

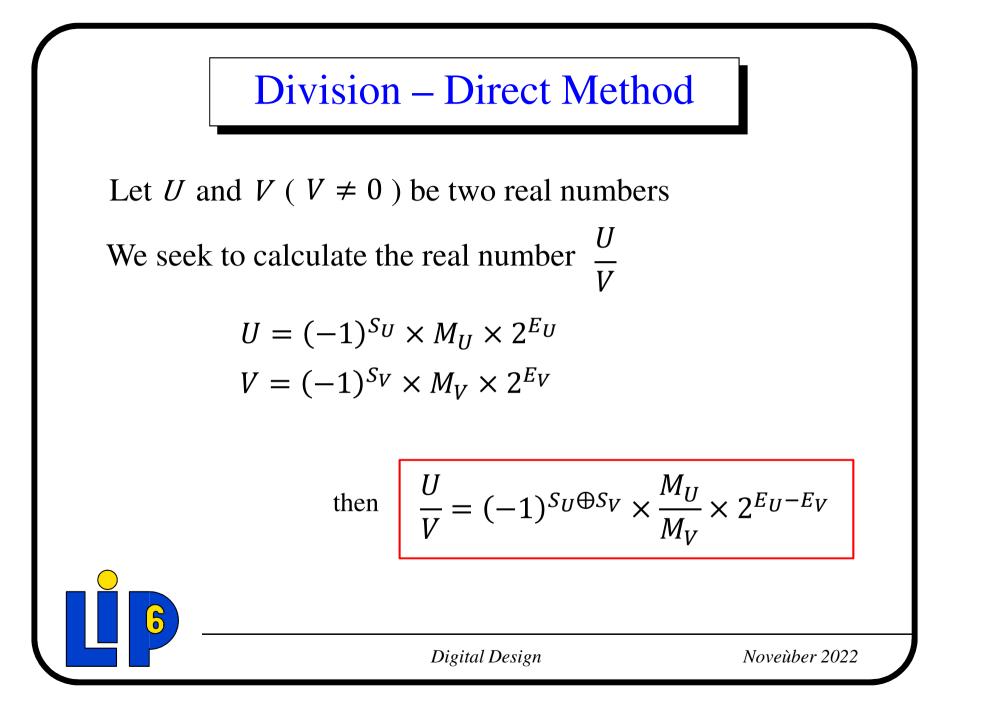
In *IEEE* floating point standard a real number is represented as :

$$(-1)^S \times M \times 2^E$$

In 32-bit representation :

$$S \in \{0, 1\}$$
 $M \in [1, 2[$ 
 $Or$ 
 $M \in [0, 2[$ 
 $E \in \{-126, \cdots, 127\}$ 
 if
  $E = -127$ 

6	normal		sub-normal	,
		Digital Design		Noveùber 2022



If *U* or *V* are sub-normal numbers, the calculation of  $\frac{U}{V}$  may lead to a loss of precision

