Analog Integrated Systems Design

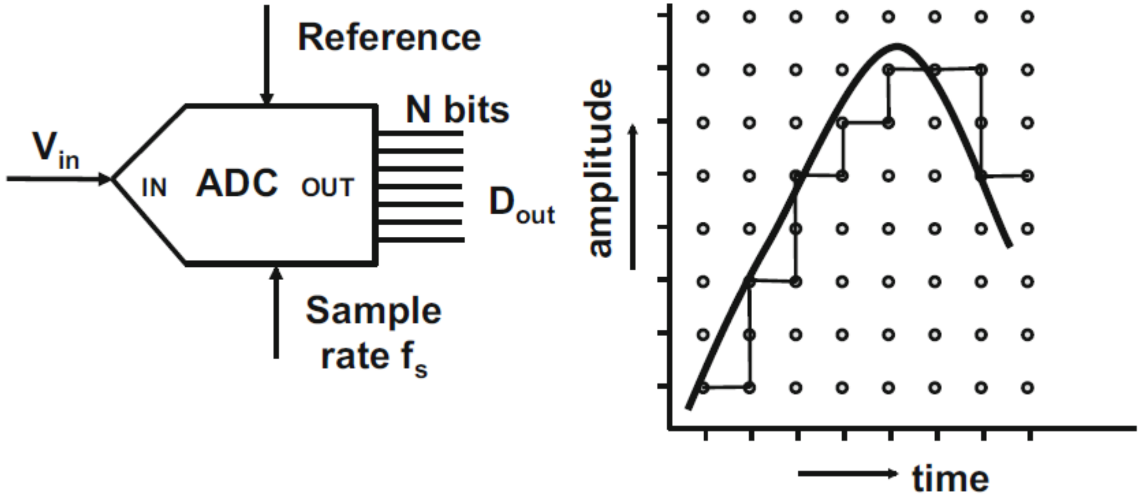
Lecture 03
Quantization

Dr. Hesham A. Omran

Integrated Circuits Laboratory (ICL)
Electronics and Communications Eng. Dept.
Faculty of Engineering
Ain Shams University

Quantization

- Sampling discretizes the analog signal in time domain.
- Quantization discretizes the analog signal in the voltage/amplitude domain.
  - Limited (finite) number of valid amplitude levels
- Quantization error: peak-to-peak $< \Delta$
  - Rounding (nearest level): $-0.5\Delta < error < 0.5\Delta$
  - Floor: $0 < error < \Delta$
  - Ceiling: $-\Delta < error < 0$

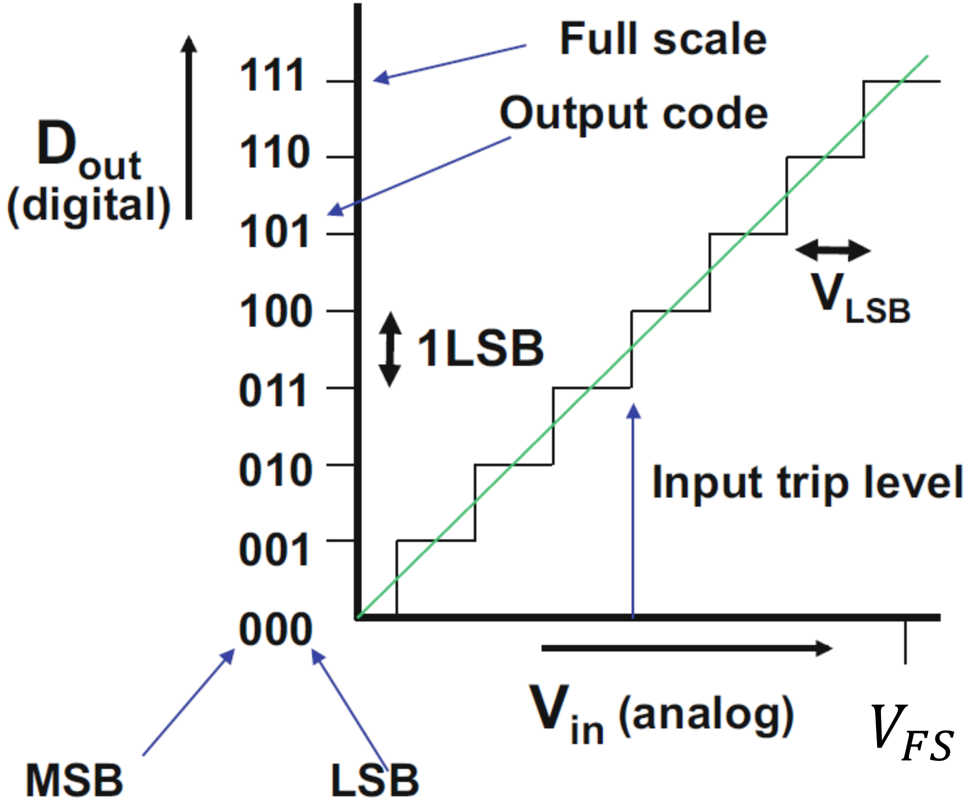


[M. Pelgrom, 2017]

Binary Representation

$$B_s = \sum_{i=0}^{i=N-1} b_i 2^i = b_0 2^0 + b_1 2^1 + b_2 2^2 \ldots + b_{N-1} 2^{N-1}$$

- N: word width, resolution, no. of bits
- Assume V_{ref} corresponds to 2^N
- No. of steps =
- b_0: Least significant bit (LSB)
- b_{N-1}: Most significant bit (MSB)
- $V_{LSB} = \Delta =$
- $N =$
- Full-scale (digital) =
- Full-scale (analog) =

