Digital Communication Systems
EES 452

Asst. Prof. Dr. Prapun Suksompong

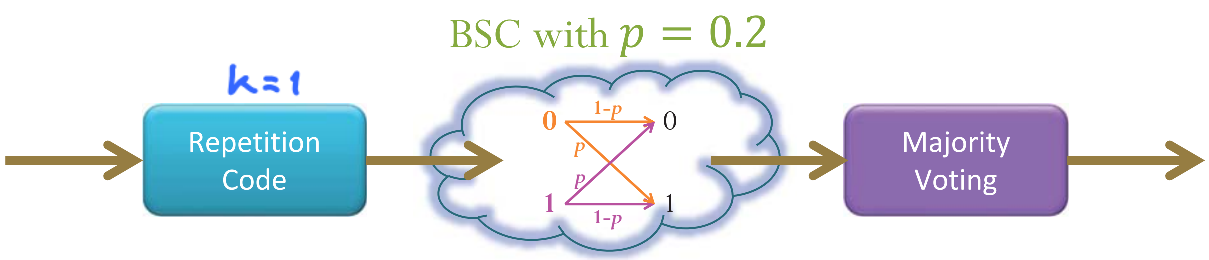
prapun@siit.tu.ac.th

4. Mutual Information and Channel Capacity

4.2 Operational Channel Capacity

Example: Repetition Code

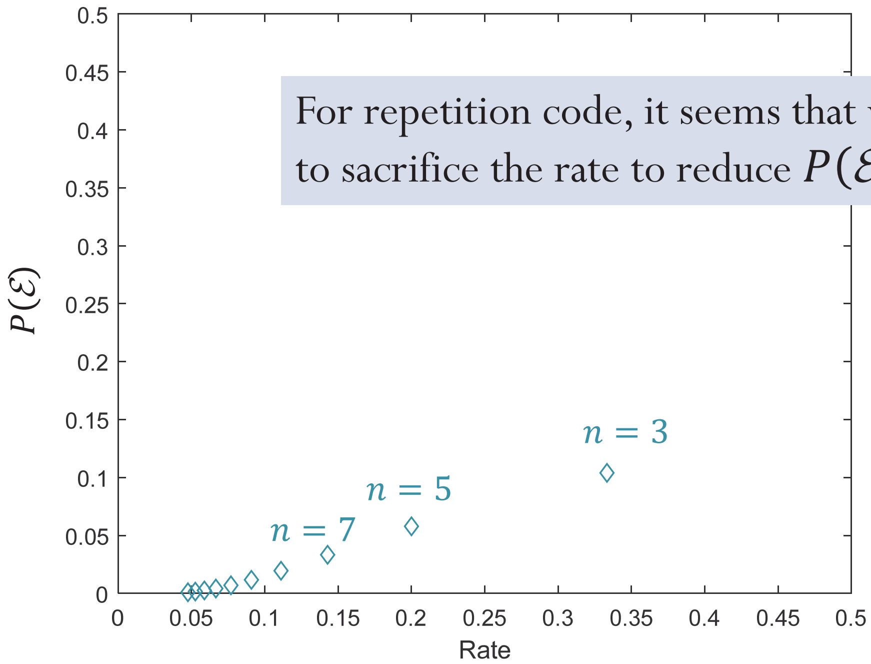
[Section 3.5]



n	$P(\mathcal{E})$ Probability that more than half of the bits are in error	Code Rate $= \frac{k}{n} = \frac{1}{n}$
1	$p = 0.2$	$\frac{1}{1} = 1$
3	$\binom{3}{2} p^2(1-p) + \binom{3}{3} p^3 \approx 0.1040$	$\frac{1}{3} \approx 0.33$
5	$\binom{5}{3} p^3(1-p)^2 + \binom{5}{4} p^4(1-p)^1 + \binom{5}{5} p^5 \approx 0.0579$	$\frac{1}{5} = 0.2$
7	≈ 0.0333	$\frac{1}{7} \approx 0.1429$
9	≈ 0.0196	$\frac{1}{9} \approx 0.1111$
11	≈ 0.0117	$\frac{1}{11} \approx 0.0909$

