DIVISION BY DIGIT RECURRENCE

SEVERAL DIVISION METHODS:

- DIGIT-RECURRENCE METHOD – studied in this chapter
- MULTIPLICATIVE METHOD (Chapter 7)
- VARIOUS APPROXIMATION METHODS (power series expansion),
- SPECIAL METHODS SUCH AS CORDIC (Chapter 11) AND CONTINUED PRODUCT METHODS.

IMPLEMENTATIONS:

- SEQUENTIAL
- COMBINATIONAL
  1. PIPELINED
  2. NONPIPELINED
- COMBINATIONAL/SEQUENTIAL

DEFINITION AND NOTATION

$$x = q \cdot d + rem$$

$$|rem| < |d| \cdot ulp \quad \text{and} \quad sign(rem) = sign(x)$$

DIVIDEND x
DIVISOR d
QUOTIENT q
REMAINDER rem

- INTEGER QUOTIENT: $ulp = 1$,
- FRACTIONAL QUOTIENT: $ulp = r^{-n}$

TWO TYPES OF DIVISION OPERATION:

1. INTEGER DIVISION, WITH INTEGER OPERANDS AND RESULT
   USUALLY REQUIRES AN EXACT REMAINDER
2. FRACTIONAL DIVISION
   TO AVOID QUOTIENT OVERFLOW: $x < d$
   QUOTIENT ROUNDED, WHICH CAN RESULT IN A NEGATIVE REMAINDER

FRACTIONAL DIVISION - ASSUMPTIONS

- OPERANDS/RESULT IN SIGN-AND-MAGNITUDE FORMAT $\Longrightarrow$CONSIDER MAGNITUDES ONLY

$$1/2 \leq d < 1; \quad x < d; \quad 0 < q < 1$$

- FOR SIMPLER SELECTION: $q_j \in \mathcal{D}_a = \{-a, -a+1, \ldots, -1, 0, 1, \ldots, a-1, a\}$

- REDUNDANCY FACTOR

$$\rho = \frac{a}{r-1}, \quad \frac{1}{2} < \rho \leq 1$$

RECURRENCE

- QUOTIENT AFTER j STEPS: $q[j] = q[0] + \Sigma_{i=1}^{j} q_i r^{-i}$
- FINAL QUOTIENT: $q = q[n] = q[0] + \Sigma_{i=1}^{n} q_i r^{-i}$
- FINAL QUOTIENT ERROR BOUND: $0 \leq \epsilon_q = \frac{x}{d} - q \; < \; r^{-n}$
- ERROR AT STEP j:

$$\epsilon[j] = \frac{x}{d} - q[j] \leq \epsilon[n] + \sum_{i=j+1}^{n} max(q_i) r^{-i} = \epsilon[n] + \frac{a}{r-1}(r^{-j} - r^{-n})$$

- FOR CONVERGENCE: $\epsilon[j] \leq \rho r^{-j}$
- FOR SIMPLER SELECTION, ALLOW NEGATIVE ERRORS

$$|\epsilon[j]| = |\frac{x}{d} - q[j]| \leq \rho r^{-j}$$

- ELIMINATE DIVISION FROM ERROR BOUND: $|x - dq[j]| \leq \rho d r^{-j}$

- DEFINE RESIDUAL (PARTIAL REMAINDER): $w[j] = r^j(x - dq[j])$
- RESIDUAL RECURRENCE: $w[j+1] = rw[j] - dq_{j+1} \qquad w[0] = x$

cont.

- BOUND ON $w[j]$: $|w[j]| \leq \rho d$

- SELECT QUOTIENT DIGIT TO KEEP $w[j+1]$ BOUNDED:

$$q_{j+1} = SEL(rw[j],\ d)$$

- REDUNDANCY IN QUOTIENT-DIGIT SET ALLOWS

$$q_{j+1} = SEL(\widehat{y},\ \widehat{d})$$

WHERE $\widehat{y}$ IS TRUNCATED $rw[j]$, etc.

IMPLEMENTATION OF DIVISION STEP

1. ONE DIGIT ARITHMETIC LEFT-SHIFT OF $w[j]$ TO PRODUCE $rw[j]$;

2. DETERMINATION OF THE QUOTIENT DIGIT q_{j+1} BY THE QUOTIENT-DIGIT SELECTION FUNCTION;

3. GENERATION OF THE DIVISOR MULTIPLE $d \times q_{j+1}$; and

4. SUBTRACTION OF dq_{j+1} from $rw[j]$.

5. UPDATE OF THE QUOTIENT $q[j]$ TO $q[j+1]$ BY THE ON-THE-FLY CONVERSION

$$q[j+1] = CONV(q[j], q_{j+1})$$

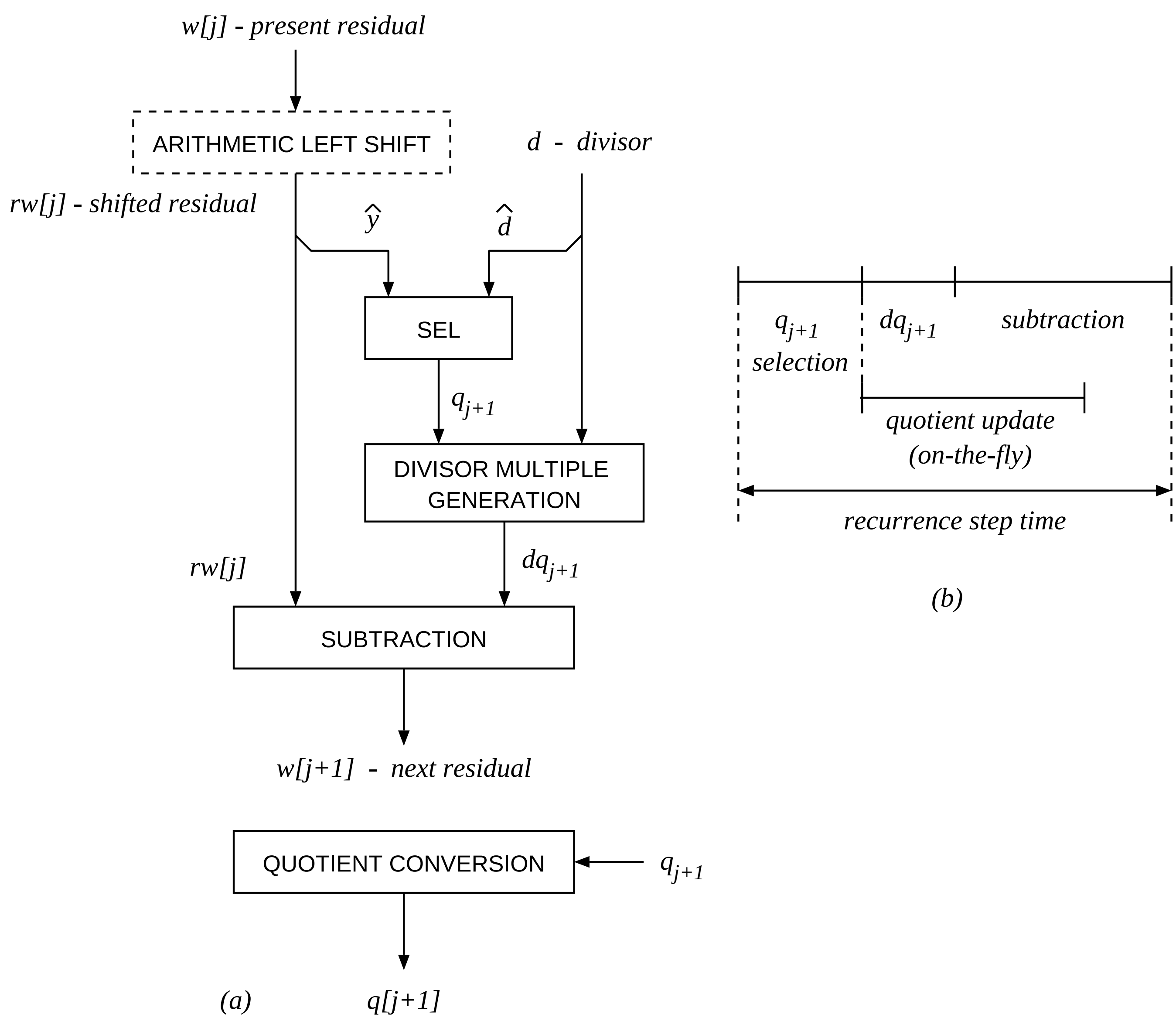


Figure 5.1: (a) COMPONENTS OF A DIVISION STEP. (b) TIMING.

MAIN FACTORS AFFECTING COST AND EXECUTION TIME

- RADIX r
- QUOTIENT-DIGIT SET
  1. CANONICAL: $0 \leq q_j \leq r-1$
  2. REDUNDANT: $q_j \in \mathcal{D}_a = \{-a, -a+1, \ldots, -1, 0, 1, \ldots, a-1, a\}$
- REDUNDANCY FACTOR: $\rho = \frac{a}{r-1}, \quad \rho > \frac{1}{2}$
- REPRESENTATION OF RESIDUAL:
  1. NONREDUNDANT (e.g., 2's complement)
  2. REDUNDANT: carry-save, signed-digit
- QUOTIENT-DIGIT SELECTION FUNCTION

INITIALIZATION AND TERMINATION

Initialization:

- $w[0] = x - dq[0]$ and $|w[0]| \leq \rho d$. Options:
  - Make $q[0] = 0$ and
    - For $\rho = 1$ we make $w[0] = x/2$.
    - For $1/2 < \rho < 1$ we make $w[0] = x/4$

    Compensated in the termination step
  - Make $q[0] = 1$ and $w[0] = x - d$. Applicable for $\rho < 1$ because $q > 1 + \rho$ not allowed.

