

Approximation of Capacity for ISI Channels with One-bit Output Quantization

Radha Krishna Ganti, Andrew Thangaraj and Arijit Mondal

Department of Electrical Engineering

Indian Institute of Technology, Madras

Chennai, India 600036

{rganti, andrew, arijit.mondal}@ee.iitm.ac.in

Abstract—Motivated by recent high bandwidth communication systems, Inter-Symbol Interference (ISI) channels with 1-bit quantized output are considered under an average-power-constrained continuous input. While the exact capacity is difficult to characterize, an approximation that matches with the exact channel output up to a probability of error is provided. The approximation does not have additive noise, but constrains the channel output (without noise) to be above a threshold in absolute value. The capacity under the approximation is computed using methods involving standard Gibbs distributions. Markovian achievable schemes approaching the approximate capacity are provided. The methods used over the approximate ISI channel result in ideas for practical coding schemes for ISI channels with 1-bit output quantization.

I. INTRODUCTION

Channels with Inter-Symbol Interference (ISI) and Additive White Gaussian Noise (AWGN) are often encountered in practice. Depending on the application, an average-input power constraint or a finite input alphabet constraint is commonly studied. Recently, in applications such as millimeter wave [1][2] or optical or intra-chip [3] communications, quantization of the output of an ISI channel has been considered because of limitations in Analog-to-Digital conversion at high speeds. In some cases, the output quantization may be as low as a single bit. Since the transmitters in some of these systems can be more complex and operate at high powers, the channel input may not have severe quantization limits.

Motivated by the above applications, we consider a noisy ISI channel with average-power constrained continuous input and 1-bit quantized output. The available literature mostly considers either continuous input/output alphabets or a finite input alphabet with a continuous output alphabet [4][5]. The quantized output case has been considered for the AWGN channel with no ISI [6], and the ISI case has been briefly addressed recently [7].

The exact capacity of an ISI channel with 1-bit quantized output appears to be difficult to characterize explicitly. In this work, we introduce an approximation to the ISI channel model with 1-bit quantized output. The approximation does not have additive noise, but constrains the channel output (without noise) to be above a threshold in absolute value. Because of the thresholding, the approximate channel output, after quantization, matches the actual channel output up to a probability of error that can be controlled by the threshold.

The main advantage of avoiding noise in the approximation is that the exact capacity can be computed for the approximate model. We show how such a computation can be carried out using Gibbs distributions. In addition, we exhibit achievable schemes with Markov input approaching the approximate capacity in some numerical examples. Since the approximation is valid up to a probability of error, a coding scheme used over the approximate channel model can be coupled with a standard error control code to derive a practical coding scheme for the exact channel.

The rest of the paper is organized as follows. Section II describes the ISI channel model and its approximation. In Section III, we provide a method for computing the approximate capacity under an average power constraint on the input, and elaborate on achievable schemes in Section IV. Numerical results are given in Section V, and concluding remarks are made in Section VI.

II. SYSTEM MODEL

We consider a discrete-time finite-tap ISI channel with average-power-constrained continuous input and one-bit quantized output as depicted in Fig. 1. The input to the channel is

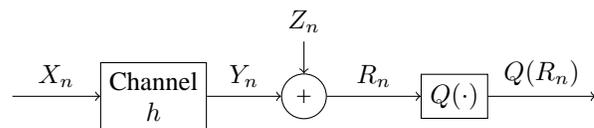


Fig. 1: ISI channel with quantized output.

denoted $X = \{X_n, 0 \leq n \leq N-1\}$, and the channel impulse response of length L is denoted $h = \{h_n, 0 \leq n \leq L-1\}$. The convolution of the input with the channel impulse response is denoted $Y = \{Y_n, 0 \leq n \leq N-1\}$, and is given by

$$Y_n = \sum_{k=0}^{L-1} h_k X_{n-k}. \quad (1)$$

The channel h is assumed to be constant, and all signals are assumed to be zero outside the specified ranges. Independent and identically distributed zero-mean Gaussian noise of variance σ^2 , denoted Z_n , is added to obtain an intermediate signal $R_n = Y_n + Z_n$. The signal R_n is quantized by a 1-bit quantizer

