Numerically Computing Achievable Rates of Memoryless Channels

Francisco Javier García-Gómez
Institute for Communications Engineering
Technical University of Munich
javier.garcia@tum.de

***Abstract*—This is a short tutorial document explaining a well-known method to numerically compute achievable rates for memoryless channels. A MATLAB® function that implements the method is also provided.**

I. Mutual information and mismatched channel models

Consider a memoryless channel with conditional probability $p_{Y|X}(y|x)$, where the channel output Y is a continuous random variable. Assume that the channel input X is drawn from a discrete, memoryless source with probability mass function $P_X(x)$ for $x \in \mathcal{X}$. Here, $\mathcal{X}$ is the finite *alphabet*, or *constellation* of source $\mathcal{X}$:

$$\mathcal{X} = \{x_n\}_{n=1}^{N} \tag{1}$$

where N is the *constellation size*, and $x_n \in \mathbb{R}^D$ are real, D-dimensional *constellation points*.

The maximum achievable communication rate of this system (in bits) is given by the *mutual information* (MI) [1], [2], [3]:

$$I(X;Y) = \mathrm{E}\left[\log_2 \frac{p_{Y|X}(y|x)}{\sum_{x'\in\mathcal{X}} P_X(x')p_{Y|X}(y|x')}\right]. \tag{2}$$

Achieving this rate usually requires coded modulation [4].

In many cases, a closed-form expression for $p_{Y|X}$ in (2) is not known. In this case, one can use a *mismatched channel model* $q_{Y|X}$ [2, Exercise 5.22], [5], [6], [7] to obtain a lower bound on the mutual information (2). Replacing $p_{Y|X}$ with $q_{Y|X}$ in (2) (while still taking the expectation over the true channel $p_{Y|X}$) gives the *mismatched mutual information*

$$I_q(X;Y) = \mathrm{E}\left[\log_2 \frac{q_{Y|X}(y|x)}{\sum_{x'\in\mathcal{X}} p_X(x')q_{Y|X}(y|x')}\right]. \tag{3}$$

The mismatched mutual information is a *lower bound on capacity*, i.e.,

$$I_q(X;Y) \leq I(X;Y). \tag{4}$$

This can be proved using error exponents [6], but a simpler proof is provided in [8, Eqs. (37)-(41)], where it is shown that

$$I_q(X;Y) = I(X;Y) - D\left(P_X p_{Y|X} \middle\| p_Y r_{X|Y}\right) \tag{5}$$

with

$$P_Y(y) = \sum_{x'\in\mathcal{X}} P_X(x')p_{Y|X}(y|x') \tag{6}$$

being the true output probability, and

$$r_{X|Y}(x|y) = \frac{P_X(x)q_{Y|X}(y|x)}{\sum_{x'\in\mathcal{X}} p_X(x')q_{Y|X}(y|x')} \tag{7}$$

being the reverse conditional probability of the mismatched model. The expression

$$D(f\|g) = \sum_{x\in\mathcal{X}} \int_y f(x,y)\log_2 \frac{f(x,y)}{g(x,y)}\,\mathrm{d}y \geq 0 \tag{8}$$

is the Kullback-Leibler divergence. In (4), equality is achieved if and only if the $D(\cdot\|\cdot)$ term in (5) becomes 0, i.e, if and only if $p_{Y|X} = K q_{Y|X}$, where K is an arbitrary constant.

Remark. The interpretation of $q_{Y|X}$ as a mismatched channel model requires

$$\int_y q_{Y|X}(y|x)\,\mathrm{d}y = 1. \tag{9}$$

In the literature, mismatched channel models are often used by *mismatched maximum-likelihood decoders*, in which case $q_{Y|X}$ is a decoding metric and does not need to satisfy (9). All the results in this document are valid without the constraint (9).

A. Numerical computation of lower bounds on capacity

Let

$$\mathbf{s} = (s_1, \ldots, s_M) \tag{10}$$

(with $s_m \in \mathcal{X}\ \forall m$) be a long sequence of M symbols drawn from source X, and let

$$\mathbf{r} = (r_1, \ldots, r_M) \tag{11}$$

be the corresponding sequence of output symbols, such that $\{(s_m, r_m)\}_{m=1}^{M}$ are distributed according to $P_X(s_m)p_{Y|X}(r_m|s_m)$. These sequences can be obtained from Monte-Carlo simulations of the source and channel, or from experiments.