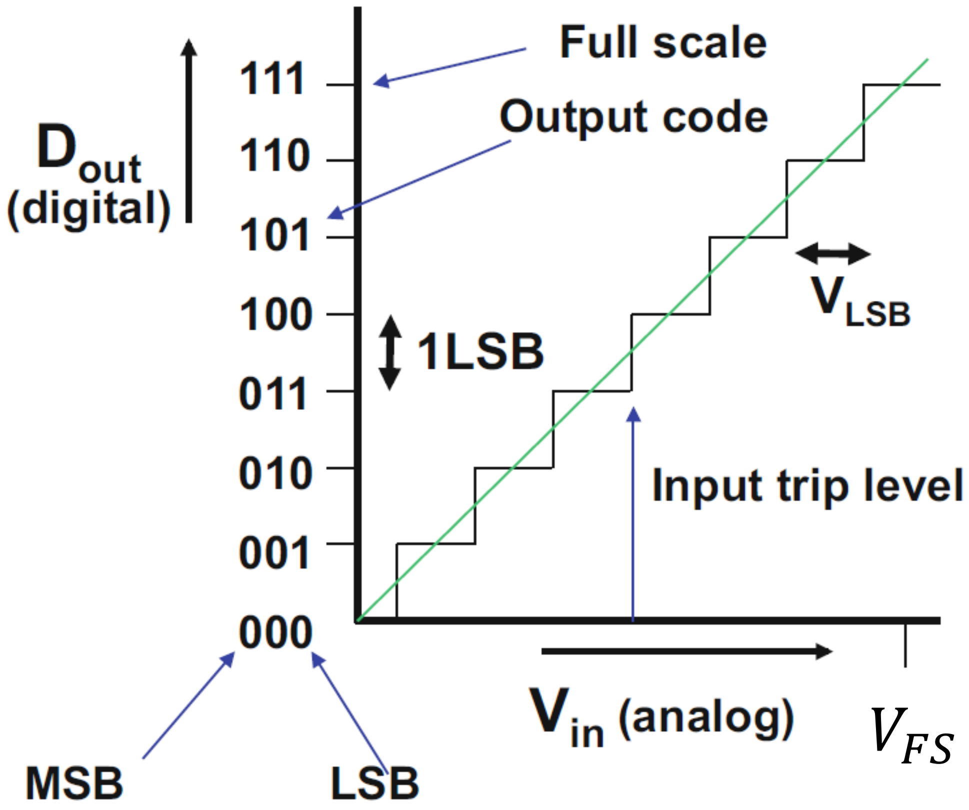


[M. Pelgrom, 2017]

Binary Representation

$$B_s = \sum_{i=0}^{i=N-1} b_i 2^i = b_0 2^0 + b_1 2^1 + b_2 2^2 \ldots + b_{N-1} 2^{N-1}$$

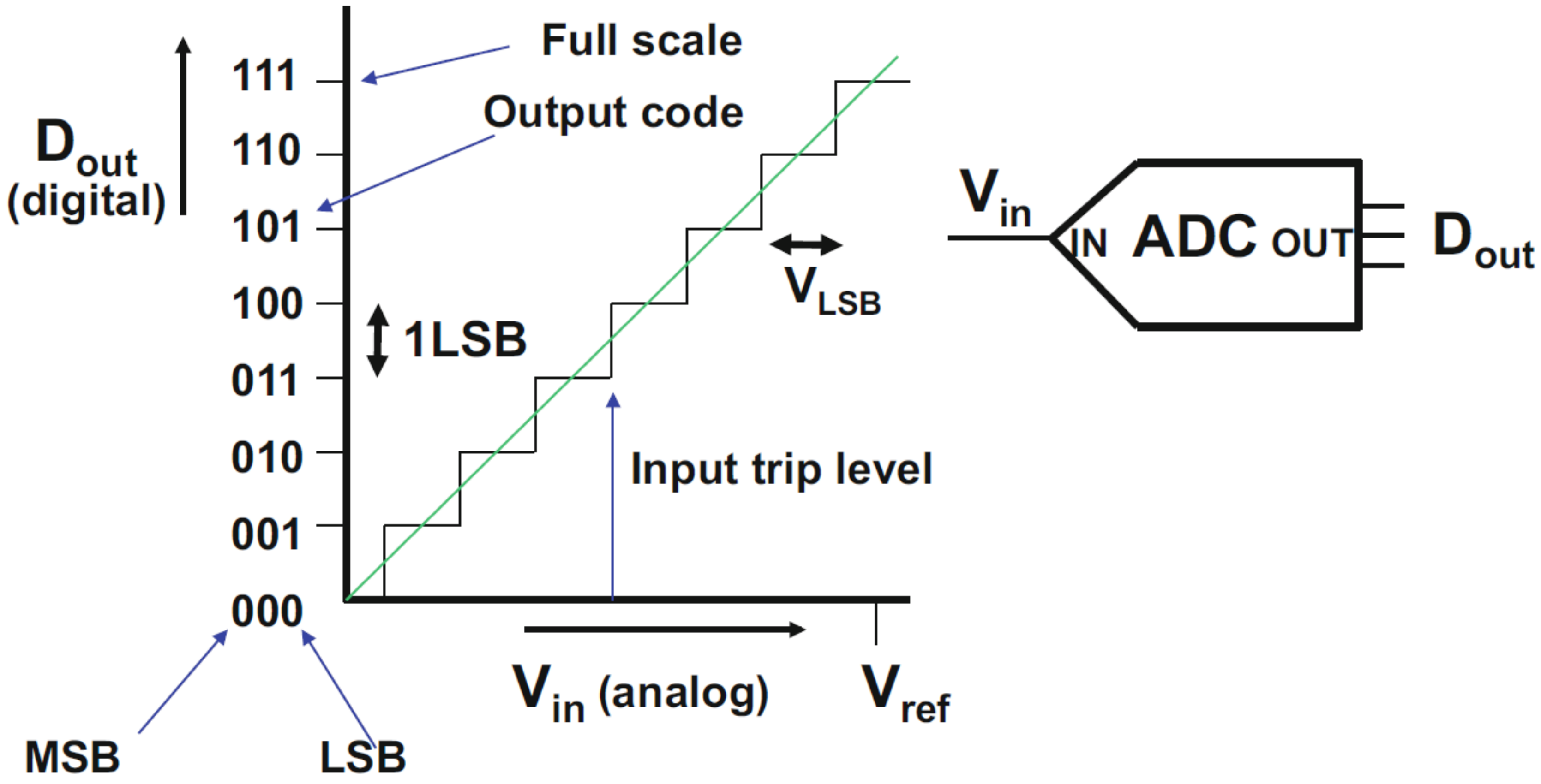
- N: word width, resolution, no. of bits
- Assume V_{ref} corresponds to 2^N
- No. of steps = $2^N - 1$
- b_0: Least significant bit (LSB)
- b_{N-1}: Most significant bit (MSB)
- $V_{LSB} = \Delta = \frac{V_{ref}}{2^N} = \frac{V_{FS}}{No.of\ steps} = \frac{V_{ref} - V_{LSB}}{2^N - 1}$
- $N = \log_2(V_{ref}/V_{LSB})$
- Full-scale (digital) = 111…111 = $2^N - 1$
- Full-scale (analog) = $V_{FS} = V_{ref} - V_{LSB}$



