CNES - Geraldine Artaud – note on mutual information 22/04/2020

Mutual Information

# Mutual information

## introduction

Mutual information is widely used in information theory to build communication systems. This note is not intended to explain rigorously this concept; it only present a way to compute it when performing simulations.

In optical coms, post-FEC performances were initially obtained using pre-FEC BER measurements: to compute the BER, without having to code and decode, Pre-FEC BER/post-FEC BER curves were used (FEC limit paradigm). The use of mutual information as a FEC design metric has been introduced [1].

Mutual Information (MI), gives consistent post-FEC FER predictions across different channel conditions and modulation formats. The mutual information quantifies the "amount of information" obtained about the transmitted data (input) through observing the received data (output of the channel). It takes into account the “Optical Transmission Channel”, composed of the Propagation channel (turbulences, absorptions) and the Receiver architecture (telescope diameter, receiver type ( APD, preamplified SMF), noises

The MI gives the corresponding achievable Code Rate: theoretically if MI >Code Rate => a code exists that allows no residual error after decoding

The mutual information is defined by the information shared between the symbol ***X*** a the input of a channel and the corresponding symbol ***Y*** at the output of the channel.

For a channel with discreet input and outputs, the IM is given by :

* the joint probability of **X** and **Y** ;
* ***P(X|Y )*** the probability of **X** conditionally to **Y**.
* ***P(Y|X )*** the probability of **Y** conditionally to **X**.

The representation 2 and 3 are easier to manipulate (see matlab code bellow).

***X*** are the binary inputs 1 or 0 and ***Y*** represents the received signal before decoding occurs.

## Estimation of mutual information

The calculation of mutual information is possible using the received symbols, when you know the sequence of bits that have been sent through the channel.

Practically the computation is done using histograms that allow to approximate the distributions of the received symbols, knowing the transmitted symbols.

In the following figure an example of histogram is given. The distribution of the received signal knowing a 0 was transmitted corresponds to the histogram in the front, and the distribution of the received signal knowing a 1 was transmitted corresponds to the histogram in the back.

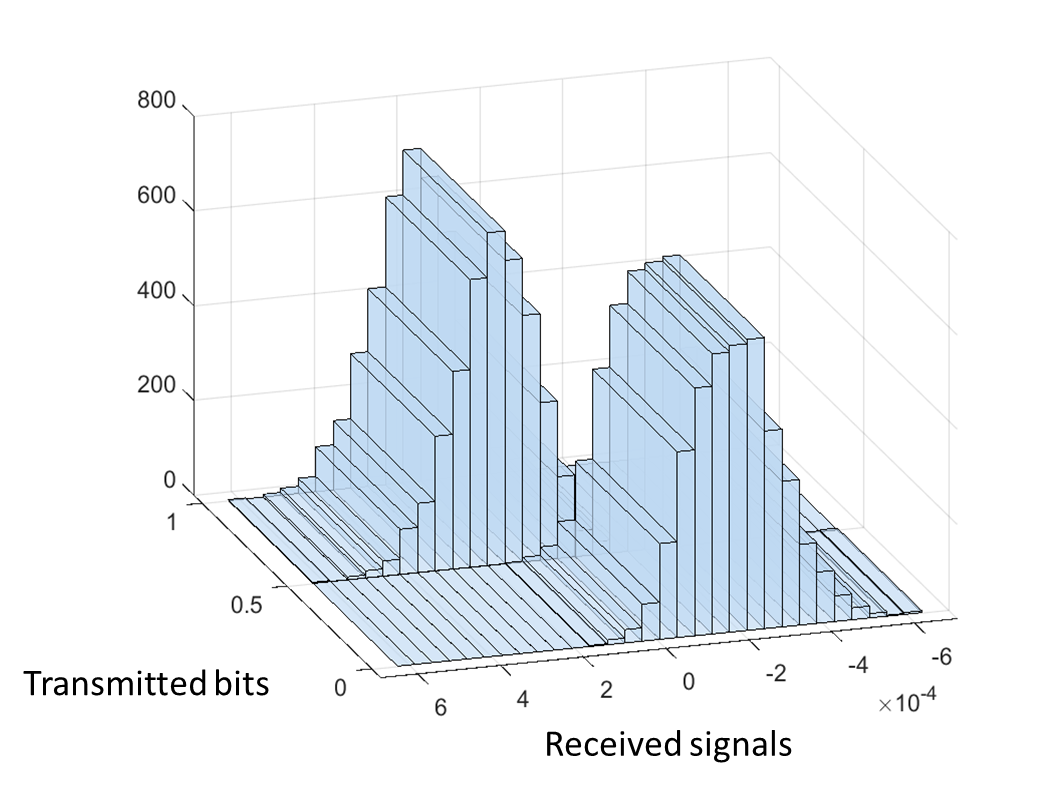


Figure 1 : Histogram of received signals vs transmitted signals

The choice of the number of symbols and the number of bins of the histogram in order to obtain a proper distribution estimation shall be done carefully.

