Capacity Region of a One-Bit Quantized Gaussian Multiple Access Channel

Borzoo Rassouli, Morteza Varasteh and Deniz Gündüz

***Abstract*—The capacity region of a two-transmitter Gaussian multiple access channel (MAC) under average input power constraints is studied, when the receiver employs a zero-threshold one-bit analog-to-digital converter (ADC). 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, upper bounds on the number of the mass points of the corresponding distributions are derived. Finally, a conjecture on the sufficiency of K mass points in a point-to-point real AWGN with a K-bin ADC front end (symmetric or asymmetric) is settled.**[1]

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, it is shown in [4] that, if the receiver is equipped with a K-bin (i.e., $\log_2 K$-bit) ADC front end, the capacity-achieving input distribution is discrete with at most $K+1$ mass points. We further tighten this to K mass points in this paper. 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 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 [5]-[7]. In [5], the authors show that for a Rayleigh fading channel with a one-bit ADC front end and perfect channel state information at the receiver (CSIR), quadrature phase shift keying (QPSK) modulation is capacity-achieving. 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 [6].

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 derived. Finally, in the proof of Theorem 1, a simple optimization trick is used that also settles a conjecture in the real AWGN channel with a K-bin ADC front end (symmetric or asymmetric).

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 denote the set $\{m, m+1, \ldots, n\}$ by $[m:n]$.

Remark 1. Some of the proofs, omitted here, can be found in the longer version of the paper available online [8].

II. System model and preliminaries

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

$$Y_i = \Gamma(X_{1,i} + X_{2,i} + Z_i),\ i \in [1:n],$$

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_1^n and X_2^n with $Z_i \sim \mathcal{N}(0,1), i \in [1:n]$. Γ 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

Borzoo Rassouli, Morteza Varasteh and Deniz Gündüz are with the Intelligent Systems and Networks group of Department of Electrical and Electronics, Imperial College London, United Kingdom. emails: {b.rassouli12, m.varasteh12, d.gunduz}@imperial.ac.uk.

[1]This research was supported in part by the European Research Council (ERC) through Starting Grant BEACON (agreement #677854), by the UK Engineering and Physical Sciences Research Council (EPSRC) through the project COPES (EP/N021738/1), and by the British Council Institutional Link Program under grant number 173605884.

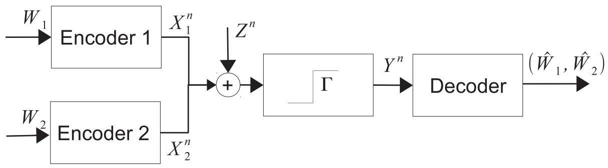

Fig. 1: A two-transmitter Gaussian MAC with a one-bit ADC at the receiver.

$(\mathcal{X}_1 \times \mathcal{X}_2, p(y|x_1, x_2), \mathcal{Y})$, where $\mathcal{X}_1, \mathcal{X}_2$ $(= \mathbb{R})$ and $\mathcal{Y}$ $(= \{0,1\})$, 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

$$p(0|x_1, x_2) = 1 - p(1|x_1, x_2) = Q(x_1 + x_2), \tag{1}$$

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

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

Average power constraints are imposed on the channel inputs as

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

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* $\mathscr{C}(P_1, P_2)$ 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 $\mathscr{C}(P_1, P_2)$ 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

$$\begin{aligned} R_1 &\leq I(X_1; Y|X_2, U), \\ R_2 &\leq I(X_2; Y|X_1, U), \\ R_1 + R_2 &\leq I(X_1, X_2; Y|U), \end{aligned} \tag{2}$$

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

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 $|\mathcal{U}| \leq 5$. This can be proved by using Carathéodory's Theorem [10] and taking into account the connectedness of the set of all product distributions on $\mathbb{R}^2$ [8]. □

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

Proof. The proof follows similarly to the proof of Proposition 1, and is provided in [8]. □

