Abstract

This paper addresses the capacity region of a two-transmitter Gaussian multiple access channel (MAC) under average input power constraints and a zero-threshold quantizer at the receiver. It is proved that the input distributions that achieve the boundary points of the capacity region are discrete. Based on the position of a boundary point, which is determined by the slope of the line tangent to the capacity region at this point, upper bounds on the number of the mass points of the corresponding distributions are proposed.

Index Terms

Gaussian multiple access channel, quantizer, capacity region.

I. INTRODUCTION

The energy consumption of an analog-to-digital converter (ADC) (measured in Joules/sample) grows exponentially with its resolution (in bits/sample) [1], [2]. When the available power is limited, for example, for mobile devices with limited battery capacity, or for wireless receivers that operate on limited energy harvested from ambient sources [3], the receiver circuitry may be constrained to operate with low resolution ADCs. The presence of a low-resolution ADC, in particular a one-bit ADC at the receiver, alters the channel characteristics significantly. Such a constraint not only limits the fundamental bounds on the achievable rate, but it also changes the nature of the communication and modulation schemes approaching these bounds. For example, in a real additive white Gaussian noise (AWGN) channel under an average power constraint on the input, if the receiver is equipped with a \( K \)-bit ADC front end, the capacity-achieving input distribution is discrete with at most \( K + 1 \) mass points [4]. This is in contrast with the optimality of the Gaussian input distribution when the receiver has infinite resolution.

Especially with the adoption of massive multiple-input multiple-output (MIMO) receivers and the millimeter wave (mmWave) technology enabling communication over large bandwidths, communication systems with limited-resolution receiver front ends are becoming of practical importance. Accordingly, there have been a growing research interest in understanding both the fundamental information theoretic limits and the design of practical communication protocols for systems with finite-resolution ADC front ends. In [5], the authors show that for a Rayleigh fading channel with a one-bit ADC and perfect channel state information at the receiver (CSIR), quadrature phase shift...
keying (QPSK) modulation is capacity-achieving. In case of no CSIR, [6] shows that (QPSK) modulation is optimal when the signal-to-noise (SNR) ratio is above a certain threshold, which depends on the coherence time of the channel, while for SNRs below this threshold, on-off QPSK achieves the capacity. For the point-to-point multiple-input multiple-output (MIMO) channel with a one-bit ADC front end at each receive antenna and perfect CSIR, [7] shows that QPSK is optimal at very low SNRs, while with perfect channel state information at the transmitter (CSIT), upper and lower bounds on the capacity are provided in [8].

To the best of our knowledge, the existing literature on communications with low-resolution ADCs focus exclusively on point-to-point systems. Our goal in this paper is to understand the impact of low-resolution ADCs on the capacity region of a multiple access channel (MAC). In particular, we consider a two-transmitter Gaussian MAC with a one-bit quantizer at the receiver. The inputs to the channel are subject to average power constraints. We show that any point on the boundary of the capacity region is achieved by discrete input distributions. Based on the slope of the tangent line to the capacity region at a boundary point, upper bounds on the cardinality of the support of these distributions are proposed.

The paper is organized as follows. Section II introduces the system model. In section III, the capacity region of a general two-transmitter memoryless MAC under input average power constraints is investigated. Through an example, it is shown that when there is input average power constraint, it is necessary to consider the capacity region with the auxiliary random variable $U$ in general. The main result of the paper is given in section ?? followed by its proof in section IV. Finally, section V concludes the paper.

**Notations.** Random variables are denoted by capital letters, while their realizations with lower case letters. $F_X(x)$ denotes the cumulative distribution function (CDF) of random variable $X$. The conditional probability mass function (pmf) $p_{Y|X_1,X_2}(y|x_1,x_2)$ will be written as $p(y|x_1,x_2)$. For integers $m \leq n$, we have $\{m:n\} = \{m,m+1,\ldots,n\}$. $H_b(t) = -t \log_2 t - (1-t) \log_2(1-t)$ denotes the binary entropy function. The unit-step function is denoted by $s(.)$.

**II. System Model and Preliminaries**

We consider a two-transmitter memoryless Gaussian MAC (as shown in Figure 1) with a one-bit quantizer $\Gamma$ at the receiver front end. Transmitter $j = 1, 2$ encodes its message $W_j$ into a codeword $X^n_j$ and transmits it over the shared channel. The signal received by the decoder is given by

$$Y^n = \Gamma(X^n_1 + X^n_2 + Z^n),$$

where $\{Z_i\}_{i=1}^n$ is an independent and identically distributed (i.i.d.) Gaussian noise process, also independent of the channel inputs $X^n_1$ and $X^n_2$ with $Z_i \sim \mathcal{N}(0,1), i \in \{1:n\}$. $\Gamma$ represents the one-bit ADC operation given by

$$\Gamma(x) = \begin{cases} 
1 & x \geq 0 \\
0 & x < 0 
\end{cases}.$$ 

This channel can be modelled by the triplet $(X_1 \times X_2, p(y|x_1,x_2), \mathcal{Y})$, where $X_1, X_2 (= \mathbb{R})$ and $\mathcal{Y} (= \{0,1\})$. 

Fig. 1: A two-transmitter Gaussian MAC with a one-bit quantizer at the receiver.

respectively, are the alphabets of the inputs and the output. The conditional pmf of the channel output $Y$ conditioned on the channel inputs $X_1$ and $X_2$ (i.e. $p(y|x_1, x_2)$) is characterized by

$$p(0|x_1, x_2) = 1 - p(1|x_1, x_2) = Q(x_1 + x_2),$$

(1)

where $Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_{x}^{+\infty} e^{-\frac{t^2}{2}} dt$.

Upon receiving the sequence $Y^n$, the decoder finds the estimates $(\hat{W}_1, \hat{W}_2)$ of the messages.

A $(2^{nR_1}, 2^{nR_2}, n)$ code for this channel consists of (as in [9])

- two message sets $[1 : 2^{nR_1}]$ and $[1 : 2^{nR_2}]$,
- two encoders, where encoder $j = 1, 2$ assigns a codeword $x_j^n(w_j)$ to each message $w_j \in [1 : 2^{nR_j}]$, and
- a decoder that assigns estimates $(\hat{w}_1, \hat{w}_2) \in [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$ or an error message to each received sequence $y^n$.

We assume that the message pair $(W_1, W_2)$ is uniformly distributed over $[1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$. The average probability of error is defined as

$$P_e^{(n)} = \Pr \left\{ (\hat{W}_1, \hat{W}_2) \neq (W_1, W_2) \right\}$$

(2)

Average power constraints are imposed on the channel inputs as

$$\frac{1}{n} \sum_{i=1}^{n} x_{j,i}^2(w_j) \leq P_j, \quad \forall m_j \in [1 : 2^{nR_j}], j \in \{1, 2\},$$

where $x_{j,i}(w_j)$ denotes the $i$th element of the codeword $x_j^n(w_j)$.

A rate pair $(R_1, R_2)$ is said to be achievable for this channel if there exists a sequence of $(2^{nR_1}, 2^{nR_2}, n)$ codes (satisfying the average power constraints) such that $\lim_{n \to \infty} P_e^{(n)} = 0$. The capacity region $\mathcal{C}(P_1, P_2)$ of this channel is the closure of the set of achievable rate pairs $(R_1, R_2)$. 
III. MAIN RESULTS

**Proposition 1.** The capacity region \( \mathcal{C}(P_1, P_2) \) of a two-transmitter memoryless MAC with average power constraints \( P_1 \) and \( P_2 \) is the set of non-negative rate pairs \((R_1, R_2)\) that satisfy

\[
R_1 \leq I(X_1; Y|X_2, U), \\
R_2 \leq I(X_2; Y|X_1, U), \\
R_1 + R_2 \leq I(X_1, X_2; Y|U),
\]

for some \( F_U(u)F_{X_1|U}(x_1|u)F_{X_2|U}(x_2|u) \), such that \( \mathbb{E}[X_j^2] \leq P_j, \ j = 1, 2 \). Also, it is sufficient to consider \(|U| \leq 5\).

It is obvious that for a fixed joint distribution \( F_U(u)F_{X_1|U}(x_1|u)F_{X_2|U}(x_2|u) \), the region in (3) is a pentagon.

