# Quantization Noise

I.Nelson
AP, ECE
SSN College of Engineering

# Finite word length effects in Digital filters

- The following errors arise due to quantization of numbers:

1. Input Quantization error
2. Product Quantization error
3. Coefficient Quantization error

## 1. Input Quantization error

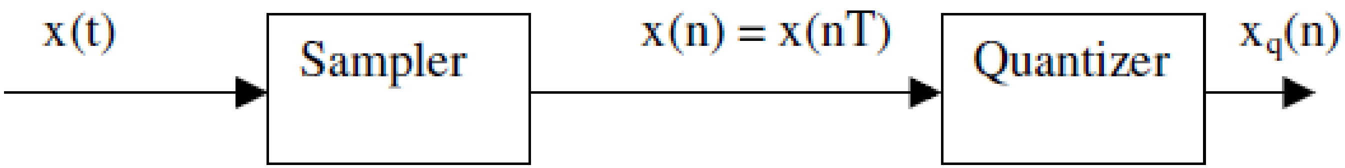
- The quantization error arises when a continuous signal is converted into digital value and it is given by,

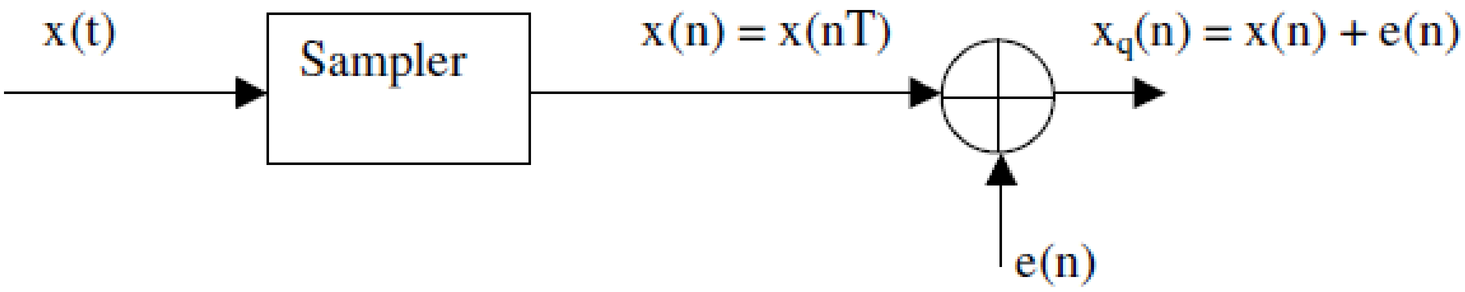
  $e(n) = x_q(n) - x(n)$

  where $x_q(n)$ is the sampled quantized value and $x(n)$ is the sampled unquantized value

## Steady State Input Noise power

- The Quantization error is commonly viewed as additive noise signal, i.e.,
  $x_q(n) = x(n) + e(n)$



- We assume that the error has the following properties:

1. The error is a random variable and the error sequence is the sample sequence of a stationary random process.
2. The error sequence is uncorrelated with x(n) and other signals in the system.
3. The error is a white noise process with uniform probability distribution over the range of quantization error.

- In case of rounding, the error e(n) lies between -Δ/2 and Δ/2 with equal probability. (Assume Δ = $2^{-b}$). The variance of e(n) is given by

$$\sigma_\varepsilon^2 = E[e^2(n)] - E^2\left[e(n)\right]$$

$$= \int_{-\Delta/2}^{\Delta/2} e^2(n)\, p(e)\ de - (0)^2$$

$$= \frac{1}{\Delta}\int_{-\Delta/2}^{\Delta/2} e^2(n)\ de = \frac{\Delta^2}{12} = \frac{2^{-2b}}{12}$$

- In case of truncating, the error e(n) lies between -Δ and 0 with equal probability. The variance of e(n) is given by

$$\sigma_\varepsilon^2 = E[e^2(n)] - E^2\left[e(n)\right]$$

$$= \int_{-\Delta}^{0} e^2(n)\,p(e)\;de - \left(-\frac{\Delta}{2}\right)^2$$

$$= \frac{1}{\Delta}\int_{-\Delta}^{0} e^2(n)\;de \quad - \quad \frac{\Delta^2}{4} = \frac{\Delta^2}{3} - \quad \frac{\Delta^2}{4} = \frac{2^{-2b}}{12}$$

- From the two cases, the steady state noise power due to input quantization is same.

