DUNG L□ỢNG KÊNH VÀ MÃ HOÁ KÊNH

Giới thiêu:

Mục đích của hệ thống viễn thông là truyền thông tin từ nơi này đến nơi khác. Môi tr- ờng mà trên đó thông tin đ- ợc truyền qua đ- ợc gọi là kênh truyền thông. Nội dung thông tin của nguồn tin đ- ợc đánh giá (đo) bởi Entropy của nguồn tin th- ờng theo đơn vị bit. Mô hình toán học thích hợp khảo sát nguồn tin là quá trình ngẫu nhiên.

Phần1: Xét mô hình toán phù hợp cho các kênh truyền thông. Ta cũng đề cập dung l- ợng kênh (khả năng thông qua kênh) mà đ- ợc định nghĩa cho bất kỳ một kênh truyền thông nào và đ- a ra giới hạn cơ bản về l- ợng thông tin mà có thể đ- ợc truyền qua kênh. Thực tế ta th- ờng xét hai loại kênh: kênh đối xứng cơ hai (Binary Symmetric Channel – BSC) và kênh tạp âm Gaussian trắng cộng (Additive White Gaussian Noise Channel AWGN).

Phần2: Xét các kỹ thuật mã hoá để truyền thông khả tin trên các kênh truyền thông. Ta đề cập hai kỹ thuật mã hoá đ- ợc dùng phổ biến nhất là: Mã hoá khối và mã hoá xoắn, các kỹ thuật mã hoá và giải mã cho các mã này và đề cập đến các hiệu năng của chúng.

Phần 1:

DUNG L□ỢNG KÊNH TRUYỀN DẪN

I. MỤC ĐÍCH, YÊU CẦU:

1. Mục đích:

- Nêu ra các bài tập nhỏ cơ bản để sinh viên làm quen, sau đó ứng dụng vào chuyên ngành vô tuyến. (Phần ứng dụng vào chuyên nghành đ- ợc đề cập sau).
- Thông qua ch- ơng trình đ- ợc viết trên Matlab sinh viên hiểu khái niệm & ý nghĩa vai trò dung l- ợng kênh truyền dẫn trong hệ thống viễn thông.
- Ph- ơng pháp xác định dung l- ợng kênh truyền dẫn và các thông số.

2.Yêu cầu:

- Hiểu thuật toán đ- ợc dùng trong ch- ơng trình Matlab.
- Biết cách thay đổi các thông số để khảo sát các công thức trong ch- ơng trình.
- Sử dụng ch- ơng trình Matlab để giải quyết các bài tập hoặc vấn đề (sinh viên tự lập ra các tình huống).
- Tại sao ng- ời ta sử dụng phân phối Gaussian trong việc mô hình hoá các kênh truyền thông ?
- Nêu ý nghĩa của ngẫu nhiên hoá tín hiệu theo quan điểm lý thuyết thông tin trong các hệ thống viễn thông.

II. NỘI DUNG:

2.1. Tóm tắt lý thuyết:

Phần tóm tắt lý thuyết đ- ợc lấy từ các tài liệu Cơ sở truyền dẫn vi ba số tác giả TS.Nguyễn Phạm Anh Dũng; Digital Communication tác giả John G.Proakis 2001; Digital Communication tác giả Simon Haykin 1988; Digital Modulation and Coding tác giả Stephen G.Wilson 1996. Nên có sự không đồng nhất về ký hiệu các công thức chẳng hạn độ rộng băng tần B = W, SNR=P/N_0...

Chi tiết phần lý thuyết dung l- ợng kênh đ- ợc nhắc lại ở phụ lục1.

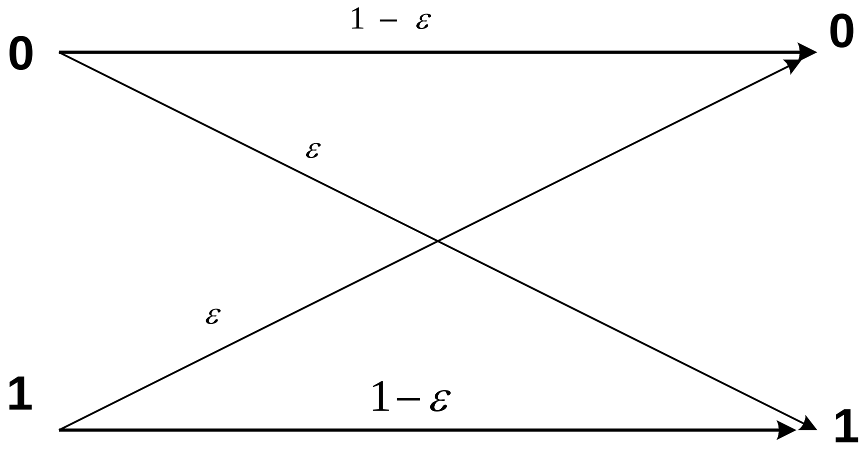
Mô hình kênh và dung l- ơng kênh.

Kênh truyền thông thực hiện truyền tín hiệu mang thông tin đến đích. Trong truyền dẫn, tín hiệu mang tin phải chựu ảnh h- ởng lớn của môi tr- ờng. Một trong các ảnh h- ởng này có tính ***tất định*** chẳng hạn suy hao, méo tuyến tính và không tuyến tính; một số có tính ***xác suất*** (ngẫu nhiên) chẳng hạn tác động của tạp âm, fading nhiều tia Multipath Fading.v.v.. ***Vì ảnh h- ởng có tính tất định đ- ợc xem là***

tr- ờng hợp đặc biệt của thay đổi ngẫu nhiên, nên một cách tổng quát xem môt hình toán học cho kênh truyền thông là sự phụ thuộc ***ngẫu nhiên*** giữa các tín hiệu vào và ra.

- **Mô hình kênh.**

Tr- ờng hợp đơn giản nhất, kênh đ- ợc mô hình hoá nh- là xác suất có điều kiện liên hệ giữa mỗi đầu ra với đầu vào t- ơng ứng của kênh. Kênh nh- vậy đ- ợc gọi là kênh không nhớ rời rạc (DMC- Discrete-Memoryless-Channel) và hoàn toàn đ- ợc biểu diễn bởi các ký hiệu vào X và ký hiệu đầu ra Y của kênh và ma trận xác suất truyền của kênh P(y|x) đối với mọi $x \in X$ và $y \in Y$. Một tr- ờng hợp cụ thể của **DMC** là **BSC**-Binary Symmetric-Channel mà đ- ợc coi là mô hình toán học cho việc truyền dẫn cơ hai trên kênh Gaussian với ***quyết định cứng*** ở đầu ra. Kênh đối xứng cơ hai BSC t- ơng ứng với tr- ờng hợp X=Y={0,1} và P(y=0|x=1) = P(y=1|x=0) = ε đ- ợc cho ở hình 2.1. Thông số ε đ- ợc gọi là xác suất giao nhau CrossOver của kênh.



Hình 2.1: Kênh đối xứng cơ hai BSC

- **Dung l- ợng kênh hay khả năng thông qua của kênh.**

Theo định nghĩa, khả năng thông qua của kênh là tốc độ cực đại mà tại tốc độ này truyền thông tin trên kênh đó đảm bảo đ- ợc độ tin cậy.

Khả năng thông qua của kênh đ- ợc ký hiệu C; theo định nghĩa này thì tại các tốc độ R<C, cho phép truyền dẫn khả tin trên kênh đó, tại tốc độ R>C thì việc truyền tin trên kênh là không thể đảm bảo tin cậy đ- ợc.

Kết quả cơ bản lý thuyết thông tin của Shanno phát biểu rằng đối các kênh rời rạc không nhớ thì khả năng thông qua đ- ợc cho bởi biểu thức sau.

$$C = \underset{p(x)}{Max}\, I[X;Y] \qquad (2.1)$$

Trong đó **I**[X;Y] là lượng thông tin chéo trung bình giữa X (đầu vào kênh) và Y (đầu ra của kênh) việc cực đại hoá được thực hiện trên toàn bộ ***phân phối xác suất đầu vào của kênh***.

Thông tin chéo trung bình giữa hai biến ngẫu nhiên X và Y được định nghĩa là.

$$I(X,Y) = \sum_{x\in X}\sum_{y\in Y} p(x)p(y \mid x)\log\frac{p(x,y)}{p(x)p(y)} \quad (2.2)$$

Trong đó:

Lượng thông tin chéo lấy theo log cơ số 2 thì được đơn vị bit .

- **Mô hình kênh BSC:**

Khả năng thông qua kênh được cho bởi mối quan hệ đơn giản hơn.

$$C = 1 - H_b(\varepsilon) \quad (2.3)$$

Trong đó ε là xác suất chéo của kênh và Hb(.) thể hiện cho hàm Entropy cơ hai.

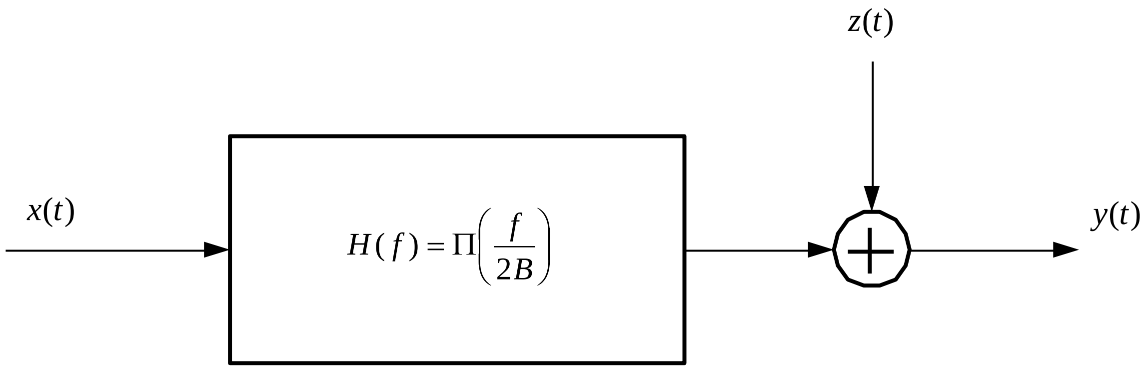
$$H_b(x) = -x\log(x) - (1-x)\log(1-x) \quad (2.4)$$

- **Mô hình kênh AWGN:**

Kênh tạp âm Gausian trắng cộng AWGN bị hạn chế băng và có công suất vào hạn chế. Kênh này được mô hình hoá như hình 2.2.

Kênh bị hạn chế băng trong khoảng [-B,B], tạp âm là Gaussian và trắng có mật độ phổ công suất (hai biên) $N_0/2$, và đầu vào kênh là quá trình mà thoả mãn công suất vào hạn chế là P. Shammon đã chỉ ra rằng khả năng thông qua của kênh này đơn vị bit/s được cho bởi.

$$C = B\log_2\left(1 + \frac{P}{N_0 B}\right) \quad \text{bit/s} \quad (2.5)$$



Hình 2.2: Kênh tạp âm Gausian trắng cộng AWGN bị hạn chế băng

Đối với kênh AWGN rời rạc theo thời gian có công suất vào hạn chế P và ph- ơng sai tạp âm σ^2, thì khả năng thông qua kênh đơn vị bit/truyền dẫn đ- ợc cho bởi.

$$C = \frac{1}{2}\log_2\left(1+\frac{P}{\sigma^2}\right) \qquad \text{bits/kênh} \qquad (2.6)$$

2.2. Bài tập và ch- ơng trình:

Dựa vào quan hệ hàm biến thông qua các ví dụ trong các tài liệu, khảo sát giải giá trị của biến số tìm đ- ợc giải giá trị của hàm ⇒ thiết kế ch- ơng trình, thực hiện ch- ơng trình trên Matlab.

Mỗi bài tập thể hiện mối quan hệ hàm và biến đ- ợc viết thành các hàm Matlab kết quả khảo sát quan hệ hàm biến đ- ợc thể hiện trực quan bằng đồ thị t- ơng ứng.

Bài tập 1:

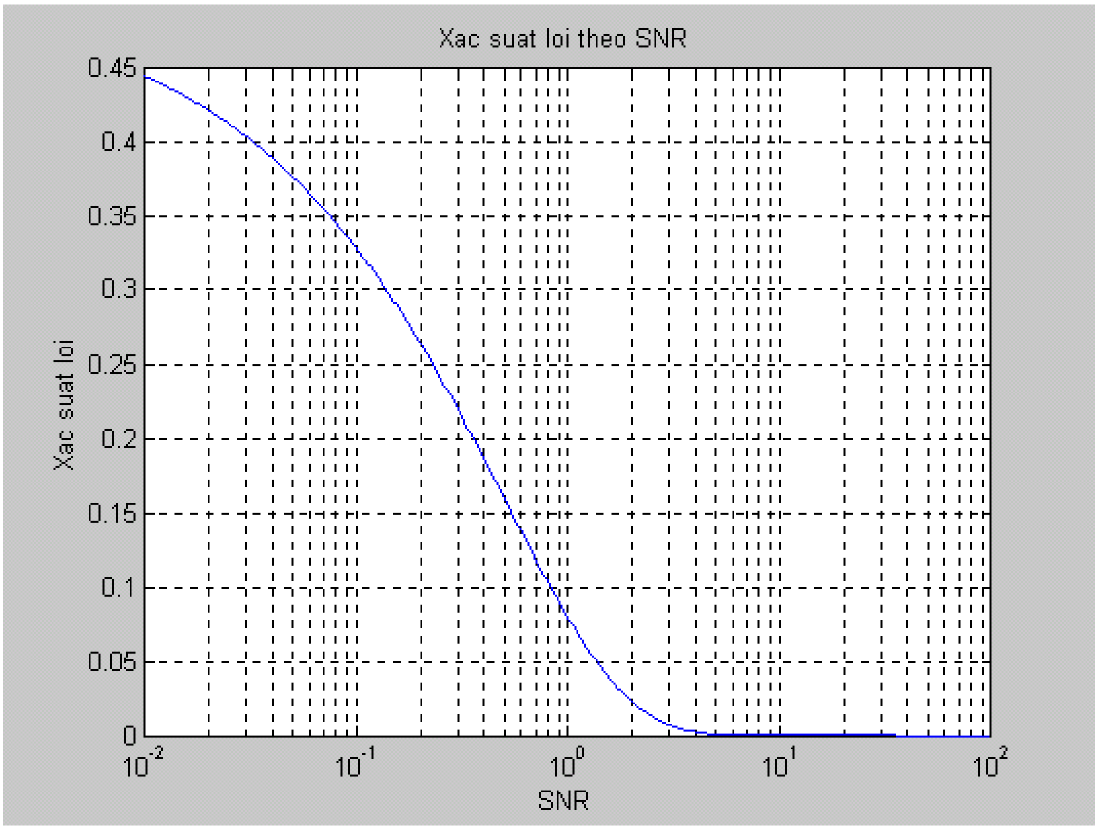
Dung l- ợng kênh BSC

Dữ liệu cơ hai đ- ợc truyền trên kênh AWGN dùng BPSK và giải mã quyết định cứng ở đầu ra dùng tách sóng lọc thích hợp tối - u.

1. Vẽ xác suất lỗi của kênh theo E_b/N_0. Trong đó E_b năng l- ợng của bit trong tín hiệu BPSK và $N_0/2$ là mật độ phổ công suất tạp âm với giả thiết E_b/N_0 thay đổi trong phạm vi − 20dB đến 20dB.
2. Vẽ dung l- ợng của kênh nh- là hàm của E_b/N_0.

Giải:

1. Xác suất lỗi của BPSK với tách sóng tối -u (biên giới quyết định tối -u) đ- ợc cho bởi $p_e = Q\left(\sqrt{2E_b/N_0}\right)$ vẽ t- ơng ứng đ- ợc cho ở hình 2.3 (xem tài liệu Cơ sở kỹ thuật truyền dẫn vi ba số).



