

3 An Introduction to Digital Communication Systems Over Discrete Memoryless Channel (DMC)

3.1 Discrete Memoryless Channel (DMC) Models

In this section, we keep our analysis of the communication system simple by considering purely digital systems. (Recall that the transmitted signal from an antenna is an analog waveform.) To do this, as shown in Figure 5, we **assume all non-source-coding parts of the system, including the physical (analog) channel, can be combined into an “equivalent channel”** which we shall simply refer to in this section as the **“channel”**.

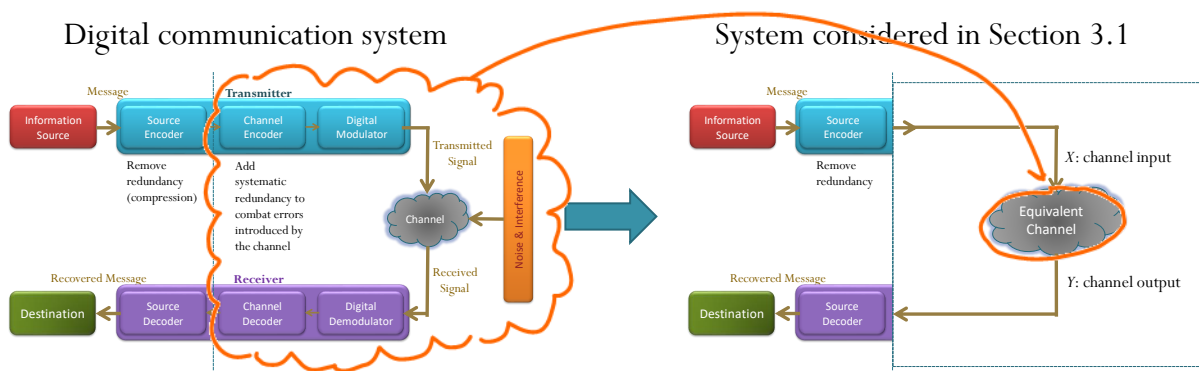


Figure 4: Equivalent Channel Considered in Section 3.1.

Example 3.1. In Chapter 2, the (equivalent) channel does not change (corrupt) its input. The channel output is assumed to be the same as the channel input.

Example 3.2. The **binary symmetric channel (BSC)**, which is the simplest model of a channel with errors, is shown in Figure 5.

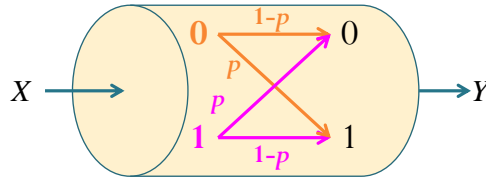


Figure 5: Binary symmetric channel and its channel diagram

- “Binary” means that there are **two possible values** for the input and also two possible values for the output. We normally use the symbols 0 and 1 to represent these two values.
- Passing through this channel, the input symbols are complemented with **crossover probability p** .
- It is simple, yet it captures most of the complexity of the general problem.
switchover bit-flip

Example 3.3. Consider a BSC whose samples of input and output are provided below

x: 1 0 1 1 1 1 1 1 1 1 1 1 0 1 0 1 1 1 1 1
y: 1 1 1 1 1 1 1 1 1 0 1 1 0 1 0 1 1 1 1 1

Estimate the following (unconditional and conditional) probabilities by their relative frequencies.

$$P[X = 0] \approx \frac{3}{20} = 0.15$$

$$P[X = 1] \approx \frac{17}{20}$$

$$P[Y = 0] \approx \frac{3}{20} = 0.15$$

$$P[Y = 1] \approx \frac{17}{20}$$

$$P[Y = 0|X = 0] \approx \frac{2}{3} \approx 0.67$$

$$P[Y = 1|X = 0] \approx \frac{1}{3}$$

$$P[Y = 0|X = 1] \approx \frac{1}{17}$$

$$P[Y = 1|X = 1] \approx \frac{16}{17}$$

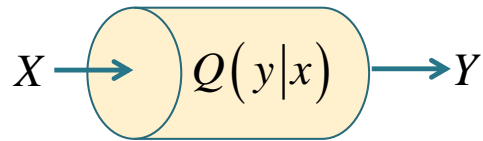
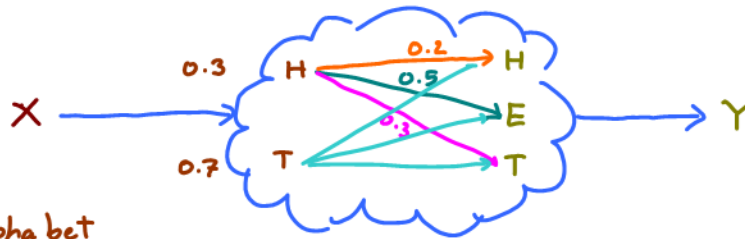