**Note:**

- If the input signal is x(n) and it's variance is $\sigma^2_x$, then the ratio of signal power to noise power (SNR) is,

$$\frac{\sigma_x^2}{\sigma_e^2} = \frac{\sigma_x^2}{2^{-2b}/12} = (12)(2^{2b})(\sigma_x^2)$$

- When it is expressed in a log scale, then,

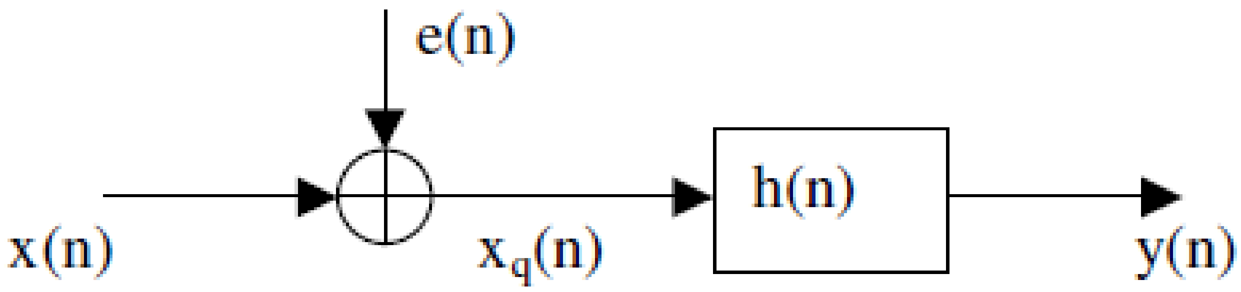
$$10\log\frac{\sigma_x^2}{\sigma_e^2} = 10\log(12) + 10\log(2^{2b}) + 10\log(\sigma_x^2)$$

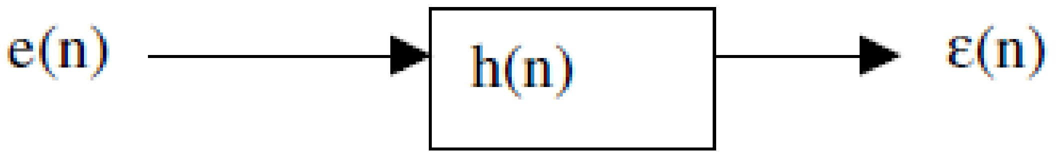
$$= 10.79 + 6.02b + 10\log(\sigma_x^2)$$

- From the above expression we can find that the SNR increases approximately 6dB for the each bit added to register length.

## Steady State Output Noise power

- Due to A/D conversion noise, the quantized input to a digital system with impulse response h(n) can be represented as,



- Let ε(n) be the output noise due to quantization of the input. Then we get,

$$\varepsilon(n) = e(n) * h(n)$$

$$= \sum_{k=0}^{n} h(k)\, e(n-k)$$

- The variance of any term in the above sum is equal to $\sigma_e^2\, h^2(n)$.

- The variance of the sum of the independent random variable is the sum of their variances.

