EAI Endorsed Transactions
on Industrial Networks and Intelligent Systems
Research Article EAI.EU

On the Capacity-Achieving Scheme and Capacity of 1-Bit ADC Gaussian-Mixture Channels

Md Hasan Rahman[1], Mohammad Ranjbar[1], and Nghi H. Tran[1,*]

[1]Department of Electrical and Computer Engineering, University of Akron, Akron OH 44325, USA

Abstract

This paper addresses the optimal signaling scheme and capacity of an additive Gaussian mixture (GM) noise channel using 1-bit analog-to-digital converters (ADCs). The consideration of GM noise provides a more realistic baseline for the analysis and design of co-channel interference links and networks. Towards that goal, we first show that the capacity-achieving input signal is $\pi/2$ circularly symmetric. By examining a necessary and sufficient Kuhn–Tucker condition (KTC) for an input to be optimal, we demonstrate that the maximum number of optimal mass points is four. Our proof relies on Dubin's theorem and the fact that the KTC coefficient is positive, i.e., the power constraint is active. By combining with the $\pi/2$ circularly symmetric property, it is then concluded the optimal input is unique, and it has exactly four mass points forming a square centered at the origin. By further checking the first and second derivatives of the modified KTC, it is then shown that the phase of the optimal mass point located in the first quadrant is $\pi/4$. Thus, the capacity-achieving input signal is QPSK. This result helps us obtain the channel capacity in closed-form.

Keywords: 1-bit ADC, Capacity, Gaussian-Mixture, Kuhn-Tucker Condition, Mutual Information

Received on 20 October 2019; accepted on 06 January 2020; published on 31 January 2020

doi:10.4108/eai.31-1-2020.162830

1. Introduction

As noted in [1, 2], current information and communication technology (ICT) sector that relies heavily on mobile applications is responsible for 3% of the total worldwide energy consumption, which accounts for 2% to 2.5% of the total CO_2 emission in the world. As a result, a possible energy crunch can be a significant bottleneck for future wireless networks. Thus, there is an urgent need to develop proper communication paradigms for emerging wireless applications that require high energy efficiency. Among different approaches, ultra-low resolution analog-to-digital-converter (ADC) such as 1-bit ADC has received significant attention, and it is considered as an attractive solution to reduce cost and power for high bandwidth wireless systems or systems having multiple RF chains (please see [3] and references therein). The use of low-resolution ADC also significantly reduces quality requirements of key elements in RF circuitry, such as low-noise amplifiers and oscillators. In addition, 1-bit ADC is simply a comparator. Therefore, automatic gain control (AGC) is not needed. Another benefit of employing low-resolution ADCs is the significant reduction in the amount of data being processed and exchanged at the receiver, which is critical for high-speed applications.

During the last decade, significant attention has been paid to information-theoretic aspects of 1-bit ADC channels under traditional additive white Gaussian noise (AWGN), and several interesting results have been obtained. Specifically, as one of the first pioneering investigations addressing the fundamental limits of low-resolution ADC systems, reference [4] showed that in a static AWGN environment, BPSK is capacity-achieving. The results have also been extended to AWGN channels under additional impairments such as fading under different channel state information (CSI) assumptions [5–11] as well as multi-antenna channels [12–14]. For example, for a complex fading single-antenna channel with perfect CSI knowledge at the transmitter and receiver, QPSK is optimal. In the case that CSI is available at the receiver only, it was shown in [8] that the use of any $\pi/2$ circularly symmetric input signal with constant amplitude achieves the channel capacity. In references [10, 11], it was demonstrated that for a general non-coherent Rician fading channel, a rotated QPSK is optimal. Recently, the detailed characterizations of the capacity region and capacity-achieving schemes for a multiple-access Gaussian channel in Rayleigh fading have also been established [15]. In addition, effective coding schemes and achievable rate regions have been established for 1-bit ADC in multi-terminal communications in a recent work in [16].

In current and future communication systems, especially cellular networks, due to heterogeneous structures, co-channel interference exists, and it is asynchronous and intermittent with the main communication. To date, the assumption of having noise plus interference as conditionally Gaussian has been widely adopted, and it has served as

*Corresponding author. Email: Nghi.tran@uakron.edu

a basic block for extensive developments of information-theoretic studies for both point-to-point links and multi-user networks. However, it has been widely recognized via both experimental results and analytical analysis that the asynchronism in a heterogeneous network leads to a non-Gaussian noise plus interference [17–21]. As an example, in a cellular network consisting of micro cells and macro cells, noise plus co-channel interference follows a Gaussian mixture (GM) distribution [17–20]. However, different from Gaussian channels, information-theoretic results for non-Gaussian links and networks are rather limited. In fact, even for a simple non-Gaussian point-to-link, numerical methods can still be used to identify a capacity-achieving signal and the corresponding capacity [22, 23].

In this work, we extend the capacity analysis and optimal signaling design for 1-bit ADC to a general complex GM channel. In the first step, we show that an optimal input signal is $\pi/2$ circularly symmetric. By examining a necessary and sufficient Kuhn–Tucker condition (KTC) for an input to be optimal, we demonstrate that the number of mass points in the optimal input is upper-bounded by four. Due to the presence of multiple Gaussian components in the GM noise, it is much more challenging to examine the KTC and the detailed characteristics of the capacity-achieving signals. Our proof relies on Dubin's theorem and the fact that the KTC coefficient is positive, i.e., the power constraint is active. Combining with the $\pi/2$ circularly symmetric property, we can then conclude that the optimal input is unique, and it has exactly 4 mass points. These four mass point forms a square centered at the origin. By checking the first and second derivatives of the modified KTC, it is then shown that the phase of the optimal mass point located in the first quadrant is $\pi/4$. As a result, QPSK is the capacity achieving scheme. The capacity can then be obtained in closed-form.

The rest of the paper is organized as follows. The channel model is first described in Section 2. In Section 3, we formulate the channel capacity and demonstrate the existence of the optimal input signal. The KTC and detailed characteristics of the optimal input are established in Section 4. In Section 5, numerical results are provided to confirm the analysis. Finally, conclusions are drawn in Section 6.

