Disclaimer

- The material provided in this document is not my original work and is a summary of some one else’s work(s).
- A simple Google search of the title of the document will direct you to the original source of the material.
- I do not guarantee the accuracy, completeness, timeliness, validity, non-omission, merchantability or fitness of the contents of this document for any particular purpose.
-

Speech Coding

SAMPLING AND QUANTIZATION

Outline

- Introduction
- Sampling
- Scalar Quantization
  - Quantization Error
  - Uniform Quantizer
  - Optimum Quantizer
  - Logarithmic Quantizer
  - Adaptive Quantizer
  - Differential Quantizer

Introduction

- The digital conversion process can be split into sampling, which discretizes the continuous time, and quantization, which reduces the infinite range of the sampled amplitudes to a finite set of possibilities.

Sampling

- The sampled waveform can be represented by

$$s(n) = s_a(nT) \qquad -\infty < n < \infty$$

- The sampling theorem states that if a signal $s_a(t)$ has a band-limited Fourier transform $S_a(j\omega)$ given by

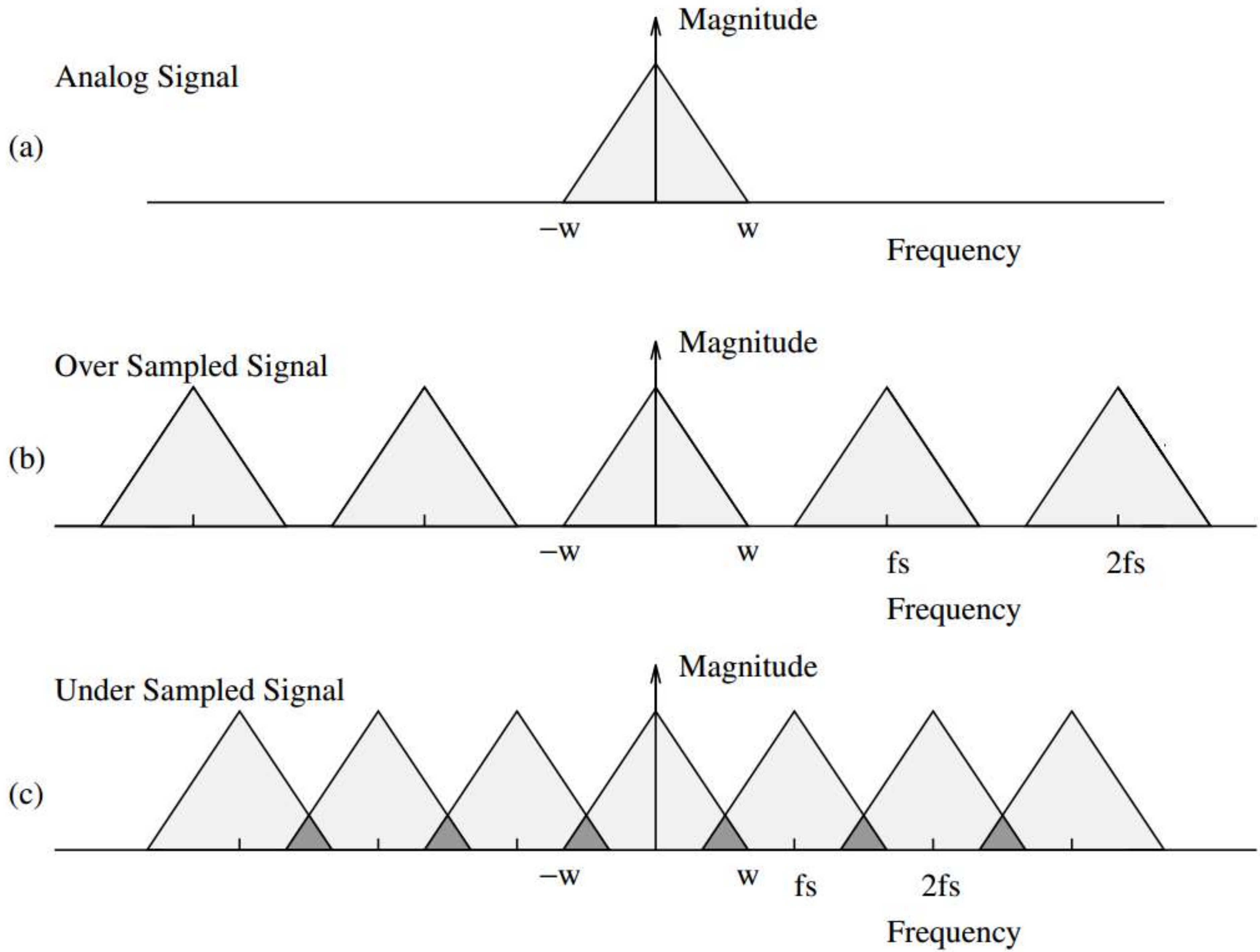
$$S_a(j\omega) = \int_{-\infty}^{\infty} s_a(t) e^{-j\omega t} dt$$

- such that $S_a(j\omega)=0$ for $|\omega| \geq 2\pi W$ then the analogue signal can be reconstructed from its sampled version if $T \leq 1/2W$. W is called the Nyquist frequency.