A numerical lower bound on the capacity of the channel $p_{Y|X}$ can be obtained by applying (3) to $\mathbf{s}$ and $\mathbf{r}$. For large M, we have

$$I_q(X;Y) \approx \frac{1}{M}\sum_{m=1}^{M} \log_2 \frac{q_{Y|X}(r_m|s_m)}{\sum_{n=1}^{N} p_X(x_n)q_{Y|X}(r_m|x_n)}. \tag{12}$$

Date of current version January 2, 2020. J. García-Gómez is with the Institute for Communications Engineering (LNT), Technical University of Munich, Germany (e-mail: javier.garcia@tum.de). His work was supported by the German Research Foundation under Grant KR 3517/8-2.

Remark. Equations (3) and (12) assume a channel with discrete input and continuous output. They can be extended to channels with continuous input and continuous output by replacing the sum in the denominator with an integral. The integral can either be computed in closed form or by Monte-Carlo integration, but care must be taken that the result is accurate.

II. Conditionally Gaussian mismatched model

One possibility for memoryless channels with discrete input and continuous output is to use a conditionally Gaussian mismatched model $q_{Y|X}$. That is, for every constellation point $x_n \in \mathcal{X}$, we approximate the unknown conditional probability $p_{Y|X}$ by

$$q_{Y|X}(y|x_n) = \frac{1}{\sqrt{\det\left(2\pi \mathbf{C}_n\right)}} e^{-\frac{1}{2}(y-\mu_n)^T \mathbf{C}_n^{-1}(y-\mu_n)} \tag{13}$$

i.e., a Gaussian distribution whose mean μ_n and covariance matrix $\mathbf{C}_n$ depend on the input constellation point x_n.

To obtain the parameters μ_n and $\mathbf{C}_n$ for $n \in \{1, \dots, N\}$ of our mismatched channel model, we use the sequences $\mathbf{s}$ and $\mathbf{r}$ obtained from simulations or experiments of the real channel. The parameters are estimated as

$$\mu_n = \frac{1}{|\mathcal{M}_n|} \sum_{m \in \mathcal{M}_n} r_m \tag{14}$$

$$\mathbf{C}_n = \left(\frac{1}{|\mathcal{M}_n|} \sum_{m \in \mathcal{M}_n} r_m r_m^T \right) - \mu_n \mu_n^T \tag{15}$$

where

$$\mathcal{M}_n = \{m \colon s_m = x_n\} \tag{16}$$

is the set of time indices m for which constellation point x_n was transmitted.

Substituting (13) in (12) yields our numerical lower bound on capacity:

$$\begin{aligned} I_q(X;Y) \approx \frac{1}{M} \sum_{m=1}^{M} \Bigg\{ & -\frac{1}{2} \log_2 \det \mathbf{C}_{n_m} \\ & - \frac{1}{2 \log 2} \left(r_m - \mu_{n_m}\right)^T \mathbf{C}_{n_m}^{-1} \left(r_m - \mu_{n_m}\right) \\ & - \log_2 \left[\sum_{n=1}^{N} \frac{P_X(x_n)}{\sqrt{\det \mathbf{C}_n}} e^{-\frac{1}{2}(r_m-\mu_n)^T \mathbf{C}_n^{-1}(r_m-\mu_n)} \right] \Bigg\} \end{aligned} \tag{17}$$

where

$$n_m = (n \colon s_m = x_n) \tag{18}$$

is the index n of the constellation point $s_m = x_n$ transmitted at time m.

A method for computing achievable rates of channels with memory is available in [8].

III. MATLAB® software tool: MI-CG

We provide in [9] a MATLAB® function `mi_cg.m` that implements the method explained in this document. Two example scripts are provided that compute achievable rates for well-known channels.

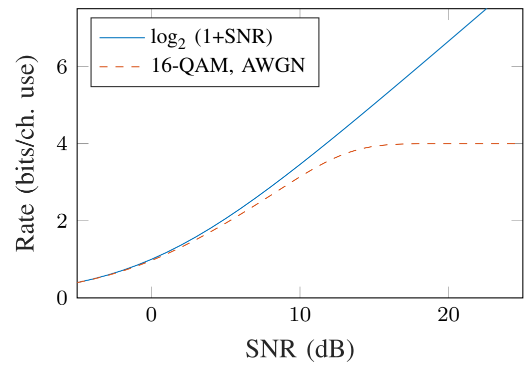

Figure 1. Achievable rate of a complex AWGN channel using uniform 16-QAM modulation, computed using the software tool in [9].