Achievable Performance

BSC with $p = 0.2$ Repetition Code ($k = 1$)



For repetition code, it seems that we have to sacrifice the rate to reduce $P(\mathcal{E})$

Designing Channel Encoder

2^k rows

$\underline{\mathbf{s}}$	$\underline{\mathbf{x}}$
00	? ? ? ? ?
01	? ? ? ? ?
10	? ? ? ? ?
11	? ? ? ? ?

n columns

Each "?" can be 0 or 1.

So, there are

$$2^{(n2^k)}$$ $= 1{,}048{,}576$ for $n = 5, k = 2$

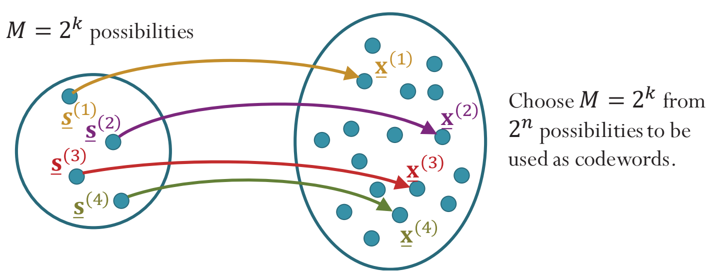
possibilities.

But we don't want to use the same codeword to represent two different info blocks.

So, actually, we need to consider

$$\binom{2^n}{2^k}$$ $= 35{,}960$ for $n = 5, k = 2$

possibilities.



MATLAB

```
close all; clear all;

% EES315 2020 Example 6.58
% EES452 2020 Examples 3.62, 3.67
C = [0 0 0 0 0; 1 1 1 1 1]; % repetition code

p = (1/100);
PE_minDist(C,p)
```

Code C is defined by putting all its (valid) codewords as its rows. For repetition code, there are two codewords: 00..0 and 11..1.

Crossover probability of the binary symmetric channel.

```
>> PE_minDist_demo1

ans =
   9.8506e-06
```

