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 \( X \):

\[
\mathcal{X} = \{x_n\}_{n=1}^N
\]

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]:

\[
I(X;Y) = \mathbb{E} \left[ \log_2 \frac{p_{Y|X}(y|x)}{\sum_{x' \in \mathcal{X}} p_X(x') p_{Y|X}(y|x')} \right].
\]

(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

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

(3)

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

\[
I_q(X;Y) \leq I(X;Y).
\]

(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

\[
I_q(X;Y) = I(X;Y) - D(P_X p_{Y|X} \| p_Y q_{Y|X})
\]

(5)

with

\[
P_Y(y) = \sum_{x' \in \mathcal{X}} P_X(x') p_{Y|X}(y|x')
\]

(6)

being the true output probability, and

\[
r_{X|Y}(x|y) = \frac{P_X(x) q_{Y|X}(y|x)}{\sum_{x' \in \mathcal{X}} P_X(x') q_{Y|X}(y|x')}
\]

(7)

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

\[
D(f\|g) = \int_{x \in \mathcal{X}} f(x) \log_2 \frac{f(x)}{g(x)} \, dy \geq 0
\]

(8)

is the Kullback-Leibler divergence. In (4), equality is achieved if and only if the \( D(f\|\cdot) \) 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

\[
\int_y q_{Y|X}(y|x) \, dy = 1.
\]

(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

\[
s = (s_1, \ldots, s_M)
\]

(10)

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

\[
r = (r_1, \ldots, r_M)
\]

(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 \( s \) and \( r \). For large \( M \), we have

\[
I_q(X;Y) \approx \frac{1}{M} \sum_{m=1}^M \log_2 \frac{q_{Y|X}(r_m | s_m)}{\sum_{n=1}^N P_X(x_n) q_{Y|X}(r_m | x_n)}.
\]

(12)
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

$$q_{Y|X}(y|x_n) = \frac{1}{\sqrt{\det{(2\pi C_n)}}} e^{-\frac{1}{2}(y-\mu_n)^T C_n^{-1}(y-\mu_n)}$$

(13)

i.e., a Gaussian distribution whose mean $\mu_n$ and covariance matrix $C_n$ depend on the input constellation point $x_n$.

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

$$\mu_n = \frac{1}{|M_n|} \sum_{m \in M_n} r_m$$

(14)

$$C_n = \left(\frac{1}{|M_n|} \sum_{m \in M_n} r_m r_m^T \right) - \mu_n \mu_n^T$$

(15)

where

$$M_n = \{m : s_m = x_n\}$$

(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:

$$I_q(X;Y) \approx \frac{1}{M} \sum_{m=1}^M \left\{ -\frac{1}{2} \log_2 \det C_{n_m}^{-1} \right. \right.$$  

$$\left. - \frac{1}{2} \log_2 \left( r_m - \mu_{n_m} \right)^T C_{n_m}^{-1} \left( r_m - \mu_{n_m} \right) \right.$$  

$$\left. - \log_2 \left[ \frac{1}{\sqrt{\det C_n}} \sum_{n=1}^N P_X(x_n) e^{-\frac{1}{2} (r_m - \mu_{n_m})^T C_{n_m}^{-1} (r_m - \mu_{n_m})} \right] \right\}$$

(17)

where

$$n_m = (n : s_m = x_n)$$

(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.

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:

$$Y = X + Z$$

(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:

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

(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]:

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

(21)
where $Z$ is CSCG with zero mean and unit variance, and $\Theta$ has a Gaussian distribution with mean $\overline{\theta}$ and variance $\sigma_\Theta^2$:

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

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 $\Theta$ 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