DIGITAL SIGNAL PROCESSING

Chapter 3:
Quantization Process with Over-Sampling and Noise Shaping

Reference:
S J.Orfanidis, "Introduction to Signal Processing", Prentice –Hall , 1996,ISBN 0-13-209172-0
M. D. Lutovac, D. V. Tošić, B. L. Evans, "Filter Design for Signal Processing Using MATLAB and Mathematica", Prentice Hall, 2001

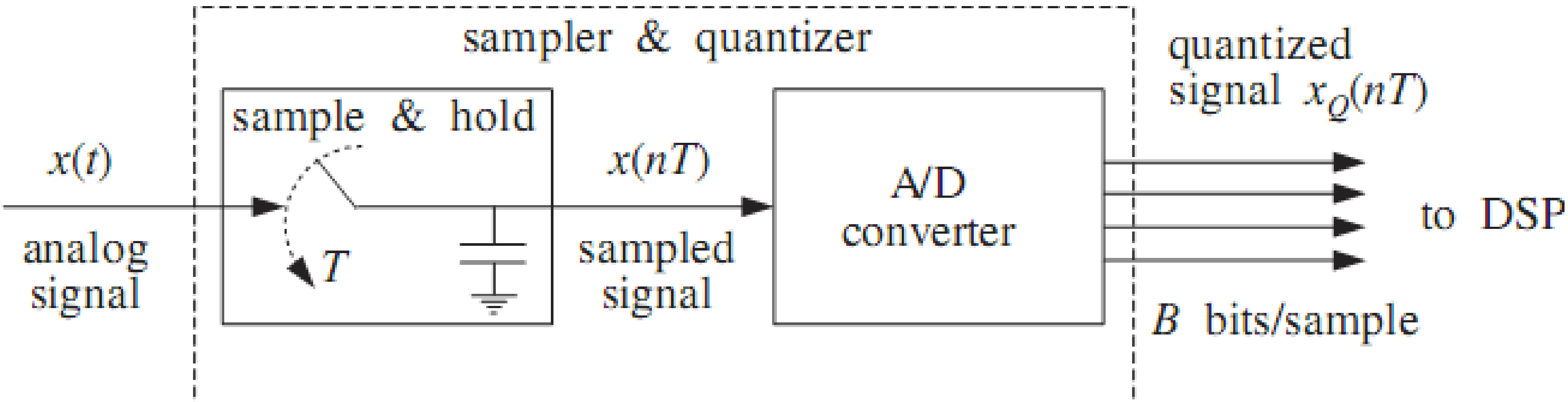
Lectured by Prof. Dr. Thuong Le-Tien
National Distinguished Lecturer
Cell: 0903 787 989
Email: ThuongLe@hcmut.edu.vn

Dated on January 2024

Quantization process and noise shaping

1. Quantization process.
2. Over sampling and Noise Shaping – Normalized and non-normalized SNR.
3. Coding: Natural Binary code; Offset Binary code; Two's Complement Code Digital to Analog conversion (DAC)
4. Analog to Digital Conversion ADC.
5. Analog and Digital Dither

1. Quantization Process



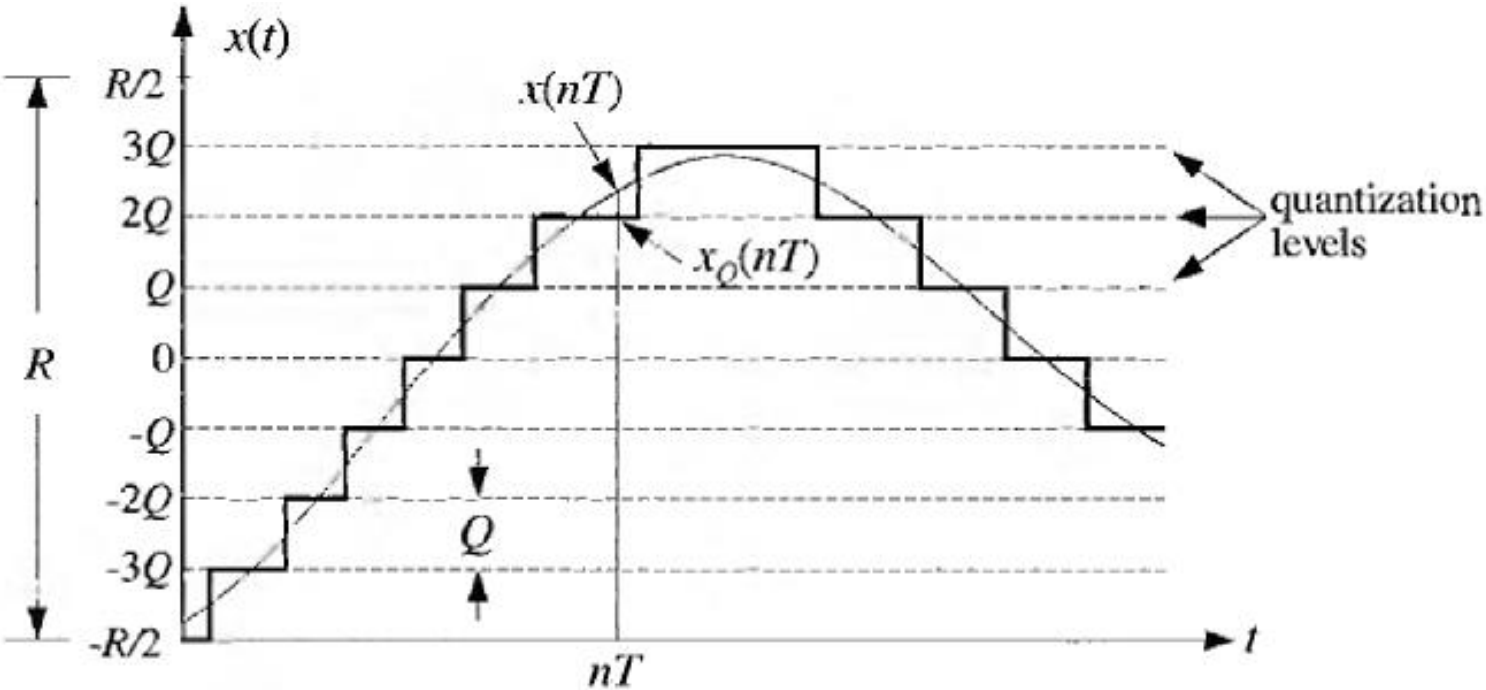
Analog to digital converter - ADC.

Quantized sample $x_Q(nT)$ represented by B bits take only one of 2^B possible value.