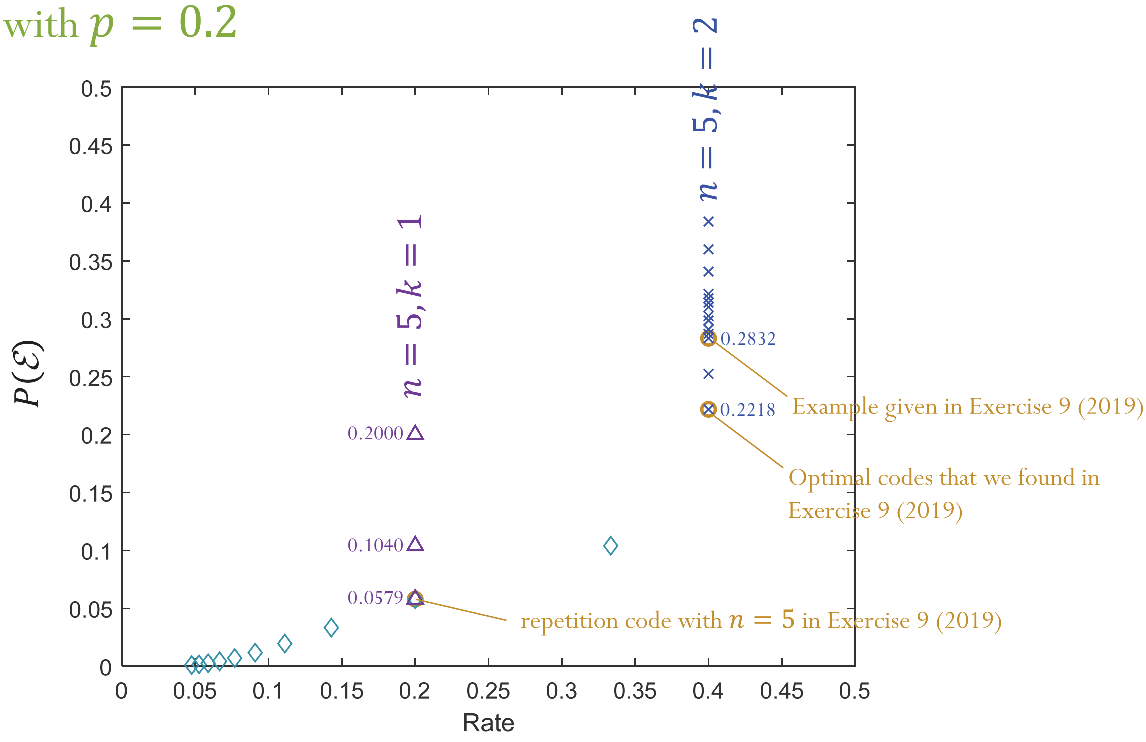

MATLAB

PE_minDist.m

```
function PE = PE_minDist(C,p)
% Function PE_minDist computes the error probability P(E) when code C
% is used for transmission over BSC with crossover probability p.
% Code C is defined by putting all its (valid) codewords as its rows.
M = size(C,1); % the number of (valid) codewords
k = log2(M);
n = size(C,2);

% Generate all possible n-bit received vectors
Y = dec2bin(0:2^n-1)-'0';

% Normally, we need to construct an extended Q matrix. However, because
% each conditional probability in there is a decreasing function of the
% (Hamming) distance, we can work with the distances instead of the
% conditional probability. In particular, instead of selecting the max in
% each column of the Q matrix, we consider min distance in each column.
dminy = zeros(1,2^n); % preallocation
for j = 1:(2^n)
    % for each received vector y,
    y = Y(j,:);
    % find the minimum distance
    % (the distance from y to the closest codeword)
    d = sum(mod(bsxfun(@plus,y,C),2),2);
    dminy(j) = min(d);
end

% From the distances, calculate the conditional probabilities.
% Note that we compute only the values that are to be selected (instead of
% calculating the whole Q first).
n1 = dminy; n0 = n-dminy;
Qmax = (p.^n1).*((1-p).^n0);
% Scale the conditional probabilities by the input probabilities and add
% the values. Note that we assume equally likely input.
PC = sum((1/M)*Qmax);
PE = 1-PC;
end
```