Sampling Contd…

- The Fourier transform of the sampled signal is evaluated at multiples of the sampling frequency which forms the relationship

$$S(e^{j\omega T}) = \frac{1}{T} \sum_{n=-\infty}^{\infty} S_a(j\omega + j2\pi n/T)$$


Magnitude
Analog Signal
(a)
−w
w
Frequency
Magnitude
Over Sampled Signal
(b)
−w
w
fs
2fs
Frequency
Magnitude
Under Sampled Signal
(c)
−w
w
fs
2fs
Frequency

Sampling Contd…

- The Fourier transform of the sampled signal is evaluated at multiples of the sampling frequency which forms the relationship

$$S(e^{j\omega T}) = \frac{1}{T}\sum_{n=-\infty}^{\infty} S_a(j\omega + j2\pi n/T)$$

Sampling Contd…

- The Fourier transform of the sampled signal is evaluated at multiples of the sampling frequency which forms the relationship

$$S(e^{j\omega T}) = \frac{1}{T}\sum_{n=-\infty}^{\infty} S_a(j\omega + j2\pi n/T)$$

- Fourier transform of the sampled sequence is proportional to the Fourier transform of the analogue signal in the base band as follows

$$S(e^{j\omega T}) = \frac{1}{T}S_a(j\omega) \quad |\omega| < \frac{\pi}{T}$$

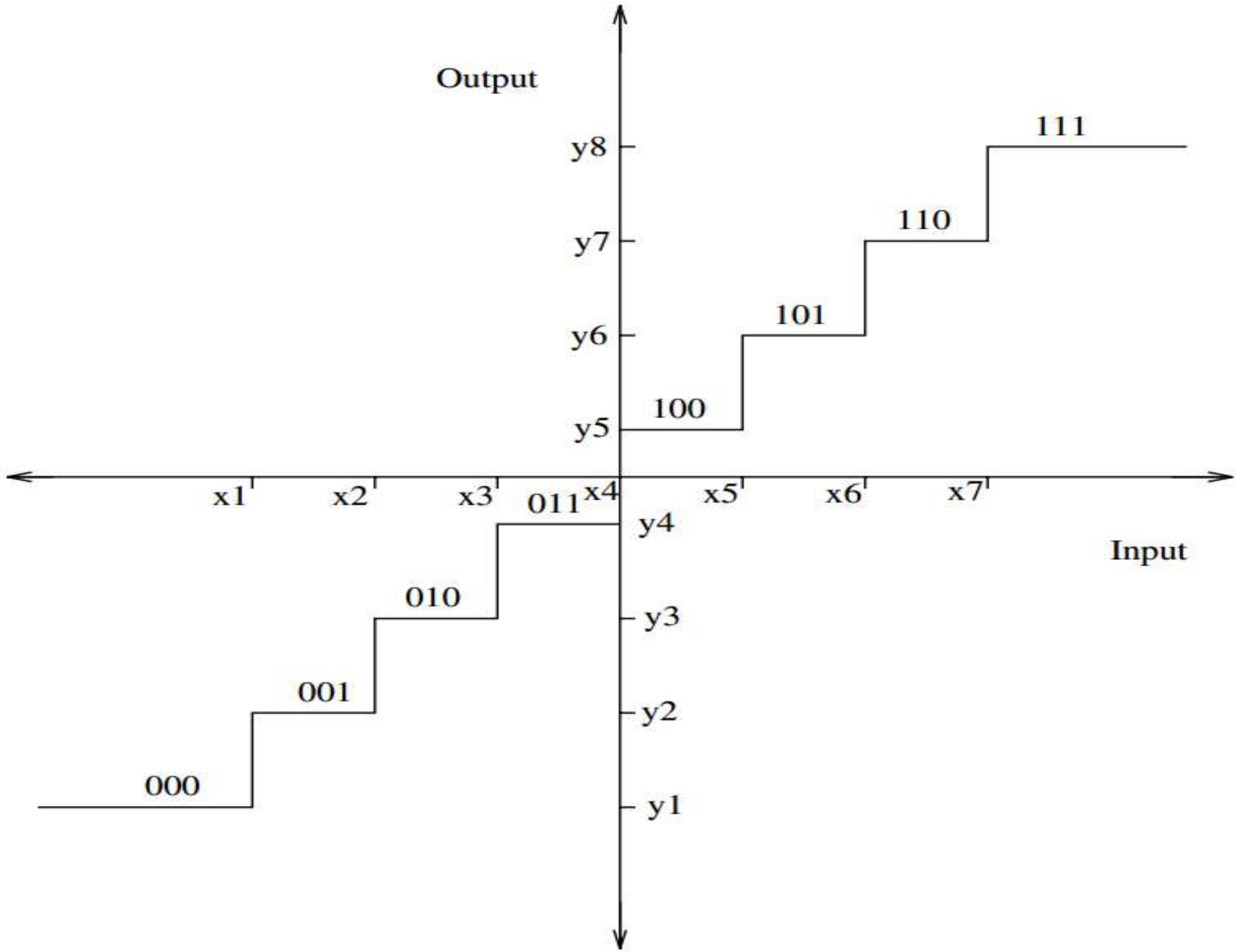
Sampling Contd…

- The original analogue signal can be obtained from the sampled sequence using interpolation given by

$$s_a(t) = \sum_{n=-\infty}^{\infty} s_a(nT) \frac{sin[\pi(t-nT)/T]}{\pi(t-nT)/T}$$