**Proof.** The capacity region of the discrete memoryless (DM) MAC with input cost constraints has been addressed in Exercise 4.8 of [9]. If the input alphabets are not discrete, the capacity region is still the same because: 1) the converse remains the same if the inputs are from a continuous alphabet; 2) the region is achievable by coded time sharing and the discretization procedure (see Remark 3.8 in [9]). Therefore, it is sufficient to show the cardinality bound \(|U| \leq 5\).

Let \( \mathcal{P} \) be the set of all product distributions on \( \mathbb{R}^2 \). Let \( g : \mathcal{P} \to \mathbb{R}^5 \) be a vector-valued mapping defined element-wise as

\[
\begin{align*}
g_1(F_{X_1}(x_1|u)F_{X_2}(x_2|u)) &= I(X_1; Y|X_2, U = u), \\
g_2(F_{X_1}(x_1|u)F_{X_2}(x_2|u)) &= I(X_2; Y|X_1, U = u), \\
g_3(F_{X_1}(x_1|u)F_{X_2}(x_2|u)) &= I(X_1, X_2; Y|U = u), \\
g_4(F_{X_1}(x_1|u)F_{X_2}(x_2|u)) &= \mathbb{E}[X_1^2|u], \\
g_5(F_{X_1}(x_1|u)F_{X_2}(x_2|u)) &= \mathbb{E}[X_2^2|u].
\end{align*}
\]

Let \( \mathcal{G} \subset \mathbb{R}^5 \) be the image of \( \mathcal{P} \) under the mapping \( g \) (i.e., \( \mathcal{G} = g(\mathcal{P}) \)). Note that although \( g \) is a continuous mapping, \( \mathcal{G} \) is not necessarily compact\(^1\) (since \( g_4, g_5 \) cannot be bounded). Given an arbitrary \((U, X_1, X_2) \sim F_UF_{X_1|U}F_{X_2|U}\), we obtain the vector \( r \) as

\[
\begin{align*}
r_1 &= I(X_1; Y|X_2, U) = \int_U I(X_1; Y|X_2, U = u)dF_U(u), \\
r_2 &= I(X_2; Y|X_1, U) = \int_U I(X_2; Y|X_1, U = u)dF_U(u), \\
r_3 &= I(X_1, X_2; Y|U) = \int_U I(X_1, X_2; Y|U = u)dF_U(u), \\
r_4 &= \mathbb{E}[X_1^2] = \int_U \mathbb{E}[X_1^2|u]dF_U(u), \\
r_5 &= \mathbb{E}[X_2^2] = \int_U \mathbb{E}[X_2^2|u]dF_U(u).
\end{align*}
\]

\(^1\)Note that compactness is not a necessary condition here, as stated in the footnote on page 267 of [10].
Therefore, r is in the convex hull of $\mathcal{G} \subset \mathbb{R}^5$. By Carathéodory’s theorem [10], r is a convex combination of 6 ($= 5 + 1$) or fewer points in $\mathcal{G}$, which states that it is sufficient to consider $|U| \leq 6$. Since $\mathcal{P}$ is a connected set and the mapping $g$ is continuous, $\mathcal{G}$ is a connected subset of $\mathbb{R}^5$. Therefore, connectedness of $\mathcal{G}$ refines the cardinality of $U$ to $|U| \leq 5$.

**Lemma 1.** For the boundary points of $\mathcal{C}(P_1, P_2)$ that are not sum-rate optimal, it is sufficient to have $|U| \leq 4$.

**Proof.** Any point on the boundary of the capacity region that does not maximize $R_1 + R_2$, is either of the form $(I(X_1; Y|X_2, U), I(X_2; Y|U))$ or $(I(X_1; Y|U), I(X_2; Y|X_1, U))$ for some $F_UF_{X_1|U}F_{X_2|U}$ that satisfies $E[X_j^2] \leq P_j, j = 1, 2$. In other words, it is one of the corner points of the corresponding pentagon in (3). As in the proof of Proposition 1, define the mapping $g : \mathcal{P} \to \mathbb{R}^4$, where $g_1$ and $g_2$ are the coordinates of this boundary point conditioned on $U = u$, and $g_3$, $g_4$ are the same as $g_4$ and $g_5$ in (4), respectively. The sufficiency of $|U| \leq 4$ in this case follows similarly to the proof of Proposition 1.

When there is no input cost constraint, the capacity region of the MAC can be characterized either through the convex hull operation as in [9, Theorem 4.2], or with the introduction of an auxiliary random variable as in [9, Theorem 4.3]. The following remark states that when there is an input cost constraint, the capacity region has only the computable characterization with the auxiliary random variable.

**Remark 2.** Let $(X_1, X_2) \sim F_{X_1}(x_1)F_{X_2}(x_2)$ such that $E[X_j^2] \leq P_j, j = 1, 2$. Let $\mathcal{R}(P_1, P_2)$ denote the set of non-negative rate pairs $(R_1, R_2)$ such that

$$R_1 \leq I(X_1; Y|X_2),$$

$$R_2 \leq I(X_2; Y|X_1),$$

$$R_1 + R_2 \leq I(X_1, X_2; Y).$$

Let $\mathcal{R}_1(P_1, P_2)$ be the convex closure of $\bigcup_{F_{X_1}, F_{X_2}} \mathcal{R}(P_1, P_2)$, where the union is over all product distributions that satisfy the average power constraints.

Let $\mathcal{R}_2(P_1, P_2)$ be the set of non-negative rate pairs $(R_1, R_2)$ such that

$$R_1 \leq I(X_1; Y|X_2, U),$$

$$R_2 \leq I(X_2; Y|X_1, U),$$

$$R_1 + R_2 \leq I(X_1, X_2; Y|U)$$

for some $F_U(u)F_{X_1|U}(x_1|u)F_{X_2|U}(x_2|u)$ that satisfies $E[X_j^2|u] \leq P_j, j = 1, 2, \forall u$.

It can be verified that $\mathcal{R}_1(P_1, P_2) = \mathcal{R}_2(P_1, P_2)$. By comparing $\mathcal{R}_2(P_1, P_2)$ to the capacity region $\mathcal{C}(P_1, P_2)$, we can conclude that $\mathcal{R}_2(P_1, P_2) \subseteq \mathcal{C}(P_1, P_2)$. This follows from the fact that in the region $\mathcal{R}_2(P_1, P_2)$, the average power constraint $E[X_j^2|u] \leq P_j$ holds for every realization of the auxiliary random variable $U$, which is a stronger condition than $E[X_j^2] \leq P_j$ used in the capacity region. The following example shows that $\mathcal{R}_1(P_1, P_2)$ and $\mathcal{R}_2(P_1, P_2)$ can be strictly smaller than $\mathcal{C}(P_1, P_2)$.
Consider the same Gaussian MAC with one-bit quantizer at the receiver (as depicted in Figure 1) with the following changes: i) $\mathcal{X}_1 = \mathcal{X}_2 = \{-\sqrt{2}, 0, \sqrt{2}\}$, ii) Besides the average power constraints of $P_1 = P_2 = 1$, we also impose per codeword average cost constraints as $\frac{1}{n} \sum_{i=1}^{n} x_{j,i}(m_j) = 0, \forall m_j \in [1: 2^{nR_i}], j \in \{1, 2\}$.

The capacity region of this channel is the set of non-negative rate pairs $(R_1, R_2)$ such that (3) holds for some $F_i(u)F_{X_1|U}(x_1|u)F_{X_2|U}(x_2|u)$ which satisfies $\mathbb{E}[X_j^2] \leq P_j$, $\mathbb{E}[X_j] = 0$, $j = 1, 2$. Also, let $\mathcal{R}_1$ be the rate region in Remark 2 with the additional constraints $\mathbb{E}[X_j] = 0$, $j = 1, 2$.

