Q2. a) Define entropy, joint entropy, mutual information, self information and obtain the condition for the maximum value of entropy.

Ans: 2.a) Self-information- Shannon derived a measure of information content called the self-information or "surprisal" of a message *m*:

$$I(m) = \log\left(\frac{1}{p(m)}\right) = -\log(p(m))$$

where $p(m) = \Pr(M = m)$ is the probability that message *m* is chosen from all possible choices in the message space M. The base of the logarithm only affects a scaling factor and, consequently, the units in which the measured information content is expressed. If the logarithm is base 2, the measure of information is expressed in units of bits. Information is transferred from a source to a recipient only if the recipient of the information did not already have the information to begin with. Messages that convey information that is certain to happen and already known by the recipient contain no real information. Infrequently occurring messages contain more information than more frequently occurring messages. This fact is reflected in the above equation - a certain message, i.e. of probability 1, has an information measure of zero. In addition, a compound message of two (or more) unrelated (or mutually independent) messages would have a quantity of information that is the sum of the measures of information of each message individually. That fact is also reflected in the above equation, supporting the validity of its derivation.

Entropy

The entropy of a discrete message space M is a measure of the amount of uncertainty one has about which message will be chosen. It is defined as the average self-information of a message m from that message space:

$$H(M) = \mathbb{E}\{I(M)\} = \sum_{m\in M} p(m)I(m) = -\sum_{m\in M} p(m)\log p(m).$$

where $\mathbb{E}\{\}$ denotes the expected value operation.

An important property of entropy is that it is maximized when all the messages in the message space are equiprobable (e.g. $p(m) = 1/M$). In this case $H(M) = \log|M|$.

Sometimes the function *H* is expressed in terms of the probabilities of the distribution:

$$H(p_1, p_2, \ldots, p_k) = -\sum_{i=1}^{k} p_i \log p_i,$$ where each $p_i \geq 0$ and $\sum_{i=1}^{k} p_i = 1$.

An important special case of this is the binary entropy function:

$$H_{\mathrm{b}}(p) = H(p, 1-p) = -p\log p - (1-p)\log(1-p).$$

Joint entropy

The joint entropy of two discrete random variables X and Y is defined as the entropy of the joint distribution of X and Y:

$$H(X,Y) = \mathbb{E}_{X,Y}[-\log p(x,y)] = -\sum_{x,y} p(x,y)\log p(x,y)$$

If X and Y are independent, then the joint entropy is simply the sum of their individual entropies.

Mutual information (transinformation)

It turns out that one of the most useful and important measures of information is the mutual information, or transinformation. This is a measure of how much information can be obtained about one random variable by observing another. The mutual information of X relative to Y (which represents conceptually the average amount of information about X that can be gained by observing Y) is given by:

$$I(X;Y) = \sum_{y\in Y} p(y)\sum_{x\in X} p(x|y)\log\frac{p(x|y)}{p(x)} = \sum_{x,y} p(x,y)\log\frac{p(x,y)}{p(x)\,p(y)}.$$

A basic property of the mutual information is that:

$$I(X;Y) = H(X) - H(X|Y).$$

That is, knowing *Y*, we can save an average of $I(X;Y)$ bits in encoding *X* compared to not knowing *Y*. Mutual information is symmetric:

$$I(X;Y) = I(Y;X) = H(X) + H(Y) - H(X,Y).$$

Mutual information can be expressed as the average Kullback–Leibler divergence (information gain) of the posterior probability distribution of *X* given the value of *Y* to the prior distribution on *X*:

$$I(X;Y) = \mathbb{E}_{p(y)}\{D_{\mathrm{KL}}(p(X|Y=y)\|p(X))\}.$$

In other words, this is a measure of how much, on the average, the probability distribution on *X* will change if we are given the value of *Y*. This is often recalculated as the divergence from the product of the marginal distributions to the actual joint distribution:

$$I(X;Y) = D_{\mathrm{KL}}(p(X,Y)\|p(X)p(Y)).$$

Mutual information is closely related to the log-likelihood ratio test in the context of contingency tables and the multinomial distribution and to Pearson's χ^2 test: mutual information can be considered a statistic for assessing independence between a pair of variables, and has a well-specified asymptotic distribution.

b) Consider the random variable:

X =

X1	x2	x3	x4	x5	x_6	x_7
0.49	0.26	0.12	0.04	0.04	0.04	0.02

(i) Find a binary Huffman code for X.