Scalar Quantization

- Quantization converts a continuous amplitude signal to a discrete-amplitude signal that is different from the continuous amplitude signal by the quantization error
- When each of a set of discrete values is quantized separately the process is known as scalar quantization


Output
Input
y8
y7
y6
y5
y4
y3
y2
y1
x1
x2
x3
x4
x5
x6
x7
000
001
010
011
100
101
110
111

SQ Contd…

- Assuming all of the discrete amplitude values in the quantizer are represented by the same number of bits B and the sampling frequency is fs, the channel transmission bit rate is given by

$$T_c = Bf_s \quad bits/second$$

- Given fixed fs, the only way to reduce bit rate is by reducing number of bits B
- In order to reduce the bit rate while maintaining good speech quality, various types of scalar quantizer have been designed

Quantization Error

- The signal lying in the i^th interval,

$$x_i - \frac{\Delta_i}{2} \leq s(n) < x_i + \frac{\Delta_i}{2}$$

is represented by the quantized amplitude x_i

- The mean squared error of the signal is given by

$$E_i^2 = \int_{x_i - \frac{\Delta_i}{2}}^{x_i + \frac{\Delta_i}{2}} (x - x_i)^2 p(x) dx$$

Quantization Error Contd…

- Assuming that Δ_i is small, we can assume that p(x) is flat within the interval $x_i-\Delta/2$ to $x_i+\Delta/2$
- Representing the flat region of p(x) by its value at the center, $p(x_i)$, the previous equation can be written as

$$E_i^2 = p(x_i)\int_{-\frac{\Delta_i}{2}}^{\frac{\Delta_i}{2}} y^2 dy = \frac{\Delta_i^3}{12} p(x_i)$$

Quantization Error Contd…

- The probability of the signal falling in the i^th interval is

$$\Gamma_i = \int_{x_i - \frac{\Delta_i}{2}}^{x_i + \frac{\Delta_i}{2}} p(x)dx = p(x_i)\Delta_i$$

- Substituting this equation in the previous equation we have

$$E_i^2 = \frac{\Delta_i^2}{12}\Gamma_i$$

- And the total means squared error

$$E^2 = \frac{1}{12}\sum_{i=1}^{N}\Gamma_i\Delta_i^2$$

Uniform Quantizer

- In a uniform quantizer, all of the quantizer intervals (steps) are the same width
- A uniform quantizer can be defined by two parameters: the number of quantizer levels and the quantizer step size Δ
- The step size is given by

$$\Delta = \frac{2X_{max}}{2^B}$$

- The quantization error is bounded by

$$-\frac{\Delta}{2} \leq e_q(n) \leq \frac{\Delta}{2}$$

Uniform Quantizer Contd…

- The only way to reduce the quantization error is by increasing the number of bits
- When a uniform quantizer is used, it is assumed that the input signal has a uniform probability density function varying between $\pm X_{max}$, with magnitude $1/2X_{max}$

Uniform Quantizer Contd...