In order to show that $\mathcal{R}_1$ can be strictly smaller than the capacity region, we show that there exists a point in the capacity region which is not in $\mathcal{R}_1$. We have,

$$\max_{(R_1, R_2) \in \mathcal{R}_1} R_1 + R_2 = \max_{F_{X_1}, F_{X_2}; \mathbb{E}[X_j^2] \leq 1, \mathbb{E}[X_j] = 0, j = 1, 2} I(X_1, X_2; Y)$$

$$= \max_{F_{X_1}, F_{X_2}; \mathbb{E}[X_j^2] \leq 1, \mathbb{E}[X_j] = 0, j = 1, 2} I(X_1 + X_2; Y)$$

$$\leq \max_{F_{X}; \mathbb{E}[X^2] \leq 2} I(X; Y)$$

$$= 1 - H_b(Q(\sqrt{2}))$$

$$\leq \max_{F_{U}, F_{X_1|U}, F_{X_2|U}; \mathbb{E}[X_j^2] \leq 1, \mathbb{E}[X_j] = 0, j = 1, 2} I(X_1, X_2; Y|U)$$

$$= \max_{(R_1, R_2) \in \mathcal{R}} R_1 + R_2,$$

where (5) is due to the fact that $X_1 + X_2$ is a function of the pair $(X_1, X_2)$, and the following Markov chain holds: $(X_1, X_2) \rightarrow X_1 + X_2 \rightarrow Y$. In (6), we use the inequality $\mathbb{E}[(X_1 + X_2)^2] = \mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] \leq 2$, since $X_1$ and $X_2$ are independent and zero mean. Also, the channel from $X$ to $Y$ is characterized by the conditional distribution $p_{Y|X}(y|x) \sim \text{Bern}(Q(x))$. (7) is due to [4], where the maximum is shown to be achieved by the CDF $F_X^*(x) = \frac{1}{2} s(x + \sqrt{2}) + \frac{1}{2} s(x - \sqrt{2})$, where $s(\cdot)$ is the unit step function. Let $U \sim \text{Bern}(\frac{1}{2})$, $F_{X_1|U}(x|1) = F_{X_2|U}(x|0) = \frac{1}{2} s(x + \sqrt{2}) + \frac{1}{2} s(x - \sqrt{2})$ and $F_{X_1|U}(x|0) = F_{X_2|U}(x|1) = s(x)$. For this joint distribution on $(U, X_1, X_2)$, we have $\mathbb{E}[X_j] = 0$, $\mathbb{E}[X_j^2] \leq 1$, $j = 1, 2$ and $I(X_1, X_2; Y|U) = 1 - H_b(Q(\sqrt{2}))$, which results in (8).

In what follows, it is proved that the inequality in (6) is strict. In other words, the sum rate of $1 - H_b(Q(\sqrt{2}))$ cannot be obtained by any rate pair in $\mathcal{R}_1$, while it is in the capacity region. Let $\tilde{X} = X_1 + X_2$, where $X_1$ and $X_2$ are two zero-mean independent random variables on $\mathcal{X}_1 = \mathcal{X}_2$ satisfying the average power constraint $\mathbb{E}[X_j^2] \leq 1, j = 1, 2$. We show that the minimum Lévy distance$^2$ between $F_{\tilde{X}}^*(x)$ and all the distributions $F_{X_j}(x)$

$^2$The Lévy distance between two distributions $F, G : \mathbb{R} \rightarrow [0, 1]$ is defined as $d_L(F, G) = \inf \{ \epsilon > 0 : F(x - \epsilon) - \epsilon \leq G(x) \leq F(x + \epsilon) + \epsilon, \forall x \in \mathbb{R} \}$. 


(induced by $F_{X_1}, F_{X_2}$) is bounded away from zero. Since $E[X] = 0$ and $E[X^2] \leq 1$, the distribution of $X_1$ is $F_{X_1}(x) = ps(x + \sqrt{2}) + (1 - p)s(x) + ps(x - \sqrt{2})$ with $p \leq \frac{1}{4}$. The same applies to $F_{X_2}$ with parameter $p'(\leq \frac{1}{4})$.

The distribution of $\tilde{X}$ induced by $F_{X_1}, F_{X_2}$ is given by

$$F_{\tilde{X}}(x) = pp's(x + 2\sqrt{2}) + p*p's(x + \sqrt{2}) + (1 - 2(pp' + p*p'))s(x) + pp's(x - \sqrt{2}) + pp's(x - 2\sqrt{2}),$$

where $p*p' \triangleq p(1 - p') + p'(1 - p)$ denotes the binary convolution operation. Let $\tilde{F}$ be the set of all distributions on $\tilde{X}$ obtained in this way. It can be easily verified that (see Figure 2) for any given $p, p' \leq \frac{1}{4}$, the Lévy distance between $F_{\tilde{X}}$ and $F_*$ is

$$d_L(F_{\tilde{X}}, F_*) = \max\{pp', \frac{1}{2} - pp' - p*p'\} = \frac{1}{2} - pp' - p*p'. \quad (9)$$

Subsequently,

$$\min_{p,p' \leq \frac{1}{4}} d_L(F_{\tilde{X}}, F_*) = \frac{1}{16}.$$

This shows that there is a neighborhood of $F_{\tilde{X}}$ whose intersection with $\tilde{F}$ is empty. Note that any neighborhood with radius less than $\frac{1}{16}$ has this property. Combined with the facts that the mutual information is continuous and $F_*$ is the unique solution, it proves that the inequality in (6) is strict. Therefore, $\mathcal{R}_1 (= \mathcal{R}_2)$ is smaller than the capacity region in general.

The main result of this paper is provided in the following theorem. It bounds the cardinality of the support set of the capacity achieving distributions.

**Theorem 1.** Let $P$ be an arbitrary point on the boundary of the capacity region $\mathcal{C}(P_1, P_2)$ of the memoryless MAC with a one-bit ADC front end (as shown in Figure 1) achieved by $F_{U}^P(u|U), F_{X_1|U}^P(x_1|u)F_{X_2|U}^P(x_2|u)$. Let $l$ be the slope of the line tangent to the capacity region at this point. For any $u \in U$, the conditional input distributions

![Fig. 2: The distributions of $X^*$ and $\tilde{X}$.](image-url)
(11), the maximizations are over all product distributions of the form \( p(U|P_1, P_2) \), can be written as \[
(R_1^b, R_2^b) = \arg \max_{(R_1, R_2) \in \mathcal{C}(P_1, P_2)} R_1 + \lambda R_2, \]
for some \( \lambda > 0 \).
Any rate pair \((R_1, R_2) \in \mathcal{C}(P_1, P_2)\) is within the pentagon defined by (3) for some \( F_{X_1|U}F_{X_2|U} \) that satisfies the power constraints. Therefore, due to the structure of the pentagon, the problem of finding the boundary points is equivalent to the following maximization problem.

(11)

where on the right hand side (RHS) of (11), the maximizations are over all \( F_{X_1|U}F_{X_2|U} \) that satisfy the power constraints. It is obvious that when \( \lambda = 1 \), the two lines in (11) are the same, which results in the sum capacity.

For any product of distributions \( F_{X_1}F_{X_2} \) and the channel in (1), let \( I_\lambda \) be defined as

(12)

With this definition, (11) can be written as

\[
\max_{(R_1, R_2) \in \mathcal{C}(P_1, P_2)} R_1 + \lambda R_2 = \max_{i=1}^5 p_U(u_i) I_\lambda(F_{X_1|U}(x_1|u_i)F_{X_2|U}(x_2|u_i)),
\]
where the maximization is over product distributions of the form \( p_U(u)F_{X_1|U}(x_1|u)F_{X_2|U}(x_2|u) \), \(|U| \leq 5 \), such that

\[
\sum_{i=1}^5 p_U(u_i)E[X_j^2|u_i] \leq P_j, \quad j = 1, 2.
\]

**Proposition 2.** For a given \( F_{X_1} \) and any \( \lambda > 0 \), \( I_\lambda(F_{X_1}, F_{X_2}) \) is a concave, continuous and weakly differentiable function of \( F_{X_2} \). In the statement of this Proposition, \( F_{X_1} \) and \( F_{X_2} \) could be interchanged.

\footnote{A point \( Z \) is said to be a point of increase of a distribution if for any open set \( \Omega \) containing \( Z \), we have \( \Pr(\Omega) > 0 \).}
Proof. The proof is provided in Appendix A. □

**Proposition 3.** Let $P'_1, P'_2$ be two arbitrary non-negative finite real numbers. For the following problem

\[
\max_{F_{X_1}, F_{X_2}} I_\lambda(F_{X_1}, F_{X_2}),
\]

the optimal inputs $F^*_1$ and $F^*_2$, which are not unique in general, have the following properties,

(i) The support sets of $F^*_1$ and $F^*_2$ are bounded subsets of $\mathbb{R}$.