When there is no input cost constraint, the capacity region of a 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 $\mathsf{E}[X_j^2] \leq P_j, j = 1, 2$. Let $\mathscr{R}(P_1, P_2)$ denote the set of non-negative rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &\leq I(X_1; Y|X_2), \\ R_2 &\leq I(X_2; Y|X_1), \\ R_1 + R_2 &\leq I(X_1, X_2; Y). \end{aligned}$$

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

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

$$\begin{aligned} R_1 &\leq I(X_1; Y|X_2, U), \\ R_2 &\leq I(X_2; Y|X_1, U), \\ R_1 + R_2 &\leq I(X_1, X_2; Y|U) \end{aligned}$$

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

It can be verified that $\mathscr{R}_1(P_1, P_2) = \mathscr{R}_2(P_1, P_2)$. By comparing $\mathscr{R}_2(P_1, P_2)$ to the capacity region $\mathscr{C}(P_1, P_2)$, we can conclude that $\mathscr{R}_2(P_1, P_2) \subseteq \mathscr{C}(P_1, P_2)$. This follows from the fact that in the region $\mathscr{R}_2(P_1, P_2)$, the average power constraint $\mathsf{E}[X_j^2|u] \leq P_j$ holds for every realization of the auxiliary random variable U, which is a stronger condition than $\mathsf{E}[X_j^2] \leq P_j$ used in the capacity region. In [8], we show through an example that $\mathscr{R}_1(P_1, P_2)$ and $\mathscr{R}_2(P_1, P_2)$ can be strictly smaller than $\mathscr{C}(P_1, P_2)$. Therefore, in the presence of input cost constraints, there are cases in which the capacity region can be characterized only with the help of an auxiliary random variable.

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

Theorem 1. Let P be an arbitrary point on the boundary of the capacity region $\mathscr{C}(P_1, P_2)$ of the memoryless MAC with a one-bit ADC front end (as shown in Figure 1) achieved by $F_U^P(u)F_{X_1|U}^P(x_1|u)F_{X_2|U}^P(x_2|u)$. Let l_P be the slope of the line tangent to the capacity region at this point. For any $u \in \mathcal{U}$, the conditional input distributions $F_{X_1|U}^P(x_1|u)$ and $F_{X_2|U}^P(x_2|u)$ have at most n_1 and n_2 points of increase[2], respectively, where

$$(n_1, n_2) = \begin{cases} (2,3) & l_p < -1 \\ (2,2) & l_p = -1 \\ (3,2) & l_p > -1 \end{cases}. \tag{3}$$

Proof. The proof is provided in Section IV. □

Proposition 1, Lemma 1 and Theorem 1 above establish upper bounds on the number of mass points of the distributions that achieve a boundary point. The significance of this result is that once it is known that the optimal inputs are discrete with at most certain number of mass points, the capacity region along with the optimal distributions can be obtained via computer programs.

IV. Proof of Theorem 1

Any point on the boundary of the capacity region, denoted by (R_1^b, R_2^b), can be written as

$$(R_1^b, R_2^b) = \arg \max_{(R_1,R_2)\in\mathscr{C}(P_1,P_2)} R_1 + \lambda R_2,$$

for some $\lambda > 0$.

Any rate pair $(R_1, R_2) \in \mathscr{C}(P_1, P_2)$ is within the pentagon defined by (2) for some distribution $F_U F_{X_1|U} F_{X_2|U}$ 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.

$$\begin{aligned} &\max_{(R_1,R_2)\in\mathscr{C}(P_1,P_2)} R_1 + \lambda R_2 \\ &= \begin{cases} \max I(X_1;Y|X_2,U) + \lambda I(X_2;Y|U) & 0<\lambda\le 1 \\ \max I(X_2;Y|X_1,U) + \lambda I(X_1;Y|U) & \lambda > 1 \end{cases}, \end{aligned} \tag{4}$$

where on the right hand side (RHS) of (4), the maximizations are over all $F_U F_{X_1|U} F_{X_2|U}$ that satisfy the power constraints.