2. 1-bit ADC under GM Noise: Channel Model and Conditional Probability Density Functions (PDFs)

We consider a communication channel under additive Gaussian mixture noise with the following input-output model:

$$Z = X + N. \tag{1}$$

In (1), X is the complex input signal, and it is subject to the power constraint $E\left[|X|^2\right] \le P$. Furthermore, N is the additive noise that follows a GM distribution. Its probability density function (pdf) is a mixture of M complex Gaussian distributions with mean 0 and variance σ_k^2, $1 \le k \le M$, which is given as

$$p_N(n) = \sum_{k=1}^{M} \frac{\varepsilon_k}{2\pi\sigma_k^2} \exp\left(-\frac{|n|^2}{2\sigma_k^2}\right), \tag{2}$$

where $\varepsilon_k > 0$ is the mixing probability satisfying $\sum_{k=1}^{M} \varepsilon_k = 1$. The conditional PDF of the received signal Z can be written as,

$$p(z|x) = \sum_{k=1}^{M} \frac{\varepsilon_k}{\pi{\sigma_k}^2} e^{-\frac{|z-x|^2}{\sigma_k^2}}. \tag{3}$$

Using the real and imaginary components of Z, the above PDF can be re-written as

$$p(z|x) = \sum_{k=1}^{M} \varepsilon_k \frac{1}{\sigma_k\sqrt{\pi}} e^{-\frac{(\Re(z-x))^2}{\sigma_k^2}} \frac{1}{\sigma_k\sqrt{\pi}} e^{-\frac{(\Im(z-x))^2}{\sigma_k^2}}. \tag{4}$$

The complex received signal Z is then fed to two identical 1-bit ADCs, and the quantized output Y can be written as

$$Y = Q(Z), \tag{5}$$

where $Q(\cdot)$ denotes the 1-bit quantization scheme. Without loss of generality, assume the threshold levels for each 1-bit ADC are $[q_0, q_1, q_2] \in R$ such that $q_0 < q_1 < q_2$, with $q_0 = -\infty, q_1 = 0, q_2 = +\infty$. Thus, Y can be expressed as,

$$Y = \Re(Y) + \Im(y) = \text{sign}(\Re(z)) + j\text{sign}(\Im(z)), \tag{6}$$

where sign$(\cdot)$ is the sign function, and $\Re$ and $\Im$ are the real and imaginary parts of a complex number, respectively. Equivalently, it is clear that Y belongs to the set $\mathcal{Y} \in \{y_{1,1} = 1 + j, y_{1,2} = 1 - j, y_{2,1} = -1 + j, y_{2,2} = -1 - j\}$. For a given pair of i and j with $1 \le i, j \le 2$, the transition probability function $W_{i,j}(x)$ for a given $X = x$ can be obtained as

$$\begin{aligned} W_{i,j}(x) &= P\left(Y = y_{i,j} \middle| X = x\right) \\ &= \sum_{k=1}^{M} \varepsilon_k \int_{q_{i-1}}^{q_i} \int_{q_{j-1}}^{q_j} \frac{1}{\sigma_k\sqrt{\pi}} e^{-\frac{(\Re(z-x))^2}{\sigma_k^2}} \frac{1}{\sigma_k\sqrt{\pi}} e^{-\frac{(\Im(z-x))^2}{\sigma_k^2}} dz \\ &= \sum_{k=1}^{M} \varepsilon_k \left[Q\left(\frac{\sqrt{2}(q_{i-1} - \Re(x))}{\sigma_k}\right) - Q\left(\frac{\sqrt{2}(q_i - \Re(x))}{\sigma_k}\right)\right] \\ &\quad \times \left[Q\left(\frac{\sqrt{2}(q_{j-1} - \Im(x))}{\sigma_k}\right) - Q\left(\frac{\sqrt{2}(q_j - \Im(x))}{\sigma_k}\right)\right]. \end{aligned}$$

Here, $Q(\cdot)$ is well-known tail distribution function of the standard normal distribution given as

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} exp^{-\frac{u^2}{2}} du. \tag{7}$$

Now, to further simplify $W_{i,j}$, Table 1 shows the relationship between q_{i-1}'th and q_i'th thresholds and $\Re(y_{i,j})$ and $\Im(y_{i,j})$. Let start with the transition probability $W_{1,1}(x)$:

$$W_{1,1}(x) = P(Y = 1 + j \mid X = x). \tag{8}$$

Table 1. Threshold limits for the real and imaginary parts in the quantization bins.

	$\Re(y_{i,j})$	$\Im(y_{i,j})$	$\Re(y_{i,j})$		$\Im(y_{i,j})$	
			Lower Threshold limit q_{i-1}	Upper Threshold limit q_i	Lower Threshold limit q_{i-1}	Upper Threshold limit q_i
$y_{1,1}$	1	1	0	∞	0	∞
$y_{1,2}$	1	-1	0	∞	$-\infty$	0
$y_{2,1}$	-1	1	$-\infty$	0	0	∞
$y_{2,2}$	-1	-1	$-\infty$	0	$-\infty$	0

In this case, both real and imaginary parts of the output fall in the range $[0, \infty)$. Hence, we can write

$$\begin{aligned} W_{1,1} &= \sum_{k=1}^{M} \varepsilon_k \left[Q\left(\frac{\sqrt{2}(0 - \Re(x))}{\sigma_k} \right) - Q\left(\frac{\sqrt{2}(\infty - \Re(x))}{\sigma_k} \right) \right] \\ &\quad \times \left[Q\left(\frac{\sqrt{2}(0 - \Im(x))}{\sigma_k} \right) - Q\left(\frac{\sqrt{2}(\infty - \Im(x))}{\sigma_k} \right) \right] \\ &= \sum_{k=1}^{M} \varepsilon_k \, Q\left(\frac{-\sqrt{2}\Re(x)}{\sigma_k} \right) Q\left(\frac{-\sqrt{2}\Im(x)}{\sigma_k} \right). \end{aligned} \tag{9}$$

Following a similar procedure, we can calculate the other transition probabilities as follows:

$$W_{1,2} = \sum_{k=1}^{M} \varepsilon_k Q\left(\frac{-\sqrt{2}\Re(x)}{\sigma_k} \right) Q\left(\frac{\sqrt{2}\Im(x)}{\sigma_k} \right), \tag{10}$$

$$W_{2,1} = \sum_{k=1}^{M} \varepsilon_k Q\left(\frac{\sqrt{2}\Re(x)}{\sigma_k} \right) Q\left(\frac{-\sqrt{2}\Im(x)}{\sigma_k} \right), \tag{11}$$

$$W_{2,2} = \sum_{k=1}^{M} \varepsilon_k Q\left(\frac{\sqrt{2}\Re(x)}{\sigma_k} \right) Q\left(\frac{\sqrt{2}\Im(x)}{\sigma_k} \right). \tag{12}$$

Therefore, $W_{i,j}$ can be generally expressed as

$$W_{i,j}(x) = \sum_{k=1}^{M} \varepsilon_k Q\left(\frac{-\sqrt{2}\Re(x)\Re(y_{i,j})}{\sigma_k} \right) Q\left(\frac{-\sqrt{2}\Im(x)\Im(y_{i,j})}{\sigma_k} \right). \tag{13}$$

3. Channel Capacity and Existence of Capacity-Achieving Input Distribution

In this section, we shall formulate the channel capacity and demonstrate that there always exists an input distribution that achieves the capacity.