(ii) $F^*_1$ and $F^*_2$ are discrete distributions that have at most $n_1$ and $n_2$ points of increase, respectively, where 

\[
(n_1, n_2) = \begin{cases} 
(3, 5) & 0 < \lambda < 1 \\
(3, 3) & \lambda = 1 \\
(5, 3) & \lambda > 1 
\end{cases}
\]

Proof. We start with the proof of the first claim. Assume that $0 < \lambda \leq 1$, and $F_{X_2}$ is given. Consider the following optimization problem:

\[
I^*_{F_{X_2}} \triangleq \sup_{F_{X_1}} I_\lambda(F_{X_1}, F_{X_2}).
\]

Note that $I^*_{F_{X_2}} < +\infty$, since for any $\lambda > 0$, from (12),

\[
I_\lambda \leq (\lambda + 1)H(Y) \leq (1 + \lambda) < +\infty.
\]

From Proposition 2, $I_\lambda$ is a continuous, concave function of $F_{X_1}$. Also, the set of all CDFs with bounded second moment (here, $P'_1$) is convex and compact\(^4\). Therefore, the supremum in (14) is achieved by a unique distribution. Since for any $F_{X_1}(x) = s(x - x_0)$ with $|x_0|^2 < P'_1$, we have $E[X_1^2] < P'_1$, the Lagrangian theorem and the Karush-Kuhn-Tucker conditions state that $F^*_1$ achieves the maximum $I^*_{F_{X_2}}$ if and only if there exists a $\theta_1 \geq 0$ such that

\[
I^*_{F_{X_2}} = \sup_{F_{X_1}} \left\{ I_\lambda(F_{X_1}, F_{X_2}) - \theta_1 \left( \int x^2 dF_{X_1}(x) - P'_1 \right) \right\}
\]

\[
= I_\lambda(F^*_1, F_{X_2}) - \theta_1 \left( \int x^2 dF^*_1(x) - P'_1 \right),
\]

and

\[
\theta_1 \left( \int x^2 dF^*_1(x) - P'_1 \right) = 0.
\]

**Lemma 1.** The Lagrangian multiplier $\theta_1$ is nonzero.

Proof. Having a zero Lagrangian multiplier means the power constraint is inactive. In other words, if $\theta_1 = 0$, (14) and (15) imply that

\[
\sup_{F_{X_1}} \frac{1}{E[X_1^2]} I_\lambda(F_{X_1}, F_{X_2}) = \sup_{F_{X_1}} I_\lambda(F_{X_1}, F_{X_2}).
\]

We prove that (18) does not hold by showing that its left hand side (LHS) is strictly lower than 1, while its RHS equals 1. The details are provided in Appendix B. □

\(^4\)The compactness follows from [11, Appendix I].
Proof. The proof is provided in Appendix C.
Note that
\[
\lim_{x_1 \to +\infty} \int_{-\infty}^{+\infty} D \left( p(y|x_1, x_2) \| p(y; F_{X_1}^*, F_{X_2}) \right) dF_{X_2}(x_2) = \int_{-\infty}^{+\infty} \lim_{x_1 \to +\infty} D \left( p(y|x_1, x_2) \| p(y; F_{X_1}^*, F_{X_2}) \right) dF_{X_2}(x_2)
\]
\[
= - \log p_Y(1; F_{X_1}^*, F_{X_2})
\]
\[
\leq - \log Q(\sqrt{P_1^*} + \sqrt{P_2^*}),
\]
where (28) is due to Lebesgue dominated convergence theorem [12] and (25), which permit the interchange of the limit and the integral; and (29) is obtained from (57) in Appendix C. Furthermore,
\[
\lim_{x_1 \to +\infty} \int_{-\infty}^{+\infty} \sum_{y=0}^{1} p(y|x_1, x_2) \log \frac{p(y; F_{X_1}^*, F_{X_2})}{p(y; F_{X_1}^*, F_{X_2})} dF_{X_2}(x_2) = \int_{-\infty}^{+\infty} \lim_{x_1 \to +\infty} \sum_{y=0}^{1} p(y|x_1, x_2) \log \frac{p(y; F_{X_1}^*, F_{X_2})}{p(y; F_{X_1}^*, F_{X_2})} dF_{X_2}(x_2)
\]
\[
= \log p_Y(1; F_{X_1}^*, F_{X_2}) - \int_{-\infty}^{+\infty} \log p(1; F_{X_1}^*, F_{X_2}) dF_{X_2}(x_2)
\]
\[
< - \log Q(\sqrt{P_1^*} + \sqrt{P_2^*}),
\]
where (30) is due to Lebesgue dominated convergence theorem along with (27); (31) is from (26) and convexity of \( \log Q(\alpha + \sqrt{t}) \) in \( t \) when \( \alpha \geq 0 \) (see Appendix D).

Therefore, from (29) and (31),
\[
\lim_{x_1 \to +\infty} \tilde{i}_\lambda(x_1; F_{X_1}^*, F_{X_2}) \leq -(1 + \lambda) \log Q(\sqrt{P_1^*} + \sqrt{P_2^*}) < +\infty.
\]
Using a similar approach, we can also obtain
\[
\lim_{x_1 \to -\infty} \tilde{i}_\lambda(x_1; F_{X_1}^*, F_{X_2}) \leq -(1 + \lambda) \log Q(\sqrt{P_1^*} + \sqrt{P_2^*}) < +\infty.
\]
From (32), (33) and the fact that \( \theta_1 > 0 \) (see Lemma 1), the left hand side of (24) goes to \(-\infty\) when \( |x_1| \to +\infty \). Since any point of increase of \( F_{X_1}^* \) must satisfy (24) with equality, and \( I_{F_{X_2}}^* \geq 0 \), it is proved that \( F_{X_1}^* \) has a bounded support, i.e., \( X_1 \in [A_1, A_2] \) for some \( A_1, A_2 \in \mathbb{R} \).\(^5\)

Similarly, for a given \( F_{X_1} \), the optimization problem
\[
I_{F_{X_2}}^* = \sup_{F_{X_2}} I_\lambda(F_{X_1}, F_{X_2}),
\]
boils down to the following necessary condition
\[
i_\lambda(x_2; F_{X_1}^*, F_{X_1}) + \theta_2(P_2^* - x_2^2) \leq I_{F_{X_1}}^*, \quad \forall x_2 \in \mathbb{R},
\]
for the optimality of \( F_{X_2}^* \), which holds with equality if and only if \( x_2 \) is a point of increase of \( F_{X_2}^* \). However, there are two main differences between (34) and (24). First is the difference between \( i_\lambda \) and \( \tilde{i}_\lambda \). Second is the fact that we do not claim \( \theta_2 \) to be nonzero, since the approach used in Lemma 1 cannot be readily applied to \( \theta_2 \).

\(^5\)Note that \( A_1 \) and \( A_2 \) are determined by the choice of \( F_{X_2} \).
Nonetheless, the boundedness of the support of $F_{X_2}^*$ can be proved by inspecting the behaviour of the LHS of (34) when $|x_2| \to +\infty$.

It can be easily verified that

$$
\lim_{x_2 \to +\infty} i_\lambda(x_2; F_{X_2}^*|F_{X_1}) = -\lambda \log p_Y(1; F_{X_1}, F_{X_2}^*) \leq -\lambda \log Q(\sqrt{P_1'} + \sqrt{P_2'}),
$$

$$
\lim_{x_2 \to -\infty} i_\lambda(x_2; F_{X_2}^*|F_{X_1}) = -\lambda \log p_Y(0; F_{X_1}, F_{X_2}^*) \leq -\lambda \log Q(\sqrt{P_1'} + \sqrt{P_2'}),
$$

(35)

If $\theta_2 > 0$, the LHS of (34) goes to $-\infty$ with $|x_2|$ which proves that $X_2^*$ is bounded. For the case of $\theta_2 = 0$, we rely on the boundedness of $X_1$, and use the fact that $i_\lambda$ approaches its limit in (35) from below. In other words, there is a real number $K$ such that $i_\lambda(x_2; F_{X_2}^*|F_{X_1}) < -\lambda \log p_Y(1; F_{X_1}, F_{X_2}^*)$ when $x_2 > K$, and $i_\lambda(x_2; F_{X_2}^*|F_{X_1}) < -\lambda \log p_Y(0; F_{X_1}, F_{X_2}^*)$ when $x_2 < -K$. This establishes the boundedness of $X_2^*$.