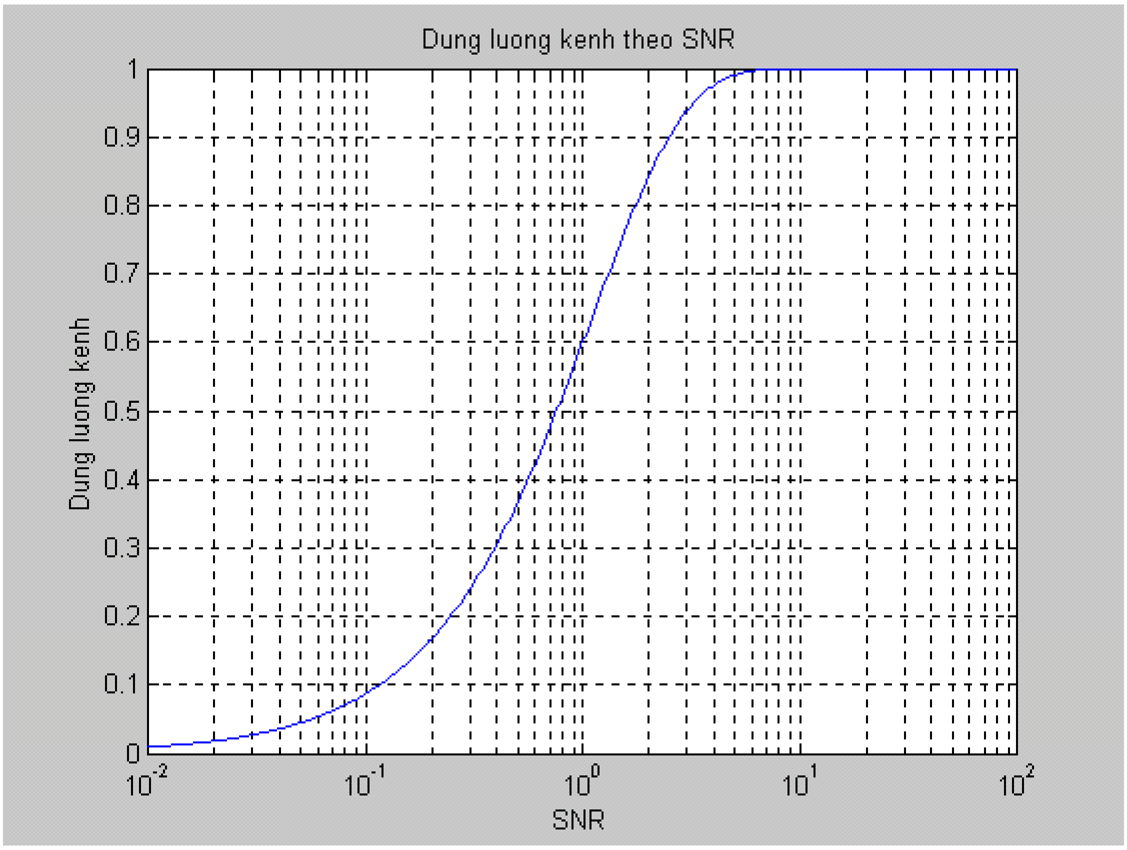
Hình 2.3: Xác suất lỗi của BPSK

2. Ta dùng mối quan hệ.

$$C = 1 - H_b(p)$$
$$= 1 - H_b\left(Q\left(\sqrt{2E_b/N_0}\right)\right)$$

H_b(.) thể hiện cho hàm Entropy cơ hai (xem phụ lục1). Đối số của hàm lúc này là $Q\left(\sqrt{2E_b/N_0}\right) \Rightarrow$ biến số E_b/N_0.

Để vẽ dung l- ợng kênh C theo E_b/N_0. Kết quả đ- ợc cho ở hình 2.4.



Hình 2.4: Dung l- ợng kênh theo E_b/N_0

Từ hai hình vẽ hãy nhận xét các quan hệ hàm biến.

Ch- ơng trình Matlab để thực hiện bài tập này đ- ợc cho ở file: CS81

```
function y=CS81()

gamma_dB = [-20:0.1:20];
gamma =10.^(gamma_dB/10);
P = Q(sqrt(2.*gamma));
if length(find(P<0))~=0
   error('không phải là vecor xác suất, thành phần âm');
end
k=length(P);
for i=1:k
  h = -P.*log2(P)-(1-P).*log2(1-P);
end
 Capacity = 1.- h;
figure(1);
semilogx(gamma,P);
xlabel('SNR');
title('Xac suat loi theo SNR');
ylabel('Xac suat loi');
grid on;
figure(2);
semilogx(gamma,Capacity);
```

```
xlabel('SNR');
title('Dung luong kenh theo SNR');
ylabel('Dung luong kenh');
grid on;
function h = entropy(p)
% H = ENTROPY(P) cho kết quả tính entropy của vector xác suất P.
if length(find(p<0))~=0
  error('không phải là vecor xác suất, thành phần âm');
end
if abs(sum(p)-1)>10e-10
  error('không phải là vecor xác suất, tổng các phần tử lớn hơn 1');
end
  h=sum(-p.*log2(p));
```

Bài tập 2:

Dung l- ợng kênh Gaussian

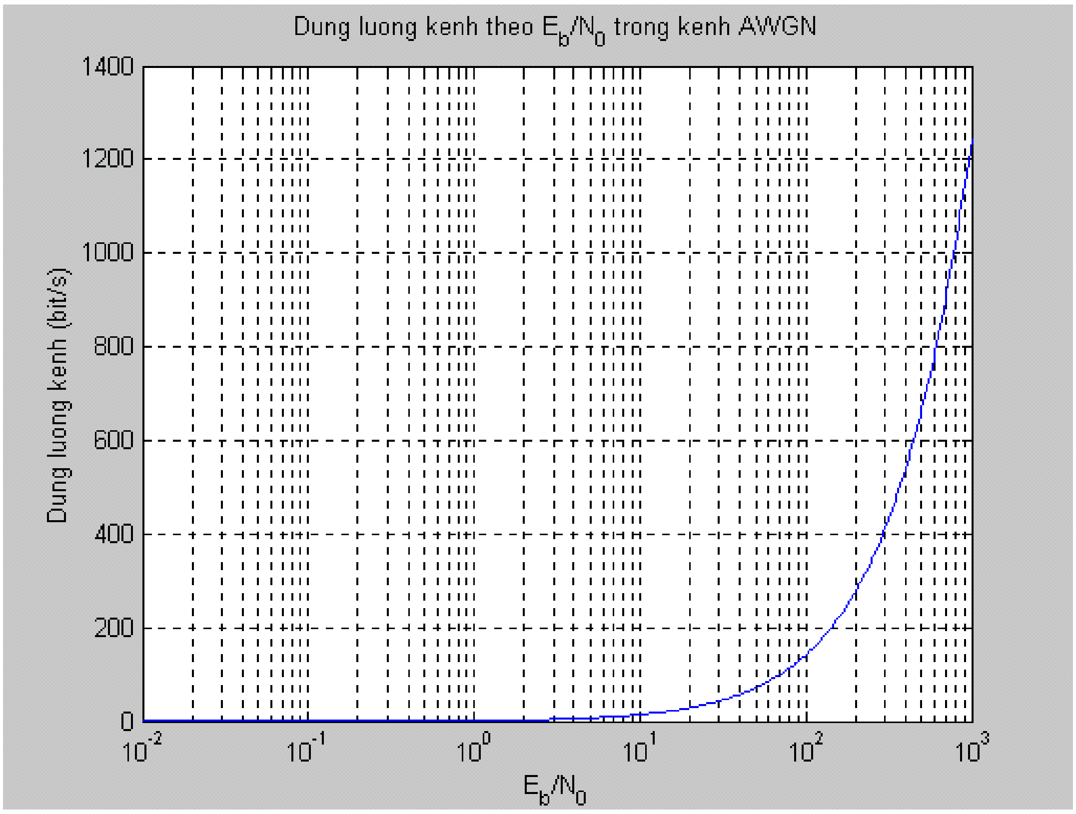
1. Vẽ dung l- ợng kênh AWGN có độ rộng băng B=3000Hz nh- là hàm của P/N_0 (E_b/N_0) khi giá trị của P/N_0 nằm trong khoảng [-20 dB đến 30dB].
2. Vẽ dung l- ợng kênh AWGN có P/N_0 (E_b/N_0) = 25dB nh- là hàm của B. Đặc biệt xét xem xuất hiện điều gì ? khi B tăng đến vô hạn.

Giải:

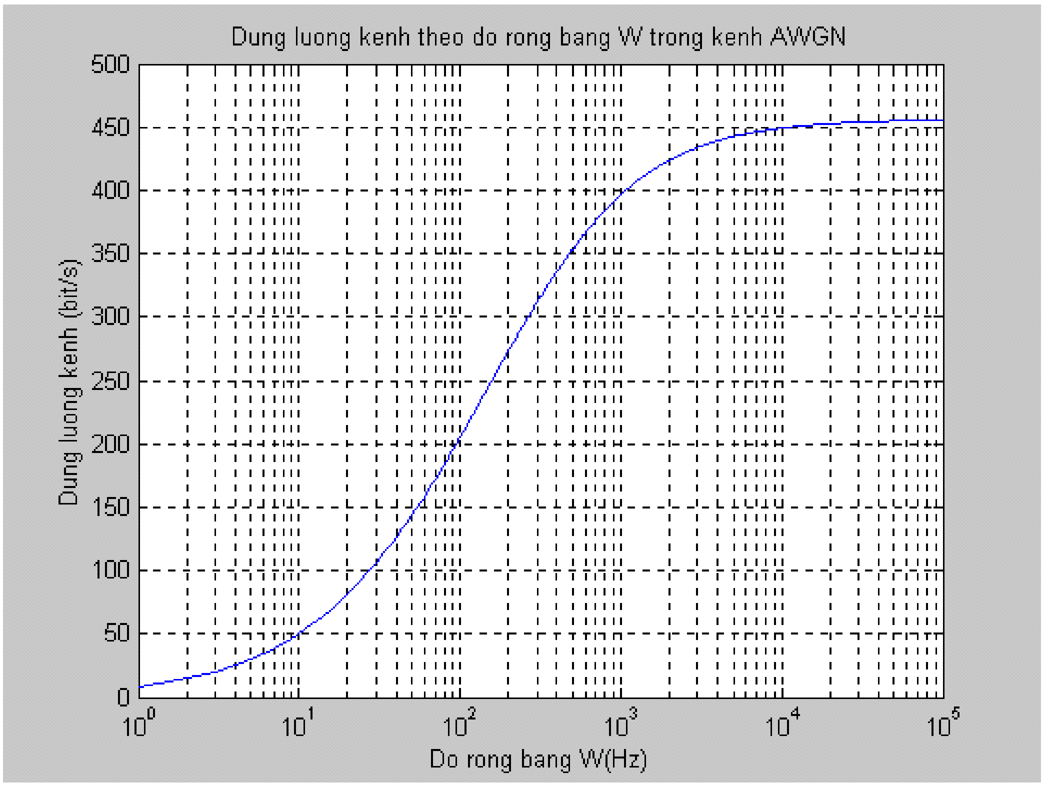
1. Từ công thức.

$$C = B\log_2\left(1 + \frac{P}{N_0 B}\right)$$

Để khảo sát $C = f\left(E_b/N_0\right)$ cố định độ rộng băng thông B của kênh. Kết quả chạy ch- ơng trình đ- ợc cho hình 2.5.



Hình 2.5: Dung l- ợng kênh AWGN có W=3000Hz nh- hàm của E_b/N_0



Hình 2.6: Dung l- ợng kênh AWGN có E_b/N_0=25 dB theo (hàm của) B

2. Kết quả xét dung l- ợng kênh nh- hàm của độ rộng băng đ- ợc cho ở hình 2.6. Để khảo sát $C = f(B)$ thì cố định giá trị E_b/N_0

Nhận xét:

Ta thấy, d- ờng nh- khi tỉ số tín hiệu trên tạp âm P/N_0 (SNR, E_b/N_0) hoặc độ rộng băng của kênh B tiến đến không, thì dung l- ợng của kênh cũng tiến đến không.

Tuy nhiên, khi P/N_0 hoặc B tiến đến vô hạn, thì dung l- ợng kênh xem xét lại khác.

$\Rightarrow$ Khi P/N_0 tiến đến vô hạn thì dung l- ợng kênh cũng tiến đến vô hạn nh- đ- ợc thấy ở hình 2.5.

$\Rightarrow$ Khi B tiến đến vô hạn thì dung l- ợng kênh tiến đến giới hạn nào đó, nó đ- ợc xác định bởi P/N_0. Để xác định giá trị giới hạn này, ta có.

$$\lim_{W\to\infty} W \log_2\left(1+\frac{P}{N_0W}\right) = \frac{P}{N_0 \ln 2}$$
$$= 1{,}4427\frac{P}{N_0}$$

Ch- ơng trình Matlab thực hiện bài tập này đ- ợc cho ở File: CS82

```
function y=CS82()

disp('Nen nhap E_b/N_0 = [-20:0.1:30], va BW = 3000Hz');

SNR_dB=input('Ban hay nhap vector E_b/N_0 = ');

SNR = 10.^(SNR_dB/10);

BW = input('Ban hay nhap Bandwidth = ');

Capacity = BW.*log2(1 + SNR/BW);

figure(1);
semilogx(SNR,Capacity);
title('Dung luong kenh theo E_b/N_0 trong kenh AWGN');
xlabel('E_b/N_0');
ylabel('Dung luong kenh (bit/s)');
grid on;

W=[1:10, 12:2:100, 105:5:500, 510:10:5000, 5025:25:20000, 20050:50:100000];
SNR_dB = 25;
SNR=10.^(SNR_dB/10);

Capacity = W.*log2(1 + SNR./W);

figure(2);
semilogx(W,Capacity);
title('Dung luong kenh theo do rong bang W trong kenh AWGN');
```

```
xlabel('Do rong bang W(Hz)');
ylabel('Dung luong kenh (bit/s)');
grid on;
```

Bài tập 3:

Dung l- ợng của kênh AWGN đầu vào cơ hai

Kênh AWGN đầu vào cơ hai đ- ợc mô hình hoá bởi hai mức tín hiệu vào cơ hai ± A và tạp âm Gaussian trung bình không có ph- ơng sai σ^2. Trong tr- ờng hợp này, $X = \{-A, A\}, Y = \Re, p(y|X = A) \sim N(A, \sigma^2)$ và $p(y|X = -A) \sim N(-A, \sigma^2)$.

Vẽ dung l- ợng của kênh này theo (hàm của) A/σ.

Giải:

Do đối xứng nên dung l- ợng kênh đạt đ- ợc từ phân phối đầu vào đồng đều - nghĩa là, có $p(X = A) = p(X = -A) = 1/2$. Vì phân phối đầu vào nh- vậy, nên phân phối đầu ra đ- ợc cho bởi.

$$p(y) = \frac{1}{2\sqrt{2\pi\sigma^2}} e^{-(y+A)^2/2\sigma^2} + \frac{1}{2\sqrt{2\pi\sigma^2}} e^{-(y-A)^2/2\sigma^2}$$

và thông tin chéo giữa các đầu vào ra đ- ợc cho bởi.

$$\begin{aligned} I(X;Y) = {} & \frac{1}{2}\int_{-\infty}^{\infty} p(y|X = A)\log_2 \frac{p(y|X = A)}{p(y)} dy \\ & + \frac{1}{2}\int_{-\infty}^{\infty} p(y|X = -A)\log_2 \frac{p(y|X = -A)}{p(y)} dy \end{aligned}$$

Thực hiện lấy tích phân và chuyển biến cho kết quả.

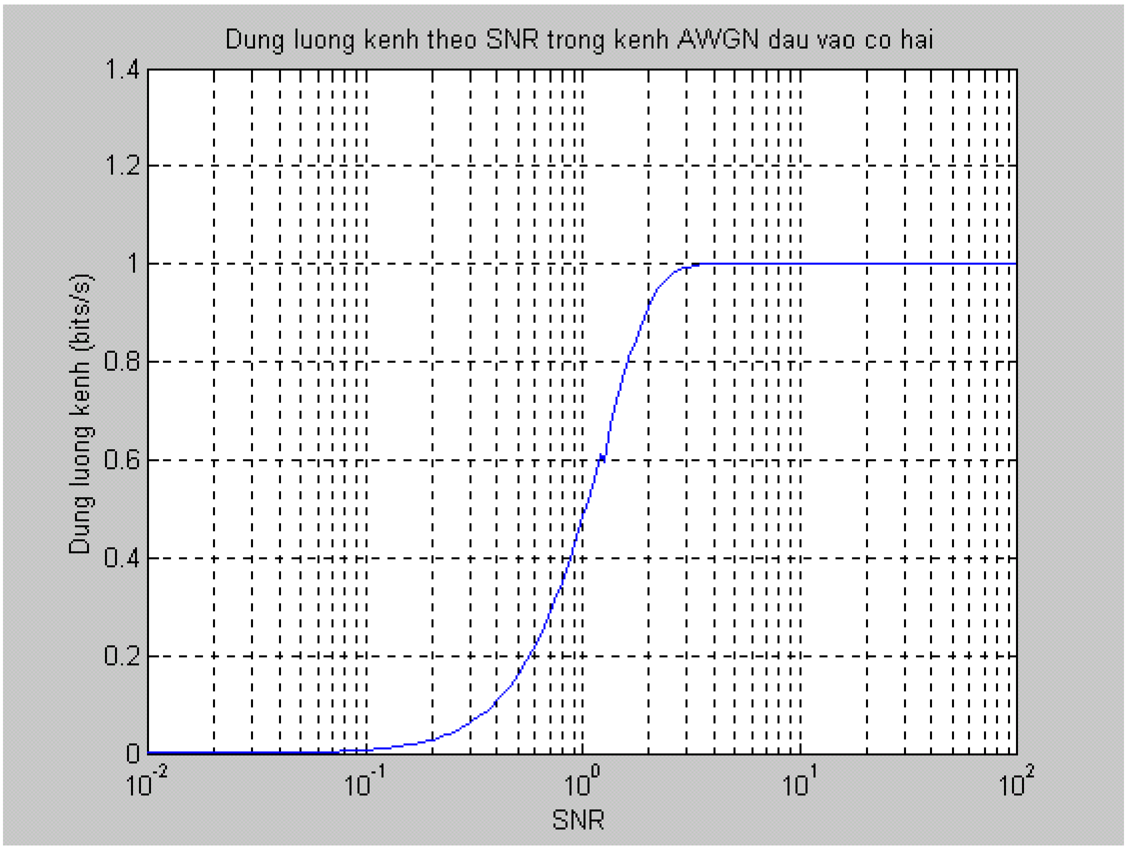
$$I(X;Y) = \frac{1}{2} f\left(\frac{A}{\sigma}\right) + \frac{1}{2} f\left(-\frac{A}{\sigma}\right)$$

Trong đó

$$f(a) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(u-a)^2/2} \log_2 \frac{2}{1 + e^{-2au}} du$$

Dùng quan hệ này để tính I(X;Y) cho các giá trị A/σ khác nhau và vẽ đồ thị quan hệ giữa chúng. Kết quả vẽ đ- ợc cho ở hình 2.7.

