Analysis of Rayleigh-Fading Channels with 1-bit Quantized Output

Amine Mezghani and Josef A. Nossek
Institute for Circuit Theory and Signal Processing
Munich University of Technology, 80290 Munich, Germany
E-Mail: {Mezghani, Nossek}@nws.ei.tum.de

***Abstract*— We consider multi-input multi-output (MIMO) Rayleigh-fading channels with coarsely quantized outputs, where the channel is unknown to the transmitter and receiver. This analysis is of interest in the context of sensor network communication with low cost devices. The key point is that the analog-to-digital converters (ADCs) for such applications should be low-resolution, in order to reduce their cost and power consumption. In this paper, we consider the extreme case of only 1-bit ADC for each receive signal component. We elaborate on some properties of the mutual information compared to the unquantized case. For the SISO case, we show that on-off QPSK signaling is the capacity achieving distribution. To our knowledge, the block-wise Rayleigh-fading channel with mono-bit detection was not studied in the literature.**

I. Introduction

Several works studied MIMO channels operating in Rayleigh fading environments [1], [2], especially in the limit of low SNR [3], [4], [5] and high SNR [6]. Unfortunately, most of these contributions assume that the receiver has access to the channel data with infinite precision. In practice, however, a quantizer (analog-to-digital converter) is applied to the receive signal, so that the channel measurements can be processed in the digital domain. The reliance on high-resolution analog-to-digital converters (ADCs) easily becomes unjustified as soon as we have to deal with high speed MIMO channels [7]. In this case, the required high resolution ADCs are expensive and even no longer feasible. In fact, in order to reduce circuit complexity and save power and area, low resolution ADCs have to be employed [8], [9]. In [7], [10], we study the effects of quantization from an information theoretical point of view for MIMO systems, where the channel is perfectly known at the receiver. It turns out that the loss in channel capacity due to coarse quantization is surprisingly small at low to moderate SNR.

Motivated by these works, we aim to study the communication performance of Rayleigh-fading channel taking into account the coarse quantization. We consider the extreme case of 1-bit quantized (hard-decision detection) MIMO channel with no CSI at the transmitter and the receiver. When a single bit is used, the implementation of the all digital receiver is considerably simplified [11], [12], [13]. In particular, automatic gain control (AGC) is not needed.

Our paper is organized as follows. Section II describes the system model and notational issues. In Section III we give the general expression of the mutual information between the inputs and the quantized outputs of the MIMO system, then we elaborate on some of its properties in Section IV. In Section V, we derive the capacity for the SISO channel case.

II. System Model And Notation

We consider a point-to-point quantized MIMO channel where the transmitter employs M antennas and the receiver has N antennas. We assume a block-Rayleigh fading model [2], in which the channel propagation matrix $\boldsymbol{H} \in \mathbb{C}^{M\times N}$ remains constant for a coherence interval of length T symbols, and then changes to a new independent value. The entries of the channel matrix are i.i.d. zero-mean complex circular Gaussian with unit variance. The channel realizations are unknown to both the transmitter and receiver. At each coherence interval, the transmitted T symbols are denoted by the matrix $\boldsymbol{X} \in \mathbb{C}^{T\times M}$, whose row vectors are the transmitted signals at each time slot via the multiple antennas. Furthermore, the total unquantized channel output is denoted by $\boldsymbol{R} \in \mathbb{C}^{T\times N}$. The transmitted and received signal matrices are related as follows

$$\boldsymbol{R} = \boldsymbol{X}\boldsymbol{H} + \boldsymbol{W}, \tag{1}$$

where $\boldsymbol{W}$ represents the additive noise, whose entries are i.i.d. and distributed as $\mathcal{CN}(0,1)$. The input matrix $\boldsymbol{X}$ satisfies the average power constraint

$$\mathrm{E}[\mathrm{tr}(\boldsymbol{X}\boldsymbol{X}^{\mathrm{H}})] = T\cdot \mathsf{SNR}, \tag{2}$$

where SNR represents the average signal-to-noise ratio at each receive antenna.

In our system, the real parts $r_{t,i,R}$ and the imaginary parts $r_{t,i,I}$ of the receive signal components $r_{t,i}$, $1\leq t\leq T$ and $1\leq i\leq N$, are each quantized by a 1-bit resolution quantizer. Thus, the resulting quantized signals read as

$$y_{t,i,c} = \mathrm{sign}(r_{t,i,c}) \in \{-1,1\},\ \text{for } c\in\{R,I\},\quad 1\leq i\leq N,\ \text{and } 1\leq t\leq T. \tag{3}$$

Throughout our paper, a_i denotes the i-th element of the vector $\boldsymbol{a}$ and $[\boldsymbol{a}]_{i,c} = a_{i,c}$ with $c\in\{R,I\}$ is the real or imaginary part of a_i. The operators $(\cdot)^{\mathrm{H}}$ and $\mathrm{tr}(\cdot)$ stand for Hermitian transpose and trace of a matrix, respectively. Vectors and matrices are denoted by lower and upper case italic bold letters.