[M. Pelgrom, 2017]

ADC Parameters

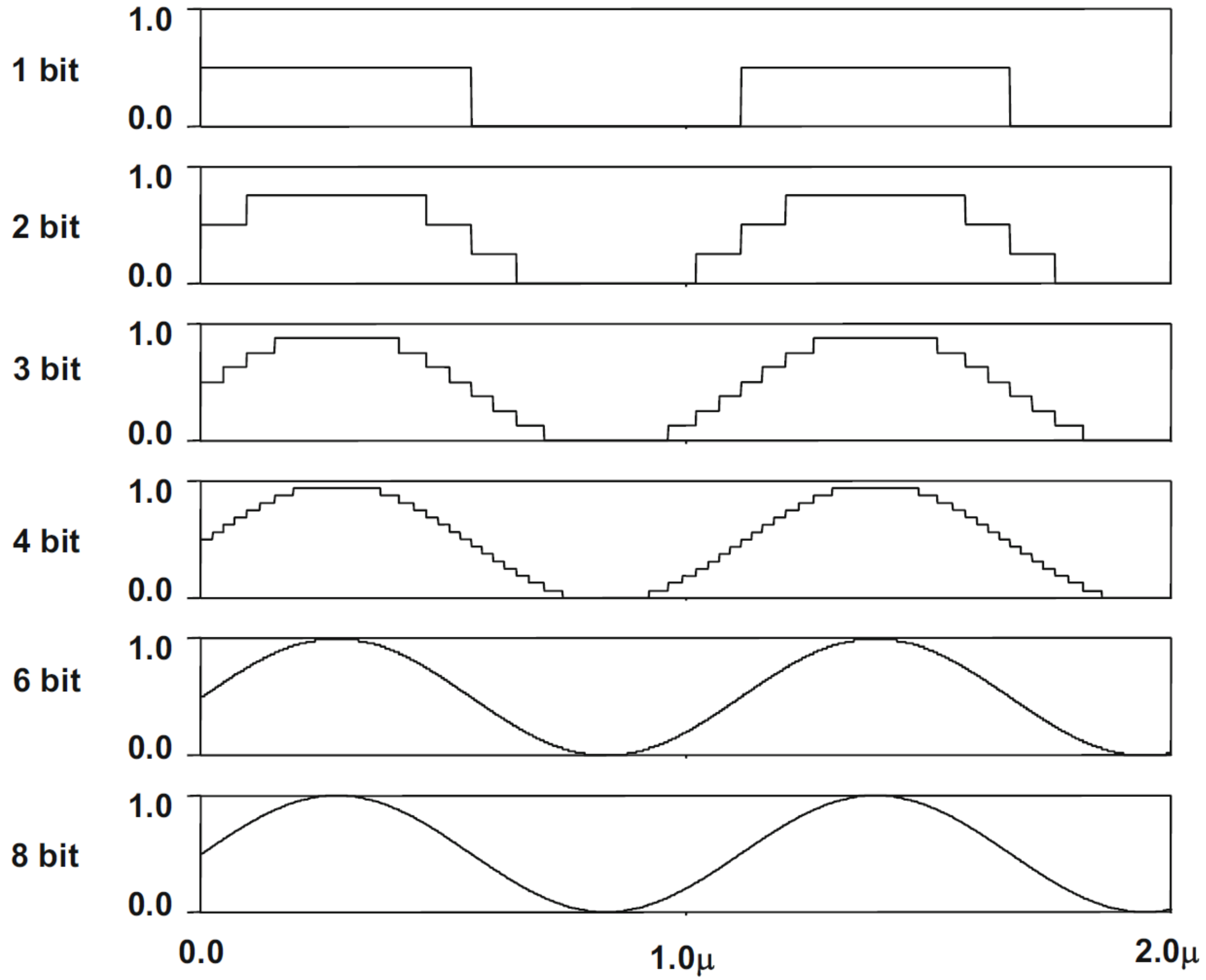
- Trip level = Decision level
  - The digital signal “trips” by one bit at the “trip level”.



 [M. Pelgrom, 2017]