Ch- ơng trình Matlab thực hiện bài tập này đ- ợc cho ở File: CS83.



Hình 2.7: Dung l- ợng kênh AWGN đầu vào cơ hai nh- hàm của SNR = A/σ.

Ch- ơng trình Matlab thực hiện bài tập này đ- ợc cho ở File: CS83

```
function y=CS83()

a_dB=[-20:0.2:20];
a=10.^(a_dB/10);

for i=1:length(a)
  f(i)=quad(@Quocanh , a(i)-5 , a(i)+5 , 1e-3 , [] , a(i) );
  g(i)=quad(@Quocanh , -a(i)-5 , -a(i)+5 , 1e-3 , [] , -a(i) );
  c(i)=0.5*f(i)+0.5*g(i);
end

semilogx(a,c);
title('Dung luong kenh theo SNR trong kenh AWGN dau vao co hai');
xlabel('SNR');
ylabel('Dung luong kenh (bits/s)');
grid on;

function y = Quocanh(u,a)
A = 1./sqrt(2*pi).*exp((-(u-a).^2)/2);
B = log2(2./(1 + exp(-2*a*u)));
y = A.*B;
```

Bài tập 4:

[So sánh ph- ơng pháp quyết định cứng và quyết định mềm]

Kênh đầu vào cơ hai dùng hai mức ±A. Đầu ra của kênh là tổng của tín hiệu vào và tạp âm AWGN có trung bình không và ph- ơng sai σ^2. Kênh này đ- ợc dùng trong hai tr- ờng hợp.

Tr- ờng hợp1: Dùng đầu ra trực tiếp mà không thực hiện l- ợng tử hoá (quyết định cứng).

Tr- ờng hợp2: Quyết định tối - u đ- ợc thực hiện trên mỗi mức đầu vào (quyết định mền).

Vẽ dung l- ợng kênh theo (hàm của) A/σ trong mỗi tr- ờng hợp.

Giải:

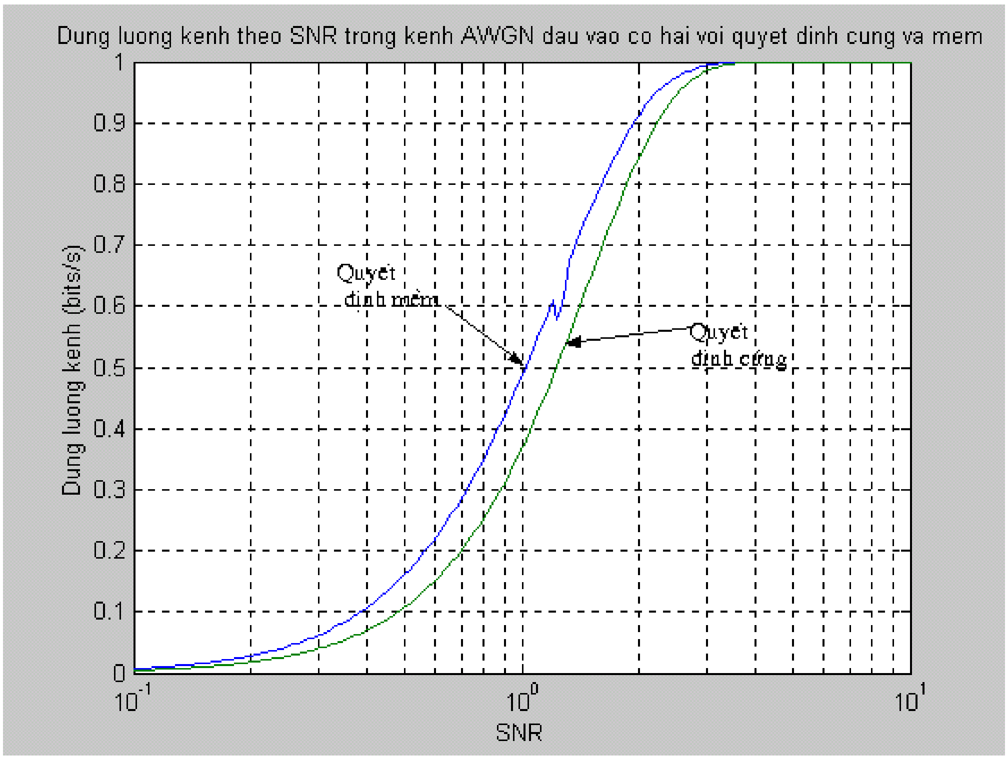
Tr- ờng hợp quyết định mềm t- ơng tự nh- bài tập 3.

Tr- ờng hợp quyết định cứng, xác suất chéo của kênh BSC là Q(A/σ). Vì vậy, dung l- ợng kênh đ- ợc cho bởi.

$$C_H = 1 - H_b\left(Q\left(\frac{A}{\sigma}\right)\right)$$

Cả C_H và C_S đều đ- ợc cho ở hình 2.8. Đầu ra giải mã quyết định mềm thực hiện tốt hơn giải mã quyết định cứng tại tất cả các giá trị A/σ, nh- đ- ợc thấy.

Ch- ơng trình Matlab thực hiện bài tập này đ- ợc cho ở File: CS84.



Hình 2.8: Dung l- ợng kênh quyết định cứng C_H và quyết định mềm C_S theo SNR =A/σ

Ch- ơng trình Matlab thực hiện bài tập này đ- ợc cho ở File: CS84

```
function y=CS84()

a_dB=[-13:0.1:13];
a=10.^(a_dB/10);

p = Q(a);

k=length(p);
for i=1:k
  h = -p.*log2(p)-(1-p).*log2(1-p);
end

 c_hard = 1.- h;

for i=1:length(a)

   f(i) = quad(@Quocanh , a(i)-5 , a(i)+5 , 1e-3 , [] , a(i) );

   g(i) = quad(@Quocanh , -a(i)-5 , -a(i)+5 , 1e-3 , [] , -a(i) );

   c_soft(i) = 0.5*f(i) + 0.5*g(i);

end

semilogx(a,c_soft,a,c_hard);

title('Dung luong kenh theo SNR trong kenh AWGN dau vao co hai voi quyet dinh cung va mem');
xlabel('SNR');
ylabel('Dung luong kenh (bits/s)');
axis([0.1 10 0 1]);
grid on;

function y = Quocanh(u,a)
A = 1./sqrt(2*pi).*exp((-(u-a).^2)/2);
B = log2(2./(1 + exp(-2*a*u)));
y = A.*B;
```

Bài tập 5:

[Dung l- ợng kênh theo độ rộng băng và SNR]

Dung l- ợng của kênh AWGN hạn chế băng có công suất không đổi P và độ rộng băng B đ- ợc cho bởi.

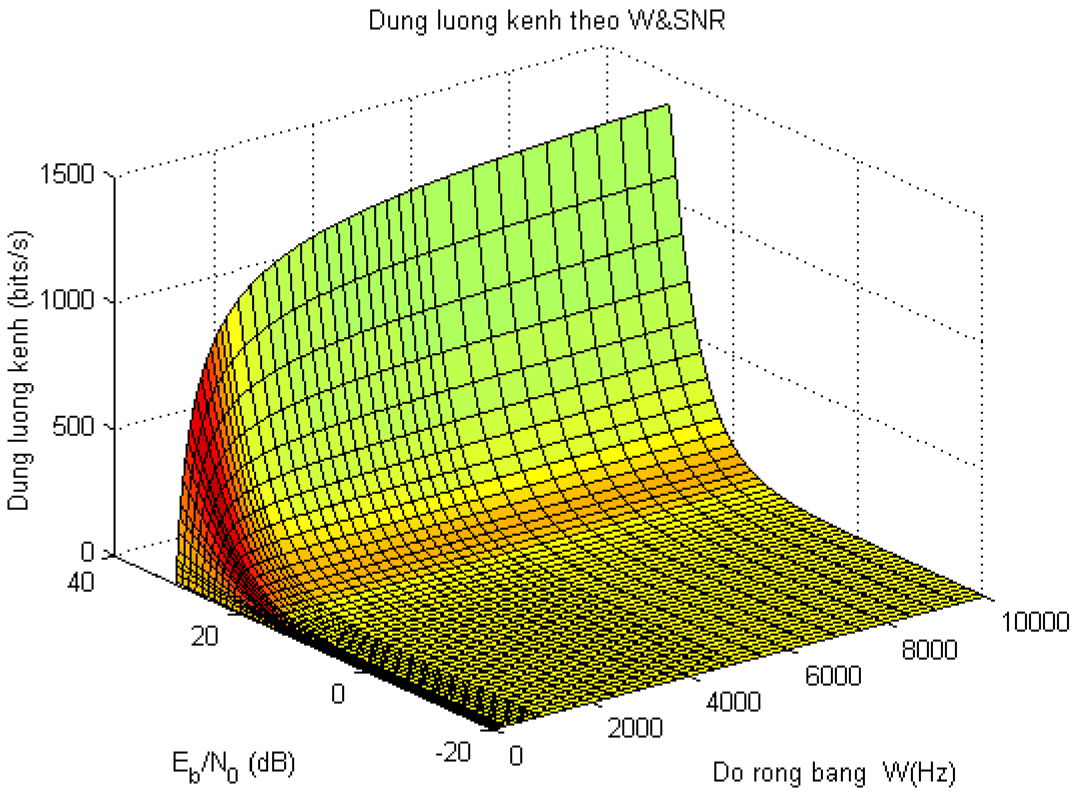
$$C = B\log_2\left(1+\frac{P}{N_0 B}\right)$$

Vẽ dung l- ợng kênh nh- hàm của cả hai thông số B và SNR (hay P/N_0).

Giải:

Kết quả vẽ đ- ợc cho ở hình 2.9. l- u ý rằng, khi SNR không đổi thì việc vẽ chuyển thành hình 2.6. Khi B không đổi thì việc vẽ chuyển thành hình 2.5.

Ch- ơng trình Matlab thực hiện bài tập này đ- ợc cho ở File: CS85.



Hình 2.9: Dung l- ợng kênh nh- hàm của hai thông số độ rộng băng B và tỉ số tín hiệu trên tạp âm SNR trong kênh AWGN

Ch- ơng trình Matlab thực hiện bài tập này đ- ợc cho ở File: CS85

function y = CS85
 	
w=[1:5:20, 25:20:100, 130:50:300, 400:100:1000, 1250:250:5000, 5500:500:10000];
 	
SNR_dB = [-20:1:30];
 	
SNR = 10.^(SNR_dB/10);
 	
for i=1:45
  for j=1:51,
    c(i,j) = w(i) * log2(1 + SNR(j) / w(i));

```
  end
end

k=[0.9, 0.8, 0.5, 0.6];
s=[-70, 35];
surfl(w,SNR_dB,c',s,k);
ylabel('E_b/N_0 (dB)');
xlabel('Do rong bang  W(Hz)')
zlabel('Dung luong kenh (bits/s)')
title('Dung luong kenh theo W&SNR');
```

Bài tập 6:

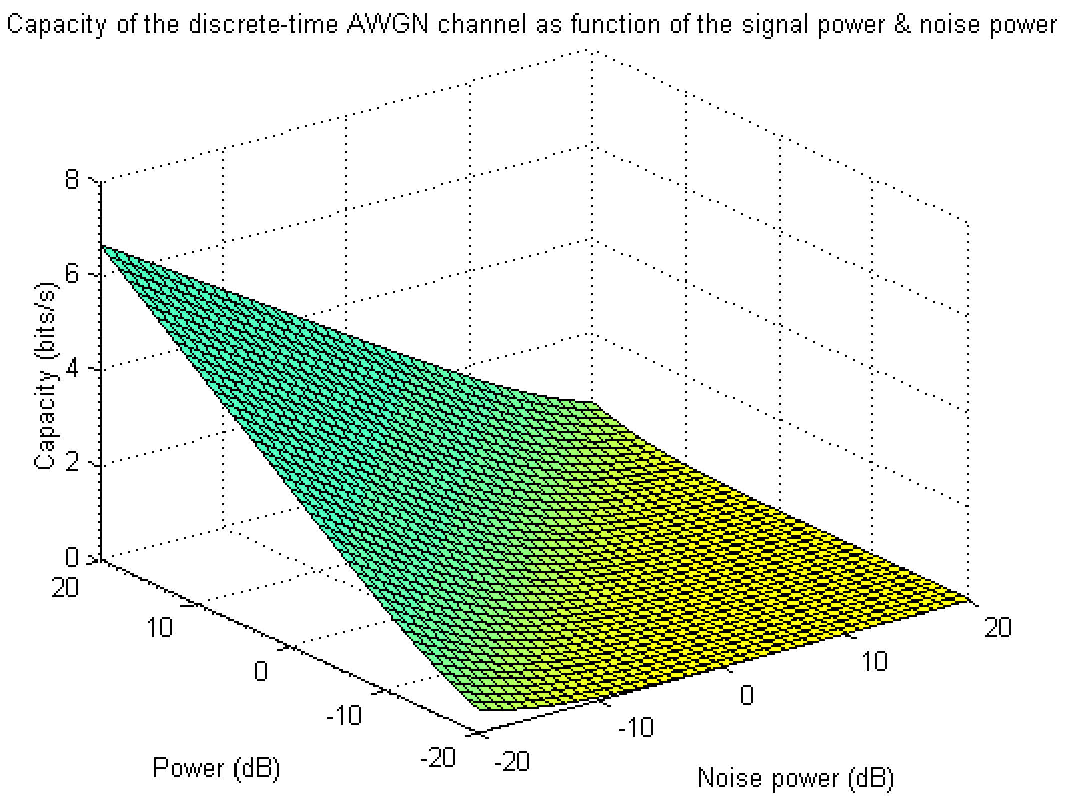
Dung l- ợng kênh AWGN rời rạc

Hãy vẽ dung l- ợng kênh AWGN rời rạc nh- là hàm của công suất đầu vào và ph- ơng sai tạp âm.

Giải:

Kết quả vẽ đ- ợc cho ở hình 2.10.

Ch- ơng trình Matlab thực hiện bài tập này đ- ợc cho ở File: CS86.



Hình 2.10: Dung l- ợng kênh AWGN rời rạc nh- hàm của công suất tín hiệu (P) và công suất tạp âm (σ^2)

Ch- ơng trình Matlab thực hiện bài tập này đ- ợc cho ở File: CS86

```
function y = CS86

p_dB=-20:1:20;
np_dB=p_dB;
p=10.^(p_dB/10);
np=p;

for i=1:41,
  for j=1:41,
    c(i,j)=0.5*log2(1+p(j)/np(i));
    echo off;
  end
end
echo on;

k=[0.9, 0.8, 0.5, 0.6];
s=[-70, 35];

surfl(np_dB,p_dB,c',s,k);

ylabel('Power (dB)');
xlabel('Noise power (dB)')
zlabel('Capacity (bits/s)')
title('Capacity of the discrete-time AWGN channel as function of the signal power & noise power');
```

Phu luc

KHẢ NĂNG THÔNG QUA HAY DUNG L□ỢNG KÊNH

(Dịch ch- ơng II tài liệu Digital Communication tác giả Simon Haykin 1988)

2.1. TH□NG TIN, □□ B□T □□NH, VÀ ENTROPY

Xét nguồn tin rời rạc đ- ợc xác định bởi.

$$\Im = \{s_0, s_1, \ldots\ldots, s_{K-1}\} \qquad (2.1)$$

Với xác suất.

$$P(S = s_k) = P_k \qquad k = 0,1,2,\ldots,K-1 \qquad (2.2)$$

Tất nhiên tập các xác suất này phải thoả mãn điều kiện.

$$\sum_{k=0}^{K-1} P_k = 1 \qquad (2.3)$$

Giả thiết các ký hiệu đ- ợc phát từ nguồn trong các khoảng thời gian tín hiệu liên tiếp là ***độc lập thống kê***. Nguồn có các thuộc tính vừa đ- ợc mô tả đ- ợc gọi là nguồn ***không nhớ rời rạc***, không nhớ có nghĩa là ký hiệu đ- ợc phát đi ở thời điểm nào đó là độc lập với tr- ớc đó.

⇒Làm thế nào đánh giá l- ợng tin có trong nguồn tin đó? ý t- ởng về thông tin đó liên quan mật thiết với độ bất định "Uncertainty" hay "sự bất ngờ Surprise" đ- ợc đề cập sau đây.

✓ Xét sự kiện $S = s_k$, thể hiện việc phát ký hiệu s_k từ nguồn tin với xác suất p_k nh- đ- ợc định nghĩa ở ph- ơng trình (2.2). Rõ ràng

+ Nếu xác suất $p_k = 1$ và $p_i = 0$ với $\forall i \neq k$, thì sẽ không có 'sự ngạc nhiên surprise' và không có 'thông tin Information' khi ký hiệu s_k đ- ợc phát.

+ Nếu xác suất xuất hiện của ký hiệu (sự kiện) từ nguồn tin p_k càng nhỏ thì l- ợng tin chứa trong đó càng lớn và ng- ợc lại.

⇒ Vì vậy "độ bất định Uncertainty", "thông tin - Information", "sự bất ngờ - Surprise" tất cả đ- ợc liên quan với nhau. Ta thấy.

+ Tr- ớc khi sự kiện $S = s_k$ xảy ra, thì có một l- ợng bất định Uncertainty (l- ợng tin tiên nghiệm từ nguồn tin ⇔ l- ợng tin có trong nguồn tin).