Achievable Performance

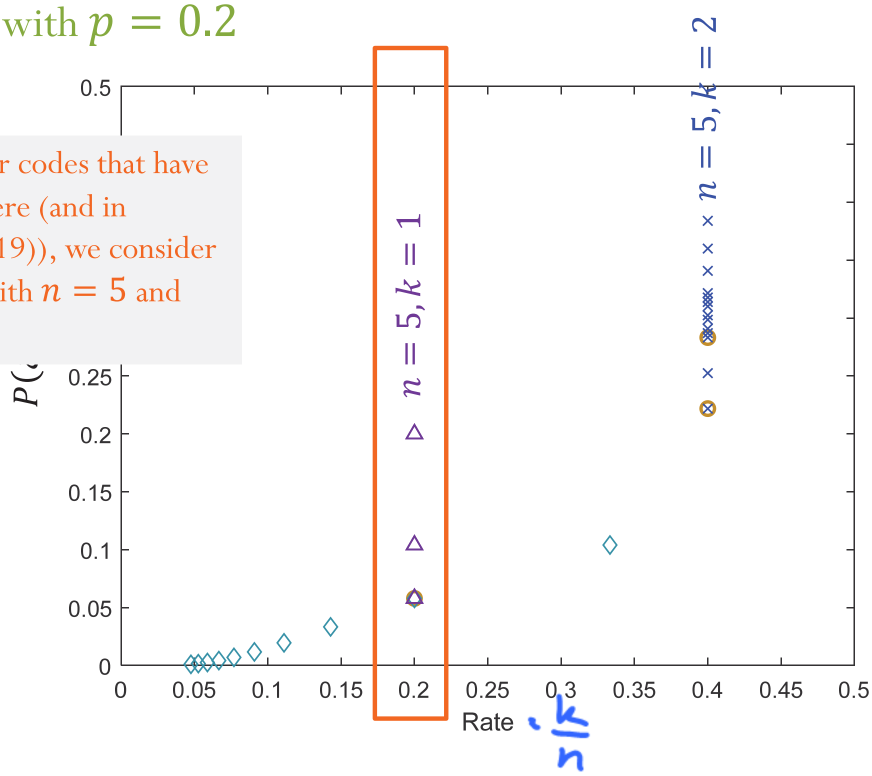
BSC with $p = 0.2$



Achievable Performance

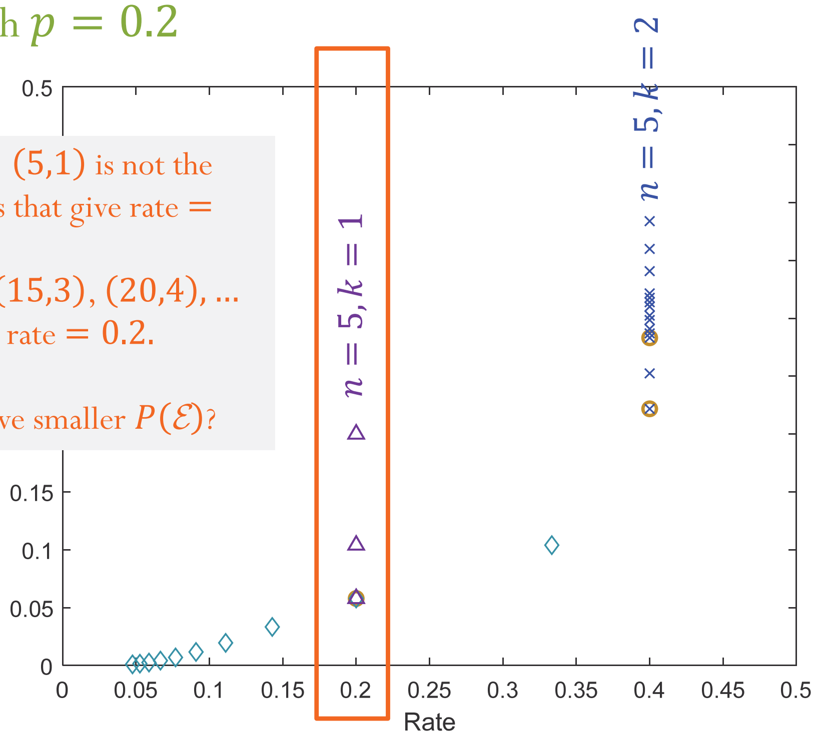
BSC with $p = 0.2$

There are other codes that have rate $= 0.2$. Here (and in Exercise 9 (2019)), we consider all the codes with $n = 5$ and $k = 1$.