- The power of the input signal can be written as

$$P_x = \int_{-X_{max}}^{X_{max}} x^2 p(x)dx = \frac{X_{max}^2}{3}$$

- The signal to Quantization error ratio is given by

$$SNR = \frac{P_x}{P_n} = \frac{X_{max}^2/3}{\Delta^2/12}$$

$$SNR = \frac{P_x}{P_n} = 2^{2B}$$

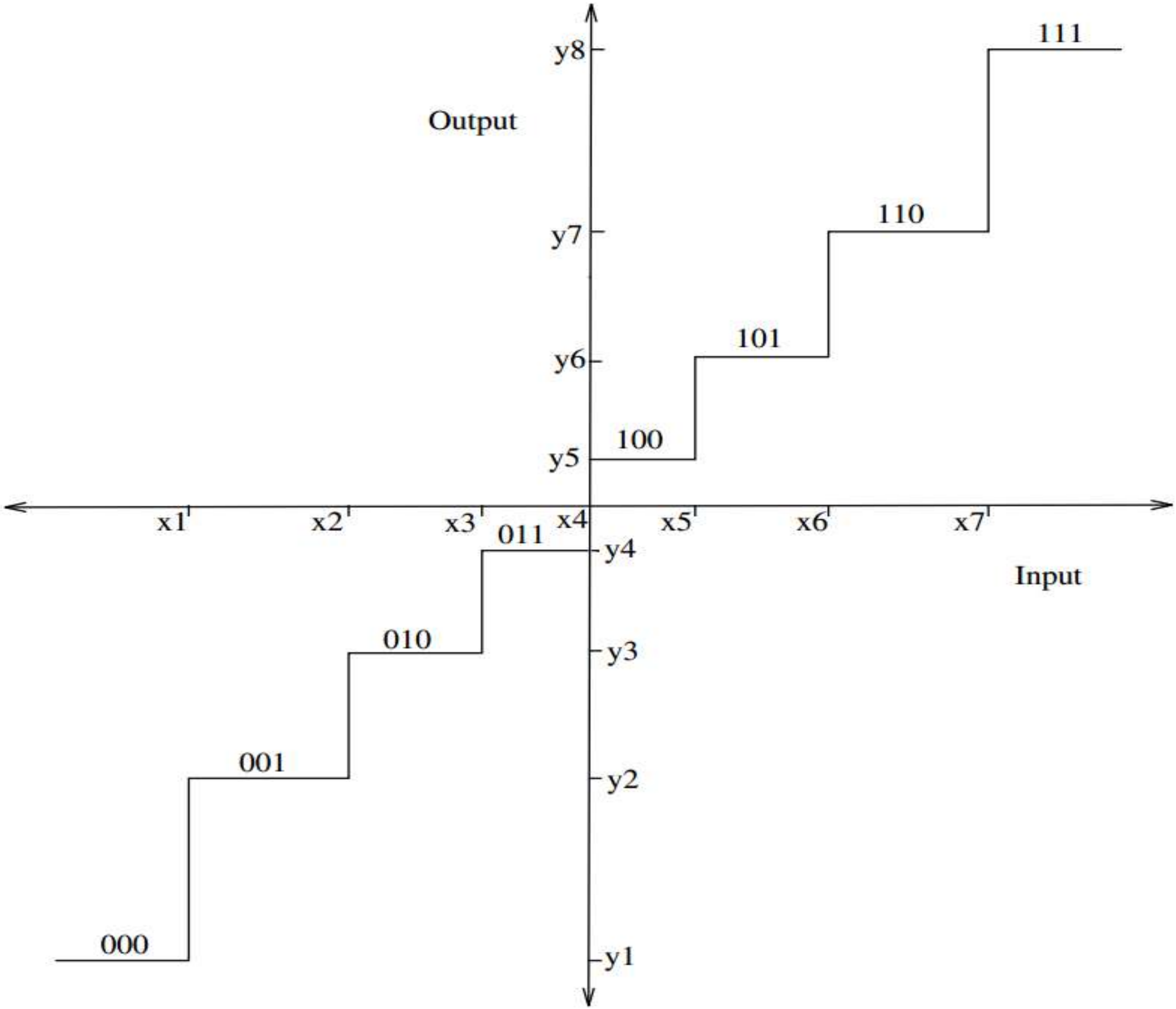
$$SNR(dB) = 10\log_{10}(2^{2B}) = 20B\log_{10}(2) = 6.02B\ dB$$

Optimum Quantizer

- In order to maximize SNR for a given number of bits per sample, levels of the quantizer must be selected to match the probability density function of the signal to be quantized
- This is because speech-like signals do not have a uniform probability density function, and the probability of smaller amplitudes occurring is much higher than that of large amplitudes

Optimum Quantizer Contd…

- Consequently, the optimum quantizer should have quantization levels with nonuniform spacing



y8
111
Output
y7
110
y6
101
100
y5
x1
x2
x3
x4
x5
x6
x7
011
y4
Input
010
y3
001
y2
000
y1

Optimum Quantizer Contd…

- Consequently, the optimum quantizer should have quantization levels with nonuniform spacing

Optimum Quantizer Contd…

- Consequently, the optimum quantizer should have quantization levels with nonuniform spacing
- The noise contribution of each interval depends on the probability of the signal falling into a certain quantization interval
- The nonuniform spacing of the quantization levels is equivalent to a nonlinear compressor C(x) followed by a uniform quantizer
- The less likely higher sample values are compressed more than the more likely low amplitude samples

Table 3.1 Max quantizer input and output levels for 1, 2, 3, 4, and 5 bit quantizers

Max quantizer thresholds									
1 bit		2 bit		3 bit		4 bit		5 bit	
i/p	*o/p*	*i/p*	*o/p*	*i/p*	*o/p*	*i/p*	*o/p*	*i/p*	*o/p*
0.0000	0.7980	0.0000	0.4528	0.0000	0.2451	0.0000	0.1284	0.0000	0.0659
		0.9816	1.5100	0.5006	0.7560	0.2582	0.3881	0.1320	0.1981
				1.0500	1.3440	0.5224	0.6568	0.2648	0.3314
				1.7480	2.1520	0.7996	0.9424	0.3991	0.4668
						1.0990	1.2560	0.5359	0.6050
						1.4370	1.6180	0.6761	0.7473
						1.8440	2.0690	0.8210	0.8947
						2.4010	2.7330	0.9718	1.0490
								1.1300	1.2120
								1.2990	1.3870
								1.4820	1.5770
								1.6820	1.7880
								1.9080	2.0290
								2.1740	2.3190
								2.5050	2.6920
								2.9770	3.2630

Optimum Quantizer Contd…

- Nonuniform quantization is advantageous in speech coding because
  - It matches the speech probability density function better and hence produces higher signal to noise ratio
  - Lower amplitudes, which contribute more to the intelligibility of speech, are quantized more accurately in a nonuniform quantizer
- In speech coding, Max's quantizer is widely used to normalize the input samples to unit variance

Logarithmic Quantizer

- An optimum quantizer is advantageous if the dynamic range (or variance) of the input signal is fixed to a small known range
- The performance of such a quantizer deteriorates rapidly as the power of the signal moves away from the value that the quantizer is designed for

Logarithmic Quantizer Contd…

- Logarithmic quantizers performances do not change significantly with changing signal variance and remain relatively constant over a wide range of input speech levels
- In a companding quantizer, quantizer levels are closely spaced for small amplitudes which progressively increase as the input signal range increases.