+ Khi sự kiện $S = s_k$ xẩy ra có một l- ợng bất ngờ Surprise.

+ Sau khi sự kiện $S = s_k$ xẩy ra, nhận đ- ợc một l- ợng tin. Rõ ràng cả ba l- ợng này nh- nhau.

⇒L- ợng tin tỷ lệ nghịch với xác suất xuất hiện.

ĐINH NGHĨA: *(l- ợng tin riêng của một sự kiện trong tập các sự kiện)*

L- ợng tin nhận đ- ợc sau khi quan sát sự kiện S = s_k, xảy ra với xác suất p_k, là hàm logarithmic.

$$I(s_k) = \log\left(\frac{1}{p_k}\right) \qquad (2.4)$$
$$= -\log p_k$$

<u>THUỘC TÍNH</u> Từ định nghĩa bộc lộ các thuộc tính sau (Property).

1. $I(s_k) = 0$ với $p_k = 1$ (2.5)

⇔ Hiển nhiên, nếu biết chắc chắn về kết cục của sự kiện, kể cả khi tr- ớc khi nó xẩy ra, thì không nhận đ- ợc thông tin gì cả.

2. $I(s_k) \geq 0$ với $0 \leq p_k \leq 1$ (2.6).

⇔ Sự xuất hiện sự kiện S=s_k cho ta thông tin hoặc không nh- ng ***không bao giờ gây ra mất thông tin***.

3. $I(s_k) \geq I(s_i)$ với $p_k < p_i$ (2.7).

⇔ Sự kiện xẩy ra có xác suất càng nhỏ, thì l- ợng tin nhận đ- ợc càng lớn.

4. $I(s_k s_l) = I(s_k) + I(s_l)$ nếu s_k và s_l dộc lập thống kê

Cơ số của hàm logarit trong ph- ơng trình (2.4) là hoàn toàn tuỳ ý. Tuy nhiên, ngày nay chuẩn theo cơ số 2. Đơn vị thông tin đ- ợc gọi là bit. Vì vậy ta viết.

$$I(s_k) = \log_2\left(\frac{1}{p_k}\right) \qquad k = 0,1,......K-1 \qquad (2.8).$$

Khi p_k=1/2, thì ta có I(s_k) = 1 bit.

⇒ **Định nghĩa bit:** Một bit là l- ợng tin mà ta nhận đ- ợc khi một trong hai sự kiện có thể có và đồng xác suất xẩy ra. *One bit is the amount of information that we gain when one of two possible & equally likely (i.e., equiprobable) events occurs.*

L- u ý, bit cũng liên quan đến số nhị phân. Trong tài liệu, ta dùng thuật ngữ "bit" là đơn vị thông tin khi liên hệ nội dung thông tin của nguồn tin hoặc đầu ra kênh kênh ***và*** là từ cấu tạo đầu cho số nhị phân khi liên hệ với chuỗi các số 0 và 1.

L- ợng tin I(s_k) đ- ợc tạo ra bởi nguồn tin trong khoảng thời gian tín hiệu nào đó phụ thuộc vào ký hiệu s_k đ- ợc phát đi bởi nguồn đó tại thời điểm đó. Thực vậy, I(s_k) là biến ngẫu nhiên rời rạc mà nhận các giá trị I(s_0), I(s_1),..., I(s_{K-1}) với các xác suất t- ơng ứng $p_0,p_1,p_2,.....p_{K-1}$. Giá trị trung bình của I(s_k) trên nguồn tin $\mathfrak{I}$ đ- ợc cho bởi.

$$\begin{aligned} H(\Im) &= E[I(s_k)] \\ &= \sum_{k=0}^{K-1} p_k I(s_k) \\ &= \sum_{k=0}^{K-1} p_k \log\left(\frac{1}{p_k}\right) \\ &= -\sum_{k=0}^{K-1} p_k \log p_k \end{aligned} \qquad (2.9)$$

H(ℑ) đ- ợc gọi Entropy của nguồn không nhớ rời rạc với ℑ. Nó đánh giá nội dung thông tin trung bình trên ký hiệu nguồn tin. L- u ý rằng Entropy H(ℑ) chỉ phụ thuộc vào xác suất của các ký hiệu trong bảng mẫu tự ℑ của nguồn tin. Vì vậy ℑ trong H(ℑ) không phải là đối số của hàm mà chỉ là nhãn cho nguồn tin.

⇒ **<u>ĐINH NGHĨA ENTROPY</u>**: Entropy của nguồn rời rạc đ- ợc xác định bởi ph- ơng trình 2.9 là trung bình thống kê của l- ợng thông tin riêng của các tin thuộc nguồn.

❖ **<u>(1) CÁC THUỘC TÍNH CỦA ENTROPY.</u>**

Xét nguồn rời rạc không nhớ mà mô hình toán học của nó đ- ợc định nghĩa bởi ph- ơng trình 2.1 và 2.2. Entropy H(ℑ) của nguồn tin đ- ợc giới hạn nh- sau.

$$0 \le H(\Im) \le \log_2 K \qquad (2.10).$$

Trong đó: K là số các ký hiệu có trong tập ℑ của nguồn tin. Hơn nữa, ta có thể phát biểu.

1. **H(ℑ) = 0** nếu và chỉ nếu xác suất p_k=1 với *một số k*, và tất cả các xác suất còn lại trong tập đều bằng không. Giới hạn d- ới Entropy t- ơng ứng với sự bất định không có (*không có tin trong nguồn tin*).

<u>Chứng minh</u>: Giới hạn d- ới- Lower Bound

Vì $p_k \le 1$ nên.

+ Nếu mỗi xác suất $p_k < 1$, ⇒ mỗi thành phần $p_k \log_2\left(1/p_k\right)$ trong ph- ơng trình (2.9) là số không âm (d- ơng), ⇒ H(ℑ) > 0.

+ Nếu p_k=1 hay p_k = 0 (nghĩa là p_k=1 với một số k thì tất cả các xác suất còn lại đều bằng 0) ⇒ $p_k \log_2\left(1/p_k\right) = 0$ ⇔ H(ℑ) = 0

⇒ Kết hợp lại ta đ- ợc H(ℑ) ≥ 0 (ĐPCM)

2. **H(ℑ) = log₂K**, nếu và chỉ nếu $p_k = 1/K$ với $\forall k$ (nghĩa là tất cả các ký hiệu trong ℑ là đồng xác suất suy ra từ 2.9 và 2.3). Giới hạn trên về Entropy này t- ơng đ- ơng với sự bất định lớn nhất (*l- ợng tin có trong nguồn tin lớn nhất*)

<u>Chứng minh</u> Giới hạn trên - Upper Bound

Từ tính chất hàm logarithm ta có

$$\ln x \le (x-1) \quad \text{với } x \ge 0 \qquad (2.11).$$

Có thể kiểm tra bất đẳng thức này bằng cách vẽ hàm **lnx** và **(x-1)** theo x, đ- ợc cho ở hình 2.1. ở đây ta thấy rằng đ- ờng thẳng (x-1) luôn nằm trên đ- ờng cong $y = \ln x$. Dấu bằng chỉ xẩy ra tại điểm x=1, tại đây đ- ờng thẳng (x-1) là đ- ờng tiếp tuyến với đ- ờng cong $y = \ln x$.

Tiếp theo chứng minh, tr- ớc hết xét hai phân phối xác suất nào đó $\{p_0, p_1, ..., p_{K-1}\}$ và $\{q_0, q_1, ..., q_{K-1}\}$ trên tập ký hiệu $\Im = \{s_1, s_2, ..., s_K\}$ của nguồn tin không nhớ rời rạc. Ta có thể viết.

$$\sum_{k=0}^{K-1} p_k \log_2\left(\frac{q_k}{p_k}\right) = \frac{1}{\log_2 e}\sum_{k=0}^{K-1} p_k \ln\left(\frac{q_k}{p_k}\right)$$

Trong đó **e** là cơ số của logarit tự nhiên. Vì vậy dùng bất đẳng thức (2.11), ta đ- ợc.

$$\begin{aligned}
\sum_{k=0}^{K-1} p_k \log_2\underbrace{\left(\frac{q_k}{p_k}\right)}_{\Updownarrow \atop x} &\le \frac{1}{\log_2 e}\sum_{k=0}^{K-1} p_k \underbrace{\left(\frac{q_k}{p_k}-1\right)}_{\Updownarrow \atop x-1} \\
&\le \frac{1}{\log_2 e}\sum_{k=0}^{K-1}(q_k - p_k) \\
&\le \frac{1}{\log_2 e}\left(\sum_{k=0}^{K-1} q_k - \sum_{k=0}^{K-1} p_k\right) = 0
\end{aligned}$$

Vậy ta đ- ợc.

$$\sum_{k=0}^{K-1} p_k \log_2\left(\frac{q_k}{p_k}\right) \le 0 \qquad (2.12).$$

Dấu bằng chỉ xảy ra nếu $q_k = p_k$ với $\forall k$.

Giả sử đặt

$$q_k = \frac{1}{K} \qquad k = 0,1,.....,K-1 \qquad (2.13)$$

T- ơng đ- ơng với $\Im$ có các ký hiệu đồng xác suất. Entropy của nguồn không nhớ rời rạc có đặc tính nh- vậy bằng.

$$H(\Im) = \sum_{k=0}^{K-1} q_k \log_2\left(\frac{1}{q_k}\right) = \log_2 K \qquad (2.14)$$

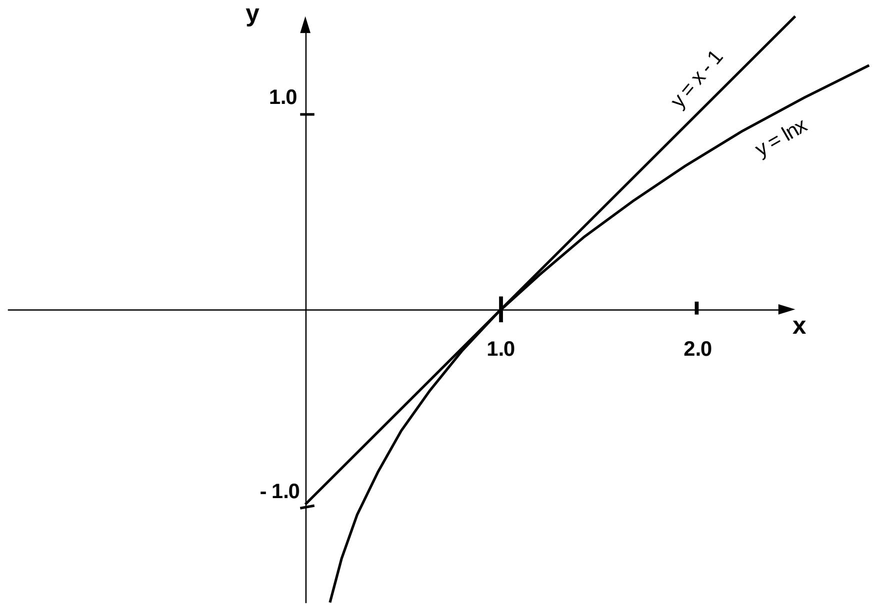
$\Rightarrow$ Dùng ph- ơng trình 2.13 vào 2.12 ta đ- ợc.

$$\sum_{k=0}^{K-1} p_k \log_2\left(\frac{1}{p_k}\right) \le \log_2 K$$

Một cách t- ơng đ- ơng, Entropy của nguồn không nhớ rời rạc có phân phối xác suất tuỳ ý đối với các ký hiệu thuộc bảng mẫu tự $\Im$ của nó đ- ợc giới hạn là.

$$H(\Im) \le \log_2 K$$

Vì vậy, $H(\Im)$ luôn nhỏ hơn hoặc bằng $\log_2 K$. Dấu bằng chỉ xảy ra nếu các ký hiệu trong $\Im$ đồng xác xuất nh- ph- ơng trình 2.13$\Rightarrow$ Vậy công thức 2.11 và 2.13 đ- ợc chứng minh.



Hình2.1: Vẽ hàm y=x-1 và y=lnx

<u>Ví du1:</u> Entropy của nguồn không nhớ cơ hai

Entropy of Binary Memoryless Source.

Để minh hoạ thuộc tính của entropy $H(\Im)$, xét nguồn tin cơ hai trong đó ký hiệu 0 xuất hiện với xác suất p_0 và ký hiệu 1 xuất hiện với xác suất $p_1 = 1 - p_0$. G*iả thiết nguồn không nhớ để các ký hiệu liên tiếp đ- ợc phát đi là độc lập thống kê nhau*.

Entropy của nguồn bằng.

$$\begin{aligned} H(\Im) &= \sum_{k=0}^{1} \mathrm{p_k} \log_2\left(\frac{1}{\mathrm{p_k}}\right) \qquad \text{do } \mathrm{K} = 2 \\ &= -p_0 \log_2 p_0 - p_1 \log_2 p_1 \qquad (2.15) \\ &= -p_0 \log_2 p_0 - (1-p_0)\log_2(1-p_0) \qquad \text{bits} \end{aligned}$$

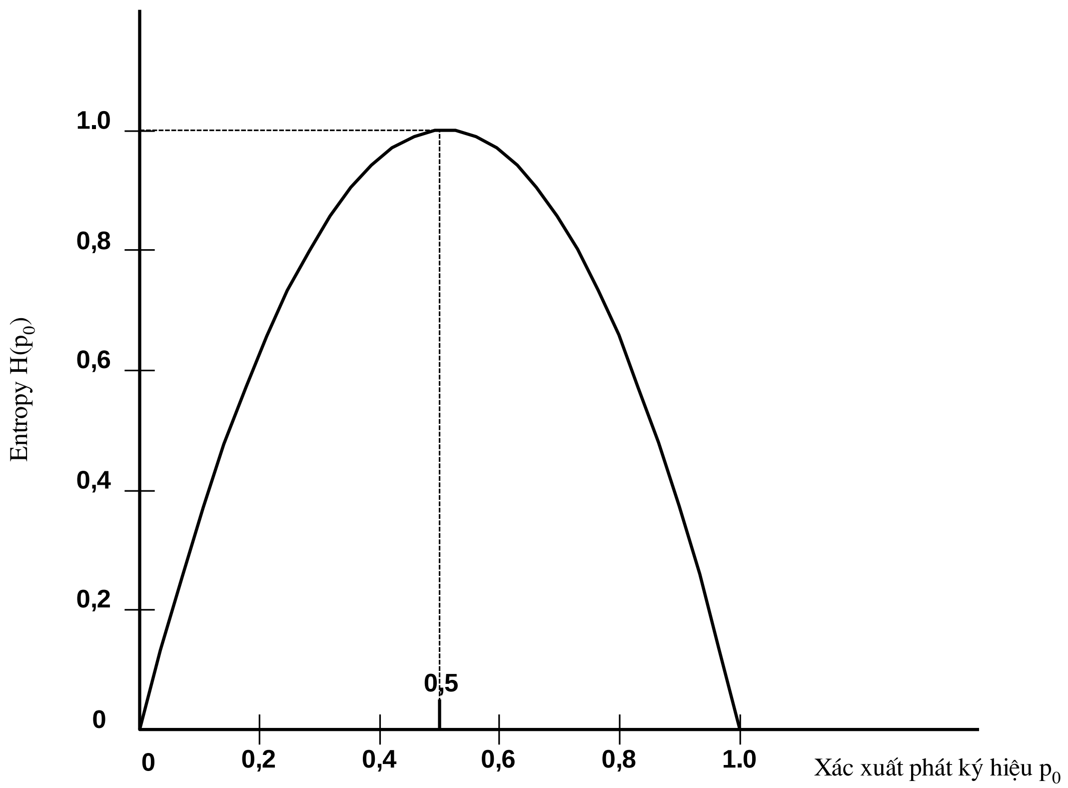
Ta l- u ý rằng:

1. Khi $p_0 = 0$ thì $H(\Im) = 0$. Điều này suy ra từ thực tế là $\lim_{x \to 0} (x \log x) = 0$.
2. Khi $p_0 = 1$ thì $H(\Im) = 0$.
3. Entropy $H(\Im)$ đạt giá trị lớn nhất ($H_{max} = 1$ bit), khi $p_0 = p_1 = 0,5$ nghĩa là xác xuất phát bít 0 và bit 1 nh- nhau và bằng 0,5.

Hàm của p_0 đ- ợc cho ở ph- ơng trình 2.15 th- ờng đ- ợc gặp phải trong các bài tập về lý thuyết thông tin. Vì vậy th- ờng ấn định các ký hiệu cụ thể cho hàm này. Cụ thể ta định nghĩa.

$$H'(p_0) = -p_0 \log_2 p_0 - (1 - p_0) \log_2 (1 - p_0) \qquad (2.16)$$

Ta coi $H'(p_0)$ là hàm Entropy. Nên cẩn thận trong việc phân biệt giữa các ph- ơng trình 2.15 & 2.16. $H(\Im)$ trong 2.15 cho biết entropy của nguồn không nhớ rời rạc có bảng mẫu tự nguồn $\Im$. Mặt khác, $H'(p_0)$ trong ph- ơng trình 2.16 là một hàm của xác suất tiên nghiệm p_0 đ- ợc xác định trên khoảng [0,1]. Theo đó ta có thể vẽ hàm entropy $H'(p_0)$ theo p_0, đ- ợc xác định trên khoảng [0,1] nh- hình 2.2. Rõ ràng đ- ờng cong trong hình 2.2 nêu bật đ- ợc các vấn đề đ- ợc l- u ý 1,2,3.