Termination:

- QUOTIENT:

$$q = \begin{cases} q[N] & \text{if } w[N] \geq 0 \\ q[N] - r^{-N} & \text{if } w[N] < 0 \end{cases}$$

N – number of iterations
If dividend shifted in initialization - shift quotient (extra iteration)

ON-THE-FLY QUOTIENT CONVERSION

- j MS DIGITS OF CONVERTED QUOTIENT

$$Q[j] = \sum_{i=1}^{j} q_i r^{-i}$$

- UPDATE

$$Q[j+1] = Q[j] + q_{j+1} r^{-(j+1)}$$

- SINCE q_{j+1} CAN BE NEGATIVE:

$$Q[j+1] = \begin{cases} Q[j] + q_{j+1} r^{-(j+1)} & \text{if } q_{j+1} \geq 0 \\ Q[j] - r^{-j} + (r - |q_{j+1}|) r^{-(j+1)} & \text{if } q_{j+1} < 0 \end{cases}$$

- DISADVANTAGE: SUBTRACTION $Q[j] - r^{-j}$
  REQUIRES THE PROPAGATION OF A BORROW – SLOW

ON-THE-FLY CONVERSION (cont.)

- DEFINE ANOTHER FORM $QM[j]$

$$QM[j] = Q[j] - r^{-j}$$

- NEW CONVERSION ALGORITHM IS

$$Q[j+1] = \begin{cases} Q[j] + q_{j+1} r^{-(j+1)} & \text{if } q_{j+1} \geq 0 \\ QM[j] + (r - |q_{j+1}|) r^{-(j+1)} & \text{if } q_{j+1} < 0 \end{cases}$$

- SUBTRACTION REPLACED BY LOADING THE FORM $QM[j]$
- UPDATE FORM $QM[j]$ AS FOLLOWS:

$$\begin{aligned} QM[j+1] &= Q[j+1] - r^{-(j+1)} \\ &= \begin{cases} Q[j] + (q_{j+1} - 1) r^{-(j+1)} & \text{if } q_{j+1} > 0 \\ QM[j] + ((r-1) - |q_{j+1}|) r^{-(j+1)} & \text{if } q_{j+1} \leq 0 \end{cases} \end{aligned}$$

- ALL ADDITIONS ARE CONCATENATIONS

ON-THE-FLY CONVERSION ALGORITHM

$$Q[j+1] = \begin{cases} (Q[j], q_{j+1}) & \text{if } q_{j+1} \geq 0 \\ (QM[j], (r - |q_{j+1}|)) & \text{if } q_{j+1} < 0 \end{cases}$$

$$QM[j+1] = \begin{cases} (Q[j], q_{j+1} - 1) & \text{if } q_{j+1} > 0 \\ (QM[j], ((r-1) - |q_{j+1}|)) & \text{if } q_{j+1} \leq 0 \end{cases}$$

INITIAL CONDITIONS $Q[0] = QM[0] = 0$ (for a positive quotient)

EXAMPLE OF RADIX-2 CONVERSION

j	q_j	$Q[j]$	$QM[j]$
0		0	0
1	1	0.1	0.0
2	1	0.11	0.10
3	0	0.110	0.101
4	1	0.1101	0.1100
5	-1	0.11001	0.11000
6	0	0.110010	0.110001
7	0	0.1100100	0.1100011
8	-1	0.11000111	0.11000110
9	1	0.110001111	0.110001110
10	0	0.1100011110	0.1100011101
11	1	0.11000111101	0.11000111100
12	0	0.110001111010	0.110001111001

IMPLEMENTATION OF THE CONVERSION

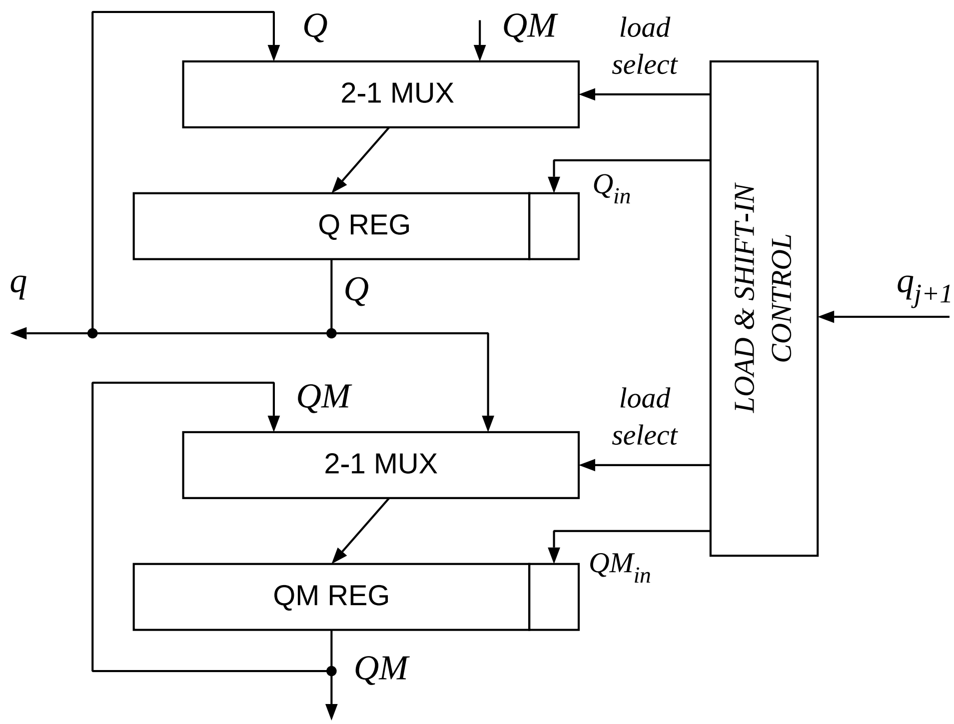


Figure 5.2: Implementation of on-the-fly conversion.

$$Q \leftarrow \begin{cases} shift\ Q\ with\ insert\ (Q_{in}) & \text{if } C_{shiftQ} = 1 \\ shift\ QM\ with\ insert\ (Q_{in}) & \text{if } C_{loadQ} = 1 \end{cases}$$

$$QM \leftarrow \begin{cases} shift\ QM\ with\ insert\ (QM_{in}) & \text{if } C_{shiftQM} = 1 \\ shift\ Q\ with\ insert\ (QM_{in}) & \text{if } C_{loadQM} = 1 \end{cases}$$

cont.

$$Q_{in} = \begin{cases} q_{j+1} & \text{if } q_{j+1} \geq 0 \\ r - |q_{j+1}| & \text{if } q_{j+1} < 0 \end{cases}$$

$$QM_{in} = \begin{cases} q_{j+1} - 1 & \text{if } q_{j+1} > 0 \\ (r-1) - |q_{j+1}| & \text{if } q_{j+1} \leq 0 \end{cases}$$

REGISTER CONTROL SIGNALS: $C_{loadQ} = C'_{shiftQ}$ and $C_{loadQM} = C'_{shiftQM}$

EXAMPLE OF RADIX-4 CONVERSION

q_{j+1}	Q_{in}	C_{shiftQ}	$Q[j+1]$	QM_{in}	$C_{shiftQM}$	$QM[j+1]$
3	3	1	$(Q[j],3)$	2	0	$(Q[j],2)$
2	2	1	$(Q[j],2)$	1	0	$(Q[j],1)$
1	1	1	$(Q[j],1)$	0	0	$(Q[j],0)$
0	0	1	$(Q[j],0)$	3	1	$(QM[j],3)$
-1	3	0	$(QM[j],3)$	2	1	$(QM[j],2)$
-2	2	0	$(QM[j],2)$	1	1	$(QM[j],1)$
-3	1	0	$(QM[j],1)$	0	1	$(QM[j],0)$

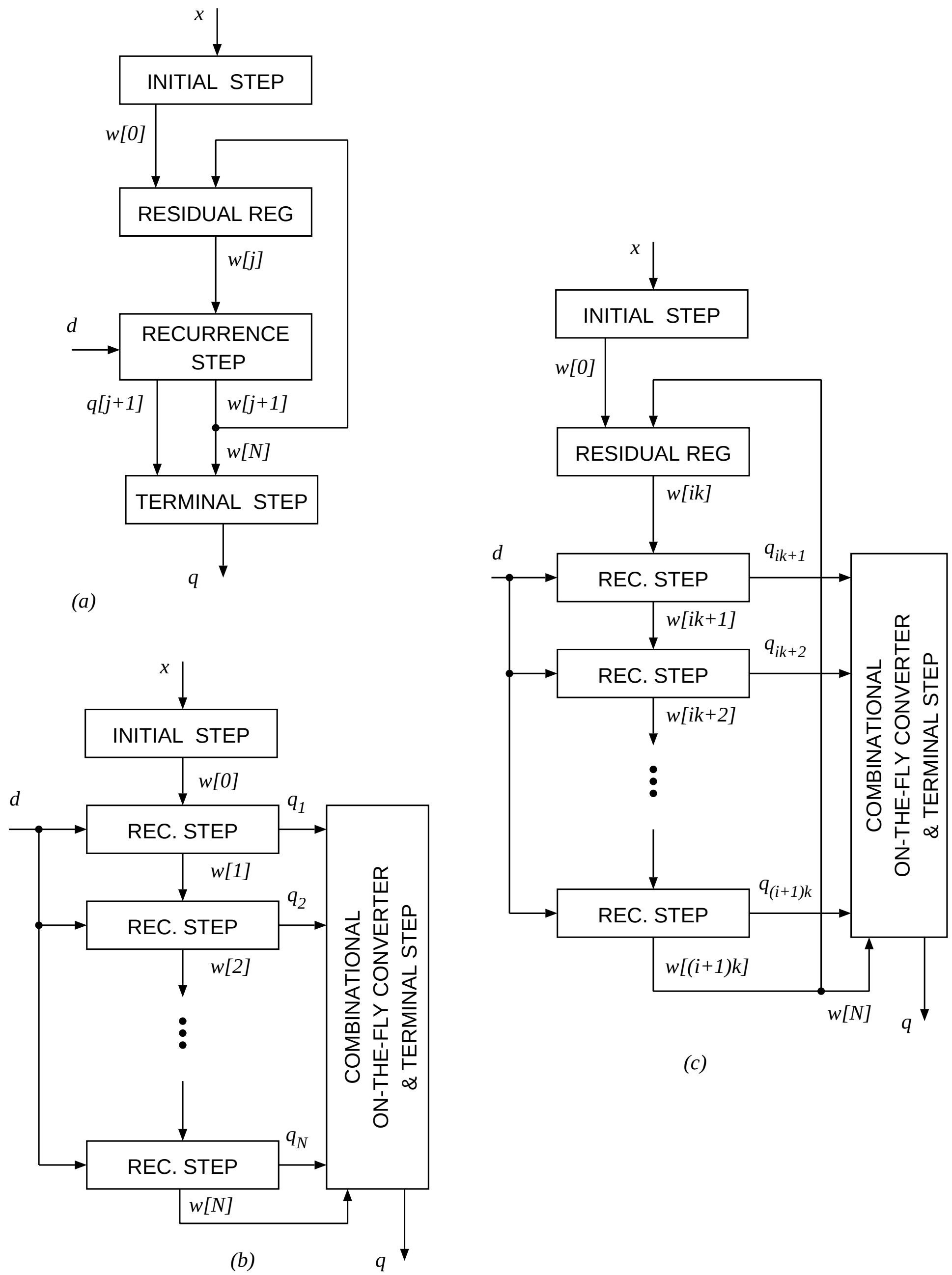


Figure 5.3: DIVISION IMPLEMENTATION: (a) TOTALLY SEQUENTIAL. (b) TOTALLY COMBINATIONAL. (c) COMBINED IMPLEMENTATION.