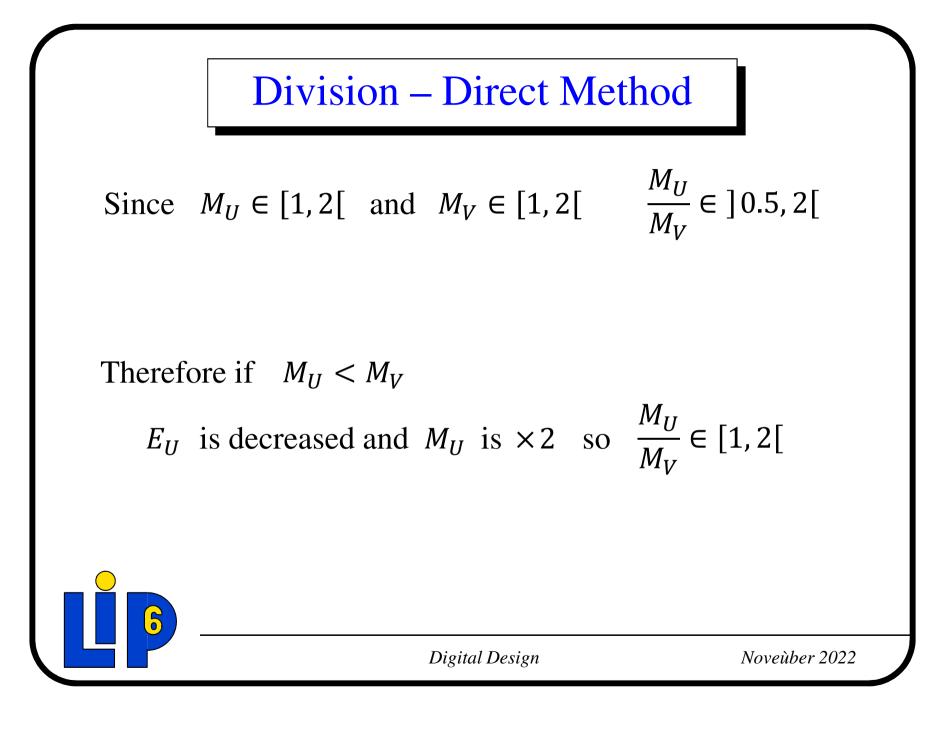
Therefore if 
$$M_U = 0$$
  $\frac{U}{V} = 0$ 

and if  $M_U$  or  $M_V \in ]0, 1[$ 

 $E_U$  or  $E_V$  are decreased and  $M_U$  or  $M_V$  are  $\times 2$  until they can fit within [1,2[

Digital Design

Noveùber 2022



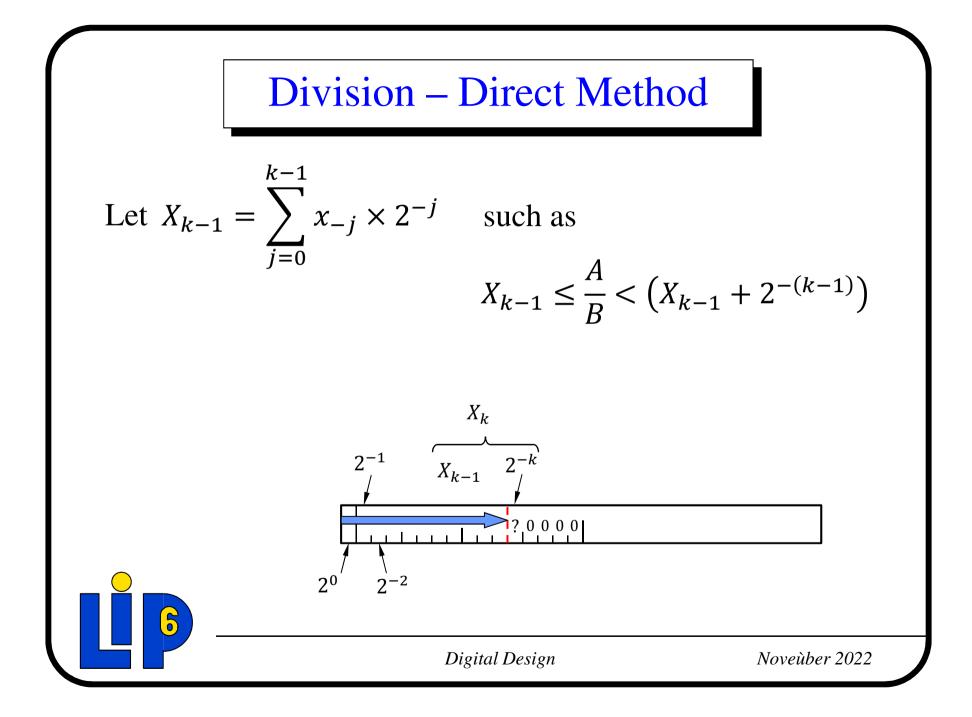
Then, the problem can be stated as :