$Q(\cdot)$ to obtain the channel output $Q(R) = \{Q(R_n), 0 \leq n \leq N + L - 2\}$. The quantizer is defined as follows:

$$Q(x) = \begin{cases} +1, & \text{if } x \geq 0, \\ -1, & \text{if } x < 0. \end{cases} \quad (2)$$

The average power of the input is constrained to be at most P . That is, we require

$$\mathbb{E}[\|X\|^2] = \sum_{n=0}^{N-1} E[|X_n|^2] \leq NP. \quad (3)$$

In this work, the overall goal is to approximate the mutual information rate $\frac{1}{N}I(X; Q(R))$ and provide computable expressions or bounds.

The approximate ISI channel model is depicted in Fig. 2. In the approximate ISI channel, there is no noise. However,

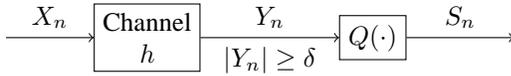


Fig. 2: Approximate ISI channel with quantized output.

the convolution output Y_n is constrained to be greater than δ in absolute value, and this provides justification for ignoring noise. We readily see that, under the constraint $|Y_n| \geq \delta$, the output of the actual model $Q(R_n)$ is approximated by $S_n = Q(Y_n)$ up to a probability of error lesser than or equal to $\mathcal{Q}(\delta/\sigma)$, where $\mathcal{Q}(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$ is the standard Q -function. Therefore, coding schemes developed for the approximate model can be used in the actual model with additional error control coding for the approximation error $\mathcal{Q}(\delta/\sigma)$. While the additional coding results in a loss in information rate, we find that the approximation, because of the removal of noise, is useful on the following two counts:

- 1) The capacity or maximum information rate $\frac{1}{N}I(X; S)$, where $S = \{S_n, 0 \leq n \leq N - 1\}$, under the power constraint of (3) and the constraint $|Y_n| \geq \delta$, has computable expressions and bounds. This provides useful estimates of the capacity of output-quantized ISI channels.
- 2) The techniques used for computing the approximate capacity provide useful intuition on coding and signaling methods for output-quantized ISI channels.

Convolution: We will use matrix notation to denote the convolution in (1) as $Y = \tilde{M}_h X$, where the entries of the $N \times N$ matrix \tilde{M}_h are either 0 or one of the channel taps h_k and X, Y are column vectors. The matrix \tilde{M}_h is not circulant. Let M_h denote the $N \times N$ circulant matrix with first column equal to h . When N becomes large (as is in our case), with L fixed, \tilde{M}_h behaves like the circulant matrix M_h in the sense that $\lim_{N \rightarrow \infty} \|M_h - \tilde{M}_h\| = 0$ where $\|\cdot\|$ is a matrix norm. In this paper¹, for simplicity, we will always assume the circular convolution $Y = M_h X$ in the channel model.

¹For all N , the results hold with a suitable cyclic prefix, for instance.

III. APPROXIMATE CAPACITY

The capacity of the approximate ISI channel is given by

$$C_\delta(P) = \lim_{N \rightarrow \infty} \sup_{\substack{\mathbb{E}[\|X\|^2] \leq NP \\ |Y_n| \geq \delta}} \frac{I(X, S)}{N} = \frac{H(S)}{N},$$

where the last equality is because the sequence S is a deterministic function of the input X in the absence of noise in the approximate channel. Since the output alphabet of the quantizer is either 1 or -1 , it is easy to observe that $C_\delta(P) \leq 1$.

