Noise Shaping - Sigma Delta Modulation

Maxx Seminario

University of Nebraska-Lincoln

Fall 2025

Why Oversampling?

Key Concept: Sample at a rate *higher* than the Nyquist rate to gain advantages in A/D and D/A conversion.

Benefits:

- **Simplified Antialiasing Filters:** Oversampling relaxes analog filter requirements
- **Reduced Quantization Noise:** Spreading noise over wider bandwidth
- **Increased Effective Resolution:** Can achieve high bit accuracy with coarse quantizers
- **Cost-Effective Design:** Trade computation for analog precision

Applications:

- High-quality audio (CD, DAC)
- Precision measurement systems
- Software-defined radio

Oversampling Ratio

Definition: The **oversampling ratio** M is:

$$M = \frac{f_s}{2f_N}$$

where:

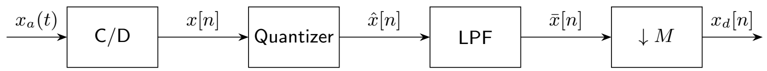
- f_s = actual sampling frequency
- f_N = Nyquist frequency (bandwidth of signal)

Key Idea: As M increases, we can:

- Reduce quantizer bits
- Improve signal-to-noise ratio (SNR)
- Use digital filtering to remove out-of-band noise

System Overview: Direct Quantization

Basic System:



$$T = \frac{\pi}{\Omega_N M}$$

Process:

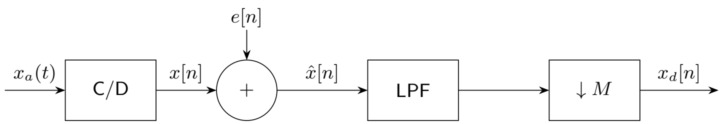
1. Sample at high rate: $T = \pi/(\Omega_N M)$
2. Quantize samples
3. Lowpass filter to remove out-of-band noise
4. Downsample by M to return to Nyquist rate

Additive Noise Model

DSP

Maxx Seminario

Quantizer Model: Replace quantizer with additive white noise



Noise Properties:

- White noise: $\phi_{ee}[m] = \sigma_e^2 \delta[m]$
- Uniform distribution over quantization step Δ
- Variance: $\sigma_e^2 = \frac{\Delta^2}{12}$
- Power spectral density: $\Phi_{ee}(e^{j\omega}) = \sigma_e^2$ for $|\omega| < \pi$

Signal and Noise Power Analysis

DSP

Maxx Seminario

Signal Power: Remains constant throughout processing

$$E\{x_a^2(t)\} = E\{x^2[n]\} = E\{x_d^2[n]\}$$

Quantization Noise Power:
Before filtering:

$$E\{e^2[n]\} = \sigma_e^2 = \frac{\Delta^2}{12}$$

After lowpass filtering and decimation:

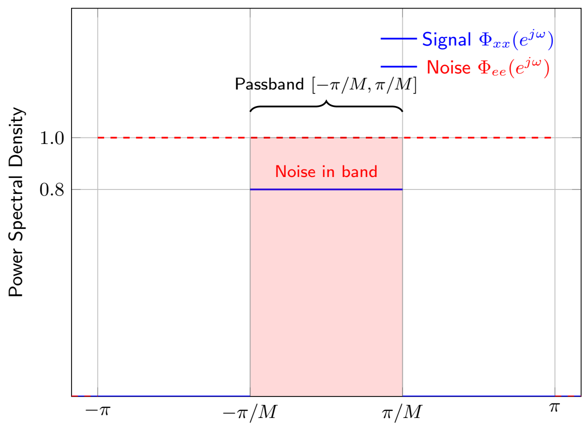
$$E\{x_{de}^2[n]\} = \frac{1}{2\pi}\int_{-\pi/M}^{\pi/M} \sigma_e^2 \, d\omega = \frac{\sigma_e^2}{M} = \frac{\Delta^2}{12M}$$

Key Result: Noise power reduced by factor of M!

