln\left(f_W(w)\right) \\ &= -\ln\left[\sum_{n=1}^{N} \varepsilon_n \int_{-\infty}^{\infty} \frac{\exp\left(-\frac{(w-a_n-u)^2}{2\sigma_n^2}\right)}{\sqrt{2\pi\sigma_n^2}} dF_U(u)\right] \\ &\leq \sum_{n=1}^{N} \varepsilon_n \left(-\ln\left[\int_{-\infty}^{\infty} \frac{\exp\left(-\frac{(w-a_n-u)^2}{2\sigma_n^2}\right)}{\sqrt{2\pi\sigma_n^2}} dF_U(u)\right]\right) \quad (45) \\ &\leq \sum_{n=1}^{N} \varepsilon_n \int_{-\infty}^{\infty} -\ln\left[\frac{\exp\left(-\frac{(w-a_n-u)^2}{2\sigma_n^2}\right)}{\sqrt{2\pi\sigma_n^2}}\right] dF_U(u) \quad (46) \\ &= \sum_{n=1}^{N} \frac{\varepsilon_n}{2\sigma_n^2} \int_{-\infty}^{\infty} \left[w^2 - 2w(u+a_n) + (u+a_n)^2\right] dF_U(u) \\ &\quad + \sum_{n=1}^{N} \varepsilon_n \ln\left(\sqrt{2\pi\sigma_n^2}\right) \quad (47) \\ &\leq \sum_{n=1}^{N} \frac{\varepsilon_n}{2\sigma_n^2} \int_{-\infty}^{\infty} \left[w^2 - 2w(u+a_n) + 2\left(u^2 + a_n^2\right)\right] dF_U(u) \\ &\quad + \sum_{n=1}^{N} \varepsilon_n \ln\left(\sqrt{2\pi\sigma_n^2}\right) \quad (48) \\ &= aw^2 - 2b_1 w + c_1 \quad (49) \end{aligned}$$

where

$$a = \sum_{n=1}^{N} \frac{\varepsilon_n}{2\sigma_n^2}, \quad (50)$$

$$b_1 = \sum_{n=1}^{N} \frac{\varepsilon_n}{2\sigma_n^2}\left(a_n + E[U]\right), \quad (51)$$

$$c_1 = \sum_{n=1}^{N} \varepsilon_n \ln\left(\sqrt{2\pi\sigma_n^2}\right) + \sum_{n=1}^{N} \frac{\varepsilon_n}{\sigma_n^2}\left(a_n^2 + E\left[U^2\right]\right). \quad (52)$$

Note that the inequality in (45) follows from the convexity of $-\ln(x)$, while we obtain (46) using Jensen inequality and (48) using the fact that $(\alpha+\beta)^2 \leq 2\left(\alpha^2+\beta^2\right)$. It is clear to see that the above upper bound for $-\ln\left(f_W(w)\right)$ is equivalent to the lower bound $f_W(w) \geq \frac{e^{-c_1}}{e^{\left(aw^2 - 2b_1 w\right)}}$. ■

It is worth mentioning that the two bounds in (43) and (44) do not satisfy the two conditions used in [30]. It is because $L(w) = \ln\left[\frac{1}{T_l(w)}\right]$ is not non-increasing for $w < 0$, and we also have $\int_{-\infty}^{\infty} T_u(w) \ln\left[T_l(w)\right] dw = -\infty$.

APPENDIX B
BOUNDS ON $f_V(\cdot)$

Lemma 6: The distribution $f_V(\cdot)$ for a given input distribution F_X can be bounded as

$$\frac{e^{-c_2}}{e^{(av^2 - 2b_2 v)}} \leq f_V(v; F_X) \leq d \quad (53)$$

for finite constants a, b_2, c_2 and d.