For a given input distribution F_X, the mutual information between the input signal X and output signal Y, denoted as $I(F_X)$, is

$$I(F_X) = I(X;Y) = H_{F_X}(Y) - H_{F_X}(Y|X), \tag{14}$$

where $H_{F_X}(Y)$ and $H_{F_X}(Y|X)$ are the output and conditional entropies, and they can be calculated as:

$$H_{F_X}(Y) = -\int_{\mathbb{C}} \sum_{i=1}^{2} \sum_{j=1}^{2} W_{i,j}(x^*) \log p\left(y_{i,j}; F_X\right) dF_X, \tag{15}$$

and

$$H_{F_X}(Y|X) = -\int_{\mathbb{C}} \sum_{i=1}^{2} \sum_{j=1}^{2} W_{i,j}(x^*) \log W_{i,j}(x^*) \, dF_X. \tag{16}$$

Here, the log operator is of base 2, and $p\left(y_{i,j}; F_X\right)$ is the probability of $Y = y_{i,j}$ when the input follows F_X, and it can be calculated as

$$p\left(y_{i,j}; F_X\right) = \int W_{i,j}(x) \, dF_X(x). \tag{17}$$

The mutual information between the input and output can then be expressed as:

$$I(F_X) = -\int_{\mathbb{C}} \sum_{i=1}^{2} \sum_{j=1}^{2} W_{i,j}(x^*) \log \frac{p\left(y_{i,j}; F_X\right)}{W_{i,j}(x^*)} dF_X. \tag{18}$$

The channel capacity is defined as the supremum of $I(F_X)$ over all input distributions F_X that satisfy the power constraint $E\left[|X|^2\right] \le P$. For convenience, let Ω be the set of such input distributions. The channel capacity can then be written as:

$$C = \sup_{F_X \in \Omega} I(F_X). \tag{19}$$

If there exists an input distribution that achieves C, this distribution, denoted as F_X^*, is referred to as a capacity-achieving distribution. Under the power constraint, it is known that Ω is convex and weakly compact with respect to weak* topology [24]. Furthermore, the mutual information $I(F_X)$ is concave with respect to F_X. Therefore, the existence of F_X^* is equivalent to the continuity of $I(F_X)$ over F_X. The proof of the continuity of $I(F_X)$ follows closely the method in [24] by showing the *weak** continuity of $H_{F_X}(Y)$ and $H_{F_X}(Y|X)$, due to the fact that $W_{i,j}(x)$ is a continuous function of x and bounded. For the brevity of the presentation, the proofs are omitted here. As a result, F_X^* exists.

4. Capacity-Achieving Distribution F_X^*

Given the existence of F_X^*, this section examines the detailed characteristics of F_X^*. Specifically, we first show that it has to be $\pi/2$ circularly symmetric before shedding light on the number of mass points of F_X^* and their locations. Note that a random variable X is $\pi/2$ circularly symmetric if it has the same distribution as $Xe^{kj\pi/2}$ for any $k \in \mathcal{R}$.

4.1. $\pi/2$ circular symmetry

For a given input distribution F_X, let's define a $\frac{\pi}{2}$ circularly symmetric distribution $F_X^{\pi/2}$ as follows

$$F_X^{\pi/2} = \frac{1}{4}\left(F_X(x) + F_X\left(xe^{j\frac{\pi}{2}}\right) + F_X\left(xe^{j\pi}\right) + F_X\left(xe^{j\frac{3\pi}{2}}\right)\right). \tag{20}$$

From (9)-(12), it is straightforward to see that $W_{1,1}(x) = W_{2,1}\left(xe^{j\frac{\pi}{2}}\right)$, $W_{2,1}(x) = W_{2,2}\left(xe^{j\frac{\pi}{2}}\right)$, $W_{2,2}(x) = W_{1,2}\left(xe^{j\frac{\pi}{2}}\right)$, $W_{1,2}(x) = W_{1,1}\left(xe^{j\frac{\pi}{2}}\right)$. Therefore,

$$\sum_{i=1}^{2}\sum_{1}^{2} w_{i,j}(x)\log W_{i,j}(x) = \sum_{i=1}^{2}\sum_{1}^{2} w_{i,j}\left(xe^{j\frac{\pi}{2}}\right)\log W_{i,j}\left(xe^{j\frac{\pi}{2}}\right). \tag{21}$$

It then follows that

$$\begin{aligned}
&H_{F_X^{\frac{\pi}{2}}}(Y|X)\\
&= -\frac{1}{4}\int\sum_{i=1}^{2}\sum_{j=1}^{2} W_{i,j}(x)\log W_{i,j}(x)\\
&\quad\times d\left(F_X(x) + F_X\left(xe^{j\frac{\pi}{2}}\right) + F_X\left(xe^{j\pi}\right) + F_X\left(xe^{j\frac{3\pi}{2}}\right)\right)\\
&= -\frac{1}{4}\left\{\int\sum_{i=1}^{2}\sum_{j=1}^{2} W_{i,j}(x)\log W_{i,j}(x)\,dF_X(x)\right\}\\
&\quad-\frac{1}{4}\left\{\int\sum_{i=1}^{2}\sum_{j=1}^{2} W_{i,j}\left(xe^{j\frac{\pi}{2}}\right)\log W_{i,j}\left(xe^{j\frac{\pi}{2}}\right)dF_X\left(xe^{j\frac{\pi}{2}}\right)\right\}\\
&\quad-\frac{1}{4}\left\{\int\sum_{i=1}^{2}\sum_{j=1}^{2} W_{i,j}\left(xe^{j\pi}\right)\log W_{i,j}\left(xe^{j\pi}\right)dF_X\left(xe^{j\pi}\right)\right\}\\
&\quad-\frac{1}{4}\left\{\int\sum_{i=1}^{2}\sum_{j=1}^{2} W_{i,j}\left(xe^{j\frac{3\pi}{2}}\right)\log W_{i,j}\left(xe^{j\frac{3\pi}{2}}\right)dF_X\left(xe^{j\frac{3\pi}{2}}\right)\right\}\\
&= -\frac{1}{4}\left\{4\int\sum_{i=1}^{2}\sum_{j=1}^{2} W_{i,j}(x)\log W_{i,j}(x)\,dF_X(x)\right\}\\
&= H_{F_X}(Y|X).
\end{aligned} \tag{22}$$

It is clear from (22) that the conditional output entropy is the same for both F_X and $F_X^{\pi/2}$.

Now, we shall examine the output entropy when $F_X^{\pi/2}$ is used. To this end, we can first calculate $p(y_{2,2}; F_X^{\frac{\pi}{2}})$ as follows:

$$\begin{aligned}
&p(y_{2,2}; F_X^{\frac{\pi}{2}})\\
&= \frac{1}{4}\left\{\oint W_{2,2}(x)\,dF_X(x) + \oint W_{2,2}(x)\,dF_X\left(xe^{j\frac{\pi}{2}}\right)\right\}\\
&\quad+\frac{1}{4}\left\{\oint W_{2,2}(x)\,dF_X\left(xe^{j\pi}\right) + \oint W_{2,2}(x)\,dF_X\left(xe^{j\frac{3\pi}{2}}\right)\right\}.
\end{aligned} \tag{23}$$

With a variable transformation, the phases can be interchanged between dF_X and $W_{i,j}$, which leads to

$$\begin{aligned}
&p(y_{2,2}; F_X^{\frac{\pi}{2}})\\
&= \frac{1}{4}\left\{\oint W_{2,2}(x)\,dF_X(x) + \oint W_{2,2}\left(xe^{j\frac{\pi}{2}}\right)dF_X(x)\right\}\\
&\quad+\frac{1}{4}\left\{\oint W_{2,2}\left(xe^{j\pi}\right)dF_X(x) + \oint W_{2,2}\left(xe^{j\frac{3\pi}{2}}\right)dF_X(x)\right\}\\
&= \frac{1}{4}\left\{\oint\left(W_{2,2}(x) + W_{1,1}(x) + W_{1,2}(x) + W_{2,1}(x)\right)dF_X(x)\right\}\\
&= \frac{1}{4}\left\{\oint dF_X(x)\right\} = \frac{1}{4}.
\end{aligned} \tag{24}$$