Logarithmic Quantizer Contd…

- The A-Law compression is defined by

$$A_{Law}(x) = \frac{Ax}{1+\log_{10}(A)} \quad for\ 0 \leq x \leq \frac{1}{A}$$

$$A_{Law}(x) = \frac{1+\log_{10}(Ax)}{1+\log_{10}(A)} \quad for\ \frac{1}{A} \leq x \leq 1$$

- The u-law compression is given by

$$\mu_{Law}(x) = sign(x)\frac{V_o \log_{10}\left[1+\frac{\mu|x|}{V_o}\right]}{\log_{10}[1+\mu]}$$

Logarithmic Quantizer Contd…

- A comparison showed that the optimum quantizer can be as much as 4dB better, however
  - An optimum quantizer may have more background noise when the channel is idle
  - Its dynamic range is limited to a smaller input signal range
- For these two reasons, logarithmic quantizers are usually preferred

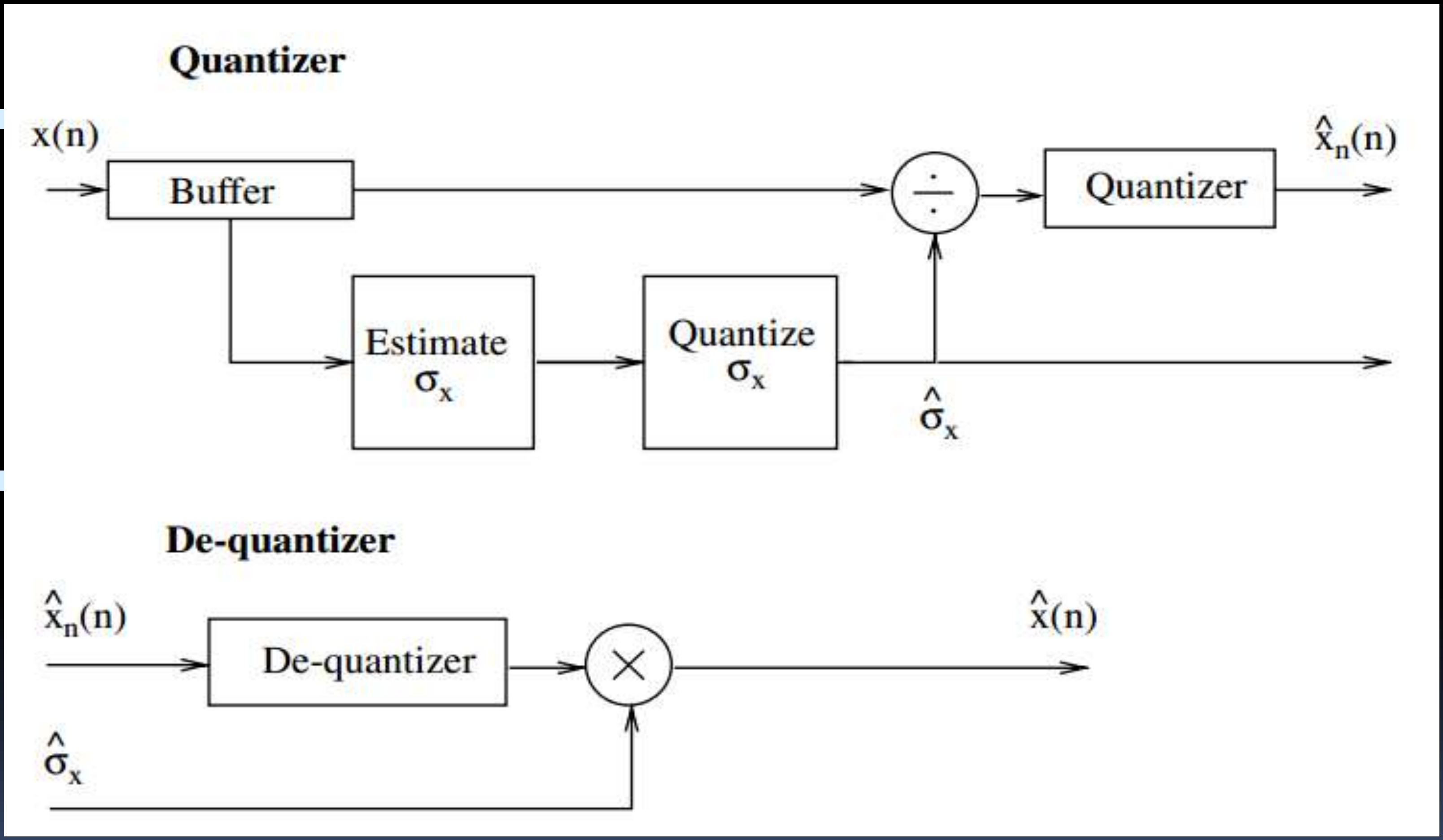
Adaptive Quantizer

- Although, the probability density function of speech can easily be estimated and used in a quantizer design process, the variations in its dynamic range, which can be as much as 30dB, reduces the performance of any quantizer
- One solution is estimating the variance of the speech segment prior to quantization and hence, adjusting the quantizer levels accordingly