Figure 6: Discrete memoryless channel

Definition 3.4. Our general model for **discrete memoryless channel (DMC)** is shown in Figure 6.

- The **channel input** is denoted by a random variable X .
 - The pmf $p_X(x)$ is usually denoted by $p(x)$.
 For example, $p(0) = p_X(0) = P[X=0]$
 $p(1) = p_X(1) = P[X=1]$
 - The **support S_X** is often denoted by \mathcal{X} .
 - * \mathcal{X} may be referred to as the **channel input alphabet**.
 - * In many DMC, $|\mathcal{X}|$ is a power of two.
 - For finite $|\mathcal{X}|$, the whole pmf $p(x)$ is usually expressed in the form of a row vector $\underline{\mathbf{p}}$ or $\underline{\pi}$.

- Similarly, the **channel output** is denoted by a random variable Y .
 - The pmf $p_Y(y)$ is usually denoted by $q(y)$ and usually expressed in the form of a row vector $\underline{\mathbf{q}}$.
 - The **support S_Y** is often denoted by \mathcal{Y} and referred to as the **channel output alphabet**.



Channel Input Alphabet

$$\mathcal{X} = \{H, T\}$$

Input Probabilities

$$P[X=H] = 0.3 \equiv p(H)$$

$$P[X=T] = 0.7 \equiv p(T)$$

Input Probability Vector

$$\underline{p} = [0.3 \quad 0.7]$$

$\underline{\quad}$ denotes row vector

Channel Transition Probabilities

$$Q(E|H) \equiv P[Y=E | X=H] = 0.5$$

$$Q(T|H) \equiv P[Y=T | X=H] = 0.2$$

$$Q(H|H) \equiv P[Y=H | X=H] = 0.2$$

Channel Matrix

$$\mathbf{Q} = \begin{matrix} & \begin{matrix} H & E & T \end{matrix} \\ \begin{matrix} H \\ T \end{matrix} & \begin{bmatrix} 0.2 & 0.5 & 0.2 \\ \dots & \dots & \dots \end{bmatrix} \end{matrix}$$

Channel Output Alphabet

$$\mathcal{Y} = \{H, E, T\}$$

Output probabilities

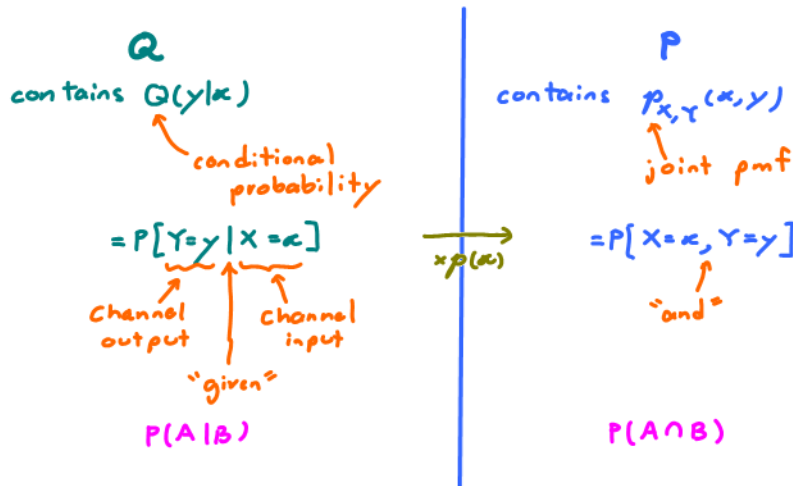
$$P[Y=H] \equiv q(H)$$

$$P[Y=E] \equiv q(E)$$

$$P[Y=T] \equiv q(T)$$

Output Probability Vector

$$\underline{q} = [\quad \quad \quad]$$



Both P and Q give the relationship between two RVs X and Y.

