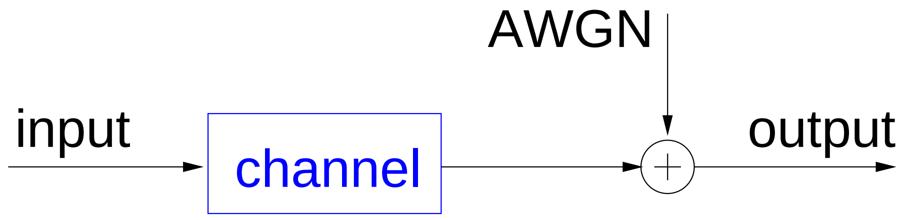
Revision of Lecture 4

- We have completed studying digital sources from information theory viewpoint
  - We have learnt all fundamental principles for source coding, provided by information theory
  - Practical source coding is guided by information theory, with practical constraints, such as performance and processing complexity/delay trade off
  - When you come to practical source coding part, you can smile – as you should know everything

- Information theory does far more, in fact it provides guiding principle for everything in communications
  - For example, **what happens to information transmitted through channel**?

Information Across Channels

- For conveying information across a transmission channel, a suitable model is:

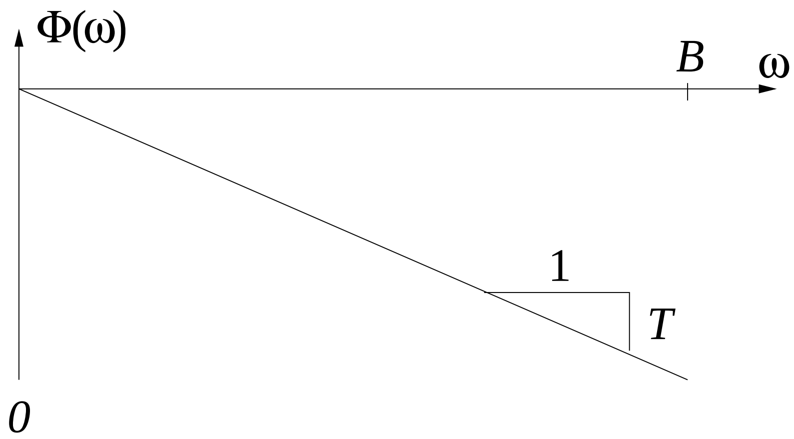


- The channel itself introduces amplitude and phase distortion, is potentially time-varying, and has a **limited bandwidth** B
- Error-free reception of symbols is additionally impeded by additive white Gaussian noise (AWGN); it severeness is described by the **signal-to-noise ratio** (SNR)
- Therefore, dependent on the above parameters, we are interested in determining the maximum possible error-free information transmission (channel capacity C)
- We will see that C depends on B and SNR

Characteristics of Channel

- The channel can be described by its impulse response $h(t)$ or equivalently its frequency response $H(j\omega) = A(\omega) \cdot e^{j\Phi(\omega)}$ with amplitude response $A(\omega)$ and phase response $\Phi(\omega)$; $h(t)$ and $H(j\omega)$ are Fourier pair

- **Ideal channel** (pure delay): $h(t) = \delta(t - T) \quad \rightarrow \quad A(\omega) = 1, \quad \Phi(\omega) = -\omega T$



  - Flat magnitude and linear phase ($=$ constant group delay $G(\omega) = -\partial\Phi(\omega)/\partial\omega$)
  - The only impairment caused by an ideal channel is AWGN

- **Non-ideal channel**: channel is *dispersive*, causing intersymbol interference