III. Mutual Information

The mutual information between the channel input and the quantized output of the presented system reads as [14]

$$\begin{aligned} I(\boldsymbol{X},\boldsymbol{Y}) &= H(\boldsymbol{Y}) - H(\boldsymbol{Y}|\boldsymbol{X}) \\ &= \mathrm{E}_{\boldsymbol{X}}\left[\sum_{\boldsymbol{Y}} P(\boldsymbol{Y}|\boldsymbol{X})\log\frac{P(\boldsymbol{Y}|\boldsymbol{X})}{P(\boldsymbol{Y})}\right], \end{aligned} \tag{4}$$

with $P(\boldsymbol{Y}) = \mathrm{E}_{\boldsymbol{X}}[P(\boldsymbol{Y}|\boldsymbol{X})]$ and $\mathrm{E}_{\boldsymbol{X}}[\cdot]$ is the expectation taken with respect to the distribution $p(\boldsymbol{X})$. Herewith,

978-1-4244-2571-6/08/$25.00 ©2008 IEEE

$H(\cdot)$ and $H(\cdot|\cdot)$ represent the entropy and the conditional entropy, respectively. Given the input $\boldsymbol{X}$, the unquantized output $\boldsymbol{R}$ is zero-mean complex Gaussian with covariance $\mathrm{E}[\boldsymbol{R}\boldsymbol{R}^{\mathrm{H}}|\boldsymbol{X}] = N(\mathbf{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathrm{H}})$, and thus we have [2]

$$p(\boldsymbol{R}|\boldsymbol{X}) = \frac{\exp(-\mathrm{tr}(\boldsymbol{R}^{\mathrm{H}}(\mathbf{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathrm{H}})^{-1}\boldsymbol{R}))}{\pi^{NT}[\det(\mathbf{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathrm{H}})]^N}. \tag{5}$$

Afterwards, we can express the conditional probability of the quantized output as

$$\begin{aligned}P(\boldsymbol{Y}|\boldsymbol{X}) &= \int_0^\infty\!\!\cdots\!\int_0^\infty p(\boldsymbol{Y}\bullet\boldsymbol{R}|\boldsymbol{X})\mathrm{d}\boldsymbol{R} = \\ &\int_0^\infty\!\!\cdots\!\int_0^\infty \frac{\exp(-\mathrm{tr}((\boldsymbol{Y}\bullet\boldsymbol{R})^{\mathrm{H}}(\mathbf{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathrm{H}})^{-1}(\boldsymbol{Y}\bullet\boldsymbol{R})))}{\pi^{NT}[\det(\mathbf{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathrm{H}})]^N}\mathrm{d}\boldsymbol{R} \\ &= \prod_{i=1}^{N}\int_0^\infty\!\!\cdots\!\int_0^\infty \frac{\exp(-(\boldsymbol{y}_i\bullet\boldsymbol{r})^{\mathrm{H}}(\mathbf{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathrm{H}})^{-1}(\boldsymbol{y}_i\bullet\boldsymbol{r}))}{\pi^{T}[\det(\mathbf{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathrm{H}})]}\mathrm{d}\boldsymbol{r},\end{aligned} \tag{6}$$

where the integration is performed over the positive orthant of the complex hyperplane and $\boldsymbol{Y}\bullet\boldsymbol{R}$ denotes an element-wise matrix product with $[\boldsymbol{Y}\bullet\boldsymbol{R}]_{i,j,c} = y_{i,j,c}r_{i,j,c}$.
The conditional probability $P(\boldsymbol{Y}|\boldsymbol{X})$ have the following properties

$$P(\boldsymbol{Y}|\boldsymbol{X}\Psi^{\mathrm{H}}) = P(\boldsymbol{Y}|\boldsymbol{X}) \tag{7}$$

$$P(\boldsymbol{Y}|\boldsymbol{P}\boldsymbol{X}) = P(\boldsymbol{P}^{\mathrm{H}}\boldsymbol{Y}|\boldsymbol{X}), \tag{8}$$

where $\boldsymbol{\Psi}$ is a unitary matrix and $\boldsymbol{P}$ is a complex permutation matrix [1]. Property (7) is obvious, property (8) follows since, for a permutation matrix, the following holds

$$\boldsymbol{P}(\boldsymbol{Y}\bullet\boldsymbol{R}) = (\boldsymbol{P}\boldsymbol{Y})\bullet(\boldsymbol{P}\boldsymbol{R}), \tag{9}$$

and by changing the variables of integration from $\boldsymbol{R}$ to $\boldsymbol{P}\boldsymbol{R}$. Note that (8) holds only for a complex permutation matrix and not for a general unitary matrix, as in the unquantized case (c.f. [2]).

IV. Properties of the Mutual Information

In this section, we derive some properties of the mutual information based on the structure of the conditional probabilities (6), where some similarities and differences to the work [2] for the unquantized case can be observed.