Sine Wave Discretization Example

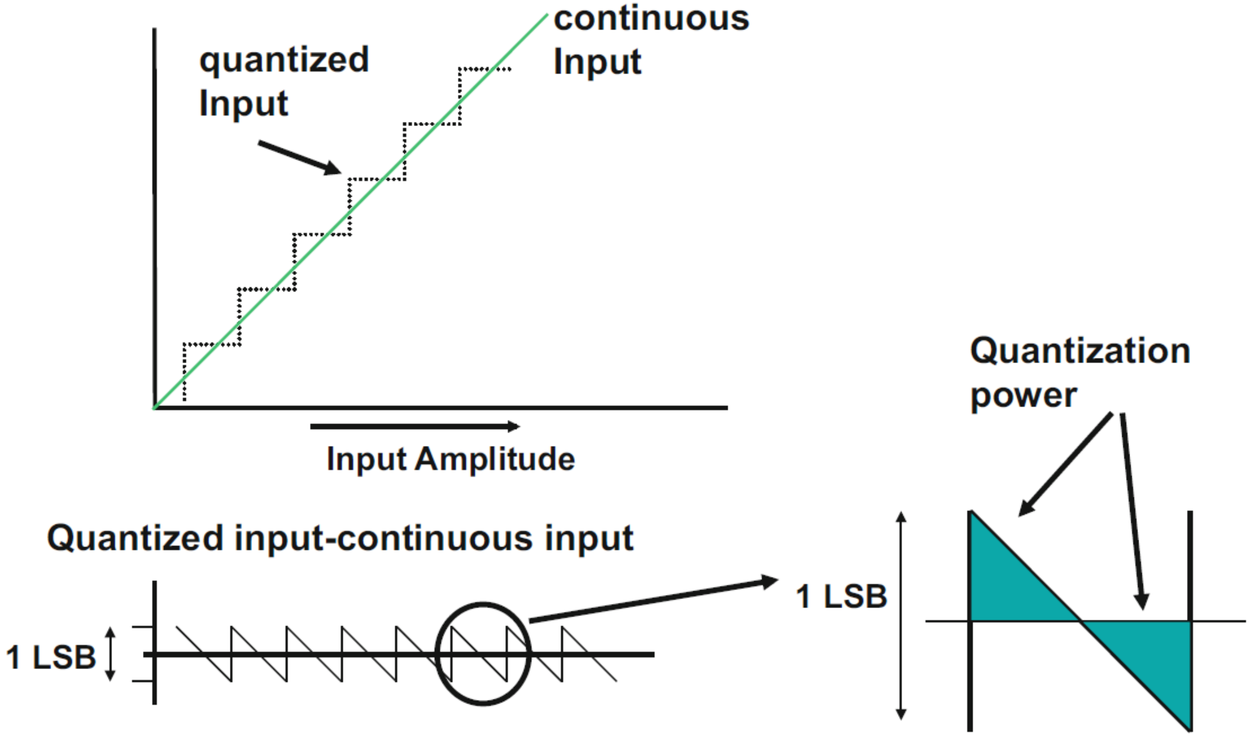
- The quantization error decreases as the number of bits increases



[M. Pelgrom, 2017]

Quantization Error (Noise)

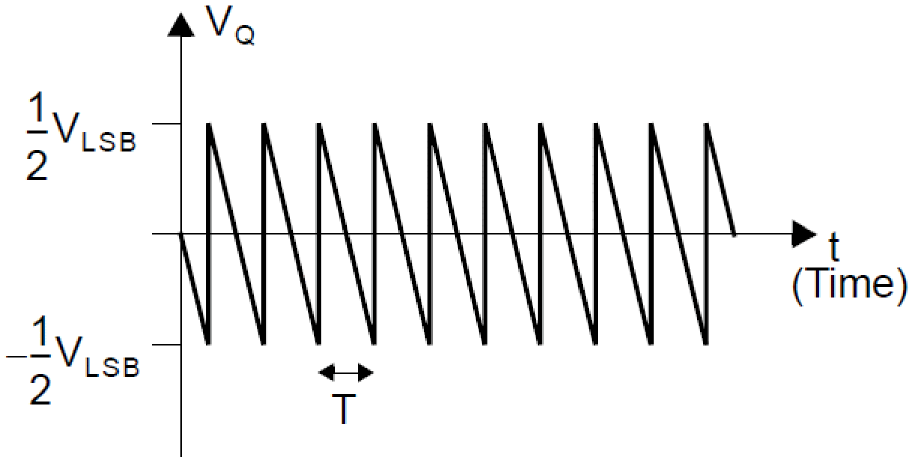
- For low-resolution ADC, the quantization error will show as distortion components (harmonics of the input).
- For N > 6-bit, the quantization error can be approximated as:
  - Uniformly distributed PDF (from -0.5 LSB to 0.5 LSB)
  - White noise in the frequency domain (from 0 to $f_s/2$)



[M. Pelgrom, 2017]

Quantization Noise: Deterministic Approach

- Assume linear ramp input → quantization error is sawtooth wave

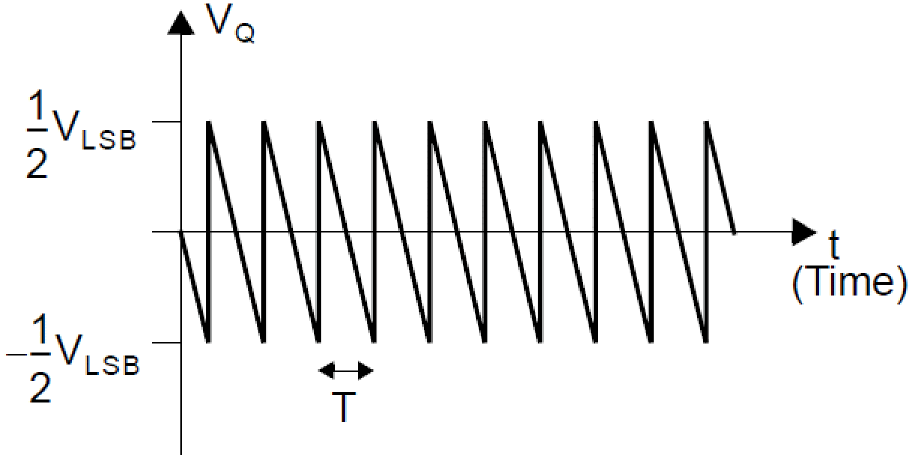


$$V_{Q(rms)} = \left[\frac{1}{T}\int_{-T/2}^{T/2} V_Q^2\, dt\right]^{1/2} = \left[\frac{1}{T}\int_{-T/2}^{T/2} V_{LSB}^2\left(\frac{-t}{T}\right)^2 dt\right]^{1/2}$$

[Johns & Martin, 2012]

Quantization Noise: Deterministic Approach

- Assume linear ramp input → quantization error is sawtooth wave



$$V_{Q(rms)} = \left[\frac{1}{T}\int_{-T/2}^{T/2} V_Q^2\,dt\right]^{1/2} = \left[\frac{1}{T}\int_{-T/2}^{T/2} V_{LSB}^2\left(\frac{-t}{T}\right)^2 dt\right]^{1/2}$$