Proof: The upper bound on $f_V(\cdot)$ comes directly from that of $f_W(\cdot)$, which is:

$$\begin{aligned} f_V(v; F_X) &= \int_{-\infty}^{\infty} f_W(v-x)dF_X(x) \\ &\leq \sum_{n=1}^{N} \frac{\varepsilon_n}{\sqrt{2\pi\sigma_n^2}} = d \quad (54) \end{aligned}$$

For the lower bound, we have:

$$\begin{aligned} &-\ln\left(f_V(v; F_X)\right) \\ &= -\ln\left[\int_{-\infty}^{\infty} f_W(v-x)dF_X(x)\right] \\ &\leq \int_{-\infty}^{\infty} -\ln\left[f_W(v-x)\right] dF_X(x) \quad (55) \\ &\leq \int_{-\infty}^{\infty} \left[a(v-x)^2 - 2b_1(v-x) + c_1\right] dF_X(x) \quad (56) \\ &= av^2 - 2b_2 v + c_2, \quad (57) \end{aligned}$$

where

$$b_2 = b_1 + aE[X], \quad (58)$$

$$c_2 = c_1 + aE\left[X^2\right] + 2b_1 E[X]. \quad (59)$$

Note that for the above, we apply Jensen inequality in (55). In addition, (56) comes from the upper bound on

$-\ln(f_W(w))$. The lower bound on $f_V(v; F_X)$ therefore follows directly from the upper bound on $-\ln(f_V(v; F_X))$ in (57). ■

APPENDIX C
ANALYTICITY OF $M_W(s)$

Because $M_W(t)$ is finite for all $|t| < t_0$, we have

$$|M_W(s)| \leq \int_{-\infty}^{\infty} \left|e^{sw}\right| dF_W(w) = \int_{-\infty}^{\infty} e^{\Re(s)w} dF_W(w) < \infty, \quad (60)$$

for any $s \in \mathcal{D}$. Therefore, $M_W(s)$ exists on $\mathcal{D}$. Let $\{s_n\}_{n\in\mathbb{N}}$, be a sequence in $\mathcal{D}$ such that $\lim_{n\to\infty} s_n = s_*$ for $s_* \in \mathcal{D}$. Then there exists a constant $b > 0$ such that $|\Re(s_n)| \leq b < t_0$ for any $n \in \mathbb{N}$. As such, for any $w \in \mathbb{R}$, we have $|e^{s_n w}| = e^{\Re(s_n)w} \leq e^{bw}$. Moreover, $\int_{\mathbb{R}} e^{bw} dF_W(w) < \infty$ since $b < t_0$ and $\int_{\mathbb{R}} e^{tw} dF_W(w) < \infty$ for all $|t| < t_0$. So by applying Dominated Convergence Theorem, we have $\lim_{n\to\infty} M_W(s_n) = M_W(s_*)$. Equivalently, $M_W(s)$ is continuous on $\mathcal{D}$. Furthermore, because of the analyticity of e^{sw}, we can apply the same procedure as in the proof of Lemma 1 using Fubini-Tonelli, Cauchy's Integral, and Morera's theorems to show the analyticity of $M_W(s)$.

APPENDIX D
DIRECTIONAL DERIVATIVE AND LAGRANGIAN THEOREM

This appendix provides some well-known concepts, which are helpful for the establishment of the KTC.

Definition 1: Let $f : \mathcal{V} \to \mathbb{R}$ be a function on a normed linear space $\mathcal{V}$. The directional derivative of f at $a \in \mathcal{V}$ along a direction $d \in \mathcal{V}$ is defied by

$$D_a(f; d) = \lim_{t\to 0} \frac{1}{t}(f(a + td) - f(a)),$$

if this limit exists.

Lemma 7: Let $f : \mathcal{A} \to \mathbb{R}$ be a concave function where $\mathcal{A}$ is a convex set. For any $a \in int(\mathcal{A})$ and for any $a' \in \mathcal{A}$, the directional derivative of f along $a' - a$

$$D_a(f; a' - a) = \lim_{t\to 0} \frac{1}{t}(f((1 - t)a + ta') - f(a))$$

exists and is finite. Moreover, a is a point of global maximum for f if and only if $D_a(f; a' - a) \leq 0$ for any $a' \in \mathcal{A}$.