Adaptive Quantizer Contd…

- The adjustment of the quantizer levels is equivalent to designing the quantizer for unit variance and normalizing the input signal before quantization
- This is called forward adaptation

Adaptive Quantizer Contd…



Adaptive Quantizer Contd…

- Assuming the speech is stationary during K samples, the rms is given by

$$\sigma_x = \sqrt{\frac{1}{K}\sum_{n=1}^{K} x(n)^2}$$

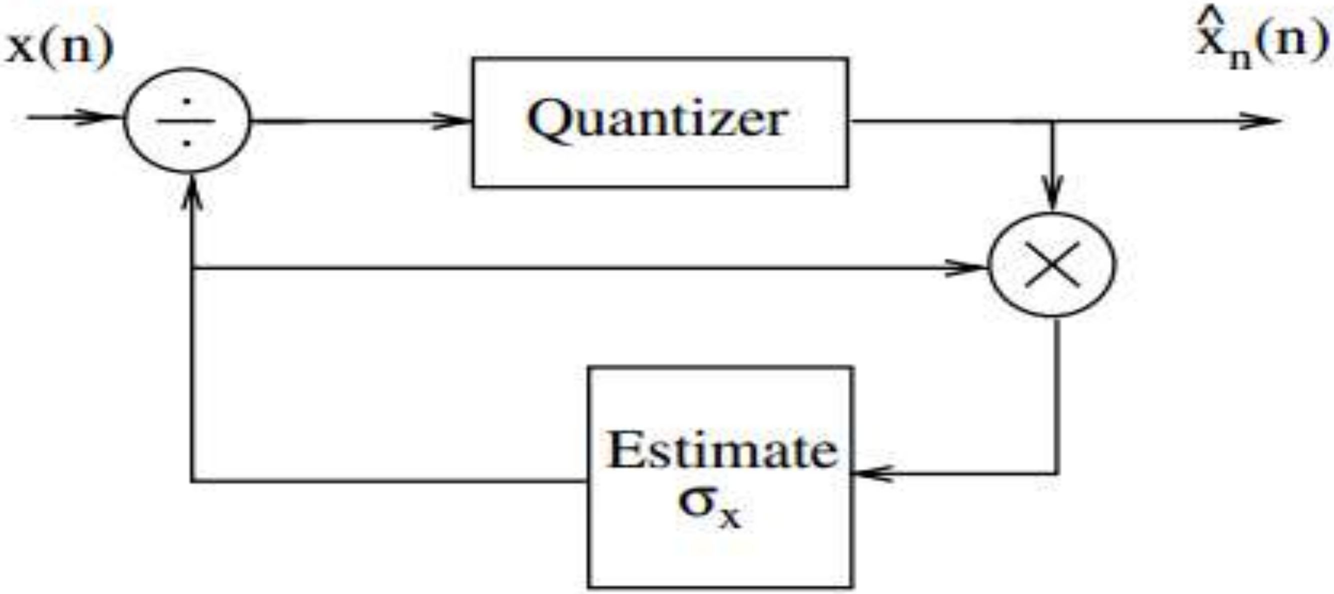
- The choice of block length K is very important because the probability density function of the normalized input signal can be affected by K