- If the quantization errors are assumed to be independent at different sampling instances, then the variance of the output is,

$$\sigma_{\varepsilon}^{2}(n)=\sigma_{e}^{2}\sum_{n=0}^{k}h^{2}(n)$$

- To find steady state variance, extend k to infinity, i.e.,

$$\sigma_{\varepsilon}^{2}(n)=\sigma_{e}^{2}\sum_{n=0}^{\infty}h^{2}(n)$$

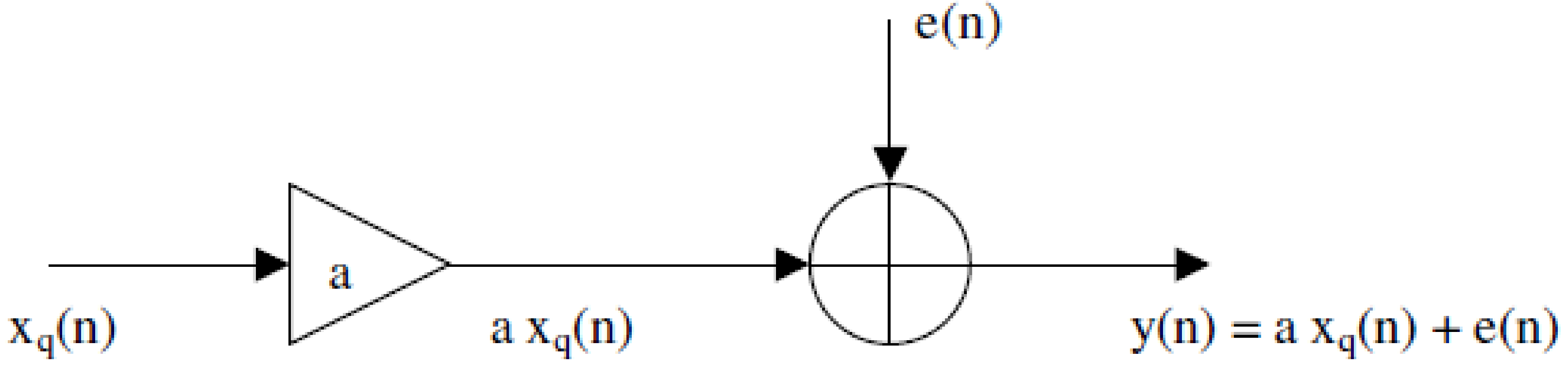
- Using the Parseval's theorem, the steady state output noise variance due to the quantization error is given by,

$$\sigma_{\varepsilon}^{2}(n)=\sigma_{e}^{2}\sum_{n=0}^{\infty}h^{2}(n)=\frac{\sigma_{e}^{2}}{2\pi j}\oint_{c}H(z)H(z^{-1})z^{-1}dz$$

where the closed contour of integration is around the unit circle $|z| = 1$ in which case only the poles that lie inside the unit circle are evaluated using the residue theorem.

## 2. Product Quantization Error

- The model for fixed point roundoff noise can be given as



- The roundoff noise sample is a zero mean random variable with a variance $2^{-2b}/12$, where 'b' is the number of bits used to represent the variables.