Achievable Performance

BSC with $p = 0.2$

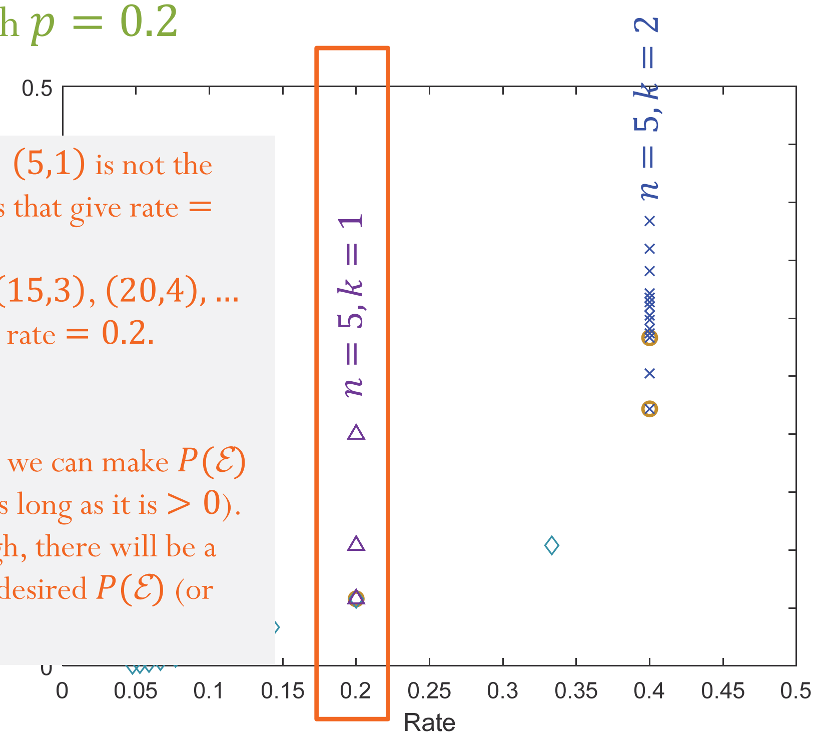


Note that $(n, k) = (5,1)$ is not the only family of codes that give rate $=$ 0.2.
$(n, k) = (10,2), (15,3), (20,4), \dots$ also corresponds to rate $= 0.2$.

Will these codes have smaller $P(\mathcal{E})$?

Achievable Performance

BSC with $p = 0.2$



Note that $(n, k) = (5,1)$ is not the only family of codes that give rate $=$ 0.2.
$(n, k) = (10,2), (15,3), (20,4), \dots$ also corresponds to rate $= 0.2$.

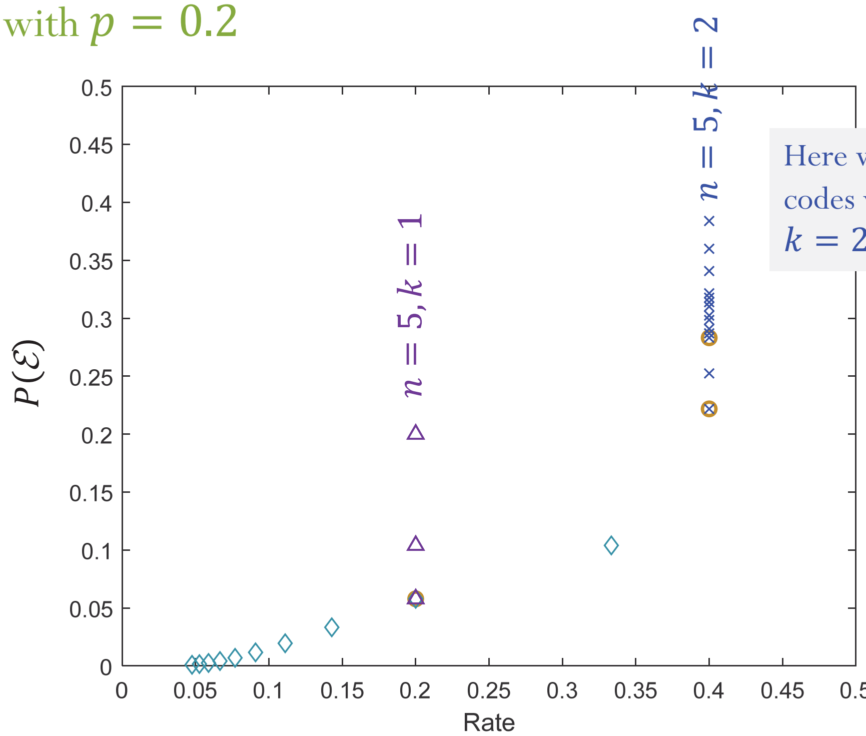
At rate $= 0.2$,
Shannon found that we can make $P(\mathcal{E})$ as small we want (as long as it is > 0).
With n large enough, there will be a code that gives the desired $P(\mathcal{E})$ (or smaller).