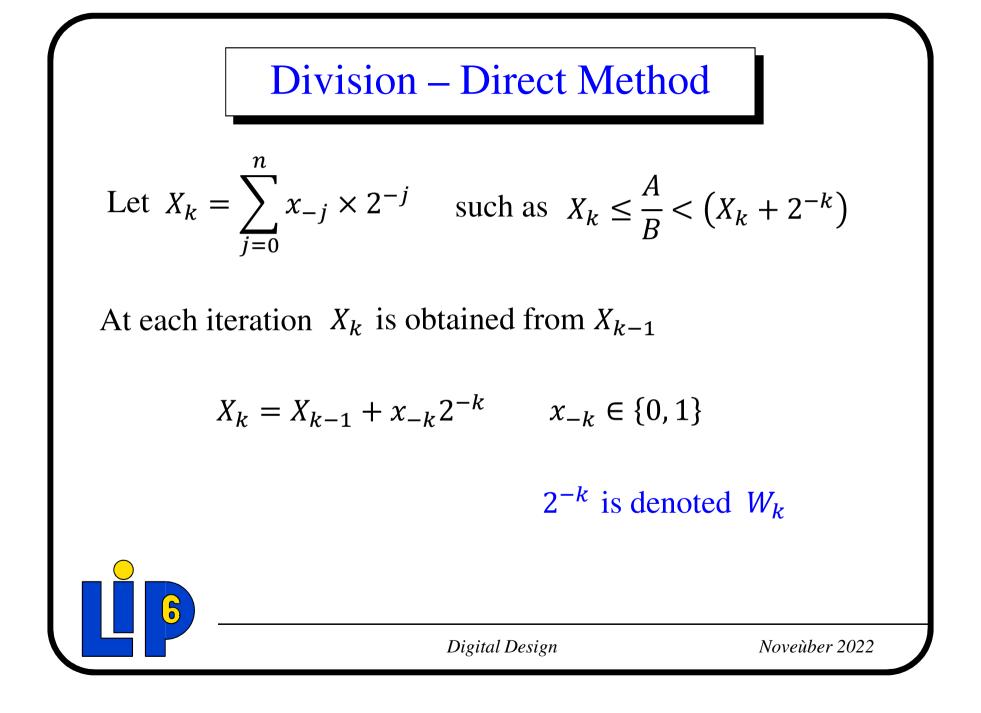
Given two positive real numbers  $A \in [1, 4[$  and  $B \in [1, 2[$ such as  $A \ge B$  we seek to calculate  $X_{Th} = \frac{A}{B}$  $X_{Th} \in [1, 2[$ 

Let  $X_n$  be an approximation of  $X_{Th}$  coded on n+1 bits  $X_n = \sum_{j=0}^n x_{-j} \times 2^{-j}$  such as  $X_n \le \frac{A}{B} < (X_n + 2^{-n})$ 