EXAMPLES OF ALGORITHMS AND IMPLEMENTATIONS

- CELLS: delay function of the load; delay and area in terms of 2-NAND.
- DEGREE OF OPTIMIZATION: the same modules have been used in all designs.
- INTERCONNECTIONS NOT INCLUDED: not considered the delay, area nor load of interconnections.
- EXECUTION TIME AND THE AREA FOR 53-BIT OPERANDS AND RESULT
- INCLUDED DELAY AND AREA OF REGISTERS

SCHEMES COMPARED

r2 Scheme Radix-2 with carry-save residual.

r4 Scheme Radix-4 with $a = 2$ and carry-save residual.

r8 Scheme Radix-8 with $a = 7$ and carry-save residual.

r16over Scheme Radix-16 with two overlapped radix-4 stages.

r512 Scheme Radix-512 with $a = 511$, carry-save residual, scaling, and quotient-digit selection by rounding.

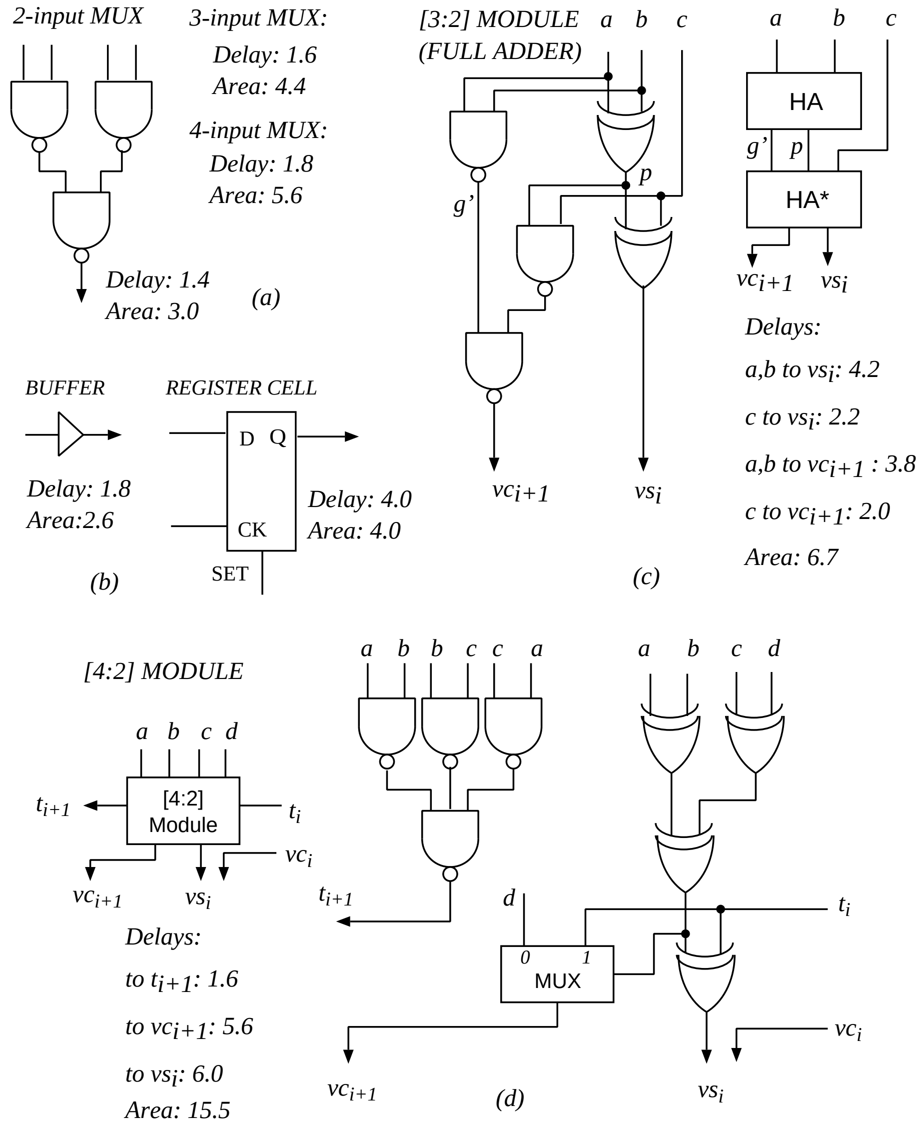


Figure 5.4: Basic modules: (a) Multiplexers. (b) Buffer and register cell. (c) Full-adder. (d) [4:2] module.

RADIX-2 DIVISION WITH CS RESIDUAL

- REDUNDANT RESIDUAL $w[j] = (WC[j],\ WS[j])$

1. $[Initialize]$
   $WS[0] \leftarrow x/2;\ WC[0] \leftarrow 0;\ q_0 \leftarrow 0;$Q[-1]=0
2. $[Recurrence]$
   for $j = 0 \ldots n+1$ ($n+2$ iterations because of initialization and guard bit)
   $q_{j+1} \leftarrow SEL(\widehat{y});$
   $(WC[j+1],\ WS[j+1]) \leftarrow CSADD(2WC[j],\ 2WS[j],\ -q_{j+1}d);$
   $Q[j] \leftarrow CONVERT(Q[j-1],\ q_j)$
   end for
3. $[Terminate]$
   If $w[n+2] < 0$ then $q = 2(CONVERT(Q[n+1],\ q_{n+2} - 1))$
   else $q = 2(CONVERT(Q[n+1],\ q_{n+2}))$

RADIX-2 ALGORITHM (cont.)

- n is the precision in bits,
- SEL is the quotient-digit selection function:

$$q_{j+1} = SEL(\widehat{y}) = \begin{cases} 1 & \text{if } 0 \leq \widehat{y} \leq 3/2 \\ 0 & \text{if } \widehat{y} = -1/2 \\ -1 & \text{if } -5/2 \leq \widehat{y} \leq -1 \end{cases}$$

  The estimate $\widehat{y}$ has four bits (three integer bits and one fractional bit) of the shifted residual in carry-save form,

- $CSADD$ is carry-save addition
- $-q_{j+1}d$ is in 2's complement form, and
- $CONVERT$ on-the-fly quotient conversion function

Dividend $x = (0.10011111)$, divisor $d = (0.11000101)$, scaled residual $2w[0] = 2(x/2) = x$, $q_{computed} = q/2$

$$
\begin{array}{rlll}
2WS[0] = & 000.10011111 & & \\
2WC[0] = & 000.00000001 \ * & \widehat{y}[0] = 0.5 & q_1 = 1 \\
-q_1 d = & \ \ 11.00111010 & & \\
\hline
2WS[1] = & 111.01001000 & & \\
2WC[1] = & 000.01101100 & \widehat{y}[0] = -1 & q_2 = -1 \\
-q_2 d = & \ \ 00.11000101 & & \\
\hline
2WS[2] = & 111.11000010 & & \\
2WC[2] = & 001.00110001 \ * & \widehat{y}[1] = 0.5 & q_3 = 1 \\
-q_3 d = & \ \ 11.00111010 & & \\
\hline
2WS[3] = & 011.10010010 & & \\
2WC[3] = & 100.11001001 \ * & \widehat{y}[2] = 0 & q_4 = 1 \\
-q_4 d = & \ \ 11.00111010 & & \\
\hline
2WS[4] = & 000.11000010 & & \\
2WC[4] = & 110.01101000 & \widehat{y}[4] = -1.5 & q_5 = -1 \\
-q_5 d = & \ \ 00.11000101 & & \\
\hline
\end{array}
$$

* for 2's complement of $q_{j+1}d$

$$q = 2(.1\bar{1}11\bar{1}) = .1101$$

Figure 5.7: Example of radix-2 division with residual in carry-save form.