Adaptive Quantizer Contd…

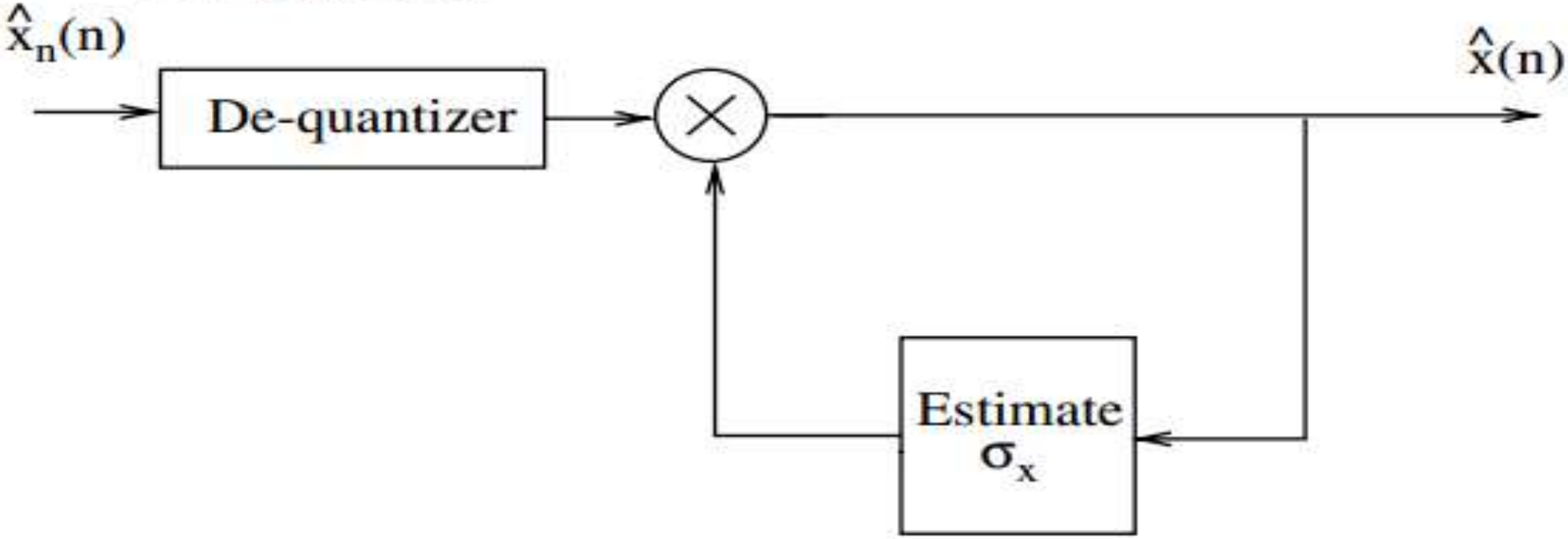
- Another adaptation scheme which does not require transmission of the speech variance to the de-quantizer is called backward adaptation
- Before quantizing each sample, the rms of the input signal is estimated from N previously quantized samples
- The normalizing factor for the n^{th} sample is