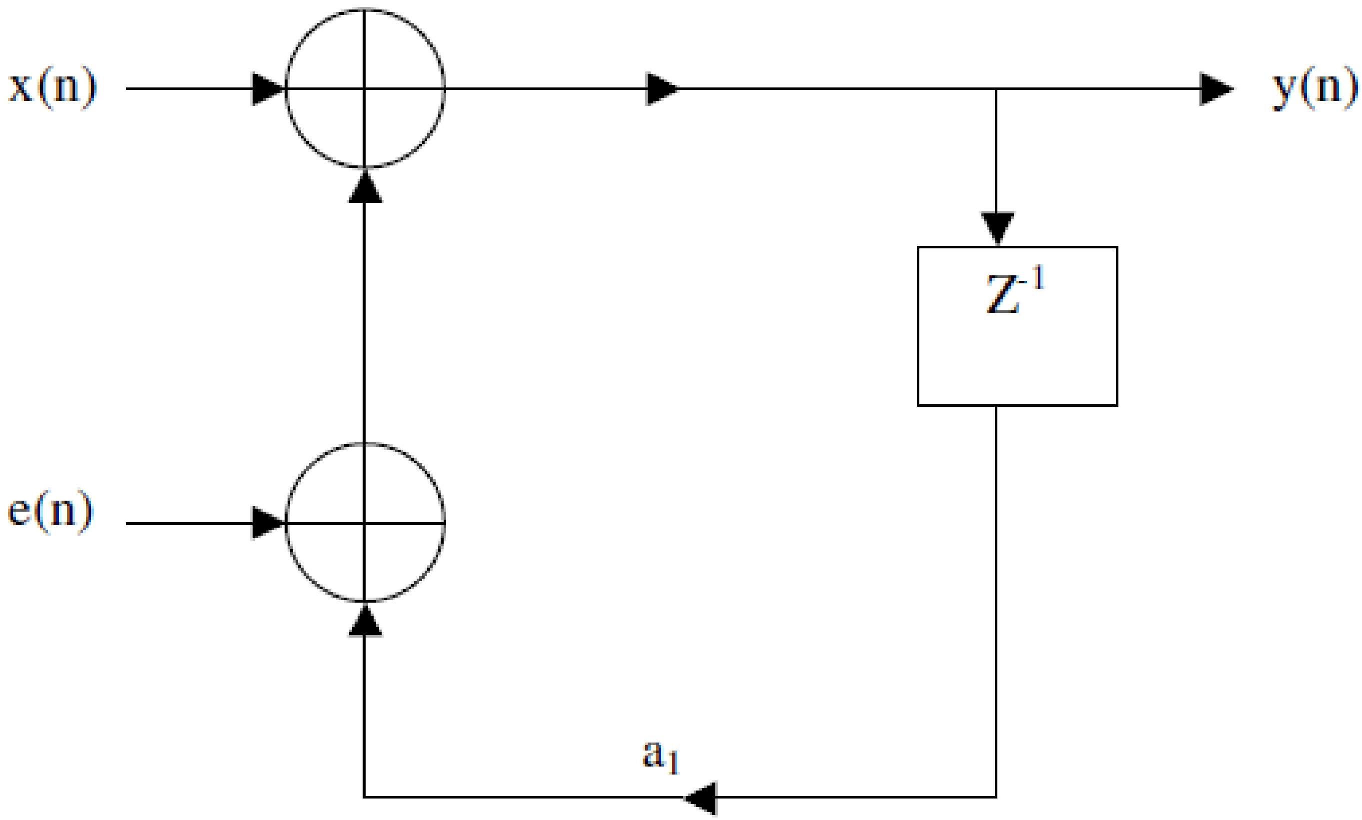
- In order to model the effects of rounding due to multiplication in digital filter, certain assumptions must be made.

1. For any n, the error sequence e(n) is uniformly distributed over the range $-2^{-2b}/2$ and $2^{-2b}/2$. This implies that mean value of e(n) is zero and its variance is $\sigma_e^2=2^{-2b}/12$.