Proof: The proof is given in [54]. ■

Theorem 3 (Optimization Theorem): Let Ω be a compact and convex metric space, and f a continuous, weakly differentiable and convex-cap functionals on Ω to $\mathbb{R}$. Define:

$$C = \sup_{x\in\Omega} f(x).$$

Then

1) *$C = \max_{x\in\Omega} f(x)$; i.e., $f(x)$ achieves its maximum on Ω.*
2) *The necessary and sufficient condition for $f(x_0) = C$ is $D_{x_0}(f; x - x_0) \leq 0$ for all $x \in \Omega$.*
3) *If f is strictly convex-cap, C is achieved by a unique x_0 in Ω.*

Proof: See [47], [55]. ■

APPENDIX E
THE PROOF OF TWO DISTINCT POINTS ACHIEVING THE SUM CAPACITY

Let $\{F^\star_{X_i}\}$ be a given set of sum-capacity-achieving distributions. If there exists only a single point $\mathcal{S}$ in the rate region that achieves the sum capacity, it is clear that

$$C_{\text{sum}} = \sum_{i=1}^{M} I\left(F^\star_{X_i} \mid \left\{X^\star_j\right\}_{j=1, j\neq i}^{M}\right). \quad (61)$$

On the other hand, based on the chain rule of MI, we have

$$C_{\text{sum}} = I\left(\{F^\star_{X_j}\}_{j=1}^{M}\right) = \sum_{i=1}^{M} I\left(F^\star_{X_i} \mid \left\{X^\star_j\right\}_{j=1}^{i-1}\right). \quad (62)$$

Let $\boldsymbol{Q}$ and $\boldsymbol{Q}'$ denote the sets $\left\{X^\star_j\right\}_{j=i+1}^{M}$ and $\left\{X^\star_j\right\}_{j=1}^{i-1}$, respectively and $\mathcal{X}$ represent the support set of $\boldsymbol{Q}$. Since $\{F^\star_{X_i}\}$ are finite and discrete, $\mathcal{X}$ has finite number of elements. It then follows that:

$$\begin{aligned}
&I\left(F^\star_{X_i} \mid \left\{X^\star_j\right\}_{j=1, j\neq i}^{M}\right) \\
&= \sum_{\boldsymbol{q}\in\mathcal{X}} \Pr(\boldsymbol{Q} = \boldsymbol{q})\, I\left(F^\star_{X_i} \mid \boldsymbol{Q} = \boldsymbol{q}, \boldsymbol{Q}'\right) \\
&= \sum_{\boldsymbol{q}\in\mathcal{X}} \Pr(\boldsymbol{Q} = \boldsymbol{q}) \\
&\quad \times D\left(f_{X^\star_i, Y^\star|\boldsymbol{Q},\boldsymbol{Q}'}(., . \mid \boldsymbol{q}, .) \,\|\, f_{X^\star_i}(.) f_{Y^\star|\boldsymbol{Q},\boldsymbol{Q}'}(. \mid \boldsymbol{q}, .)\right) \\
&\geq D\left(\sum_{\boldsymbol{q}\in\mathcal{X}} \Pr(\boldsymbol{Q} = \boldsymbol{q}) f_{X^\star_i, Y^\star|\boldsymbol{Q},\boldsymbol{Q}'}(., . \mid \boldsymbol{q}, .)\right. \\
&\qquad \left.\| \sum_{\boldsymbol{q}\in\mathcal{X}} \Pr(\boldsymbol{Q} = \boldsymbol{q}) f_{X^\star_i}(.) f_{Y^\star|\boldsymbol{Q},\boldsymbol{Q}'}(. \mid \boldsymbol{q}, .)\right) \\
&= I\left(F^\star_{X_i} \mid \boldsymbol{Q}'\right) = I\left(F^\star_{X_i} \mid \left\{X^\star_j\right\}_{j=1}^{i-1}\right). \quad (63)
\end{aligned}$$