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

$$I_\lambda(F_{X_1}F_{X_2}) \triangleq \begin{cases} I(X_1;Y|X_2) + \lambda I(X_2;Y) & 0<\lambda\le 1 \\ I(X_2;Y|X_1) + \lambda I(X_1;Y) & \lambda > 1 \end{cases}. \tag{5}$$

With this definition, (4) can be written as

$$\max \sum_{i=1}^{5} p_U(u_i) I_\lambda(F_{X_1|U}(x_1|u_i)F_{X_2|U}(x_2|u_i)),$$

where the maximization is over product distributions of the form $p_U(u)F_{X_1|U}(x_1|u)F_{X_2|U}(x_2|u)$, $|\mathcal{U}| \le 5$, such that

$$\sum_{i=1}^{5} p_U(u_i)\mathsf{E}[X_j^2|u_i] \le P_j, \; j=1,2.$$

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.

Proof. The proof is provided in [8, Appendix A]. □

Proposition 3. Let P_1', P_2' be two arbitrary non-negative finite real numbers. For the following problem

$$\max_{\substack{F_{X_1}F_{X_2}: \\ \mathsf{E}[X_j^2]\le P_j',\; j=1,2}} I_\lambda(F_{X_1}F_{X_2}), \tag{6}$$

the optimal input distributions $F_{X_1}^*$ and $F_{X_2}^*$, which are not unique in general, have the following properties,

(i) The support sets of $F_{X_1}^*$ and $F_{X_2}^*$ are bounded subsets of $\mathbb{R}$.

(ii) $F_{X_1}^*$ and $F_{X_2}^*$ are discrete distributions that have at most n_1 and n_2 points of increase, respectively, where

$$(n_1, n_2) = \begin{cases} (3,2) & 0<\lambda<1 \\ (2,2) & \lambda = 1 \\ (2,3) & \lambda > 1 \end{cases}.$$

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

$$I_{F_{X_2}}^* \triangleq \sup_{\substack{F_{X_1}: \\ \mathsf{E}[X_1^2]\le P_1'}} I_\lambda(F_{X_1}F_{X_2}). \tag{7}$$

From Proposition 2, I_λ 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[3]. Therefore, the supremum in (7) is achieved by a unique distribution $F_{X_1}^*$. Since for any $F_{X_1}(x) = s(x - x_0)$ with $|x_0|^2 < P_1'$, where $s(\cdot)$ denotes the unit step function, we have $\mathsf{E}[X_1^2] < P_1'$, the Lagrangian theorem and the Karush-Kuhn-Tucker conditions state that there exists a $\theta_1 \ge 0$ such that

$$I_{F_{X_2}}^* = \sup_{F_{X_1}} \left\{ I_\lambda(F_{X_1}F_{X_2}) - \theta_1 \left(\int x^2 dF_{X_1}(x) - P_1' \right) \right\}. \tag{8}$$

Furthermore, the supremum in (8) is achieved by $F_{X_1}^*$, and

$$\theta_1 \left(\int x^2 dF_{X_1}^*(x) - P_1' \right) = 0. \tag{9}$$

Lemma 2. The Lagrangian multiplier θ_1 is nonzero.

Proof. Having a zero Lagrangian multiplier means that the power constraint is inactive. In other words, if $\theta_1 = 0$, (7) and (8) imply that

$$\sup_{\substack{F_{X_1}: \\ \mathsf{E}[X_1^2]\le P_1'}} I_\lambda(F_{X_1}F_{X_2}) = \sup_{F_{X_1}} I_\lambda(F_{X_1}F_{X_2}). \tag{10}$$

[2]A point Z is said to be a point of increase of a distribution if for any open set Ω containing Z, we have $\Pr\{\Omega\} > 0$.

[3]The compactness follows from [11, Appendix I].

