Research Article

Achieving Maximum Possible Speed on Constrained Block Transmission Systems

Obianuju Ndili and Tokunbo Ogunfunmi

Department of Electrical Engineering, Santa Clara University, Santa Clara, CA 95053, USA

We develop a theoretical framework for achieving the maximum possible speed on constrained digital channels with a finite alphabet. A common inaccuracy that is made when computing the capacity of digital channels is to assume that the inputs and outputs of the channel are analog Gaussian random variables, and then based upon that assumption, invoke the Shannon capacity bound for an additive white Gaussian noise (AWGN) channel. In a channel utilizing a finite set of inputs and outputs, clearly the inputs are not Gaussian distributed and Shannon bound is not exact. We study the capacity of a block transmission AWGN channel with quantized inputs and outputs, given the simultaneous constraints that the channel is frequency selective, there exists an average power constraint P at the transmitter and the inputs of the channel are quantized. The channel is assumed known at the transmitter. We obtain the capacity of the channel numerically, using a constrained Blahut-Arimoto algorithm which incorporates an average power constraint P at the transmitter. Our simulations show that under certain conditions the capacity approaches very closely the Shannon bound. We also show the maximizing input distributions. The theoretical framework developed in this paper is applied to a practical example: the downlink channel of a dial-up PCM modem connection where the inputs to the channel are quantized and the outputs are real. We test how accurate is the bound 53.3 kbps for this channel. Our results show that this bound can be improved upon.

1. INTRODUCTION

The performance of all digital modems is affected by the precision of analog-to-digital (A/D) and digital-to-analog (D/A) conversions. Quantization distortion which limits the performance of the system is introduced as a result of analog-to-digital conversions. There are two different situations: one consists in designing a modem together with the A/D, D/A converters that interface a given analog channel and the other consists in designing a modem to face a channel which is part analog and part digital with a preexistent D/A and/or A/D conversion included. An example of this last case can be found in use when the modem sends or receives digital data across the public switched telephone network (PSTN). The core network of the PSTN today has evolved into an all-digital transport medium supported by optical communications. The access is mostly through twisted pairs of copper wires that are terminated by a PCM conversion. "Dial-up" is a technology that allows users to do this. In the uplink connection, the user's data is converted to an appropriately band-limited analog signal by the user's network interface hardware. Common examples of network interface hardware

include PCM modems and ADSL modems. The standards governing the design of these modems are the ITU-T V.90 standards for PCM modems and T1.413 standards for DSL modems.

After the D/A conversion of the user's data, the resulting analog signal is transmitted via an analog channel (twisted pair of copper wires) to the network service provider (NSP) [1, 2]. Here the analog signal is converted into a digital signal and transmitted via a digital link (optical fiber) to the PSTN. At the NSP, the modem used is a PCM modem which utilizes a nonlinear amplitude modulation scheme designed for acceptable voice communication over the digital PSTN. In the USA this nonlinear amplitude modulation scheme is called the μ-law encoding rule, while in Europe a similar encoding rule called A-law encoding rule is used. Communication in the downlink direction is the reverse of communication in the uplink direction. Due to the finite alphabet of the μ-law and A-law encoding rules and the avoidance of an A/D conversion, the theoretical capacity of downlink, V.90 dial-up communication is 56 kbps [3]. However this capacity is further limited by AWGN in the channel and the federal communications commission (FCC) restriction on

the average transmit power P, where $P \leq -12$ dBm. Some papers indicate that 53.3 kbps is the expected bit rate but they do not give details on how this bound was obtained [1, 3].

A common inaccuracy made when computing the capacity of digital channels is in making the assumption that the inputs and outputs of the channel are analog Gaussian random variables and then using the Shannon capacity bound for an AWGN channel (refer to Section 2, (4)) [1, 4]. Since the DSP hardware used in digital modems utilize a finite signal set with finite precision, it is clear that the inputs of the channel are not Gaussian and Shannon bound is not exact. The question that naturally arises is in what region and for what parameters of the A/D, D/A converters we can rely upon the analog channel approximation? Our purpose in this paper is to propose these conditions given the following constraints.

(a) First we consider a channel whose inputs $\mathbf{x} \in \mathcal{X}$ and outputs $\mathbf{y} \in \mathcal{Y}$ are chosen from a finite set of possibilities. Next we consider a special case of this channel, one with a finite set of inputs and an infinite set of outputs.

(2) There exists an average power constraint P on the input signals (see Section 2, (3)).

(3) The channel is an ISI channel represented by the circulant matrix $\mathbf{H}$, whose rows are circular shifts of the ISI channel fading coefficients. The channel is assumed known at the transmitter.

Our conclusions are that the performance of the quantized block transmission channel approaches that of the analog channel when we constrain the quantized channel to approximate the analog channel, by increasing peak-to-average power ratio. We will apply the theoretical framework developed in this paper, to a practical numerical example which is the downlink dial-up connection. Using this example we aim to test how accurate is the bound of 53.3 kbps for this channel, under a reasonable scenario for the twisted pair connection. The results show that the bound of 53.3 kbps can be improved upon.

Note that the block transmission systems we have described can be modelled as MIMO systems where one user communicates with an NSP. As the size of the block goes to ∞, the throughput of the block transmission technique will give the capacity of the channel. In a generalized MIMO system (involving multiple users and the NSP), by adding a cyclic prefix to each user's block, the matrix $\mathbf{H}$ would be block circulant. In this paper we have sometimes used the terminology "MIMO" in place of "block transmission" especially where we want to conserve space.

The problem of obtaining the capacity of a quantized MIMO channel has been preceded by such work as [4], in which Shannon obtained the capacity of an AWGN channel and showed that this capacity is achievable by a Gaussian input distribution. Arimoto [5] and Blahut [6], derived a numerical method for computing the capacity of discrete memoryless channels. In their work, Kavcic [7] and Varnica

et al. [8] presented an equivalent expectation-maximization version of the Blahut-Arimoto algorithm. In [8] furthermore, the Blahut-Arimoto algorithm is modified to incorporate an average power constraint. In [9], Honary et al. investigated the capacity of a scalar, quantized, AWGN channel numerically. Ungerboeck [10] showed numerical results that the performance of a memoryless, quantized, AWGN channel approached the performance of a memoryless, unquantized, AWGN channel, with a certain number of input levels and the work of Ozarow and Wyner [11], provided analytically bounds that support the numerical results of [10]. In [12], Shamai et al. obtained bounds on the average mutual information rates of a discrete-time, peak power limited ISI channel with additive white Gaussian noise. In Varnica et al. [13], Varnica [14] consider Markov sources transmitted over memoryless and ISI channels with an average power constraint and a peak-to-average power ratio constraint. They obtained lower bounds on the capacity of the ISI channel. In [15], Bellorado et al. obtain the capacity of a Rayleigh flat-fading MIMO channel with QAM constellations independent across antennas and dimensions. In our work, we seek to obtain the exact numerical capacity of the quantized MIMO system with average power constraint. This system is obtained by the inclusion of a cyclic prefix to blocks of data symbols in order to supress edge effects. Therefore the capacity of the quantized MIMO system obtained is a lower bound on the capacity of the ISI channel. We compare this capacity to the capacity of the unquantized MIMO system and propose, as a result of our comparisons, conditions under which we can come arbitrarily close to the Shannon bound of (4) at low SNR operating regions.

To achieve our purpose, we use the constrained Blahut-Arimoto algorithm presented in [11], which incorporates the average power constraint P on the channel inputs. However, we replace the interval-halving procedure in [8] by a Newton-Raphson method. We derive this constrained Blahut-Arimoto algorithm in Section 2. In Section 3 we present and discuss results considering the SISO channel. Section 3 provides some useful insights for the block transmission channel, whose results we present in Section 4. We implement a practical example and give the results obtained, in Section 5. Finally we draw our conclusions in Section 6.

In the notations used in this paper, boldface font (e.g., $\mathbf{x}$) is used to denote vectors and matrices (and the corresponding random variables). Calligraphic font (e.g., $\mathcal{X}$) is used to denote the alphabet of the channel inputs or outputs. Summations such as $\sum_{\mathbf{x}}$ refer to summations taken over all the elements in a set under consideration, in this case $\mathbf{x} \in \mathcal{X}$. Unless otherwise stated, natural logarithms are used, thus the unit of capacity is in nats per channel use. We consider the real-valued ISI channel, however the results we obtain apply (mainly with changes in notation) to the complex-valued ISI channel representative of passband systems, where the inputs, outputs, and ISI channel coefficients are complex-valued.