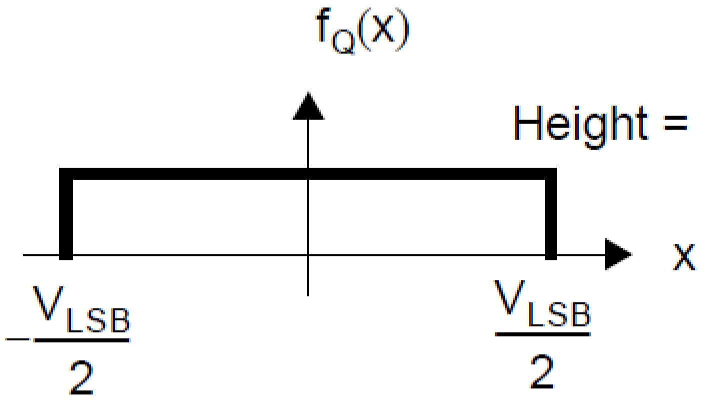
$$= \left[\frac{V_{LSB}^2}{T^3}\left(\frac{t^3}{3}\Big|_{-T/2}^{T/2}\right)\right]^{1/2}$$

$$V_{Q(rms)} = \frac{V_{LSB}}{\sqrt{12}}$$

[Johns & Martin, 2012]

Quantization Noise: Stochastic Approach

- Assume uniformly distributed random error



$$\text{Height} = \frac{1}{V_{LSB}} \quad \text{so that} \quad \int_{-\infty}^{\infty} f_Q(x)\,dx = 1$$

$$V_{Q(avg)} = \int_{-\infty}^{\infty} x f_Q(x)\,dx = \frac{1}{V_{LSB}}\left(\int_{-V_{LSB}/2}^{V_{LSB}/2} x\,dx\right) = 0$$

$$V_{Q(rms)} = \left[\int_{-\infty}^{\infty} x^2 f_e(x)\,dx\right]^{1/2} = \left[\frac{1}{V_{LSB}}\left(\int_{-V_{LSB}/2}^{V_{LSB}/2} x^2\,dx\right)\right]^{1/2} = \frac{V_{LSB}}{\sqrt{12}}$$

[Johns & Martin, 2012]

Signal-to-Quantization Noise Ratio

$$SQNR = 10\log\left(\frac{Signal\ Power}{Quantization\ Power}\right) = 20\log\left(\frac{V_{sigrms}}{V_{Qnrms}}\right)$$

$$Signal\ Power = \frac{\left(\frac{2^N V_{LSB}}{2}\right)^2}{2} = \frac{2^{2N} V_{LSB}^2}{8}$$

$$Quantization\ Power = \frac{V_{LSB}^2}{12}$$

$$SQNR = 10\log\left(\frac{Signal\ Power}{Quantization\ Power}\right) = 10\log\left(\frac{3}{2}2^{2N}\right)$$

Signal-to-Quantization Noise Ratio

$$SQNR = 10\log\left(\frac{Signal\ Power}{Quantization\ Power}\right) = 20\log\left(\frac{V_{sigrms}}{V_{Qnrms}}\right)$$

$$Signal\ Power = \frac{\left(\frac{2^N V_{LSB}}{2}\right)^2}{2} = \frac{2^{2N} V_{LSB}^2}{8}$$

$$Quantization\ Power = \frac{V_{LSB}^2}{12}$$

$$SQNR = 10\log\left(\frac{Signal\ Power}{Quantization\ Power}\right) = 10\log\left(\frac{3}{2}2^{2N}\right)$$

$$\boldsymbol{SQNR = 6.02 \times N + 1.76 \quad [dB]}$$

Verifying the SQNR Formula

- The approximate formula gives good results, especially for high-resolution ADC
- Quantization noise dominates up to around N = 14-bit

Resolution	Simulated SN_QR	$6.02N + 1.76\,dB$
1	6.31 dB	7.78 dB
2	13.30 dB	13.80 dB
3	19.52 dB	19.82 dB
4	25.60 dB	25.84 dB
5	31.66 dB	31.86 dB
6	37.71 dB	37.88 dB
7	43.76 dB	43.90 dB
8	49.80 dB	49.92 dB

[M. Pelgrom, 2017]

Resolution, LSB, and dBFS

- Each bit adds 6 dB to SNR.

RESOLUTION N	2^N	VOLTAGE (10V FS)	ppm FS	% FS	dB FS
2-bit	4	2.5 V	250,000	25	– 12
4-bit	16	625 mV	62,500	6.25	– 24
6-bit	64	156 mV	15,625	1.56	– 36
8-bit	256	39.1 mV	3,906	0.39	– 48
10-bit	1,024	9.77 mV (10 mV)	977	0.098	– 60
12-bit	4,096	2.44 mV	244	0.024	– 72
14-bit	16,384	610 μV	61	0.0061	– 84
16-bit	65,536	153 μV	15	0.0015	– 96
18-bit	262,144	38 μV	4	0.0004	– 108
20-bit	1,048,576	9.54 μV (10 μV)	1	0.0001	– 120
22-bit	4,194,304	2.38 μV	0.24	0.000024	– 132
24-bit	16,777,216	596 nV*	0.06	0.000006	– 144