We propose to calculate  $X_n$  digit-by-digit

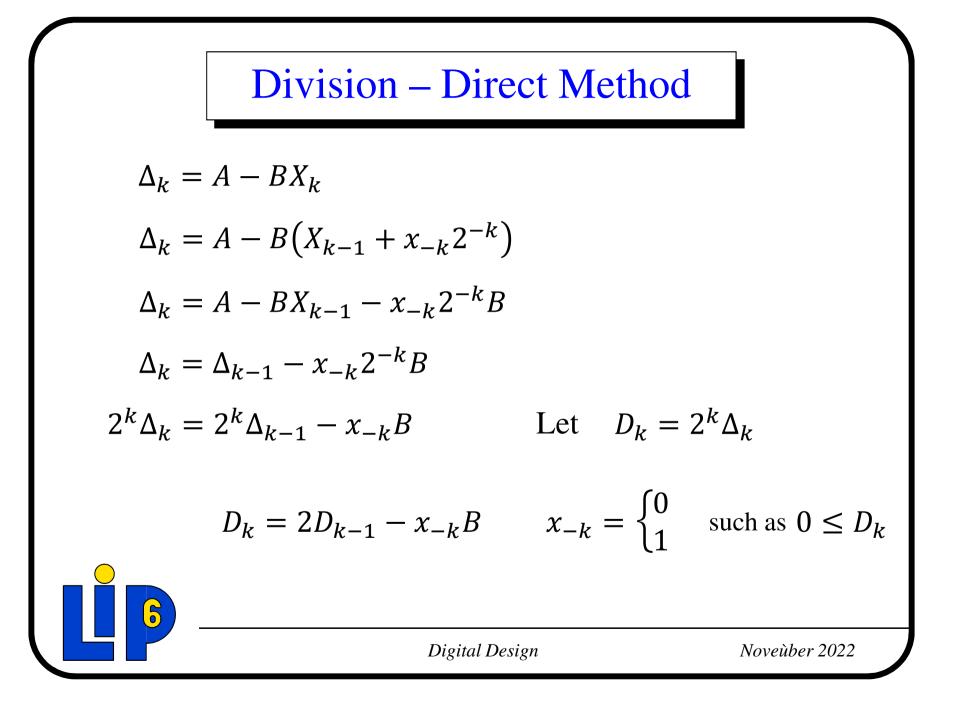
Digital Design

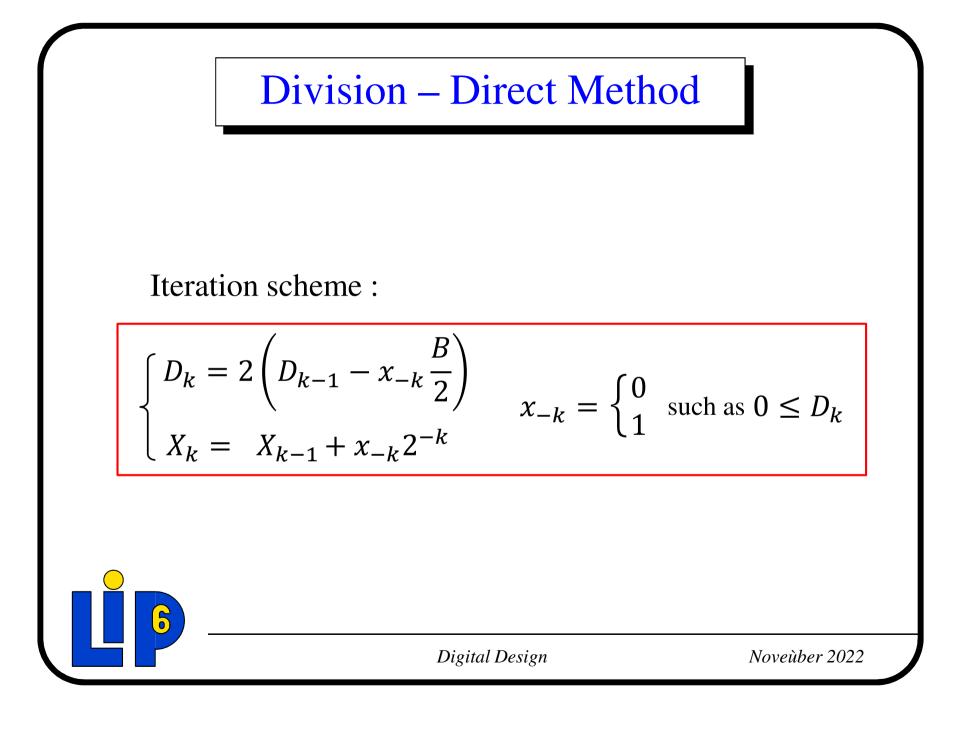


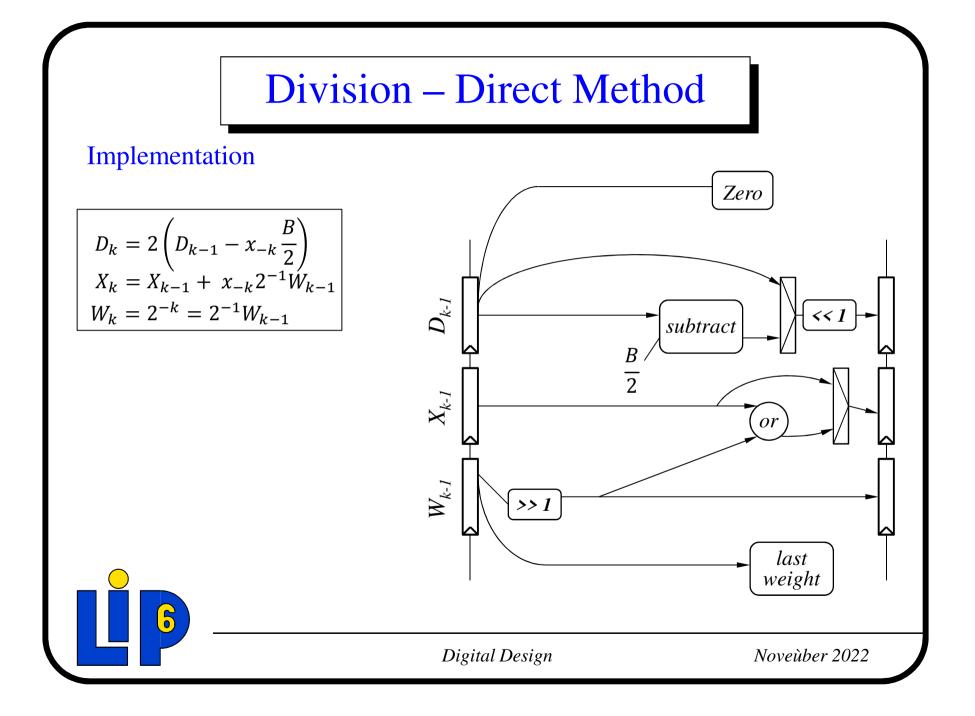


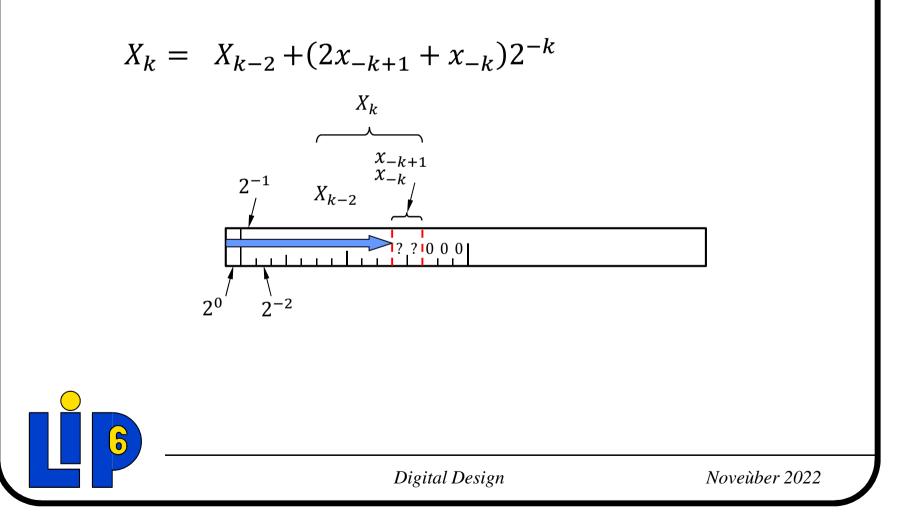
$$\begin{aligned} & \text{Division} - \text{Direct Method} \\ & X_k \leq \frac{A}{B} < (X_k + 2^{-k}) \\ & BX_k \leq A < B(X_k + 2^{-k}) \\ & 0 \leq A - BX_k < 2^{-k}B \\ & 0 \leq \Delta_k < 2^{-k}B \end{aligned} \qquad \text{Let } \Delta_k = A - BX_k \\ & 0 \leq \Delta_k < 2^{-k}B \\ & \text{yet } B < 2 \end{aligned}$$
then  $0 \leq \Delta_k < 2^{-k} \times 2$ 