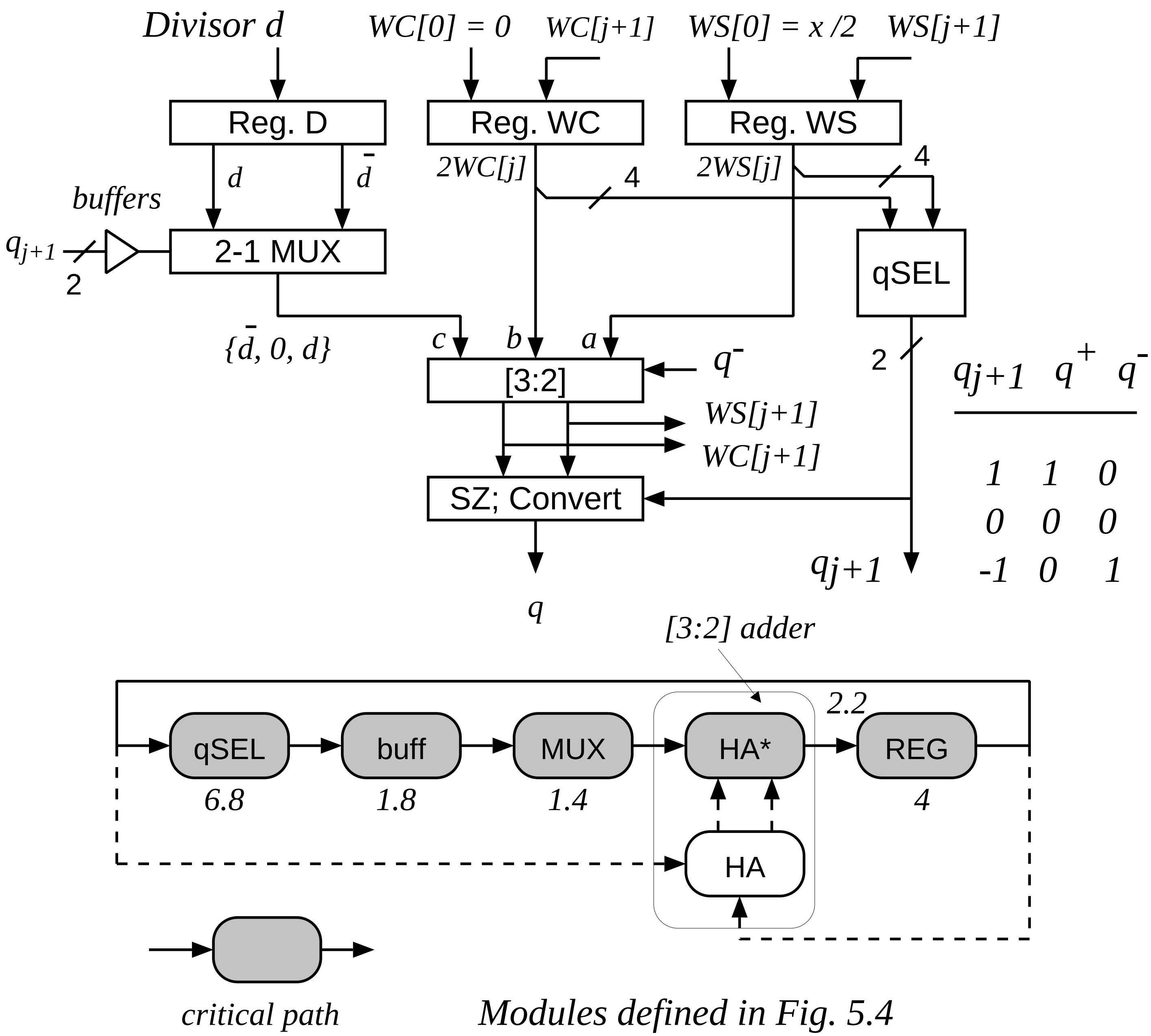


Figure 5.8: IMPLEMENTATION OF RADIX-2 SCHEME.

DELAY AND AREA OF RADIX-2 STAGE

element	delay	area
q-digit selection	6.8	50
buffers	1.8	5
MUX	1.4	160
CSA	2.2	360
registers (3)	4.0	650
Convert & Round	(NC)	1360
Cycle time	16.2	
Total area		2585

NC denotes a delay not in the critical path

RADIX-4 DIVISION WITH CS RESIDUAL

1. QUOTIENT DIGIT SET {-2,-1,0,1,2}
2. $\rho < 1$ initialize $WS[0] \leftarrow x/4$
3. THE NEXT RESIDUAL
   $(WC[j+1],\ WS[j+1]) \leftarrow CSADD(4WC[j],\ 4WS[j],\ -q_{j+1}d)$
4. QUOTIENT-DIGIT SELECTION DEPENDS ON ESTIMATES OF SHIFTED RESIDUAL AND DIVISOR described in terms of SELECTION CONSTANTS $m_k(i)$

$$q_{j+1} = k \text{ if } m_k(i) \leq \hat{y} < m_{k+1}(i)$$

5. FINAL QUOTIENT = 4 x OBTAINED QUOTIENT

SELECTION CONSTANTS

- $q_{j+1} = k$ if $m_k(i) \leq \hat{y} < m_{k+1}(i)$
- $i = 16\hat{d}$ and $\hat{d}$ divisor truncated to the 4th fractional bit and
- $\hat{y}$ is $4w[j]$ truncated to the 4th fractional bit.

i	8	9	10	11	12	13	14	15
$m_2(i)^+$	**12**	**14**	**15**	**16**	**18**	**20**	**20**	**24**
$m_1(i)^+$	**4**	**4**	**4**	**4**	**6**	**6**	**8**	**8**
$m_0(i)^+$	**-4**	**-6**	**-6**	**-6**	**-8**	**-8**	**-8**	**-8**
$m_{-1}(i)^+$	**-13**	**-15**	**-16**	**-18**	**-20**	**-20**	**-22**	**-24**

$+$: real value = shown value/16

Dividend $x = (0.10101111)$, divisor $d = (0.11000101)$ $(i = 16(0.1100)_2 = 12)$
scaled residual $4w[0] = 4(x/4) = x$, $q_{computed} = q/4$

$$
\begin{array}{rlll}
4WS[0]^{+} = & 000.10101111 & & \\
4WC[0]^{+} = & 000.00000001 \;\; * & \widehat{y}[0] = 10/16 & q_1 = 1 \\
-q_1d^{+} = & \;\;11.00111010 & & \\
\hline
WS[1] = & \quad 1.10010100 & & \\
WC[1] = & \quad 0.01010110 & & \\
\hline
4WS[1]^{+} = & 110.01010000 & & \\
4WC[1]^{+} = & 001.01011000 & \widehat{y}[1] = -6/16 & q_2 = 0 \\
-q_2d^{+} = & \;\;00.00000000 & & \\
\hline
WS[2] = & \quad 1.00001000 & & \\
WC[2] = & \quad 0.10100000 & & \\
\hline
4WS[2]^{+} = & 100.00100000 & & \\
4WC[2]^{+} = & 010.10000000 & \widehat{y}[2] = -22/16 & q_3 = -2 \\
-q_3d^{+} = & \;\;01.10001010 & & \\
\hline
w[3] = & \quad 0.00101010 & &
\end{array}
$$

* least-significant 1 for 2's complement of $q_{j+1}d$
+ only one integer bit used in the recurrence, because of the range of $w[j+1]$.

$$q[3] = .10\bar{2}_4 = .032_4$$

Figure 5.9: Example of radix-4 division with residual in carry-save form.(On-the-fly conversion and termination not shown)

DELAY AND AREA OF RADIX-4 STAGE

element	delay	area
q-digit selection	10.8	160
buffers	1.8	10
MUX	1.8	300
CSA	2.2	360
registers (3)	4.0	650
Convert & Round	(NC)	1360
Cycle time	20.6	
Total area		2840

NC denotes a delay not in the critical path

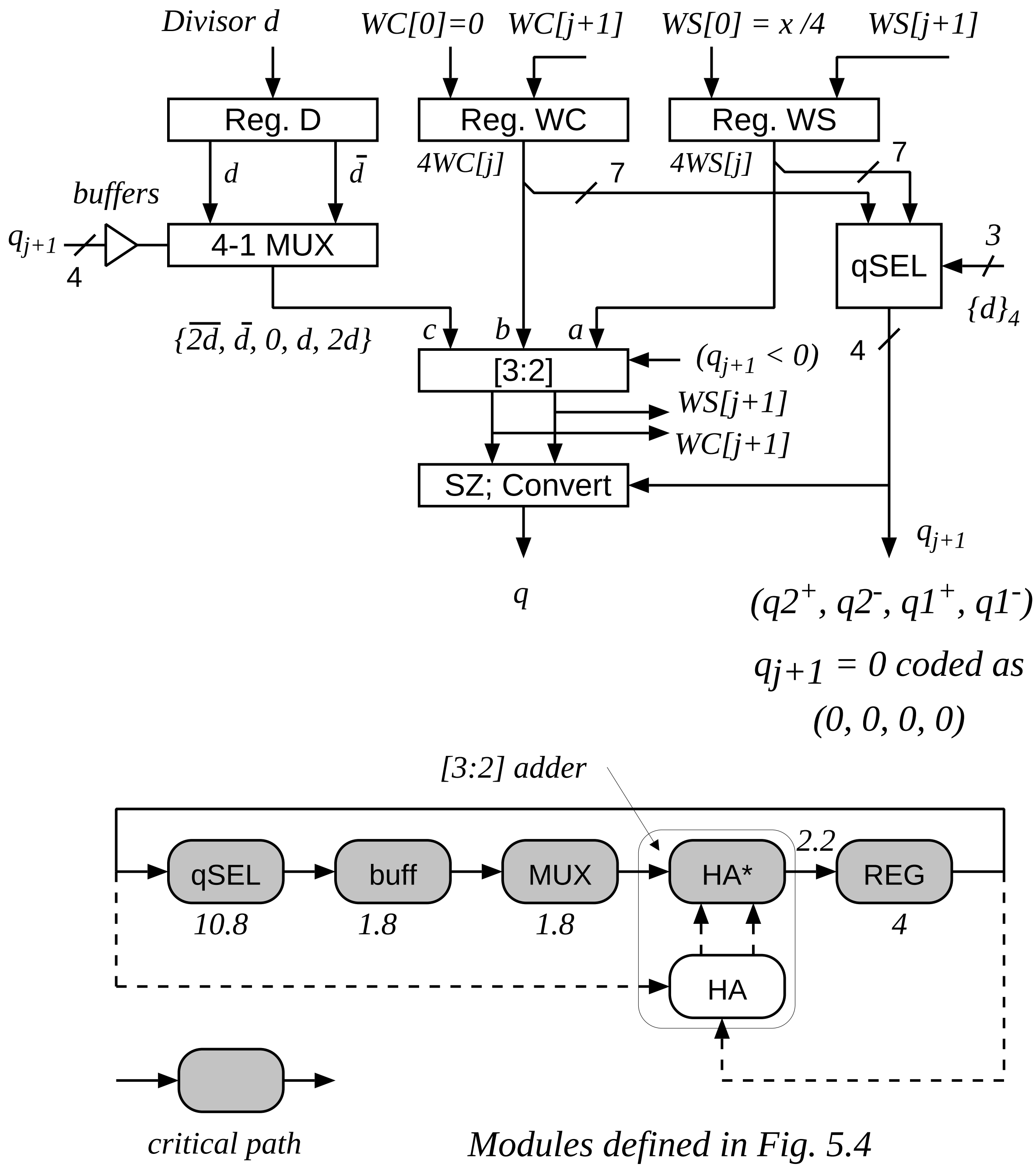


Figure 5.10: IMPLEMENTATION OF RADIX-4 SCHEME.