Power Spectral Density Illustration

DSP

Maxx Seminario



Bit Reduction Trade-off

DSP

Maxx Seminario

Quantizer Step Size: For $(B+1)$-bit quantizer with maximum input $\pm X_m$:

$$\Delta = \frac{X_m}{2^B}$$

Output Noise Power:

$$P_{de} = E\{x_{de}^2[n]\} = \frac{1}{12M}\left(\frac{X_m}{2^B}\right)^2$$

Required Bits for Given Noise Power:

$$B = -\frac{1}{2}\log_2 M - \frac{1}{2}\log_2(12) - \frac{1}{2}\log_2 P_{de} + \log_2 X_m$$

Every doubling of oversampling ratio M saves **0.5 bits** of quantization!

Example: Bit Savings

Question: How much oversampling is needed to use a 12-bit quantizer instead of a 16-bit quantizer, while maintaining the same effective noise performance?

Solution:

- Bit reduction needed: $16 - 12 = 4$ bits
- Each doubling of M saves 0.5 bit
- Doublings needed: $4/0.5 = 8$
- Required oversampling: $M = 2^8 = 256$

Practical Consideration:

- $M = 256$ is too high for many applications
- Need better approach $\to$ **Noise Shaping**

Motivation for Noise Shaping

Problem with Direct Quantization:

- Quantization noise has *flat* spectrum
- All frequencies equally contaminated
- Need high M for significant bit reduction

Better Idea: Noise Shaping

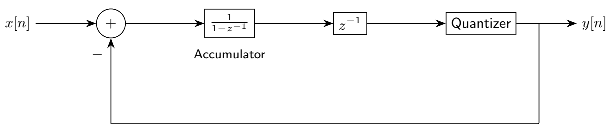
- **Shape** the noise spectrum using feedback
- Push noise power to *high frequencies*
- More noise removed by lowpass filter
- Same M gives better performance

Core Concept

Use feedback to make quantization error spectrum non-uniform, concentrating noise outside the signal band.

First-Order Delta-Sigma Modulator

System Structure:

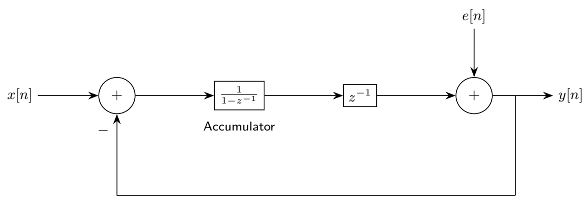


Also known as:

- Delta-Sigma ($\Delta\Sigma$) modulator
- Sigma-Delta ($\Sigma\Delta$) modulator
- 1-bit ADC with noise shaping

Linear Noise Model Analysis

Replace quantizer with additive noise:



Transfer Functions:
Signal path:

$$H_x(z) = \frac{Y(z)}{X(z)}\bigg|_{e[n]=0} = 1$$

Noise path:

$$H_e(z) = \frac{Y(z)}{E(z)}\bigg|_{x[n]=0} = 1 - z^{-1}$$

DSP

Maxx Seminario

Noise Shaping Effect

Output:

$$y[n] = x[n] + \hat{e}[n]$$

where $\hat{e}[n] = e[n] - e[n-1]$ (first difference of quantization noise)

Noise Power Spectrum:

$$\Phi_{\hat{e}\hat{e}}(e^{j\omega}) = \sigma_e^2 |H_e(e^{j\omega})|^2 = \sigma_e^2 |1 - e^{-j\omega}|^2 = \sigma_e^2 \cdot 4\sin^2(\omega/2)$$

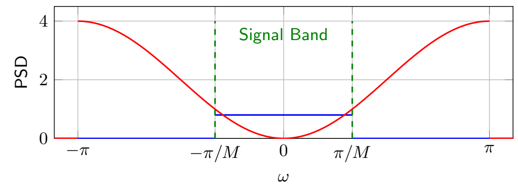
Key Properties:

- At DC ($\omega = 0$): $\Phi_{\hat{e}\hat{e}}(1) = 0$ (noise suppressed)
- At Nyquist ($\omega = \pi$): $\Phi_{\hat{e}\hat{e}}(e^{j\pi}) = 4\sigma_e^2$ (noise amplified)
- Total noise power: $E\{\hat{e}^2[n]\} = 2\sigma_e^2$ (increased from σ_e^2)
- But most noise is at high frequencies