A. Power constraint

We begin by bounding the power of the input sequence X required for a given output sequence S . Given the output symbol sequence $S = s \in \{-1, 1\}^N$, we have the constraint that $|Y_n| = s_n Y_n \geq \delta$. So, the minimum energy, denoted $\mathcal{E}(s)$, required for a given quantized sequence s is given by the following optimization problem:

$$\mathcal{E}(s) = \min_{\text{diag}(s)M_h x \succeq \delta \mathbf{1}} \sum_{n=0}^{N-1} |x_n|^2, \quad (4)$$

where $\text{diag}(s)$ is an $N \times N$ diagonal matrix with s on the main diagonal and $\mathbf{1}$ denotes the all-1 column. The inequalities $\text{diag}(s)M_h x \succeq \delta \mathbf{1}$ are linear and the feasible space for x is the intersection of hyperplanes and, hence, convex. So, the above optimization problem is a convex optimization problem in N variables that essentially finds the closest point from the origin to the convex set $\{\text{diag}(s)M_h x \succeq \delta \mathbf{1}\}$. We have

$$\begin{aligned} \mathbb{E}[\|X\|^2] &= \sum_{s \in \{-1, 1\}^N} \mathbb{P}(S = s) \mathbb{E}[\|X\|^2 | S = s], \\ &\geq \sum_{s \in \{-1, 1\}^N} \mathbb{P}(S = s) \mathcal{E}(s) \end{aligned}$$

using (4). Because of the average power constraint on the input X , we have

$$\sum_{s \in \{-1, 1\}^N} \mathbb{P}(S = s) \mathcal{E}(s) \leq NP. \quad (5)$$

B. Entropy maximization and Gibbs distribution

Let $\mathcal{E}_{\min} = \min_s \mathcal{E}(s)$, $\mathcal{E}_{\max} = \max_s \mathcal{E}(s)$, $\bar{\mathcal{E}} = \frac{1}{2^N} \sum_s \mathcal{E}(s)$. Since the constraints (5) are linear on the probabilities, it is well-known [8] that the Gibbs distribution maximizes the entropy $H(S)$ for $\mathcal{E}_{\min} \leq NP \leq \mathcal{E}_{\max}$. The optimal distribution is the Gibbs distribution given by

$$\mathbb{P}(S = s) = \frac{e^{-\frac{\beta \mathcal{E}(s)}{N}}}{Z}, \quad s \in \{-1, 1\}^N, \quad (6)$$

where Z is the normalizing constant, and β is the unique value for which (5) is met with equality. The maximum entropy is given by

$$H(S) = \beta P + \ln(Z) \quad (7)$$

$$= \beta P - \ln \left(\sum_{s \in \{-1, 1\}^N} e^{-\frac{\beta \mathcal{E}(s)}{N}} \right). \quad (8)$$

For $NP < \mathcal{E}_{\min}$, there exists no probability distribution that satisfies (5). It is also known [8][9] that $\beta = 0$ when $NP = \bar{\mathcal{E}}$ and we observe that the corresponding Gibbs distribution is the uniform distribution on $\{-1, 1\}^N$ and the maximum $H(S) = N$. For $P = \mathcal{E}_{\min}$, we have $\beta = \infty$ and the maximum entropy is given by $\log_2(|\mathcal{S}_m|)$, where $|\mathcal{S}_m|$ denotes the number of sequences in $\{-1, 1\}^N$ that achieve the minimum energy. Hence, the maximum entropy when $P = \mathcal{E}_{\min}$ is $H(S) = \frac{\log_2(|\mathcal{S}_m|)}{N}$.

In summary, we see that the capacity of the approximate ISI channel is given by the Gibbs distribution whenever the power constraint is above \mathcal{E}_{\min}/N . For $P \geq \bar{\mathcal{E}}/N$, we can achieve the maximum possible capacity $C_\delta(P) = 1$. For powers lower than \mathcal{E}_{\min}/N , capacity goes to zero. So, the interesting range of calculation is for $NP \in (\mathcal{E}_{\min}, \bar{\mathcal{E}})$.