We prove that (10) does not hold by showing that

$$\text{L.H.S of}(10) \leq 1 - Q\left(\sqrt{P_1'} + \sqrt{P_2'}\right) < 1 = \text{R.H.S of}(10).$$

The details are provided in [8, Appendix B]. □

Let $\tilde{i}_\lambda(x_1; F_{X_1}|F_{X_2})$ and $i_\lambda(x_2; F_{X_2}|F_{X_1})$ be defined as

$$\tilde{i}_\lambda(x_1; F_{X_1}|F_{X_2}) \triangleq \int_{-\infty}^{+\infty} \Bigg(D\left(p(y|x_1,x_2)||p(y;F_{X_1}F_{X_2})\right) + (1-\lambda)\sum_{y=0}^{1} p(y|x_1,x_2)\log\frac{p(y;F_{X_1}F_{X_2})}{p(y;F_{X_1}|x_2)}\Bigg) dF_{X_2}(x_2),$$

$$i_\lambda(x_2; F_{X_2}|F_{X_1}) \triangleq \int_{-\infty}^{+\infty} D\left(p(y|x_1,x_2)||p(y;F_{X_1}F_{X_2})\right) dF_{X_1}(x_1) - (1-\lambda) D\left(p(y;F_{X_1}|x_2)||p(y;F_{X_1}F_{X_2})\right),$$

where $p(y;F_{X_1}F_{X_2})$ is nothing but the pmf of Y with the emphasis that it has been induced by F_{X_1} and F_{X_2}. Likewise, $p(y;F_{X_1}|x_2)$ is the conditional pmf $p(y|x_2)$ when X_1 is drawn according to F_{X_1}. It can be verified that

$$\begin{aligned} I_\lambda(F_{X_1}F_{X_2}) &= \int_{-\infty}^{+\infty} \tilde{i}_\lambda(x_1; F_{X_1}|F_{X_2}) dF_{X_1}(x_1) \\ &= \int_{-\infty}^{+\infty} i_\lambda(x_2; F_{X_2}|F_{X_1}) dF_{X_2}(x_2). \end{aligned}$$

Note that (8) is an unconstrained optimization problem over the set of all CDFs, and a necessary condition for the optimality of $F^*_{X_1}$ is

$$\int \{\tilde{i}_\lambda(x_1; F^*_{X_1}|F_{X_2}) + \theta_1(P_1' - x_1^2)\} dF_{X_1}(x_1) \leq I^*_{F_{X_2}}, \quad \forall F_{X_1}, \tag{11}$$

which is equivalent to

$$\tilde{i}_\lambda(x_1; F^*_{X_1}|F_{X_2}) + \theta_1(P_1' - x_1^2) \leq I^*_{F_{X_2}}, \quad \forall x_1 \in \mathbb{R}, \tag{12}$$

with equality if and only if x_1 is a point of increase of $F^*_{X_1}$.

In what follows, we prove that in order to satisfy (12), $F^*_{X_1}$ must have a bounded support by showing that the left hand side (LHS) of (12) goes to $-\infty$ with x_1.

It can be verified that (see [8]),

$$\lim_{|x_1| \to +\infty} \tilde{i}_\lambda(x_1; F^*_{X_1}|F_{X_2}) < +\infty. \tag{13}$$

From (13), and the fact that $\theta_1 > 0$ (see Lemma 2), the LHS of (12) goes to $-\infty$ when $|x_1| \to +\infty$. Since any point of increase of $F^*_{X_1}$ must satisfy (12) 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 $A_1, A_2 \in \mathbb{R}$.[4]