Lemma 1: If the coherence time is equal 1, then the mutual information is equal zero, independently of the input density and the number of antennas.

Proof: For $T = 1$, we have from (6)

$$\begin{aligned}P(\boldsymbol{Y}|\boldsymbol{X}) &= \prod_{i,c}\int_0^\infty \frac{\exp(-y_{i,c}^2 r^2/(1+\sum_{i,c}x_{i,c}^2))}{\sqrt{\pi(1+\sum_{i,c}x_{i,c}^2)}}\mathrm{d}r \\ &= \prod_{i,c}\frac{1}{2} = \frac{1}{2^{2N}}.\end{aligned} \tag{10}$$

Obviously all outputs are equiprobable independently from the input density and thus $I(\boldsymbol{Y},\boldsymbol{X}) = 0$. Intuitively, for $T = 1$, all information is contained in the magnitude of the signal $\boldsymbol{R}$, and simple mono-bit conversion leads to no mutual information between the input and the quantized output and thus, to a zero capacity.

■

Lemma 2: The mutual information is invariant to a transformation of the input probability density from $p_0(\boldsymbol{X})$ to $p_1(\boldsymbol{X}) = p_0(\boldsymbol{P}\boldsymbol{X}\boldsymbol{\Psi}^{\mathrm{H}})$ with given complex permutation matrix $\boldsymbol{P}$ and unitary matrix $\boldsymbol{\Psi}^{\mathrm{H}}$.

Proof: Let I_0 and I_1 denote the mutual information corresponding to the input distribution $p_0(\boldsymbol{X})$ and $p_1(\boldsymbol{X})$ respectively, by means of equalities (7) and (8), we have

$$\begin{aligned}I_1 &= \int p_0(\boldsymbol{X})\left[\sum_{\boldsymbol{Y}} P(\boldsymbol{Y}|\boldsymbol{P}^{\mathrm{H}}\boldsymbol{X}\boldsymbol{\Psi})\ln\frac{P(\boldsymbol{Y}|\boldsymbol{P}^{\mathrm{H}}\boldsymbol{X}\boldsymbol{\Psi})}{\mathrm{E}_{\boldsymbol{X}}[P(\boldsymbol{Y}|\boldsymbol{P}^{\mathrm{H}}\boldsymbol{X}\boldsymbol{\Psi})]}\right]\mathrm{d}\boldsymbol{X} \\ &= \int p_0(\boldsymbol{X})\left[\sum_{\boldsymbol{Y}} P(\boldsymbol{P}\boldsymbol{Y}|\boldsymbol{X})\ln\frac{P(\boldsymbol{P}\boldsymbol{Y}|\boldsymbol{X})}{\mathrm{E}_{\boldsymbol{X}}[P(\boldsymbol{P}\boldsymbol{Y}|\boldsymbol{X})]}\right]\mathrm{d}\boldsymbol{X} \\ &= \int p_0(\boldsymbol{X})\left[\sum_{\boldsymbol{Y}'} P(\boldsymbol{Y}'|\boldsymbol{X})\ln\frac{P(\boldsymbol{Y}'|\boldsymbol{X})}{\mathrm{E}_{\boldsymbol{X}}[P(\boldsymbol{Y}'|\boldsymbol{X})]}\right]\mathrm{d}\boldsymbol{X} \\ &= I_0,\end{aligned} \tag{11}$$

where $\mathrm{E}_{\boldsymbol{X}}[\cdot]$ is the expectation taken with respect to the distribution $p_0(\boldsymbol{X})$, and we have substituted $\boldsymbol{Y}' = \boldsymbol{P}\boldsymbol{Y}$ in the third step.

■

Lemma 3: Let us consider the MISO case ($N = 1$). Starting from any input distribution $p_0(\boldsymbol{X})$, we can find a distribution $p_1(\boldsymbol{X})$ that generates as least as much mutual information with symmetrical (equiprobable) outputs, i.e., $H(\boldsymbol{y}) = 2T$.

Proof: This lemma is simply justified by symmetry considerations. Let $p_1(\boldsymbol{X})$ be a mixture involving the transformations of the $p_0(\boldsymbol{X})$ by all possible diagonal complex permutation matrices namely

$$p_1(\boldsymbol{X}) = \frac{1}{4^T}\sum_{k=1}^{4^T} p_0(\boldsymbol{P}_k\boldsymbol{X}). \tag{12}$$