Quantization width or quantizer resolution Q

$$Q = \frac{R}{2^B} \qquad \frac{R}{Q} = 2^B$$

R is the full-scale range



R is in the symmetrical range:

$$-\frac{R}{2} \leq x_Q(nT) < \frac{R}{2}$$

Quantization error:

$$e(nT) = x_Q(nT) - x(nT)$$

In general case: $e = x_Q - x$

where, x_Q is the quantized value

$$-\frac{Q}{2} \leq e \leq \frac{Q}{2}$$

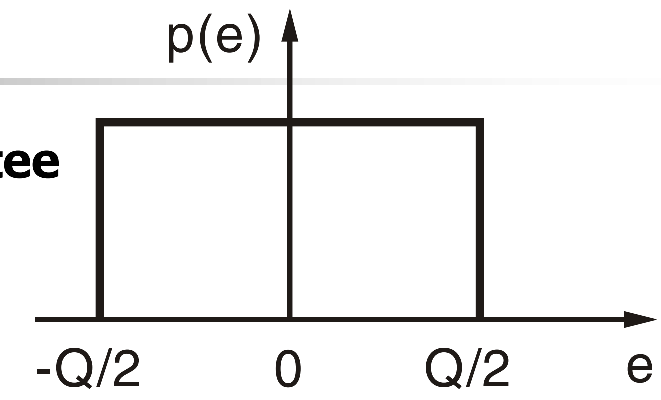
Mean: $$\bar{e} = \frac{1}{Q}\int_{-Q/2}^{Q/2} e\,de = 0$$

$$\overline{e^2} = \frac{1}{Q}\int_{-Q/2}^{Q/2} e^2\,de = \frac{Q^2}{12}$$

Root Mean Square error: $$e_{rms} = \sqrt{\overline{e^2}} = \frac{Q}{\sqrt{12}}$$

Quantization error e can be assumed as a random variable which is distributed uniformly over the range [-Q/2, Q/2] then having probability density:

$$p(e) = \begin{cases} \frac{1}{Q}, & -\frac{Q}{2} \le e \le \frac{Q}{2} \\ 0, & other \end{cases}$$



Normalization 1/Q needed to guarantee

$$\int_{-Q/2}^{Q/2} p(e)de = 1$$

The statistical expectation

$$E[e] = \int_{-Q/2}^{Q/2} ep(e)de \qquad E[e^2] = \int_{-Q/2}^{Q/2} e^2 p(e)de$$

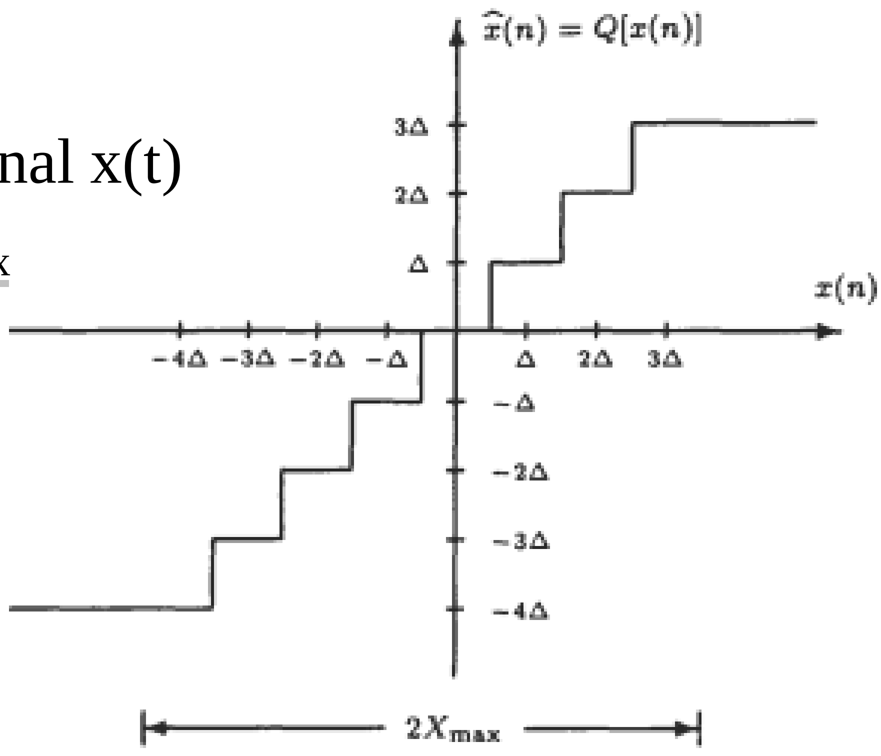
SNR (Normalized Signal-to-noise ratio):

$20\log_{10}(R/Q) = 20\log_{10}(2^B) = 20B\log_{10}(2)$

$$SNR = 20\log_{10}\left(\frac{R}{Q}\right) = 6B$$

Non-Normalized SNR

Define step size Q for the signal x(t)
With the max value to be X_{max}



$$Where \quad \Delta \equiv Q; \qquad B = B' + 1$$

$$Q = \frac{X_{max}}{2^{B-1}} \quad and \quad \sigma_e^2 = \frac{Q^2}{12}$$

then the SQNR

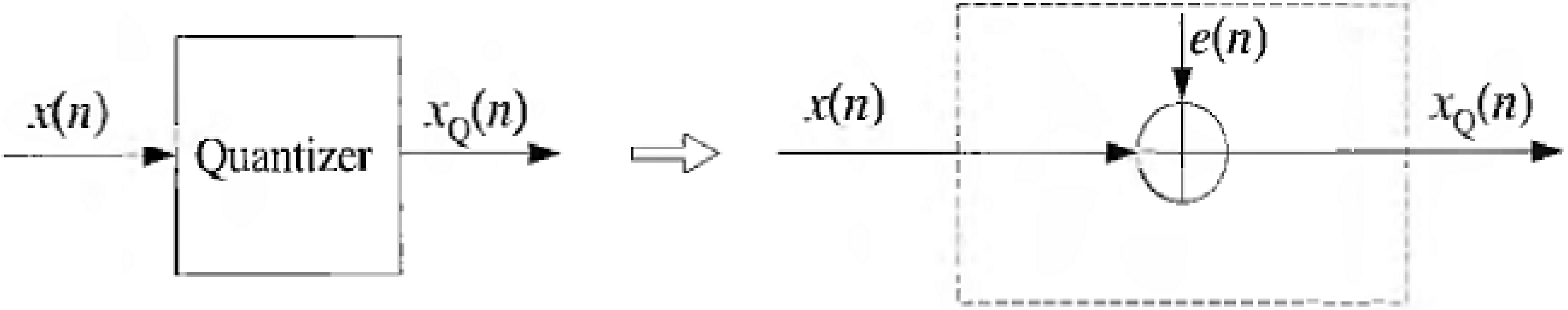
$$SQNR = 10\log\frac{\sigma_x^2}{\sigma_e^2} = 6B + 4.81 - 20\log\frac{X_{max}}{\sigma_x}$$