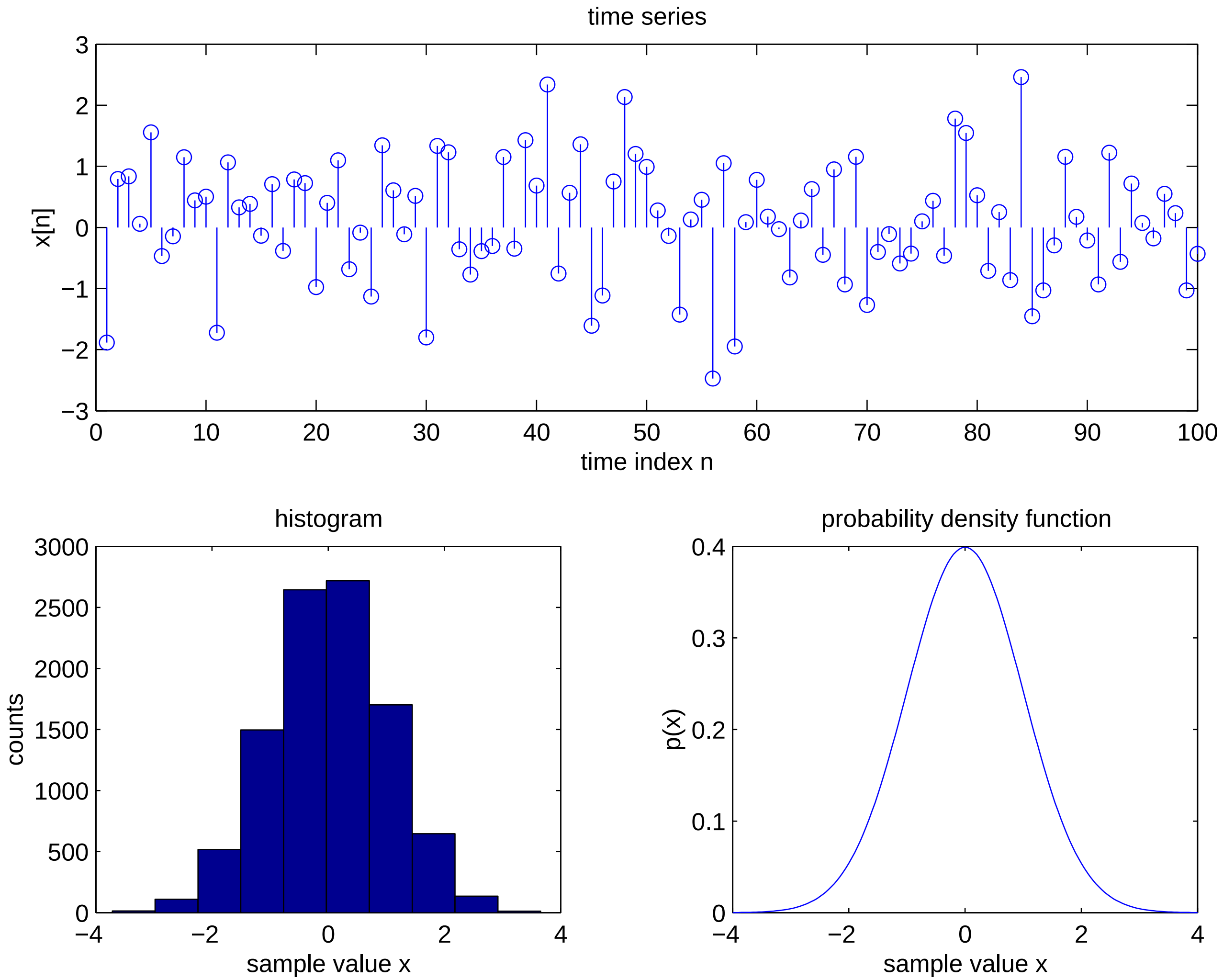
Additive White Gaussian Noise

- Noise is uncorrelated with the signal
- Gaussian noise has a bell shaped probability density function (normally distributed)