However, P contains more information.

Knowing P, we can calculate any probability statement about X and Y.

To get P from Q, we need \underline{p} .

↳ additional info

To get Q from P, it may seem that we need \underline{p} .

However, \underline{p} can be calculated from the matrix P itself by summing along each row.

- The channel corrupts its input X in such a way that when the input is $X = x$, its output Y is randomly selected from the conditional pmf $p_{Y|X}(y|x)$.

$$Q(y|x) \equiv p_{Y|X}(y|x) \equiv \underbrace{P[Y=y]}_A \underbrace{|X=x]}_B = P(A|B) \equiv \frac{P(A \cap B)}{P(B)} = \frac{P[Y=y, X=x]}{P[X=x]}$$

- In this context, each conditional probability $p_{Y|X}(y|x)$ is usually referred to as the **channel transition probability**.
- The conditional pmf $p_{Y|X}(y|x)$ is usually denoted by $Q(y|x)$.

and usually expressed in the form of a (probability) **transition matrix Q**:

$$Q = \begin{matrix} & \begin{matrix} x_1 & x_2 & \dots & x_n \end{matrix} \\ \begin{matrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{matrix} & \begin{bmatrix} \dots & & & \dots \\ \dots & P[Y=y_j|X=x_i] & \dots & \dots \\ \dots & & \dots & \dots \end{bmatrix} \end{matrix}$$

This matrix is often referred to as the “matrix of transition probabilities” or simply the “**channel matrix**”.

- The channel is called **memoryless**¹⁰ because its channel output at a given time is a function of the channel input at that time and is not a function of previous channel inputs.
- Here, the transition probabilities are assumed constant. However, in many commonly encountered situations, the transition probabilities are time varying. An example is the wireless mobile channel in which the transmitter-receiver distance is changing with time.

¹⁰Mathematically, the condition that the channel is memoryless may be expressed as [12, Eq. 6.5-1 p. 355]

$$p_{Y_1^n|X_1^n}(y_1^n|x_1^n) = \prod_{k=1}^n Q(y_k|x_k)$$

where x_1^n denotes the vector (x_1, x_2, \dots, x_n) .

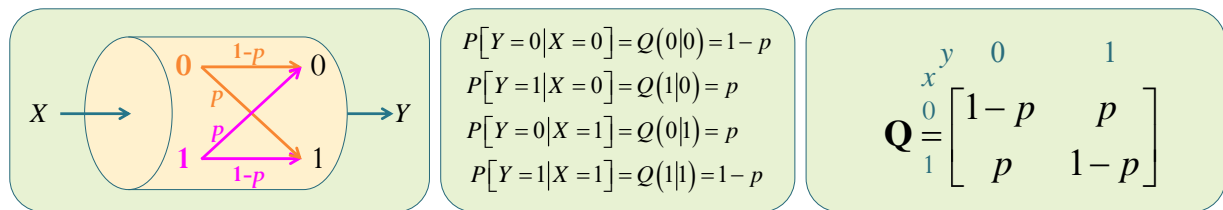
Channel Input Alphabet $S_X \equiv \mathcal{X}$	$P[X = x]$ $p_X(x) \equiv p(x)$ $\xrightarrow{\text{row vector}}$ $\underline{\mathbf{p}}$	$P[Y = y X = x]$ $p_{Y X}(y x) \equiv Q(y x)$ $\xrightarrow{\text{matrix}}$ \mathbf{Q}	Put together All values at corresponding positions in the matrix.
$S_Y \equiv \mathcal{Y}$ Channel Output Alphabet	$p_Y(y) \equiv q(y)$ $\xrightarrow{\text{row vector}}$ $\underline{\mathbf{q}}$ $P[Y = y]$	$p_{X,Y}(x, y) \equiv p(x, y)$ $\xrightarrow{\text{matrix}}$ \mathbf{P} $P[X = x, Y = y]$	"AND"

Figure 7: Notation involved in defining and describing characteristics of digital communication channels

Example 3.5. Recap: Let's try to rewrite each of the probabilities calculated in Example 3.3 using our new notation:

$$\begin{aligned}
 P[X = 0] &\equiv p(0) & P[X = 1] &\equiv p(1) \\
 P[Y = 0] &\equiv q(0) & P[Y = 1] &\equiv q(1) \\
 P[Y = 0|X = 0] &\equiv Q(0|0) & P[Y = 1|X = 0] &\equiv Q(1|0) \\
 P[Y = 0|X = 1] &\equiv Q(0|1) & P[Y = 1|X = 1] &\equiv Q(1|1)
 \end{aligned}$$

3.6. For a binary symmetric channel (BSC) defined in 3.2, we now have three equivalent ways to specify the relevant probabilities:



Definition 3.7. When the alphabets are collections of integers, we usually write $p(x)$ and $q(y)$ as p_x and q_y respectively.