By rewriting $i_\lambda$, we have

$$
i_\lambda(x_2; F_{X_2}^*|F_{X_1}) = -\lambda p(1; F_{X_1}|x_2) \log p_Y(1; F_{X_1}, F_{X_2}^*)
$$

$$
- \int_{-A_1}^{A_2} H_b(Q(x_1 + x_2)) dF_{X_1}(x_1) + (1 - \lambda) \log p_Y(0; F_{X_1}, F_{X_2}^*)
$$

(36)

Note that without loss of generality, we have assumed $X_1 \in [-A_1, A_2]$, where $A_1, A_2$ are some non-negative real numbers. Also, we restrict our attention to $x_2 \to +\infty$, since the case $x_2 \to -\infty$ follows similarly. It is obvious that the first term in the right hand side of (36) approaches $-\lambda \log p_Y(1; F_{X_1}, F_{X_2}^*)$ from below when $x_2 \to +\infty$, since $p(1; F_{X_1}|x_2) \leq 1$. It is also obvious that the remaining terms go to zero when $x_2 \to +\infty$. Hence, it is sufficient to show that the second line of (36) approaches zero from below, which is shown in Appendix E. This proves that $X_2^*$ has a bounded support.

**Remark 3.** We remark here that the order of showing the boundedness of the supports is important. First, for a given $F_{X_2}$ (not necessarily bounded), it is proved that $F_{X_1}^*$ is bounded. Then, for a given bounded $F_{X_1}$, it is shown that $F_{X_2}^*$ is also bounded. The order is reversed when $\lambda > 1$, and it follows the same steps as in the case of $\lambda \leq 1$. Therefore, it is omitted.

We next prove the second claim in Proposition 3. We assume that $0 < \lambda < 1$, and a bounded $F_{X_1}$ is given. We already know that for a given bounded $F_{X_1}$, $F_{X_2}^*$ has a bounded support denoted by $[A_1, A_2]$. Let $P_2^*$ denote the average power of $X_2^*$. Also, let $\mathcal{F}_2$ denote the set of all probability distributions on the Borel sets of $[A_1, A_2]$ and $p_0^* = p_Y(0; F_{X_1}, F_{X_2}^*)$ be the probability of the event $Y = 0$, induced by $F_{X_2}^*$ and the given $F_{X_1}$. The set

$$
\mathcal{F}_2 = \left\{ F_{X_2} \in \mathcal{F}_2 | \int_{-\infty}^{+\infty} p(0|x_2) dF_{X_2}(x_2) = p_0^* , \int_{-\infty}^{+\infty} x_2^2 dF_{X_2}(x_2) = P_2^* \right\}
$$

(37)

is the intersection of $\mathcal{F}_2$ with two hyperplanes. We can write

$$
I_\lambda(F_{X_1}, F_{X_2}^*) = \sup_{F_{X_2} \in \mathcal{F}_2} I_\lambda(F_{X_1}, F_{X_2}^*).
$$

$^6$Note that $\mathcal{F}_2$ is convex and compact.
Note that having \( F_{X_2} \in \mathcal{F}_2 \) results in
\[
I_\lambda = \lambda H(Y) + (1 - \lambda)H(Y|X_2) - H(Y|X_1, X_2).
\]
(38)
Since the linear part is continuous and \( \mathcal{F}_2 \) is compact, \( I_\lambda \) attains its maximum at an extreme point of \( \mathcal{F}_2 \), which, by Dubins’ theorem, is a convex combination of at most 3 extreme points of \( \mathcal{F}_2 \). Since the extreme points of \( \mathcal{F}_2 \) are the CDFs having only one point of increase in \([A_1, A_2]\), we conclude that given any bounded \( F_{X_1} \), the maximum of \( I_\lambda(F_{X_1}, F_{X_2}) \) over \( F_{X_2} \), which satisfies the power constraint, is achieved by an input \( F_{X_2}^* \) having at most 3 mass points.

Now, assume that an arbitrary \( F_{X_2} \) is given with at most three mass points denoted by \( \{x_{2,i}\}_{i=1}^3 \). It is already known that the support of \( F_{X_1}^* \) is bounded, which is denoted by \([A_1', A_2']\). Since the Lagrangian multiplier \( \theta_1 \) is nonzero (see Lemma 1), \( X_1^* \) has the largest permissible power \( P_1' \). Let \( \mathcal{F}_1 \) denote the set of all probability distributions on the Borel sets of \([A_1', A_2']\). The set
\[
\mathcal{F}_1 = \{F_{X_1} \in \mathcal{F}_1 | \int_{-\infty}^{+\infty} p(0|x_1, x_{2,1})dF_{X_1}(x_1) = p(0; F_{X_1}^*|x_{2,1}),
\int_{-\infty}^{+\infty} p(0|x_1, x_{2,2})dF_{X_1}(x_1) = p(0; F_{X_1}^*|x_{2,2}),
\int_{-\infty}^{+\infty} p(0|x_1, x_{2,3})dF_{X_1}(x_1) = p(0; F_{X_1}^*|x_{2,3}),
\int_{-\infty}^{+\infty} x_1^2dF_{X_1}(x_1) = P_1' \},
\]
(39)
is the intersection of \( \mathcal{F}_1 \) with four hyperplanes. In a similar way,
\[
I_\lambda(F_{X_1}^*, F_{X_2}) = \sup_{F_{X_1} \in \mathcal{F}_1} I_\lambda(F_{X_1}, F_{X_2}),
\]
(40)
and having \( F_{X_1} \in \mathcal{F}_1 \) results in
\[
I_\lambda = \lambda H(Y) + (1 - \lambda) \sum_{i=1}^{3} H(Y|X_2 = x_{2,i}) - H(Y|X_1, X_2)
\]
linear in \( F_{X_1} \),
(41)
Therefore, given any \( F_{X_2} \) with at most three points of increase, the maximum of \( I_\lambda(F_{X_1}, F_{X_2}) \) over \( F_{X_1} \) is achieved by an input \( F_{X_1}^* \) having at most five mass points.

When \( \lambda = 1 \), the term with summation in (41) disappears, which means that \( \mathcal{F}_1 \) could be replaced by
\[
\left\{ F_{X_1} \in \mathcal{F}_1 | \int_{-\infty}^{+\infty} p(0|x_1)dF_{X_1}(x_1) = \tilde{p}_0', \int_{-\infty}^{+\infty} x_1^2dF_{X_1}(x_1) = P_1' \right\},
\]
where \( \tilde{p}_0' = p_Y(0; F_{X_1}^*, F_{X_2}) \) is the probability of the event \( Y = 0 \), which is induced by \( F_{X_1}^* \) and the given \( F_{X_2} \).
Since the number of intersecting hyperplanes has been reduced to 2, it is concluded that \( F_{X_1}^* \) has at most three points of increase.

**Remark 4.** Note that the order of showing the discreteness of the supports is also important. First, for a given bounded \( F_{X_1} \) (not necessarily discrete), it is proved that \( F_{X_2}^* \) is discrete with at most three mass points. Then, for a given discrete \( F_{X_2} \) with at most three mass points, it is shown that \( F_{X_1}^* \) is also discrete with at most five mass
points (three mass points) when \( \lambda < 1 \) (when \( \lambda = 1 \)). When \( \lambda > 1 \), the order is reversed and it follows the same steps as in the case of \( \lambda < 1 \). Therefore, it is omitted.

\[ \square \]

V. Conclusion

In this paper, the capacity region of a two-transmitter Gaussian MAC under average input power constraints and one-bit ADC front end at the receiver is considered. It is shown that an auxiliary random variable is necessary for characterizing the capacity region. An upper bound for the cardinality of this auxiliary variable is derived, and it is proved that the distributions that achieve the boundary points of the capacity region, are finite and discrete.

APPENDIX A

PROOF OF PROPOSITION 2

A. Concavity

When \( 0 < \lambda \leq 1 \), we have

\[
I_{\lambda}(F_{X_1}, F_{X_2}) = \lambda H(Y) + (1 - \lambda)H(Y|X_2) - H(Y|X_1, X_2).
\]

(42)

For a given \( F_{X_1} \), \( H(Y) \) is a concave function of \( F_{X_2} \), while \( H(Y|X_2) \) and \( H(Y|X_1, X_2) \) are linear in \( F_{X_2} \). Therefore, \( I_{\lambda} \) is a concave function of \( F_{X_2} \). For a given \( F_{X_2} \), \( H(Y) \) and \( H(Y|X_2) \) are concave functions of \( F_{X_1} \), while \( H(Y|X_1, X_2) \) is linear in \( F_{X_1} \). Since \( (1 - \lambda) \geq 0 \), \( I_{\lambda} \) is a concave function of \( F_{X_1} \). The same reasoning applies to the case \( \lambda > 1 \).