Thus, the practical Signal –to-Quantization-Noise increases approximately 6dB for each bit
(can be compared to the Normalized SNR)

$$\boxed{x_Q(n) = x(n) + e(n)}$$

The average power or variance of e(n)

$$\sigma_e^2 = E\left[e^2(n)\right] = \frac{Q^2}{12}$$



Assumed $e(n)$ is white noise then the autocorrelation function is the delta function

$$R_{ee}(k) = E[e(n+k)e(n)] = \sigma_e^2 \delta(k)$$

Example: in digital audio application, signal sampled at 44kHz and each sample quantized using a ADC having full scale of 10volts. Determine number of bits B if the rms quantization error must be kept below 50 microvolts. Then determine the actual rms error and bit rate

Sol:

$$e_{\rm rms} = Q/\sqrt{12} = R2^{-B}/\sqrt{12}$$

$$B = \log_2\left[\frac{R}{e_{\rm rms}\sqrt{12}}\right] = \log_2\left[\frac{10}{50\cdot 10^{-6}\sqrt{12}}\right] = 15.82$$

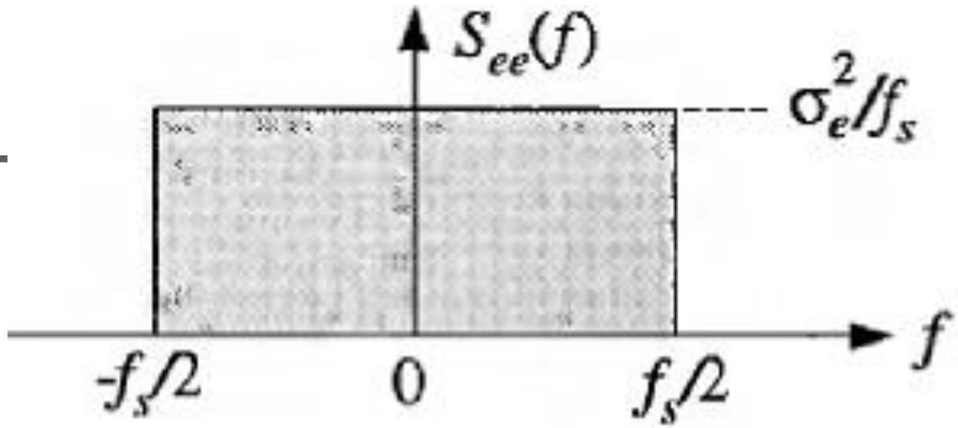
Which is rounded to B=16

$$e_{\rm rms} = R2^{-B}/\sqrt{12} = 44 \text{ microvolts}$$

Then bit rate: $Bf_s = 16 \cdot 44 = 704 \text{ kbits/sec}$

The Nomalized-SNR of the quantizer is 6B = 96 dB

2. Oversampling and noise shaping



Power spectrum of white quantization noise

Power spectrum density of e(n)

$$S_{ee}(f) = \frac{\sigma_e^2}{f_s}, \qquad \text{for} \quad -\frac{f_s}{2} \leq f \leq \frac{f_s}{2}$$

The noise power within at Nyquist sub-interval [f_a, f_b] with $\Delta f = f_b - f_a$:

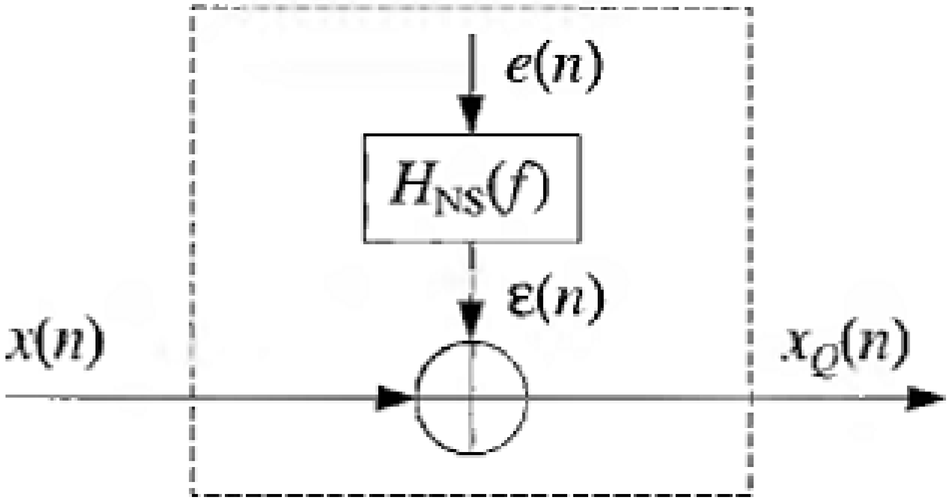
$$S_{ee}(f)\,\Delta f = \sigma_e^2 \frac{\Delta f}{f_s} = \sigma_e^2 \frac{f_b - f_a}{f_s}$$

The noise power over the entire interval $\Delta f = f_s$

$$\frac{\sigma_e^2}{f_s} f_s = \sigma_e^2$$

Noise shaping quantizers reshape the spectrum of the quantization noise into more convenient shape. This accomplished by filtering the white noise sequence $e(n)$ by a noise shaping filter $H_{NS}(f)$.

$$x_Q(n) = x(n) + \varepsilon(n)$$



Power spectral density

$$S_{\varepsilon\varepsilon}(f) = \left|H_{NS}(f)\right|^2 S_{ee}(f) = \frac{\sigma_e^2}{f_s}\left|H_{NS}(f)\right|^2$$

Noise power within a given interval

$$\int_{f_a}^{f_b} S_{\varepsilon\varepsilon}(f)df = \frac{\sigma_e^2}{f_s}\int_{f_a}^{f_b}\left|H_{NS}(f)\right|^2 df$$

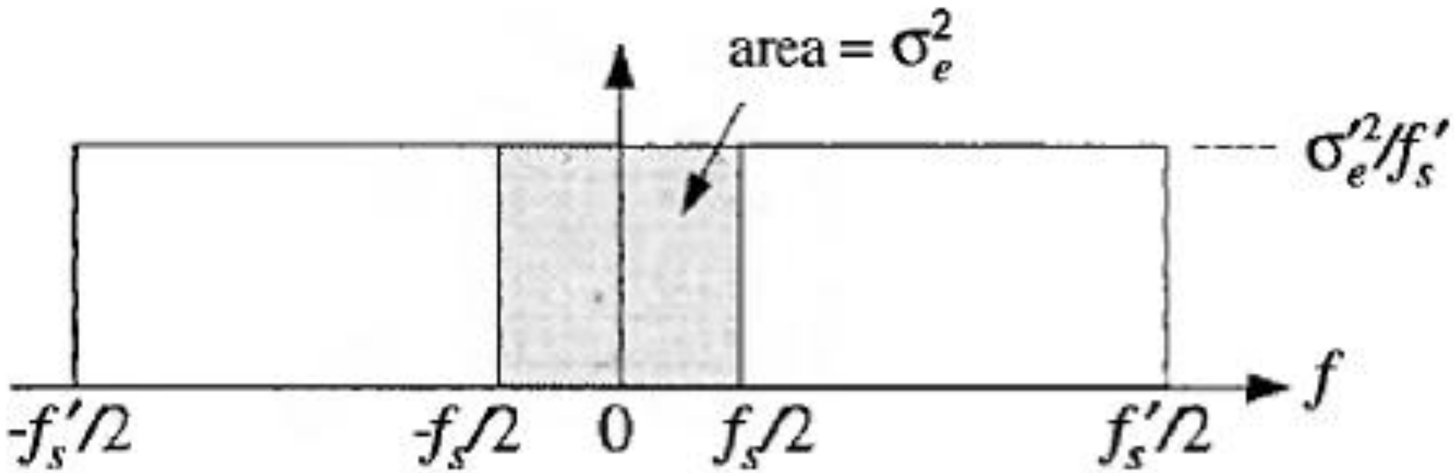
Over Sampling ratio $L = \dfrac{f_s'}{f_s}$