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)/2\sigma^2}$$

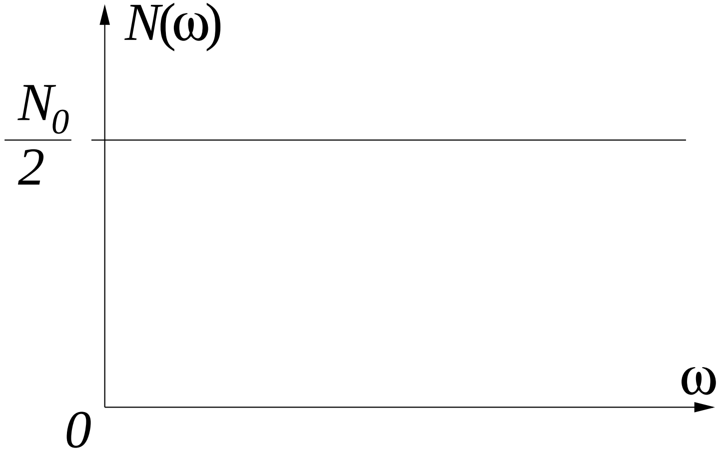
with mean μ and variance σ^2

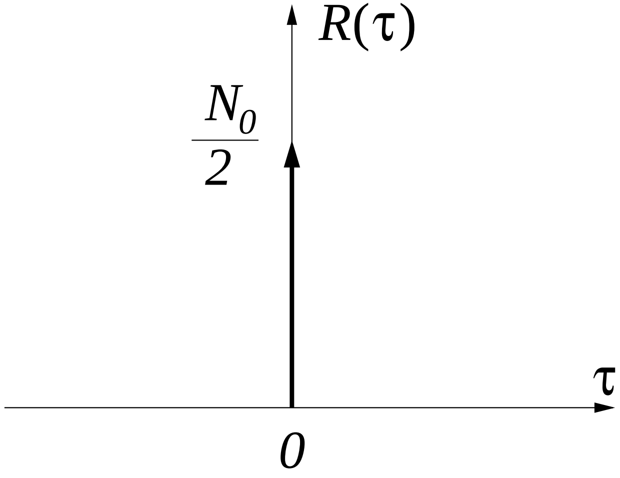
- White noise has zero mean, and channel noise is usually modelled as an AWGN



White Noise

- **White noise** is characterised by a **flat** power spectral density function, $N(\omega)$, or equivalently, its **impulse-shaped** auto-correlation function, $R(\tau)$





- $N(\omega)$ and $R(\tau)$ are a Fourier pair:

$$R(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} N(\omega)e^{j\omega\tau}d\omega \qquad N(\omega) = \int_{-\infty}^{\infty} R(\tau)e^{-j\omega\tau}d\tau$$

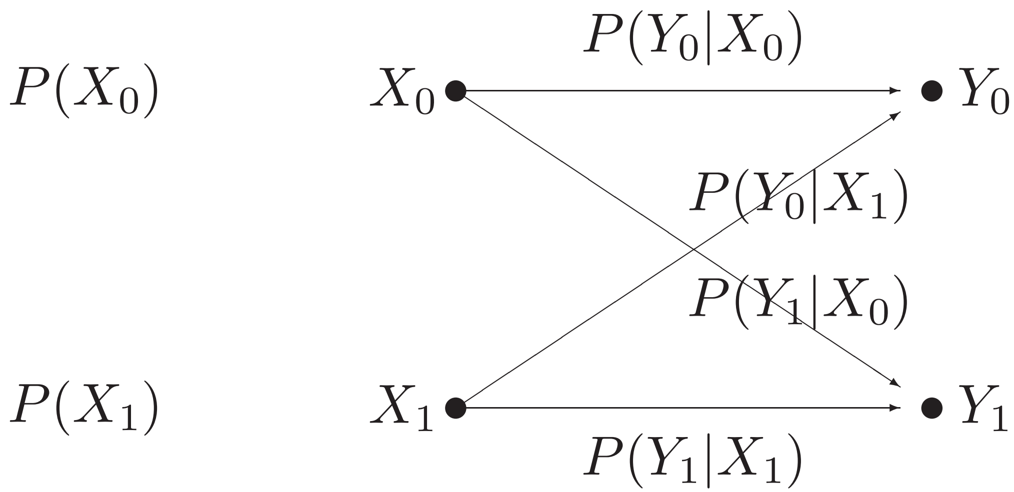
- Two-side spectrum is usually used for convenience, and N_0 is the noise power

Physic Basis of Channel

- Transmitted signal is amplified to required power level and launched from transmit antenna to channel
  - Signal power is attenuated, as it travels in distance – path loss
  - Copies of signal arrive at receiver with different attenuation and delays, which may cause dispersive and fading (power level fluctuates rapidly) effects
- Received signal at receiver is very weak, and needs to be amplified to required power level in order to detect digital information contained
  - While receiver amplifier amplifies receive signal, it also introduces thermal noise
  - AWGN in our channel model in fact models this noise
  - How much noise introduced by amplifier is specified by its noise figure
- Depending on communication carrier frequency, channel bandwidth and actual communication conditions, channel may be modelled as:
  - AWGN channel, i.e. channel is nondispersive or memoryless
  - Or dispersive channel, i.e. channel has memory

Binary Symmetric Channel (BSC)

- **BSC** is the simplest model for information transmission via a discrete channel (channel is ideal, no amplitude and phase distortion, only distortion is due to AWGN):



$P(X_i)$: probability of occurrence of X_i at source output, *a priori* probability