B. Continuity

When \( \lambda \leq 1 \), the continuity of the three terms on the right hand side of (42) is investigated. Let \( \{F_{X_2,n}\} \) be a sequence of distributions which is weakly convergent\(^7\) to \( F_{X_2} \). For a given \( F_{X_1} \), we have

\[
\lim_{x_2 \to x_2^0} p(y; F_{X_1}|x_2) = \lim_{x_2 \to x_2^0} \int Q(x_1 + x_2) dF_{X_1}(x_1)
\]

\[
= \int \lim_{x_2 \to x_2^0} Q(x_1 + x_2) dF_{X_1}(x_1)
\]

\[
= p(y; F_{X_1}|x_2^0),
\]

(44)

(45)

\[^7\text{The weak convergence of } \{F_n\} \text{ to } F \text{ (also shown as } F_n(x) \xrightarrow{w} F(x)) \text{ is equivalent to}
\]

\[
\lim_{n \to \infty} \int_{\mathbb{R}} \psi(x) dF_n(x) = \int_{\mathbb{R}} \psi(x) dF(x),
\]

(43)

for all continuous and bounded functions \( \psi(\cdot) \) on \( \mathbb{R} \). Note that \( F_n(x) \xrightarrow{w} F(x) \) if and only if \( d_L(F_n, F) \to 0 \)
where (44) is due to the fact that the \( U \) function can be dominated by 1, which is an absolutely integrable function over \( F_{X_1} \). Therefore, \( p(y; F_{X_1|x_2}) \) is continuous in \( x_2 \), and combined with the weak convergence of \( \{F_{X_2,n}\} \), we can write

\[
\lim_{n \to \infty} p(y; F_{X_1,F_{X_2,n}}) = \lim_{n \to \infty} \int p(y; F_{X_1|x_2})dF_{X_2,n}(x_2) = \int p(y; F_{X_1|x_2})dF_{X_2}(x_2) = p(y; F_{X_1,F_{X_2}}).
\]

This allows us to write

\[
\lim_{n \to \infty} - \sum_{y=0}^{1} p(y; F_{X_1,F_{X_2,n}}) \log p(y; F_{X_1,F_{X_2,n}}) = - \sum_{y=0}^{1} p(y; F_{X_1,F_{X_2,n}}) \log p(y; F_{X_1,F_{X_2}}),
\]

which proves the continuity of \( H(Y) \) in \( F_{X_2} \). \( H(Y|X_2 = x_2) \) is a bounded \((\in [0,1])\) continuous function of \( x_2 \), since it is a continuous function of \( p(y; F_{X_1|x_2}) \), and the latter is continuous in \( x_2 \) (see (45)). Therefore,

\[
\lim_{n \to \infty} \int H(Y|X_2 = x_2)dF_{X_2,n}(x_2) = \int H(Y|X_2 = x_2)dF_{X_2}(x_2),
\]

which proves the continuity of \( H(Y|X_2) \) in \( F_{X_2} \). In a similar way, it can be verified that \( \int H(Y|X_1 = x_1, X_2 = x_2)dF_{X_1}(x_1) \) is a bounded and continuous function of \( x_2 \) which guarantees the continuity of \( H(Y|X_1, X_2) \) in \( F_{X_2} \). Therefore, for a given \( F_{X_1} \), \( I_\lambda \) is a continuous function of \( F_{X_2} \). Exchanging the roles of \( F_{X_1} \) and \( F_{X_2} \) and also the case \( \lambda > 1 \) can be addressed similarly, and are omitted for the sake of brevity.

### C. Weak Differentiability

For a given \( F_{X_1} \), the weak derivative of \( I_\lambda \) at \( F_{X_2}^0 \) is given by

\[
I'_\lambda(F_{X_1,F_{X_2}})|_{F_{X_2}^0} = \lim_{\beta \to 0^+} \frac{I_\lambda(F_{X_1}((1 - \beta)F_{X_2}^0 + \beta F_{X_2})) - I_\lambda(F_{X_1,F_{X_2}})}{\beta},
\]

if the limit exists. It can be verified that

\[
I'_\lambda(F_{X_1,F_{X_2}})|_{F_{X_2}^0} = \lim_{\beta \to 0^+} \frac{\int i_\lambda(x_2; (1 - \beta)F_{X_2}^0 + \beta F_{X_2}|F_{X_1})d((1 - \beta)F_{X_2}^0(x_2) + \beta F_{X_2}(x_2)) - \int i_\lambda(x_2; F_{X_2}^0|F_{X_1})dF_{X_2}^0(x_2)}{\beta}.
\]

In a similar way, for a given \( F_{X_2} \), the weak derivative of \( I_\lambda \) at \( F_{X_1}^0 \) is

\[
I'_\lambda(F_{X_1,F_{X_2}})|_{F_{X_1}^0} = \int i_\lambda(x_1; F_{X_1}^0,F_{X_2})dF_{X_1}(x_1) - I_\lambda(F_{X_1,F_{X_2}}).
\]

The case \( \lambda > 1 \) can be addressed similarly.
APPENDIX B

PROOF OF LEMMA 1

We have

$$\sup_{F_{X_1}} \ I_x(F_{X_1}, F_{X_2}) \leq \sup_{F_{X_1}, F_{X_2}: E[X_j^2] \leq P'_j} I_x(F_{X_1}, F_{X_2})$$

$$\leq \sup_{F_{X_1}, F_{X_2}: E[X_j^2] \leq P'_j, j=1,2} I(X_1, X_2; Y)$$

$$\leq \sup_{F_{X_1}, F_{X_2}: E[X_j^2] \leq P'_j, j=1,2} H(Y) - \inf_{F_{X_1}, F_{X_2}: E[X_j^2] \leq P'_j, j=1,2} H(Y | X_1, X_2)$$

$$= 1 - \inf_{F_{X_1}, F_{X_2}: E[X_j^2] \leq P'_j, j=1,2} \int \int H_b(Q(x_1 + x_2)) dF_{X_1}(x_1) dF_{X_2}(x_2)$$

$$= 1 - \inf_{F_{X_1}, F_{X_2}: E[X_j^2] \leq P'_j, j=1,2} \int \int H_b \left( Q \left( \sqrt{x_1^2 + x_2^2} \right) \right) dF_{X_1}(x_1) dF_{X_2}(x_2)$$

$$\leq 1 - \inf_{F_{X_1}, F_{X_2}: E[X_j^2] \leq P'_j, j=1,2} \int \int Q \left( \sqrt{x_1^2 + x_2^2} \right) dF_{X_1}(x_1) dF_{X_2}(x_2)$$

$$= 1 - Q \left( \sqrt{P'_1 + P'_2} \right)$$

$$< 1,$$

where (47) is from the assumption that $0 < \lambda \leq 1$; (48) is justified by the fact that since the $U$ function is monotonically decreasing and the sign of the inputs does not affect the average power constraints, $X_1$ and $X_2$ can be assumed non-negative (or alternatively non-positive) without loss of optimality; in (49), we use the fact that $Q \left( \sqrt{x_1^2 + x_2^2} \right) \leq \frac{1}{2}$, and for $t \in [0, \frac{1}{2}]$, $H_b(t) \geq t$; (50) is based on the convexity and monotonicity of the function $Q(\sqrt{u} + \sqrt{v})$ in $(u, v)$, which is shown in Appendix D. Therefore, the left hand side of (18) is strictly lower than 1.