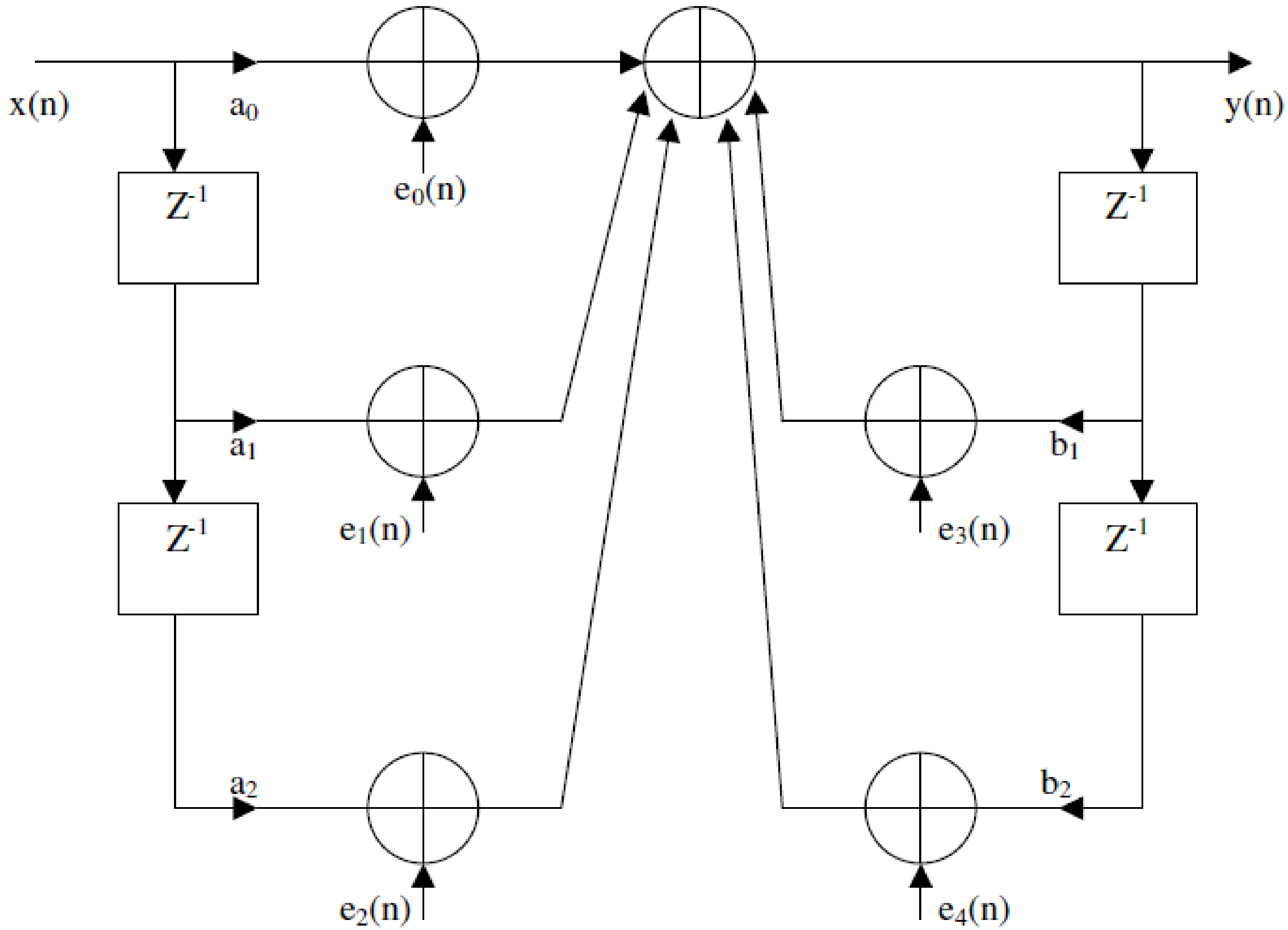
2. The error sequence e(n) is a stationary white noise sequence.

3. The error sequence e(n) is uncorrelated with the signal sequence x(n). Thus each noise source is modeled as a discrete stationary white random process with a power density spectrum of $2^{-2b}/12$.

# Quantization noise model for a first order system



# Quantization noise model for a second order system



- For a second order direct form digital filter every finite precision multiplier is replace by an ideal multiplier and an additive roundoff noise.