Quantization noise powers $\sigma_e^2 = \dfrac{Q^2}{12}$

To maintain the same quality required the power spectral density remain the same

$$\frac{\sigma_e^2}{f_s} = \frac{\sigma_e'^2}{f_s'}$$

$$\sigma_e^2 = f_s \frac{\sigma_e'^2}{f_s'} = \frac{\sigma_e'^2}{L}$$

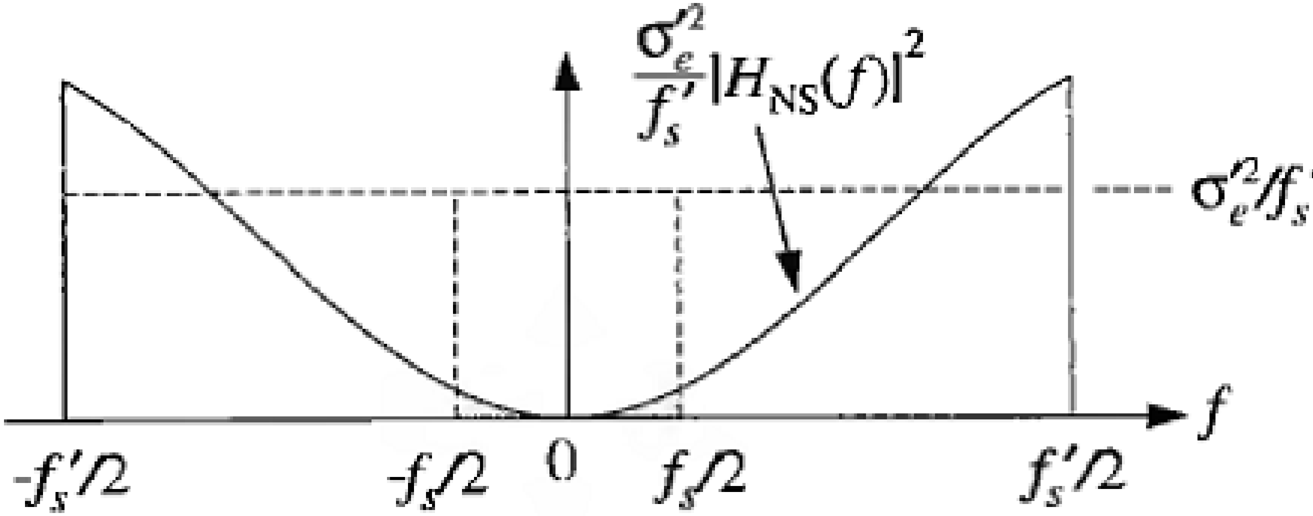


$$L = \frac{\sigma'^2_e}{\sigma^2_e} = 2^{2(B-B')} = 2^{2\Delta B}$$

ΔB = B-B', or ΔB = 0.5 $\log_2 L$

The total noise power in the Nyquist interval:

$$\sigma_e^2 = \frac{\sigma_e'^2}{f_s'} \int_{-f_s/2}^{f_s/2} |H_{NS}(f)|^2 \, df$$



$$|H_{NS}(f)|^2 = \left| 2\sin\left(\frac{\pi f}{f_s'}\right) \right|^{2p} \qquad -\frac{f_s'}{2} \le f \le \frac{f_s'}{2}$$

$$|H_{NS}(f)|^2 = \left(\frac{2\pi f}{f_s'}\right)^{2p} \quad \text{for } |f| \ll f_s'/2$$

$$\sigma_e^2 = \frac{\sigma'^2_e}{f_s'}\int_{-f_s/2}^{f_s/2}\left(\frac{2\pi f}{f_s'}\right)^{2p} df = \sigma'^2_e \frac{\pi^{2p}}{2p+1}\left(\frac{f_s}{f_s'}\right)^{2p+1}$$

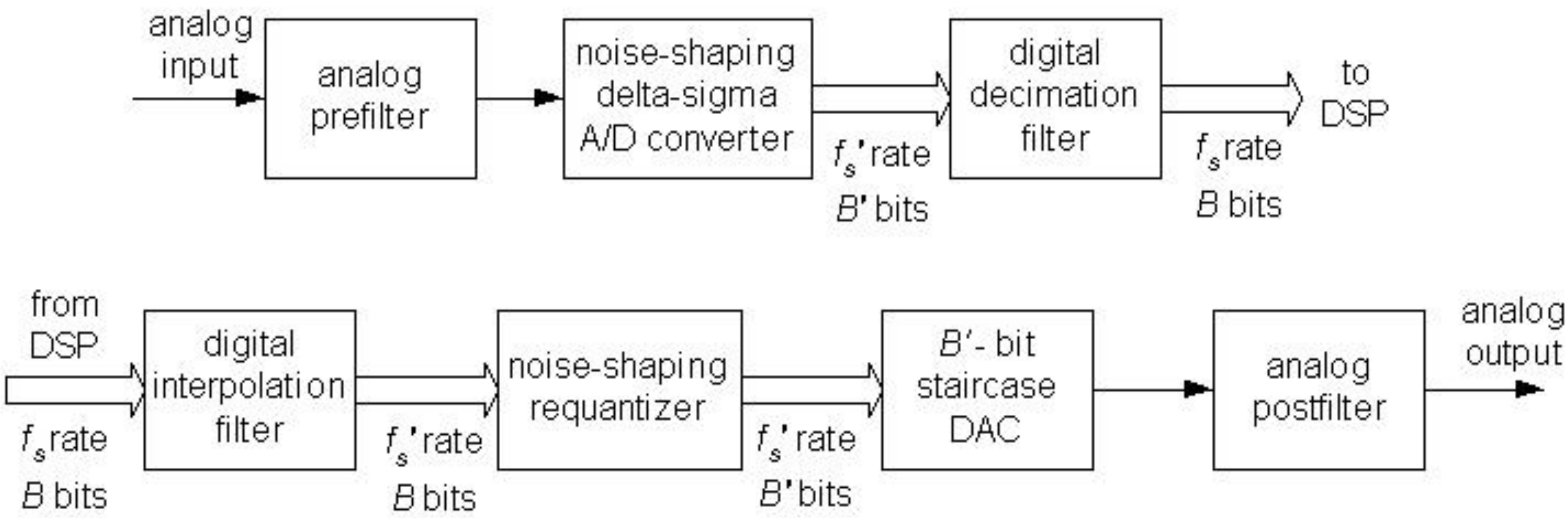
$$= \sigma'^2_e \frac{\pi^{2p}}{2p+1}\left(\frac{f_s}{f_s'}\right)^{2p+1} = \sigma'^2_e \frac{\pi^{2p}}{2p+1}\left(\frac{1}{L^{2p+1}}\right)$$