Since $X_2$ has a finite second moment ($E[X_2^2] \leq P'_2$), from Chebyshev inequality, we have

$$P(|X_2| \geq M) \leq \frac{P'_2}{M^2}, \quad \forall M > 0.$$
Fix \( M > 0 \) and consider \( X_1 \sim F_{X_1}(x_1) = \frac{1}{2}[s(x_1 + 2M) + s(x_1 - 2M)] \). By this choice of \( F_{X_1} \), we get

\[
I_\lambda(F_{X_1}, F_{X_2}) = I(X_1; Y|X_2) + \lambda I(X_2; Y)
\]

\[
\geq I(X_1; Y|X_2)
\]

\[
= \int_{-\infty}^{+\infty} I(X_1; Y|X_2 = x_2)dF_{X_2}(x_2)
\]

\[
\geq \int_{-M}^{+M} I(X_1; Y|X_2 = x_2)dF_{X_2}(x_2)
\]

\[
\geq \inf_{F_{X_2}} \int_{-M}^{+M} H(Y|X_2 = x_2)dF_{X_2}(x_2) - \sup_{F_{X_2}} \int_{-M}^{+M} H(Y|X_1, X_2 = x_2)dF_{X_2}(x_2)
\]

\[
\geq \left( 1 - \frac{P^2}{M^2} \right) H_b \left( \frac{1}{2} - \frac{1}{2} (Q(3M) + Q(M)) \right) - H_b \left( Q(2M) \right),
\] (53)

where (53) is due to the fact that \( H(Y|X_2 = x_2) = H_b\left( \frac{1}{2} Q(2M + x_2) + \frac{1}{2} Q(-2M + x_2) \right) \) is minimized over \([-M, M]\) at \( x_2 = M \) (or, alternatively at \( x_2 = -M \)), and \( H(Y|X_1, X_2 = x_2) = \frac{1}{2} H_b(Q(2M + x_2)) + \frac{1}{2} H_b(Q(-2M + x_2)) \) is maximized at \( x_2 = 0 \). (53) shows that \( I_\lambda \) can become arbitrarily close to 1 given that \( M \) is large enough. Hence, its supremum over all distributions \( F_{X_1} \) is 1. This means that (18) cannot hold, and \( \theta_1 \neq 0 \).

**APPENDIX C**

**PROOF OF LEMMA 2**

(25) is obtained as follows.

\[
\left| D(p(y|x_1, x_2)||p(y; F_{X_1}, F_{X_2})) \right| = \left| \sum_{y=0}^{1} p(y|x_1, x_2) \log \frac{p(y|x_1, x_2)}{p(y; F_{X_1}, F_{X_2})} \right|
\]

\[
\leq \left| H(Y|X_1 = x_1, X_2 = x_2) \right| + \left| \sum_{y=0}^{1} p(y|x_1, x_2) \log p(y; F_{X_1}, F_{X_2}) \right|
\]

\[
\leq 1 + \left| \sum_{y=0}^{1} \log p(y; F_{X_1}, F_{X_2}) \right|
\]

\[
= 1 - \left| \sum_{y=0}^{1} \log p(y; F_{X_1}, F_{X_2}) \right|
\]

\[
\leq 1 - 2 \min \left\{ \log p_Y(0; F_{X_1}, F_{X_2}), \log p_Y(1; F_{X_1}, F_{X_2}) \right\}
\]

\[
\leq 1 - 2 \log Q(\sqrt{P_1} + \sqrt{P_2})
\]

< \infty, (55)
where (54) is due to the fact that the binary entropy function is upper bounded by 1. (55) is justified as follows.

\[
\min \left\{ p_Y(0; F_X, F_{X_2}), p_Y(1; F_X, F_{X_2}) \right\} \geq \inf_{E[X_j^2] \leq P_j} \min \left\{ p_Y(0; F_X, F_{X_2}), p_Y(1; F_X, F_{X_2}) \right\} \\
= \inf_{E[X_j^2] \leq P_j} p_Y(0; F_X, F_{X_2}) \\
= \inf_{E[X_j^2] \leq P_j} \int \int Q(x_1 + x_2) dF_{X_1}(x_1) dF_{X_2}(x_2) \\
= \inf_{E[X_j^2] \leq P_j} \int \int Q \left( \sqrt{x_1^2 + x_2^2} \right) dF_{X_1}(x_1) dF_{X_2}(x_2) \tag{56}
\]

\[
\geq Q \left( \sqrt{P_1^2 + P_2^2} \right), \tag{57}
\]

where (57) is based on the convexity and monotonicity of the function \( Q(\sqrt{u + v}) \), which is shown in appendix D.

(26) is obtained as follows.

\[
p(y; F_{X_1}, x_2) \geq \min \left\{ p(0; F_{X_1}, x_2), p(1; F_{X_1}, x_2) \right\} \\
\geq \int Q \left( |x_1| + |x_2| \right) dF_{X_1}(x_1) \\
= \int Q \left( \sqrt{x_1^2 + x_2^2} \right) dF_{X_1}(x_1) \\
\geq Q \left( \sqrt{P_1^2 + |x_2|} \right), \tag{58}
\]

where (58) is due to convexity of \( Q(\alpha + \sqrt{x}) \) in \( x \) for \( \alpha \geq 0 \).

(27) is obtained as follows.

\[
\int_{-\infty}^{+\infty} \sum_{y=0}^{1} p(y|x_1, x_2) \log \frac{p(y; F_{X_1}, F_{X_2})}{p(y; F_{X_1}, x_2)} dF_{X_2}(x_2) \leq -\int_{-\infty}^{+\infty} \sum_{y=0}^{1} \log p(y; F_{X_1}, x_2) dF_{X_2}(x_2) \\
\leq -2 \int_{-\infty}^{+\infty} \log Q \left( \sqrt{P_1^2 + |x_2|} \right) dF_{X_2}(x_2) \tag{59}
\]

\[
= -2 \int_{-\infty}^{+\infty} \log Q \left( \sqrt{P_1^2 + x_2^2} \right) dF_{X_2}(x_2) \leq -2 \log Q \left( \sqrt{P_1^2 + P_2^2} \right) \tag{60}
\]

\[
< +\infty \tag{61}
\]

where (73) is from (58) and (60) is due to concavity of \(- \log Q(\alpha + \sqrt{x}) \) in \( x \) for \( \alpha \geq 0 \) as shown in Appendix D.

**APPENDIX D**

**TWO CONVEX FUNCTIONS**

Let \( f(x) = \log Q(\alpha + \sqrt{x}) \) for \( x, \alpha \geq 0 \). We have,

\[
f'(x) = -\frac{e^{-(\alpha + \sqrt{x})^2}}{2\sqrt{2\pi}xQ(\alpha + \sqrt{x})}.
\]
and
\[ f''(x) = \frac{e^{-(a+x)^2}}{4x\sqrt{2\pi}U^2(a+x)} \left( (a + \sqrt{x} + \frac{1}{\sqrt{x}})Q(a + \sqrt{x}) - \phi(a + \sqrt{x}) \right), \]  
(62)
where \( \phi(u) = \frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}} \). Not that
\[(1 + at + t^2)Q(a + t) + a\phi(a + t) > (1 + (a + t)^2)Q(a + t) \]
(63)
\[ > (a + t)\phi(a + t), \quad \forall a, t > 0, \]
(64)
where (63) and (64) are, respectively, due to \( \phi(x) > xQ(x) \) and \( (1 + x^2)Q(x) > x\phi(x) \) (\( x > 0 \)). Therefore,
\[(a + \sqrt{x} + \frac{1}{\sqrt{x}})Q(a + \sqrt{x}) > \phi(a + \sqrt{x}), \]
which makes the second derivative in (62) positive and proves the (strict) convexity of \( f(x) \).

Let \( f(u, v) = Q(\sqrt{u} + \sqrt{v}) \) for \( u, v \geq 0 \). By simple differentiation, the Hessian matrix of \( f \) is
\[ H = \frac{e^{-\frac{(\sqrt{x}+\sqrt{y})^2}{2}}}{\sqrt{2\pi}} \begin{bmatrix}
\frac{1}{2v\sqrt{v}} + \frac{\sqrt{u} + \sqrt{v}}{4u} & \frac{\sqrt{u} + \sqrt{v}}{4\sqrt{v}\sqrt{u}} \\
\frac{\sqrt{u} + \sqrt{v}}{4\sqrt{v}\sqrt{u}} & \frac{1}{2v\sqrt{v}} + \frac{\sqrt{u} + \sqrt{v}}{4v}
\end{bmatrix}. \]
(65)
It can be verified that \( \det(H) > 0 \) and \( \text{trace}(H) > 0 \). Therefore, both the eigenvalues of \( H \) are positive, which makes the matrix positive definite. Hence, \( Q(\sqrt{u} + \sqrt{v}) \) is (strictly) convex in \( (u, v) \).

**APPENDIX E**

The following lemma is helpful in the sequel.