Ans. The Huffman tree for this distribution is:

Codeword								
1	x1	0:49	0:49	0:49	0:49	0:49	0:51	1
00	x2	0:26	0:26	0:26	0:26	0:26	0:49	
011	x3	0:12	0:12	0:12	0:13	0:25		
01000	x4	0:04	0:05	0:08	0:12			
01001	x5	0:04	0:04	0:05				
01010	x6	0:03	0:04					
01011	x7	0:02						

(ii) Find the expected codelength for this encoding.

Ans. The expected length of the codeword for the binary Hu_man code E(X)=2.02 bits.

(iii) Find a ternary Huffman code for X.

Ans.

Codeword					
0	x1	0:49	0:49	0:49	1
1	x2	0:26	0:26	0:26	
20	x3	0:12	0:12	0:25	
22	x4	0:04	0:09		
210	x5	0:04	0:04		
211	x6	0:03			
212	x7	0:02			

3. (a) On the basis of your study on Sampling Process throw some light on the Practical Aspects of Sampling and Signal Recovery.

Ans: Page no. 154 of Textbook 'Digital Communications' by Siman Haykin

(b) A digital communication system is to carry a single voice signal using linearly quantised PCM. what PCM bit rate will be required if an ideal anti-aliasing filter with a cut-off frequency of 3.4KHz is used at the transmitter and the signal to quantization noise ratio is to be kept above 50dB.

Ans:

$$SN_qR = 4.8 + 6n - \alpha_{dB}$$

For voice signals $\alpha = 10$ dB, i.e.:

$$n = \frac{50 + 10 - 4.8}{6} = 9.2$$

10 bit/sample are therefore required. The sampling rate required is given by Nyquist's rule, $f_s \geq 2f_H$. Taking a practical version of the sampling theorem, equation (5.3), gives:

$$f_s = 2.2 \times 3.4 \text{ kHz} = 7.48 \text{ kHz (or k samples/s)}$$

The PCM bit rate (or more strictly binary baud rate) is therefore:

$$R_b = f_s n$$
$$= 7.48 \times 10^3 \times 10 \text{ bit/s}$$
$$= 74.8 \text{ kbit/s}$$

3(a) On the basis of your study on Sampling Process throw some light on the Practical aspects of sampling and signal recovery.

Ans: **Article 4.5 on Pg 154 of Textbook 'Digital Communications' by Simon Haykin.**

4. (a) List the main attributes of a speech codec.

Ans: The following are the speech codec attributes.

Transmission Bit Rate
Since the speech codec shares the communications channel with other data, the peak bit rate should be as low as possible so as not to use a disproportionate share of the channel. The codecs below 64 kbps are primarily developed to increase the capacity of circuit multiplication equipment used for narrow bandwidth links. For the most part, they are fi xed bit rate codecs, meaning they operate at the same rate regardless of the input. In the variable bit rate codecs, network loading and voice activity determine the instantaneous rate assigned to a particular voice channel.
Any of the fi xed rate speech codecs can be combined with a voice activity detector (VAD) and made into a simple two-state variable bit rate system. The lower rate could be either zero or some low rate needed to characterize slowly changing background noise characteristics. Either way, the bandwidth of the communications channel is only used for active speech.

Delay
The delay of a speech codec can have a great impact on its suitability for a particular application. For a one-way delay of conversation greater than 300 ms, the conversation becomes more like a half-duplex or push-to-talk experience, rather than an ordinary conversation. The components of total system delay include: (1) frame size, look ahead, multiplexing delay, (2) processing delay for computations, and (3) transmission delay.
Most low bit rate speech codecs process a frame of speech data at a time. The speech parameters are updated and transmitted for every frame. In addition, to analyze the data properly it is sometimes necessary to analyze data beyond the frame boundary; hence, before the speech can be analyzed it is necessary to buffer a frame's worth of data. The resulting delay is referred to as algorithmic delay. This delay component cannot be reduced by changing the implementation, but all other delay components can. The second major contribution for delay comes from the time taken by the encoder to analyze the speech and the decoder to reconstruct the speech. This part of the delay is referred to as processing delay. It depends on the speed of the hardware used to implement the coder. The sum of the algorithmic and processing delays is called the one-way codec delay. The third component of delay is due to transmission. It is the time taken for an entire frame of data to be transmitted from the encoder to the decoder. The total of the three delays is the oneway system delay. In addition, frame interleaving delay adds an additional frame delay to the total transmission delay. Frame interleaving is necessary to combat channel fading, and is part of the channel coding process.

