Last time, we discussed the channel capacity for a point-to-point channel and moreover the channel capacity region for the multiple-access channel.

Now we want to extend in two ways

- Can we be universal, so we don't need to know the channel statistics in advance, yet still achieve the capacity / capacity region

  (That is we won't know the typical sets, so can't base a coding scheme on them).

- Coding theorems showed that an arbitrarily small error probability is possible for rate pairs within the capacity region, but does not tell how large the block length must be to achieve a specified error probability

  → a partial answer is provided by examining error exponents

  → a finer characterization follows from second-order and higher order characterizations, which we won't pursue today.

  Capacity: law of large numbers

  Second-Order: quantified central limit theorem

  Error Exponent: large deviations

<u>Definition</u>: A number $E \geq 0$ is called an achievable error exponent at rate pair (R_x, R_y) for a discrete memoryless multiple access channel $W: \mathcal{X} \times \mathcal{Y} \mapsto \mathcal{Z}$ if, for every $\delta > 0$ and all sufficiently large n, there exists a multiuser code C of block length n such that:

$$M_X \geq \exp\{n(R_X-\delta)\}$$
$$M_Y \geq \exp\{n(R_Y-\delta)\}$$

and $$e(c,w) \leq \exp\{-n(E-\delta)\}$$

The largest achievable error exponent at rate pair (R_X, R_Y) considered as a function of R_X and R_Y, is called the reliability function of the MAC and denoted by $E(R_X, R_Y, W)$.

[recall that here M_X is the # codewords for X; R_X is $\frac{1}{n}\log M_X$
M_Y is the # codewords for Y; R_Y is $\frac{1}{n}\log M_Y$]

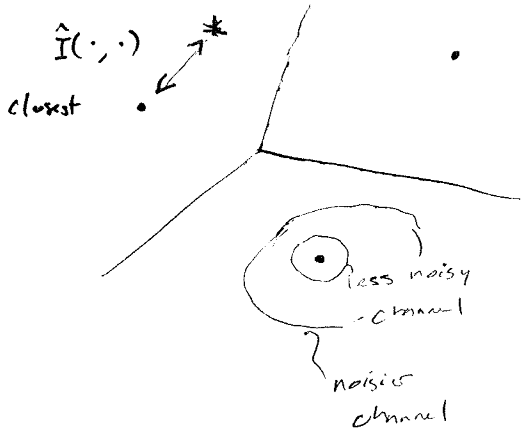
Exponentially decreasing error probability is also generally true for single-user and all kinds of other channels:

→ random coding arguments with refined analysis of performance can provide bounds on the reliability function: generally not tight. (In fact matching upper and lower bounds on the error exponent are not generally known).

What about universality?

In the single-user setting, Goppa (1975) introduced the maximum mutual information (MMI) decoder to universally achieve capacity.

The basic idea is to define universal decoding sets around each codeword that are based on which codeword has the most empirical mutual information with the received sequence.



will work for either

notice that computing the empirical mutual information does not require knowing the typical set. Yet it is something that will be closest to the transmitted/received codeword with high probability since mutual information is a measure of correlation or distance from axiomatic derivation.

In fact the MMI decoder and similar techniques are often used in signal processing and machine learning, without really invoking information-theoretic coding theorems. Example is mutual information-based image registration, or signal separation.

[opportunity for clearer performance analysis].

Consider "A new universal random coding bound for the multiple-access channel," by Yu-sun Liu and Brian L. Hughes, IEEE Trans. IT, Mar. 1996.

Before getting into it, define a multivariate information measure called multiinformation, e.g. among three random variables.

$$I(X;Y;Z) = H(X) + H(Y) + H(Z) - H(X,Y,Z).$$

which naturally extends mutual information $I(X;Y) = H(X) + H(Y) - H(X,Y)$.

Also the conditional multiinformation

$$I(X;Y;Z|U) = H(X|U) + H(Y|U) + H(Z|U) - H(X,Y,Z|U).$$

[These multivariate information measures are also useful elsewhere.
See e.g. "Illum Information" by Raman, Yu, and Varshney (ITA, 2017).

A two-user, discrete memoryless MAC is defined by stochastic matrix
$W: \mathcal{X}\times\mathcal{Y} \mapsto \mathcal{Z}$. The channel transition probability for n-sequences is given by $W^n(\vec{z}|\vec{x},\vec{y}) = \prod_{i=1}^{n} W(z_i|x_i,y_i)$.

A multiuser code of blocklength n is a collection

$$C = \{(\vec{x}_i, \vec{y}_i, D_{ij}) : i=1,\ldots,M_X \text{ and } j=1,\ldots,M_Y\}.$$

where there are $M_X M_Y$ disjoint decoding sets $D_{ij} \subset \mathcal{Z}^n$
where $\bigcup_{ij} D_{ij} = \mathcal{Z}^n$.

[Let $\mathcal{M}(\mathcal{X})$ be set of all probability measures on alphabet $\mathcal{X}$]

<u>Thm</u> (universal random coding bound on error exponent for MAC)

preliminaries. For every finite set $\mathcal{U}$, let $\mathcal{P}(\mathcal{U})$ be set of all joint probability measures $P_{XYU} \in \mathcal{M}(\mathcal{X}\times\mathcal{Y}\times\mathcal{U})$.

The collection of <u>types</u> contained within each sets is denoted

$$\mathcal{P}_n(\mathcal{U}) = \mathcal{P}(\mathcal{U}) \cap \mathcal{M}_n(\mathcal{X}\times\mathcal{Y}\times\mathcal{Z})$$

Thm

For every finite set $\mathcal{U}$, $P_{XYU} \in \mathcal{P}_n(\mathcal{U})$, $R_x \geq 0$, $R_y \geq 0$, $\delta > 0$

and $\vec{u} \in \mathcal{T}^n_{P_u}$ (where $\mathcal{T}^n_{P_u}$ is set of types P_u of length n).

there exists a multiuser code

$$C = \{(\vec{x}_i, \vec{y}_i, D_{ij}) : i = 1, \ldots, M_x;\ j = 1, \ldots, M_y\}.$$

with a particular reliability function expression given in the paper.

$$E_r(R_x, R_y, W).$$

This expression is fairly complicated, but we can get some more insight as follows.

The exponent $E_r(R_x, R_y, W)$ is positive for all rate pairs in the interior of the capacity region, which corresponds to

$$R_x < I(X; Y, Z \mid U)$$
$$R_y < I(Y; X, Z \mid U)$$
$$R_x + R_y < I(X; Y; Z \mid U).$$

where Z, X, Y, U are random variables with joint distribution $W \times P_{XYU}$.

Proof of achievability is from using the MMI decoder, and its slight simplification, the minimum emirical equivocation decoder.

In ~~the~~ the thm, the sum rate is bounded by the multiinformation. which is the tight bound in many cases.