2. ALGORITHM

The channel model for an AWGN baseband ISI channel is

$$y_i = \sum_{l=0}^{L-1} h_l x_{i-l} + z_i, \tag{1}$$

where $\{y_i\}$ and $\{x_i\}$ are, respectively, real-valued channel output and input symbols. $\{h_l\}$, $l = 0, 1, \ldots, L-1$, are real-valued ISI coefficients of the channel of memory $L - 1$ symbols long. $\{z_i\}$ are independent and identically distributed Gaussian noise samples, with zero mean and variance σ_z^2. By adding a cyclic prefix, we can describe the channel of (1) with matrix notation as

$$\mathbf{y} = \mathbf{Hx} + \mathbf{z} \quad z \sim \mathcal{N}(0, \sigma_z^2 \mathbf{I}), \tag{2}$$

where $\mathbf{y} = (y_0, y_1, \ldots, y_{M-1})^T$, $\mathbf{x} = (x_0, x_1, \ldots, x_{M-1})^T$, and $\mathbf{z} = (z_0, z_1, \ldots, z_{M-1})^T$ are vectors whose elements represent, respectively, real channel outputs, real channel inputs, and real noise samples. $\mathbf{H}$ is the channel matrix, whose rows are circular shifts of the ISI coefficients $\{h_i\}$, thanks to the inclusion of the cyclic prefix. For analytical purposes, we can choose an appropriate value for L, and by specifying $M \geq L$ and maintaining the circulant structure of $\mathbf{H}$, we suppress edge effects and ensure that the inputs and outputs of the channel are independent from block to block as it is done in OFDM systems [2]. The vectors $\mathbf{y}$ and $\mathbf{x}$ are such that $\mathbf{y} \in \mathcal{Y}$, $\mathbf{x} \in \mathcal{X}$, where $|\mathcal{Y}| \leq \infty$ and $|\mathcal{X}| \leq \infty$ for a quantized block transmission channel. The inputs of the channel are constrained by an average power constraint P such that

$$E_{\mathbf{x}}[\|\mathbf{x}\|^2] \leq P. \tag{3}$$

The Shannon capacity bound for the channel modelled in (2) (with $\mathbf{x} \in \mathbb{R}$ and $\mathbf{y} \in \mathbb{R}$) is given by [4]

$$C = \lim_{M \to \infty} \frac{1}{M} \sup_{R_{xx}} \frac{1}{2} \log |\mathbf{I} + \sigma_z^{-2} \mathbf{H} \mathbf{R}_{xx} \mathbf{H}^T|, \tag{4}$$

where σ_z^2 is the noise power of the additive white Gaussian noise in the channel. $\mathbf{H}$ is the channel matrix already described and $\mathbf{R}_{xx}$ is the correlation matrix $E[\mathbf{xx}^T]$ of the inputs $\mathbf{x}$. To achieve capacity C, water-filling is done on each of the M subchannels. If we diagonalize $\mathbf{R}_{xx}$ and $\mathbf{H}$ such that $\mathbf{R}_{xx} = \mathbf{F}^H \Lambda_x \mathbf{F}$ and $\mathbf{H} = \mathbf{F}^H \Lambda_H \mathbf{F}$, where $\mathbf{F}^H$ and $\mathbf{F}$ are the IFFT and FFT matrices, respectively, and $\Lambda_x = [\lambda_x]_{ii}$, $\Lambda_H = [\lambda_H]_{ii}$, then

$$\lambda_{x_{ii}} = \left(\left(P + \frac{\sigma_z^2}{\sum_i^M \lambda_{H_{ii}}} \right) - \frac{1}{\sigma_z^{-2} \lambda_{H_{ii}}} \right)^+, \tag{5}$$

where $(\eta)^+$ means that $(\eta)^+ = \eta$ if $\eta \geq 0$ and $(\eta)^+ = 0$ if $\eta < 0$.

If P is the average power constraint on the input signals, such that (3) holds, then we ask the following question: what is the capacity of the quantized block transmission channel described earlier and how does it differ from the capacity
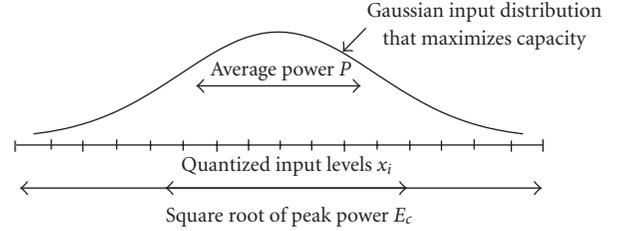
FIGURE 1: Peak power of quantizer is larger than average power imposed.
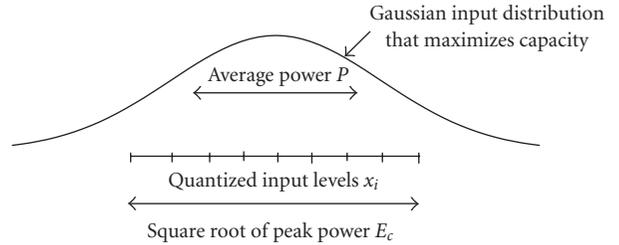
FIGURE 2: Peak power of quantizer is smaller than average power imposed.

given in (4), valid for an unquantized Gaussian channel with average signal power constraint P? We also examine the input distribution that achieves capacity in the quantized block transmission channel.

We consider only the average power constraint P and as a result we use two definitions for SNR in the presentation of our simulation results in Section 3. The first is the nominal SNR (P/σ_z^2) already defined in (3), and the second is the actual SNR ($\sum_{x \in \mathcal{X}'} \|x\|^2 r(x)/\sigma_z^2$), where $\{\mathcal{X}' \subseteq \mathcal{X} : r(x) \neq 0 \text{ for all } x \in \mathcal{X}'\}$. From our simulation results in Section 3 we observe that the actual SNR is equal to the nominal SNR if the peak power $E_c > P$. Interestingly in [14], Varnica proposed an approach that avoids the issue of nominal versus actual SNR given that a subset $\mathcal{X}'$ of the inputs is already chosen.

To examine the effect on capacity of the interaction between E_c and P, we refer to Figures 1 and 2. Figure 1 shows a regime where the peak power of the quantizer E_c is much larger than the average power imposed P. Because we are interested in approaching the capacity of the unquantized MIMO channel, which is achieved by a Gaussian distribution shown as the curve in the figure, to observe an approximate Gaussian optimum input distribution in our quantized system, we deliberately set $E_c > P$. The so-called high-resolution theory [16] covers the case where $E_c \gg P$ and there are fine quantization levels. However, it is unclear how a coarse quantization affects further the performance. Our results (see Section 3, Figure 6) show that whereas at high SNR performance degrades more with precision loss than with saturation loss, at low SNR and in the regime where $E_c > P$, we closely approach the Shannon bound inspite of having a coarse quantization. In Figure 2 instead, $E_c < P$.