$P(Y_j|X_i)$: conditional probability of Rx Y_j given Tx X_i

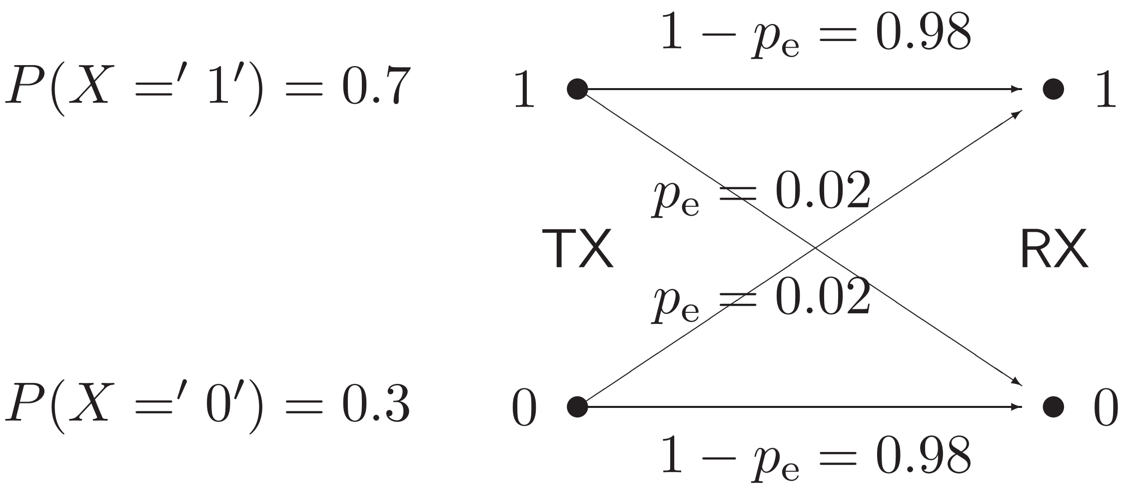
"Symmetric" refers to $P(Y_1|X_0) = P(Y_0|X_1) = p_e$, p_e being channel error probability

- The joint probability $P(Y_j, X_i)$ (Tx X_i and Rx Y_j) is linked with the conditional probabilities $P(Y_j|X_i)$ by Bayes' rule:

$$
\begin{aligned}
P(Y_j, X_i) &= P(X_i) \cdot P(Y_j|X_i) = P(Y_j) \cdot P(X_i|Y_j) \\
&= P(X_i, Y_j)
\end{aligned}
$$

Binary Symmetric Channel – Example

- Consider a BSC:



- This has a non-equiprobable source with $P(X='1') = 0.7$ and $P(X='0') = 0.3$: on average, 70% of transmitted bits are '1' and 30% are '0'

- Channel's error probability $p_e = 0.02$: on average, bit error rate is 2%

Binary Symmetric Channel – Example (continue)

- Probability of correct reception: $P_{\text{correct}} = P(Y =' 1', X =' 1') + P(Y =' 0', X =' 0') = 0.98$, as

$$\begin{aligned} P(Y =' 1', X =' 1') &= P(X =' 1') \cdot P(Y =' 1' | X =' 1') = 0.7 \cdot 0.98 = 0.686 \\ P(Y =' 0', X =' 0') &= P(X =' 0') \cdot P(Y =' 0' | X =' 0') = 0.3 \cdot 0.98 = 0.294 \end{aligned}$$

- Probability of erroneous reception: $P_{\text{error}} = P(Y =' 1', X =' 0') + P(Y =' 0', X =' 1') = 0.02$, as

$$\begin{aligned} P(Y =' 1', X =' 0') &= P(X =' 0') \cdot P(Y =' 1' | X =' 0') = 0.3 \cdot 0.02 = 0.006 \\ P(Y =' 0', X =' 1') &= P(X =' 1') \cdot P(Y =' 0' | X =' 1') = 0.7 \cdot 0.02 = 0.014 \end{aligned}$$

- Total probability of receiving a '1' (or a '0'):