Reliable communication

- **Reliable communication** (at a particular rate) means arbitrarily small error probability can be achieved (at that rate).
- In our example, Shannon showed that reliable communication is achievable at rate $= 0.2$.
- Turn out that reliable communication is <u>not</u> achievable at rate $= 0.4$.

Achievable Performance

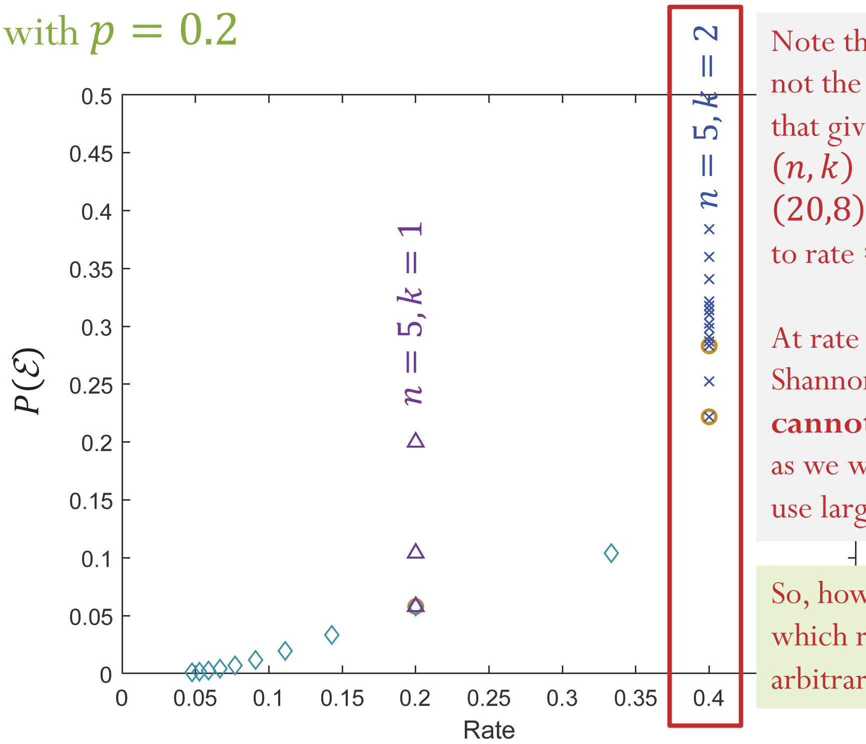
BSC with $p = 0.2$



Here we consider all the codes with $n = 5$ and $k = 2$.

Achievable Performance

BSC with $p = 0.2$



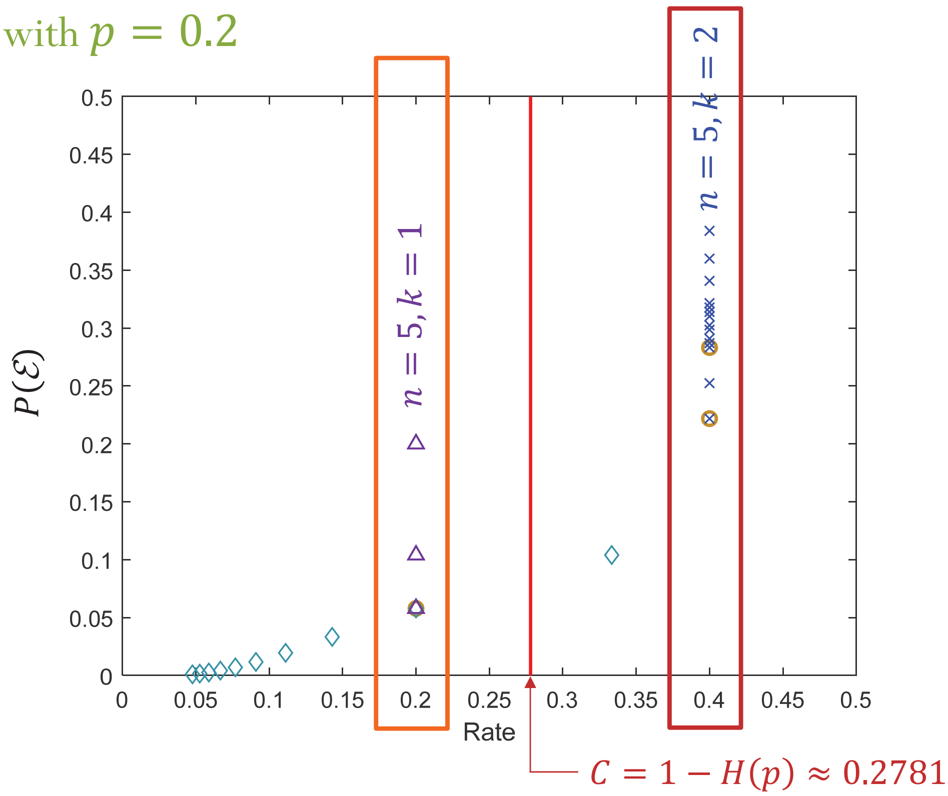
Note that $(n, k) = (5,2)$ is not the only family of codes that give rate $= 0.4$. $(n, k) = (10,4), (15,6), (20,8), \ldots$ also corresponds to rate $= 0.4$.