C. Diagonally-dominant channels

A matrix $A = (a_{ij})$ is said to be row-diagonally-dominant or simply diagonally-dominant if $|a_{ii}| \geq \sum_{j, j \neq i} |a_{ij}|$. Let us call channels h for which the matrix $(M_h M_h^T)^{-1}$ exists and is diagonally-dominant as diagonally-dominant channels. For such channels, the minimum energy values $\mathcal{E}(s)$ can be characterized as follows.

Lemma 1. *When the matrix $(M_h M_h^T)^{-1}$ is row-diagonally-dominant, $\mathcal{E}(s)$ for $s \in \{-1, 1\}^N$ is achieved at x^* that satisfies the equality constraints*

$$\text{diag}(s) M_h x^* = \delta \mathbf{1}. \quad (9)$$

Proof: Since the constraint set is linear, the optimization problem (4) is strongly dual. We now solve the optimization problem by forming its dual. The Lagrangian is given by

$$L(\lambda, x) = \|x\|^2 + \lambda^T (\delta \mathbf{1} - \text{diag}(x) M_h x).$$

The gradient of the Lagrangian with respect to x is given by $2x - (\text{diag}(x) M_h)^T \lambda$, which gives

$$x^* = \frac{(\text{diag}(x) M_h)^T \lambda}{2}.$$

Substituting x^* in the Lagrangian, the dual problem is given by

$$\max_{\lambda \geq 0} \underbrace{\frac{-\lambda^T \text{diag}(x) M_h M_h^T \text{diag}(x) \lambda}{4}}_{g(\lambda)} + \delta \lambda^T \mathbf{1}. \quad (10)$$

We have

$$\nabla g(\lambda) = \frac{-\text{diag}(s) M_h M_h^T \text{diag}(s) \lambda}{2} + \delta \mathbf{1}. \quad (11)$$

Setting the gradient to zero in (11), we obtain the optimal x^* as the solution to

$$-\text{diag}(s) M_h x^* + \delta \mathbf{1} = 0,$$

which is exactly equivalent to (9). The only caveat is that the λ obtained from $g(\lambda) = 0$ should be in the positive orthant. Solving $g(\lambda) = 0$, we obtain

$$\lambda^* = 2\delta \text{diag}(s) (M_h M_h^T)^{-1} \text{diag}(s) \mathbf{1}.$$

We want $\lambda^* \geq 0$ for every s . We can easily observe that this is indeed true if $(M_h M_h^T)^{-1}$ is diagonally-dominant. ■

Hence, from Lemma 1, when $(M_h M_h^T)^{-1}$ is diagonally-dominant, $x^* = \delta M_h^{-1} s$. Hence,

$$\mathcal{E}(s) = \|x^*\|^2 = \delta^2 s^T G s, \quad (12)$$

where $G = (M_h M_h^T)^{-1}$.

In the next lemma, the mean energy $\bar{\mathcal{E}}$ is characterized in terms of the channel matrix M_h for diagonally-dominant channels.

Lemma 2. *The mean energy for diagonally-dominant channels is given by*

$$\bar{\mathcal{E}} = \delta^2 \text{tr}(M_h^{-T} M_h^{-1}).$$

Proof: The energy $\mathcal{E}(s)$ in (12) can be expanded as

$$\mathcal{E}(S) = \delta^2 \sum_{i=1}^N G_{ii} + \sum_{i, j, i \neq j} G_{ij} s_i s_j. \quad (13)$$

Hence,

$$\begin{aligned} \frac{1}{2^N} \sum_{s \in \{-1, 1\}^N} \mathcal{E}(s) &= \\ \frac{\delta^2}{2^N} \sum_{s \in \{-1, 1\}^N} \sum_{i=1}^N G_{ii} &+ \underbrace{\sum_{i \neq j} G_{ij} \frac{1}{2^N} \sum_{s \in \{-1, 1\}^N} s_i s_j}_{=0} \end{aligned}$$