Complexity
Speech codecs are implemented on special purpose hardware, such as digital signal processing (DSP) chips. DSP attributes are the computing speed in millions of instructions per second (MIPS), random access memory (RAM), and read only memory (ROM). For a speech codec, the system designer makes a choice about how much of these resources are to be allocated to the speech codec. Speech codecs using less than 15 MIPS are considered low-complexity codecs; those requiring 30 MIPS or more are thought of as high-complexity codecs. More complexity results in higher costs and greater power usage; for portable applications, greater power usage means reduced time between battery recharges or use of larger batteries, which means more expense and weight.

Quality
Of all the attributes, quality has the most dimensions. In many applications there are large amounts of background noise (car noise, street noise, offi ce noise, etc.). How well does the codec perform under these adverse conditions? What happens when there are channel errors during transmission? Are the errors detected or undetected? If undetected, the codec must perform even more robustly than when it is informed that entire frames are in error. How good does the codec sound when speech is encoded and decoded twice? All these questions must be carefully evaluated during the testing phase of a speech codec. The speech quality is often based on the fi ve-point MOS scale as defi ned by the International Telecommunication Union-Technical (ITU-T).

(b) Specify a baseband Nyquist channel which has a piecewise linear amplitude response, an absolute bandwidth of 10 KHz, and its appropriate for a baud rate of 16Kbaud. What is the channel's excess bandwidth?

Ans:

The cut-off frequency of the parent rectangular frequency response is given by:

$$f_\chi = R_s/2 = 16 \times 10^3/2 = 8 \times 10^3 \text{ Hz}$$

The simplest piecewise linear roll-off therefore starts at 8 − 2 = 6 kHz, is 6 dB down at f_χ = 8 kHz and is zero (−∞ dB down) at 8 + 2 = 10 kHz. (The amplitude response, below the start of roll-off, is flat and the phase is linear.) Thus:

$$|H_N(f)| = \begin{cases} 1.0, & |f| < 6000 \text{ (Hz)} \\ 2.5 - 0.25 \times 10^{-3} f, & 6000 \le |f| \le 10000 \text{ (Hz)} \\ 0, & |f| > 10000 \text{ (Hz)} \end{cases}$$

The channel's excess bandwidth is 10 kHz −8 kHz = 2 kHz.

5. (a) **(i)** The BPSK modulation is used in a channel that adds white noise with single-sided PSD $N_0 = 10^{-10}$ W/Hz. Calculate the amplitude *A* of the carrier signal to give $Pe = 10^{-6}$ for a data rate of 100 kbps.

Ans:

$$\therefore A = \left[2N_0\left(\frac{E_b}{N_0}\right)R\right]^{1/2}$$

$$A = [2 \times 10^{-10} \times 11.32 \times 10^5]^{1/2} = 1.505\,\text{mv}$$

(ii) Find E_b/N_0 in dB to provide $P_e = 10^{-6}$ for BPSK and coherent FSK.

Ans:

For BPSK

$$P_e \approx \frac{e^{-E_b/N_0}}{2\sqrt{\pi\left(\frac{E_b}{N_0}\right)}}$$

$$10^{-6} = \frac{e^{-\xi}}{2\sqrt{\pi\xi}}$$

where:

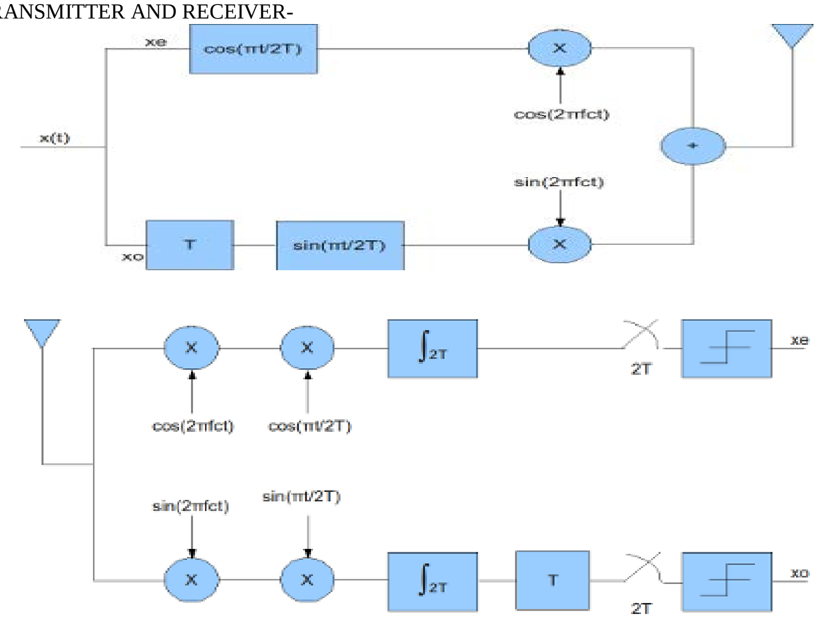
$\xi = E_b/N_0$

$$\therefore \xi = 11.32 = 10.54\,\text{dB}$$

FSK requires 3 dB more in terms of E_b/N_0 to give the same P_e as BPSK, i.e., 13.54 dB.