Hình 2.2: Hàm Entropy $H'(p_0)$

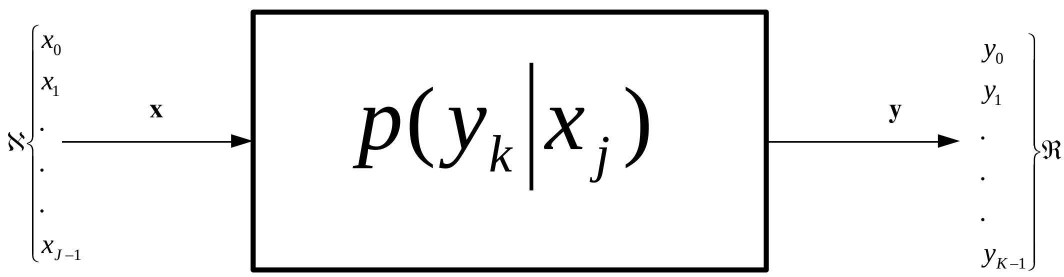
❖ **(2) MỞ RỘNG CỦA NGUỒN KHÔNG NHỚ RỜI RẠC:**

Trong việc thảo luận các khái niệm lý thuyết thông tin, ta th- ờng tìm công cụ hữu hiệu để xét các khối hơn là xét các ký hiệu riêng lẻ, với mỗi khối gồm ***n*** ký hiệu nguồn liên tiếp. Ta có thể xem mỗi khối nh- vậy khi nó đang đ- ợc tạo ra bởi một nguồn mở rộng có $\Im^n$ nguồn sao cho có K^n khối riêng, trong đó K là số các ký hiệu riêng có trong $\Im$ của nguồn ban đầu. Trong tr- ờng hợp nguồn không nhớ rời rạc, thì các ký hiệu ***độc lập thống kê nhau***. Vì vậy xác suất của ký hiệu nguồn trong $\Im^n$ bằng tích các xác suất của ***n*** ký hiệu nguồn trong $\Im$ tạo thành ký hiệu nguồn cụ thể trong $\Im^n$. Vì vậy bằng trực giác mong muốn $H(\Im^n)$, Entropy của nguồn mở rộng bằng ***n*** lần $H(\Im)$, Entropy nguồn ban đầu. Nghĩa là, ta có thể viết.

$$H(\Im^n) = n\, H(\Im) \qquad (2.17)$$

2.4. C□C K□NH KH□NG NH□ R□I R□C - DMC
DESCRETE MEMORYLESS CHANNELS

Ta thay việc khảo sát tạo tin bằng việc truyền tin, d- ới góc độ nhấn mạnh độ tin cậy. Tr- ớc hết đề cập kênh không nhớ rời rạc, bản sao của nguồn không nhớ rời rạc.



Hình 2.7: Kênh không nhớ rời rạc Discrete Memoryless Channel

Kênh không nhớ rời rạc là mô hình thống kê với: Đầu vào X. Đầu ra Y là phiên bản tạp âm của X. Trong đó X và Y đều là biến ngẫu nhiên.

Mỗi đơn vị thời gian, kênh nhận ký hiệu đầu vào X đ- ợc chọn từ $\aleph$ và đáp ứng ra bằng cách nó phát ký hiệu Y từ $\Re$.

Kênh đ- ợc gọi là " rời rạc- Descrete" khi cả $\aleph$ và $\Re$ đều có kích th- ớc hữu hạn và đ- ợc gọi là " không nhớ Memoryless" khi ký hiệu đầu ra hiện thời chỉ phụ thuộc ký hiệu vào hiện thời và không phụ thuộc vào bất kỳ một ký hiệu vào tr- ớc đó.

Hình 2.7 cho minh hoạ kênh rời rạc không nhớ. Kênh đ- ợc mô tả d- ới dạng các đầu vào và đầu ra nh- sau.

+ Các đầu vào: $\aleph = \{x_0, x_1, ..., x_{J-1}\}$ (2.29)

+ Các đầu ra: $\Re = \{y_0, y_1, ..., y_{J-1}\}$ (2.30)

+ Tập các xác suất truyền: $P(y_k|x_j) = P(Y = y_k|X = x_j)$ Với $\forall$ j & k (2.31)

Hiển nhiên ta có.

$$0 \le P(y_k|x_j) \le 1 \qquad \text{với mọi j và k} \qquad (2.32)$$

Bảng mẫu tự đầu vào $\aleph$ và đầu ra $\Re$ không nhất thiết phải có cùng kích th- ớc. *Ví dụ*: trong mã hoá kênh kích th- ớc đầu ra K của $\Re$ có thể lớn hơn kích th- ớc J của đầu vào $\aleph$ vì thế mà K≥J. Mặt khác, ta gặp phải tình huống trong đó kênh phát ra cùng một ký hiệu khi một trong hai ký hiệu đầu vào đ- ợc gửi đi khi đó K≤J.

Xác suất truyền $p(y_k|x_j)$ là xác suất có điều kiện mà đầy ra $Y=y_k$, đã biết đầu vào $X=x_j$. Do hạn chế về vật lý làm ảnh h- ởng đến độ tin cậy khi truyền tin qua kênh gây ra lỗi. Vì vậy, ***khi k=j, thì xác suất truyền $p(y_k|x_j)$ thể hiện xác suất thu có điều kiện đúng, và khi k≠j, thể hiện xác suất lỗi có điều kiện***.

Ph- ơng pháp phổ biến để mô tả kênh không nhớ rời rạc là xắp xếp các xác suất truyền của kênh d- ới dạng ma trận nh- sau.

$$P = \begin{bmatrix} p(y_0|x_0) & p(y_1|x_0) & \cdots\cdots & p(y_{K-1}|x_0) \\ p(y_0|x_1) & p(y_1|x_1) & \cdots\cdots & p(y_{K-1}|x_1) \\ . & . & & . \\ & & \cdots\cdots & \\ . & . & & . \\ p(y_0|x_{J-1}) & p(y_1|x_{J-1}) & \cdots\cdots & p(y_{K-1}|x_{J-1}) \end{bmatrix}_{J\times K} \qquad (2.33)$$

Ma trận P kích th- ớc J×K đ- ợc gọi là ma trận kênh (Channel Matrix). L- u ý mỗi hàng của ma trận P t- ơng ứng với đầu vào kênh cố định (Fixed Channel Input), còn mỗi cột của ma trận P t- ơng ứng với đầu ra kênh cố định (Fixed Channel Output). Cũng cần l- u ý rằng, thuộc tính cơ bản của ma trận kênh P là tổng các phần tử dọc theo một hàng nào đó của ma trận luôn bằng 1, nghĩa là.

$$\sum_{k=0}^{K-1} p(y_k|x_j) = 1 \qquad \text{Với mọi j} \qquad (2.34).$$

Giả sử đầu vào kênh rời rạc không nhớ đ- ợc chọn t- ơng ứng với phân phối xác suất $\{p(x_j), j=0,1,...J-1\}$. Nói cách khác, sự kiện đầu vào $X=x_j$ xuất hiện với xác suất

$$P(x_j) = P(X=x_j) \qquad \text{với } j = 0,1,...,J-1 \qquad (2.35)$$

Xác định biến ngẫu nhiên X biểu thị cho đầu vào kênh, xác định biến ngẫu nhiên Y biểu thị cho đầu ra kênh. Phân phối xác suất hợp của các biến ngẫu nhiên X và Y đ- ợc cho bởi.

$$\begin{aligned} p(x_j, y_k) &= P(X = x_j, Y = y_k) \\ &= P(Y = y_k | X = x_j) P(X = x_j) \qquad (2.36) \\ &= p(y_k | x_j) p(x_j) \end{aligned}$$

Phân phối xác suất mép (*Marginal Probability Distribution*) của biến nhẫu nhiên ra Y đạt đ- ợc bằng cách lấy trung bình phụ thuộc của $p(x_j, y_k)$ trên x_j nh- sau.

$$\begin{aligned} p(y_k) &= P(Y = y_k) \\ &= \sum_{j=0}^{J-1} P(Y = y_k | X = x_j) P(X = x_j) \qquad (2.37) \\ &= \sum_{j=0}^{J-1} P(y_k | x_j) P(x_j) \qquad \text{Với } k = 0, 1, ..., K - 1 \end{aligned}$$

Với J=K, thì xác suất lỗi ký hiệu trung bình Pe đ- ợc xác định là xác suất mà biến ngẫu nhiên đầu ra Y_k khác với biến ngẫu nhiên đầu vào X_j, đ- ợc lấy trung bình trên toàn bộ k≠j. Vì vậy ta viết.

$$\begin{aligned} P_e &= \sum_{\substack{k=0 \\ k \neq j}}^{K-1} P(Y = y_k) \\ &= \sum_{\substack{k=0 \\ k \neq j}}^{K-1} \sum_{j=0}^{J-1} P(y_k | x_j) P(x_j) \end{aligned} \qquad (2.38)$$

⇒ Hiệu (1- Pe) là xác suất thu đúng trung bình.

Các xác suất $p(x_j)$ với j = 0,1,...,J-1, đ- ợc coi là xác suất tiên nghiệm (*Priori-probability*) của các ký hiệu vào.

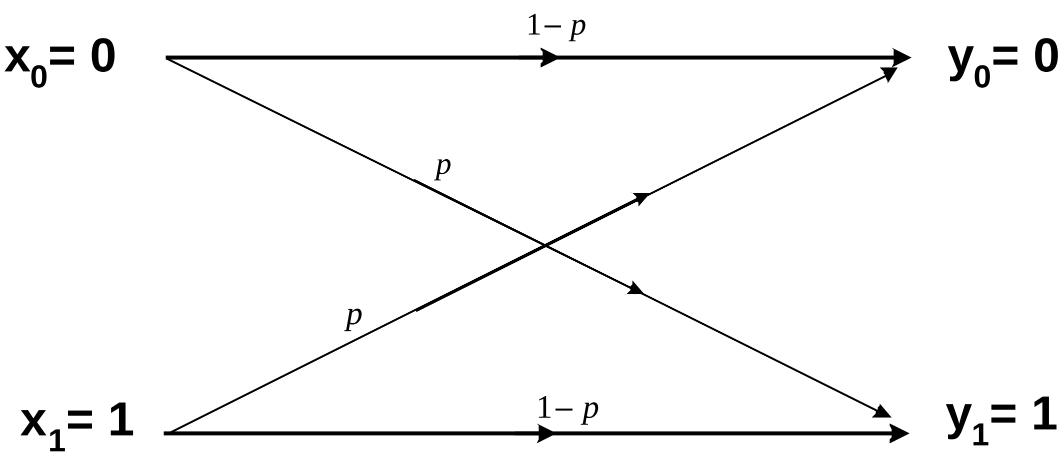
Ph- ơng trình 2.37 cho biết: nếu biết đầu vào có xác suất tiên nghiệm $p(x_j)$ và biết ma trận kênh, ma trận xác suất truyền $p(y_k|x_j)$ ⇒ thì tính đ- ợc xác suất ký hiệu ra $p(y_k)$. Phần sau ta tổng kết lại, ta cho $p(x_j)$, $p(y_k|x_j)$ trong tr- ờng hợp này ta có thể tính $p(y_k)$ bằng ph- ơng trình 2.37.

<u>**Ví du 5:**</u> Kênh nhị phân đối xứng BSC Binary Symmetric Channel.

Kênh BSC là tr- ờng hợp cụ thể của kênh không nhớ rời rạc DMC với J=K=2. Kênh có hai ký hiệu đầu vào ($x_0 = 0$, $x_1 = 1$) và hai ký hiệu đầu ra ($y_0 = 0$, $y_1 = 1$). Kênh đ- ợc gọi là đối xứng vì xác suất thu đ- ợc 1 nếu 0 đ- ợc gửi đi bằng xác xuất thu đ- ợc bit 0 nếu bit 1 đ- ợc gửi đi tức là $p(1|0) = p(0|1)$ ⇒ ma trận kênh là.

$$P = \begin{bmatrix} p(0|0) & p(1|0) \\ p(0|1) & p(1|1) \end{bmatrix}$$

Xác suất truyền hoặc xác suất lỗi có điều kiện đ- ợc ký hiệu là p. Sơ đồ xác suất truyền của BSC đ- ợc cho ở hình 2.8.



Hình 2.8: Kênh đối xứng cơ hai BSC

2.5: THÔNG TIN CHÉO MUTUAL INFORMATION.

Ta hãy nghĩ về đầu ra kênh Y (đ- ợc chọn từ bảng mẫu tự ℜ) nh- là phiên bản tạp âm của đầu vào kênh X (đ- ợc chọn từ bảng mẫu tự ℵ) và Entropy H(ℵ) là l- ợng bất định tiên nghiệm (*prior uncertainty*) về X, làm thế nào có thể đánh giá l- ợng bất định về X sau khi quan sát đ- ợc Y? Để trả lời câu hỏi này, ta triển khai ý t- ởng đã đ- ợc đề cập trong phần 2.1 bằng cách xác định Entropy có điều kiện của X đ- ợc chọn từ bảng mẫu tự ℵ khi biết Y=y_k đã xảy ra nh-

$$\mathrm{H}(\aleph|\mathrm{Y}=y_k) = \sum_{j=0}^{J-1} p(x_j|y_k)\log_2\left(\frac{1}{p(x_j\mid y_k)}\right) \qquad (2.39)$$

L- ợng này là bản thân biến ngẫu nhiên nhận các giá trị H(ℵ|Y=y_0),H(ℵ|Y=y_1),....,H(ℵ|Y=y_{K-1}) với các xác suất p(y_0),p(y_1),...,p(y_{K-1}). Vì vậy, giá trị trung bình của H(ℵ|Y=y_k) trên bảng mẫu tự đầu ra ℜ đ- ợc cho bởi.

$$\begin{aligned} H(\aleph|\Re) &= \sum_{k=0}^{K-1} H(\aleph|Y=y_k)p(y_k) \\ &= \sum_{k=0}^{K-1}\sum_{j=0}^{J-1} p(x_j\mid y_k)p(y_k)\log_2\left(\frac{1}{p(x_j\mid y_k)}\right) \qquad (2.40) \\ &= \sum_{k=0}^{K-1}\sum_{j=0}^{J-1} p(x_j, y_k)\log_2\left(\frac{1}{p(x_j\mid y_k)}\right) \end{aligned}$$

ở dòng cuối, ta đã dùng ph- ơng trình 2.36 viết lại nh- sau.

$$p(x_j, y_k) = p(x_j\mid y_k)p(y_k).$$

L- ợng $H(\aleph|\Re)$ đ- ợc gọi là Entropy có điều kiện (Conditional Entropy). Nó thể hiện cho l- ợng bất định còn lại ở đầu vào kênh sau khi đầu ra kênh đ- ợc quan sát.

Vì Entropy $H(\aleph)$ biểu thị cho sự bất định của ta về đầu vào kênh tr- ớc khi quan sát đầu ra kênh, và Entropy có điều kiện $H(\aleph|\Re)$ biểu thị sự bất định cho đầu vào kênh sau khi quan sát đầu ra kênh, theo đó sự chếnh lệch $H(\aleph)$- $H(\aleph|\Re)$ phải biểu thị sự bất định cho đầu vào kênh mà đ- ợc ***thủ tiêu*** *bằng cách quan sát đầu ra kênh. L- ợng quan trọng này đ- ợc gọi là thông tin chéo (Mutual Information) của kênh đ- ợc ký hiệu bởi $I(\aleph;\Re)$, vì vậy ta có thể viết.*

$$I(\aleph;\Re) = H(\aleph) - H(\aleph|\Re) \qquad (2.41)$$

(1) CÁC THUÔC TÍNH THÔNG TIN CHÉO.

Thuôc tính 1: Thông tin chéo của kênh có tính đối xứng, nghĩa là.

$$I(\aleph;\Re) = I(\Re;\aleph) \qquad (2.42)$$

Trong đó thông tin chéo I($\aleph$,$\Re$) là số đo mức độ bất định đầu vào kênh mà bị thủ tiêu bằng cách quan sát đầu ra kênh, và thông tin chéo I($\Re$,$\aleph$) là số đo của sự bất định về đầu ra kênh mà bị thủ tiêu bằng cách gửi đến (sending) đầu vào kênh.

Thuôc tính 2: Thông tin chéo luôn là số không âm, nghĩa là.

$$I(\aleph;\Re) \geq 0 \qquad (2.47).$$

Thuộc tính 2 phát biểu rằng, ta không thể làm mất thông tin, trên ph- ơng diện trung bình bằng cách quan sát đầu ra của kênh. Vả lại, thông tin trung bình bằng không, nếu và chỉ nếu các ký hiệu *vào ra* của kênh là *độc lập thống kê tức là* $p(x_j, y_k) = p(x_j)p(y_k)$.

Thuôc tính 3: Thông tin chéo của kênh có thể đ- ợc biểu diễn d- ới dạng Entropy đầu ra của kênh nh- sau.

$$I(\aleph;\Re) = H(\Re) - H(\Re|\aleph) \qquad (2.51)$$

Trong đó: $H(\Re|\aleph)$ là Entropy có điều kiện. Thuộc tính này đ- ợc suy ra từ định nghĩa 2.41 về thông tin chéo của kênh và thuộc tính 1.