The second term in the above sum is zero since the summation spans over all the sequences on $\{1, -1\}^N$. Hence

$$\bar{\mathcal{E}} = \delta^2 \sum_{i=1}^N G_{ii} \stackrel{(a)}{=} \delta^2 \text{tr}(M_h^{-T} M_h^{-1}),$$

where (a) follows from the definition of the matrix G . ■

We now characterize $\bar{P}_h \triangleq \lim_{N \rightarrow \infty} \bar{\mathcal{E}}/N$, which is the minimum average power needed for capacity of 1 bit, in terms of the Fourier transform of the channel h . Let the discrete-Fourier transform of the channel be

$$f(\lambda) = \sum_{k=0}^{L-1} h_k e^{jk\lambda}. \quad (14)$$

Since M_h is a circulant matrix, it is easy to see that

$$\bar{\mathcal{E}} = \delta^2 \text{tr}(M_h^{-T} M_h^{-1}) = \delta^2 \sum_{k=1}^N \frac{1}{|f(\frac{2\pi k}{N})|^2}. \quad (15)$$

Using standard arguments, it can be shown that [10]

$$\bar{P}_h \rightarrow \frac{\delta^2}{2\pi} \int_0^{2\pi} \frac{1}{|f(\lambda)|^2} d\lambda. \quad (16)$$

Observe that \bar{P}_h is the energy of the inverse of the channel scaled by δ^2 , and is related to the power needed for zero-forcing.

In summary, for diagonally-dominant channels, the approximate ISI capacity is given by

$$C_\delta(P) = \begin{cases} 1 & \text{if } P \geq \bar{P}_h, \\ \beta P + \ln(Z) & \text{if } \underline{P}_h \leq P \leq \bar{P}_h, \\ 0 & \text{if } P < \underline{P}_h, \end{cases}$$

where $\underline{P}_h = \lim_{N \rightarrow \infty} \mathcal{E}_{\min}/N$ (assuming limit exists) and Z is the normalizing constant for the Gibbs distribution.

IV. ACHIEVABLE SCHEMES

We now consider achievable schemes for the approximate ISI channel under that assumption that the channel matrix M_h is invertible. In achievable schemes, an information sequence $B \in \{-1, 1\}^N$ with a well-chosen distribution is encoded into a channel input x that satisfies $|y_n| \geq \delta$. The rate of transmission over the approximate ISI channel is $H(B)/N$.

A. Zero-forcing with Gibbs distribution

Let $b = \{b_1, b_2, \dots, b_N\} \in \{-1, 1\}^N$ be an instance of the information sequence B . Choose the input to the channel as

$$x = \delta M_h^{-1} b, \quad (17)$$

which implies that the output of the ISI channel is $y = M_h x = \delta b$. This is referred to as zero-forcing because it involves channel inversion. Hence, the output of the quantizer s equals the information sequence b . When the channel is diagonally dominant, Lemma 1 and Section III-B imply that a Gibbs distribution on B and the choice $x = \delta M_h^{-1} s$ as the input to the channel results in a capacity-achieving scheme. Hence, for diagonally-dominant h , the scheme in (17) is optimal when b is sampled from the Gibbs distribution given in (8).

However, b is a sequence of length N , and it is well known that sampling from a Gibbs distribution has exponential complexity in N . For achieving capacity, N should be very large which makes this scheme impractical. In the next subsection, we compute the entropy rate when b is sampled from a Markov chain instead of a Gibbs distribution.

B. Zero-forcing with Markov input

As before we choose $x = \delta M_h^{-1} b$, where b is the information sequence. The sequence b is sampled from a two state Markov chain shown in Fig. 3 with transition matrix

$$P = \begin{pmatrix} \alpha & 1 - \alpha \\ 1 - \alpha & \alpha \end{pmatrix}, \quad (18)$$

where $0 \leq \alpha \leq 1$. Observe that $s_n = b_n$ and the achievable rate of this scheme equals the entropy $H(B) = H_2(\alpha)$, where $H_2(x)$ is the binary entropy function. The average transmit power, denoted $P_{zm}(\alpha)$, is given by

$$\begin{aligned} P_{zm}(\alpha) &= \frac{1}{N} \mathbb{E}[\|X\|^2] = \frac{\delta^2}{N} \mathbb{E}[B^T M_h^{-T} M_h^{-1} B] \\ &= \frac{\delta^2}{N} \text{tr}(M_h^{-T} \mathbb{E}[BB^T] M_h^{-1}). \end{aligned}$$