Obviously, this transformation does not change the average power of the transmit signal. Besides, due to lemma 2 and the concavity of the mutual information with respect to the input distribution, it can be easily shown that higher mutual information is achieved by $p_1(\boldsymbol{X})$ than by $p_0(\boldsymbol{X})$. On the other hand, we have the output distribution under $p_1(\boldsymbol{X})$ as

$$\begin{aligned}P(\boldsymbol{y}) &= \int P(\boldsymbol{y}|\boldsymbol{X})\frac{1}{4^T}\sum_{k=1}^{4^T} p_0(\boldsymbol{P}_k\boldsymbol{X})\mathrm{d}\boldsymbol{X} \\ &= \frac{1}{4^T}\sum_{k=1}^{4^T}\int P(\boldsymbol{y}|\boldsymbol{X})p_0(\boldsymbol{P}_k\boldsymbol{X})\mathrm{d}\boldsymbol{X} \\ &= \frac{1}{4^T}\int\sum_{k=1}^{4^T} P(\boldsymbol{y}|\boldsymbol{P}_k^{\mathrm{H}}\boldsymbol{X}')p_0(\boldsymbol{X}')\mathrm{d}\boldsymbol{X}' \\ &= \frac{1}{4^T}\int\sum_{k=1}^{4^T} P(\boldsymbol{P}_k\boldsymbol{y}|\boldsymbol{X}')p_0(\boldsymbol{X}')\mathrm{d}\boldsymbol{X}' \\ &= \frac{1}{4^T}\int\sum_{\boldsymbol{y}} P(\boldsymbol{y}|\boldsymbol{X}')p_0(\boldsymbol{X}')\mathrm{d}\boldsymbol{X}' \\ &= \frac{1}{4^T}\int p_0(\boldsymbol{X}')\mathrm{d}\boldsymbol{X}' = \frac{1}{4^T},\end{aligned} \tag{13}$$

[1] A complex permutation-matrix $\boldsymbol{P}$ has all of its elements zero, except for a single one in each row and column drawn from the set {-1, -j, j, 1}.

where we substituted $\boldsymbol{X}' = \boldsymbol{P}\boldsymbol{X}$. This completes the proof. ■

This lemma means that, for the MISO case, there exists a capacity achieving distribution that generates equiprobable outputs. Thus, the MISO capacity can be found by performing the following optimization

$$C = 2T - \min_{p(\boldsymbol{X})} H(\boldsymbol{y}|\boldsymbol{X}), \tag{14}$$

over permutation invariant distributions. This result holds not necessarily for the general MIMO case, and we expect that the 4^{NT} possible outputs are not equally likely when achieving the capacity.

In the rest of the paper, we will treat the SISO ($M = N = 1$) case in more details.

V. Capacity and Optimal Input for the SISO Case

In this section we consider the SISO case, $N = M = 1$. Since the covariance matrix $\mathrm{E}[\boldsymbol{r}\boldsymbol{r}^{\mathrm{H}}|\boldsymbol{x}] = (\mathbf{I}_T + \boldsymbol{x}\boldsymbol{x}^{\mathrm{H}})$ is the sum of a diagonal matrix (unit matrix) and a rank one matrix, we can use a general form of the Dunnett-Sobel decomposition [15] to replace the multidimensional integral in (6) by a double integral. That is

$$P(\boldsymbol{y}|\boldsymbol{x}) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\frac{u^2+v^2}{2}} \prod_{t=1}^{T} \Phi((x_{t,R}u - x_{t,I}v)y_{t,R}) \cdot \Phi((x_{t,R}v + x_{t,I}u)y_{t,I})\mathrm{d}u\mathrm{d}v, \tag{15}$$

with $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt$ is the cumulative normal distribution function. For $T = 2$ and $T = 3$, closed form solutions for this integral are found in [16], which will be used later.

Now, we state a theorem on the capacity achieving distribution of this mono-bit quantized SISO channel. For this, we define

$$\begin{aligned} R^{\mathrm{QPSK}}(A) &= \frac{1}{4^T}\sum_{l=0}^{4^T-1}\sum_{\boldsymbol{y}} P(\boldsymbol{y}|\boldsymbol{x}^l)\log(4^T P(\boldsymbol{y}|\boldsymbol{x}^l)) \\ &= \sum_{\boldsymbol{y}} P(\boldsymbol{y}|\boldsymbol{x}^0)\log(4^T P(\boldsymbol{y}|\boldsymbol{x}^0)), \end{aligned} \tag{16}$$

as the achievable rate with QPSK symbols of power A. Here, $\boldsymbol{x}^l$, $0 \le l \le 4^T - 1$, are all possible sequences of T equally likely QPSK data symbols, i.e., $x_t^l \in \{\sqrt{A}, -\sqrt{A}, \mathrm{j}\sqrt{A}, -\mathrm{j}\sqrt{A}\}$, $1 \le t \le T$. $\boldsymbol{x}^0$ denotes the sequence with $x_t^0 = \sqrt{A}, \forall t$. The second equality in (17) follows due using symmetry of QPSK and expression (15). Furthermore, we use (15) to get a simpler expression for $P(\boldsymbol{y}|\boldsymbol{x}^0)$ as

$$\begin{aligned} &P(\boldsymbol{y}|\boldsymbol{x}^0) = \\ &\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} \prod_{t=1}^{T}\Phi(\sqrt{A}y_{t,R}u)\mathrm{d}u \int_{-\infty}^{+\infty} e^{-\frac{v^2}{2}} \prod_{t=1}^{T}\Phi(\sqrt{A}y_{t,I}v)\mathrm{d}v \\ &= P(\mathrm{Re}\{\boldsymbol{y}\}|\boldsymbol{x}^0)\cdot P(\mathrm{Im}\{\boldsymbol{y}\}|\boldsymbol{x}^0). \end{aligned} \tag{17}$$