Typical values can be:

* a minimum number of 213=8192 bits is required, and 30 bins for the histogram
* a minimum number of 216= 65536 bits is required, and 40 bins for the histogram

The mutual information can be computed for a multiple of the codeword length. For each codeword if it is long enough (for example 30K codes), or for a group of 5 codewords for a 2K long code.

## matlab code

U= [Transmitted\_bits Received\_Symbols];

% - - Mutual Information computing formulation 2 - -

%create a 2D-histogram

counts=histcounts2(U(:,1),U(:,2),'Normalization','probability');

%remove 0 entris in the histogram

counts = max(counts,10e-10);

%marginal and joint pdfs

pXY = counts./sum(counts(:));

pX = sum(pXY,2);

pY = sum(pXY,1);

%compute MI

MI = pXY.\*log2(pXY./(pX\*pY));

C= sum(MI(:))

% - - Mutual Information computing formulation 3 - -

%create histograms of conditional probability

[pYX2,edges]=histcounts(U(:,2),40);

pYX2 =pYX2./sum(pYX2(:));%normalization

x=logical(U(:,1));

pY1=histcounts(y(x) ,edges);

pY1=pY1./sum(pY1(:));%normalization

pY0=histcounts(y(~x) ,edges);

pY0= pY0./sum(pY0(:));%normalization

%remove 0 entris in the histogram

pYX2 = max(pYX2,10e-10);

pY1 = max(pY1,10e-10);

pY0 = max(pY0,10e-10);

MI2=1/2\*(pY0.\*log2(pY0./pYX2)+pY1.\*log2(pY1./pYX2));

I2(i)=sum(MI2(:));

### Exemple for an AWGN chanel

clear all;close all

AdB=[-10:0.5:10];

A=10.^(AdB/10);

for(i=1:length(A))

nb=2^17;

n=randn(nb,1); % gaussian noise with variance sigma^2=N0/2=1

x=[sign(randn(nb,1))]; x(find(x<0))=0; % transmitted bits

xa=x\*A(i);

y=xa+n;

% compute histogram

[N,Xedges,Yedges] = histcounts2(x,y,[2,40], 'Normalization','probability');

%remove 0 entris in the histogram

N = max(N,10e-10);

pXY=N./sum(N(:)); %joint probability

pX= sum(pXY,2);

pY = sum(pXY,1);

MI = pXY.\*log2(pXY./(pX\*pY));

I1(i)=sum(MI(:));

%2e method

% Ik=1/2\*(pY|0\*log2(pY|0/pY|X)+pY|1\*log2(pY|1/pY|X))

[pYX2,edges]=histcounts(y,40);

pYX2 =pYX2./sum(pYX2(:));%normalization

x=logical(x);

pY1=histcounts(y(x) ,edges);

pY1=pY1./sum(pY1(:));%normalization

pY0=histcounts(y(~x) ,edges);

pY0= pY0./sum(pY0(:));%normalization

pYX2 = max(pYX2,10e-10);

pY1 = max(pY1,10e-10);

pY0 = max(pY0,10e-10);

MI2=1/2\*(pY0.\*log2(pY0./pYX2)+pY1.\*log2(pY1./pYX2));

I2(i)=sum(MI2(:));

end

figure(1)

%SNR=2\*(Es/N0)=(A^2/2)/1

SNR=A.^2/2;

plot(10\*log10(SNR),I1,10\*log10(SNR),I2,'\*')

title('mutual information')

ylabel('Mutual information C (bit/symbol)')

xlabel('SNR (dB)')

ylabel('Mutual information (bit/symbol)')

figure(2)

%Np=sqrt(Eb/N0/R)=sqrt(Es/N0)/R=sqrt(SNR/2)/R

Np1=sqrt(SNR/2)./I1;

Np2=sqrt(SNR/2)./I2;

plot(10\*log10(Np1),I1,10\*log10(Np2),I2,'\*')

title('mutual information')

ylabel('Mutual information C (bit/symbol)')

xlabel('photon per bit (dB)')

ylabel('Mutual information (bit/symbol)')

# References

[1] Alex Alvarado & all, « Replacing the Soft-Decision FEC Limit Paradigm in the Design of Optical Communication Systems », JOURNAL OF LIGHTWAVE TECHNOLOGY, VOL. 34, NO. 2, JANUARY 15, 2016 707

[2] Laurent Schmalen, Alex Alvarado, and Rafael Rios-Müler, “Performance Prediction of Nonbinary Forward Error Correction in Optical Transmission Experiments” JOURNAL OF LIGHTWAVE TECHNOLOGY, VOL. 35, NO. 4, FEBRUARY 15, 2017 1015