At rate $= 0.4$, Shannon found that we **cannot** make $P(\mathcal{E})$ as small as we want; even when we use large n.

So, how can we determine which rate can have arbitrarily small $P(\mathcal{E})$?

Achievable Performance

BSC with $p = 0.2$



Channel Capacity

[Section 4.2]

Channel Capacity

"**Operational**": max rate at which **reliable** communication is possible

→ Arbitrarily small error probability can be achieved.

"**Information**": $\max_{\mathbf{p}} I(X;Y)$ [bpcu]

[Section 4.3]

Shannon [1948] showed that these two quantities are actually the same.

4.2 Operational Channel Capacity

4.16. In Chapter 3, we have studied how to compute the error probability $P(\mathcal{E})$ for digital communication systems over DMC. At the end of that chapter, we studied block encoding where the channel is used n times to transmit a k-bit info-block.

In this section, our consideration is "reverse".

4.17. In this and the next sections, we introduce a quantity called channel capacity which is crucial in benchmarking communication system. Recall that, in Chapter 2 where source coding was discussed, we were interested in the minimum rate (in bits per source symbol) to represent a source. Here, we are interested in the maximum rate (in bits per channel use) that can be sent through a given channel *reliably*.

4.18. Here, **reliable communication** means *arbitrarily* small error probability *can be achieved.*

- This seems to be an impossible goal.
  - If the channel introduces errors, how can one correct them all?
    - Any correction process is also subject to error, ad infinitum.

Definition 4.19. Given a DMC, its **"operational" channel capacity** is the maximum rate at which *reliable* communication over the channel *is possible*.

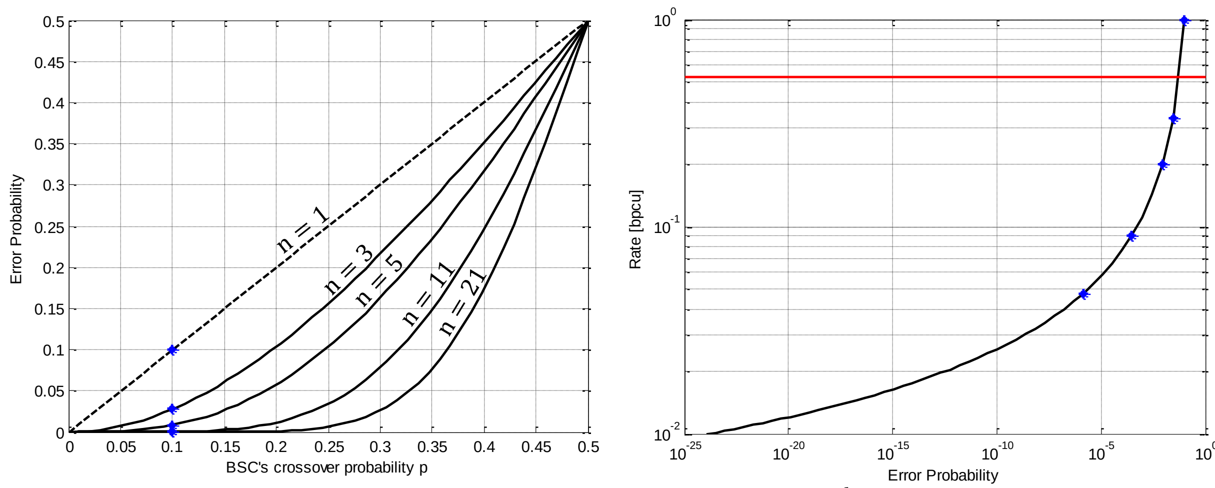
- The channel capacity is the maximum rate in bits per channel use at which information *can be* sent with **arbitrarily low** error probability.