Step (1) Compute $p(\mathbf{y} \mid \mathbf{x})$.
Initialize:
(1) Choose any $r(\mathbf{x})$ such that $0 < r(\mathbf{x}) < 1$ and
  $\sum_{\mathbf{x}} r(\mathbf{x}) = 1$.
(2) Initialize capacity C_0, C_{-1}.
Repeat until $C_n - C_{n-1} \leq \varepsilon'$, for some $\varepsilon' \geq 0$
  Step (2) Compute:
  (1) $C_{n-1} = C_n$.
  (2) $q^n(\mathbf{x} \mid \mathbf{y}) = r^{n-1}(\mathbf{x})p(\mathbf{y} \mid \mathbf{x})/\sum_{\mathbf{x}} r^{n-1}(\mathbf{x})p(\mathbf{y} \mid \mathbf{x})$.
  (3) $C_n = \sum_{\mathbf{x}} \sum_{\mathbf{y}} r^{n-1}(\mathbf{x})p(\mathbf{y} \mid \mathbf{x}) \log q^n(\mathbf{x} \mid \mathbf{y})/r^{n-1}(\mathbf{x})$.
  Step (3) Initialize the parameter $\beta : \beta_{0,n}, \beta_{-1,n}$.
  Repeat until $\beta_{i,n} - \beta_{i-1,n} \leq \varepsilon''$, for some $\varepsilon'' \geq 0$
  Step (4) Compute:
  (1) $\beta_{i-1,n} = \beta_{i,n}$.
  (2) $\beta_{i,n} = \beta_{i-1,n} - \sum_{\mathbf{x}} e^{\beta_{i-1,n}\|\mathbf{x}\|^2} [1-\|\mathbf{x}\|^2/P] \prod_{\mathbf{y}} q^n(\mathbf{x}|\mathbf{y})^{p(\mathbf{y}|\mathbf{x})}/$
    $\sum_{\mathbf{x}} \|\mathbf{x}\|^2 e^{\beta_{i-1,n}\|\mathbf{x}\|^2} [1 - \|\mathbf{x}\|^2/P]$
    $\cdot \prod_{\mathbf{y}} q^n(\mathbf{x} \mid \mathbf{y})^{p(\mathbf{y}|\mathbf{x})}$
  end
  Step (5) Compute:
    $r^n(\mathbf{x}) = e^{\beta^n\|\mathbf{x}\|^2} \prod_{\mathbf{y}} q^n(\mathbf{x} \mid \mathbf{y})^{p(\mathbf{y}|\mathbf{x})}/$
    $\sum_{\mathbf{x}'} e^{\beta^n\|\mathbf{x}'\|^2} \prod_{\mathbf{y}} q^n(\mathbf{x}' \mid \mathbf{y})^{p(\mathbf{y}|\mathbf{x}')}$.
end

ALGORITHM 1: The constrained Blahut-Arimoto algorithm.

In this regime the average power constraint is loose and the modified Blahut-Arimoto algorithm utilizes all inputs and assigns input probabilities as if the average power constraint was not in place. This results in a maximizing input distribution which departs from the Gaussian one. Simulation results in Sections 3 and 4 show as expected that the performance degrades compared to the analog channel as we move away from the Gaussian distribution case, because the ratio of peak-to-average power reduces. Thus by increasing the ratio of peak-to-average power, we are still within the constraints of (3), yet we come arbitrarily close to achieving the Shannon capacity bound at low SNR.

2.1. Algorithm for computing the capacity of a block transmission channel with quantized inputs and outputs

Let $r(\mathbf{x})$ denote the input distribution of the channel symbols and let $p(\mathbf{y} \mid \mathbf{x})$ denote the channel transition probability which is a function of SNR, where SNR is defined as P/σ_z^2.[1] $q(\mathbf{x} \mid \mathbf{y})$ denotes the conditional distribution of $\mathbf{x}$ given $\mathbf{y}$. The constrained Blahut-Arimoto algorithm we use for computing the capacity of the quantized block transmission channel is derived below and summarized in Algorithm 1.

---

[1] Since we are given $\mathbf{H}$ and $\mathbf{x}$, the channel transition probability is actually $p(\mathbf{y} \mid \mathbf{Hx})$. Given $\mathbf{Hx}$, $\mathbf{y} \sim \mathcal{N}(\mathbf{Hx}, \sigma_z^2\mathbf{I})$, thus knowing the quantization levels and appropriate decision regions, the complementary error function can be used to compute $p(\mathbf{y} \mid \mathbf{Hx})$ [9].

Derivation of algorithm

Using the earlier defined quantities $r(\mathbf{x})$, $p(\mathbf{y} \mid \mathbf{x})$, and $q(\mathbf{x} \mid \mathbf{y})$, we want to obtain the capacity C of the channel given by

$$C = \max_{r(\mathbf{x})} I(\mathbf{X}; \mathbf{Y})$$
$$= \max_{q(\mathbf{x}|\mathbf{y})} \max_{r(\mathbf{x})} \sum_{\mathbf{x}} \sum_{\mathbf{y}} r(\mathbf{x})p(\mathbf{y} \mid \mathbf{x}) \log \frac{q(\mathbf{x} \mid \mathbf{y})}{r(\mathbf{x})} \quad (6)$$

subject to the constraints

$$\sum_{x} r(x) = 1, \quad (7)$$

$$\sum_{x} \|x\|^2 r(x) \leq P. \quad (8)$$

Start with an initial guess for $r(\mathbf{x})$. The maximizing conditional distribution $q(\mathbf{x} \mid \mathbf{y})$ is given by [5–8, 14, 17]

$$q(\mathbf{x} \mid \mathbf{y}) = \frac{r(\mathbf{x})p(\mathbf{y} \mid \mathbf{x})}{\sum_{\mathbf{x}} r(\mathbf{x})p(\mathbf{y} \mid \mathbf{x})}. \quad (9)$$

Given constraints (7) and (8), we obtain (6) using Lagrange multipliers as a maximization of

$$I(\mathbf{X}; \mathbf{Y}) = \sum_{\mathbf{x}} \sum_{\mathbf{y}} r(\mathbf{x})p(\mathbf{y} \mid \mathbf{x}) \log \frac{q(\mathbf{x} \mid \mathbf{y})}{r(\mathbf{x})}$$
$$+ \lambda \left(\sum_{\mathbf{x}} r(\mathbf{x}) - 1 \right) + \beta \left(\sum_{\mathbf{x}} \|\mathbf{x}\|^2 r(\mathbf{x}) - P \right)$$
$$= \sum_{\mathbf{x}} \sum_{\mathbf{y}} r(\mathbf{x})p(\mathbf{y} \mid \mathbf{x}) \log \frac{q(\mathbf{x} \mid \mathbf{y})}{r(\mathbf{x})}$$
$$+ \lambda \sum_{\mathbf{x}} r(\mathbf{x}) - \lambda + \beta \sum_{\mathbf{x}} \|\mathbf{x}\|^2 r(\mathbf{x}) - \beta P. \quad (10)$$

Maximizing $I(\mathbf{X}; \mathbf{Y})$ with respect to $r(\mathbf{x})$, we obtain

$$\frac{\partial I(\mathbf{X}; \mathbf{Y})}{\partial r(\mathbf{x})} = \sum_{\mathbf{y}} p(\mathbf{y} \mid \mathbf{x}) \log \frac{q(\mathbf{x} \mid \mathbf{y})}{r(\mathbf{x})}$$
$$- \sum_{\mathbf{y}} r(\mathbf{x})p(\mathbf{y} \mid \mathbf{x}) \frac{1}{r(\mathbf{x})} + \lambda + \beta\|\mathbf{x}\|^2 = 0 \quad (11)$$

which implies that

$$\sum_{\mathbf{y}} p(\mathbf{y} \mid \mathbf{x}) \log \frac{q(\mathbf{x} \mid \mathbf{y})}{r(\mathbf{x})} - 1 + \lambda + \beta\|\mathbf{x}\|^2 = 0. \quad (12)$$

Thus

$$e^{1-\lambda}e^{-\beta\|\mathbf{x}\|^2} = e^{\sum_{\mathbf{y}} \log [q(\mathbf{x}|\mathbf{y})/r(\mathbf{x})]^{p(\mathbf{y}|\mathbf{x})}}$$
$$= r(\mathbf{x})^{(-\sum_{\mathbf{y}} p(\mathbf{y}|\mathbf{x}))} \prod_{\mathbf{y}} q(\mathbf{x} \mid \mathbf{y})^{p(\mathbf{y}|\mathbf{x})}, \quad (13)$$

$$r(\mathbf{x}) = \frac{\prod_{\mathbf{y}} q(\mathbf{x} \mid \mathbf{y})^{p(\mathbf{y}|\mathbf{x})}}{e^{1-\lambda}e^{-\beta\|\mathbf{x}\|^2}}. \quad (14)$$

If we substitute for $r(\mathbf{x})$ in (7), we obtain

$$1 = \sum_{\mathbf{x}} \frac{\prod_{\mathbf{y}} q(\mathbf{x} \mid \mathbf{y})^{p(\mathbf{y}|\mathbf{x})}}{e^{1-\lambda} e^{-\beta \|\mathbf{x}\|^2}} \implies e^{1-\lambda} = \sum_{\mathbf{x}} \frac{\prod_{\mathbf{y}} q(\mathbf{x} \mid \mathbf{y})^{p(\mathbf{y}|\mathbf{x})}}{e^{-\beta \|\mathbf{x}\|^2}}. \tag{15}$$