Then (17) simplifies to

$$\begin{aligned} R^{\mathrm{QPSK}}(A) &= 2\sum_{t, y_t=\pm 1}\left(\int_{-\infty}^{+\infty}\frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}\prod_{t=1}^{T}\Phi(\sqrt{A}y_t u)\mathrm{d}u\right)\cdot \\ &\quad \log\left(2^T\int_{-\infty}^{+\infty}\frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}\prod_{t=1}^{T}\Phi(\sqrt{A}y_t u)\mathrm{d}u\right) \\ &= 2\sum_{k=0}^{T}\binom{T}{k}\left(\int_{-\infty}^{+\infty}\frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}\Phi(-\sqrt{A}u)^k\Phi(\sqrt{A}u)^{T-k}\mathrm{d}u\right)\cdot \\ &\quad \log\left(2^T\int_{-\infty}^{+\infty}\frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}\Phi(-\sqrt{A}u)^k\Phi(\sqrt{A}u)^{T-k}\mathrm{d}u\right). \end{aligned} \tag{18}$$

Theorem 1: For $N = M = 1$, on-off QPSK is optimal[2], where the sender transmits, during each coherence time block T, either a zero sequence with probability $1 - p$, or QPSK symbols with probability p and power SNR/p. Here, the duty cycle p is given by

$$p = \begin{cases} \frac{\mathsf{SNR}}{\mathsf{SNR}_{\mathrm{crit}}} & \mathsf{SNR} < \mathsf{SNR}_{\mathrm{crit}} \\ 1 & \text{otherwise.} \end{cases} \tag{19}$$

The critical SNR, $\mathsf{SNR}_{\mathrm{crit}}$, is the threshold below which a higher capacity can be achieved by signaling with a duty cycle less than unity. $\mathsf{SNR}_{\mathrm{crit}}$ results from the following optimization

$$\mathsf{SNR}_{\mathrm{crit}} = \operatorname*{argmax}_{A} \frac{1}{A} R^{\mathrm{QPSK}}(A) \quad \text{st.} \quad A \ge 0. \tag{20}$$

The capacity per block is then given by

$$C = \begin{cases} \frac{\mathsf{SNR}}{\mathsf{SNR}_{\mathrm{crit}}} \overbrace{R^{\mathrm{QPSK}}(\mathsf{SNR}_{\mathrm{crit}})}^{C_{\mathrm{crit}}} & \mathsf{SNR} \le \mathsf{SNR}_{\mathrm{crit}} \\ R^{\mathrm{QPSK}}(\mathsf{SNR}) & \text{otherwise.} \end{cases} \tag{21}$$

Proof: First, as it can be shown that the function $R^{\mathrm{QPSK}}(A)/A$ is concave in A (see later for the special cases $T = 2$ and $T = 3$), the maximizing value given in (22) is unique and can be obtained by differentiating with respect to A. Therefore, $\mathsf{SNR}_{\mathrm{crit}}$ that solves (21) is the unique solution of

$$\frac{R^{\mathrm{QPSK}}(\mathsf{SNR}_{\mathrm{crit}})}{\mathsf{SNR}_{\mathrm{crit}}} = \left.\frac{\mathrm{d}R^{\mathrm{QPSK}}(A)}{\mathrm{d}A}\right|_{A=\mathsf{SNR}_{\mathrm{crit}}}. \tag{22}$$

The proof is based on solving the Kuhn-Tucker condition. Since we have to do with a convex optimization problem in (14)[3], a sufficient condition for an input distribution $p^*(\boldsymbol{x})$ to achieve the capacity C claimed in (22) is to be invariant to any complex permutation (see lemma 3) and $\exists\gamma \ge 0$ such that

$$\gamma(||\boldsymbol{x}||^2 - T\mathsf{SNR}) + C - \sum_{\boldsymbol{y}} P(\boldsymbol{y}|\boldsymbol{x})\log(4^T P(\boldsymbol{y}|\boldsymbol{x})) \ge 0, \tag{23}$$

[2] Actually, when the transmitter is on, any rotated version of QPSK is optimal due to property (7). Besides the first symbol x_1 of $\boldsymbol{x}$ can be fixed to $\sqrt{\mathsf{SNR}/p}$, which gives the well-known differential QPSK.