Thuôc tính 4: Thông tin chéo của kênh liên quan với Entropy hợp của đầu vào ra kênh.

$$I(\aleph;\Re) = H(\aleph) + H(\Re) - H(\aleph,\Re) \qquad (2.52)$$

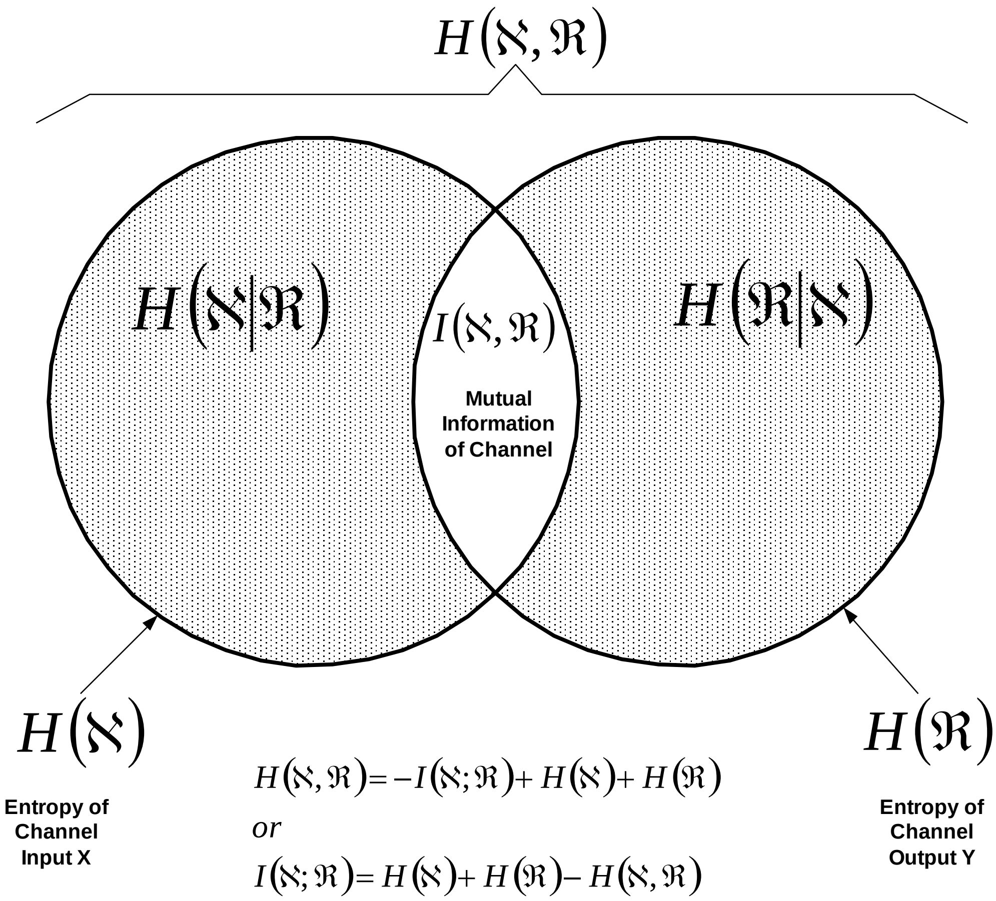
$$\Rightarrow H(\aleph,\Re) = -I(\aleph;\Re) + H(\aleph) + H(\Re) \qquad (2.56)$$

Trong đó Entropy hợp $H(\aleph,\Re)$đ- ợc định nghĩa bởi.

$$H(\aleph,\Re) = \sum_{j=0}^{J-1}\sum_{k=0}^{K-1} p(x_j, y_k)\log_2\left(\frac{1}{p(x_j, y_k)}\right) \qquad (2.53).$$

Ta kết luận l- ợng thông tin chéo của kênh bằng cách thể hiện bằng sơ đồ cho các ph- ơng trình 2.41 và 2.51 và 2.56 đ- ợc cho ở hình vẽ sau.

$$\begin{cases} I(\aleph;\Re) = H(\aleph) - H(\aleph|\Re), & (2.41) \\ I(\aleph;\Re) = H(\Re) - H(\Re|\aleph), & (2.51) \\ H(\aleph,\Re) = -I(\aleph;\Re) + H(\aleph) + H(\Re), & (2.56) \end{cases}$$



Hình 2.9: Minh hoạ mối quan hệ giữa các thông số kênh

2.6: DUNG L- □NG K□NH

CHANNEL CAPACITY

Cơ sở xét:

Để đánh giá kênh truyền phải dựa trên cơ sở chất l- ợng truyền dẫn (Pe) ⇔ xét mô hình kênh sau đ- ợc suy ra từ $I(\aleph;\Re) = H(\aleph) - H(\aleph|\Re)$

✓ Mô hình kênh không nhiễu: là mô hình kênh đ- ợc xác định bởi

$$I(\aleph;\Re) = H(\aleph)$$

$$H(\aleph) = H(\Re) \Rightarrow p(x_j) = p(y_k)$$

✓ Mô hình kênh bị đứt: là mô hình kênh đ- ợc xác định bởi

$$I(\aleph;\Re) = 0$$

Tin thu khác hoàn toàn với tin phát. Nh- vậy coi $\aleph$ & $\Re$ là độc lập nhau nên $p(x_j|y_k) = p(x_j)$; $p(y_k|x_j) = p(y_k)$ và $p(x_j,y_k) = p(x_j)p(y_k)$. khi này ta có

$$H(\aleph|y_k) = H(\aleph)$$

$$H(\Re|x_j) = H(\Re)$$

$$H(\aleph|\Re) = H(\aleph)$$

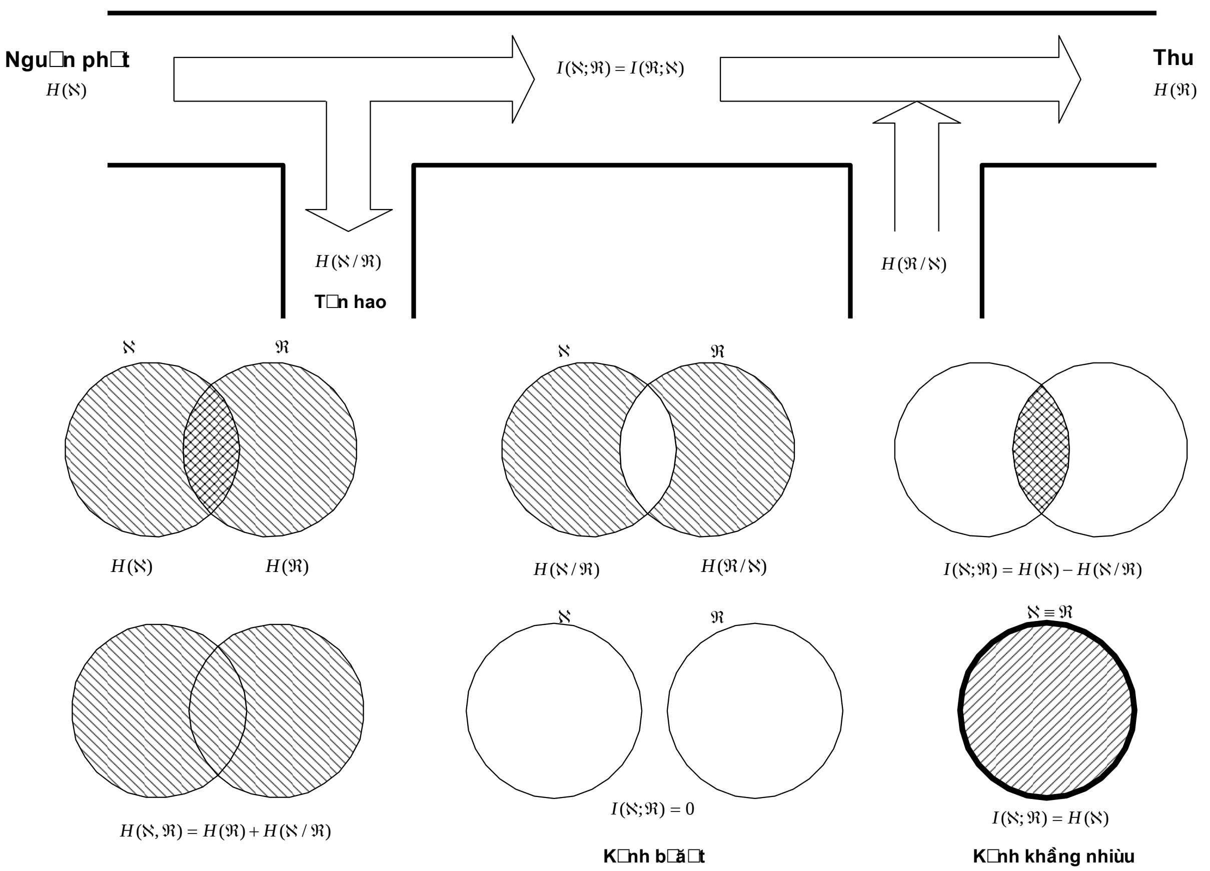
$$H(\Re|\aleph) = H(\Re)$$

✓ Mô hình kênh có nhiễu là mô hình kênh đ- ợc xác định bởi

$$I(\aleph;\Re) = H(\aleph) - H(\aleph|\Re) \text{ khi } H(\aleph|\Re) \neq 0$$

$H(\aleph|\Re)$ l- ợng thông tin tổn hao trung bình của mỗi tin ở phía phát khi phía thu đã thu đ- ợc một tin (dấu) nào đó.

$H(\aleph)$ l- ợng tin trung bình phía phát gửi đi.



Xét kênh rời rạc không nhớ DMC với tập các đầu vào $\aleph$, đầu ra $\Re$ và xác suất truyền $p(y_k|x_j)$. Thông tin chéo của kênh đ- ợc định nghĩa bởi dòng thứ nhất của ph- ơng trình 2.46, để tiện lợi viết lại.

$$I(\aleph;\Re)=\sum_{j=0}^{J-1}\sum_{k=0}^{K-1}p(x_j,y_k)\log_2\left(\frac{p(y_k\,|\,x_j)}{p(y_k)}\right) \tag{2.57}$$

Trong đó l- u ý theo ph- ơng trình 2.36.

$$\begin{aligned}p(x_j,y_k)&=P(X=x_j,Y=y_k)\\&=P(Y=y_k\,|\,X=x_j)P(X=x_j)\\&=\mathrm{p}(\mathrm{y}_k\,|\,\mathrm{x}_j)\mathrm{p}(\mathrm{x}_j)\end{aligned} \tag{2.36}$$

Từ ph- ơng trình 2.37 ta có.

$$\begin{aligned}p(y_k)&=P(Y=y_k)\\&=\sum_{j=0}^{J-1}P(Y=y_k\,|\,X=x_j)P(X=x_j)\\&=\sum_{j=0}^{J-1}P(y_k\,|\,x_j)P(x_j)\qquad \text{Với k = 0,1,..., K - 1}\end{aligned} \tag{2.37}$$

Từ các ph- ơng trình này cho thấy để tính đ- ợc ***thông tin chéo I($\aleph$,$\Re$)*** của kênhcần phải biết:

+ Phân phối xác suất đầu vào {$p(x_j)$, j=0,1,2,...,J-1}.

+ Xác suất truyền của kênh $p(y_k|x_j)$.

⇒ Vì vậy thông tin chéo I($\aleph$,$\Re$) của kênh không chỉ chỉ phụ thuộc vào kênh mà còn phụ thuộc vào cách mà kênh đó đ- ợc dùng.

Hiển nhiên thấy phân phối xác suất đầu vào {$p(x_j)$} độc lập với kênh. Vì vậy ta có thể cực đại hoá thông tin chéo trung bình I($\aleph$,$\Re$) của kênh theo {$p(x_j)$}.

<u>⇒ Định nghĩa dung l- ợng kênh rời rạc không nhớ DMC:</u>

Dung l- ợng kênh C của DMC là thông tin chéo trung bình cực đại MaxI($\aleph$,$\Re$) ở một kênh đơn nào đó (nghĩa là, khoảng thời gian tín hiệu), trong đó việc lấy cực đại hoá đ- ợc thực hiện trên toàn bộ phân phối xác suất đầu vào {$p(x_j)$} trên $\aleph$. Khả năng thông qua kênh ký hiệu C vì vậy ta có thể viết.

$$C=\max_{p(x_j)} I(\aleph;\Re) \tag{2.58}$$

Khả năng thông qua của kênh đ- ợc đo bởi đơn vị bits/kênh (bits per channel use)

<u>*Ý nghĩa:*</u> Đánh giá chất l- ợng kênh truyền dựa trên khả năng truyền đúng ⇒ cho $\aleph$ đầu vào kênh đầu ra kênh nhận đ- ợc là $\Re$. Nếu $\aleph$ = $\Re$ thì kênh tốt ⇒

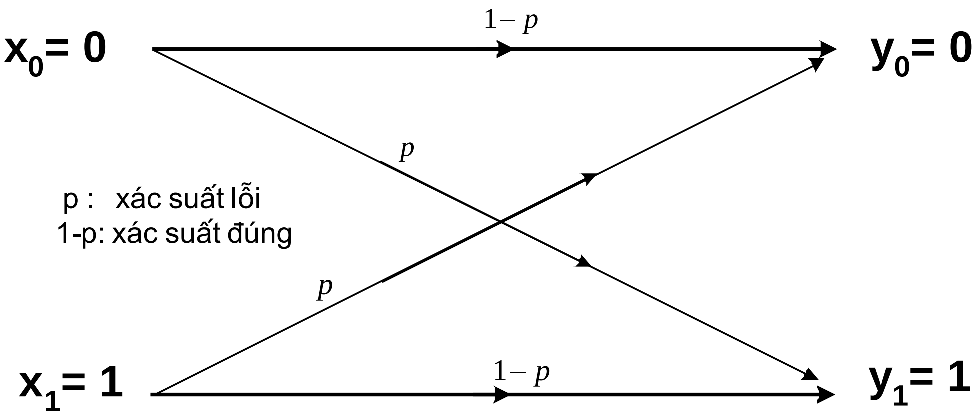
$I(\aleph,\Re)\Rightarrow$max. Ng- ợc lại thì xấu (xem hình 2.9). Khi kênh bị đứt thì $\aleph$ và $\Re$ độc lập nhau $\Rightarrow I(\aleph,\Re)=0 \Rightarrow$ Vậy đánh giá khả năng thông qua của kênh dự trên thông tin chéo giữa các ký hiệu đầu vào ra của kênh.

L- u ý rằng dung l- ợng kênh C chỉ là một hàm của xác suất truyền $p(y_k/x_j)$, mà xác định kênh đó. Việc tính dung l- ợng kênh C bao gồm thực hiện cực đại hoá thông tin chéo trung bình $I(\aleph,\Re)$ trên các biến J [nghĩa là, các xác suất đầu vào $p(x_0),p(x_1),...,p(x_{J-1})$] cho cả bất đẳng thức $p(x_j)\geq 0$ với $\forall$j và đẳng thức $\sum_j p(x_j)=1$.

Nói chung, vấn đề tìm ra khả năng thông qua kênh C có thể hoàn toàn có tính thách thức.

<u>VÍ DỤ 6:</u> Kênh nhị phân đối xứng BSC.

Xét lại kênh BSC, đ- ợc mô tả bởi sơ đồ xác suất truyền của kênh hình 2.8



Hình 2.8: Kênh đối xứng cơ hai BSC

Sơ đồ này xác định hoàn toàn xác suất lỗi có điều kiện p.

Do tính đối xứng, dung l- ợng kênh C đối với BSC đạt đ- ợc với xác suất đầu vào kênh $p(x_0)=p(x_1)=1/2$ (xác suất phát ký hiệu của nguồn tin).

$$C = I(\aleph;\Re)\Big|_{p(x_0)=p(x_1)=\frac{1}{2}} \qquad (2.59)$$

Từ hình 2.8, ta có

Xác suất thu sai: $p(y_0|x_1) = p(y_1|x_0) = p$

Xác suất thu đúng: $p(y_0|x_0) = p(y_1|x_1) = 1- p$.

Vì vậy thay các xác suất truyền kênh này vào ph- ơng trình 2.57 với J=K=2.

$$I(\aleph;\Re) = \sum_{j=0}^{1}\sum_{k=0}^{1} p(x_j,y_k)\log_2\left(\frac{p(y_k\mid x_j)}{p(y_k)}\right) \qquad (2.57)$$

Và thay các giá trị $p(x_0)=p(x_1)=1/2$ vào 2.16.

$$H(p_0) = -p_0\log_2 p_0 - (1-p_0)\log_2(1-p_0) \qquad (2.16)$$

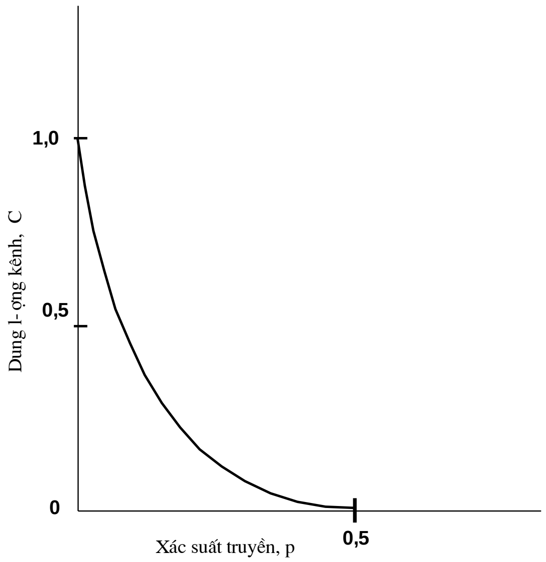
$\Rightarrow$ Ta tìm đ- ợc dung l- ợng kênh của kênh BSC theo xác suất lỗi **p**

$$C = 1 + p\log_2 p + (1-p)\log_2(1-p) \tag{2.60}$$

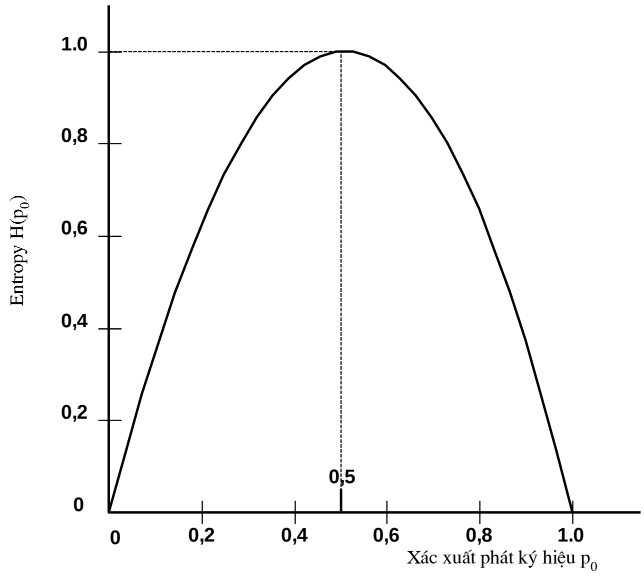
Dựa trên ph- ơng trình 2.16 ta xác định đ- ợc hàm Entropy.

$$H(p) = p\log_2\left(\frac{1}{p}\right) + (1-p)\log_2\left(\frac{1}{1-p}\right) \tag{2.62}$$

Dung l- ợng kênh C thay đổi theo xác suất lỗi **p** nh- đ- ợc cho ở hình 2.10 so sánh đ- ờng cong này với hình 2.2, ta có các nhận xét sau.



Hình 2.10: Sự thay đổi dung l- ợng kênh của kênh BSC theo xác suất truyền của kênh



Hình 2.2: Hàm Entropy $H'(p_0)$

✓ Khi kênh là kênh tạp âm tự do (Noise-Free không có tạp âm), cho phép ta đặt xác suất **lỗi p** = 0, thì dung l- ợng kênh C tiến đến giá trị cực đại của nó là 1bit/kênh, là thông tin chính xác ở mỗi đầu vào. Tại giá trị p = 0 này, thì Entropy H(p) tiến đến giá trị nhỏ nhất của nó là bằng 0.

✓ Khi kênh là kênh tạp âm, tạo ra xác suất truyền **lỗi p** =1/2, thì dung l- ợng kênh C tiến đến giá trị cực tiểu của nó bằng 0, trong khí đó hàm Entropy H(p) tiến đến giá trị cực đại của nó bằng 1.

1.8. ENTROPY VI PHÂN & TH□NG TIN CH□O □□I V□I TOÀN B□ LI□N T□C

DIFFERENTIAL ENTROPY & MUTUAL INFORMATION FOR CONTINOUS ENSEMBLES.

Các nguồn và kênh đ- ợc xét trong phần thảo luận về các khái niệm lý thuyết thông tin hơn nữa phải chứa toàn bộ các biến ngẫu nhiên mà có biên độ rời rạc. Trong phần này, ta mở rộng các khái niệm đó cho các biến ngẫu nhiên liên tục và các vector ngẫu nhiên. Động cơ thúc đẩy đến việc mở rộng này là chuẩn bị ph- ơng pháp để mô tả hạn chế cơ bản khác trong lý thuyết thông tin.

Xét biến ngẫu nhiên liên tục X có hàm mật độ xác suất $f_X(x)$. T- ơng tự với Entropy của biến ngẫu nhiên rời rạc, ta đ- a ra định nghĩa sau.

$$h(X)=\int_{-\infty}^{\infty} f_X(x)\log_2\left(\frac{1}{f_X(x)}\right)dx \qquad (2.68)$$

h(X) là Entropy vi phân của X để phân biệt nó với Entropy tuyệt đối hoặc thông th- ờng (*Ordinary or Absolute Entropy*). Ta thừa nhận thực tế mặc dù h(X) là l- ợng chính xác hữu hiệu để biết, nh- ng không làm mất tính ngẫu nhiên của X. Tuy nhiên, ta đã chứng minh dùng ph- ơng trình 2.68 nh- sau. Ta bắt đầu bằng nhìn nhận biến ngẫu nhiên liên tục X nh- là dạng giới hạn của biến ngẫu nhiên rời rạc mà nó nhận các giá trị $x_k = k\Delta x$ trong đó $k=0,\pm1, \pm2,...$, và $\Delta x\Rightarrow 0$.

Theo định nghĩa thì biến ngẫu nhiên X nhận giá trị trong khoảng $[x_k,x_k+\Delta x]$ với xác suất $f_X(x)\Delta x$. Vì vậy, cho $\Delta x\Rightarrow 0$, Entropy thông th- ờng của biến ngẫu nhiên liên tục X có thể đ- ợc viết lại theo giới hạn nh- sau.

$$\begin{aligned}
H(X)&=\lim_{\Delta x\to 0}\sum_{k=-\infty}^{\infty} f_X(x_k)\Delta x\log_2\left(\frac{1}{f_X(x_k)\Delta x}\right)\\
&=\lim_{\Delta x\to 0}\left[\sum_{k=-\infty}^{\infty} f_X(x_k)\Delta x\log_2\left(\frac{1}{f_X(x_k)}\right)\Delta x-\log_2\Delta x\sum_{k=-\infty}^{\infty} f_X(x_k)\Delta x\right]\\
&=\int_{-\infty}^{\infty} f_X(x)\log_2\left(\frac{1}{f_X(x_k)}\right)dx-\lim_{\Delta x\to 0}\log_2\Delta x\int_{-\infty}^{\infty} f_X(x)dx\\
&=h(X)-\lim_{\Delta x\to 0}\log_2\Delta x
\end{aligned} \qquad (2.69)$$

Trong đó:

ở dòng cuối cùng ta đã dùng ph- ơng trình 2.68 và thực tế toàn bộ vùng diện tích d- ới đ- ờng cong hàm mật độ xác suất bằng $f_X(x) = 1$. Khi lấy giới hạn, $\Delta x\Rightarrow 0$ thì $\log_2\Delta x\Rightarrow\infty$. ***Nghĩa là entropy của biến ngẫu nhiên liên tục là lớn vô cùng***. Bằng trực giác, ta mong muốn là đúng, vì biến ngẫu nhiên liên tục có thể nhận giá trị bất kỳ trong khoảng $[-\infty,\infty]$ và sự bất định (l- ợng thông tin) t- ơng ứng với biến đó là vô hạn.

⇒ Ta ngăn ngừa vấn đề liên quan đến thành phần log₂Δx bằng cách chấp nhận h(X) nh- là Entropy vi phân, với thành phần - log₂Δx xem nh- chuẩn (tham khảo).

Hơn nữa, do thông tin đ- ợc phát lên kênh *là thực* nên sự khác nhau giữa hai thành phần entropy mà có chuẩn chung ⇒ thông tin sẽ giống nh- sự khác nhau giữa hai thành phần entropy vi phân t- ơng ứng. Vì vậy, ta hoàn toàn chứng minh đ- ợc bằng cách dùng thành phần h(X), đ- ợc định nghĩa ở ph- ơng trình 2.68, nh- là Entropy vi phân của biến ngẫu nhiên liên tục X.

❖ Khi ta có vector ngẫu nhiên liên tục X chứa **n** biến ngẫu nhiên $X_1, X_2,..,X_n$, ta định nghĩa entropy vi phân của vector X nh- tích phân bậc n.

$$h(X)=\int_{-\infty}^{\infty} f_X(x)\log 2\left[\frac{1}{f_X(x)}\right]dx \qquad (2.70)$$

Trong đó: $f_X(x)$ là hàm mật độ xác suất hợp của X.

<u>VÍ DU:</u> Maximum Differential Entropy for Specified Variance.

Trong ví dụ này, ta tìm ra lời giải cho vấn đề tối - u hoá đ- ợc hạn chế quan trọng (Important Constrained Optimization Problem). Ta xác định dạng mà hàm mật độ xác suất của biến ngẫu nhiên X phải có để Entropy vi phân của X nhận giá trị lớn nhất của nó đối với một số giá trị ph- ơng sai đ- ợc quy định. D- ới dạng toán học, ta phát biểu lại vấn đề nh- sau.

Với entropy vi phân của biến ngẫu nhiên X đ- ợc định nghĩa bởi.

$$h(X)=-\int_{-\infty}^{\infty} f_X(x)\log_2 f_X(x)dx$$

Tìm hàm mật độ xác suất $f_X(x)$ để h(X) đạt giá trị max, với giả thiết hai hằng số sau.

$$\int_{-\infty}^{\infty} f_X(x)dx=1 \qquad (2.71)$$

và

$$\int_{-\infty}^{\infty}(x-\mu)^2 f_X(x)dx=\sigma^2=\text{hằng số} \qquad (2.72)$$

Trong đó: μ là trung bình và σ^2 là ph- ơng sai của X. Công thức h(X) trong ph- ơng trình 2.68, đ- ợc tái tạo ở đây có sự thay đổi nhỏ. Công nhận ph- ơng sai của X có một giá trị quy định, là quan trọng (có ý nghĩa) vì σ^2 là giá trị đánh giá công suất trung bình và để thực hiện cực đại hoá entropy vi phân h(X) phải giả thiết công suất không đổi. Kết quả của việc tối - u hoá ép buộc (Constrained Optimization) sẽ đ- ợc khai thác ở phần 2.9.

Ta dùng ph- ơng pháp nhân Lagrange để giải bài toán tối - u hoá. Cụ thể, Entropy vi phân h(X) sẽ tiến tới giá trị cực đại chỉ khi tích phân

$$\int_{-\infty}^{\infty}\left[-f_X(x)\log_2 f_X(x)+\lambda_1 f_X(x)+\lambda_2(x-\mu)^2 f_X(x)\right]dx \text{ là dừng}$$

Các thông số λ_1 và λ_2 đ- ợc biết nh- là số nhân Lagrange (Lagrange Multiplier). Nghĩa là, h(X) đạt max chỉ khi đạo hàm của tích phân.

$$-f_X(x)\log_2 f_X(x)+\lambda_1 f_X(x)+\lambda_2(x-\mu)^2 f_X(x) \text{ theo } f_X(x) \text{ bằng không.}$$

⇒ Kết quả.

$$\begin{aligned}-\log_2 e+\lambda_1+\lambda_2(x-\mu)^2 &= \log_2 f_X(x)\\ &= (\log_2 e)\ln f_X(x)\end{aligned}$$

Trong đó e là cơ số của logarit tự nhiên. Giải đối với $f_X(x)$ ta đ- ợc.

$$f_X(x)=e^{\left[-1+\frac{\lambda_1}{\log_2 e}+\frac{\lambda_2}{\log_2 e}(x-\mu)^2\right]} \tag{2.73}$$

L- u ý rằng: λ_2 phải là số âm nếu tích phân của $f_X(x)$ và $(x-\mu)^2 f_X(x)$ theo x là hội tụ. Thay ph- ơng trình 2.73 vào ph- ơng trình 2.71 và 2.72, sau đó giải để tìm λ_1 và λ_2.

$$\lambda_1=\frac{1}{2}\log_2\left(\frac{e}{2\pi\sigma^2}\right)$$

và

$$\lambda_2=\frac{\log_2 e}{2\sigma^2}$$

Vì vậy, dạng mong muốn đối với $f_X(x)$ đ- ợc mô tả bởi.

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)=\frac{1}{\sqrt{2\pi}\sigma}e^{\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)} \tag{2.74}$$

Nó đ- ợc nhìn nhận nh- hàm mật độ xác suất của biến ngẫu nhiên Gaussian X có trung bình μ và ph- ơng sai σ^2. Giá trị max của entropy vi phân của biến ngẫu nhiên nh- vậy đạt đ- ợc bằng cách thay ph- ơng trình 2.74 vào ph- ơng trình 2.68⇒Kết quả đ- ợc.

$$h(X)=\frac{1}{2}\log_2\left(2\pi e\sigma^2\right) \tag{2.75}$$

Ta có thể tổng kết lại ví dụ này:

- Khi cho ph- ơng sai σ^2, thì biến ngẫu nhiên Gaussian có một entropy vi phân lớn nhất có thể đạt đ- ợc bởi một biến ngẫu nhiên nào đó. Nghĩa là, nếu X là một biến ngẫu nhiên Gausian và Y là một biến ngẫu nhiên khác có cùng ph- ơng sai và giá trị trung bình, thì với ∀Y luôn có.

$$h(X) \geq h(Y) \qquad \text{với } \forall Y \qquad (2.76).$$

Trong đó dấu bằng xảy ra khi và chỉ khi X=Y.

- Entropy của biến ngẫu nhiên Gausian X đ- ợc xác định duy nhất bởi ph- ơng sai của X (nghĩa là ph- ơng sai của X độc lập với trung bình của X).

Thực vậy, vì thuộc tính 1 mà mô hình kênh Gaussian đ- ợc dùng phổ biến trong việc nghiên cứu các hệ thống viễn thông.

(1). Thông tin chéo- Mutual Information.

Xét cặp biến ngẫu nhiên liên tục X và Y. T- ơng tự ph- ơng trình 2.44. Ta định nghĩa thông tin chéo giữa biến ngẫu nhiên liên tục X và Y nh- sau.

$$I(X;Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\log_2\left[\frac{f_X(x|y)}{f_X(x)}\right]dxdy \qquad (2.77).$$

Trong đó: $f_{X,Y}(x,y)$ là hàm mật độ xác suất hợp của các biến ngẫu nhiên liên tục X & Y, và $f_X(x|y)$ là hàm mật độ xác suất có điều kiện của X khi đã biết Y=y. Cũng t- ơng tự nh- các ph- ơng trình 2.42, 2.47, 2.41 & 2.51, ta tìm đ- ợc thông tin chéo I(X;Y) có các thuộc tính sau.

1. $I(X;Y) = I(Y;X)$ (2.78)
2. $I(X;Y) \geq 0$ (2.79)
3. $I(X;Y) = h(X) - h(X|Y)$ (2.80)
4. $I(X;Y) = h(Y) - h(Y|X)$. (2.81)

Thông số h(X) là entropy vi phân của biến ngẫu nhiên liên tục X & h(Y) là entropy vi phân của biến ngẫu nhiên liên tục Y. Thông số h(X|Y) là entropy vi phân có điều kiện của biến ngẫu nhiên liên tục X khi đã cho Y; nó đ- ợc định nghĩa bởi tích phân kép sau.

$$h(X;Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\log_2\left[\frac{1}{f_X(x|y)}\right]dxdy \qquad (2.82).$$

Thông số h(Y|X) là entropy vi phân có điều kiện của biến ngẫu nhiên liên tục Y khi đã cho X; nó đ- ợc định nghĩa t- ơng tự nh- h(X|Y).

1.9. □□NH L□ DUNG L- □NG K□NH

CHANNEL CAPACITY THEOREM

Trong phần này ta biểu diễn giới hạn thứ ba trong lý thuyết thông tin, nó đ- ợc định nghĩa bởi định lý dung l- ợng kênh cho các kênh Gaussian có băng tần hạn chế và công suất hữu hạn.

Xét quá trình dừng trung bình không X(t) mà băng tần đ- ợc giới hạn đến B (Hz). Ký hiệu X_k, k = 1,2,...,n, cho các biến ngẫu nhiên liên tục đạt đ- ợc bằng cách lấy mẫu đồng đều quá trình X(t) tại tốc độ 2B mẫu trên giây (trong ch- ơng 4, chỉ ra rằng nếu quá trình X(t) đ- ợc giới hạn băng trong khoảng tần số $-B \le f \le B$, đ- ợc lấy mẫu tại tốc độ lấy mẫu lớn hơn hoặc bằng 2B mẫu trên giây, thì quá trình gốc có thể đ- ợc khôi phục lại từ các mẫu của nó). Các mẫu này đ- ợc phát đi trên kênh tạp âm trong T giây và cũng đ- ợc hạn chế băng B (Hz) (nghĩa là: kênh bị hạn chế băng B Hz). Vì vậy số mẫu n đ- ợc cho bởi

$$\mathbf{n} = 2BT \qquad (2.83)$$

Coi X_k là mẫu tín hiệu phát. Đầu ra kênh bị nhiễu loạn (tác động xấu) bởi tạp âm Gaussian trắng cộng có trung bình không và mật độ phổ công suất $N_0/2$. Tạp âm bị hạn chế băng tới B (Hz). Ký hiệu các biến ngẫu nhiên liên tục Y_k , k=1,2,....,n cho các mẫu tín hiệu thu nh- đ- ợc cho bởi.

$$Y_k = X_k + N_k \qquad k=1,2,...,n \qquad (2.84)$$

Mẫu tạp âm N_k là Gaussian có trung bình 0 và ph- ơng sai đ- ợc cho bởi.

$$\sigma^2 = N_0 \times B \qquad (2.85).$$

Ta giả thiết các mẫu Y_k, k=1,2,...,n là *độc lập thống kê*.