In a similar manner, it can be proved that $p(y_{1,1}; F_X^{\frac{\pi}{2}}) = p(y_{1,2}; F_X^{\frac{\pi}{2}}) = p(y_{2,1}; F_X^{\frac{\pi}{2}}) = \frac{1}{4}$. It means that when $F_X^{\pi/2}$ is used, the output Y is uniformly distributed, and it results in a maximum output entropy. Thus, a $\pi/2$ circularly symmetric $F_X^{\pi/2}$ leads to a better input-output mutual information than F_X. As a result, F_X^* is $\pi/2$ circularly symmetric. Hence the output entropy can be calculated as

$$H_{F_X}(Y) = -\sum_{i=1}^{2}\sum_{j=1}^{2} p(y_{i,j}; F_X^{\frac{\pi}{2}})\log p(y_{i,j}; F_X^{\frac{\pi}{2}}) = \log 4. \tag{25}$$

The results state that the output probability distribution resulting from $F_X^{\frac{\pi}{2}}$ is uniform, which maximizes the output entropy. Equivalently, $F_X^{\frac{\pi}{2}}$ always leads to better mutual information than F_X. As a result, it can be concluded that F_X^* belongs to a set of $pi/2$ circularly symmetric distributions, denoted as $\Omega_{\frac{\pi}{2}}$. Note that for any distribution F_X in $\Omega_{\frac{\pi}{2}}$, the corresponding input-output mutual information can be calculated as

$$I(F_X) = E_X[-d(x)] + \log(4). \tag{26}$$

where $d(x) = \left[-\sum_{i=1}^{2}\sum_{j=1}^{2} W_{i,j}(x)\log W_{i,j}(x)\right]$.

4.2. Kuhn-Tucker Condition (KTC) and the Number of Mass Points in F_X^*

In this subsection, we shall further characterize important properties of the optimal distribution F_X^* by first establishing a necessary and sufficient condition for an input distribution to be optimal, which is called Kuhn-Tucker condition (KTC). The KTC is then exploited to find the number of mass points in F_X^*.

KTC. We first state the following lemma.

Lemma 1. For a given $F_X^0 \in \Omega_{\frac{\pi}{2}}$, the limit

$$\lim_{\theta\to 0} \frac{I\left((1-\theta)\,F_X^0 + \theta F_X\right) - I(F_X^0)}{\theta} = I(F_X) - I(F_X^0) \quad (27)$$

exists, and it is finite for all $F_X \in \Omega_{\frac{\pi}{2}}$ and $\theta \in [0,1]$.

Proof. Let's define F_X^θ as, $F_X^\theta = (1-\theta)F_X^0 + \theta F_X$. Therefore, $dF_X^\theta = (1-\theta)dF_X^0 + \theta dF_X$ and

$$p\left(y_{i,j}; F_X^\theta\right) = (1-\theta)\, p\left(y_{i,j}; F_X^0\right) + \theta\, p\left(y_{i,j}; F_X\right). \quad (28)$$

Then the mutual information for any F_θ can be written as

$$I\left(F_X^\theta\right) = -\int \sum_{i=1}^{2}\sum_{j=1}^{2} W_{i,j}(x) \log \frac{p\left(y_{i,j}; F_X^\theta\right)}{W_{i,j}(x)} dF_X^\theta. \quad (29)$$

As a result, we have

$$\begin{aligned} & I\left(F_X^\theta\right) - I\left(F_X^0\right) \\ &= -(1-\theta)\int \sum_{i=1}^{2}\sum_{j=1}^{2} W_{i,j}(x) \log \frac{p\left(y_{i,j}; F_X^\theta\right)}{W_{i,j}(x)} dF_X^0 \\ &\quad - \theta \int \sum_{i=1}^{2}\sum_{j=1}^{2} W_{i,j}(x) \log \frac{p\left(y_{i,j}; F_X^\theta\right)}{W_{i,j}(x)} dF_X \\ &\quad + \int \sum_{i=1}^{2}\sum_{j=1}^{2} W_{i,j}(x) \log \frac{p\left(y_{i,j}; F_X^0\right)}{W_{i,j}(x)} dF_X^0 \end{aligned} \quad (30)$$

By rearranging the terms in the above equation, it can be re-written as,

$$\begin{aligned} & I\left(F_X^\theta\right) - I\left(F_X^0\right) \\ &= -\int \sum_{i=1}^{2}\sum_{j=1}^{2} W_{i,j}(x) \log \frac{p\left(y_{i,j}; F_X^\theta\right)}{p\left(y_{i,j}; F_X^0\right)} dF_X^0 \\ &\quad - \theta \int \sum_{i=1}^{2}\sum_{j=1}^{2} W_{i,j}(x) \log \frac{p\left(y_{i,j}; F_X^\theta\right)}{W_{i,j}(x)} dF_X \\ &\quad + \theta \int \sum_{i=1}^{2}\sum_{j=1}^{2} W_{i,j}(x) \log \frac{p\left(y_{i,j}; F_X^\theta\right)}{W_{i,j}(x)} dF_X^0. \end{aligned} \quad (31)$$

Now, considering the limit $\theta \to 0$, we have

$$\begin{aligned} & \lim_{\theta\to 0} \frac{I\left(F_X^\theta\right) - I\left(F_X^0\right)}{\theta} \\ &= -\lim_{\theta\to 0} \frac{1}{\theta} \int \sum_{i=1}^{2}\sum_{j=1}^{2} W_{i,j}(x) \log \frac{p\left(y_{i,j}; F_X^\theta\right)}{p\left(y_{i,j}; F_0\right)} dF_X^0 \\ &\quad + I_{F_X}\left(F_X^0\right) - I\left(F_X^0\right). \end{aligned} \quad (32)$$

Since $\frac{p(Y_{i,j};F_\theta)}{p(Y_{i,j};F_0)} \to 1$ when $\theta \to 0$, it then follows that

$$\lim_{\theta\to 0} \frac{I\left((1-\theta)\,F_X^0 + \theta F_X\right) - I(F_X^0)}{\theta} = I_{F_X}(F_X^0) - I(F_X^0). \quad (33)$$

The limit therefore exists, and it is finite. It is because both terms on the right-hand side of (33) are finite. □

Based on Lemma 1, we can conclude that $I(F_X)$ is weakly differentiable. In addition, due to the concavity of $I(F_X)$, and also the convexity and compactness of $\Omega_{\frac{\pi}{2}}$, we can use the theorem of Lagrangian multipliers. Specifically, there exists a non-negative μ such that

$$C = \sup_{\substack{F_X\in\Omega_{\frac{\pi}{2}} \\ \mathbb{E}[|X|^2]\le P}} I(F_X) = \sup_{F_X\in\Omega_{\frac{\pi}{2}}} I(F_X) - \mu\phi(F_X), \quad (34)$$

where $\phi(F_X) = \int |x|^2 dF_X - P$. It is not difficult to verify that $\phi(F_X)$ is also weak differentiable with weak derivative, i.e.,

$$\phi'_{F_X^0}(F_X) = \phi(F_X) - \phi(F_X^0). \quad (35)$$

Therefore, by following the arguments as in [25], we can conclude that $F_X^* \in \Omega_{\frac{\pi}{2}}$ is optimal if and only if

$$\begin{aligned} & I'_{F_X^*}(F_X) - \mu\phi'_{F_X^*}(F_X) \\ &= I_{F_X}\left(F_X^*\right) - I\left(F_X^*\right) - \mu\left(\phi(F_X) - \phi\left(F_X^*\right)\right) \le 0, \end{aligned} \quad (36)$$

where the equality is achieved when $x \in E_X^*$, the set of points of increase of F_X^*. After some manipulations, the KTC can be established as

$$-d(x) + \log(4) \le C + \mu(|x|^2 - P), \quad (37)$$

where the equality is achieved when x is a mass point of F_X^*.