[3] Note that $H(\boldsymbol{y}|\boldsymbol{x}) = \int p(\boldsymbol{x})\sum_{\boldsymbol{y}} P(\boldsymbol{y}|\boldsymbol{x})\log(P(\boldsymbol{y}|\boldsymbol{x}))d\boldsymbol{x}$ is a linear functional of $p(\boldsymbol{x})$.

for all $\boldsymbol{x}$, with equality if $\boldsymbol{x}$ is in the support of $p^*(\boldsymbol{x})$.

Let us define the function $f(\boldsymbol{x})$ as

$$f(\boldsymbol{x}) = \gamma(||\boldsymbol{x}||^2 - T\mathsf{SNR}) + C - \sum_{\boldsymbol{y}} P(\boldsymbol{y}|\boldsymbol{x}) \log(4^T P(\boldsymbol{y}|\boldsymbol{x})) \tag{24}$$

for fixed γ. Now, choosing

$$\gamma = \begin{cases} \frac{C_{\text{crit}}}{T\mathsf{SNR}_{\text{crit}}} & \mathsf{SNR} < \mathsf{SNR}_{\text{crit}} \\ \frac{1}{2\sqrt{\mathsf{SNR}}} \left.\frac{\mathrm{d}R^{\text{QPSK}}(A)}{\mathrm{d}A}\right|_{A=\mathsf{SNR}} & \text{otherwise,} \end{cases} \tag{25}$$

we can then verify that the point $\hat{\boldsymbol{x}}$ with $\hat{x}_t = \sqrt{\mathsf{SNR}/p}$, $\forall t$, satisfies

$$f(\hat{\boldsymbol{x}}) = 0, \tag{26}$$

and is a stationary point of the function $f(\boldsymbol{x})$, i.e.,

$$\nabla_{\boldsymbol{x}} f(\boldsymbol{x})|_{\boldsymbol{x}=\hat{\boldsymbol{x}}} = \mathbf{0}. \tag{27}$$

It can be further shown that $\hat{\boldsymbol{x}}$ is a global minimizer of $f(\boldsymbol{x})$ and the corresponding minimum value is 0, and thus the KKT condition (24) is verified. The details of the proof is omitted due to lake of space and since the result seems quite intuitive.[4]

■

Actually, it is well known that in the absence of CSI at both the transmitter and the receiver, and for vanishing SNR, the capacity can be achieved by on-off (peaky) signaling schemes. In the unquantized increasingly peaky signals are needed, i.e., the peak level goes to infinity as the SNR goes to 0 [3]. In the quantized case, however, the peak level is fixed to $\sqrt{\mathsf{SNR}_{\text{crit}}}$ as long as $\mathsf{SNR} \leq \mathsf{SNR}_{\text{crit}}$, and the capacity grows linearly up to C_{crit}. In this regime, the normalized energy per bit can be obtained from (22) as

$$\frac{E_b}{N_0} = T \left.\frac{\mathsf{SNR}}{C}\right|_{\mathsf{SNR} \leq \mathsf{SNR}_{\text{crit}}} = T\frac{\mathsf{SNR}_{\text{crit}}}{C_{\text{crit}}}. \tag{28}$$

For the special case $T = 2$, we found in [16] the following closed form solution for the integral in (19)

$$\begin{aligned} &\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} \prod_{t=1}^{2} \Phi(\sqrt{A} y_{t,c} u) \mathrm{d}u = \\ &\qquad \frac{1}{4}\left[1 + \frac{2}{\pi} \arcsin(\frac{A}{1+A}) y_{1,c} y_{2,c}\right]. \end{aligned} \tag{29}$$

Thus (19) becomes

$$\begin{aligned} &R^{\text{QPSK}}(A)|_{T=2} = \\ &= \left(1 + \frac{2}{\pi}\arcsin(\frac{A}{1+A})\right) \log\left(1 + \frac{2}{\pi}\arcsin(\frac{A}{1+A})\right) \\ &+ \left(1 - \frac{2}{\pi}\arcsin(\frac{A}{1+A})\right) \log\left(1 - \frac{2}{\pi}\arcsin(\frac{A}{1+A})\right). \end{aligned} \tag{30}$$

[4]Note that any point equal to $\boldsymbol{P}\hat{\boldsymbol{x}}e^{\mathrm{j}\varphi}$, for given complex permutation matrix $\boldsymbol{P}$ and real number φ, is a global minimizer of $f(\boldsymbol{x})$. The origin $\boldsymbol{x} = \mathbf{0}$ is also a stationary point and is a global minimizer if $\mathsf{SNR} \leq \mathsf{SNR}_{\text{crit}}$. These are the only stationary point of the function $f(\boldsymbol{x})$ due to special structure of the conditional probability given in (15), and thus the only potential global minimizers, with a minimal value of 0.