If we substitute for $r(\mathbf{x})$ in (8), we obtain

$$P \geq \sum_{\mathbf{x}} \|\mathbf{x}\|^2 \frac{\prod_{\mathbf{y}} q(\mathbf{x} \mid \mathbf{y})^{p(\mathbf{y}|\mathbf{x})}}{e^{1-\lambda} e^{-\beta \|\mathbf{x}\|^2}} \tag{16}$$

which implies that

$$1 \geq \sum_{\mathbf{x}} \frac{\|\mathbf{x}\|^2}{P} \cdot \frac{\prod_{\mathbf{y}} q(\mathbf{x} \mid \mathbf{y})^{p(\mathbf{y}|\mathbf{x})}}{e^{1-\lambda} e^{-\beta \|\mathbf{x}\|^2}}. \tag{17}$$

Combining (15) and (17) we obtain

$$\sum_{\mathbf{x}} e^{\beta \|\mathbf{x}\|^2} \prod_{\mathbf{y}} q(\mathbf{x} \mid \mathbf{y})^{p(\mathbf{y}|\mathbf{x})} \geq \sum_{\mathbf{x}} \frac{\|\mathbf{x}\|^2}{P} e^{\beta \|\mathbf{x}\|^2} \prod_{\mathbf{y}} q(\mathbf{x} \mid \mathbf{y})^{p(\mathbf{y}|\mathbf{x})}. \tag{18}$$

Thus

$$\sum_{\mathbf{x}} \left[1 - \frac{\|\mathbf{x}\|^2}{P} \right] e^{\beta \|\mathbf{x}\|^2} \prod_{\mathbf{y}} q(\mathbf{x} \mid \mathbf{y})^{p(\mathbf{y}|\mathbf{x})} \geq 0, \tag{19}$$

where (19) is a nonlinear equation in β which we solve numerically using the Newton-Raphson method. This yields an iterative solution for β given by

$$\beta_{n+1} = \beta_n - \frac{\sum_{\mathbf{x}} e^{\beta_n \|\mathbf{x}\|^2} \left[1 - \|\mathbf{x}\|^2/P \right] \prod_{\mathbf{y}} q(\mathbf{x} \mid \mathbf{y})^{p(\mathbf{y}|\mathbf{x})}}{\sum_{\mathbf{x}} \|\mathbf{x}\|^2 e^{\beta_n \|\mathbf{x}\|^2} \left[1 - \|\mathbf{x}\|^2/P \right] \prod_{\mathbf{y}} q(\mathbf{x} \mid \mathbf{y})^{p(\mathbf{y}|\mathbf{x})}}, \tag{20}$$

where n is the index of iteration. In Section 2.4 we will determine a reasonable initial guess for β.[2] With a solution for β, the optimum input distribution is then given as

$$r(\mathbf{x}) = \frac{e^{\beta \|\mathbf{x}\|^2} \prod_{\mathbf{y}} q(\mathbf{x} \mid \mathbf{y})^{p(\mathbf{y}|\mathbf{x})}}{\sum_{\mathbf{x}'} e^{\beta \|\mathbf{x}'\|^2} \prod_{\mathbf{y}} q(\mathbf{x}' \mid \mathbf{y})^{p(\mathbf{y}|\mathbf{x}')}} \tag{21}$$

by combining (14) and (15).

2.2. A specific case

When computing the capacity of a block transmission channel with quantized inputs and real outputs as in a downlink dial-up channel, (9) remains unchanged while (14) becomes

$$r(\mathbf{x}) = \frac{e^{\int \log q(\mathbf{x}|\mathbf{y})^{p(\mathbf{y}|\mathbf{x})} d\mathbf{y}}}{e^{1-\lambda} e^{-\beta \|\mathbf{x}\|^2}}. \tag{22}$$

---

[2] Note that because we use the Newton-Raphson method for the numerical solution of β, the right initial guess of β is crucial to avoid the convergence of β to some unreasonable value that would yield unreasonable results or in some cases, infinite iterations of the algorithm.

This follows by rewriting the first line of (13) as

$$\begin{aligned} e^{1-\lambda} e^{-\beta \|\mathbf{x}\|^2} &= e^{\int \log [q(\mathbf{x}|\mathbf{y})/r(\mathbf{x})]^{p(\mathbf{y}|\mathbf{x})} d\mathbf{y}} \\ &= e^{\int \log q(\mathbf{x}|\mathbf{y})^{p(\mathbf{y}|\mathbf{x})} d\mathbf{y} - \int \log r(\mathbf{x})^{p(\mathbf{y}|\mathbf{x})} d\mathbf{y}} \\ &= \frac{e^{\int \log q(\mathbf{x}|\mathbf{y})^{p(\mathbf{y}|\mathbf{x})} d\mathbf{y}}}{r(\mathbf{x})}. \end{aligned} \tag{23}$$

Simple manipulation of (23) yields (22).

Substituting this value for $r(\mathbf{x})$ in (14) and using similar computations as was done from (15) to (21), we finally obtain $r(\mathbf{x})$ as

$$r(\mathbf{x}) = \frac{e^{\beta \|\mathbf{x}\|^2} e^{\int \log q(\mathbf{x}|\mathbf{y})^{p(\mathbf{y}|\mathbf{x})} d\mathbf{y}}}{\sum_{\mathbf{x}'} e^{\beta \|\mathbf{x}'\|^2} e^{\int \log q(\mathbf{x}|\mathbf{y})^{p(\mathbf{y}|\mathbf{x})} d\mathbf{y}}}. \tag{24}$$

For the purposes of implementation, it is acceptable to quantize the output $\mathbf{Y}$ into bins of length $\Delta \mathbf{Y}$, where $\Delta \mathbf{Y}$ is small compared to the variance of the noise σ_z^2 and $\Delta \mathbf{Y} \leq \Delta \mathbf{X}$. Consider

$$\begin{aligned} \zeta &= \int \log q(\mathbf{x} \mid \mathbf{y})^{p(\mathbf{y}|\mathbf{x})} d\mathbf{y} \\ &= \int p(\mathbf{y} \mid \mathbf{x}) \log \frac{p(\mathbf{y} \mid \mathbf{x}) r(\mathbf{x})}{\sum_{\mathbf{x}} p(\mathbf{y} \mid \mathbf{x}) r(\mathbf{x})} d\mathbf{y} \\ &= \int p(\mathbf{y} \mid \mathbf{x}) \log \frac{p(\mathbf{y} \mid \mathbf{x})}{\sum_{\mathbf{x}} p(\mathbf{y} \mid \mathbf{x}) r(\mathbf{x})} d\mathbf{y} + \log r(\mathbf{x}). \end{aligned} \tag{25}$$

Noting that

$$p(\mathbf{y} \mid \mathbf{x}) = \left[\frac{1}{\sqrt{2\pi\sigma_z^2}} \right]^M e^{-\|\mathbf{y}-\mathbf{Hx}\|^2/2\sigma_z^2} = \xi e^{-\|\mathbf{y}-\mathbf{x}\|^2/2\sigma_z^2}, \tag{26}$$

we see that

$$\begin{aligned} \zeta &= \log \xi + \xi \int e^{-\|\mathbf{y}-\mathbf{Hx}\|^2/2\sigma_z^2} \cdot \left[-\frac{\|\mathbf{y}-\mathbf{Hx}\|^2}{2\sigma_z^2} \right] d\mathbf{y} \\ &\quad - \int p(\mathbf{y} \mid \mathbf{x}) \log \sum_{\mathbf{x}} p(\mathbf{y} \mid \mathbf{x}) r(\mathbf{x}) d\mathbf{y} + \log r(\mathbf{x}) \\ &= \log \xi - \sigma_z^2 + \log r(\mathbf{x}) - \int p(\mathbf{y} \mid \mathbf{x}) \log \sum_{\mathbf{x}} p(\mathbf{y} \mid \mathbf{x}) r(\mathbf{x}) d\mathbf{y}. \end{aligned} \tag{27}$$

Finally we consider $\int p(\mathbf{y} \mid \mathbf{x}) \log \sum_{\mathbf{x}} p(\mathbf{y} \mid \mathbf{x}) r(\mathbf{x}) d\mathbf{y}$. We note that $\sum_{\mathbf{x}} p(\mathbf{y} \mid \mathbf{x}) r(\mathbf{x})$ is a weighted sum of exponentials that are shifted in their mean. This yields a function that is Reimann integrable provided the variance $\mathrm{Var}(\mathbf{Y} \mid \mathbf{X}) = \sigma_z^2$ is finite > 0. If the quantization of $\mathbf{X}$ is fine enough, the smoothness of $\sum_{\mathbf{x}} p(\mathbf{y} \mid \mathbf{x}) r(\mathbf{x}) d\mathbf{y}$ increases especially at low SNR. Therefore we can approximate the continuous output $\mathbf{Y}$ by a quantized output as long as the number of quantization levels is greater than or equal to the number of quantization levels of the input.