RADIX-8 DIVISION WITH CS RESIDUAL

- QUOTIENT DIGIT SET {-7, ..., 7} DECOMPOSED INTO

$$q_H = \{-8, -4, 0, 4, 8\}$$

AND

$$q_L = \{-2, -1, 0, 1, 2\}$$

element	delay	area
q-digit selection	(q_h) 12.2	610
buffers	1.8	20
MUXes	1.8	600
CSAh	2.2	360
CSAl	4.2	360
registers (3)	4.0	650
Convert & Round	(NC)	1360
Cycle time	26.2	
Total area		3960

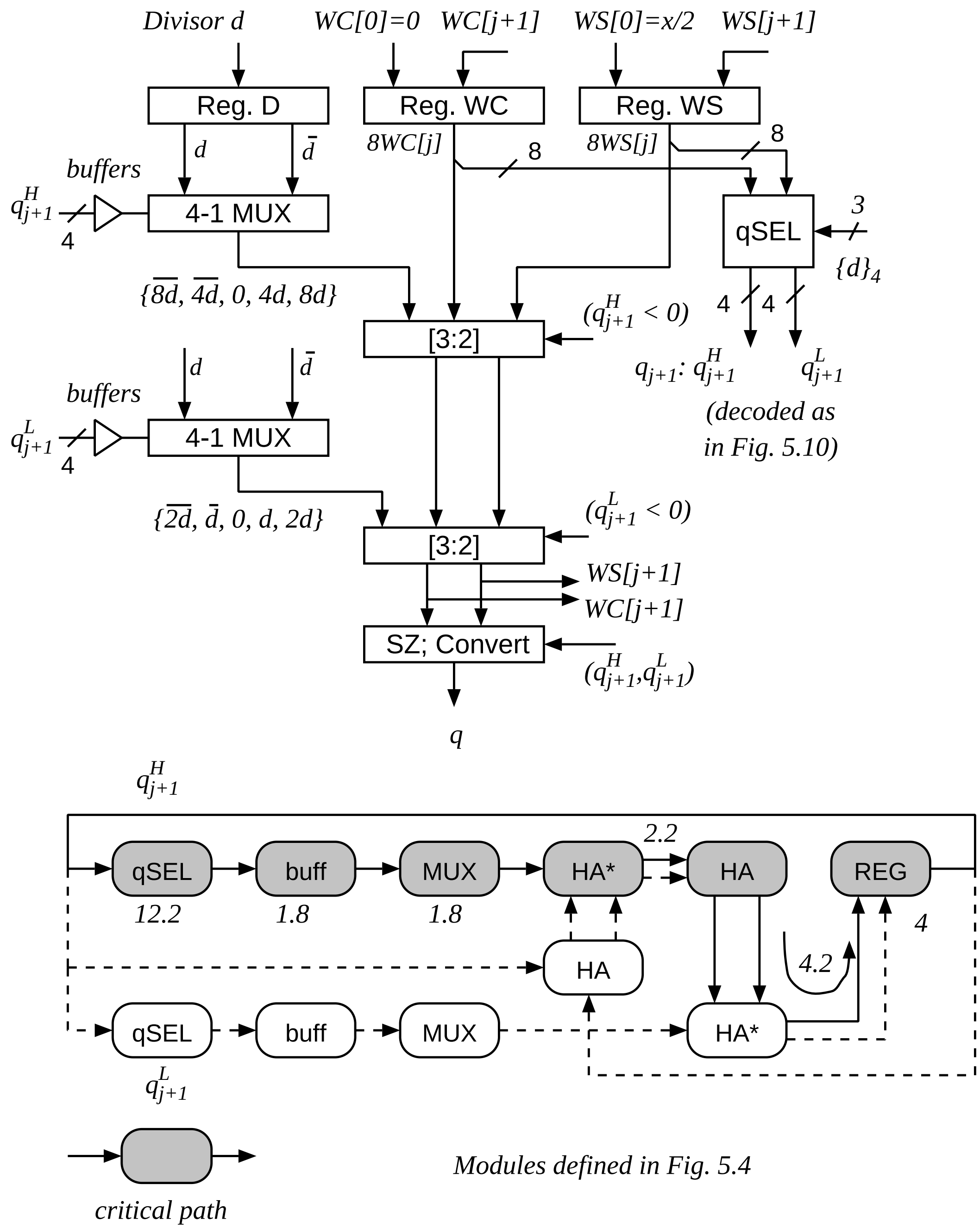


Figure 5.11: Implementation of radix-8 scheme.

RADIX-16 DIVISION WITH TWO RADIX-4 OVERLAPPED STAGES

element	delay	area
CSA	4.2	220
q-digit selection	11.2	820
MUX	1.4	
buffers	1.8	20
MUXes	1.8	600
CSA1	(NC)	360
CSA2	2.1	360
registers (3)	4.0	650
Convert & Round	(NC)	1360
Cycle time	26.6	
Total area		4390

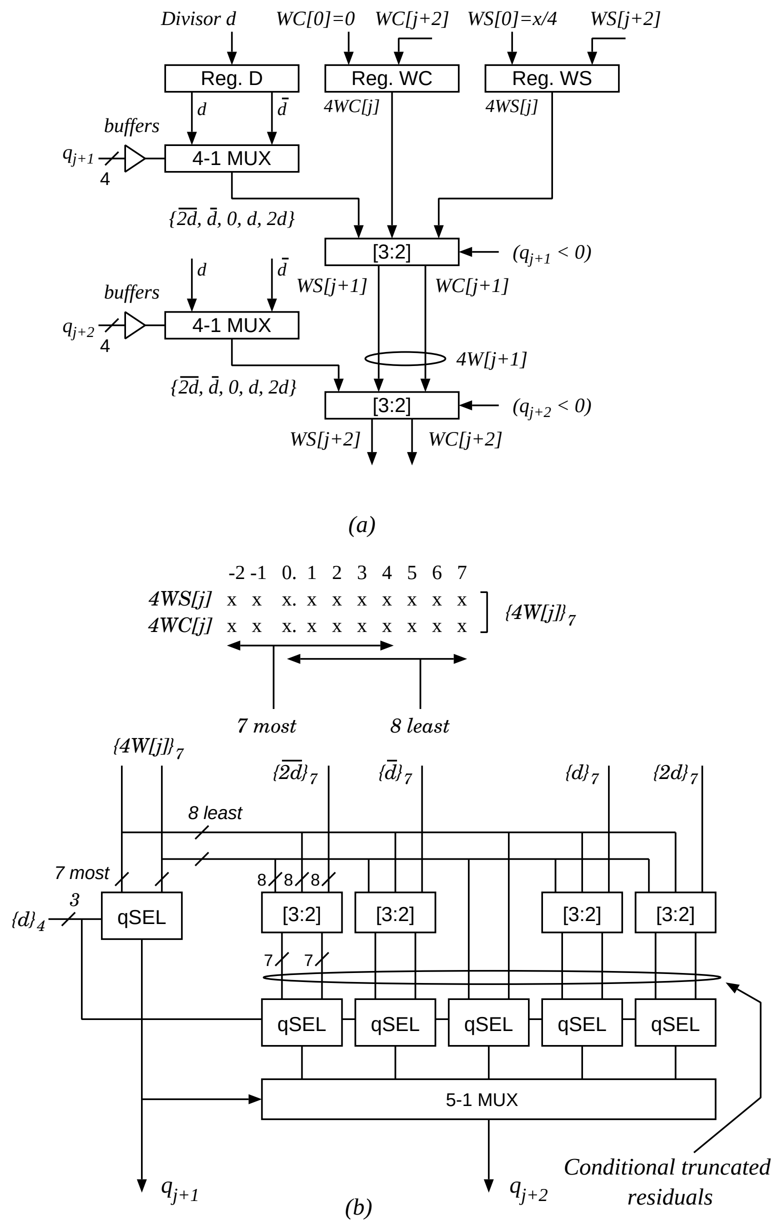


Figure 5.12: Implementation of radix-16 with radix-4 stages.

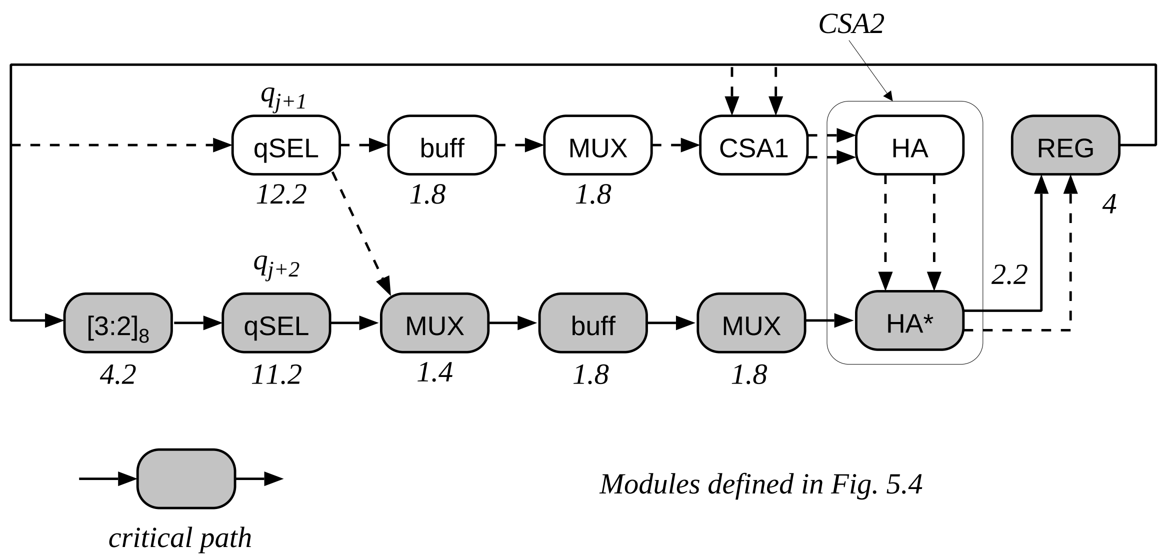


Figure 5.13: Critical path in radix-16 scheme.

RADIX-512 DIVISION WITH PRESCALING AND SELECTION BY ROUNDING

Cycle 1 : COMPUTE $M \approx 1/d$; compare d and x and set g

Cycle 2 : COMPUTE $z = Md$ (in c-s form); $v = 2^{-g}x$

Cycle 3 : INITIALIZE $w[0] = Mv$ (in c-s form); ASSIMILATE z;

Cycles 4 to 9 : ITERATE

$$q_{j+1} = round(\hat{y});\ \ w[j+1] = 512w[j] - q_{j+1}z$$

Cycle 10 : CORRECTING AND ROUNDING.