(b) What is a CPFSK modulation scheme. How it is related with the MSK modulation scheme. Explain MSK transmission and reception.

Ans: MSK TRANSMITTER AND RECEIVER-



Q6(a) With the help of block diagram Elaborate Baseband binary data transmission system.

Ans. Fig. 6.5 on Page 244 of TextBook 'Digital Communications' by Simon Haykin

7. (a) Obtain the transfer function and impulse response of matched filter.

Ans:

Let $S(f)$ be the Fourier transform of $s(t)$. Paserval's theorem states that the average signal energy is

$$E = \int_{-\infty}^{\infty} s(t)^2 \, dt = \int_{-\infty}^{\infty} |S(f)|^2 \, df$$

The signal at the output of the filter is

$$s_o(t) = \int_{-\infty}^{\infty} S(f)H(f) \, e^{j2\pi ft} \, df$$

At time $t = t_0$, we let

$$A = |s_o(t)| = \left| \int_{-\infty}^{\infty} S(f)H(f) \, e^{j2\pi f t_0} \, df \right|$$

Since the input noise power spectral density is given by

$$G_n(f) = n_0/2$$

the noise power spectral density at the output of the filter is

$$G_{n_o}(f) = G_n(f) \, |H(f)|^2 = \frac{n_0}{2} \, |H(f)|^2$$

and the average output noise power is

$$N = \int_{-\infty}^{\infty} G_{n_o}(f) \, df = \frac{n_0}{2} \int_{-\infty}^{\infty} |H(f)|^2 \, df$$

At $t = t_0$, we have

$$\frac{A^2}{EN} = \frac{\left| \int_{-\infty}^{\infty} S(f)H(f)e^{j2\pi f t_0} \, df \right|^2}{\int_{-\infty}^{\infty} |S(f)|^2 \, df \, \frac{n_0}{2} \int_{-\infty}^{\infty} |H(f)|^2 \, df}$$

$$\left| \int_{-\infty}^{\infty} X(f)Y(f)df \right|^2 \le \int_{-\infty}^{\infty} |X(f)|^2 df \int_{-\infty}^{\infty} |Y(f)|^2 df$$

Equality holds when

$$Y(f) = KX^*(f)$$

where K is a constant and $X^*(f)$ is the complex conjugate of $X(f)$.

We can apply Schwarz's inequality in our matched-filter case by lett[ing] $X(f) = S(f)\, e^{j2\pi f t_0}$ and $Y(f) = H(f)$, when equation (27.9) may be written as

$$\left| \int_{-\infty}^{\infty} S(f)\, e^{j2\pi f t_0} H(f) df \right|^2 \le \int_{-\infty}^{\infty} |S(f)|^2 df \int_{-\infty}^{\infty} |H(f)|^2 df \qquad (27.11)$$

Equality holds when

$$\begin{aligned} H(f) &= K\,[S(f)\, e^{j2\pi f t_0}]^* \\ &= K\, S^*(f)\, e^{-j2\pi f t_0} \end{aligned} \qquad (27.12)$$

and equation (27.8) becomes

$$\frac{A^2}{EN} = \frac{2}{n_0}$$

$$\sqrt{\frac{2E}{n_0}} = \frac{A}{\sqrt{N}} \qquad (27.13)$$

Taking the inverse Fourier transform of $H(f)$ in equation (27.12), we get

$$h(t) = \int_{-\infty}^{\infty} K\, S^*(f)\, e^{j2\pi f(t - t_0)} df$$

For real $h(t)$, $S(f) = S^*(f)$. Thus

$$h(t) = \int_{-\infty}^{\infty} K\, S(f) e^{j2\pi f(t_0 - t)} df$$

$$h(t) = K\, s(t_0 - t)$$

(b) What is the sampling instant signal to noise ratio at the output of a filter matched to a triangular pulse of height 10mV and width 1msec if the noise at the input to the filter is white with a power spectral density of $10\text{nV}^2/\text{Hz}$.