Using the eigenvalue decomposition of P , it can be shown

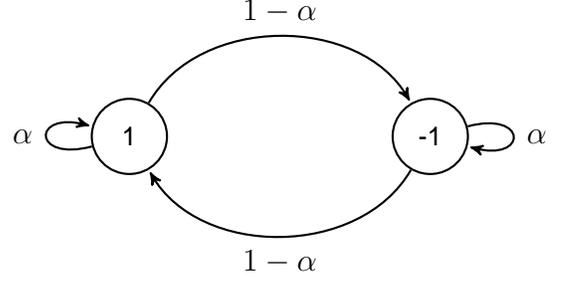


Fig. 3: Markov chain with transition matrix given in (18)

that

$$P^d = \begin{pmatrix} \frac{1+(2\alpha-1)^d}{2} & \frac{1-(2\alpha-1)^d}{2} \\ \frac{1-(2\alpha-1)^d}{2} & \frac{1+(2\alpha-1)^d}{2} \end{pmatrix}.$$

Using $(0.5, 0.5)$ as stationary distribution and P^d , we get

$$\mathbb{E}[b_n b_{n+d}] = \beta^d, \quad (19)$$

where $\beta = 2\alpha - 1$. Hence the correlation matrix $R = \mathbb{E}[BB^T]$ is given by $R_{ij} = \beta^{|j-i|}$. So for finite N

$$P_{zm}(\alpha) = \frac{\delta^2}{N} \text{tr}(R M_h^{-1} M_h^{-T}).$$

When N is large, the Toeplitz matrix R can be approximated by a circulant matrix [10] and

$$P_{zm}(\alpha) \rightarrow \frac{\delta^2}{2\pi} \int_0^{2\pi} \frac{1}{|f(\lambda)|^2} \left(\frac{2(1 - \beta \cos(\lambda))}{1 + \beta^2 - 2\beta \cos(\lambda)} - 1 \right) d\lambda. \quad (20)$$

The value of α is chosen so as to maximize the entropy rate and the rate achieved at power P , denoted $\mathcal{R}_m(P)$, is obtained as

$$\mathcal{R}_m(P) = \max_{\alpha: P_{zm}(\alpha) \leq P} H_2(\alpha). \quad (21)$$

V. NUMERICAL EXAMPLES

In this section, we evaluate the approximate ISI capacity and the rate achieved by the Markov scheme in Section IV-B for some sample channels. For numerical evaluation, we assume $\delta = 0.3$.

The channel $(1, \epsilon)$, $|\epsilon| < 1$ is a diagonally-dominant channel and zero forcing with Gibbs distribution is an optimal strategy. For this channel,

$$f(\lambda) = 1 + \epsilon e^{j\lambda}.$$

Using (16), the minimum power required for zero-forcing is given by

$$\bar{P}_h = \frac{\delta^2}{2\pi} \int_0^{2\pi} \frac{1}{1 + \epsilon^2 + 2\epsilon \cos(\lambda)} d\lambda = \frac{\delta^2}{1 - \epsilon^2}.$$