Alternatively, when the members of the alphabets are explicitly indexed (as x_1, x_2, \dots and y_1, y_2, \dots), we often define

$$p_i \equiv p(x_i) \quad \text{and} \quad q_j = q(y_j).$$

Example 3.8. Suppose, for a DMC, we have $\mathcal{X} = \{x_1, x_2\}$ and $\mathcal{Y} = \{y_1, y_2, y_3\}$. Then, its probability transition matrix \mathbf{Q} is of the form

$$\mathbf{Q} = \begin{bmatrix} Q(y_1|x_1) & Q(y_2|x_1) & Q(y_3|x_1) \\ Q(y_1|x_2) & Q(y_2|x_2) & Q(y_3|x_2) \end{bmatrix}.$$

You may wonder how this \mathbf{Q} happens in physical system. Let's suppose that the input to the channel is binary; hence, $\mathcal{X} = \{0, 1\}$ as in the BSC. However, in this case, after passing through the channel, some bits can be lost¹¹ (rather than corrupted). In such case, we have three possible outputs of the channel: 0, 1, e where the "e" represents the case in which the bit is erased by the channel.

Example 3.9. Consider a DMC whose samples of input and output are provided below

x: 1 1 1 1 1 1 1 1 1 1 0 0 1 1 1 0 1 1 0 1
 y: 1 3 2 2 1 2 1 2 2 3 1 1 1 3 1 3 2 3 1 2

Estimate its input probability vector $\underline{\mathbf{p}}$, output probability vector $\underline{\mathbf{q}}$, and \mathbf{Q} matrix.

$$\underline{\mathbf{p}} = [p(0) \quad p(1)] \approx \begin{bmatrix} \frac{4}{20} & \frac{16}{20} \end{bmatrix} = [0.2 \quad 0.8]$$

$$\underline{\mathbf{q}} = [q(1) \quad q(2) \quad q(3)] \approx \begin{bmatrix} \frac{8}{20} & \frac{7}{20} & \frac{5}{20} \end{bmatrix} = [0.4 \quad 0.35 \quad 0.25]$$

$$\mathbf{Q} = \begin{matrix} & \begin{matrix} 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 0 \\ 1 \end{matrix} & \begin{bmatrix} \frac{3}{4} & \frac{0}{4} & \frac{1}{4} \\ \frac{5}{16} & \frac{7}{16} & \frac{4}{16} \end{bmatrix} \end{matrix}$$

3.10. Observe that the sum along any row of the \mathbf{Q} matrix is 1.

- This is different from the \mathbf{P} matrix (the joint probability matrix) that was the main focus in basic probability class. Recall that, for \mathbf{P} matrix, the sum of *all* elements in the matrix is 1.

◦ See 3.15 for more discussion about the \mathbf{P} matrix.

3.11. The channel matrix \mathbf{Q} is often defined or visualized in the form of the **channel diagram** as shown in Figure 8. Note that each arrow should be labeled with the transition probability $Q(y|x)$. See also Example 3.12.

¹¹The receiver knows which bits have been erased.

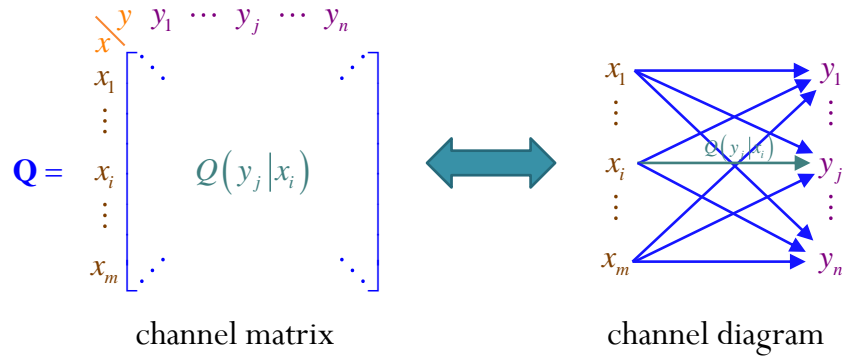


Figure 8: Conversion between the channel matrix and the channel diagram.