The number of mass points in F_X^*. Given the above KTC, in the following, we will examine further the number of mass points in F_X^*. To do so, we consider the following two cases of the KTC coefficient μ.

case 1. $\mu > 0$. In this case, the term $\mu\left(|x|^2 - P\right) \to \infty$ when $|x| \to \infty$. Furthermore, we have

$$0 \le \log 4 + \left[\sum_{i=1}^{2}\sum_{j=1}^{2} W_{i,j}(x) \log W_{i,j}(x)\right] \le \log 4. \quad (38)$$

As a result, the equality in (37) cannot be achieved when $|x| \to \infty$. That means for any point $|x| \in E_X^*$, the magnitude $|x|$ is finite.

case 2. $\mu = 0$. In this case, the KTC can be re-written as follows:

$$d(x) \ge \log(4) - C. \quad (39)$$

In the following, we will show that $\lim_{|x|\to\infty} d(x)$ is achieved from above, i.e.,

$$\exists M \in \mathbb{R}^+ \quad \Big| \quad \forall |x| > M : d(x) > \lim_{|x|\to\infty} d(x), \quad (40)$$

or equivalently, F_X^* is bounded. To do so, let first consider the case that $\phi_x \neq i\frac{\pi}{2}$ for all $i \in \mathbb{R}$, where ϕ_x denotes the phase of x. We then rewrite $W_{i,j}$ as

$$W_{i,j}(x) = \sum_{k=1}^{M} \varepsilon_k Q\left(\frac{-\sqrt{2}\,|x|\cos(\phi_x)\Re(y_{i,j})}{\sigma_k}\right) \times Q\left(\frac{-\sqrt{2}\,|x|\sin(\phi_x)\Im(y_{i,j})}{\sigma_k}\right). \tag{41}$$

It is then straightforward to verify that $\lim_{|x|\to\infty} W_{1,1}(x) = 1$, which is equivalent to $\lim_{|x|\to\infty} d(x) = 0$. Moreover, it is clear that $d(x) > 0$, due to the fact that $W_{i,j}$'s are bounded away from zero (please refer to the equations (9)-(12)), and $-x\log x > 0$ for $x > 0$. In the case that $\phi_x = i\frac{\pi}{2}$, without loss of generality, we assume $i = 1$. It is then not difficult to show that

$$\lim_{|x|\to\infty} W_{1,1}(x) = \lim_{|x|\to\infty} W_{2,1}(x) = 0.5, \tag{42}$$

or $\lim_{|x|\to\infty} d(x) = 1$. It is also clear form (41), that for $\phi_x = \frac{\pi}{2}$, we have $W_{1,1}(x) < 0.5$ and $W_{2,1}(x) < 0.5$. Combining this fact with (42), we have

$$\exists M \in \mathbb{R}^+ \quad \Big| \quad \forall\, |x| > M : e^{-1} < W_{1,1}(x), W_{2,1}(x) < 0.5. \tag{43}$$

On the other hand, by calculating the derivative of $-x\log x$ and checking the sign of the derivative, it is easy to verify that $-x\log x$ is decreasing for $x > e^{-1}$. Thus

$$\exists M \in \mathbb{R}^+ \quad \Big| \quad \forall\, |x| > M : d(x) > 1 = \lim_{|x|\to\infty} d(x). \tag{44}$$

As a result, for the case of $\mu = 0$, the magnitude $|x|$ is finite for any point $|x| \in E_X^*$.

Combining the results from Cases 1 and 2, it is then clear that the optimal input distribution must have a bounded amplitude. Given that, we have the following proposition regarding the number of mass points in F_X^*:

Proposition 1. The support set of an optimal input distribution contains at most 4 mass points.

Proof. Let $P_0 \leq P$ and $R^* = [p_{1,1}^*, p_{1,2}^*, p_{2,1}^*, p_{2,2}^*]$ be the power and the output probabilities, respectively, under the optimal input. Moreover, let $B(l)$ be a Borel set of complex numbers $\{x\}$ with $-\leq \Re(x), \Im(x) \leq l$. Since F_X^* has a bounded support, $supp(F_X^*) \subset B(L)$, where L is a finite number. Then defining a new convex set as follows:

$$\mathcal{L} = \{F_X | supp(F_X) \in B(L)\}. \tag{45}$$

It is clear that $\mathcal{L}$ is convex and compact, and $F_X^* \in \mathcal{L}$. As such, we have:

$$C = \sup_{\substack{F_X\in\mathcal{L}\\ \mathbb{E}[|X|^2]\leq P}} I(F_X) = \sup_{F_X\in\mathcal{L}} I(F_X) - \mu\left(\int |x|^2 dF_X - P\right), \tag{46}$$

where μ is a non-negative number. Let a subset $\mathcal{U}$ of $\mathcal{L}$ be defined as:

$$\mathcal{U} = \{F_X \in \mathcal{L} | p(y; F_X) = R^*\}. \tag{47}$$

The optimal input F_X^* therefore belongs to $\mathcal{U}$. It then follows that

$$C = \max_{F_X\in\mathcal{U}} I(F_X) - \mu\left(\int |x|^2 dF_X - P\right). \tag{48}$$

It is not difficult to show that the objective function of the above equation is a linear function of F_X over the set $\mathcal{U}$. It is is because $H_{F_X}(Y)$ is constant over this set and $H_{F_X}(Y|X)$, $\mu\left(\int |x|^2 dF_X - P\right)$ are linear functions of F_X. As a result, the solution of (48) can be found at an extreme point of $\mathcal{U}$. Hence, F_X^* is an extreme point of $\mathcal{U}$. Furthermore, the set $\mathcal{U}$ is the intersection of $\mathcal{L}$ and 3 hyperplanes defined as:

$$H_{i,j} : \int W_{i,j}(x)\, dF_X = P_{i,j}^*, \quad (i,j) \neq (2,2). \tag{49}$$

Following Dubin's theorem [11], the extreme points of $\mathcal{U}$ are a convex combination of at most 4 extreme points of the set $\mathcal{L}$. In addition, the extreme points of $\mathcal{L}$ have only one mass point. So the capacity can be achieved by a discrete distribution having at most 4 mass points. □

4.3. Optimal Solution

Given the result from Proposition 1, and from the fact that F_X^* is $\frac{\pi}{2}$ circularly symmetric, it is clear that F_X^* must have exactly 4 mass points, each belonging to a quadrant, and they have the same amplitude. In the following, we will show that the power constraint is active, and the amplitude is $\sqrt{P}$. The result is stated in the following lemma.