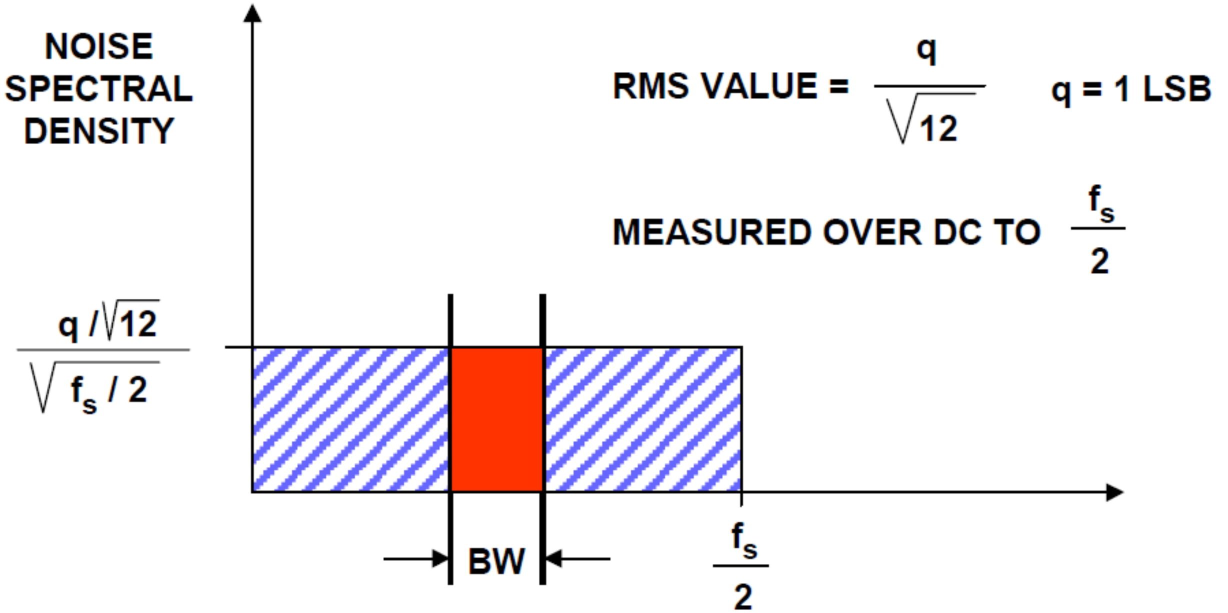
[W. Kester, 2005]

Oversampling/Processing Gain (1)

- Quantization power is uniformly spread from 0 to $f_s/2$.
- If only part of the spectrum is useful, some quantization power can be filtered out (digital filtering).

$$P_{Qn-total} = \frac{V_{LSB}^2}{12} = S_Q(f) \times \frac{f_s}{2} \quad \rightarrow \quad S_Q(f) = \frac{V_{LSB}^2}{12} \times \frac{2}{f_s}$$

$$P_{Qn-red} = S_Q(f) \times BW = \frac{V_{LSB}^2}{12} \times \frac{BW}{f_s/2}$$



[W. Kester, 2005]

Oversampling/Processing Gain (2)

- Quantization power is uniformly spread from 0 to $f_s/2$.
- If only part of the spectrum is useful, some quantization power can be filtered out (digital filtering).
- Select a bandwidth (BW) out of the available spectrum (0 to $f_s/2$):

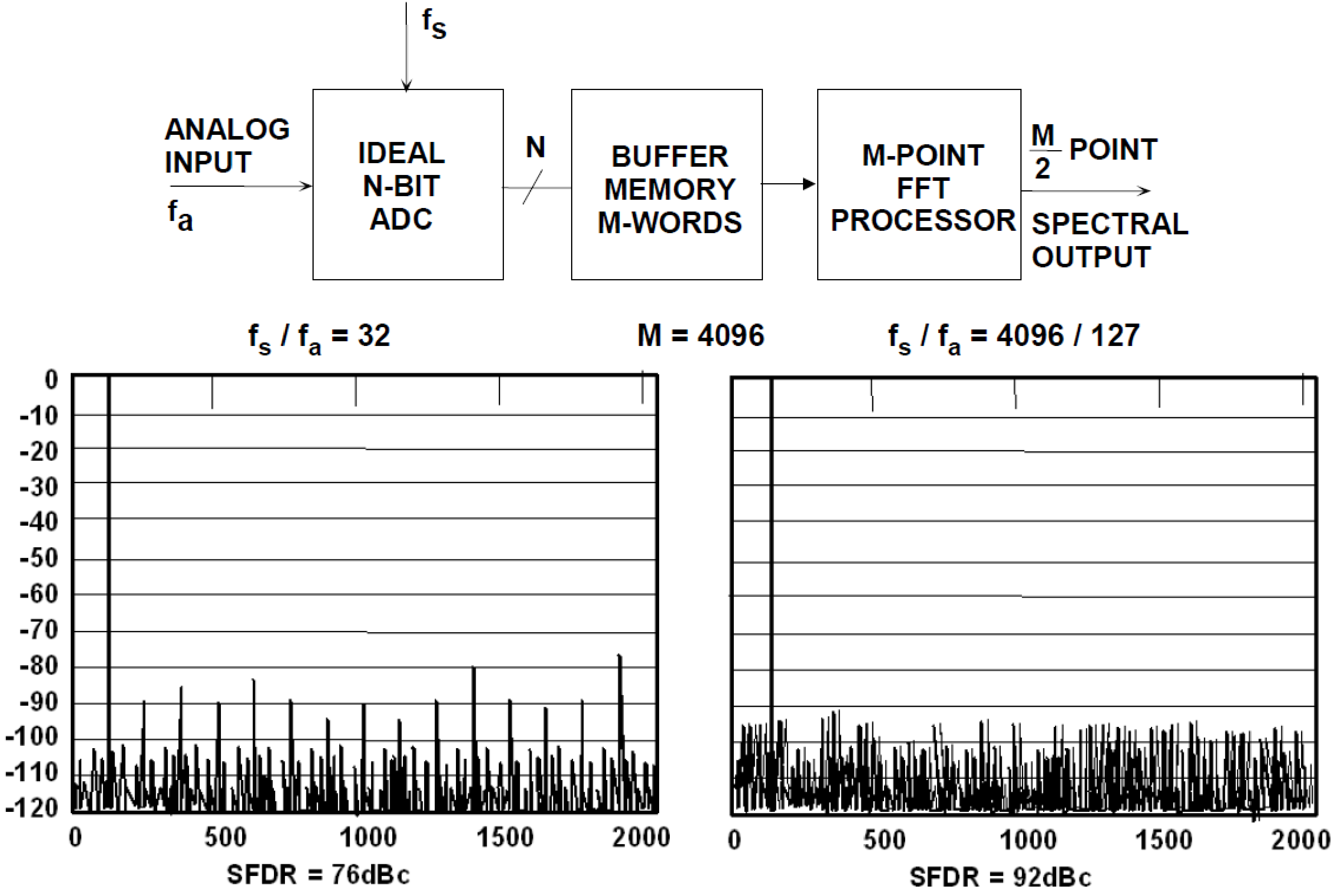
$$SQNR = 10\log\left(\frac{Signal\ Power}{Quantization\ Power \times \frac{BW}{f_s/2}}\right)$$

$$SQNR = 6.02 \times N + 1.76 + 10\log\left(\frac{f_s/2}{BW}\right)$$

Quantization Noise Spectrum

- In most practical applications, the input to the ADC is a band of frequencies + noise
  - The quantization noise tends to be random white noise.
  - Uniformly distributed from 0 to $f_s/2$.
- In ADC testing/simulation, a pure single-tone sine wave is used.
  - The sampling frequency may be correlated to the test-tone.
    - The quantization noise power may also become correlated to the input signal.
    - Quantization noise may appear as harmonic distortion.
- This is a testing artifact.
  - It should be avoided so that the true ADC distortion is measured, rather than the correlated quantization noise.