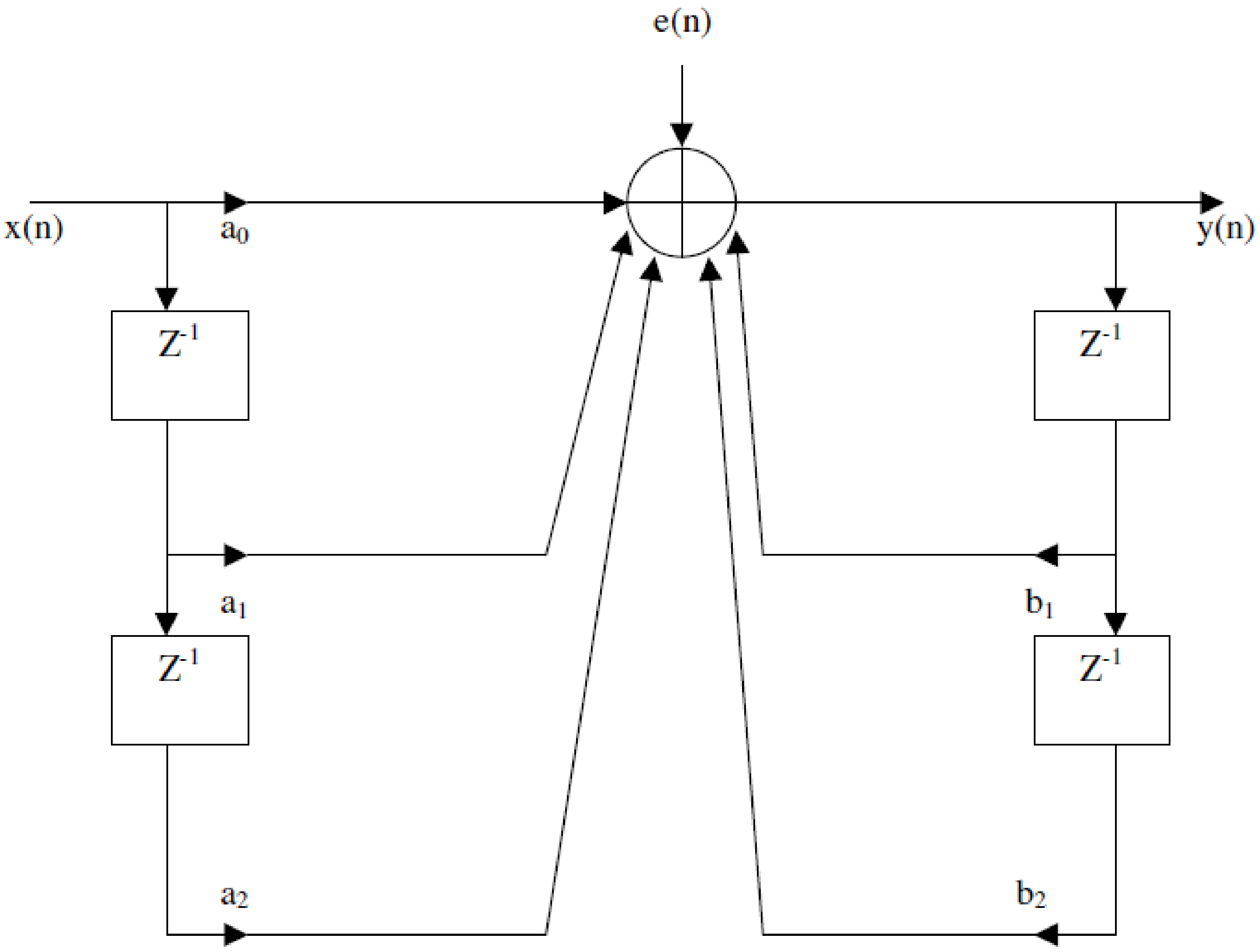
- Since all the noise sources are added at the same point in the filter, all these sources can be replaced by single noise source