Example 3.12. The channel diagram for a channel whose

$$\mathcal{X} = \{0, 1\}, \quad \mathcal{Y} = \{1, 2, 3\}, \quad \underline{\mathbf{p}} = [0.2 \ 0.8], \quad \text{and} \quad \mathbf{Q} = \begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.4 & 0.3 \end{bmatrix}$$

is shown in Figure 9.

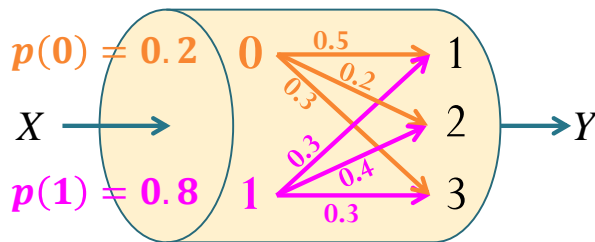


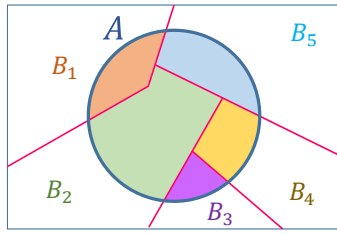
Figure 9: Channel diagram for Example 3.12.

3.13. Knowing the input probability vector $\underline{\mathbf{p}}$ and the channel (probability transition) matrix \mathbf{Q} , we can calculate the output probabilities $\underline{\mathbf{q}}$ from

$$\underline{\mathbf{q}} = \underline{\mathbf{p}}\mathbf{Q} \tag{5}$$

To see this, recall the **total probability theorem**: If a (finite or infinitely) countable collection of events $\{B_1, B_2, \dots\}$ is a partition of Ω , then

$$P(A) = \sum_i P(A \cap B_i) = \sum_i P(A|B_i)P(B_i). \tag{6}$$



$$P(A) = P(A \cap B_1) + P(A \cap B_2) + P(A \cap B_3) + P(A \cap B_4) + P(A \cap B_5)$$

For us, event A is the event $[Y = y]$. Applying this theorem to our variables, we get

$$\begin{aligned} q(y) &= P[Y = y] = \sum_x P[X = x, Y = y] \\ &= \sum_x P[Y = y|X = x] P[X = x] = \sum_x Q(y|x)p(x). \end{aligned}$$

This calculation, illustrated below, is exactly the same as the matrix multiplication calculation performed to find each element of $\underline{\mathbf{q}}$:

$$\underline{\mathbf{q}} = \begin{bmatrix} \cdots & q(y_j) & \cdots \end{bmatrix} = \underline{\mathbf{p}} \mathbf{Q} = \underbrace{\begin{bmatrix} & \end{bmatrix}}_{\underline{\mathbf{p}}} \underbrace{\begin{bmatrix} \phantom{Q_{j0}} \\ \phantom{Q_{j1}} \end{bmatrix}}_{\mathbf{Q}}$$

Example 3.14. For a binary symmetric channel (BSC) defined in 3.2,

$$\begin{aligned} q(0) &= P[Y = 0] = P[Y = 0, X = 0] + P[Y = 0, X = 1] \\ &= P[Y = 0|X = 0] P[X = 0] + P[Y = 0|X = 1] P[X = 1] \\ &= Q(0|0)p(0) + Q(0|1)p(1) \\ &= \begin{bmatrix} p(0) & p(1) \end{bmatrix} \begin{bmatrix} Q(0|0) \\ Q(0|1) \end{bmatrix} = \underline{\mathbf{p}} \begin{bmatrix} Q(0|0) \\ Q(0|1) \end{bmatrix} \\ &= (1-p) \times p_0 + p \times p_1 \\ q(1) &= P[Y = 1] = P[Y = 1, X = 0] + P[Y = 1, X = 1] \\ &= P[Y = 1|X = 0] P[X = 0] + P[Y = 1|X = 1] P[X = 1] \\ &= Q(1|0)p(0) + Q(1|1)p(1) \\ &= \begin{bmatrix} p(0) & p(1) \end{bmatrix} \begin{bmatrix} Q(1|0) \\ Q(1|1) \end{bmatrix} = \underline{\mathbf{p}} \begin{bmatrix} Q(1|0) \\ Q(1|1) \end{bmatrix} \\ &= p \times p_0 + (1-p) \times p_1 \quad 32 \end{aligned}$$