where $D(.)$ is the relative entropy between two distributions, $Y^\star = \sum_{i=1}^{M} X^\star_i + Z$ and the inequality comes from the log-sum inequality. It can be verified that the equality in (63) can only be achieved if and only if $f_{Y^\star|X^\star_i,\boldsymbol{Q},\boldsymbol{Q}'}(. \mid ., ., .) = f_{Y^\star|\boldsymbol{Q},\boldsymbol{Q}'}(. \mid ., .)$, which is equivalent to $E\left[Y^\star \mid X^\star_i = x, \boldsymbol{Q}, \boldsymbol{Q}'\right] = E\left[Y^\star \mid \boldsymbol{Q}, \boldsymbol{Q}'\right]$ for any mass point x in the support set of $X^\star_i$. Therefore, the equality in (63) only happens if $E\left[X^\star_i\right] = x$ for any mass point x in the support set of $X^\star_i$, or equivalently, $X^\star_i$ has only a single mass point, which is not possible. As a result, $I\left(\{F^\star_{X_j}\}_{j=1}^{M}\right) < \sum_{i=1}^{M} I\left(F^\star_{X_i} \mid \left\{X^\star_j\right\}_{j=1, j\neq i}^{M}\right)$, which contradicts with (61).

ACKNOWLEDGMENT

This paper was presented in part at the IEEE International Conference on Communications, Kansas, USA, May 2018 [1].

REFERENCES

[1] M. Ranjbar, N. H. Tran, T. Nguyen, and M. C. Gursoy, "Optimal inputs of single-user and multi-user non-Gaussian aggregate interference channels," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–6.

[2] T. Q. S. Quek, G. de la Roche, I. Güvenç, and M. Kountouris, *Small Cell Networks: Deployment, PHY Techniques, and Resource Management*, 1st ed. Cambridge, U.K.: Cambridge Univ. Press, 2013.

[3] A. K. Mishra and D. L. Johnson, *White Space Communication: Advances, Developments and Engineering Challenges*, 1st ed. New York, NY, USA: Springer, 2014.

[4] J. Lin, M. Nassar, and B. L. Evans, "Impulsive noise mitigation in powerline communications using sparse Bayesian learning," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, pp. 1172–1183, Jul. 2013.

[5] G. Ozcan, M. C. Gursoy, and S. Gezici, "Error rate analysis of cognitive radio transmissions with imperfect channel sensing," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1642–1655, Mar. 2014.

[6] D. J. Jakubisin and R. M. Buehrer, "Approximate joint MAP detection of co-channel signals in non-Gaussian noise," *IEEE Trans. Commun.*, vol. 64, no. 10, pp. 4224–4237, Oct. 2016.

[7] G. Ozcan, M. C. Gursoy, N. Tran, and J. Tang, "Energy-efficient power allocation in cognitive radio systems with imperfect spectrum sensing," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3466–3481, Dec. 2016.

[8] T. Bai *et al.*, "Discrete multi-tone digital subscriber loop performance in the face of impulsive noise," *IEEE Access*, vol. 5, pp. 10478–10495, 2017.

[9] T. Y. Al-Naffouri, A. A. Quadeer, and G. Caire, "Impulsive noise estimation and cancellation in DSL using orthogonal clustering," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul./Aug. 2011, pp. 2841–2845.

[10] R. Pighi, M. Franceschini, G. Ferrari, and R. Raheli, "Fundamental performance limits of communications systems impaired by impulse noise," *IEEE Trans. Commun.*, vol. 57, no. 1, pp. 171–182, Jan. 2009.

[11] C. Soltanpur, K. M. Rabie, B. Adebisi, and A. Wells, "Masreliez-equalized VOFDM in non-Gaussian channels: Power line communication systems," *IEEE Syst. J.*, vol. 11, no. 1, pp. 1–9, Jan. 2017.

[12] M. Korki, J. Zhang, C. Zhang, and H. Zayyani, "Block-Sparse impulsive noise reduction in OFDM systems—A novel iterative Bayesian approach," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 271–284, Jan. 2016.