**Lemma 3.** Let \( X_1 \) be distributed on \([-A, A]\) according to \( F_{X_1}(x_1) \). We have
\[ \lim_{x_2 \to +\infty} \frac{\int_{-A}^{A} H_b(Q(x_1 + x_2))dF_{X_1}(x_1)}{H_b\left( \int_{-A}^{A} Q(x_1 + x_2)dF_{X_1}(x_1) \right)} = 1, \]
(66)
\[ \lim_{x \to 0} \frac{H_b(x)}{cx} = +\infty, \]
(67)
where \( c \) is a positive constant.

**Proof.** It is easy to verify that
\[ \int_{-A}^{A} Q(x_1 + x_2)dF_{X_1}(x_1) = \beta Q(x_2 + A) + (1 - \beta)Q(x_2 - A) \]
(68)
for some \( \beta \in [0, 1] \). Note that \( \beta \) is a function of \( x_2 \). Also, due to concavity of \( H_b(\cdot) \), we have
\[ H_b\left( \int_{-A}^{A} Q(x_1 + x_2)dF_{X_1}(x_1) \right) \geq \int_{-A}^{A} H_b(Q(x_1 + x_2))dF_{X_1}(x_1) \]
(69)
\[ \geq \beta H_b(Q(x_2 + A)) + (1 - \beta)H_b(Q(x_2 - A)), \]
(70)
as shown in Figure 3. Therefore,
\[ \frac{\beta H_b(Q(x_2 + A)) + (1 - \beta)H_b(Q(x_2 - A))}{H_b\left( \beta Q(x_2 + A) + (1 - \beta)Q(x_2 - A) \right)} \leq \frac{\int_{-A}^{A} H_b(Q(x_1 + x_2))dF_{X_1}(x_1)}{H_b\left( \int_{-A}^{A} Q(x_1 + x_2)dF_{X_1}(x_1) \right)} \leq 1 \]
(71)
for some $\beta \in [0, 1]$. Let

$$
\beta^* = \arg \min_\beta \frac{\beta H_b(Q(x_2 + A)) + (1 - \beta) H_b(Q(x_2 - A))}{H_b(\beta Q(x_2 + A) + (1 - \beta) Q(x_2 - A))}.
$$

(72)

Therefore,

$$
\frac{d}{d\beta} \left( \frac{\beta H_b(Q(x_2 + A)) + (1 - \beta) H_b(Q(x_2 - A))}{H_b(\beta Q(x_2 + A) + (1 - \beta) Q(x_2 - A))} \right) \bigg|_{\beta = \beta^*} = 0.
$$

(73)

After some manipulation, (73) becomes equivalent to

$$
\frac{\beta^* H_b(Q(x_2 + A)) + (1 - \beta^*) H_b(Q(x_2 - A))}{H_b(\beta^* Q(x_2 + A) + (1 - \beta^*) Q(x_2 - A))} = \frac{H_b(Q(x_2 - A)) - H_b(Q(x_2 + A))}{H_b(Q(x_2 - A) - Q(x_2 + A))}
$$

(74)

$$
\geq \frac{H_b(Q(x_2 - A)) - H_b(Q(x_2 + A))}{H_b'(Q(x_2 + A))}
$$

(75)

where $H_b'(t) = \log\left(\frac{1-t}{t}\right)$ is the derivative of the binary entropy function; (75) is due to the fact that $H_b'(t)$ is a decreasing function.
Applying L’Hospital’s rule multiple times, we have

\[
\lim_{x \to +\infty} \frac{H_b(Q(x_2 - A) - H_b(Q(x_2 + A)))}{Q(x_2 - A) - Q(x_2 + A)} = \lim_{x \to +\infty} \frac{H_b(Q(x_2 - A))}{Q(x_2 - A)} \left(1 - \frac{H_b(Q(x_2 + A))}{H_b(Q(x_2 - A))}\right)
\]

\[
= \lim_{x \to +\infty} -\frac{H_b(Q(x_2 - A))}{Q(x_2 - A) \log(Q(x_2 + A))}
\]

\[
= \lim_{x \to +\infty} -\frac{e^{(-Q(x_2 - A))}}{Q(x_2 - A) \log(Q(x_2 + A)) + 1}
\]

\[
= \lim_{x \to +\infty} \frac{Q(x_2 + A)e^x}{Q(x_2 - A)e^{-x}}
\]

\[
= 1
\]

Also,

\[
\lim_{x \to 0} \frac{H_b(x)}{c x} = \lim_{x \to 0} -\frac{\log x}{c} = +\infty, \; c > 0.
\]

From (66), we can write

\[
\int_{-A}^{A} H_b(Q(x_1 + x_2))dF_{X_1}(x_1) = \gamma(x_2)H_b \left(\int_{-A}^{A} Q(x_1 + x_2)dF_{X_1}(x_1)\right),
\]

where \(\gamma(x_2) \leq 1\), and \(\gamma(x_2) \to 1\) when \(x_2 \to +\infty\). Also, from (67),

\[
H_b \left(\int_{-A}^{A} Q(x_1 + x_2)dF_{X_1}(x_1)\right) = -\eta(x_2) \log p_Y(0; F_{X_1} F_{X_2}) \int_{-A}^{A} Q(x_1 + x_2)dF_{X_1}(x_1),
\]

where \(\eta(x_2) > 0\) and \(\eta(x_2) \to +\infty\) when \(x_2 \to +\infty\). From (82) and (83), the second line of (36) becomes

\[
\left(1 - \gamma(x_2) + \frac{\lambda}{\eta(x_2)} - \lambda\right) -\eta(x_2) \log p_Y(0; F_{X_1} F_{X_2}) \int_{-A}^{A} Q(x_1 + x_2)dF_{X_1}(x_1)
\]

\[
\geq 0
\]

Since \(\gamma(x_2) \to 1\) and \(\eta(x_2) \to +\infty\) as \(x_2 \to +\infty\), there exists a real number \(K\) such that \(1 - \gamma(x_2) + \frac{\lambda}{\eta(x_2)} - \lambda < 0\) when \(x_2 > K\). Therefore, the second line of (36) approaches zero from below.

**REFERENCES**

[1] R. Walden, “Analog-to-digital converter survey and analysis,” *IEEE Journal on Selected Areas in Comm.*, vol. 17, no. 4, pp. 539–550, Apr. 1999.

[2] B. Murmann, “ADC performance survey,” *CoRR*, vol. abs/1404.7736, 1997-2014. [Online]. Available: http://web.stanford.edu/~murmann/adc_survey.html

[3] D. Gunduz, K. Stamatiou, N. Michelusi, and M. Zorzi, “Designing intelligent energy harvesting communication systems,” *IEEE Comm. Magazine*, vol. 52, pp. 210–216, January 2014.

[4] J. Singh, O. Daheer, and U. Madhow, “On the limits of communication with low-precision analog-to-digital conversion at the receiver,” *IEEE Trans. Communications*, vol. 57, no. 12, pp. 3629–3639, December 2009.
[5] S. Krone and G. Fettweis, “Fading channels with 1-bit output quantization: Optimal modulation, ergodic capacity and outage probability,” in Proc. IEEE Inf. Theory Workshop (ITW), Aug. 2010, pp. 1–5.

[6] A. Mezghani and J. Nossek, “Analysis of Rayleigh-fading channels with 1-bit quantized output,” IEEE Int. Sym. Inf. Theory, pp. 260–264, Jul. 2008.

[7] ——, “On ultra-wideband MIMO systems with 1-bit quantized outputs: Performance analysis and input optimization,” IEEE Int. Sym. Inf. Theory, pp. 1286–1289, Jun. 2007.

[8] J. Mo and R. Heath, “Capacity analysis of one-bit quantized MIMO systems with transmitter channel state information,” IEEE Trans. on Signal Processing, vol. 63, no. 20, pp. 5498–5512, Oct. 2015.

[9] A. E. Gamal and Y.-H. Kim, Network Information Theory. Cambridge University Press, 2012.

[10] H. Witsenhausen, “Some aspects of convexity useful in information theory,” IEEE Trans. Inf. Theory, vol. 26, no. 3, pp. 265–271, May 1980.

[11] I. Abou-Faycal, M. Trott, and S. Shamai, “The capacity of discrete-time memoryless rayleigh-fading channels,” IEEE Trans. Inf. Theory, vol. 47, no. 4, pp. 1290–1300, May 2001.

[12] W. Rudin, Principles of Mathematical Analysis. McGraw-Hill, 1976.