$$\sigma_x(n) = \sqrt{\frac{a_1}{N}\sum_{i=1}^{N}\hat{x}^2(n-i)}$$

Quantizer



De-quantizer



Adaptive Quantizer Contd…

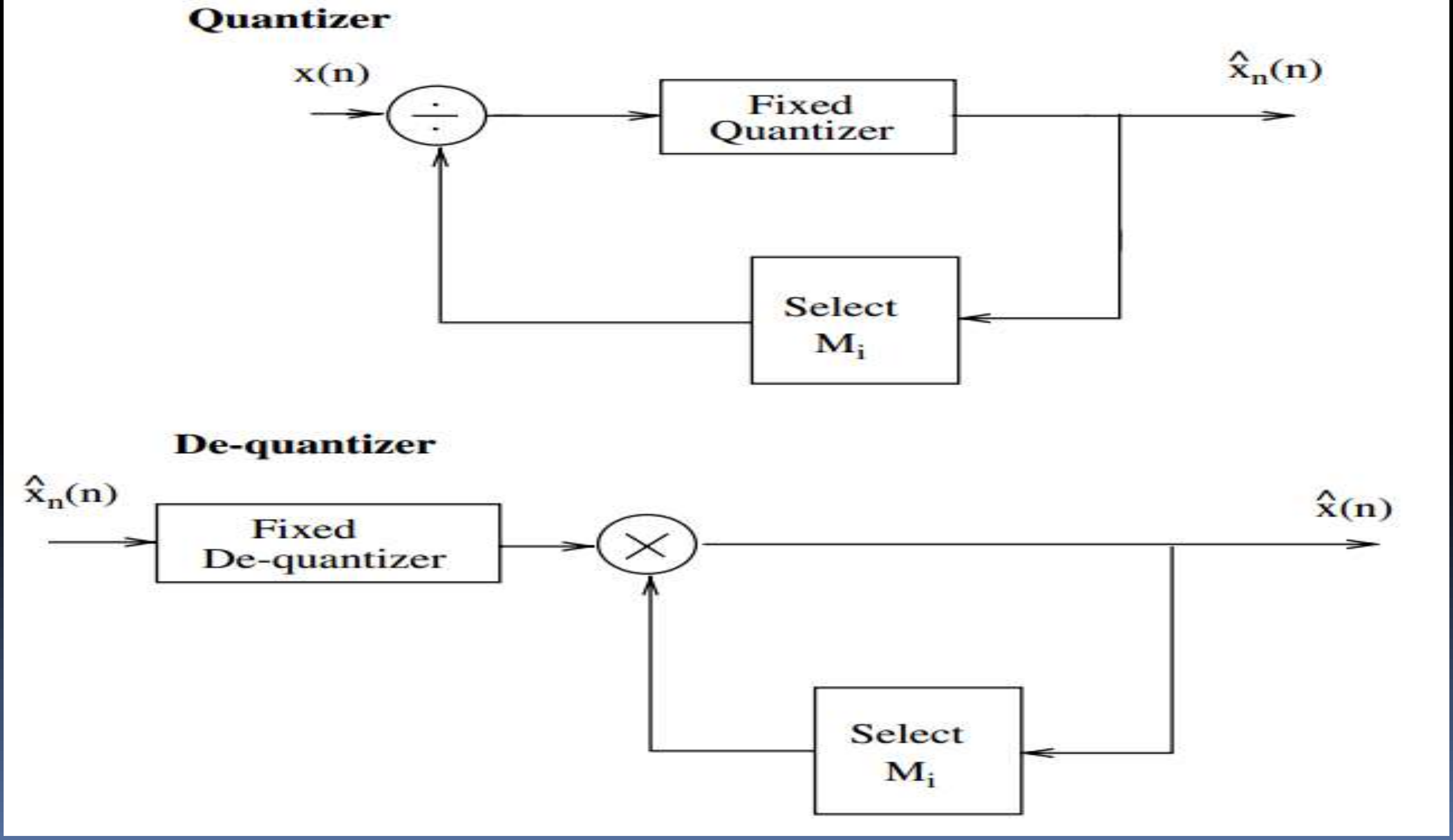
- One word memory

$$\Delta_{n+1} = \Delta_n M_i(|\hat{x}(n)|)$$

- where, Mi is one of i fixed coefficients corresponding to quantizer levels which control the expansion compression processes
- For large quantized previous samples, multiplier values are greater than one and for small previously quantized samples multiplier values are less than one

Adaptive Quantizer Contd…

- One word memory



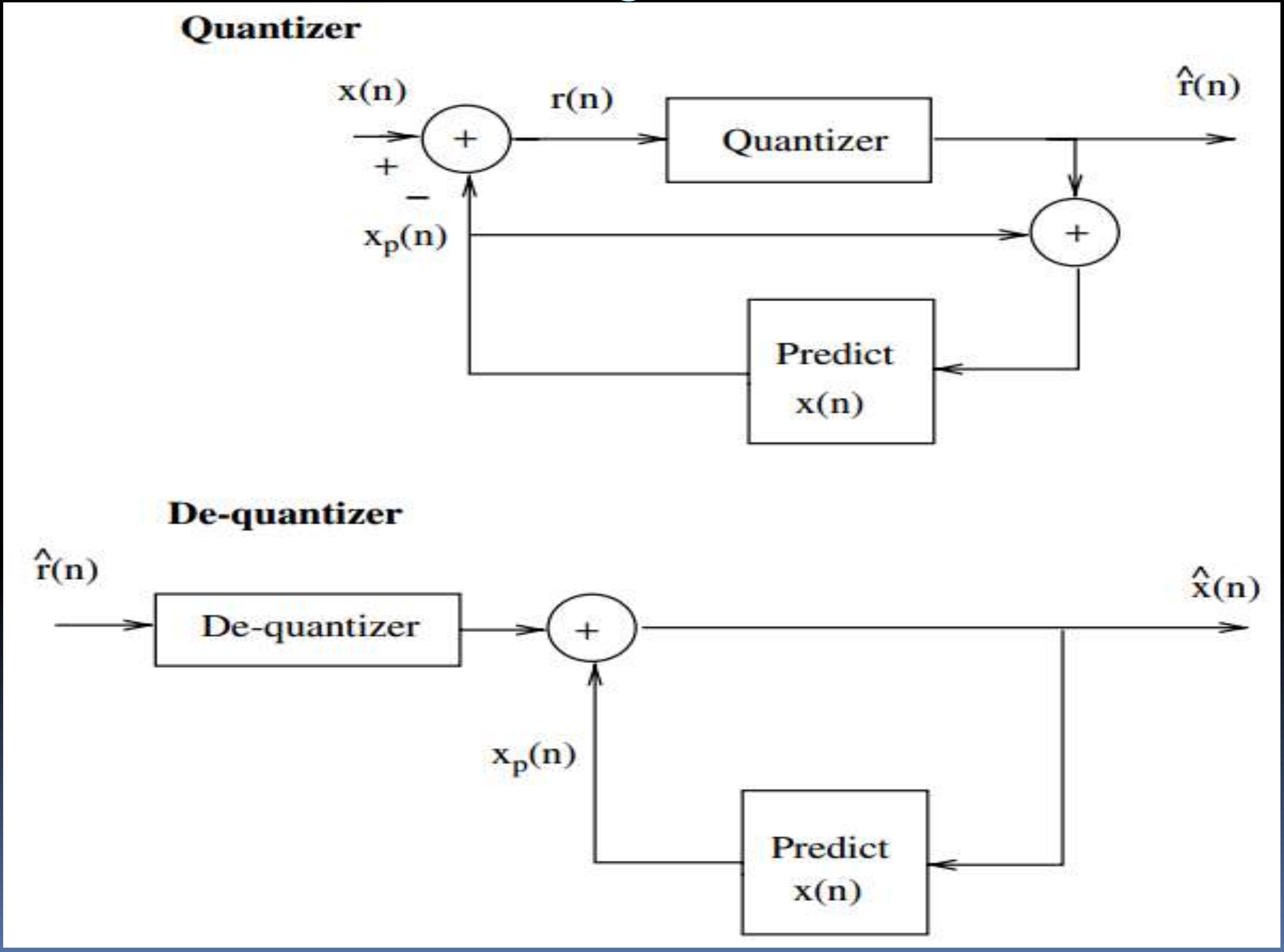
Differential Quantizer

- In a differential quantizer, the final quantized signal, r(n)is the difference between the input samples x(n) and their estimates $x_p(n)$

$$r(n) = x(n) - x_p(n)$$

$$x_p(n) = \sum_{k=1}^{p} \hat{x}(n-k)a_k$$

Differential Quantizer



Thank You