$$\sigma_e^2 / \sigma'^2_e = 2^{-2(B-B')} = 2^{-2\Delta B}$$

$$\Delta B = (p+0.5)\log_2 L - 0.5\log_2\left(\frac{\pi^{2p}}{2p+1}\right)$$

p	L	4	8	16	32	64	128
0	$\Delta B = 0.5\log_2 L$	1.0	1.5	2.0	2.5	3.0	3.5
1	$\Delta B = 1.5\log_2 L - 0.86$	2.1	3.6	5.1	6.6	8.1	9.6
2	$\Delta B = 2.5\log_2 L - 2.14$	2.9	5.4	7.9	10.4	12.9	15.4
3	$\Delta B = 3.5\log_2 L - 3.55$	3.5	7.0	10.5	14.0	17.5	21.0
4	$\Delta B = 4.5\log_2 L - 5.02$	4.0	8.5	13.0	17.5	22.0	26.5
5	$\Delta B = 5.5\log_2 L - 6.53$	4.5	10.0	15.5	21.0	26.5	32.0

Oversampling and noise shaping system



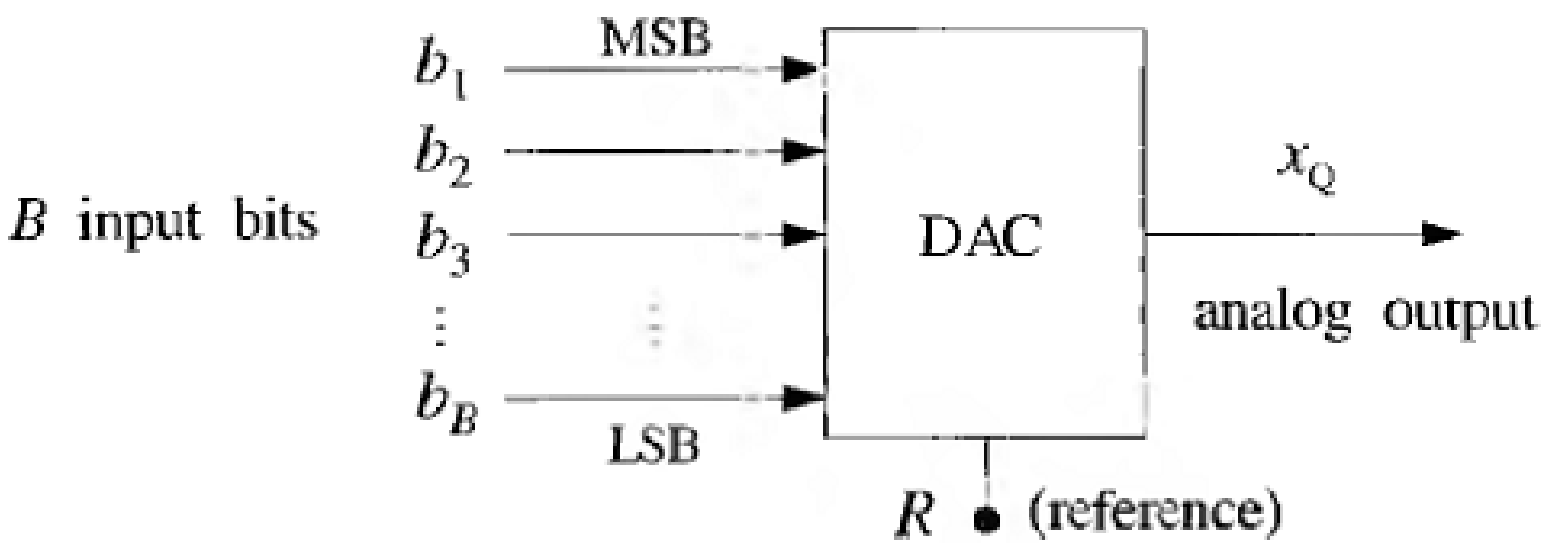
3. Digital to Analog Converter DAC

***B* bit 0 and 1 at input, b = [b_1, b_2,…, b_B],**

(a) unipolar natural binary,

(b) bipolar offset binary,

(c) bipolar 2's complement.



Unipolar natural binary

$$x_Q = R(b_1 2^{-1} + b_2 2^{-2} + \dots + b_B 2^{-B})$$

$$x_Q = R2^{-B}(b_1 2^{B-1} + b_2 2^{B-2} + \dots + b_{B-1} 2^1 + b_B)$$

Bipolar offset binary

$$x_Q = R(b_1 2^{-1} + b_2 2^{-2} + \dots + b_B 2^{-B} - 0.5)$$

Two's complement

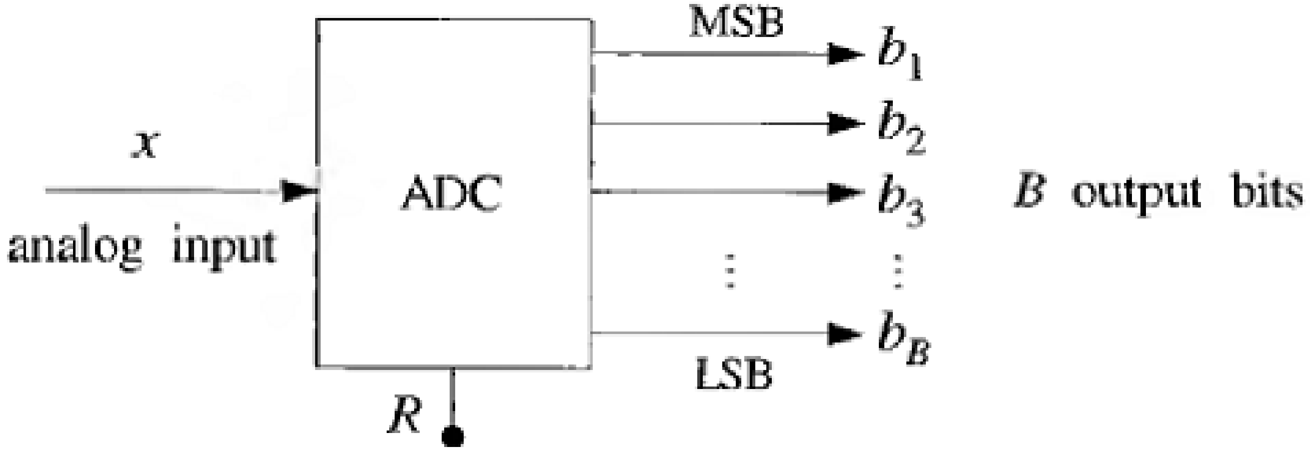
$$x_Q = R(\bar{b}_1 2^{-1} + b_2 2^{-2} + \dots + b_B 2^{-B} - 0.5)$$