$$e(n) = e_1(n) + e_2(n) + e_3(n) + e_4(n) + e_5(n)$$

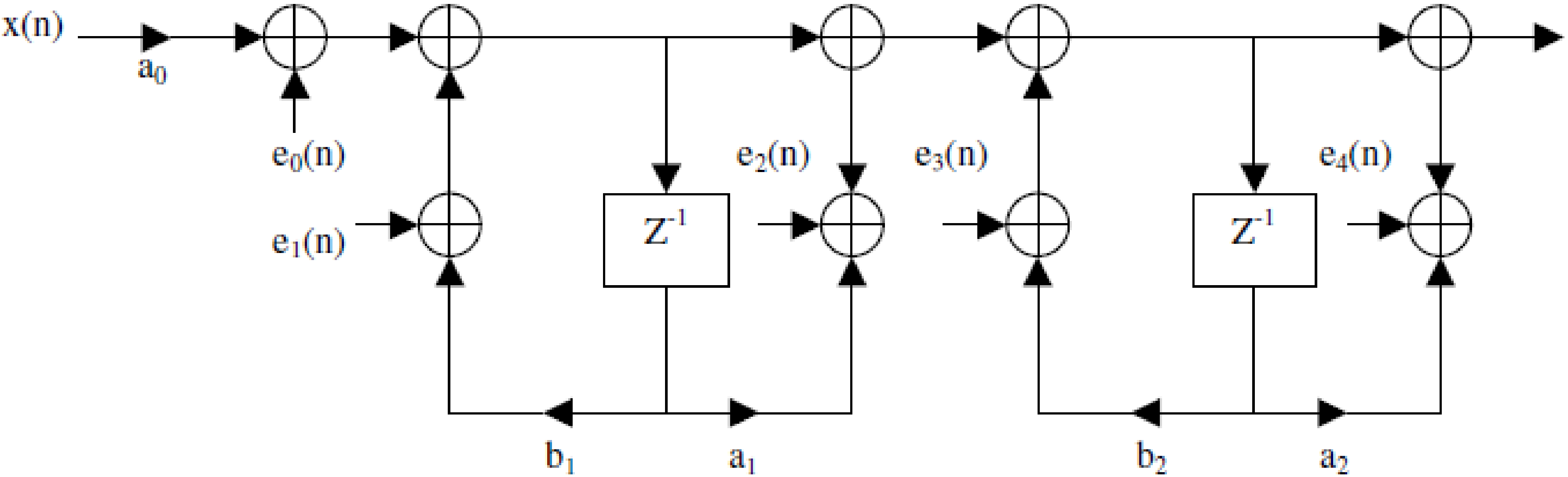
with zero mean and variance is

$$\sigma^2 = \sigma_0^2 + \sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2 = 5 * \sigma_0^2$$

Hence, the second order system can be realized as,



If the second order filter is realized as a cascade of two first order systems as shown below.



- There are five noise sources in this realization. The noise sources are added at different points and they do not see the same noise transfer function, that is, the transfer function from the noise source to the filter output.

- The total output variance can be obtained by adding the individual variances.

- Consider the $k^{th}$ noise source $e_k(n)$. If $h_k(n)$ is the filter's impulse response from the noise source to the filter output, the response due to noise source $e_k(n)$ can be obtained by convolution as,

$$\varepsilon_k(n) = \sum_{m=0}^{\infty} h_k(m) e_k(n-m)$$

- The variance of $\varepsilon_k(n)$ is,

$$\sigma_{0k}^2 = E\left[\varepsilon^2(n)\right] - E^2\left[\varepsilon(n)\right]$$

$$= E\left[\sum_{m=0}^{\infty} h_k(m)e_k(n-m)\sum_{l=0}^{\infty} h_k(l)e_k(n-l)\right] - (0)^2$$

$$= \sigma_e^2 \sum_{m=0}^{n} h_k^2(m)$$

- The total steady state noise variance is,

$$\sigma_0^2 = \sum_k \sigma_{0k}^2$$

## 3. Coefficient Quantization error

- When the filter coefficients are quantized, the frequency response of the actual filter deviates from that which would have been obtained with an infinite word length representation and the filter may fail to meet the desired specifications.

- If the poles of the desired filter are close to the unit circle, then those of the filter with quantized coefficients may lie just outside the unit circle leading to stability.