$$& \stackrel{\circ}{\text{O}} \leq \Delta_0 < 2 \\ & \text{Othermal Action the upper bound of } \Delta_k \text{ is divided by } 2 \end{aligned}$$









$$\Delta_{k} = A - BX_{k}$$

$$\Delta_{k} = A - B(X_{k-2} + (x_{-k+1}2^{-k+1} + x_{-k}2^{-k}))$$

$$\Delta_{k} = \Delta_{k-2} - B(x_{-k+1}2^{-k+1} + x_{-k}2^{-k})$$

$$2^{k}\Delta_{k} = 2^{k}\Delta_{k-2} - B(2x_{-k+1} + x_{-k}) \qquad \text{Let} \quad D_{k} = 2^{k}\Delta_{k}$$

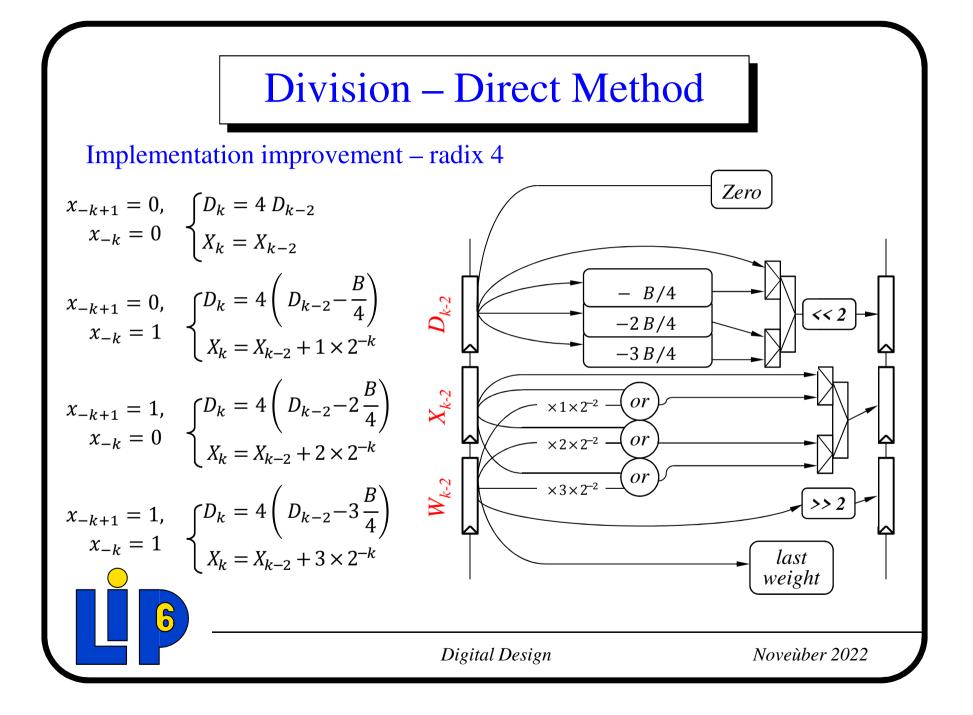
$$D_{k} = 4\left(D_{k-2} - \frac{B}{4}(2x_{-k+1} + x_{-k})\right)$$
Digital Design Noveibber 2022