Lemma 2. The KTC coefficient μ is positive. Equivalently, $P_0 = P$.

Proof. Assume that $\mu = 0$, i.e., the power constraint is inactive with $P_0 < P$. It implies that

$$\sup_{F_X, \mathbb{E}\left[|X|^2\right]\leq P} I(F_X) = \sup_{F_X} I(F_X). \tag{50}$$

Let first examine the LHS of (50). We have:

$$\sup_{\substack{F_X\\ \mathbb{E}\left[|X|^2\right]\leq P}} I(F_X) \leq 2 - \inf_{\substack{F_X\\ \mathbb{E}\left[|X|^2\right]\leq P}} H_{F_X}(Y|X). \tag{51}$$

Furthermore,

$$H_{F_X}(Y|X) = \left[-\sum_{y\in\mathcal{Y}} p(y\mid x)\log p(y\mid x)\right] dF_X \tag{52}$$

$$\geq \int \left[H_b\left(\min_{y\in\mathcal{Y}} p(y\mid x)\right)\right] dF_X$$

$$\geq \int \left[\min_{y\in\mathcal{Y}} p(y\mid x)\right] dF_X, \tag{53}$$

where $H_b(p) = -p\log\ p - (1-p)\log\ (1-p)$, and (53) follows from the fact that $H_b(p) \geq p$ for $0 \leq p \leq \frac{1}{2}$. In addition, due to monotonicity of $Q(\cdot)$, we have $\min_{y\in\mathcal{Y}} p(y \mid x) \geq \sum_{k=1}^{M} Q\left(\frac{|x|}{\sqrt{0.5}\sigma_k}\right)^2$. Thus,

$$H_{F_X}(Y|X) \geq \int\int\int Q\left(\frac{|x|}{\sqrt{0.5}\sigma_k}\right)^2 dF_X(x_1). \tag{54}$$

Moreover, because $Q\left(\sqrt{u}\right)^2$ is a convex function of $u \geq 0$, by applying Jensen's inequality to (54), we obtain

$$H_{F_X}(Y|X) \geq Q\left(\frac{\sqrt{P}}{\sqrt{0.5}\sigma}\right)^2. \tag{55}$$

Then combining (51) and (55), we have

$$\sup_{\substack{F_X \\ \mathbb{E}\left[|X|^2\right]\leq P}} I\left(F_{X_1}F_{X_2}\right) \leq 2 - Q\left(\frac{\sqrt{P}}{\sqrt{0.5}\sigma}\right)^2. \tag{56}$$

Now, consider the RHS of (50). As we do not have any power constraint, let us consider the following distribution of X

$$F_X^k(x) = \frac{1}{4}\sum_{l=1}^{4}\delta\left(x - (k+jk)\,e^{jl\pi/2}\right). \tag{57}$$

By using this input, it can be verified that $\lim_{k\to\infty} I\left(F_X^k\right) = 2$. It is then obvious that by using F_X in the form of (57) with a sufficiently large k, $I(F_X)$ can be made arbitrarily close to 2. Therefore, the equality in (50) is not achievable. □

Following the result of Lemma 2, it is clear that the optimal support set E_X^* consists of 4 mass points, each point is in one quadrant having the amplitude $\sqrt{P}$. Let $x^* = \sqrt{P}\exp(j\phi^*)$ be the optimal mass point in the first quadrant, with $0 \leq \phi^* \leq \pi/2$. While it has been shown earlier that ϕ^* exists, it is also unique. It is because if there exists another solution ϕ' corresponding to another optimal distribution F_X', from the concavity of the mutual information over the input distribution, we then have

$$I(\alpha F_X^* + (1-\alpha)F_X') \geq \alpha I(F_X^*) + (1-\alpha)I(F_X'), \tag{58}$$

for $0 < \alpha < 1$. As a result, the distribution $\alpha F_X^* + (1-\alpha)F_X'$ is also optimal. However, this distribution has eight mass points, which is not possible.

Given the uniqueness of ϕ^*, we will address its solution explicitly. In particular, from the KTC, we have:

$$-\left[\sum_{i=1}^{2}\sum_{j=1}^{2} W_{i,j}\left(\sqrt{P}\exp(j\phi^*)\right) log W_{i,j}\left(\sqrt{P}\exp(j\phi^*)\right)\right] = \log(4) - C. \tag{59}$$

For convenience, and with a slight abuse of notation, let $W_{i,j}(\phi) = W_{i,j}\left(\sqrt{P}\exp(j\phi)\right)$. The optimal ϕ^* is therefore a solution of

$$\phi^* = \operatorname{argmin}\min_{0\leq\phi\leq\frac{\pi}{2}} S(\phi), \tag{60}$$

where

$$S(\phi) = -\left[\sum_{i=1}^{2}\sum_{j=1}^{2} W_{i,j}(\phi)\, log W_{i,j}(\phi)\right]. \tag{61}$$

From the first order necessary condition (FONC), we then have:

$$\frac{dS(\phi)}{d\phi}_{|\phi=\phi^*} = 0. \tag{62}$$

The derivative in (62) can be calculated as:

$$\frac{dS(\phi)}{d\phi} = \frac{d}{d\phi}\left(-\sum_{i=1}^{2}\sum_{j=1}^{2} W_{i,j}(\phi)\, log W_{i,j}(\phi)\right) \tag{63}$$

$$= -\sum_{i=1}^{2}\sum_{j=1}^{2} W_{i,j}'(\phi)\left(\frac{1}{\ln(2)} + \log\left(W_{i,j}(\phi)\right)\right) \tag{64}$$

where $W_{i,j}'(\phi) = \frac{dW_{i,j}(\phi)}{d\phi}$. Now, let $A_k = \frac{-\sqrt{2P}}{\sigma_k}$. We can first write $W_{1,1}'(\phi)$ as:

$$W_{1,1}'(\phi) = \frac{d}{d\phi}\left(\sum_{k=1}^{M}\mathcal{E}_k Q(A_k\cos\phi)\, Q(A_k\sin\phi)\right) \tag{65}$$

$$= \sum_{k=1}^{M}\mathcal{E}_k\frac{d}{d\phi}\left(Q(A_k\cos\phi)\, Q(A_k\sin\phi)\right)$$

$$= \frac{1}{\log 2}\left(\sum_{k=1}^{M}\varepsilon_k\left\{\frac{A_k\cos(\phi)e^{-\frac{a^2\sin^2(\phi)}{2}} erfc\left(\frac{A_k\cos(\phi)}{\sqrt{2}}\right)}{2\sqrt{2\pi}}\right\}\right) \tag{66}$$

$$\times\left(1 + \log\left[\sum_{k=1}^{M}\varepsilon_k\frac{erfc\left(\frac{A_k\cos(\phi)}{\sqrt{2}}\right)erfc\left(\frac{A_k\sin(\phi)}{\sqrt{2}}\right)}{4}\right]\right) \tag{67}$$

$$-\frac{1}{\log 2}\left(\sum_{k=1}^{M}\varepsilon_k\left\{\frac{A_k\sin(\phi)e^{-\frac{a^2\cos^2(\phi)}{2}} erfc\left(\frac{A_k\sin(\phi)}{\sqrt{2}}\right)}{2\sqrt{2\pi}}\right\}\right) \tag{68}$$

$$\times\left(1 + \log\left[\sum_{k=1}^{M}\varepsilon_k\frac{erfc\left(\frac{A_k\cos(\phi)}{\sqrt{2}}\right)erfc\left(\frac{A_k\sin(\phi)}{\sqrt{2}}\right)}{4}\right]\right). \tag{69}$$