[13] F. Chiti, R. Fantacci, D. Marabissi, and A. Tani, "Performance evaluation of an efficient and reliable multicast power line communication system," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 7, pp. 1953–1964, Jul. 2016.

[14] H. Zhang, L.-L. Yang, and L. Hanzo, "Compressed impairment sensing-assisted and interleaved-double-FFT-aided modulation improves broadband power line communications subjected to asynchronous impulsive noise," *IEEE Access*, vol. 4, pp. 81–96, 2016.

[15] X. Kuai, H. Sun, S. Zhou, and E. Cheng, "Impulsive noise mitigation in underwater acoustic OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8190–8202, Oct. 2016.

[16] D. W. Stein, "Statistical characteristics of moving acoustic sources in ocean waveguides," *J. Acoust. Soc. Amer.*, vol. 98, no. 3, pp. 1486–1496, Sep. 1995.

[17] Z. Wang, S. Zhou, J. Catipovic, and P. Willett, "Parameterized cancellation of partial-band partial-block-duration interference for underwater acoustic OFDM," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1782–1795, Apr. 2012.

[18] K. Gulati, B. L. Evans, J. G. Andrews, and K. R. Tinsley, "Statistics of co-channel interference in a field of poisson and poisson-poisson clustered interferers," *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 6207–6222, Dec. 2010.

[19] S. Godsill, P. Rayner, and O. Cappé, "Digital audio restoration," in *Applications of Digital Signal Processing to Audio and Acoustics* (The International Series in Engineering and Computer Science), vol. 437. Boston, MA, USA: Springer, Apr. 2002, pp. 133–194.

[20] S. W. Perry, H.-S. Wong, and L. Guan, *Adaptive Image Processing: A Computational Intelligence Perspective*. Bellingham, WA, USA: SPIE, 2001.

[21] R. Garnett, T. Huegerich, C. Chui, and W. He, "A universal noise removal algorithm with an impulse detector," *IEEE Trans. Image Process.*, vol. 14, no. 11, pp. 1747–1754, Nov. 2005.

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

[23] J. Fahs, N. Ajeeb, and I. Abou-Faycal, "The capacity of average power constrained additive non-Gaussian noise channels," in *Proc. Int. Conf. Telecommun. (ICT)*, Apr. 2012, pp. 1–6.

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

[25] W. Oettli, "Capacity-achieving input distributions for some amplitude-limited channels with additive noise (Corresp.)," *IEEE Trans. Inf. Theory*, vol. 20, no. 3, pp. 372–374, May 1974.

[26] J. Cao, S. Hranilovic, and J. Chen, "Capacity-achieving distributions for the discrete-time poisson channel—Part I: General properties and numerical techniques," *IEEE Trans. Commun.*, vol. 62, no. 1, pp. 194–202, Jan. 2014.

[27] H. V. Vu, N. H. Tran, M. C. Gursoy, T. Le-Ngoc, and S. Hariharan, "Capacity-achieving input distributions of additive quadrature Gaussian mixture noise channels," *IEEE Trans. Commun.*, vol. 63, no. 10, pp. 3607–3620, Oct. 2015.

[28] M. Ranjbar, N. H. Tran, M. C. Gursoy, and H. R. Bahrami, "Energy efficiency of channels under additive Gaussian-mixture noise in the low-power regime," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.

[29] A. Dytso, R. Bustin, H. V. Poor, and S. Shitz (Shitz), "On additive channels with generalized Gaussian noise," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 426–430.

[30] J. Fahs and I. Abou-Faycal, "On properties of the support of capacity-achieving distributions for additive noise channel models with input cost constraints," *IEEE Trans. Inf. Theory*, vol. 64, no. 2, pp. 1178–1198, Feb. 2018.

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

[32] 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.

[33] 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.

[34] M. C. Gursoy, H. V. Poor, and S. Verdú, "Noncoherent Rician fading channel—Part II: Spectral efficiency in the low-power regime," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2207–2221, Sep. 2005.

[35] M. C. Gursoy, "On the capacity and energy efficiency of training-based transmissions over fading channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4543–4567, Oct. 2009.