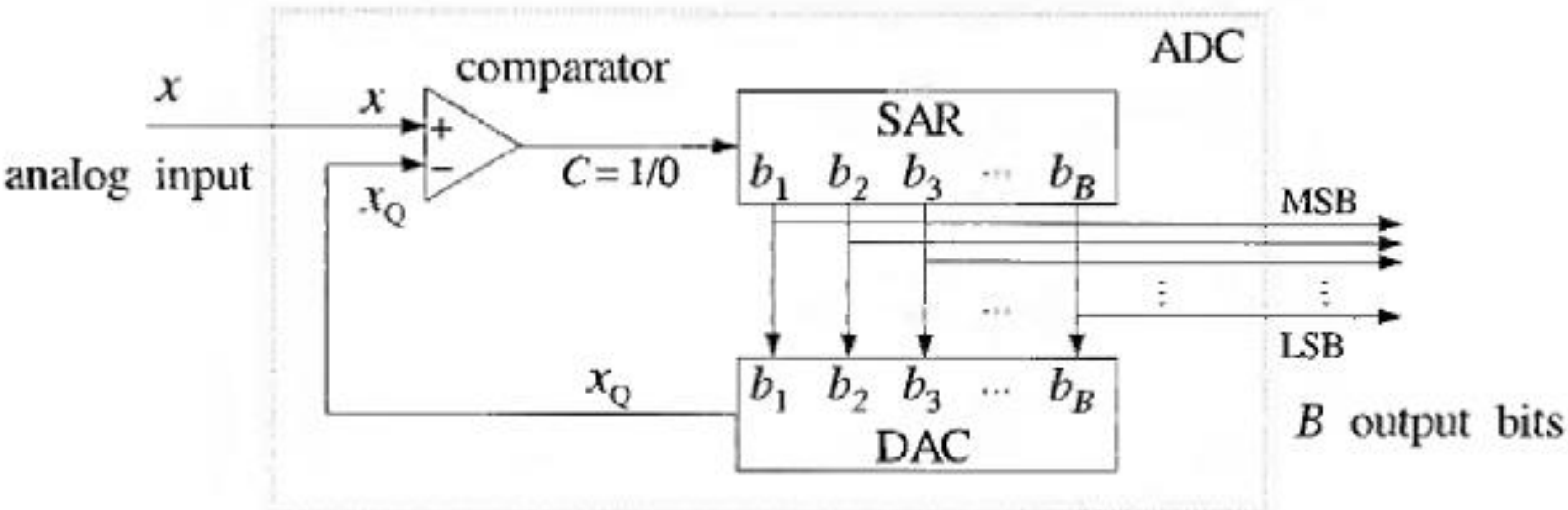
Converter type	I/O relationship
natural binary	$x_Q = R(b_1 2^{-1} + b_2 2^{-2} + \cdots + b_B 2^{-B})$
offset binary	$x_Q = R(b_1 2^{-1} + b_2 2^{-2} + \cdots + b_B 2^{-B} - 0.5)$
two's complement	$x_Q = R(\bar{b}_1 2^{-1} + b_2 2^{-2} + \cdots + b_B 2^{-B} - 0.5)$

Converter code for B=4bits, R=10volts

$b_1b_2b_3b_4$	natural binary		offset binary		2's C
	m	$x_Q = Qm$	m'	$x_Q = Qm'$	$b_1b_2b_3b_4$
—	16	10.000	8	5.000	—
1 1 1 1	15	9.375	7	4.375	0 1 1 1
1 1 1 0	14	8.750	6	3.750	0 1 1 0
1 1 0 1	13	8.125	5	3.125	0 1 0 1
1 1 0 0	12	7.500	4	2.500	0 1 0 0
1 0 1 1	11	6.875	3	1.875	0 0 1 1
1 0 1 0	10	6.250	2	1.250	0 0 1 0
1 0 0 1	9	5.625	1	0.625	0 0 0 1
1 0 0 0	8	5.000	0	0.000	0 0 0 0
0 1 1 1	7	4.375	−1	−0.625	1 1 1 1
0 1 1 0	6	3.750	−2	−1.250	1 1 1 0
0 1 0 1	5	3.125	−3	−1.875	1 1 0 1
0 1 0 0	4	2.500	−4	−2.500	1 1 0 0
0 0 1 1	3	1.875	−5	−3.125	1 0 1 1
0 0 1 0	2	1.250	−6	−3.750	1 0 1 0
0 0 0 1	1	0.625	−7	−4.375	1 0 0 1
0 0 0 0	0	0.000	−8	−5.000	1 0 0 0

4. Analog to Digital Converter (ADC)



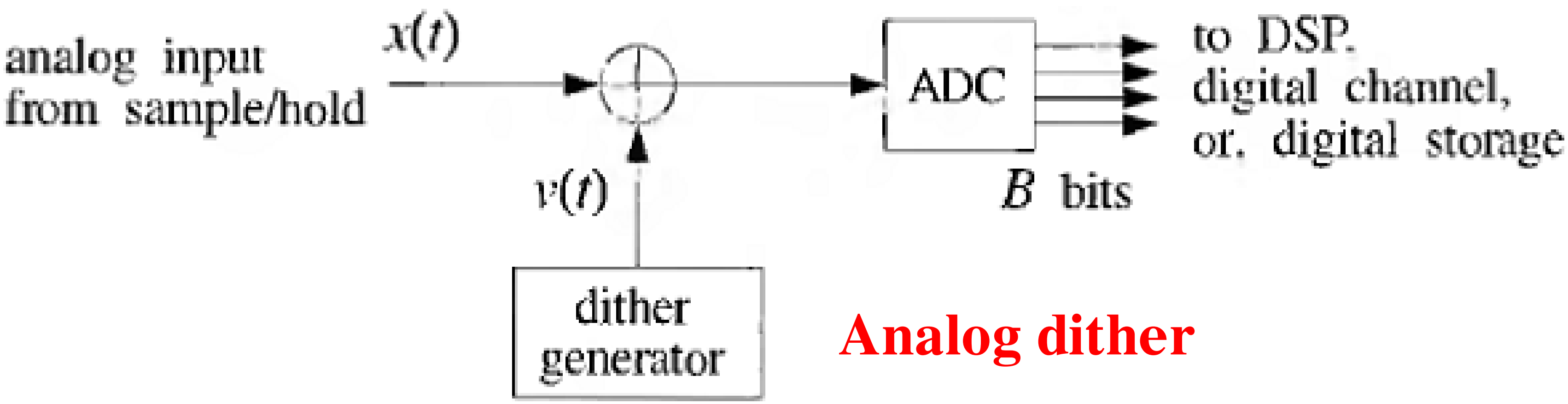


Example: A sampled sinusoid $x(n)=A\cos(2\pi fn)$, A=3volts and f=0.04 cycles/sample. The sinusoid is evaluated at the ten Sampling times n=0,1,2…9 and x(n) is quantized using a 4-bit ADC with R=10volts. The following table shows the sampled and quantized values and its codes