While difficult to prove theoretically, by careful simulations, it can be observed that the minimum energy \mathcal{E}_{\min} is obtained for the sequences $\pm(1, 1, \dots, 1)$. Using this observation, $\underline{P}_h = \lim_{N \rightarrow \infty} \mathcal{E}_{\min}/N$ is obtained as

$$\underline{P}_h = \frac{\delta^2}{(1 + \epsilon)^2}.$$

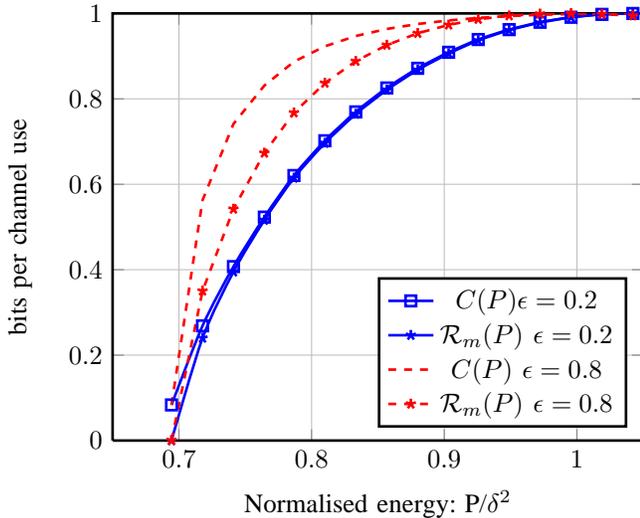


Fig. 4: $\mathcal{R}_m(P)$ and $C(P)$ versus normalised energy P/δ^2 for $\{1, \epsilon\}$ channel

The transmit power required for the two state Markov scheme (20) is

$$P_{zm}(\alpha) = \frac{\delta^2}{1 - \epsilon^2} \left[\frac{1 + \epsilon(1 - 2\alpha)}{1 - \epsilon(1 - 2\alpha)} \right].$$

Hence the maximum entropy problem for the two state Markov chain translates to

$$\mathcal{R}_m(P) = \max H_2(\alpha),$$

such that $\alpha > \frac{1}{2} + \frac{1 - P\delta^{-2}(1 - \epsilon^2)}{2\epsilon(1 + P\delta^{-2}(1 - \epsilon^2))}$.

The solution of the above problem is given by

$$\mathcal{R}_m(P) = \begin{cases} 1 & \text{if } P \geq \bar{P}_h \\ H_2\left(\frac{1}{2} + \frac{1 - P\delta^{-2}(1 - \epsilon^2)}{2\epsilon(1 + P\delta^{-2}(1 - \epsilon^2))}\right) & \text{if } \frac{\delta^2}{(1 + \epsilon)^2} \leq P < \bar{P}_h \\ 0 & \text{if } P < \frac{\delta^2}{(1 + \epsilon)^2}. \end{cases}$$

We first observe that the capacity $C_\delta(P)$ and the achievable rate $\mathcal{R}_m(P)$ match at \underline{P}_h and \bar{P}_h being equal to 0 and 1, respectively. In Fig. 4, the approximate capacity and the achievable rate of the Markov scheme are plotted as a function of normalized power P/δ^2 for $\epsilon = 0.2$ and 0.8. We observe that $\mathcal{R}_m(P)$ is very close to capacity for $\epsilon = 0.2$ and the gap increases with ϵ .

In Fig. 5, $\mathcal{R}_m(P)$ and $C_\delta(P)$ are plotted for a non-diagonally dominant channel $h = (-0.3, 1, 0.6)$. The energies $\mathcal{E}(s)$ are obtained by numerically solving (4). For this channel $\bar{P}_h/\delta^2 \approx 0.838$ and $\underline{P}_h/\delta^2 \approx 0.56$ and $\mathcal{R}_m(P) = 0$ for $P/\delta^2 < 0.59$. We observe that at higher powers the Markov scheme is close to capacity even for this non-diagonally dominant example.