2.3. Convergence

References [8, 18, 19] have shown that the Blahut-Arimoto algorithm converges with a speed that is at least inversely proportional to the approximation error (ε'). The Newton-Raphson method used to obtain β in (20) has also been shown to converge to a local root, given that the initial value is sufficiently close to the desired root [20]. While we offer no formal proof that our algorithm converges, empirical results from our simulations show that it does. In the following subsection, we define an appropriate initial guess for β. The convergence speed of the Newton-Raphson method is quadratic in ε''.

2.4. Analysis

In this section we determine a reasonable initial guess for β. In our bid to design systems with information rates arbitrarily close to the Shannon capacity bound shown in (4), it is useful for us to also examine how the capacity of the channel is affected by the average power constraint P and by the number of levels of the quantizer.

Result

Given the parameters λ and β from the constrained optimization of $I(\mathbf{X};\mathbf{Y})$, where λ and β are both functions of $p(\mathbf{y}\mid\mathbf{x})$, the following result holds:

$$1 - \lambda - \beta P = I(\mathbf{X};\mathbf{Y}). \qquad (28)$$

Proof. We start by rewriting (14) as

$$\log r(\mathbf{x}) = \log \prod_{\mathbf{Y}} q(\mathbf{x}\mid\mathbf{y})^{p(\mathbf{y}\mid\mathbf{x})} - \log e^{1-\lambda-\beta\|\mathbf{x}\|^2}$$

$$= \lambda - 1 + \beta\|\mathbf{x}\|^2 + \sum_{\mathbf{Y}} p(\mathbf{y}\mid\mathbf{x}) \log \frac{p(\mathbf{y}\mid\mathbf{x})r(\mathbf{x})}{p(\mathbf{y})}$$

$$= \lambda - 1 + \beta\|\mathbf{x}\|^2 + \sum_{\mathbf{Y}} p(\mathbf{y}\mid\mathbf{x}) \log p(\mathbf{y}\mid\mathbf{x})$$

$$+ \log r(\mathbf{x}) \sum_{\mathbf{Y}} p(\mathbf{y}\mid\mathbf{x}) - \sum_{\mathbf{Y}} p(\mathbf{y}\mid\mathbf{x}) \log p(\mathbf{y})$$

$$\qquad (29)$$

which means that

$$1 - \lambda - \beta\|\mathbf{x}\|^2 = \sum_{\mathbf{Y}} p(\mathbf{y}\mid\mathbf{x}) \log p(\mathbf{y}\mid\mathbf{x})$$

$$- \sum_{\mathbf{Y}} p(\mathbf{y}\mid\mathbf{x}) \log p(\mathbf{y}). \qquad (30)$$

If we multiply both sides of (30) by $r(\mathbf{x})$ and sum over all $\mathbf{x}$, we obtain (28). $\square$

As $P \to \infty$, $I(\mathbf{X};\mathbf{Y})$ tends to the capacity achieved by the unconstrained Blahut-Arimoto algorithm which can be at most $\log_2 |\mathcal{X}|$ bits/use of the channel. Since β is introduced in the maximization of (10) through the power constraint P,
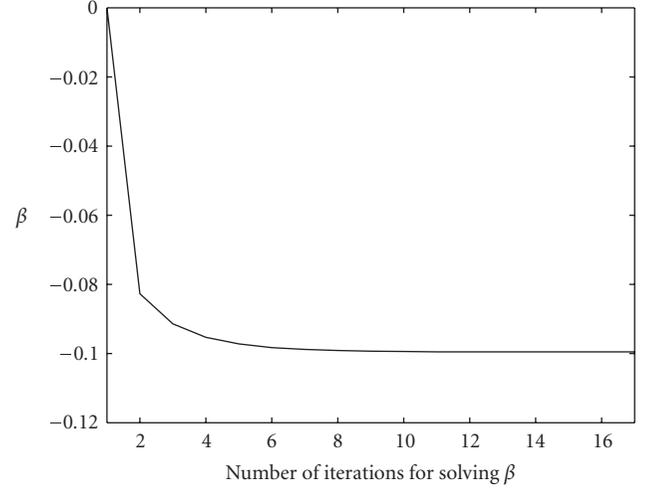
FIGURE 3: Typical convergence pattern of β.

this implies that as P increases, β decreases. In other words, the effect of β is greater as the power constraint becomes stricter. Because of this, a reasonable initial guess of β_0 for (20) is 0. For $\beta_0 = 0$, (20) becomes

$$\beta_1 = 0 - \frac{\sum_{\mathbf{x}}\left[1 - \|\mathbf{x}\|^2/P\right]\prod_{\mathbf{Y}} q(\mathbf{x}\mid\mathbf{y})^{p(\mathbf{y}\mid\mathbf{x})}}{\sum_{\mathbf{x}}\|\mathbf{x}\|^2\left[1 - \|\mathbf{x}\|^2/P\right]\prod_{\mathbf{Y}} q(\mathbf{x}\mid\mathbf{y})^{p(\mathbf{y}\mid\mathbf{x})}}. \qquad (31)$$

Note that we can conclude from (31) that β is negative as $P \to \infty$ because at some stage, $\|\mathbf{x}\|^2 \leq P$ for all $\mathbf{x} \in \mathcal{X}$. Indeed β is negative for all values of P because $r(\mathbf{x})$ and β satisfy the first order Kuhn-Tucker conditions for obtaining the unique, global optimum of a concave function subject to concave constraints. These conditions are

$$\frac{\partial I(\mathbf{X};\mathbf{Y})}{\partial r(\mathbf{x})} = 0; \quad r(\mathbf{x}) \geq 0,$$

$$\frac{\partial I(\mathbf{X};\mathbf{Y})}{\partial \beta} \leq 0; \quad \beta \leq 0, \qquad (32)$$

from which we see that β is negative. Figure 3 is a plot showing the typical convergence pattern of β from our simulation results. Values of β are seen to be negative. Figure 4 shows that β decreases as SNR increases.

At the low SNR region, the maximizing input distribution tends towards a Gaussian distribution. We can see this by utilizing high-resolution analysis [16]

$$I(\mathbf{X}^{\Delta};\mathbf{Y}^{\Delta}) \approx I(\mathbf{X};\mathbf{Y}), \quad \text{as } \Delta \longrightarrow 0, \qquad (33)$$

where Δ is the quantization step, $I(\mathbf{X}^{\Delta};\mathbf{Y}^{\Delta})$ is the mutual information between the quantized versions of the random variables $\mathbf{X}$ and $\mathbf{Y}$, and $I(\mathbf{X};\mathbf{Y})$ is the information rate of the unquantized block transmission channel which is maximized by a Gaussian input distribution. It is permissible to use high-resolution analysis for the low SNR region because in this region we are constrained from peaks that are so far from the mean of the distribution that in essence $E_c \to \infty$.

Figure 4: Variation of β with SNR.