And for the case $T = 3$, we substitute

$$\begin{aligned} &\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}} \prod_{t=1}^{3} \Phi(\sqrt{A} y_{t,c} u) \mathrm{d}u = \\ &\frac{1}{8}\left[1 + \frac{2}{\pi}\left(\arcsin(\frac{A}{1+A}) y_{1,c} y_{2,c} + \right.\right. \\ &\left.\left. + \arcsin(\frac{A}{1+A}) y_{1,c} y_{3,c} + \arcsin(\frac{A}{1+A}) y_{2,c} y_{3,c}\right)\right] \end{aligned} \tag{31}$$

(see also [16]) in (19) and we obtain

$$\begin{aligned} &R^{\text{QPSK}}(A)|_{T=3} = \\ &= \left(\frac{1}{2} + \frac{3}{\pi}\arcsin(\frac{A}{1+A})\right) \log\left(1 + \frac{6}{\pi}\arcsin(\frac{A}{1+A})\right) \\ &+ \left(\frac{3}{2} - \frac{3}{\pi}\arcsin(\frac{A}{1+A})\right) \log\left(1 - \frac{2}{\pi}\arcsin(\frac{A}{1+A})\right). \end{aligned} \tag{32}$$

In both case, $R^{\text{QPSK}}(A)/A$ is a concave function of A . Solving the optimization (21) numerically, we found the following values for $\mathsf{SNR}_{\text{crit}}$ and C_{crit} at $T = 2$

$$\begin{aligned} \mathsf{SNR}_{\text{crit},T=2} &= 1.361648, \\ C_{\text{crit},T=2} &= 0.226813. \end{aligned} \tag{33}$$

And for $T = 3$, we obtained

$$\begin{aligned} \mathsf{SNR}_{\text{crit},T=3} &= 1.139946, \\ C_{\text{crit},T=3} &= 0.474754. \end{aligned} \tag{34}$$

Obviously the critical SNR, $\mathsf{SNR}_{\text{crit}}$ decreases as the coherence block length T increases and goes to zero as $T \to \infty$. A closed form solution for the integral in (18) is for $T > 3$ unknown and only approximations are found in the literature [16]. Finally, formula (22) can be used to obtain the capacity.

In Fig. 1, we plotted the capacity of the mono-bit SISO Rayleigh-fading channel per channel use over the SNR for the two cases where the coherence time $T = 2$ and $T = 3$, both normalized by T. The squares indicates the threshold $(\mathsf{SNR}_{\text{crit}}, C_{\text{crit}})$, up to which the capacity grows linearly with the SNR. The coherent capacity achieved by quantized QPSK, which is an upper bound on our non-coherent capacity, is also shown. It is obtained assuming that the receiver knows the channel coefficient h. That is the average capacity achieved by QPSK over the quantized coherent Rayleigh-fading channel

$$\begin{aligned} C_{\text{coh}} &= \mathrm{E}_h[H(x, y|h)] \\ &= \mathrm{E}_h\left[2 - \sum_x \sum_y P(y|x,h) \log P(y|x,h)\right], \end{aligned} \tag{35}$$

where we substituted $H(y) = 2$, since all four possible outputs are equiprobable, and

$$P(y|x,h) = \Phi\left(\frac{\mathrm{Re}[y]\mathrm{Re}[hx]}{\sqrt{1/2}}\right) \Phi\left(\frac{\mathrm{Im}[y]\mathrm{Im}[hx]}{\sqrt{1/2}}\right). \tag{36}$$

In the coherent case, the ratio of bit-energy to noise density at vanishing SNR is given by [3], [10]

$$\left.\frac{E_{b,\text{coh}}}{N_0}\right|_{\mathsf{SNR}\to 0} = \frac{\pi}{2}\ln 2 \geq T\frac{\mathsf{SNR}_{\text{crit}}}{C_{\text{crit}}}, \quad \text{from (29).} \tag{37}$$

In Fig. 2, the normalized capacity of the mono-bit Rayleigh-fading SISO channel versus coherence interval T for SNR=10dB. The coherent upper-bound is also plotted.

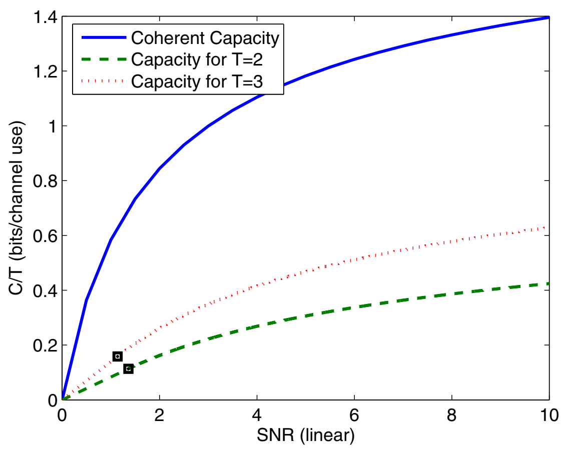

Fig. 1. Capacity of the mono-bit Rayleigh-fading SISO channel for $T = 2$, $T = 3$, $T = \infty$.