Furthermore,

$$W'_{2,1}\phi = \frac{1}{\log 2}\left(\sum_{k=1}^{M}\varepsilon_k\left\{-\frac{A_k\cos(\phi)e^{-\frac{a^2\sin^2(\phi)}{2}}erfc\left(\frac{A_k\cos(\phi)}{\sqrt{2}}\right)}{2\sqrt{2\pi}}\right\}\right)$$
$$\times\left(1+\log\left[\sum_{k=1}^{M}-\varepsilon_k\frac{erfc\left(\frac{A_k\cos(\phi)}{\sqrt{2}}-2\right)erfc\left(\frac{A_k\sin(\phi)}{\sqrt{2}}\right)}{4}\right]\right)$$
$$+\frac{1}{\log 2}\left(\sum_{k=1}^{M}\varepsilon_k\left\{\frac{A_k\sin(\phi)e^{-\frac{a^2\cos^2(\phi)}{2}}erfc\left(\frac{A_k\sin(\phi)}{\sqrt{2}}\right)}{2\sqrt{2\pi}}\right\}\right)$$
$$\times\left(1+\log\left[\sum_{k=1}^{M}-\varepsilon_k\frac{erfc\left(\frac{A_k\cos(\phi)}{\sqrt{2}}-2\right)erfc\left(\frac{A_k\sin(\phi)}{\sqrt{2}}\right)}{4}\right]\right). \quad (70)$$

$W'_{2,1}(\phi)$ and $W'_{2,2}(\phi)$ can also be calculated in a similar manner. It is then not difficult to see that all $W'_{i,j}(\phi) = 0$ when $\phi = \pi/4$. Therefore, $\frac{dS(\phi)}{d\phi} = 0$. It can also be verified that $\frac{d^2S(\phi)}{d^2\phi} = 0$ when $\phi = \pi/4$. Thus, $\phi = \pi/4$ is a solution of (60). As a result, QPSK is a capacity-achieving scheme. The channel capacity can then be calculated in closed-form as:

$$C = \log\left[4\sum_{k=1}^{M}\varepsilon_k\left(1-Q\left(\frac{|x|}{\sigma_k}\right)\right)^2\right]$$
$$+2\left(\sum_{k=1}^{M}\varepsilon_k Q\left(\frac{|x|}{\sigma_k}\right)\right)\log\left[\frac{\sum_{k=1}^{M}\varepsilon_k\left(Q\left(\frac{|x|}{\sigma_k}\right)-Q\left(\frac{|x|}{\sigma_k}\right)^2\right)}{\sum_{k=1}^{M}\varepsilon_k\left(1-Q\left(\frac{|x|}{\sigma_k}\right)\right)}\right]$$
$$+\left(\sum_{k=1}^{M}\varepsilon_k Q\left(\frac{|x|}{\sigma_k}\right)^2\right)\log\left[\frac{\sum_{k=1}^{M}\varepsilon_k Q\left(\frac{|x|}{\sigma_k}\right)^2\sum_{k=1}^{M}\varepsilon_k\left(1-Q\left(\frac{|x|}{\sigma_k}\right)\right)^2}{\left\{\sum_{k=1}^{M}\varepsilon_k\left(Q\left(\frac{|x|}{\sigma_k}\right)-Q\left(\frac{|x|}{\sigma_k}\right)^2\right)\right\}^2}\right]. \quad (71)$$

5. Numerical Results

In this section, numerical results are provided to demonstrate the optimality of QPSK signaling scheme. While the above analysis applies to any GM channel, for simplicity, we will only consider a 2-term GM channel and a 3-term GM channel with the following parameters: 1) 2-term GM with $\epsilon_1 = 0.2$, $\epsilon_2 = 0.8$, $\sigma_1^2 = 0.6643$, and $\sigma_2^2 = 1.3286$); and 2) 3-term GM with $\epsilon_1 = 0.2$, $\epsilon_2 = 0.3$, $\epsilon_3 = 0.5$, $\sigma_1^2 = 0.5296$, $\sigma_2^2 = 0.5296$, and $\sigma_3^2 = 1.0591$.

Let first examine the KTC. Figs. 1 and 2 show the KTC over the complex plane for the 2-term and 3-term GM channels, respectively, when the signal-to-noise ratio (SNR) is set at 1dB. It is clear in both cases that the KTC is zero at four points that constitute the QPSK signal.

To demonstrate the information-theoretical superiority of QPSK, Fig. 3 shows the information rates achieved by different input signaling schemes obtained by Monte Carlo simulations. These include QPSK, 8-PSK, 16-QAM, as well as a Gaussian input. The channel capacity calculated using (71) is also provided. It is clear from Fig. 3 that the information rate achieved by QPSK is same with the channel capacity, and it is significantly higher than those achieved by other input signals over a wide range of SNRs.

Similar results can also be obtained for the 3-term GM channel, which are shown in Fig. 4.

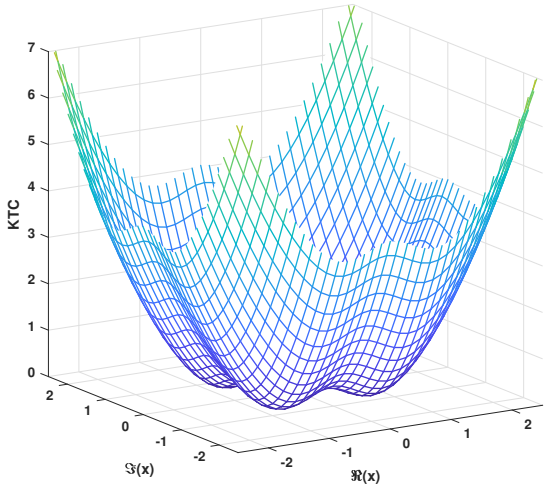

Figure 1. KTC at SNR=1dB for the 2-term GM channel.

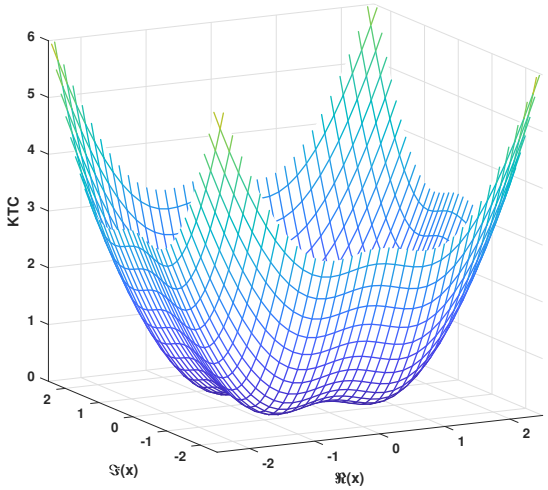

Figure 2. KTC at SNR=1dB for the 3-term GM channel.