n	$x(n)$	$x_Q(n)$	2's C	offset
0	3.000	3.125	0101	1101
1	2.906	3.125	0101	1101
2	2.629	2.500	0100	1100
3	2.187	1.875	0011	1011
4	1.607	1.875	0011	1011
5	0.927	0.625	0001	1001
6	0.188	0.000	0000	1000
7	−0.562	−0.625	1111	0111
8	−1.277	−1.250	1110	0110
9	−1.912	−1.875	1101	0101

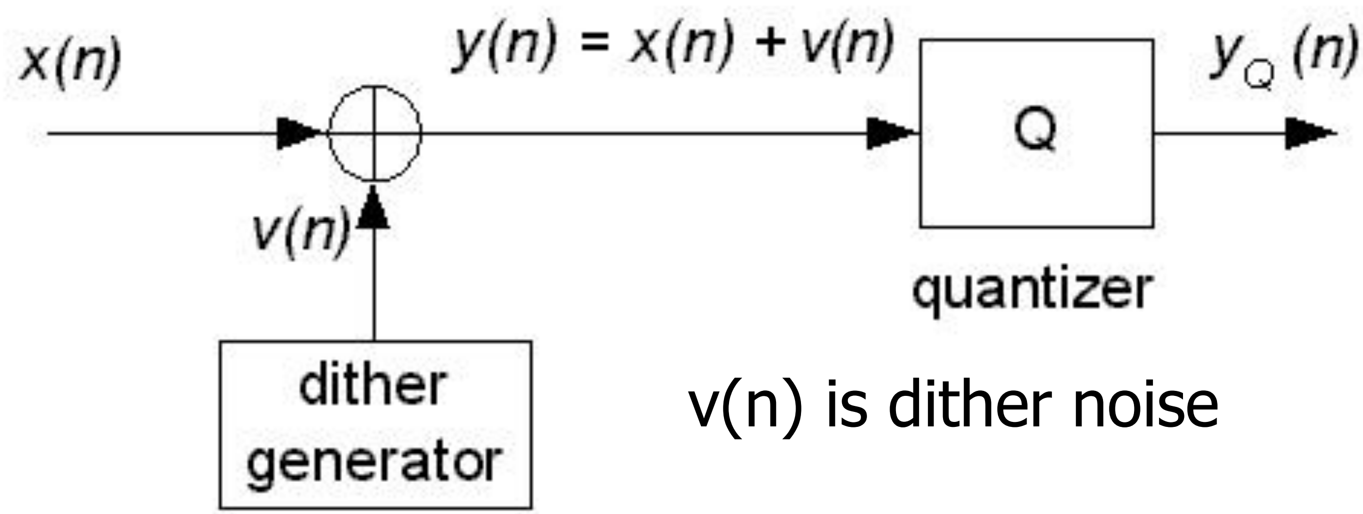
5. Analog and Digital Dither

Dither is a low-level white noise signal added to the input before quantization for eliminating granulation or quantization distortion and making the total quantization error behave like white noise



Analog dither

Digital dither can be added to a digital prior to a requantization operation that reduces the number of bits representing the signal.



v(n) is dither noise

Nonsubtractive dither process and quantization (Analog and digital dithers)

$y(n) = x(n) + v(n)$

Quantization error: $e(n) = y_Q(n) - y(n)$

Total error resulting from dithering and quantization:

$$\varepsilon(n) = y_Q(n) - x(n)$$

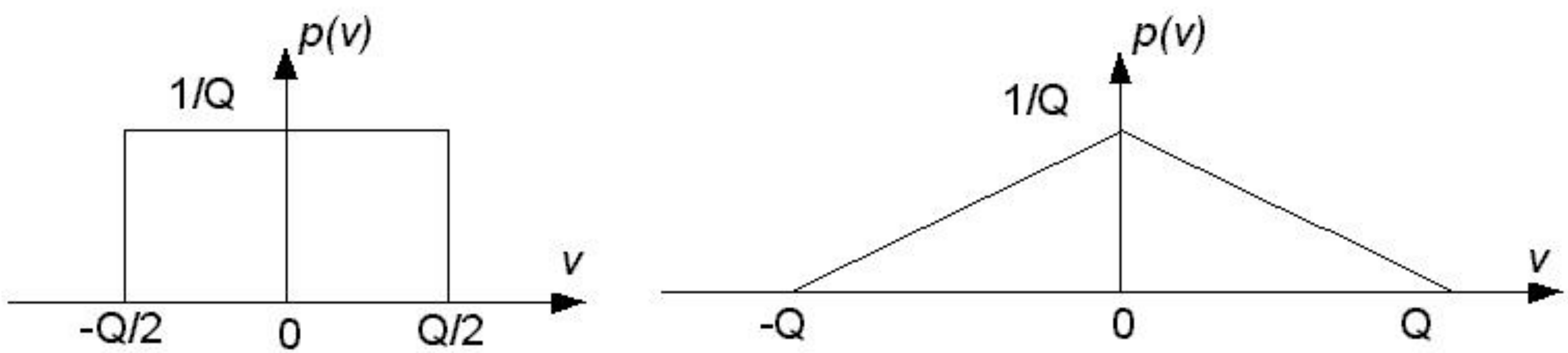
$\varepsilon(n) = (y(n) + e(n)) - x(n) = x(n) + v(n) + e(n) - x(n)$

or

$$\varepsilon(n) = yQ(n) - x(n) = e(n) + v(n)$$

Total error noise power

$$\sigma_\varepsilon^2 = \sigma_e^2 + \sigma_v^2 = \frac{1}{12}Q^2 + \sigma_v^2$$