RADIX-512 DIVISON (cont.)

- RECTANGULAR MULTIPLIER-ACCUMULATOR
- DELAY-AREA

element	delay	area
M-module	(NC)	1800
MUX	1.4	
recoder	6.0	70
buffer	1.8	
MUX	1.8	3000
2 levels of 4-2 CSA	12.0	3100
registers(3)	4.0	650
Convert & Round	(NC)	1360
Cycle time	27	
Total area		9980

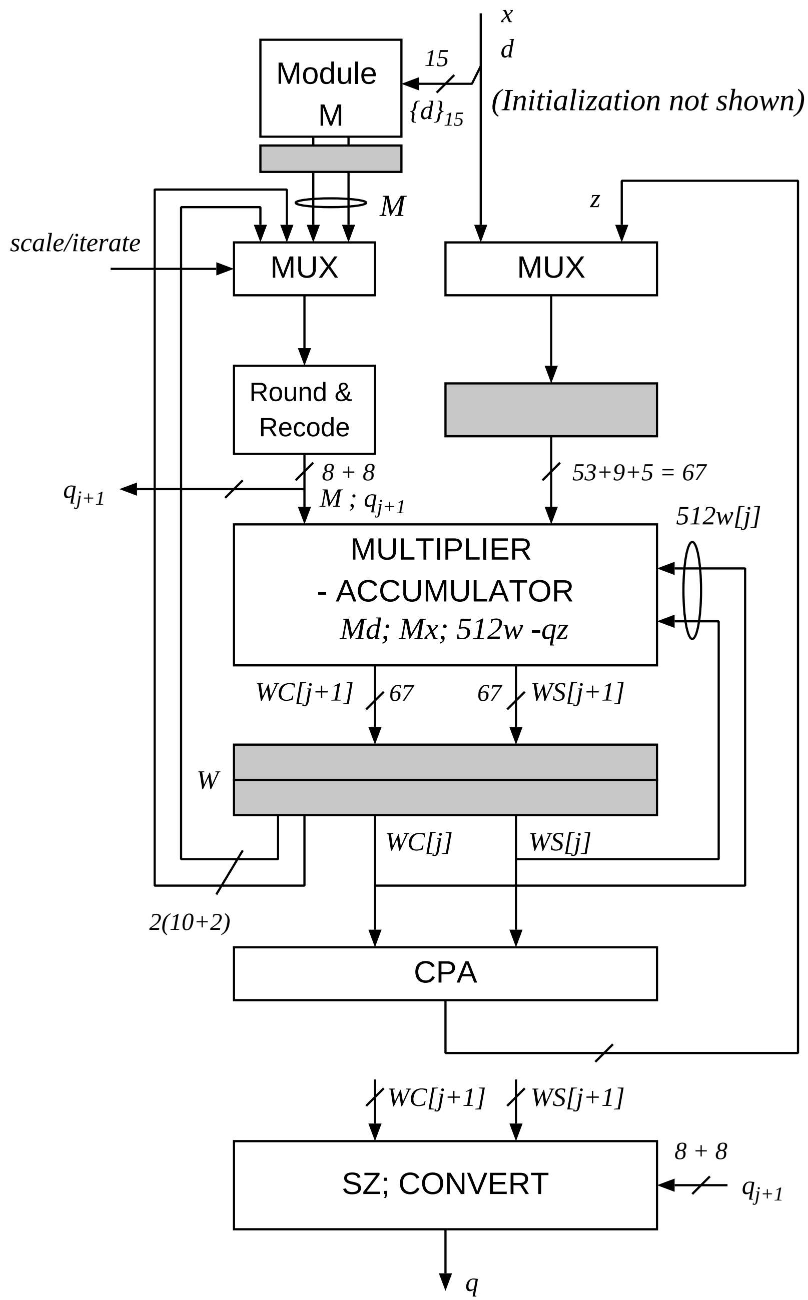


Figure 5.14: Implementation of radix-512 scheme.

OVERALL COMPARISONS

Scheme	r2	r4	r8	r16 (overlapped)	r512
Cycle-time factor	1.0	1.3	1.6	1.6	1.7
Number of cycles†	57	29	20	15	10
Speedup	1.0	1.5	1.8	2.4	3.4
Area factor	1.0	1.1	1.5	1.7	3.9

†Correction: two cycles for radix-2, one cycle for other cases.

QUOTIENT-DIGIT SELECTION FUNCTION

Quotient-digit set:

$$q_{j+1} \in \mathcal{D}_a = \{-a, -a+1, \ldots, -1, 0, 1, \ldots, a-1, a\}$$

Redundancy factor:

$$\rho = \frac{a}{r-1}, \quad \frac{1}{2} < \rho \leq 1$$

- TWO FUNDAMENTAL CONDITIONS FOR q-SELECTION
- CONTAINMENT – must guarantee bounded residual
- CONTINUITY – there must exist a valid choice of q_{j+1} in the range of shifted residual

CONTAINMENT AND SELECTION INTERVALS

- RESIDUAL RECURRENCE

$$w[j+1] = rw[j] - dq_{j+1} \quad |w[j]| \leq \rho d$$

$$\rho = a/(r-1) \quad -a \leq q_j \leq a$$

- SELECTION INTERVALS
- If $rw[j] \in [L_k, U_k]$ then $q_{j+1} = k$ makes $w[j+1]$ bounded

$$L_k \leq rw[j] \leq U_k \;\Rightarrow\; \rho d \leq w[j+1] = rw[j] - k \cdot d \leq \rho d$$