Quantization Error/Noise Spectrum

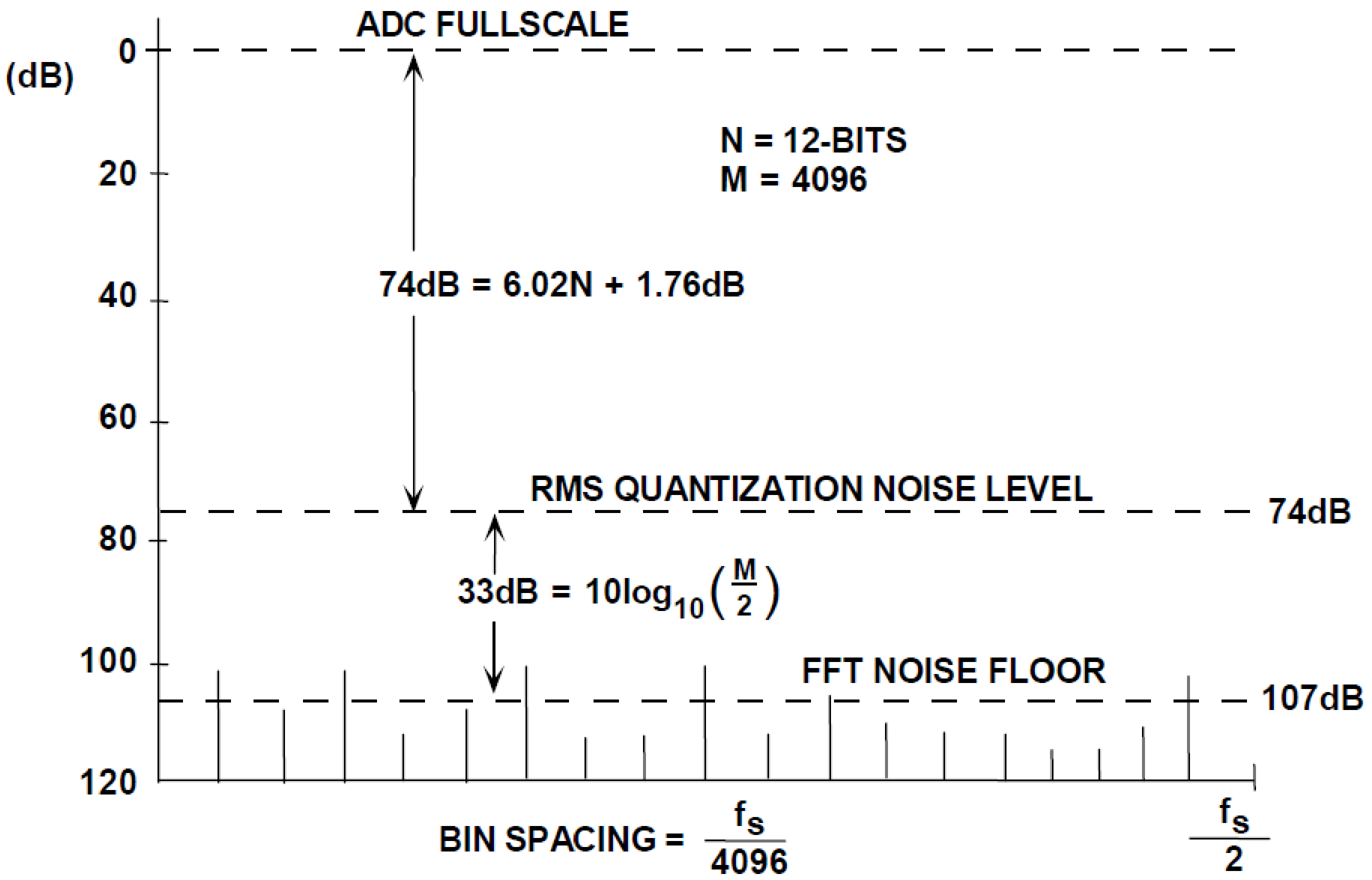


[W. Kester, 2005]

FFT Noise Floor

$$\text{Quantization Power} = \frac{V_{LSB}^2}{12} = S_Q \times \frac{f_s}{2} = S_{Q,FFT} \times \frac{M}{2} \rightarrow S_{Q,FFT} = \frac{V_{LSB}^2}{12} / \left(\frac{M}{2}\right)$$