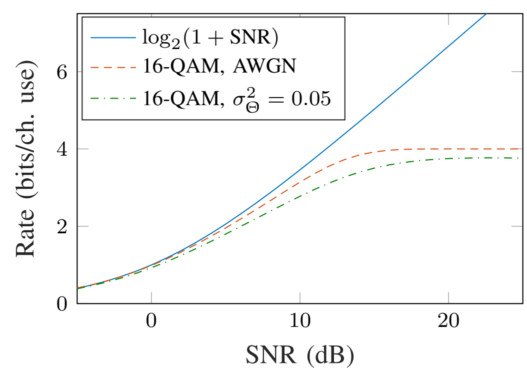

Figure 2. Achievable rate of a complex AWGN channel with and without phase noise using uniform 16-QAM modulation, computed using the software tool in [9].

A. Complex AWGN channel with 16-QAM constellation

The script `example_mi_cg_complex_16qam.m` simulates a complex additive white Gaussian noise (AWGN) channel:

$$Y = X + Z \tag{19}$$

where Z has a circularly symmetric complex Gaussian (CSCG) distribution with zero mean and unit variance. The constellation of the source is 16-QAM:

$$\mathcal{X} = \left\{ \sqrt{\frac{P}{10}} (n + jm) \right\}, \quad n, m \in \{-3, -1, 1, 3\} \tag{20}$$

where P is the transmit power and $j = \sqrt{-1}$ is the imaginary unit. By varying P, the script tests different values of the signal-to-noise ratio (SNR) and generates Fig. 1.

B. Partially coherent complex AWGN channel with 16-QAM constellation

The script `example_mi_cg_real_2D_awgn_vs_phasenoise.m` simulates a partially coherent complex AWGN channel [10]:

$$Y = X e^{j\Theta} + Z \tag{21}$$

where Z is CSCG with zero mean and unit variance, and Θ has a Gaussian distribution with mean $\overline{\theta}$ and variance σ_Θ^2:

$$p_\Theta(\theta) = \frac{1}{\sigma_\Theta\sqrt{2\pi}} e^{-\frac{(\theta-\overline{\theta})^2}{2\sigma_\Theta^2}}. \tag{22}$$

The program uses $\overline{\theta} = 0.2$ and $\sigma_\Theta^2 = 0.05$ and generates Fig. 2, which agrees with [11, Fig. 3, triangular markers]. As expected, the nonzero mean of Θ does not have any effect on the achievable rate.

IV. Acknowledgment

The author wishes to thank Prof. Gerhard Kramer for useful suggestions and proofreading this document.

References

[1] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, July 1948.

[2] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY, USA: John Wiley & Sons, Inc., 1968.

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: John Wiley & Sons, Inc., 2006.

[4] G. Ungerboeck, "Trellis-coded modulation with redundant signal sets part i: Introduction," *IEEE Communications Magazine*, vol. 25, no. 2, pp. 5–11, February 1987.

[5] T. R. M. Fischer, *Some Remarks on the Role of Inaccuracy in Shannon's Theory of Information Transmission*. Dordrecht: Springer Netherlands, 1978, pp. 211–226. [Online]. Available: https://doi.org/10.1007/978-94-009-9857-5_20

[6] G. Kaplan and S. Shamai, "Information rates and error exponents of compound channels with application to antipodal signaling in a fading environment," *AEÜ. Archiv für Elektron. und Übertragungstech.*, vol. 47, no. 4, pp. 228–239, 1993.

[7] N. Merhav, G. Kaplan, A. Lapidoth, and S. Shamai Shitz, "On information rates for mismatched decoders," *IEEE Transactions on Information Theory*, vol. 40, no. 6, pp. 1953–1967, Nov 1994.

[8] D. M. Arnold, H. . Loeliger, P. O. Vontobel, A. Kavcic, and W. Zeng, "Simulation-based computation of information rates for channels with memory," *IEEE Transactions on Information Theory*, vol. 52, no. 8, pp. 3498–3508, Aug 2006.

[9] F. J. Garcia-Gomez, "MI-CG: A function to numerically compute an achievable rate for a memoryless channel," December 2019. [Online]. Available: https://gitlab.lrz.de/javier-garcia-gomez/mutual-information

[10] B. Goebel, R. Essiambre, G. Kramer, P. J. Winzer, and N. Hanik, "Calculation of mutual information for partially coherent gaussian channels with applications to fiber optics," *IEEE Transactions on Information Theory*, vol. 57, no. 9, pp. 5720–5736, Sep. 2011.

[11] T. Fehenberger and N. Hanik, "Mutual Information as a Figure of Merit for Optical Fiber Systems," *arXiv e-prints*, p. arXiv:1405.2029, Apr 2014.