Similarly, for a given F_{X_1}, the optimization problem

$$I^*_{F_{X_1}} = \sup_{\substack{F_{X_2}:\\ \mathsf{E}[X_2^2] \leq P_2'}} I_\lambda(F_{X_1}F_{X_2}),$$

boils down to the following necessary condition

$$i_\lambda(x_2; F^*_{X_2}|F_{X_1}) + \theta_2(P_2' - x_2^2) \leq I^*_{F_{X_1}}, \quad \forall x_2 \in \mathbb{R}, \tag{14}$$

for the optimality of $F^*_{X_2}$, which holds with equality if and only if x_2 is a point of increase of $F^*_{X_2}$. Note that there are two main differences between (14) and (12). First is the difference between i_λ and $\tilde{i}_\lambda$. Second is the fact that we do not claim θ_2 to be nonzero, since the approach used in Lemma 2 cannot be readily applied to θ_2. Nonetheless, the boundedness of the support of $F^*_{X_2}$ can be proved by inspecting the behaviour of the LHS of (14) when $|x_2| \to +\infty$. More specifically, if $\theta_2 > 0$, the LHS of (14) goes to $-\infty$ with $|x_2|$ which proves that X_2^* is bounded. For the case of $\theta_2 = 0$, we rely on the fact that i_λ approaches its limit from below, as shown in [8, Appendix E]. This proves that X_2^* must have 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 the proof 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 $[A_1, A_2]$. Therefore,

$$\begin{aligned} I^*_{F_{X_1}} &= \sup_{F_{X_2}} \left\{ I_\lambda(F_{X_1}F_{X_2}) - \theta_2\left(\int x^2 dF_{X_2}(x) - P_2'\right)\right\} \\ &= \sup_{F_{X_2} \in \mathscr{S}_2} \left\{ I_\lambda(F_{X_1}F_{X_2}) - \theta_2\left(\int x^2 dF_{X_2}(x) - P_2'\right)\right\}, \end{aligned}$$

where $\mathscr{S}_2$ denotes the set of all probability distributions on the Borel sets of $[A_1, A_2]$. Let $p_0^* = p_Y(0; F_{X_1}F^*_{X_2})$ denote the probability of the event $Y = 0$, induced by $F^*_{X_2}$ and the given F_{X_1}. The set

$$\mathscr{F}_2 = \left\{ F_{X_2} \in \mathscr{S}_2 \middle| \int p(0|x_2) dF_{X_2}(x_2) = p_0^* \right\}$$

is the intersection of $\mathscr{S}_2$ with one hyperplane[5]. We can write

$$I^*_{F_{X_1}} = \sup_{F_{X_2} \in \mathscr{F}_2} \left\{ I_\lambda(F_{X_1}F_{X_2}) - \theta_2\left(\int x^2 dF_{X_2}(x) - P_2'\right)\right\}. \tag{15}$$

Note that having $F_{X_2} \in \mathscr{F}_2$, the objective function in (15) becomes

$$\underbrace{\lambda H(Y)}_{\text{constant}} + \underbrace{(1-\lambda)H(Y|X_2) - H(Y|X_1,X_2) - \theta_2\left(\int x^2 dF_{X_2}(x) - P_2'\right)}_{\text{linear in } F_{X_2}}.$$

Since the linear part is continuous and $\mathscr{F}_2$ is compact, the objective function in (15) attains its maximum at an extreme point of $\mathscr{F}_2$, which, by Dubins' theorem [10], is a convex combination of at most two extreme points of $\mathscr{S}_2$. Since the extreme points of $\mathscr{S}_2$ are the CDFs having only one point of increase in $[A_1, A_2]$, we conclude that, given any bounded F_{X_1}, $F^*_{X_2}$ has at most two mass points.

[4]Note that A_1 and A_2 are determined by the choice of F_{X_2}.

[5]Note that $\mathscr{S}_2$ is convex and compact.

Now, assume that an arbitrary F_{X_2} is given with at most two mass points denoted by $\{x_{2,i}\}_{i=1}^2$. It is already known that the support of $F^*_{X_1}$ is bounded, which is denoted by $[A'_1, A'_2]$. Let $\mathscr{S}_1$ denote the set of all probability distributions on the Borel sets of $[A'_1, A'_2]$. The set

$$\mathscr{F}_1 = \left\{ F_{X_1} \in \mathscr{S}_1 \middle| \int p(0|x_1, x_{2,j}) dF_{X_1}(x_1) = p(0; F^*_{X_1}|x_{2,j}), \right. \\ \left. j \in [1:2] \right\},$$