4.20. Claude Shannon showed, in his 1948 landmark paper, that this operational channel capacity is the same as the information channel capacity which we will discuss in the next section. From this, we can omit the words "operational" and "information" and simply refer to both quantities as the *channel capacity*.

Example 4.21. In Example 4.35, we will find that the capacity of a BSC with crossover probability $p = 0.1$ is approximately 0.531 bits per channel use. This means that for any rate $R < 0.531$ and any error probability $P(\mathcal{E})$ that we desire, as long as it is greater than 0, we can find a suitable n, a rate R encoder, and a corresponding decoder which will yield an error probability that is at least as low as our set value.

- Usually, for small desired value of $P(\mathcal{E})$, we may need large value of n.

Example 4.22. Repetition code is not good enough.



(a) (b) $p = 0.1$

Figure 17: Performance of repetition coding with majority voting at the decoder

- Continue from Example 4.21.
- In Figure 17b, with repetition code, trying to reduce the error probability to be less than the original p even a little bit already causes the rate to drop far below the capacity level indicated by the red horizontal line.
- In fact, for any rate > 0, we can see from Figure 17b that communication system based on repetition coding is not "reliable" according to Definition 4.18. For example, for rate $= 0.02$ bits per channel use,

repetition code can't satisfy the requirement that the error probability must be less than 10^{-15}. In fact, Figure 17b shows that as we reduce the error probability to 0, the rate also goes to 0 as well. Therefore, there is no positive rate that works for all error probability.

- However, because the channel capacity is 0.531 [bpcu], there must exist other encoding techniques which give better error probability than repetition code.
  - Although Shannon's result gives us the channel capacity, it does not give us any explicit instruction on how to construct codes which can achieve that value.

4.3 Information Channel Capacity

4.23. In Section 4.1, we have studied how to compute the value of mutual information $I(X;Y)$ between two random variables X and Y. Recall that, here, X and Y are the channel input and output, respectively. We have also seen, in Example 4.14, how to compute $I(X;Y)$ when the joint pmf matrix $\mathbf{P}$ is given. Furthermore, we have also worked on Example 4.15 in which the value of mutual information is computed from the prior probability vector $\underline{\mathbf{p}}$ and the channel transition probability matrix $\mathbf{Q}$. This second type of calculation is crucial in the computation of channel capacity. This kind of calculation is so important that we may write the mutual information $I(X;Y)$ as $I(\underline{\mathbf{p}},\mathbf{Q})$.

Definition 4.24. Given a DMC channel, we define its "information" channel capacity as

$$C = \max_{\underline{\mathbf{p}}} I\left(X;Y\right) = \max_{\underline{\mathbf{p}}} I\left(\underline{\mathbf{p}},\mathbf{Q}\right), \tag{34}$$

where the maximum is taken over all possible input pmfs $\underline{\mathbf{p}}$.

- Again, as mentioned in 4.20, Shannon showed that the "information" channel capacity defined here is equal to the "operational" channel capacity defined in Definition 4.19.
  - Thus, we may drop the word "information" in most discussions of channel capacity.