DSP

Maxx Seminario

Shaped Noise Spectrum



Result: Much less noise in signal band compared to flat spectrum

DSP

Maxx Seminario

Noise Power After Decimation

After lowpass filtering and downsampling:

$$P_{de} = \frac{1}{2\pi} \int_{-\pi/M}^{\pi/M} \sigma_e^2 \cdot 4\sin^2(\omega/2)\, d\omega$$

For large M**:** $\sin(\omega/2M) \approx \omega/2M$

$$P_{de} \approx \frac{1}{2\pi} \int_{-\pi/M}^{\pi/M} \sigma_e^2 \cdot \left(\frac{\omega}{M}\right)^2 d\omega = \frac{\sigma_e^2 \pi^2}{12M^3} = \frac{\Delta^2}{12} \cdot \frac{\pi^2}{M^3}$$

Bit requirement:

$$B = -\frac{3}{2}\log_2 M + \log_2(\pi/\sqrt{6}) - \frac{1}{2}\log_2 P_{de} + \log_2 X_m$$

Result: Every doubling of M saves **1.5 bits** (vs. 0.5 bits without noise shaping)

DSP

Maxx Seminario

Comparison: Direct vs. Noise Shaping

Oversampling Ratio M	Direct Quantization	1st-Order Noise Shaping
4	1.0	2.2
8	1.5	3.7
16	2.0	5.1
32	2.5	6.6
64	3.0	8.1

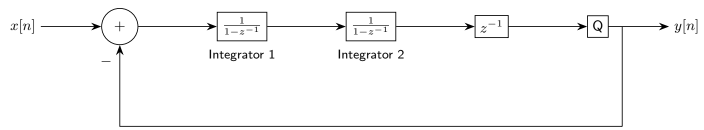
Equivalent Bit Savings (relative to no oversampling)

Example: With $M = 64$:

- Direct: 3 extra bits $\rightarrow$ 13-bit effective from 16-bit quantizer
- Noise shaping: 8.1 extra bits $\rightarrow$ **8-bit effective from 16-bit quantizer**

Second-Order Delta-Sigma Modulator

Add another integrator stage:



Noise Transfer Function:

$$H_e(z) = (1 - z^{-1})^2$$

Shaped Noise PSD:

$$\Phi_{\hat{e}\hat{e}}(e^{j\omega}) = \sigma_e^2 \cdot 16 \sin^4(\omega/2)$$

Even stronger high-frequency emphasis!

General p-th Order Noise Shaping

Order p modulator:

$$H_e(z) = (1 - z^{-1})^p$$

Noise PSD:

$$\Phi_{\hat{e}\hat{e}}(e^{j\omega}) = \sigma_e^2 \cdot [2\sin(\omega/2)]^{2p}$$

Noise power (large M approximation):

$$P_{de} \approx \frac{\sigma_e^2 \pi^{2p}}{(2p+1)M^{2p+1}}$$

Scaling Law

Order p noise shaping: doubling M saves $(p + 0.5)$ bits

- $p = 0$ (direct): 0.5 bits per doubling
- $p = 1$: 1.5 bits per doubling
- $p = 2$: 2.5 bits per doubling

Performance Comparison Table

Bit reduction relative to no oversampling:

Order	Oversampling Factor M				
p	4	8	16	32	64
0	1.0	1.5	2.0	2.5	3.0
1	2.2	3.7	5.1	6.6	8.1
2	2.9	5.4	7.9	10.4	12.9
3	3.5	7.0	10.5	14.0	17.5
4	4.1	8.5	13.0	17.5	22.0

Remarkable Example:

- 2nd-order, $M = 64$: 12.9 bit improvement
- Can achieve 14-bit accuracy with **1-bit quantizer**!
- (16 bits - 12.9 bits $\approx$ 3 bits $\rightarrow$ use simple comparator)

Practical Considerations

Advantages of Higher-Order Modulators:

- Dramatic improvement in resolution
- Can use very simple (1-bit) quantizers
- Excellent linearity

Challenges:

- **Stability:** Higher-order loops can oscillate
- **Overload:** Large signals can cause instability
- Require careful design and analysis

Alternative: MASH (Multi-stAge noise SHaping)

- Cascade multiple 1st-order stages
- Each stage shapes the previous stage's quantization error
- More stable than single high-order loop