[36] J. Huang and S. P. Meyn, "Characterization and computation of optimal distributions for channel coding," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2336–2351, Jul. 2005.

[37] 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.

[38] L. Zhang, H. Li, and D. Guo, "Capacity of Gaussian channels with duty cycle and power constraints," *IEEE Trans. Inf. Theory*, vol. 60, no. 3, pp. 1615–1629, Mar. 2014.

[39] M. Egan, S. M. Perlaza, and V. Kungurtsev, "Capacity sensitivity in additive non-Gaussian noise channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 416–420.

[40] 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.

[41] D.-A. Le, H. V. Vu, N. H. Tran, M. C. Gursoy, and T. Le-Ngoc, "Approximation of achievable rates in additive Gaussian mixture noise channels," *IEEE Trans. Commun.*, vol. 64, no. 12, pp. 5011–5024, Dec. 2016.

[42] S. Stergiopoulos, *Advanced Signal Processing Handbook: Theory and Implementation for Radar, Sonar, and Medical Imaging Real Time Systems*. Boca Raton, FL, USA: CRC Press, 2001.

[43] M. F. Huber, T. Bailey, H. Durrant-Whyte, and U. D. Hanebeck, "On entropy approximation for Gaussian mixture random vectors," in *Proc. IEEE Conf. Multisensor Fusion Integration Intell. Syst.*, Aug. 2008, pp. 181–188.

[44] V. Maz'ya and G. Schmidt, "On approximate approximations using Gaussian kernels," *IMA J. Numer. Anal.*, vol. 16, no. 1, pp. 13–29, 1996.

[45] D. Reynolds, "Gaussian mixture models," in *Encyclopedia Biometrics*. New York, NY, USA: Springer, 2009, pp. 659–663.

[46] F. Moghimi, A. Nasri, and R. Schober, "Adaptive L_p-norm spectrum sensing for cognitive radio networks," *IEEE Trans. Commun.*, vol. 59, no. 7, pp. 1934–1945, Jul. 2011.

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

[48] R. Bartle, *The Elements of Integration and Lebesgue Measure* (Wiley Classics Library). Hoboken, NJ, USA: Wiley, 1995.
[49] G. Shorack, *Probability for Statisticians* (Springer Texts in Statistics). New York, NY, USA: Springer, 2006.
[50] H. Sohrab, *Basic Real Analysis*. New York, NY, USA: Springer, 2014.
[51] E. Stein and R. Shakarchi, *Complex Analysis* (Complex Analysis). Princeton, NJ, USA: Princeton Univ. Press, 2010.
[52] D. Sarason, *Complex Function Theory*, vol. 49. Providence, RI, USA: American Mathematical Society, 2007.
[53] L. Koralov and Y. Sinai, *Theory of Probability and Random Processes*. New York, NY, USA: Springer, 2007.
[54] J. M. Borwein and A. S. Lewis, *Convex Analysis and Nonlinear Optimization, Theory and Examples*. New York, NY, USA: Springer-Verlag, 2010.
[55] D. G. Luenberger, *Optimization by Vector Space Methods*. New York, NY, USA: Wiley, 1969.

MOHAMMAD RANJBAR received the B.Sc. degree in electrical engineering from the Iran University of Science and Technology in 2011, and the M.S. degree from City University London in 2012. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Akron, Akron, OH, USA. His research interests include energy-efficient wireless networks, cognitive radio, and full-duplex radio. He received the Graduate Thesis Award for his M.S. degree.