Therefore,

$$\begin{aligned} \underline{\mathbf{q}} &= [q(0) \quad q(1)] = \left[\underline{\mathbf{p}} \begin{bmatrix} Q(0|0) \\ Q(0|1) \end{bmatrix} \quad \underline{\mathbf{p}} \begin{bmatrix} Q(1|0) \\ Q(1|1) \end{bmatrix} \right] \\ &= \underline{\mathbf{p}} \begin{bmatrix} Q(0|0) & Q(1|0) \\ Q(0|1) & Q(1|1) \end{bmatrix} = \underline{\mathbf{p}}\mathbf{Q} \end{aligned}$$

3.15. Recall, from ECS315, that there is another matrix called the **joint probability matrix \mathbf{P}** . This is the matrix whose elements give the joint probabilities $P_{X,Y}(x, y) = P[X = x, Y = y]$:

$$x \begin{bmatrix} \cdots & \vdots & \cdots \\ \cdots & P[X = x, Y = y] & \cdots \\ \cdots & \vdots & \cdots \end{bmatrix}.$$

Recall also that we can get $p(x)$ by adding the elements of \mathbf{P} in the row corresponding to x . Similarly, we can get $q(y)$ by adding the elements of \mathbf{P} in the column corresponding to y .

By definition, the relationship between the conditional probability $Q(y|x)$ and the joint probability $P_{X,Y}(x, y)$ is

$$\begin{aligned} Q(y|x) &\equiv P[\underbrace{Y=y}_A | \underbrace{X=x}_B] \\ &= P(A|B) \equiv \frac{P(A \cap B)}{P(B)} \end{aligned} \quad Q(y|x) = \frac{P_{X,Y}(x, y)}{p(x)}.$$

Equivalently,

$$P_{X,Y}(x, y) = p(x)Q(y|x).$$

Therefore, to get the matrix \mathbf{P} from matrix \mathbf{Q} , we need to multiply each row of \mathbf{Q} by the corresponding $p(x)$. This is illustrated in Figure 10. The same calculation could be done easily in MATLAB by first constructing a diagonal matrix from the elements in $\underline{\mathbf{p}}$ and then multiply this to the matrix \mathbf{Q} :

$$\mathbf{P} = (\text{diag}(\underline{\mathbf{p}})) \mathbf{Q}. \quad (7)$$

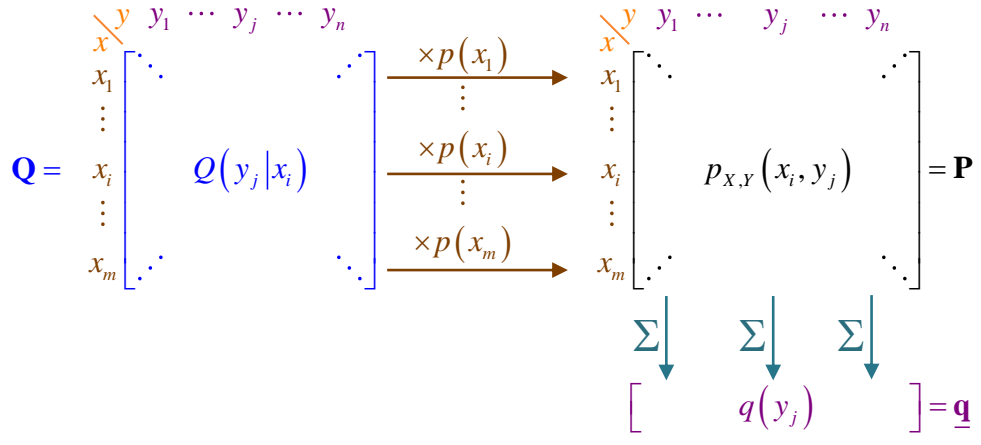


Figure 10: Conversion from the \mathbf{Q} matrix to the \mathbf{P} matrix and the output probability vector \underline{q} .