is the intersection of $\mathscr{S}_1$ with two hyperplanes. In a similar way,

$$I^*_{F_{X_2}} = \sup_{F_{X_1} \in \mathscr{F}_1} \left\{ I_\lambda(F_{X_1} F_{X_2}) - \theta_1 \left(\int x^2 dF_{X_1}(x) - P'_1 \right) \right\}, \tag{16}$$

and having $F_{X_1} \in \mathscr{F}_1$, the objective function in (16) becomes

$$\underbrace{\lambda H(Y) + (1-\lambda) \sum_{i=1}^{2} p_{X_2}(x_{2,i}) H(Y|X_2 = x_{2,i})}_{\text{constant}} \\ \underbrace{- H(Y|X_1, X_2) - \theta_1 \left(\int x^2 dF_{X_1}(x) - P'_1 \right)}_{\text{linear in } F_{X_1}}. \tag{17}$$

Therefore, given any F_{X_2} with at most two points of increase, $F^*_{X_1}$ has at most three mass points.

When $\lambda = 1$, the term with summation in (17) disappears, which means that $\mathscr{F}_1$ could be replaced by

$$\left\{ F_{X_1} \in \mathscr{S}_1 | \int_{-\infty}^{+\infty} p(0|x_1) dF_{X_1}(x_1) = \tilde{p}^*_0 \right\},$$

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 one, it is concluded that $F^*_{X_1}$ has at most two points of increase. □

Remark 4. Note that the order of showing the discreteness of the support sets 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 two mass points. Then, for a given discrete F_{X_2} with at most two mass points, it is shown that $F^*_{X_1}$ is also discrete with at most three mass points (two 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.

Remark 5. (Settling a conjecture) Consider a point-to-point real AWGN channel with a K-bin (i.e., $\log_2 K$-bit) ADC front end. It is shown in [4] that the capacity-achieving input distribution for this channel (with average input power constraint), has at most $K+1$ mass points, while in the numerical results, K mass points always appear to be sufficient, which leaves the sufficiency of K mass points as a conjecture. Therefore, it has been an open problem whether K mass points are indeed sufficient or not. The answer is positive. If the average power constraint, which is a linear function of its corresponding input distribution, is treated as an intersecting hyperplane, Dubins' theorem states that $K+1$ mass points is sufficient. A simple trick, as used in the proof of Theorem 1, is to take the average power constraint into the objective function and take into account the uniqueness of the solution. This reduces the number of intersecting hyperplanes by one, and results in the sufficiency of K mass points. This is also the case for asymmetric quantizers (e.g., [12]), since this reduction of the number of hyperplanes does not rely on the structure of the quantizer. In conclusion, the number of mass points is not affected by any number of linear constraints (e.g., $\mathsf{E}[X^4] \le K$, etc) in the optimization.

V. Conclusion

We have studied the capacity region of a two-transmitter Gaussian MAC under average input power constraints at the transmitters and one-bit ADC front end at the receiver. We have shown that an auxiliary random variable is necessary for characterizing the capacity region. We have derived an upper bound on the cardinality of this auxiliary variable, and proved that the distributions that achieve the boundary points of the capacity region are finite and discrete. Based on this result, the evaluation of the capacity region and finding efficient suboptimal signaling schemes are subjects of our ongoing research. Finally, we settled the conjecture of the sufficiency of K mass points in a point to point AWGN channel with a K-bin quantizer at the receiver.

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/adcsurvey.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. Dabeer, 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] 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.

[7] A. Mezghani and J. Nossek, "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] B. Rassouli, M. Varasteh, and D. Gündüz, "Gaussian multiple access channels with one-bit quantizer at the receiver." [Online]. Available: https://arxiv.org/pdf/1703.02324.pdf

[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] T. Koch and A. Lapidoth, "At low SNR, asymmetric quantizers are better," *IEEE Trans. on Information Theory*, vol. 59, no. 9, pp. 5498–5512, Sep. 2013.