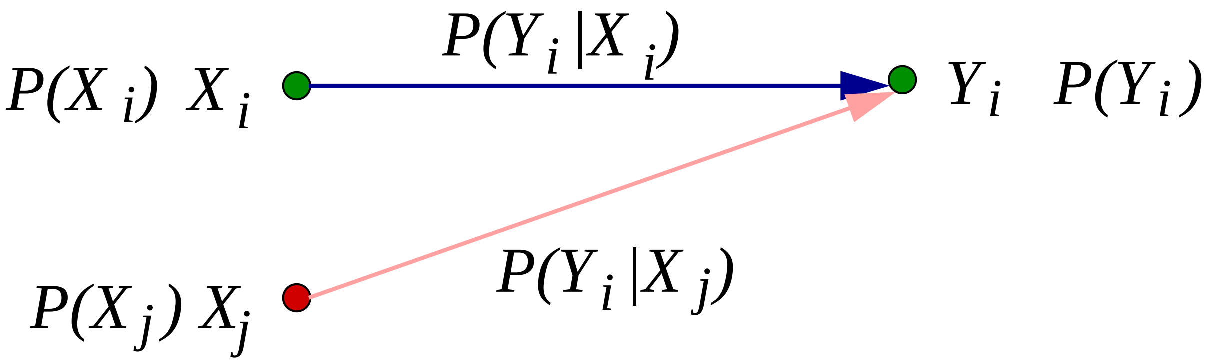
$$\begin{aligned} P(Y =' 1') &= P(X =' 1') \cdot P(Y =' 1' | X =' 1') + P(X =' 0') \cdot P(Y =' 1' | X =' 0') \\ &= 0.7 \cdot 0.98 + 0.3 \cdot 0.02 = 0.692 \\ P(Y =' 0') &= P(X =' 0') \cdot P(Y =' 0' | X =' 0') + P(X =' 1') \cdot P(Y =' 0' | X =' 1') \\ &= 0.3 \cdot 0.98 + 0.7 \cdot 0.02 = 0.308 \end{aligned}$$

A Close Look at Channel

- For above (binary symmetric channel) example

- Something happens in channel to "information"
  - How to describe this ?

- Y_i and X_j are connected $\rightarrow$ They have something in common or "mutual"



Mutual Information

- Definition of **mutual information** of X_i and Y_j:

$$I(X_i, Y_j) = \log_2 \frac{P(X_i|Y_j)}{P(X_i)} \qquad \text{(bits)}$$

- Perfect, noiseless channel: $Y_i = X_i$, i.e. $P(X_i|Y_i) = 1$ and

$$I(X_i, Y_i) = \log_2 \frac{1}{P(X_i)}$$

  – This is the information of X_i, hence no information is lost in the channel

- Extremely noisy channel with error probability $0.5 \rightarrow Y_i$ is independent of X_i, hence

$$P(X_i|Y_i) = \frac{P(X_i, Y_i)}{P(Y_i)} = \frac{P(X_i) \cdot P(Y_i)}{P(Y_i)}$$

  – Therefore $I(X_i, Y_i) = \log_2 1 = 0$, meaning all information is lost in the channel

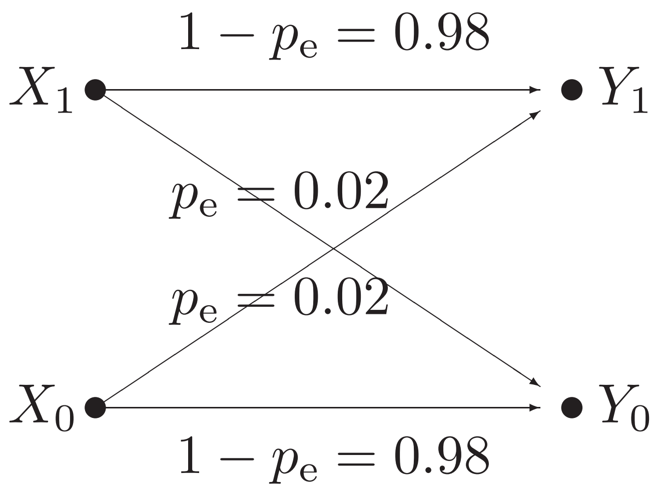
- In general, $I(X_i) > I(X_i, Y_i)$, some information is lost in the channel

Mutual Information – Example

- Consider the earlier BSC example:

$P(X_1) = 0.7$

$P(X_0) = 0.3$



$P(X_1|Y_1) = 0.9913$

$P(X_0|Y_1) = 0.0087$

$P(X_1|Y_0) = 0.0455$

$P(X_0|Y_0) = 0.9545$

- Here, the mutual information results in:

	$I(X_1, Y_1) = 0.502$ bits	$I(X_0, Y_0) = 1.670$ bits
(source info:	$I(X_1) = 0.515$ bits	$I(X_0) = 1.737$ bits)
(destin info:	$I(Y_1) = 0.531$ bits	$I(Y_0) = 1.699$ bits)
	$I(X_0, Y_1) = -5.113$ bits	$I(X_1, Y_0) = -3.945$ bits