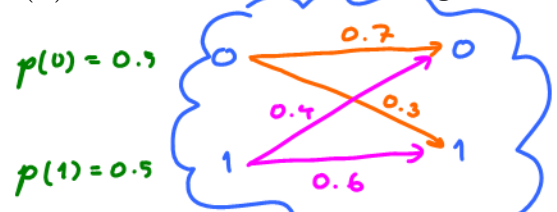
Once the \mathbf{P} matrix is obtained, we can calculate the output probability vector \underline{q} by adding the elements of \mathbf{P} along each column; this gives

$$\underline{q} = [1 \ 1 \ \dots \ 1] \mathbf{P} = [1 \ 1 \ \dots \ 1] \text{diag}(\underline{p}) \mathbf{Q} = \underline{p} \mathbf{Q}.$$

Example 3.16. Binary Asymmetric Channel (BAC): Consider a binary input-output channel whose matrix of transition probabilities is

$$\mathbf{Q} = \begin{matrix} & \begin{matrix} 0 & 1 \end{matrix} \\ \begin{matrix} 0 \\ 1 \end{matrix} & \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \end{matrix}.$$

(a) Draw the channel diagram.



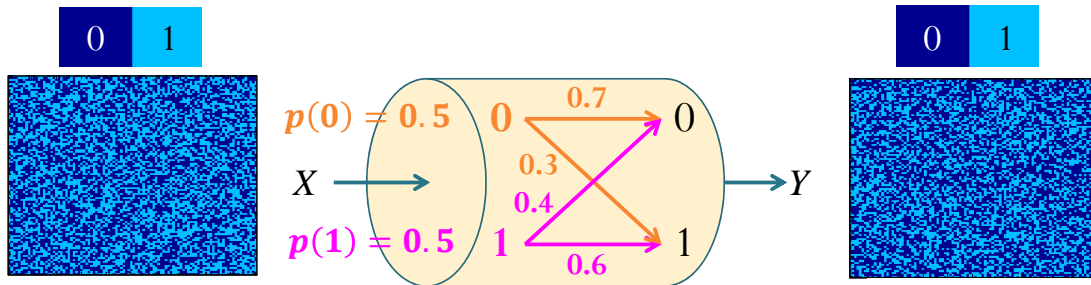
(b) If the two inputs are equally likely, find the corresponding output probability vector \underline{q} and the joint probability matrix \mathbf{P} for this channel.

$$\mathbf{Q} = \begin{matrix} & \begin{matrix} 0 & 1 \end{matrix} \\ \begin{matrix} 0 \\ 1 \end{matrix} & \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \end{matrix} \xrightarrow{\begin{matrix} \times 0.5 \\ \times 0.5 \end{matrix}} \mathbf{P} = \begin{bmatrix} 0.35 & 0.15 \\ 0.20 & 0.30 \end{bmatrix}$$

$$\underline{q} = [0.55 \ 0.45]$$

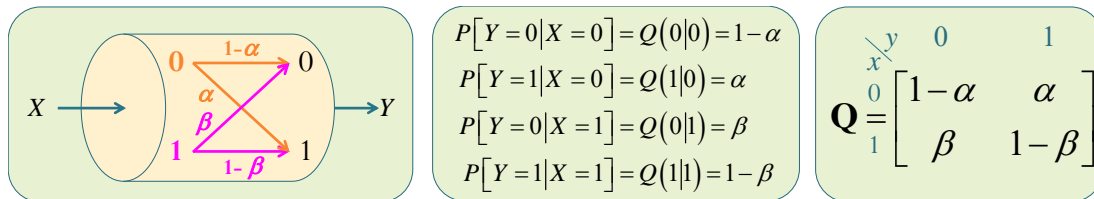
Alternatively,

$$\underline{q} = \underline{p} \mathbf{Q} = [0.5 \ 0.5] \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} = [0.55 \ 0.45]$$



[18, Ex. 11.3]

Example 3.17 (General BAC). Similar to Example 3.6 where we have three equivalent ways to specify BSC. We also have three different ways to describe a general BAC:



Example 3.18. Find the output probability vector \mathbf{q} and the joint probability matrix \mathbf{P} for the DMC defined in Example 3.12:

