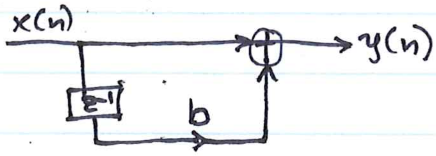
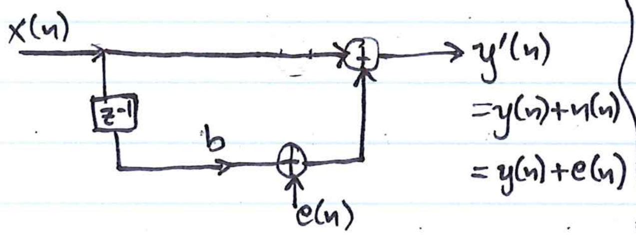
**Example 1** Let $y(n) = x(n) + b\,x(n-1)$ i.e. first-order FIR system

DIRECT REALIZATION:




**Ex.1** Fixed point roundoff noise analysis

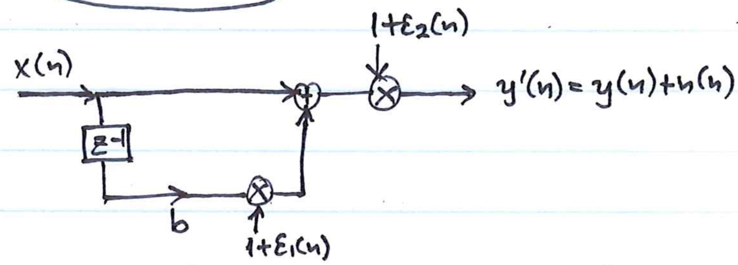



Error at output: $n(n) = e(n)$

So mean $m_n = m_e = 0$

variance: $\sigma_n^2 = \sigma_e^2 = \frac{1}{12} 2^{-2b}$

**Ex.1** Floating point roundoff noise analysis




$$y'(n) = [1+\varepsilon_2(n)]\,[x(n) + b\,x(n-1)(1+\varepsilon_1(n))] =$$

$$= x(n) + bx(n-1) + b\varepsilon_1(n)x(n-1) + \varepsilon_2(n)x(n) + b\varepsilon_2(n)x(n-1) + \cancel{b\varepsilon_1(n)\varepsilon_2(n)x(n-1)}$$

(Remove higher order error terms as being negligible)

$$\simeq \underbrace{x(n) + b\,x(n-1)}_{y(n)} + \underbrace{\cdots\cdots\cdots}_{n(n)}$$

where $n(n) = bx(n-1)[\varepsilon_1(n)+\varepsilon_2(n)] + \varepsilon_2(n)x(n)$

$m_n = E\{n(n)\} = \cdots\cdots = 0$

$$\sigma_n^2 = E\{n^2(n)\} = E\{ b^2x^2(n-1)\varepsilon_1^2(n) + b^2x^2(n-1)\varepsilon_2^2(n) + \underbrace{b^2x^2(n-1)\,2\varepsilon_1(n)\varepsilon_2(n)}_{\to 0} + \cancel{bx^2(n-1)\varepsilon_1(n)\varepsilon_2(n)} + \varepsilon_2^2(n)x^2(n) + 2bx(n)x(n-1)\varepsilon_2^2(n)\}$$

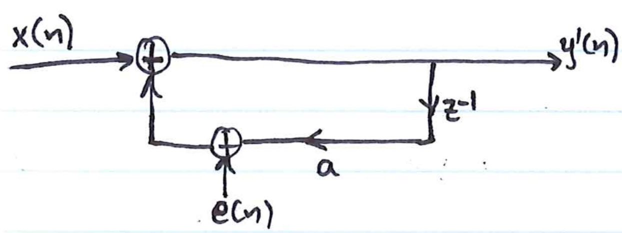
Let $E\{x(n)\} = 0$ $E\{x^2(n)\} = \sigma_x^2$ for all $n$

$$\sigma_n^2 = 2b^2\sigma_x^2\sigma_\varepsilon^2 + \sigma_x^2\sigma_\varepsilon^2 + 2\sigma_\varepsilon^2 \underbrace{E\{x(n)x(n-1)\}}_{\simeq \sigma_x^2}$$

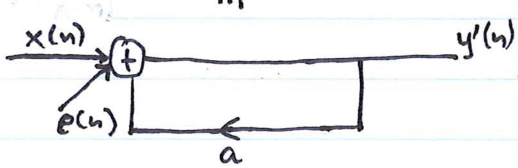
$$= \sigma_x^2\sigma_\varepsilon^2\,[2b^2 + 3] \simeq \cdots$$

**Ex.2** Let $y(n) = a\,y(n-1) + x(n) = h(n) * x(n)$

$(a<1)$ FIXED POINT ANALYSIS




$\equiv$




So $y'(n) = [x(n)+e(n)] * h(n) = \underbrace{x(n)*h(n)}_{y(n)} + \underbrace{e(n)*h(n)}_{n(n)}$

Now $h(n) = Z^{-1}\left[\frac{1}{1-az^{-1}}\right] = a^n u(n)$

Thus:

## Ex. 2

Assume that scaling is applied ~~prior to~~ prior to filtering to avoid overflow in addition.

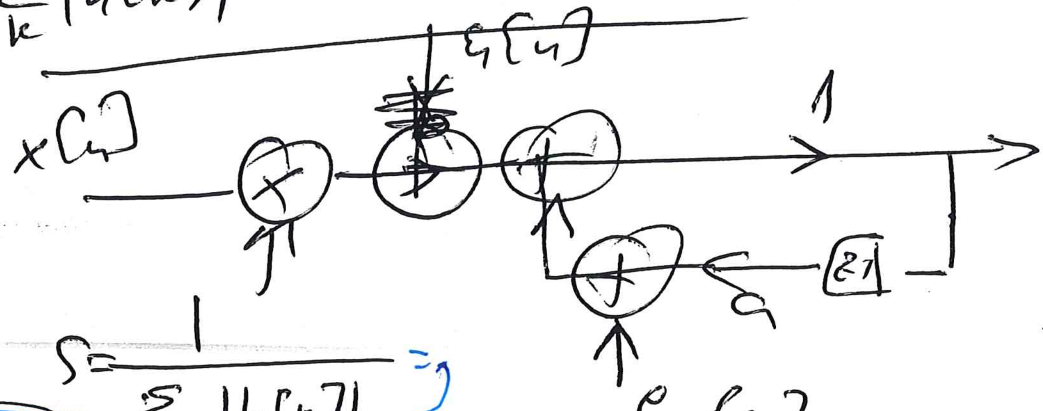
Since we want $|y[n]| < X_{MAX} \Rightarrow$

Then $\left| \sum_{k} h[k] x[n-k] \right| < X_{MAX} \Rightarrow$

$$< \sum_k |h[k]| \underbrace{|x[n-k]|} \leq \underbrace{X_{MAX}}_{1} \sum_k |h[k]|$$

Then. $\dfrac{|y[n]|}{\sum_k |h[k]|} \leq X_{MAX}$

So now




$$s = \frac{1}{\sum_k |h[k]|} \Rightarrow$$

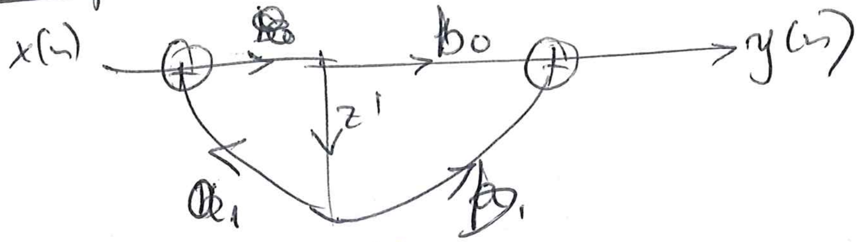
$$\boxed{\frac{1}{\sum_k |a|^n} = \frac{1}{\frac{1}{1-|a|}} = 1-|a|}$$

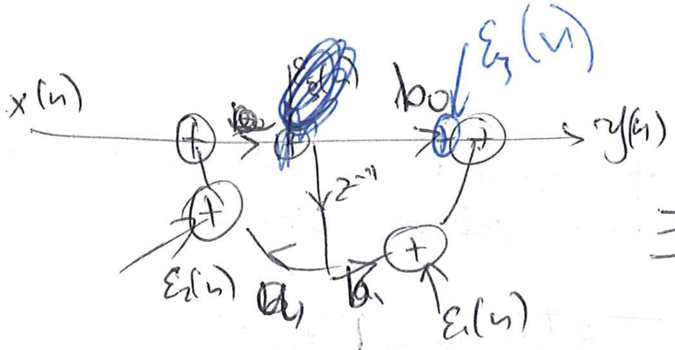
$$y[n] = \underbrace{s\,y[n]}_{\text{signal}} + \underbrace{e[n] * h[n]}_{\text{noise}}$$

$$SNR = \frac{P_y \cdot (1-|a|)^2}{2\sigma_e^2 \cdot \frac{1}{1-|a|^2}} = \frac{P_y}{2\sigma_e^2}(1-|a|)^2(1-|a|^2)$$

Example (fixed point)

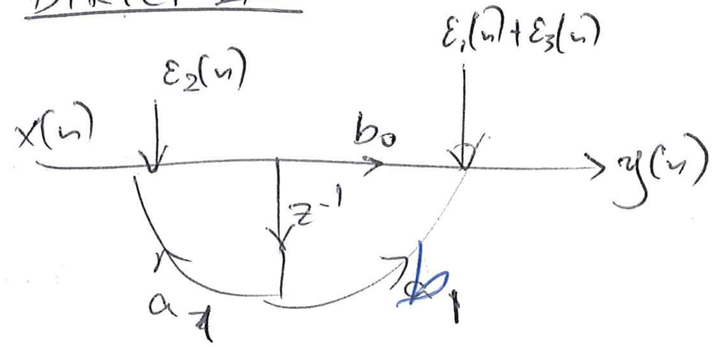
DIRECT II




$\equiv$

DIRECT II

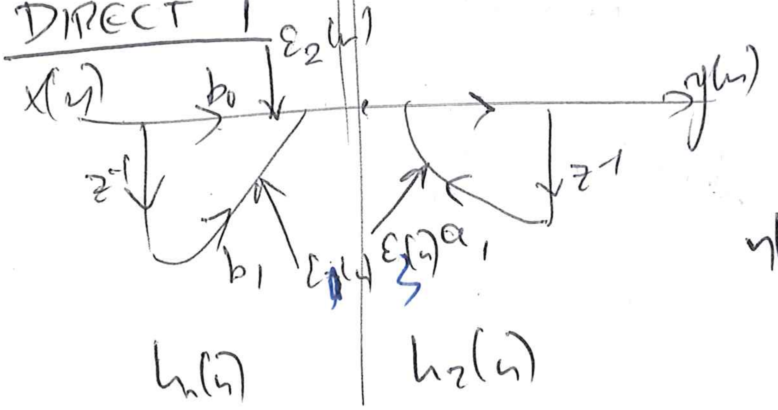



So $n(n) = \varepsilon_1(n) + \varepsilon_3(n) + \varepsilon_2(n) * h(n)$

$$\Rightarrow \boxed{\begin{aligned} m_n &= 0 \\ \sigma_n^2 &= 2\sigma_\varepsilon^2 + \sigma_\varepsilon^2 \cdot \sum h^2(n) = 2\sigma_\varepsilon^2 + \sigma_\varepsilon^2\left[b_0^2 + (b_0 a + b_1)^2 \sum_{n=1}^{\infty} a^{2(n-1)}\right] \end{aligned}}$$

$$= 2\sigma_\varepsilon^2 + \sigma_\varepsilon^2 \cdot \left\{b_0^2 + \frac{(b_0 a + b_1)^2}{\cancel{a}(1-a^2)}\right\}$$

DIRECT I




$$n(n) = \varepsilon_3(n) * h_2(n) + \left(\varepsilon_2(n) + \varepsilon_1(n)\right) * h_2(n)$$

So $m_n = 0$

$$\sigma_n^2 = 3\sigma^2 \cdot \sum h_2^2(n) = 3\sigma^2 \frac{1}{1-a^2}$$