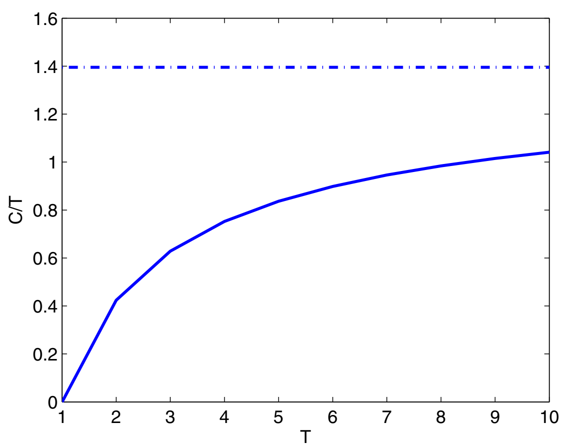

Fig. 2. Capacity of the mono-bit Rayleigh-fading SISO channel versus coherence interval T for SNR=10dB.

VI. Conclusion

We studied the mutual information of the Rayleigh-fading MIMO channels with one-bit ADC, in the absence of channel knowledge at the transmitter, and some of its properties compared to the unquantized case [2]. Then, we analysed the SISO case in details. Contrary to the unquantized channel, the non-coherent SISO capacity is achieved by on-off QPSK for the whole SNR range, where the optimal duty cycle depends only on the coherence time. We developed expressions for the achievable rates, with closed form for $T = 2$ and $T = 3$. Deriving the structure of the capacity for the general MIMO case is a more difficult task and could be a research topic for the future, especially, asymptotic analysis (at low SNR) could be investigated. Comparison to pilot-aided schemes and extension to general correlated fading channels, under such coarse quantization, are also interesting issues.

References

[1] Ibrahim C. Abou-Faycal, Mitchell D. Trott, and Shlomo Shamai, "The capacity of discrete-time memoryless Rayleigh-fading channels," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1290–1301, May 2001.

[2] T.L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 45, no. 1, pp. 139–157, January 1999.

[3] S. Verdú, "Spectral Efficiency in the Wideband Regime," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1319–1343, June 2002.

[4] L. Zheng, D. N. C. Tse, and M. Medard, "Channel Coherence in the low SNR Regime," *IEEE International Symposium on Information Theory (ISIT), Chicago, IL*, June 2004.

[5] X. Wu and R. Srikant, "MIMO Channels in the Low SNR Regime: Communication Rate, Error Exponent and Signal Peakiness," in *IEEE Information theory workshop*, San Antonio, TX, 2004.

[6] L. Zheng and D. N. C. Tse, "Communication on the Grassman manifold: A geometric approach to the non-coherent multiple-antenna channel," *IEEE Trans. Inform. Theory*, vol. 48, no. 2, pp. 359–383, February 2002.

[7] J. A. Nossek and M. T. Ivrlač, "Capacity and coding for quantized MIMO systems," in *Intern. Wireless Commun. and Mobile Computing Conf. (IWCMC)*, Vancouver, Canada, July 2006, pp. 1387–1392, invited.

[8] R. Schreier and G. C. Temes, "Understanding Delta-Sigma Data Converters," *IEEE Computer Society Press*, 2004.

[9] D. D. Wentzloff, R. Blázquez, F. S. Lee, B. P. Ginsburg, J. Powell, and A. P. Chandrakasan, "System design considerations for ultra-wideband communication," *IEEE Commun. Mag.*, vol. 43, no. 8, pp. 114–121, Aug. 2005.

[10] A. Mezghani and J. A. Nossek, "On Ultra-Wideband MIMO Systems with 1-bit Quantized Outputs: Performance Analysis and Input Optimization," *IEEE International Symposium on Information Theory (ISIT), Nice, France*, June 2007.

[11] S. Hoyos, B. M. Sadler, and G. R. Arce, "Mono-bit digital receivers for ultra-wideband communications," *IEEE Transactions Letters on Wireless Communications*, vol. 4, no. 4, pp. 1337–1344, July 2005.

[12] I. D. O'Donnell and R. W. Brodersen, "An Ultra-Wideband Transceiver Architecture for Low Power, Low Rate, Wireless systems," *IEEE Trans. on Vehicular Technology*, vol. 54, no. 5, pp. 1623–1631, September 2005.

[13] R. Blázquez, P. P. Newaskar, F. S. Lee, and A. P. Chandrakasan, "A baseband processor for impulse ultra-wideband communications," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 9, pp. 1821–1829, September 2005.

[14] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley and Son, New York, 1991.

[15] C.W. Dunnett and M. Sobel, "Approximations to the probability integral and certain percentage points of a multivariate analogue of student's t-distribution," *Biometrika*, vol. 42, pp. 258–260, 1955.

[16] Ralph Hoyt Bacon, "Approximations to Multivariate Normal Orthant Probabilities," *The Annals of Mathematical Statistics*, vol. 34, no. 1, pp. 191–198, March 1963.