VI. CONCLUDING REMARKS

The capacity of the approximate output-quantized ISI channel is characterized using Gibbs distribution, and Markov schemes are shown to approach capacity. The characterization

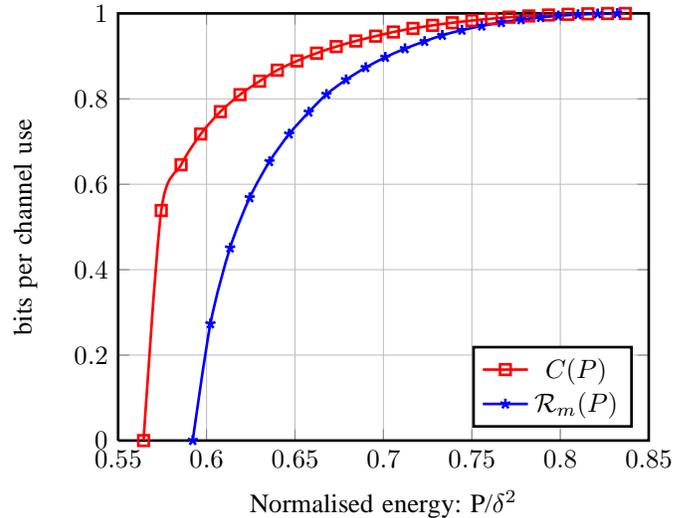


Fig. 5: $\mathcal{R}_m(P)$ and $C(P)$ versus normalised energy P/δ^2 for $\{-\epsilon, 1, 2\epsilon\}$ channel with $\epsilon = 0.3$

is complete for the case of diagonally-dominant channels. Extensions to more general channels and better achievable schemes are interesting problems for future study. Another important problem is bounding the error in the information rate because of the approximation, which is complicated by the dependencies introduced by the Gibbs distribution.

REFERENCES

- [1] S. Sun, T. Rappaport, R. Heath, A. Nix, and S. Rangan, "MIMO for millimeter-wave wireless communications: beamforming, spatial multiplexing, or both?" *Communications Magazine, IEEE*, vol. 52, no. 12, pp. 110–121, December 2014.
- [2] A. Alkhateeb, J. Mo, N. Gonzalez-Prelcic, and R. Heath, "MIMO precoding and combining solutions for millimeter-wave systems," *Communications Magazine, IEEE*, vol. 52, no. 12, pp. 122–131, December 2014.
- [3] M. Harwood, N. Warke, R. Simpson, T. Leslie, A. Amerasekera, S. Batty, D. Colman, E. Carr, V. Gopinathan, S. Hubbins, P. Hunt, A. Joy, P. Khandelwal, B. Killips, T. Krause, S. Lytollis, A. Pickering, M. Saxton, D. Sebastio, G. Swanson, A. Szczepanek, T. Ward, J. Williams, R. Williams, and T. Willwerth, "A 12.5gb/s serdes in 65nm cmos using a baud-rate adc with digital receiver equalization and clock recovery," in *Solid-State Circuits Conference, 2007. ISSCC 2007. Digest of Technical Papers. IEEE International*, Feb 2007, pp. 436–591.
- [4] S. Shamai and R. Laroia, "The intersymbol interference channel: lower bounds on capacity and channel precoding loss," *Information Theory, IEEE Transactions on*, vol. 42, no. 5, pp. 1388–1404, Sep 1996.
- [5] P. Sadeghi, P. Vontobel, and R. Shams, "Optimization of information rate upper and lower bounds for channels with memory," *Information Theory, IEEE Transactions on*, vol. 55, no. 2, pp. 663–688, Feb 2009.
- [6] J. Singh, O. Dabeer, and U. Madhow, "On the limits of communication with low-precision analog-to-digital conversion at the receiver," *Communications, IEEE Transactions on*, vol. 57, no. 12, pp. 3629–3639, December 2009.
- [7] J. Mo and R. Heath, "High SNR capacity of millimeter wave MIMO systems with one-bit quantization," in *Information Theory and Applications Workshop (ITA), 2014*, Feb 2014, pp. 1–5.
- [8] E. T. Jaynes, "Information theory and statistical mechanics," *Physical review*, vol. 106, no. 4, p. 620, 1957.
- [9] K. Conrad, "Probability distributions and maximum entropy, <http://www.math.uconn.edu/~kconrad/blurbs/entropy.pdf>."
- [10] R. M. Gray, *Toeplitz and circulant matrices: A review*. Now Publishers inc, 2006.