NGHI H. TRAN (SM'15) received the B.Eng. degree from the Hanoi University of Technology, Vietnam, in 2002, and the M.Sc. and Ph.D. degrees from the University of Saskatchewan, Canada, in 2004 and 2008, respectively, all in electrical and computer engineering. From 2008 to 2010, he was with McGill University as a Post-Doctoral Scholar under the prestigious Natural Sciences and Engineering Research Council of Canada Post-Doctoral Fellowship, where he was a Research Associate from 2010 to 2011. He was also a Consultant in the satellite industry. He joined the Department of Electrical and Computer Engineering, University of Akron, Akron, OH, USA, in 2011, as an Assistant Professor, where he was promoted to an Associate Professor in 2017. His research interests include signal processing and communication and information theories for wireless systems and networks. He has been serving as a TPC member for a number of flagship IEEE conferences. He received the Graduate Thesis Award for his M.Sc. degree. He was a TPC Co-Chair of the Workshop on Trusted Communications with Physical Layer Security for the IEEE GLOBECOM 2014 and a Publicity Chair of the Workshop on Full-Duplex Communications for Future Wireless Networks for the IEEE ICC 2017 and the Second Workshop on Full-Duplex Communications for the IEEE Globecom 2017. He is currently an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE COMMUNICATIONS LETTERS, and *Elsevier Physical Communication*, and a Lead Guest Editor of the EURASIP *Journal on Wireless Communications and Networking*, Special Issue on Full-Duplex Radio: Theory, Design, and Applications.

TRUYEN V. NGUYEN received the B.S. degree in mathematics and computer science from the National University of Vietnam, Ho Chi Minh City, in 1998, and the M.Sc. and Ph.D. degrees from Temple University, USA, in 2002 and 2005, respectively, all in mathematics. From 2005 to 2006, he was a Post-Doctoral Fellow with the Mathematical Sciences Research Institute, Berkeley. From 2006 to 2007, he holds a post-doctoral position at the Georgia Institute of Technology. Since 2007, he has been a Faculty Member with the Department of Mathematics, University of Akron, Akron, OH, USA. He has published numerous papers in prestigious journals in pure and applied mathematics. His research interests include partial differential equations and calculus of variations.

MUSTAFA CENK GURSOY (SM'16) received the B.S. degree (Hons.) in electrical and electronics engineering from Bogazici University, Istanbul, Turkey, in 1999, and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, USA, in 2004. From 2004 to 2011, he was a Faculty Member with the Department of Electrical Engineering, University of Nebraska-Lincoln. He is currently a Professor with the Department of Electrical Engineering and Computer Science, Syracuse University. His research interests include wireless communications, information theory, communication networks, and signal processing. He is currently a member of the editorial boards of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING, the IEEE TRANSACTIONS ON COMMUNICATIONS, and the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He also served as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2010 to 2015, the IEEE COMMUNICATIONS LETTERS from 2012 to 2014, the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS—SERIES ON GREEN COMMUNICATIONS AND NETWORKING from 2015 to 2016, and *Physical Communication* (Elsevier) from 2010 to 2017. He also served as a Co-Chair of the Communication QoS and System Modeling Symposium and the 2017 International Conference on Computing, Networking and Communications. He was a recipient of the Gordon Wu Graduate Fellowship from Princeton University from 1999 to 2003. He received an NSF CAREER Award in 2006. More recently, he received the EURASIP Journal of Wireless Communications and Networking Best Paper Award, the 2017 IEEE PIMRC Best Paper Award, the 2017 IEEE Green Communications and Computing Technical Committee Best Journal Paper Award, the UNL College Distinguished Teaching Award, and the Maude Hammond Fling Faculty Research Fellowship. He is the Aerospace/Communications/Signal Processing Chapter Co-Chair of the IEEE Syracuse Section.

HUNG NGUYEN-LE received the B.Eng. and M.Eng. degrees in electrical engineering from Ho Chi Minh City University of Technology, Vietnam, in 2001 and 2003, respectively, and the Ph.D. degree in electrical engineering from the National University of Singapore in 2008. From 2008 to 2010, he was a Post-Doctoral Research Fellow with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada. Since 2010, he has been with the Department of Electronics and Telecommunications Engineering, The University of Danang, University of Science and Technology, Vietnam. Since 2013, he has been with the Department of Science, Technology and Environment, The University of Danang, where he is currently an Associate Professor. His research interests include array signal processing, multiuser/multicell transmissions, channel estimation, and synchronization in broadband wireless communications.