Figure 5: Comparison of capacity versus nominal SNR for different SISO channels.
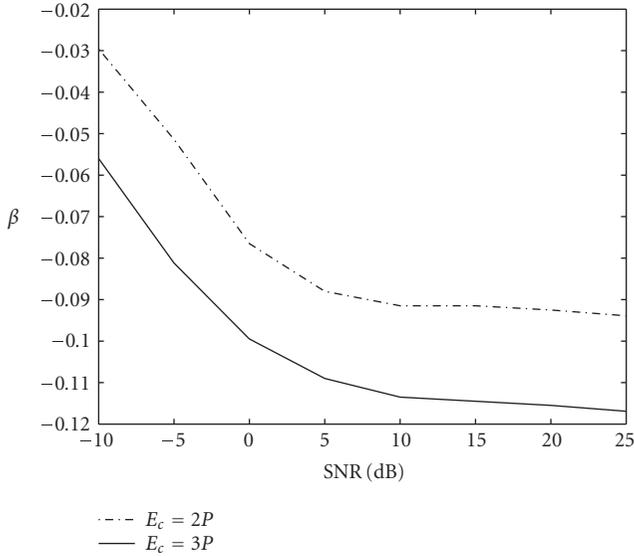
Also $\Delta \to 0$ compared to the standard deviation of the noise because at low SNR, the standard deviation of the noise is large and the input is independent of the noise. The combined conditions of $\Delta \to 0$ and $E_c \to \infty$ yield the high-resolution scenario.

Figures 7 and 10 are simulation results which show that the maximizing input distribution at low SNR is concave like the Gaussian distribution. We conclude that at low SNR, the capacity of the quantized block transmission channel will be closer to the Shannon bound.

As to the effect on information rate, of number of quantization levels of the quantizer, high-resolution analysis [16, 21] implies that as $\Delta \to 0$ (i.e., as number of quantization levels increases), approximately optimal performance on a Gaussian channel can be obtained.

In our case however, because SNR presents a constraint, the high-resolution scenario cannot be realized. Another way of looking at this is stating that we cannot arbitrarily increase the number of quantization levels and still utilize all the available input signals because at some stage, satisfactory error performance can no longer be achieved (by uncoded modulation) [10]. Hence the goal of our analysis was to find a way to approximate the high-resolution scenario given the dual constraints of a permissible SNR and a specific channel alphabet.

From our analysis we propose that given average power P, peak power E_c, and the quantization levels, the performance of the quantized block transmission channel approaches that of the analog channel when we constrain the quantized block transmission channel to approximate the analog channel, by increasing peak-to-average power ratio.

In the next sections we present simulation results which support our claims, and thereby present important underlying principles behind designing a quantized block transmission system that achieves capacity close to that of an unquantized block transmission system in the low SNR region.
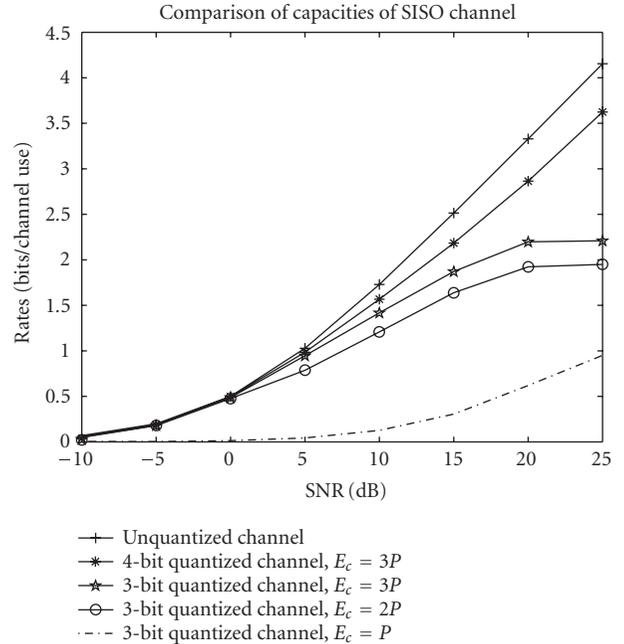
3. RESULTS: CAPACITY OF SISO CHANNEL

We run simulations in which we arbitrarily generate the channel impulse response h_l in (1) and then run Algorithm 1. Our simulations show the comparison of the capacities of the following channels: 4 bit (16-level) quantized SISO channel with $E_c = 3P$, 3 bit quantized SISO channel with $E_c = 3P$, 3 bit quantized SISO channel with $E_c = 2P$, 3 bit quantized SISO channel with $E_c = P$, and the unquantized SISO channel. Figure 5 shows the capacity curves for these channels plotted against nominal SNR (P/σ_z^2). Figure 5 shows the same curves plotted against actual average SNR achieved ($\sum_{\mathbf{x} \in \mathcal{X}'} \|\mathbf{x}\|^2 r(\mathbf{x})/\sigma_z^2$), from which we see that the actual SNR achieved is affected by the value of E_c used and the nominal SNR is more likely to be achieved when $E_c > P$.

From Figure 5 we see that as the ratio of peak-to-average power increases, we approach more closely the capacity achieved by the unquantized channel. Also the performance improves with increased number of input levels. High-resolution analysis does not provide information on how reduced peak-to-average ratio (saturation loss) and reduced number of quantization levels (precision loss) affect the performance of the channel, relative to each other. From Figure 6, we see that performance degrades more with precision loss than with saturation loss. This is expected because for any given value of E_c, you can do better by increasing the number of quantization levels whereas if you fix the number of quantization levels and increase E_c, you cannot do better beyond a certain value due to the fact that the number of inputs are only so many. Note also that the precision gain depends on the SNR and is much lower at low SNR. We can see in addition that at low SNR the information rate is almost
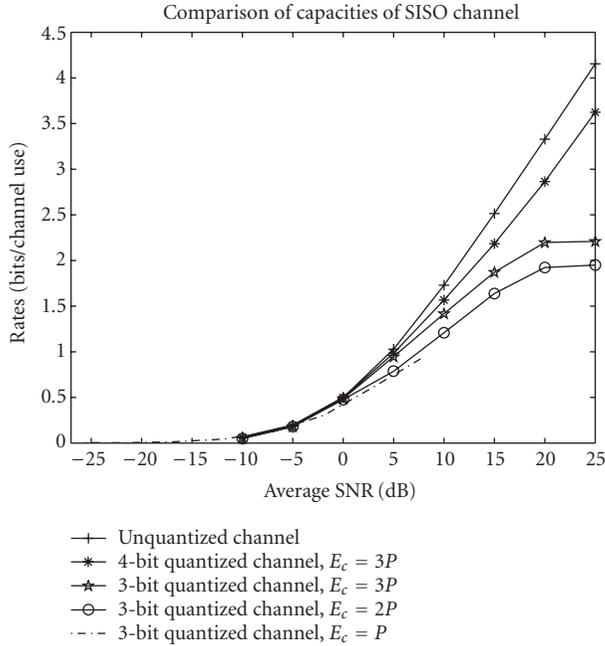
FIGURE 6: Comparison of capacity versus actual SNR for different SISO channels.
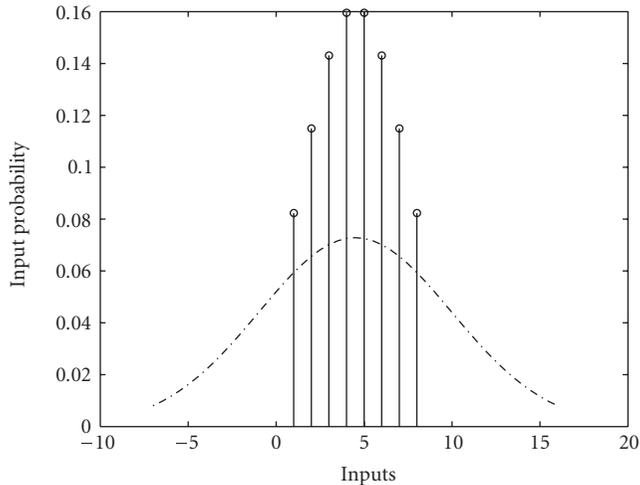
FIGURE 7: Maximizing input distribution at low SNR for the SISO channel.

insensitive to precision loss and is dominated instead by saturation loss (the slope of the curve is higher for a higher peak-to-average ratio). Note that the slope of the curves is a function of the number of quantization levels (it is the same for equal number of quantization levels). The results obtained from our simulations therefore support our earlier discussions.