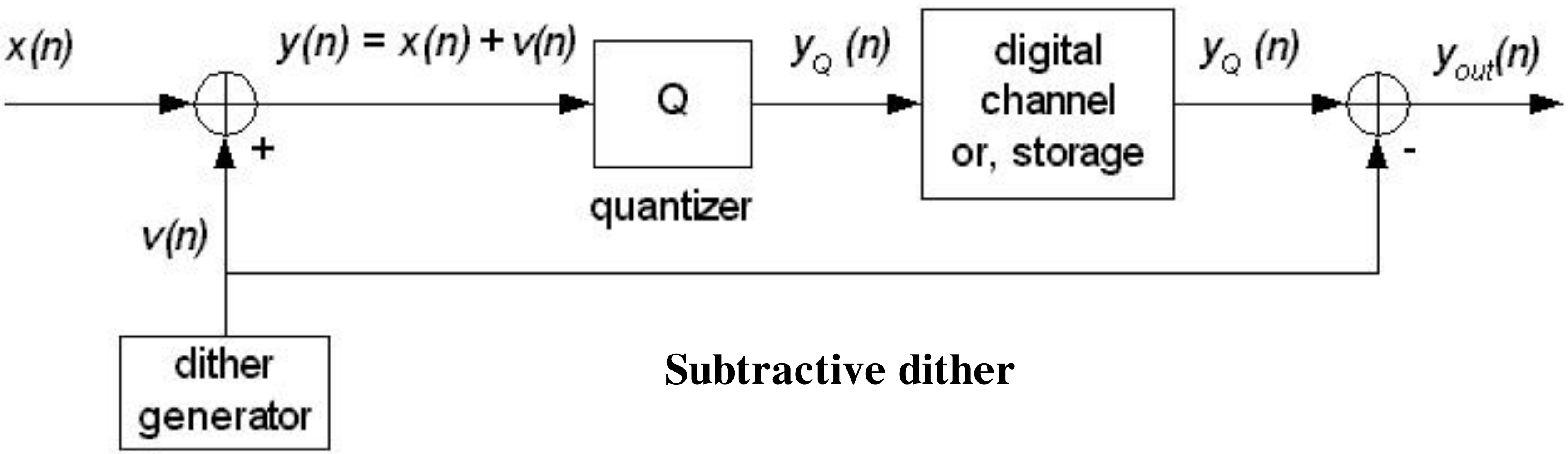


The two common Rectangular and triangular dither probability densities

Total error variance (the noise penalty in using dither)

$$\sigma_\epsilon^2 = \begin{cases} Q^2/12, & \text{undithered} \\ 2Q^2/12, & \text{rectangular dither} \\ 3Q^2/12, & \text{triangular dither} \\ 4Q^2/12, & \text{Gaussian dither} \end{cases}$$

10log2 = 3 dB
10log3 = 4.8 dB
10log4 = 6 dB

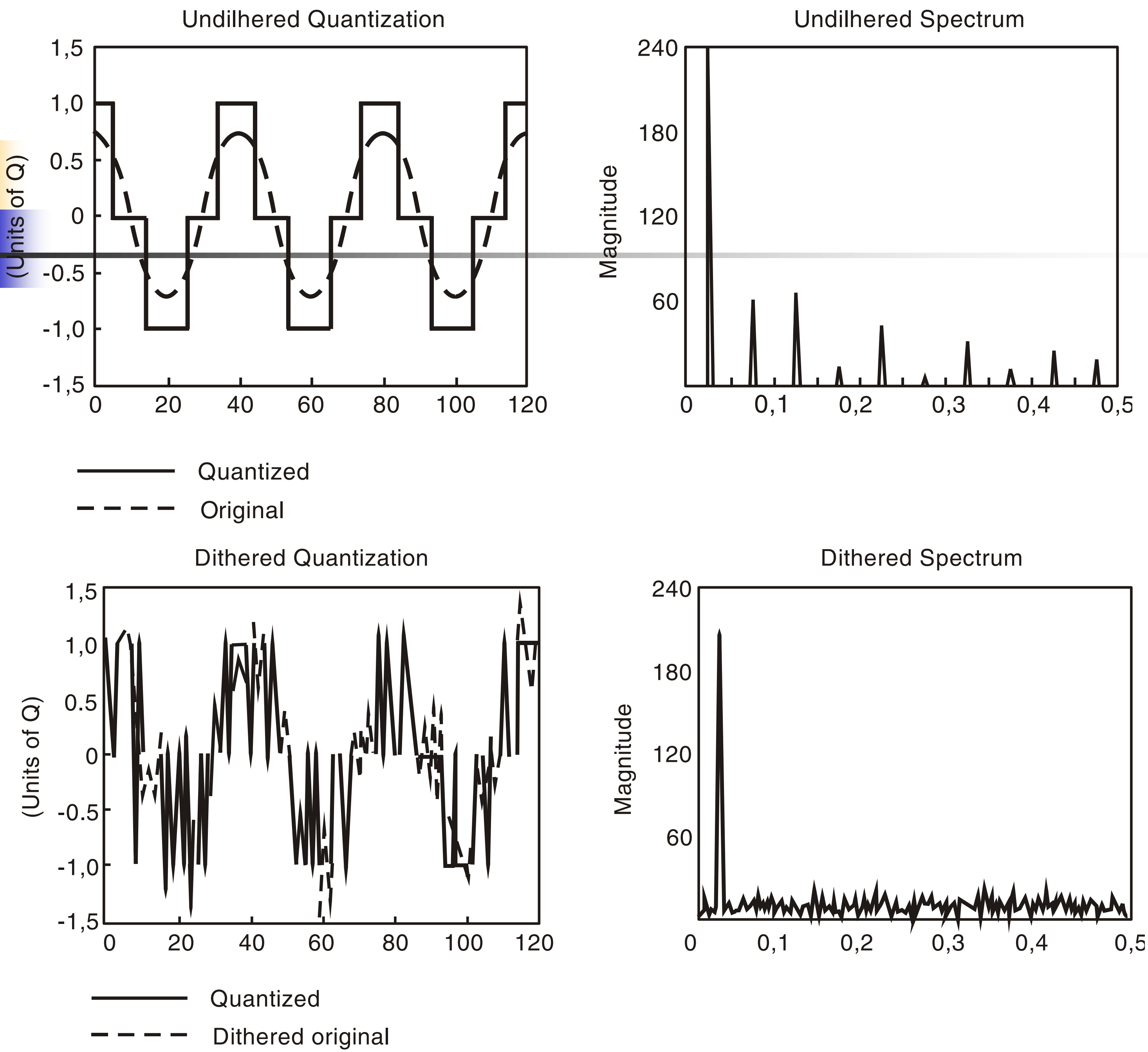


Subtractive dither

Total error

$$\varepsilon(n) = y_{out}(n) - x(n) = (y_Q(n) - v(n)) - x(n) = y_Q(n) - (x(n) + v(n))$$

$$\varepsilon(n) = y_Q(n) - y(n) = e(n)$$


Undilhered Quantization
(Units of Q)
Quantized
Original
Undilhered Spectrum
Magnitude
Dithered Quantization
(Units of Q)
Quantized
Dithered original
Dithered Spectrum
Magnitude