Kênh mà trong đó tín hiệu thu và tạp âm nh- đã đ- ợc mô tả gọi là kênh Gaussian không nhớ, rời rạc theo thời gian (Discrete-time, memoryless gaussian). Để phát biểu đầy đủ ý nghĩa về kênh đó, ta phải ấn định chi phí (cost) cho mỗi đầu vào kênh. Một cách điển hình, *máy phát đ- ợc giới hạn công suất*. Vì vậy, nó là lý do để xác định chi phí nh- .

$$E[X^2_k] = P \qquad k = 1,2,...,n \qquad (2.86)$$

Trong đó P là công suất phát trung bình. Cho nên, ta định nghĩa dung l- ợng kênh nh- sau.

$$C = \max_{f_{X_k}(x)} \left\{ I(X_k;Y_k) : E\left[X^2{}_k\right] = P \right\} \qquad (2.87)$$

Trong đó: $I(X_k;Y_k)$ là thông tin chéo trung bình giữa mẫu tín hiệu phát X_k, và mẫu tín hiệu thu Y_k t- ơng ứng. Giá trị Max đ- ợc chỉ ra ở ph- ơng trình 2.87 *đ- ợc thực hiện theo hàm mật độ xác suất của X_k, $f_{Xk}(x)$.*

Thông tin chéo trung bình $I(X_k;Y_k)$ có thể đ- ợc biểu diễn ở một trong hai dạng t- ơng đ- ơng phù hợp với các ph- ơng trình 2.80 và 2.81. Với mục đích, ta dùng dạng ph- ơng trình 2.81, phát biểu lại nh- sau.

$$I(X_k;Y_k) = h(Y_k) - h(Y_k|X_k) \qquad (2.88)$$

Vì X_k và N_k là các *biến ngẫu nhiên độc lập*, và tổng của chúng là Y_k theo 2.84, ta thấy rằng entropy *vi phân* có điều kiện của Y_k, khi đã rõ về X_k bằng với entropy có điều kiện của N_k (xem bài tập đề 2.8.4)

$$h(Y_k|X_k) = h(N_k) \qquad (2.89)$$

Vì vậy ta có thể viết lại ph- ơng trình 2.88 nh- sau.

$$I(X_k;Y_k) = h(Y_k) - h(N_k) \qquad (2.90)$$

Vì $h(N_k)$ độc lập với phân phối của X_k, thực hiện đại hoá $I(X_k;Y_k)$ theo ph- ơng trình 2.87 cần phải cực đại hoá Entropy vi phân mẫu tín hiệu thu Y_k, $h(Y_k)$. Để $h(Y_k)$ là cực đại, thì Y_k phải là biến ngẫu nhiên Gaussian (xem ví dụ 8). Nghĩa là các mẫu tín hiệu thu thể hiện quá trình tựa tạp âm (Noise-like process). Ta cũng thấy rằng, do N_k là Gaussian theo giả thiết, nên mẫu X_k của tín hiệu phát cũng phải là Gaussian. *Vì vậy ta có thể phát biểu rằng, việc thực hiện cực đại hoá đ- ợc chỉ ra ở 2.87 đạt đ- ợc bằng cách chọn các mẫu tín hiệu phát từ quá trình tựa tạp âm công suất trung bình P. Vì lẽ đó, ta có thể thành lập lại công thức 2.87 nh- sau.*

$$C = I(X_k;Y_k) : X_k \ Gaussian. \quad E\left[X^2{}_k\right] = P \qquad (2.91)$$

Trong đó thông tin chéo $I(X_k;Y_k)$ đ- ợc cho trong ph- ơng trình 2.89.

Để - ớc tính dung l- ợng kênh C, ta xử lý 3 giai đoạn sau.

1. Ph- ơng sai của mẫu tín hiêu thu Y_k bằng $P + \sigma^2$. Vì vậy, dùng ph- ơng trình 2.75 đ- ợc entropy vi phân của Y_k nh- sau.

$$h(Y_k) = \frac{1}{2}\log_2\left[2\pi e\left(P + \sigma^2\right)\right] \qquad (2.92)$$

2. Ph- ơng sai của mẫu tạp âm N_k bằng σ^2. Vì vậy, dùng ph- ơng trình 2.75 đ- ợc entropy vi phân của N_k nh- sau.

$$h(N_k) = \frac{1}{2}\log_2\left(2\pi e\sigma^2\right) \qquad (2.93)$$

3. Thay ph- ơng trình 2.92 và 2.93 vào ph- ơng trình 2.90 và chấp nhận định nghĩa dung l- ợng kênh C đ- ợc cho ở ph- ơng trình 2.91, ta thu đ- ợc kết quả mong muốn.

$$C = \frac{1}{2}\log_2\left(1 + \frac{P}{\sigma^2}\right) \ \text{bits/channel use} \qquad \mathbf{(2.94)}$$

Với kênh đó đ- ợc dùng **n** lần để truyền **n** mẫu của quá trình X(t) trong T giây, thì ta tìm đ- ợc dung l- ợng kênh trên thời gian đơn vị là (**n/T**) lần kết quả đ- ợc cho ở ph- ơng trình 2.94. Con số $n = 2BT$, nh- ph- ơng trình 2.83. Theo đó, ta có thể biểu diễn dung l- ợng kênh C trên thời gian đơn vị nh- sau.

$$C = B\log_2\left(1 + \frac{P}{N_0 B}\right) \quad \text{bits/s} \qquad (2.95)$$

Trong đó ta đã dùng ph- ơng trình 2.85 đối với ph- ơng sai tạp âm σ^2.

Dựa vào công thức 2.95, ta có thể phát biểu định lý thứ ba của Shannon (nổi tiếng nhất)- định lý dung l- ợng kênh nh- sau.

Định lý dung l- ơng kênh truyền:

Dung l- ợng kênh có độ rộng băng B (Hz), bị nhiễu loạn bởi tạp âm Gaussian trắng cộng AWGN có mật độ phổ công suất $N_0/2$ và bị giới hạn băng thông B, đ- ợc cho bởi

$$C = B\log_2\left(1 + \frac{P}{N_0 B}\right) \qquad \text{bits/s}$$

Trong đó P là công suất phát trung bình.

Định lý dung l- ợng kênh là một trong các thành quả quan trọng nhất của lý thuyết thông tin vì nó nêu bật lên đ- ợc sự tác động lẫn nhau mạnh mẽ nhất giữa ba thông số quan trọng: Độ rộng băng thông kênh Channel Bandwidth, Công suất phát trung bình Average Transmitted Power (một cách t- ơng đ- ơng công suất tín hiệu thu trung bình Average Received Signal Power), và mật độ phổ công suất tạp âm Noise Power Spectral Density tại đầu ra kênh truyền dẫn.

Định lý hàm ý rằng:

Nếu cho công suất phát trung bình P, độ rộng băng thông của kênh B ⇒ ta có thể truyền thông tin tại tốc độ C bit/s nh- đ- ợc xác định ở ph- ơng trình 2.95 với xác suất lỗi nhỏ tuỳ ý bằng cách dùng các hệ thống mã hoá phức tạp thích hợp. Không thể truyền tin tại tốc độ lớn hơn C bit/s bởi một hệ thống mã hoá nào đó mà không có lỗi.

⇒ Vì vậy định lý dung l- ợng kênh truyền dẫn xác định giới hạn cơ bản về tốc độ truyền dẫn không lỗi với kênh Gaussian hạn chế băng, công suất giới hạn. Tuy nhiên, để tiến tới giới hạn này, thì tín hiệu phát phải có thuộc tính thống kê xấp xỉ với các thuộc tính của tạp âm Gaussian trắng.

(2) HỆ THỐNG LÝ T□ỞNG: IDEAL SYSTEM

Ta định nghĩa hệ thống lý t- ởng là hệ thống mà dữ liệu phát tại tốc độ bit R_b bằng với dung l- ợng kênh C ($R_b = C$).

Ta biểu diễn công suất phát trung bình nh- sau.

$$P = E_b \times C = E_b \times R_b \qquad (2.96)$$

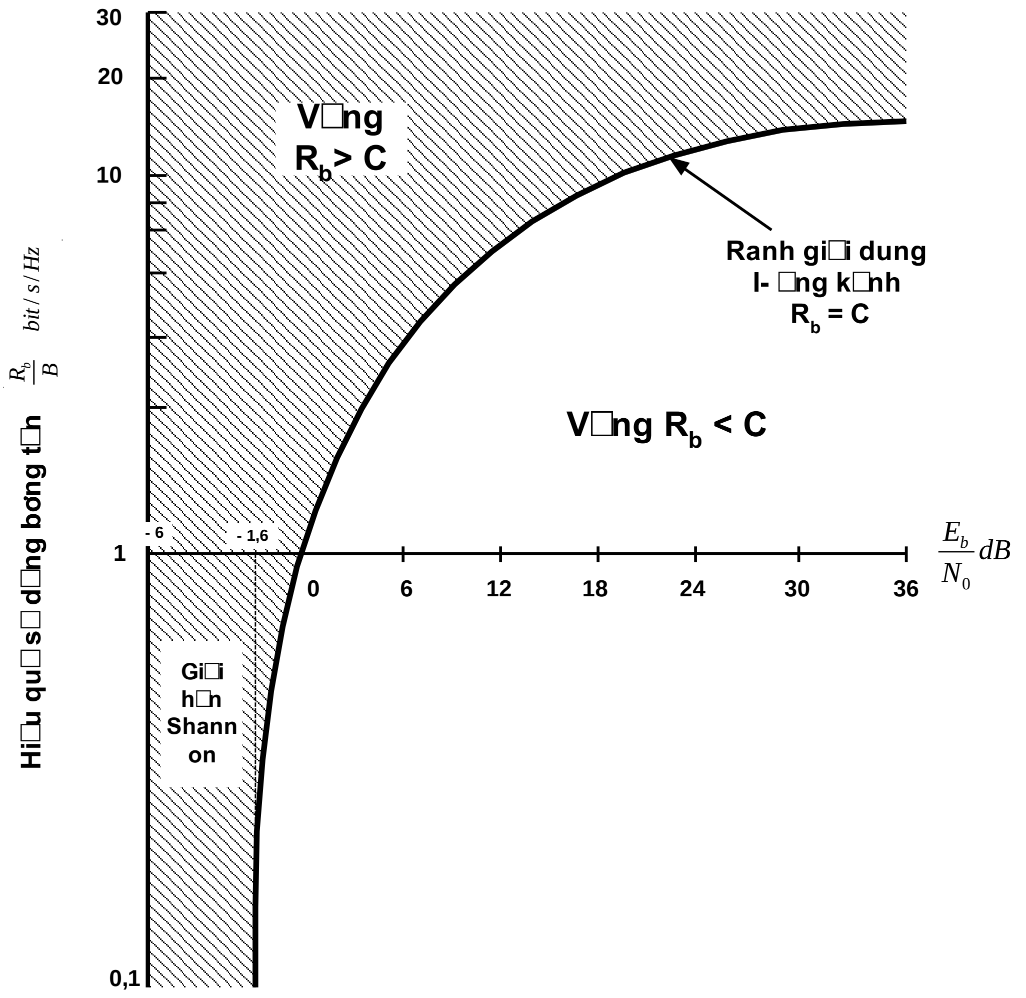
Trong đó: Eb là năng l- ợng tín hiệu phát trên bit. Theo đó, hệ thống lý t- ởng đ- ợc định nghĩa bởi.

$$\frac{C}{B} = \log_2\left(1 + \frac{Eb}{N_0} \times \frac{C}{B}\right) \qquad (2.97)$$

Một cách t- ơng đ- ơng, ta có thể định nghĩa tỉ số giữa năng l- ợng trên bit (E_b) trên mật độ phổ công suất tạp âm (N_0), d- ới dạng hiệu quả sử dụng băng thông (Bandwidth Efficiency), C/B cho hệ thống lý t- ởng nh- sau.

$$\frac{E_b}{C} = \frac{2^{C/B} - 1}{C/B} \tag{2.98}$$

Ph- ơng trình này đ- ợc thể hiện nh- đ- ờng cong đ- ợc đánh nhãn “biên giới dung l- ợng kênh Capacity Boundary” ở hình 2.14.



H□nh 2.14: S□ ă□ hi□u qu□ s□ d□ng bơng thầng Bandwidth-Efficiency Diagram

Vẽ hiệu quả sử dụng băng thông R_b/B theo E_b/N_0 nh- đ- ợc cho ở hình 2.14 đ- ợc gọi là sơ đồ hiệu quả sử dụng băng thông. Dựa trên sơ đồ này ta có một số nhận xét sau.

1. Với độ rộng băng thông vô hạn, thì tỉ số giữa năng l- ợng tín hiệu trên mật độ tạp âm E_b/N_0 tiến đến giá trị hữu hạn.

$$\begin{aligned}\left(\frac{E_b}{N_0}\right)_\infty &= \lim_{B\to\infty}\left(\frac{E_b}{N_0}\right)\\ &= \ln 2 \\ &= 0{,}693\end{aligned} \qquad (2.99)$$

Giá trị này đ- ợc gọi là giới hạn Shannon (Shannon limit), biểu diễn theo đơn vị dB thì nó băng − 1,6 dB. Giá trị giới hạn dung l- ợng kênh t- ơng ứng đạt đ- ợc bằng cách cho độ rộng băng thông kênh B ⇒∝ ta tìm đ- ợc nh- sau.

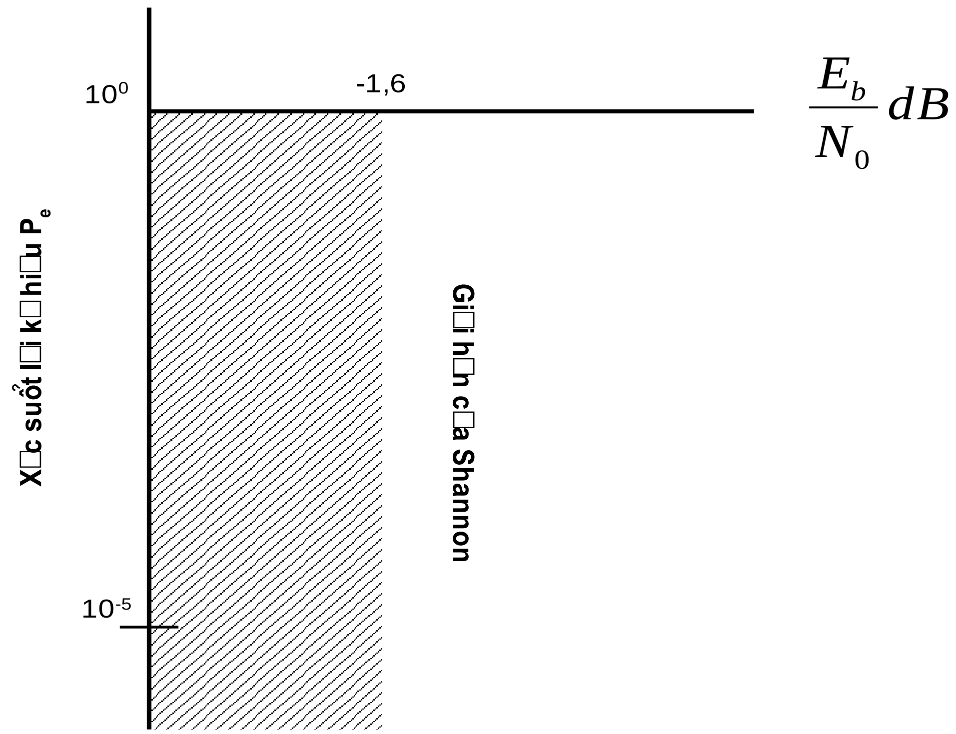
$$\begin{aligned}C_\infty &= \lim_{B\to\infty} C \\ &= \frac{P}{N_0}\log_2 e\end{aligned} \qquad (2.100)$$

2. Ranh giới dung l- ợng kênh (Capacity Boundary), đ- ợc xác định bởi đ- ờng cong với tốc độ bit tới hạn R_b = C, ngăn cách sự kết hợp các thông số hệ thống mà có khả năng trợ giúp cho việc truyền dẫn không lỗi (R_b < C), khi (R_b > C) thì truyền dẫn không lỗi là không thể đ- ợc.
3. Sơ đồ nêu bật sự dung hoà (trade-off) giữa E_b/N_0, R_b/B và xác suất lỗi ký hiệu Pe. Thực tế, ta có thể nhìn nhận sự di chuyển của các điểm làm việc dọc theo đ- ờng nằm ngang nh- quan hệ Pe theo E_b/N_0 khi cố định R_b/B. Mặt khác, ta có thể xét sự di chuyển các điểm làm việc dọc theo đ- ờng thẳng đứng nh- quan hệ giữa Pe theo R_b/B khi cố định giá trị E_b/N_0.

Giới hạn Shannon cũng có thể đ- ợc xác định d- ới dạng E_b/N_0 theo yêu cầu của hệ thống lý t- ởng đối với truyền dẫn không lỗi có thể đ- ợc. Ta có thể viết cho hệ thống lý t- ởng .

$$P_e = \begin{cases} 0 & \left(E_b/N_0\right) \ge \ln 2 = -1{,}6dB \\ 1 & \left(E_b/N_0\right) < \ln 2 = -1{,}6dB \end{cases} \qquad (2.101)$$

Quan điểm này về hệ thống lý tởng đ- ợc mô tả trong hình 2.15. Ranh giới giữa truyền dẫn không lỗi và truyền dẫn không khả tin có các lỗi có thể xảy ra đ- ợc định nghĩa bởi giới hạn Shannon trong hình 2.15 t- ơng tự nh- ranh giới dung l- ợng kênh trong hình 2.14.



H□nh 2.15: S□ ă□ t□□ l□i