Figure 7 shows a typical maximizing input distribution at low SNR and Figure 8 shows a maximizing input distribution at high SNR. These figures show that the maximizing
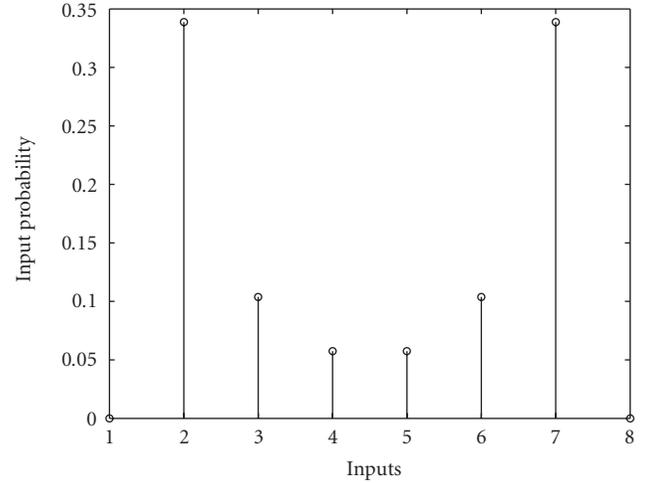
FIGURE 8: A maximizing input distribution at high SNR for the SISO channel.

TABLE 1: Accuracy of predicted value of capacity as SNR increases.

SNR	C	C'
-10 dB	0.0621	0.0410
-5 dB	0.1838	-0.0827

input distribution tends further away from Gaussian as P increases as was discussed in Section 2. If we approximate the maximizing distribution at low SNR by quantizing an appropriate Gaussian distribution (as shown in Figure 7), then we can predict the capacity at this SNR and our predictions will be fair.

This is shown by Table 1 where we see that the predicted capacity C' is close to the actual capacity C at low SNR but the prediction is poorer as SNR increases because the maximizing input distribution is far from Gaussian. In the next section, we present results of similar simulations for the block transmission (MIMO) channel.

4. RESULTS: CAPACITY OF MIMO CHANNEL

Similar to our simulations for the SISO channel, we run several simulations for the MIMO channel and for each simulation, we randomly generate the circulant matrix $\mathbf{H}$, then implement Algorithm 1. Again we take the average capacity obtained from our simulations. For reasons of computational complexity we consider channel length $L = 2$ and symbol block length $M = 3$. Channel memory order 1 is reasonable for our example application, the dial-up channel, because the frequency response of the analog twisted copper pair is almost flat over the 4 KHz bandwidth used for transmission.

We show the results of comparing the capacities of the following channels: 4-level, 3-dimensional ($M = 3$), quantized MIMO channel with $E_c = 3P$, 4-level, 3-dimensional quantized MIMO channel with $E_c = 2P$, 4-level,
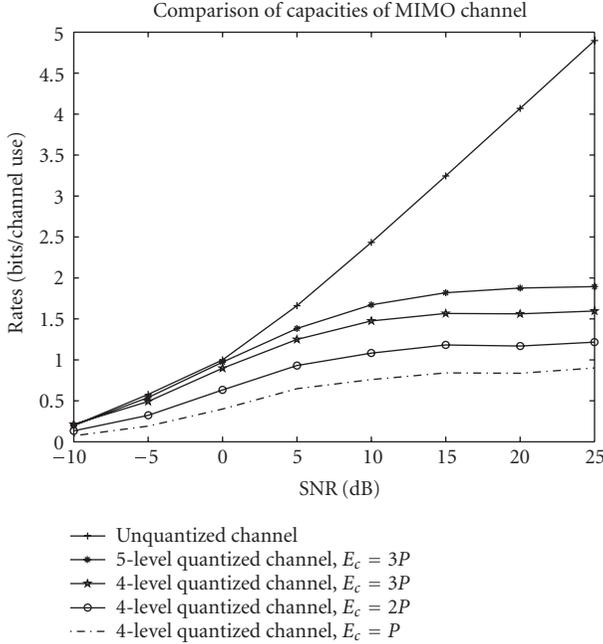
Figure 9: Comparison of capacity versus nominal SNR for different MIMO channels.
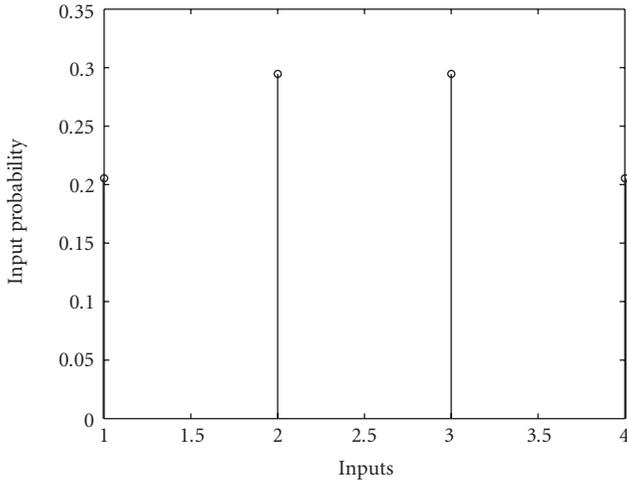
Figure 10: A maximizing input distribution at low SNR for the MIMO channel.

3-dimensional quantized MIMO channel with $E_c = P$, 5-level, 3-dimensional quantized MIMO channel with $E_c = 3P$, and the unquantized MIMO channel. Figure 9 shows the capacity curves for these channels plotted against nominal SNR (P/σ_z^2). We observe again that as with the SISO case, performance improves with increased peak-to-average ratio and increased number of quantization levels. The marginal distribution of a typical maximizing input distribution for the quantized MIMO channel at low SNR is shown in Figure 10. As we would expect, the marginal distribution is concave like the Gaussian distribution.
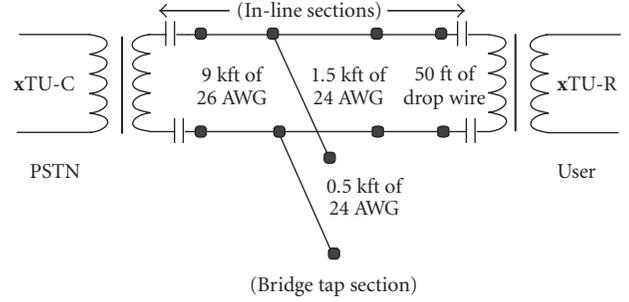
Figure 11: A typical end-to-end loop.
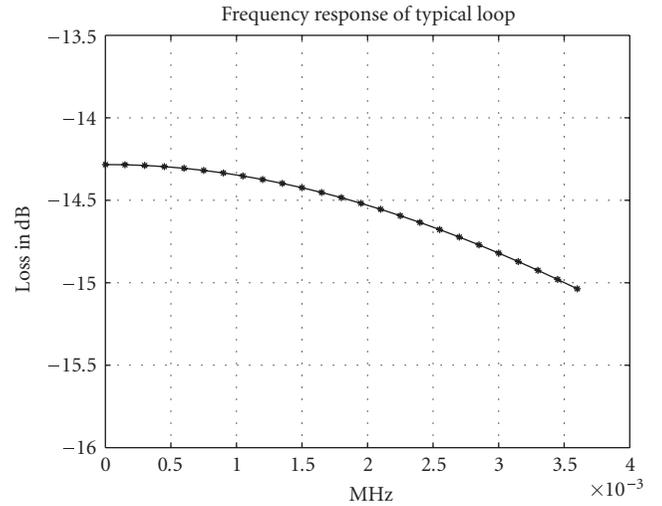
Figure 12: Frequency response of a typical end-to-end loop.

5. A PRACTICAL EXAMPLE

It is interesting to apply the capacity bounds developed in this paper to a practical example which is the downlink channel of a dial-up connection, where the inputs are quantized and the outputs are real. In this section we will simulate practical line conditions for a typical downlink dial-up channel [22]. The end-to-end loop we analyze is shown in Figure 11. The transfer function of the loop is given in [22] and can be calculated using published tables which are also provided in [22]. The bandwidth of interest is 3600 Hz between 150 Hz and 3750 Hz. This is the bandwidth that allows optimum performance. The frequency response obtained is shown in Figure 12. Figure 13 is the impulse response of the end-to-end loop. In our simulations, we assume the input $\mathbf{X}$ is uniformly quantized into 128 levels. In practice, the input constellation is picked to approach a uniform quantization as closely as possible [23]. The channel is sampled at a rate of 8000 Hz. We choose as average power P, the FCC-imposed average transmit power $P = -12$ dBm, we set $E_c = 3P$ and, we vary the noise power σ_z^2. An SNR value of around 55 dB is expected under normal operating conditions [23]. We use block length $M = 3$ and the length of the channel impulse response $L = 2$. The result we obtain is shown in Figure 14.