6. Conclusion

In this work, we have explicitly addressed the capacity-achieving signaling scheme and the capacity of an additive Gaussian mixture (GM) noise channel employing 1-bit

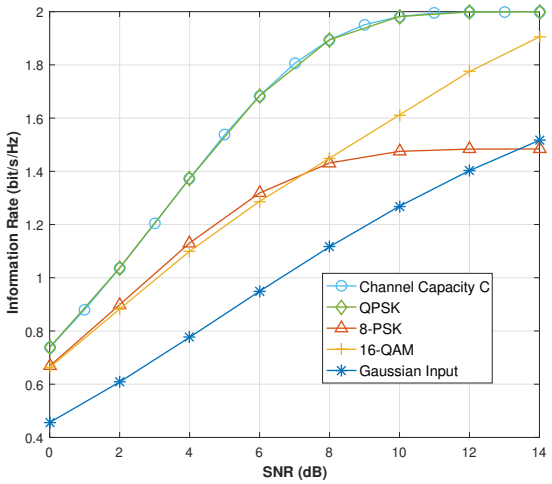

Figure 3. Information rates achieved by various signaling schemes and the channel capacity of the 2-term GM channel.

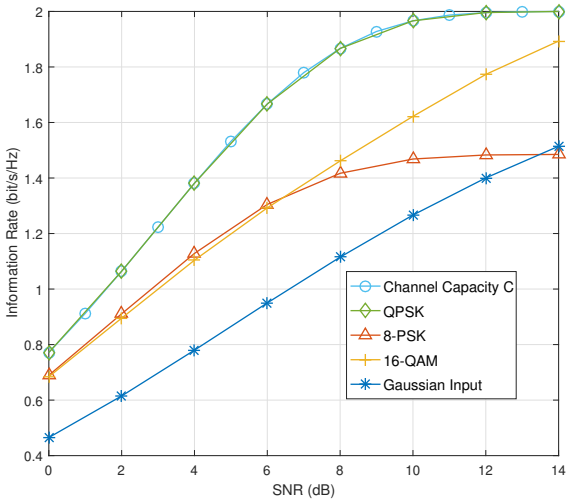

Figure 4. Information rates achieved by various signaling schemes and the channel capacity of the 3-term GM channel.

analog-to-digital converters (ADCs). Towards this end, it was first shown that an optimal input is $\pi/2$ circularly symmetric. Then by establishing and examining the KTC on the capacity-achieving input, we demonstrated that this input is unique, and it consists of exactly four mass points. It was then demonstrated that the phase of the optimal mass point in the first quadrant is $\pi/4$. Thus, QPSK is the capacity-achieving signal. Numerical results were also provided to confirm the optimality of QPSK signal.

References

[1] "European technology platform for communication networks (emobility) newsletter," 2008.

[2] Greentouch, "ICT industry combats climate change."

[3] J. Zhang, L. Dai, X. Li, Y. Liu, and L. Hanzo, "On low-resolution ADCs in practical 5G millimeter-wave massive MIMO systems," *IEEE Commun. Magazine*, vol. 56, pp. 205–211, July 2018.

[4] J. Singh, O. Dabeer, and U. Madhow, "On the limits of communication with low-precision analog-to-digital conversion at the receiver," *IEEE Trans. Commun.*, vol. 57, pp. 3629–3639, December 2009.

[5] J. Mo and R. W. Heath, "Capacity analysis of one-bit quantized MIMO systems with transmitter channel state information," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5498–5512, 2015.

[6] J. Mo and R. W. Heath, "High SNR capacity of millimeter wave MIMO systems with one-bit quantization," in *2014 Information Theory and Applications Workshop (ITA)*, pp. 1–5, Feb 2014.

[7] S. Krone and G. Fettweis, "Fading channels with 1-bit output quantization: Optimal modulation, ergodic capacity and outage probability," *Proc. IEEE Inform. Theory Workshop*, pp. 1–5, Aug 2010.

[8] A. Mezghani and J. A. Nossek, "On ultra-wideband MIMO systems with 1-bit quantized outputs: Performance analysis and input optimization," in *IEEE International Symposium on Information Theory*, pp. 1286–1289, June 2007.

[9] A. Mezghani and J. A. Nossek, "Analysis of Rayleigh-fading channels with 1-bit quantized output," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, pp. 260–264, July 2008.

[10] M. Vu, G. Dissanayakage, N. H. Tran, and K. Pham, "Optimal signaling scheme and capacity of non-coherent Rician fading channels with 1-bit output quantization," in *IEEE International Conference on Communications (ICC) - Signal Processing Symposium*, pp. 1–6, May 2018.

[11] M. Vu, N. H. Tran, G. Dissanayakage, K. Pham, K.-S. Lee, and D. H. N. Nguyen, "Optimal signaling schemes and capacity of non-coherent Rician fading channels with low-resolution output quantization," *IEEE Trans. Wireless Commun.*, vol. 18, pp. 2989–3004, June 2019.

[12] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, "Throughput analysis of massive MIMO uplink with low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, pp. 4038–4051, June 2017.

[13] S. Jacobsson, G. Durisi, M. Coldrey, T. Goldstein, and C. Studer, "Quantized precoding for massive MU-MIMO," *IEEE Trans. on Commun.*, vol. 65, pp. 4670–4684, Nov. 2017.

[14] Y. Li, C. Tao, G. Seco-Granados, A. Mezghani, A. L. Swindlehurst, and L. Liu, "Channel estimation and performance analysis of one-bit massive MIMO systems," *IEEE Trans. Signal Processing*, vol. 65, pp. 4075–4089, Aug 2017.

[15] M. Ranjbar, M. Vu, N. H. Tran, K. Pham, and D. H. N. Nguyen, "On the sum-capacity-achieving distributions and sum-capacity of 1-bit ADC multiple access channels in Rayleigh fading," in *IEEE International Conference on Communications (ICC) - Wireless Commun. Symposium*, pp. 1–6, May 2019.

[16] A. Khalili, F. Shirani, E. Erkip, and Y. Eldar, "On multiterminal communication over MIMO channels with one-bit ADCs at the

receivers," in *IEEE International Symposium on Information Theory*, pp. 602–606, July 2019.

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

[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, pp. 6207–6222, Dec. 2010.

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

[20] 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, pp. 1642–1655, Mar. 2014.

[21] P. C. Pinto and M. Z. Win, "Communication in a Poisson field of interferers–Part I: Interference distribution and error probability," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 2176–2186, July 2010.

[22] 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, pp. 3607–3620, Oct. 2015.

[23] A. D. 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, pp. 5011–5024, Dec. 2016.

[24] I. Abou-Faycal, M. Trott, and S. Shamai, "The capacity of discrete-time memoryless Rayleigh-fading channels," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1290–1301, May 2001.

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