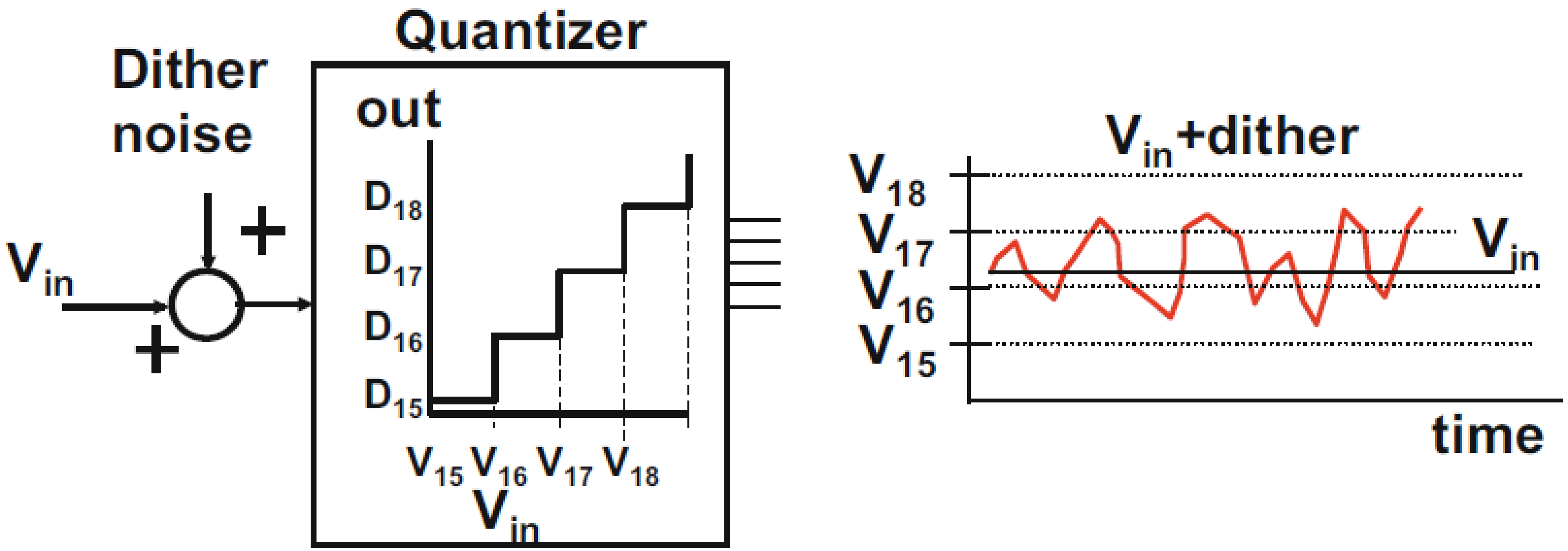
$$\text{FFT Noise Floor} = 10\log S_{Q,FFT} = 10\log\frac{V_{LSB}^2}{12} - 10\log\frac{M}{2}$$



 [W. Kester, 2005]

Dithering: Is no noise good noise?

- The addition of a random signal allows to determine the value of a DC-signal at greater accuracy than the quantization process allows.
- Additional signal processing like averaging can then lead to resolution improvement for low-frequency signals.



- FYI: http://www.analog.com/en/analog-dialogue/articles/adc-input-noise.html

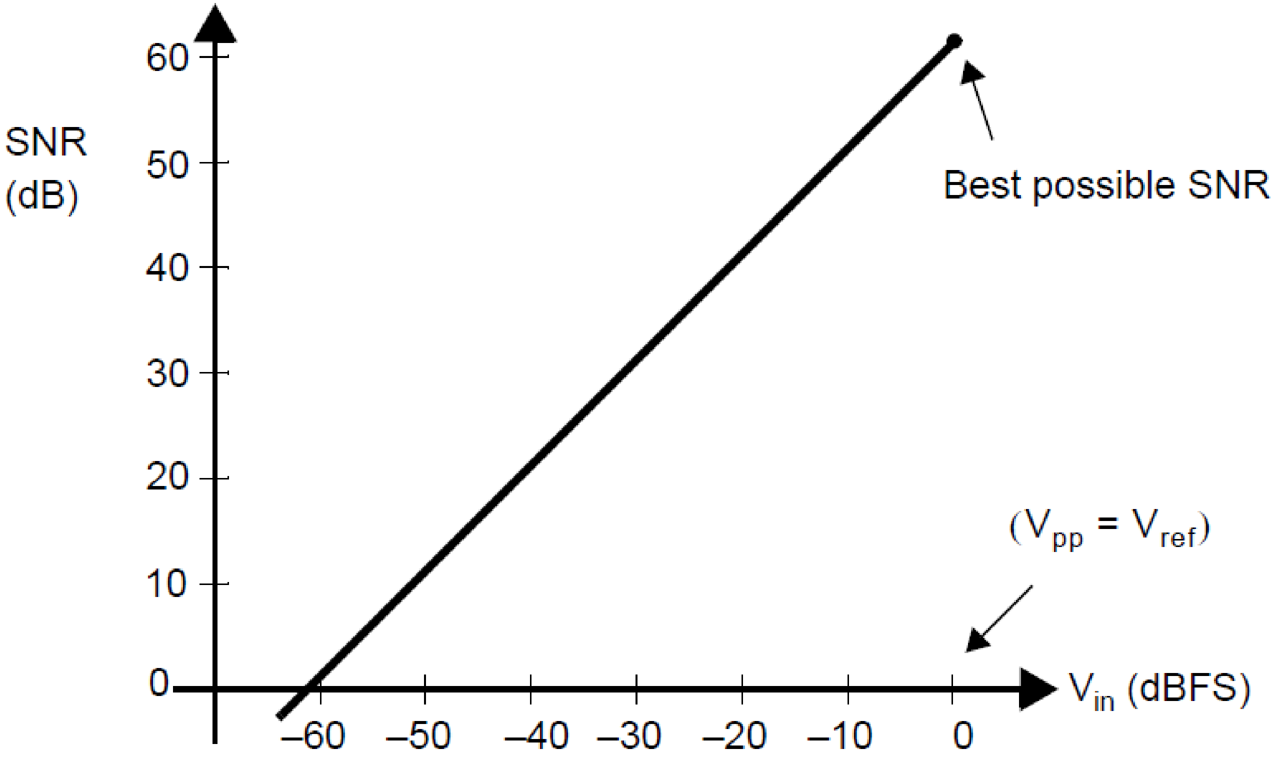
References

- M. Pelgrom, Analog-to-Digital Conversion, Springer, 3rd ed., 2017.
- W. Kester, The Data Conversion Handbook, ADI, Newnes, 2005.
- B. Boser and H. Khorramabadi, EECS 247 (previously EECS 240), Berkeley.
- B. Murmann, EE 315, Stanford.
- Y. Chiu, EECT 7327, UTD.

Thank you!

SNR vs Input Level

- SNR depends on input level
  - Ex: Ideal 10-bit ADC



[Johns & Martin, 2012]