  - Destination info contents are more balanced then source info contents
  - The negative quantities represent "mis-information"

Average Mutual Information

- Based on received symbols Y_j given transmitted symbols X_i through a BSC, average mutual information is defined as:

$$I(X,Y) = \sum_i \sum_j P(X_i,Y_j) \cdot I(X_i,Y_j)$$
$$= \sum_i \sum_j P(X_i,Y_j) \cdot \log_2 \frac{P(X_i|Y_j)}{P(X_i)} \quad \text{(bits/symbol)}$$

- This gives the average amount of **source information acquired per received symbol** by the receiver, and should be distinguished form the average source information (entropy $H(X)$)

- Note that due to Bayes:

$$\frac{P(X_i|Y_j)}{P(X_i)} = \frac{P(X_i,Y_j)}{P(X_i) \cdot P(Y_j)} = \frac{P(Y_j|X_i)}{P(Y_j)}$$

Imperfect Channel: Information Loss

- Consider re-arranging the mutual information between transmitted symbol X_i and received symbol Y_j:

$$\begin{aligned} I(X_i, Y_j) &= \log_2 \frac{P(X_i|Y_j)}{P(X_i)} = \log_2 \frac{1}{P(X_i)} - \log_2 \frac{1}{P(X_i|Y_j)} \\ &= I(X_i) - I(X_i|Y_j) \end{aligned}$$

  $I(X_i, Y_j)$ is the amount of information conveyed to receiver when transmitting X_i and receiving Y_j, $I(X_i)$ is the source information of X_i, and $I(X_i|Y_j)$ can be regarded as the information loss due to the channel

- Therefore,

$$\underbrace{I(X_i)}_{\text{Source Inf.}} - \underbrace{I(X_i, Y_j)}_{\text{Inf. conveyed to rec.}} = \underbrace{I(X_i|Y_j)}_{\text{Inf. loss}}$$

- $0 \le I(X_i|Y_j) \le I(X_i)$, see for example the previous cases of $p_e = 0$ and $p_e = 0.5$

Imperfect Channel: Average Mutual Information

- **Average mutual information** is given by:

$$I(X,Y) = \sum_i \sum_j P(X_i, Y_j) \cdot \log_2 \frac{P(X_i|Y_j)}{P(X_i)} \qquad \text{(bits/symbol)}$$

  – But this average conveyed information

$$I(X,Y) = \sum_i \sum_j P(X_i, Y_j) \cdot \log_2 \frac{1}{P(X_i)} - \sum_i \sum_j P(X_i, Y_j) \cdot \log_2 \frac{1}{P(X_i|Y_j)}$$

$$= \sum_i \left(\sum_j P(X_i, Y_j) \right) \cdot \log_2 \frac{1}{P(X_i)} - \sum_j P(Y_j) \cdot \left(\sum_i P(X_i|Y_j) \cdot \log_2 \frac{1}{P(X_i|Y_j)} \right)$$

$$= \sum_i P(X_i) \cdot \log_2 \frac{1}{P(X_i)} - \sum_j P(Y_j) \cdot I(X|Y_j) = H(X) - H(X|Y)$$

$$\underbrace{I(X,Y)}_{\text{av. converyed information}} = \underbrace{H(X)}_{\text{av. source information}} - \underbrace{H(X|Y)}_{\text{av. information lost}}$$

- A similar re-arrangement leads to:

$$\underbrace{I(X,Y)}_{\text{av. converyed information}} = \underbrace{H(Y)}_{\text{destination entropy}} - \underbrace{H(Y|X)}_{\text{error entropy}}$$

Summary

- General consideration for transferring information across channels
  - Channel characteristics or channel model
  - Binary symmetric channel $\rightarrow$ assumptions
- Mutual information between channel input $X(k) \in \{X_i\}$ and channel output $Y(k) \in \{Y_i\}$ characterises how information is transferring across channel
  - Average mutual information $I(X,Y)$

$$\underbrace{I(X,Y)}_{\text{av. conveyed information}} = \underbrace{H(X)}_{\text{source entropy}} - \underbrace{H(X|Y)}_{\text{av. information lost}}$$

$$\underbrace{I(X,Y)}_{\text{av. conveyed information}} = \underbrace{H(Y)}_{\text{destination entropy}} - \underbrace{H(Y|X)}_{\text{error entropy}}$$