- EXPRESSIONS FOR SELECTION INTERVALS

$$U_k = (k+\rho)d \qquad L_k = (k-\rho)d$$

ROBERTSON'S AND P-D DIAGRAMS

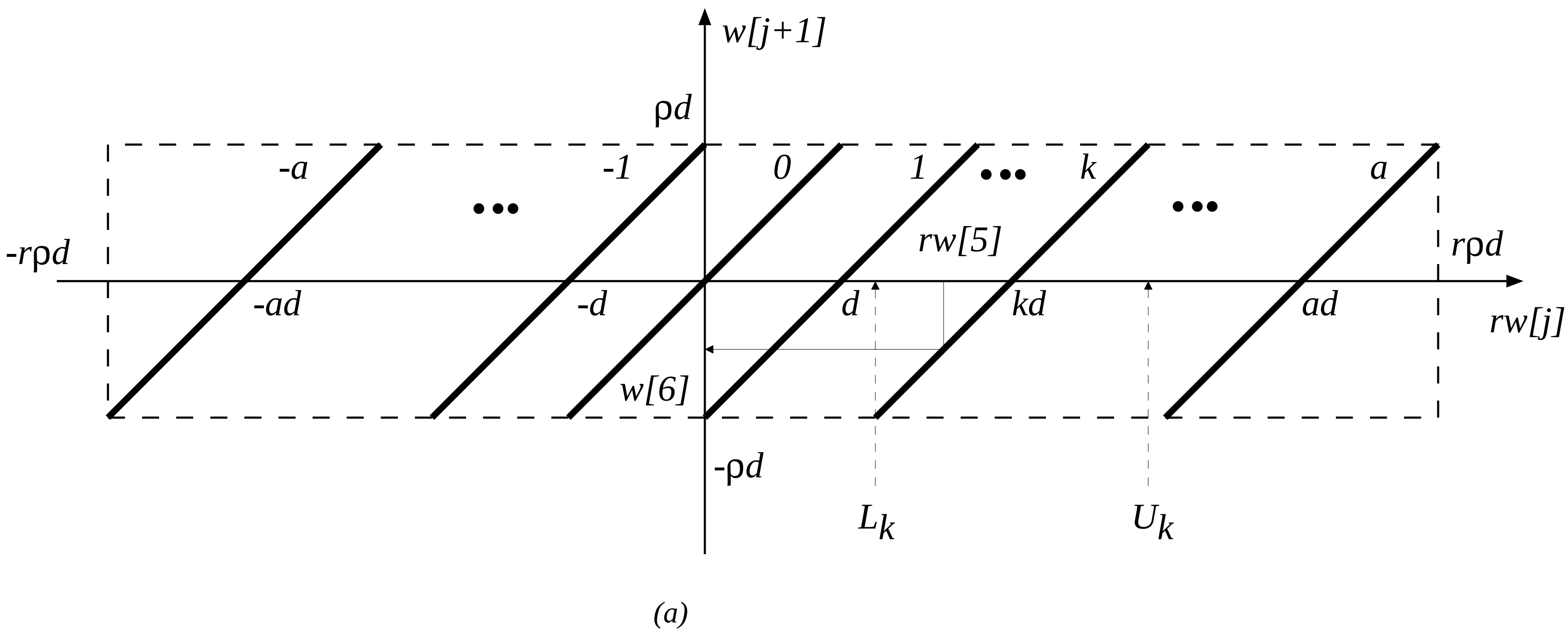


Figure 5.16: ROBERTSON'S DIAGRAM

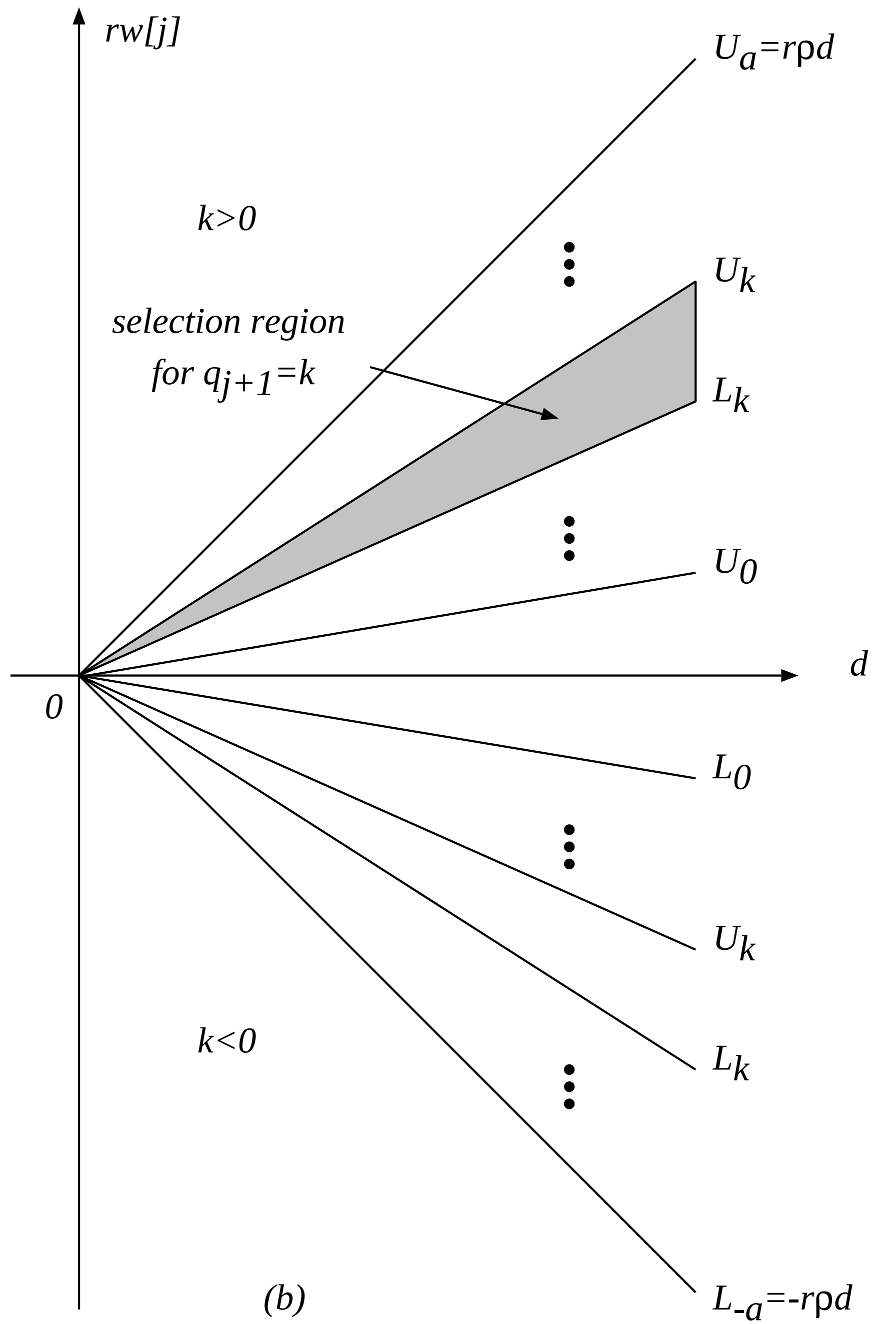


Figure 5.16: P-D DIAGRAM (for $d > 0$).

CONTINUITY CONDITION, OVERLAP, AND Q-DIGIT SELECTION

$$q_{j+1} = SEL(w[j], d)$$

- SEL represented by the set $\{s_k\}, -a \leq k \leq a$,

$$q_{j+1} = k \quad \text{if } s_k \leq rw[j] \leq s_{k+1} - ulp$$

- s_k defined as the **minimum** value of $rw[j]$ for which $q_{j+1} = k$
- s_k's are functions of the divisor d
- CONTAINMENT: $L_k \leq s_k \leq U_k$
- CONTINUITY: $q_{j+1} = k - 1$ for $rw[j] = s_k - ulp \leq U_{k-1}$

$$U_k \geq U_{k-1} + ulp \quad \rightarrow L_k \leq s_k \leq U_{k-1} + ulp \text{ or } L_k \leq s_k \leq U_{k-1}$$

- OVERLAP

$$U_{k-1} - L_k = (k - 1 + \rho)d - (k - \rho)d = (2\rho - 1)d$$

RESULTING IN

$$\rho \geq 2^{-1}$$

- REDUNDANCY IN q-DIGIT SET $\rightarrow$ OVERLAP BETWEEN SELECTION INTERVALS - SIMPLER SELECTION

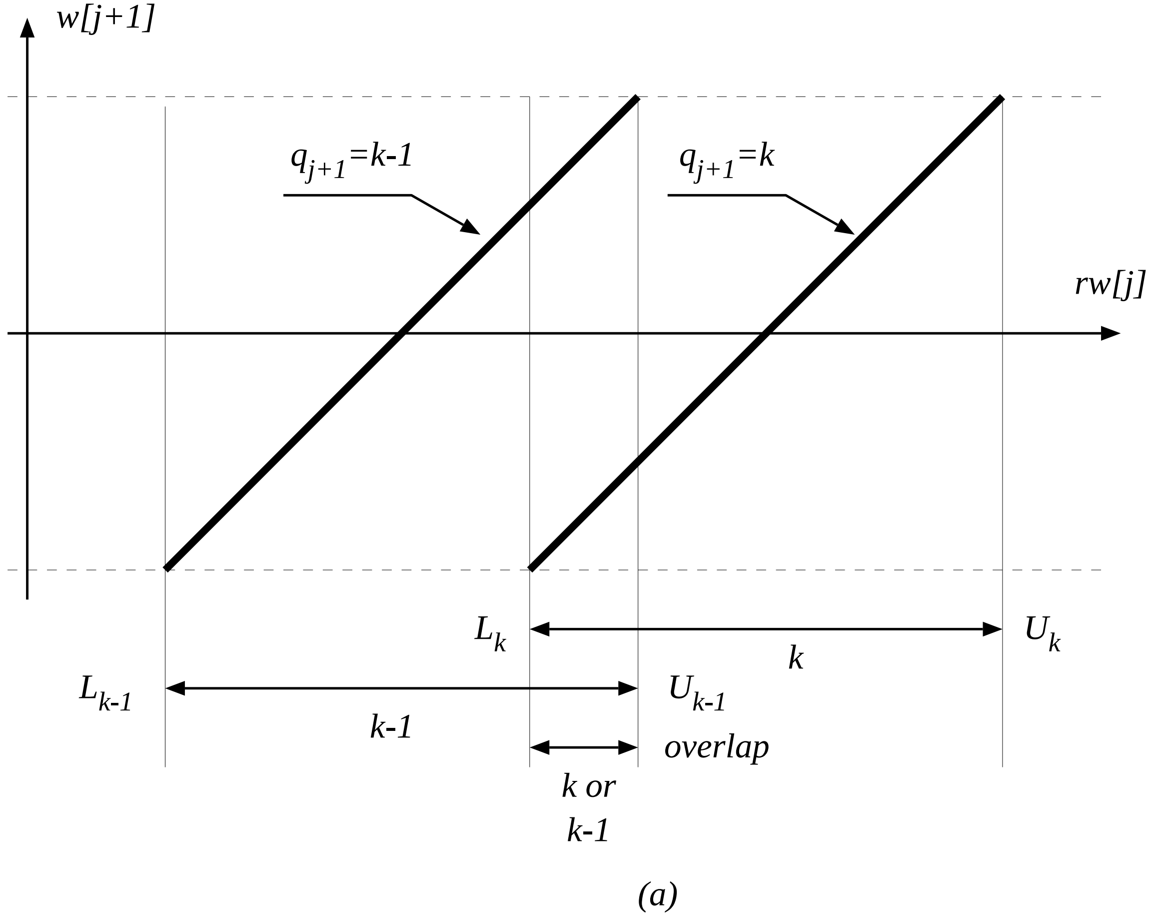


Figure 5.17: Overlap between selection intervals.

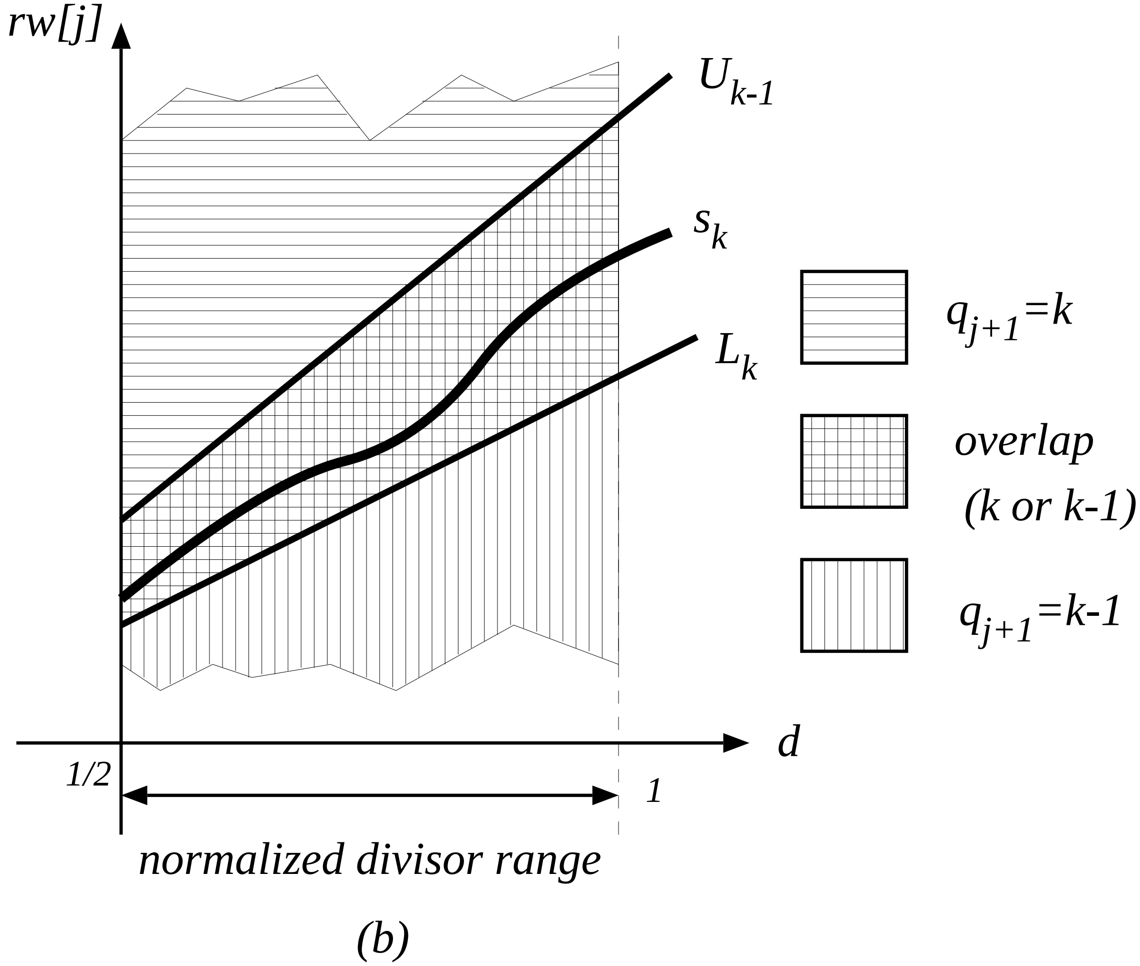


Figure 5.17: Selection function.