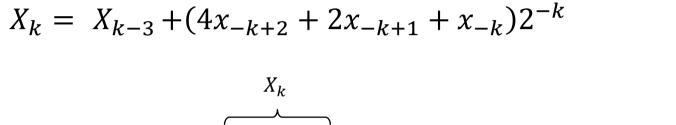
Implementation improvement – radix 4

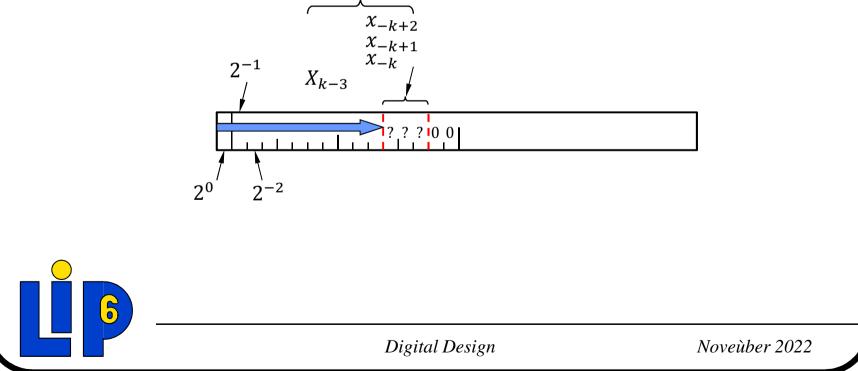
Iteration scheme :

$$\begin{cases} D_{k} = 4\left(D_{k-2} - \frac{B}{4}(2x_{-k+1} + x_{-k})\right) \\ X_{k} = X_{k-2} + (2x_{-k+1} + x_{-k})2^{-k} \\ x_{-k+1}, x_{-k} = \begin{cases} 00 \\ 01 \\ 10 \\ 11 \end{cases} \text{ such as } 0 \le D_{k} \end{cases}$$

Implementation improvement – radix 4  $x_{-k+1} = 0, \ x_{-k} = 0$   $\begin{cases}
D_k = 4 D_{k-2} \\
X_k = X_{k-2}
\end{cases}$  $x_{-k+1} = 0, \ x_{-k} = 1 \quad \begin{cases} D_k = 4\left(D_{k-2} - \frac{B}{4}\right) \\ X_k = X_{k-2} + 1 \times 2^{-k} \end{cases}$  $x_{-k+1} = 1, \ x_{-k} = 0$   $\begin{cases} D_k = 4\left(D_{k-2} - 2\frac{B}{4}\right) \\ X_k = X_{k-2} + 2 \times 2^{-k} \end{cases}$  $x_{-k+1} = 1, \ x_{-k} = 1 \quad \begin{cases} D_k = 4\left(D_{k-2} - 3\frac{B}{4}\right) \\ X_k = X_{k-2} + 3 \times 2^{-k} \end{cases}$ 6 Digital Design Noveùber 2022







$$\Delta_{k} = A - BX_{k}$$

$$\Delta_{k} = A - B(X_{k-3} + (x_{-k+2}2^{-k+2} + x_{-k+1}2^{-k+1} + x_{-k}2^{-k}))$$

$$\Delta_{k} = \Delta_{k-3} - B(x_{-k+2}2^{-k+2} + x_{-k+1}2^{-k+1} + x_{-k}2^{-k})$$

$$2^{k}\Delta_{k} = 2^{k}\Delta_{k-3} - B(4x_{-k+2} + 2x_{-k+1} + x_{-k}) \quad \text{Let} \quad D_{k} = 2^{k}\Delta_{k}$$

$$D_{k} = 8\left(D_{k-3} - \frac{B}{8}(4x_{-k+2} + 2x_{-k+1} + x_{-k})\right)$$
Digital Design Novewber 2022

Implementation improvement – radix 8

Iteration scheme :

$$\begin{cases} D_{k} = 8\left(D_{k-3} - \frac{B}{9}\left(4x_{-k+2} + 2x_{-k+1} + x_{-k}\right)\right) \\ X_{k} = X_{k-3} + \left(4x_{-k+2} + 2x_{-k+1} + x_{-k}\right)2^{-k} \\ x_{-k+2}, x_{-k+1}, x_{-k} = \begin{cases} 000\\001\\010\\011\\000\\011 \\010\\011 \end{cases} \text{ such as } 0 \le D_{k} \\ 001\\011\\010\\011 \end{cases}$$