Impulse response of a typical loop
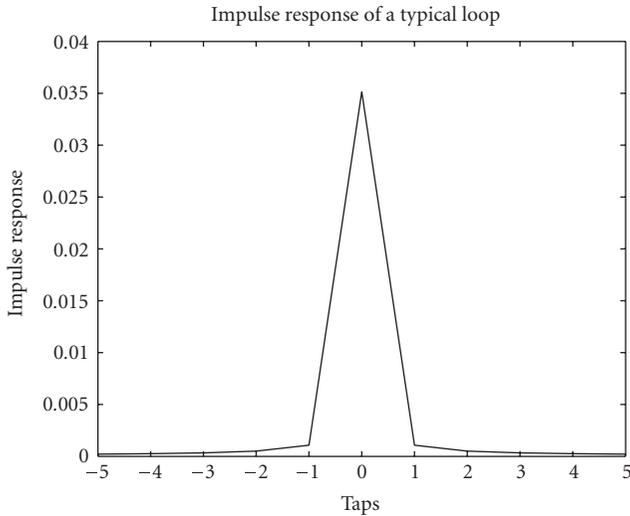
FIGURE 13: Impulse response of a typical end-to-end loop.

Capacity of V.90 downstream
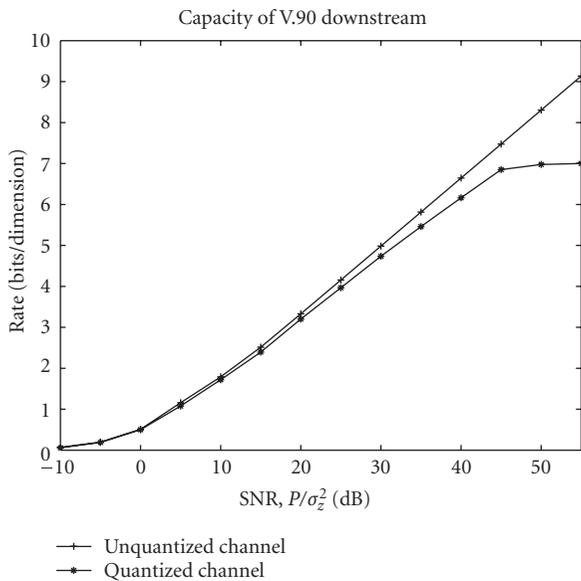
FIGURE 14: Capacity of the downstream link.

From Figure 14 we can see that at low SNR, the capacity of the quantized channel approaches the capacity of the unquantized channel very closely and at an SNR of about 45dB, the Nyquist rate of the channel 56 kbps (= 7 bits/dimension × 8000 dimensions/s) is achieved under the prevailing line conditions. This shows that the limit of 53.3 kbps can be improved upon.

6. CONCLUSION

In this paper we have in general, proposed useful guidelines for the design of block transmission systems whose

performance at low SNR is arbitrarily close to the Shannon bound. Specifically, we have tested our proposals by applying them to the downlink channel of a dial-up system and found that we improved upon the limit of 53.3 kpbs.

REFERENCES

[1] E. Ayanoglu, N. R. Dagdeviren, G. D. Golden, and J. E. Mazo, "An equalizer design technique for the PCM modem: a new modem for the digital public switched network," *IEEE Transactions on Communications*, vol. 46, no. 6, pp. 763–774, 1998.

[2] D. J. Rauschmayer, *ADSL/VDSL Principles : A Practical and Precise Study of Asymmetric Digital Subscriber Lines and Very High Speed Digital Subscriber Lines*, Macmillan, New York, NY, USA, 1999.

[3] D. S. Lawyer, "Modem-HOWTO," May 2003, http://www.tldp.org/HOWTO/Modem-HOWTO-1.html.

[4] C. E. Shannon, "A mathematical theory of communications," *Bell Systems Technical Journal*, vol. 27, pp. 379–423 (pt I), 623–656 (pt II), 1948.

[5] S. Arimoto, "An algorithm for computing the capacity of arbitrary discrete memoryless channels," *IEEE Transactions on Information Theory*, vol. 18, no. 1, pp. 14–20, 1972.

[6] R. E. Blahut, "Computation of channel capacity and rate-distortion functions," *IEEE Transactions on Information Theory*, vol. 18, no. 4, pp. 460–473, 1972.

[7] A. Kavcic, "On the capacity of Markov sources over noisy channels," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '01)*, vol. 5, pp. 2997–3001, San Antonio, Tex, USA, November 2001.

[8] N. Varnica, X. Ma, and A. Kavcic, "Capacity of power constrained memoryless AWGN channels with fixed input constellations," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '02)*, vol. 2, pp. 1339–1343, Taipei, Taiwan, November 2002.

[9] B. Honary, F. Ali, and M. Darnell, "Information capacity of additive white Gaussian noise channel with practical constraints," *IEE Proceedings, Part I: Communications, Speech and Vision*, vol. 137, no. 5, pp. 295–301, October 1990.

[10] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Transactions on Information Theory*, vol. 28, no. 1, pp. 55–67, 1981.

[11] L. H. Ozarow and A. D. Wyner, "On the capacity of the Gaussian channel with a finite number of input levels," *IEEE Transactions on Information Theory*, vol. 36, no. 6, pp. 1426–1428, 1990.

[12] S. Shamai (Shitz), L. H. Ozarow, and A. D. Wyner, "Information rates for a discrete-time Gaussian channel with intersymbol interference and stationary inputs," *IEEE Transactions on Information Theory*, vol. 37, no. 6, pp. 1527–1539, 1991.

[13] N. Varnica, X. Ma, and A. Kavcic, "Power-constrained memoryless and intersymbol interference channels with finite input alphabets: capacities and concatenated code constructions," to appear in *IEEE Transactions on Communications*, http://hrl.harvard.edu/~varnica/publications.htm.

[14] N. Varnica, "Iteratively decodable codes for memoryless and intersymbol interference channels," Ph.D. dissertation, Harvard University, Cambridge, Mass, USA, 2005, http://hrl.harvard.edu/~varnica/pic/PhDThesis.pdf.

[15] J. Bellorado, S. Ghassemzadeh, and A. Kavcic, "Approaching the capacity of the MIMO Rayleigh flat-fading channel with QAM constellations, independent across antennas and dimensions," to appear in *IEEE Transactions on Wireless Communications*, http://people.deas.harvard.edu/~kavcic/recent.html.

[16] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, NY, USA, 1991.

[17] R. W. Yeung, *A First Course in Information Theory*, Kluwer Academic/ Plenum, New York, NY, USA, 2002.

[18] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Academic Press, London, UK, 1981.

[19] P. O. Vontobel, "A generalized Blahut-Arimoto algorithm," in *Proceedings of the IEEE International Symposium on Information Theory*, p. 53, Yokohama, Japan, July 2003.

[20] R. W. Hamming, *Numerical Methods for Scientists and Engineers*, Dover, New York, NY, USA, 2nd edition, 1987.

[21] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2325–2383, 1998.

[22] J. A. C. Bingham, *ADSL, VDSL, and Multicarrier Modulation*, Wiley Series in Telecommunications and Signal Processing, John Wiley & Sons, New York, NY, USA, 2000.

[23] L. M. Caraballo, "System level design and simulation of a PCM voiceband modem compliant with the ITU V.90 standard," M.S. thesis, Texas A&M University, Canyon, Tex, USA, 2000.

Obianuju Ndili received the B.S. degree in electrical engineering from the University of Lagos, Lagos, Nigeria, in 1998, and the M.S. degree in electrical engineering from Cornell University, Ithaca, New York, in 2003. She is currently working towards the Ph.D. degree in electrical engineering at Santa Clara University, Santa Clara, California. Her research interests include applied signal processing, information theory, and video compression.

Tokunbo Ogunfunmi received the B.S. (first class honors) degree in electrical and electronic engineering from the University of Ife, Ile-Ife, Nigeria, in 1980, the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, California, in 1984 and 1990, respectively. He is currently an Associate Professor of electrical engineering and the Director of the Signal Processing Research Lab. at Santa Clara University, Santa Clara, California. His current research interests include digital and adaptive signal processing, nonlinear signal processing, and speech and multimedia (audio, video) compression.