Ans:

$$E_x = \int_0^{T_o} v^2(t)\, dt = \int_0^{\frac{T_o}{2}} v^2(t)\, dt + \int_{\frac{T_o}{2}}^{T_o} v^2(t)\, dt$$

$$= \int_0^{0.5\times10^{-3}} [20\,t]^2\, dt + \int_{0.5\times10^{-3}}^{1\times10^{-3}} [2\times10^{-2} - 20\,t]^2\, dt$$

$$= 400\left[\frac{t^3}{3}\right]_0^{0.5\times10^{-3}} + 4\times10^{-4}\left[\frac{t}{1}\right]_{0.5\times10^{-3}}^{1\times10^{-3}}$$

$$- 80\times10^{-2}\left[\frac{t^2}{2}\right]_{0.5\times10^{-3}}^{1\times10^{-3}} + 400\left[\frac{t^3}{3}\right]_{0.5\times10^{-3}}^{1\times10^{-3}}$$

$$= 0.33\times10^{-7}\ (\text{V}^2\ \text{s})$$

$$\frac{S}{N} = \frac{2E_x}{N_0} = \frac{2\times0.33\times10^{-7}}{10\times10^{-9}} = 6.67 = 8.2\ \text{dB}$$

8. (a) A baseband binary communication system transmits a positive rectangular pulse for digital ones and a negative triangular pulse for digital zeros. If the (absolute) widths, peak pulse voltages, and noise power spectral density at the input of an ideal correlation receiver. Find the probability of bit error.

Ans:

$$E_{s0} = 0.33 \times 10^{-7}\ \text{V}^2\text{s}$$

The energy in the rectangular pulse is:

$$E_{s1} = v^2 T_o = (10 \times 10^{-3})^2 \times 1 \times 10^{-3} = 1 \times 10^{-7}\ \text{V}^2\ \text{s}$$

$$\rho = \frac{1}{\sqrt{E_{s0}\, E_{s1}}} \int_0^{T} v_1(t)\, v_0(t)\, dt$$

$$= \frac{1}{\sqrt{0.33 \times 10^{-7} \times 1 \times 10^{-7}}} \int_0^{10^{-3}} \left[-10 \times 10^{-3}\, \Pi\left(\frac{t - 0.5 \times 10^{-3}}{10^{-3}} \right) \right.$$

$$\left. \times\, 10 \times 10^{-3}\, \Lambda\left(\frac{t - 0.5 \times 10^{-3}}{5 \times 10^{-4}} \right) \right] dt$$

$$= -1.74 \times 10^3 \times 2 \left[2 \times 10^3\, t^2/2 \right]_0^{0.5 \times 10^{-3}} = -0.87$$

$$\frac{\Delta V}{\sigma} = \left[\frac{2}{N_0} \left(E_{s1} + E_{s0} - 2\, \rho \sqrt{(E_{s1}\, E_{s0})} \right) \right]^{1/2}$$

$$= \left[\frac{2}{10 \times 10^{-9}} \left(0.33 \times 10^{-7} + 1.0 \times 10^{-7} - 2\,(-0.87) \sqrt{0.33 \times 10^{-14}}\right) \right]^{1/2}$$

$$= 6.83$$

$$P_e = \frac{1}{2} \left[1 - \text{erf}\, \frac{\Delta V}{2\sqrt{2}\,\sigma} \right]$$

$$= \frac{1}{2} \left[1 - \text{erf}\, \frac{6.83}{2\sqrt{2}} \right] = 3.2 \times 10^{-4}$$

(b) Consider an FH/MFSK system, let the PN generator be defined by a 20-stage linear feedback shift register with a maximal length sequence. Each state of register dictates a new centre frequency within the hopping band. The minimum step size between centre frequencies is 200Hz. The register clock rate is 2KHz. Assume that 8-ary FSK modulation is used and that the data rate is 1.2Kbits/sec.

(i) What is the hopping bandwidth.
(ii) What is the chip rate.
(iii) How many chips are there in there in each data symbol.
(iv) What is the processing gain.

Ans: The hopping bandwidth is

$W_{ss} = (2^{20} - 1)$ states $\times 200$ Hz $= 2.1 \times 10^8$ Hz.

(b) Chip rate = hop rate = 2000 chips/s

(c) Chips/symbol: $R_s = \frac{R}{k} = \frac{1200}{3} = 400$

$$\frac{2000 \text{ chips/s}}{400 \text{ symbols/s}} = 5 \text{ chips/symbol}$$

(d) Processing gain: $G_p = W_{ss}/R$

$$= \frac{2.1 \times 10^8 \text{ Hz}}{1200 \text{ bits/s}} = 175{,}000 = 52.4 \text{ dB}$$

Q9. Write a short note on:

(iii) M12 multiplexer

Ans: Article 5.8 on Page no. 222 of Textbook 'Digital Communications' by Siman Haykin