SELECTION USING CONSTANTS

- USE CONSTANTS m_k, INDEPENDENT OF DIVISOR

$$max(L_k) \leq m_k \leq min(U_{k-1}) + ulp$$

max and min for the range $2^{-1} \leq d < 1$

- For $k > 0$

$$(k - \rho) \leq m_k \leq (k - 1 + \rho)2^{-1} + ulp$$

which requires

$$\rho \geq \frac{k+1}{3}$$

- For $k \leq 0$

$$(k - \rho)2^{-1} \leq k - 1 + \rho$$

which requires

$$\rho \geq \frac{(-k)+2}{3}$$

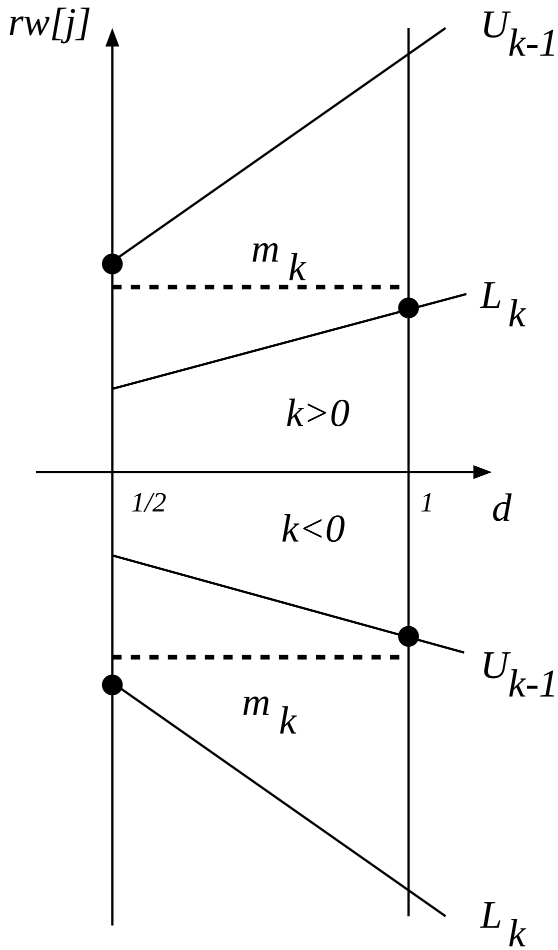


Figure 5.18: BOUNDS ON m_k.

RADIX-2 DIVISION WITH NON-REDUNDANT RESIDUAL

- EXTENSION OF NON-RESTORING DIVISION: {-1,0,1} SRT
- ALLOWS SKIPPING OVER ZEROS

$$\begin{array}{ll} U_1 = 2d & L_1 = 0 \\ U_0 = d \geq 1/2 & L_0 = -d \leq -1/2 \\ U_{-1} = 0 & L_{-1} = -2d \end{array}$$

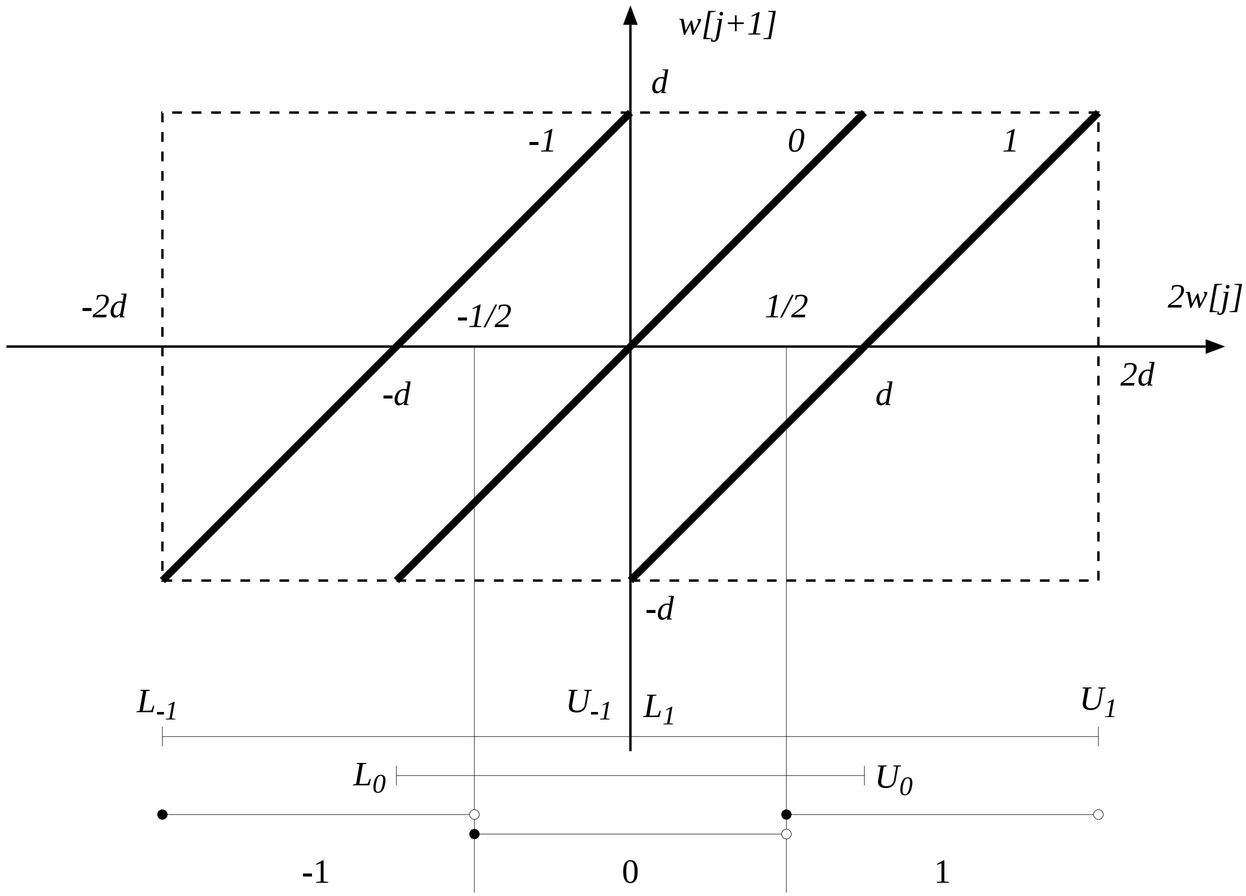
- SELECTION CONSTANTS:

$$0 \leq m_1 \leq 1/2, \qquad -1/2 \leq m_0 \leq 0$$

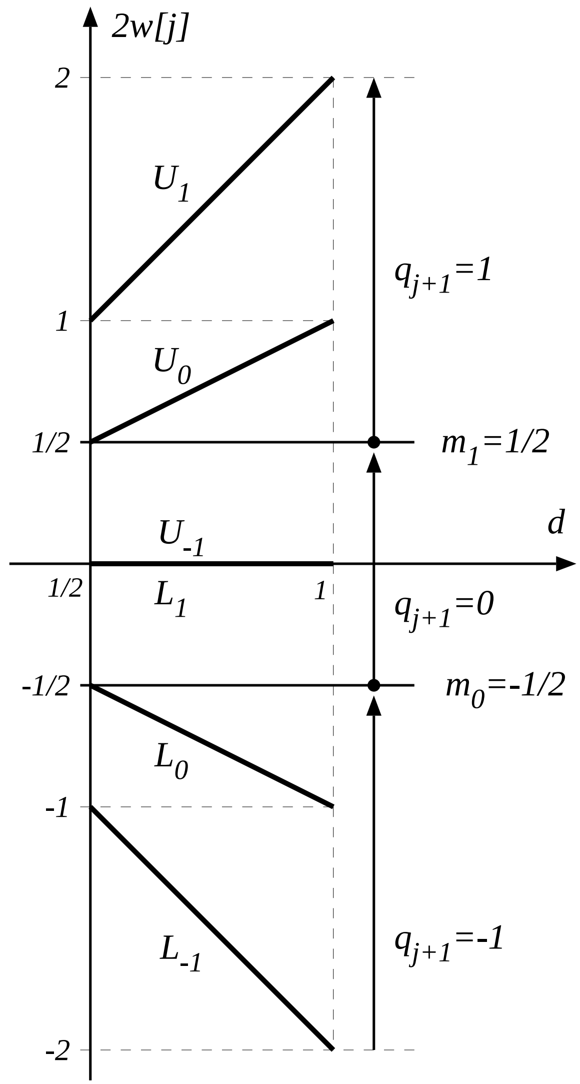
Choose: $m_1 = 1/2$ and $m_0 = -1/2$

- THE QUOTIENT-DIGIT SELECTION FUNCTION

$$q_{j+1} = \begin{cases} 1 & \text{if } 1/2 \leq 2w[j] \\ 0 & \text{if } -1/2 \leq 2w[j] < 1/2 \\ -1 & \text{if } 2w[j] < -1/2 \end{cases}$$



(a)


2w[j]
2
U1
qj+1=1
1
U0
1/2
m1=1/2
d
U-1
1/2
L1
1
qj+1=0
-1/2
m0=-1/2
L0
-1
L-1
qj+1=-1
-2

STAIRCASE SELECTION FUNCTION

- FOR $r > 2$, m_k DEPENDS ON DIVISOR

- DIVIDE RANGE OF DIVISOR INTO INTERVALS $[d_i, d_{i+1})$ with

$$d_0 = \frac{1}{2}, \qquad d_{i+1} = d_i + 2^{-\delta}$$

δ MS FRACTIONAL BITS OF DIVISOR REPRESENT THE INTERVAL

- FOR EACH INTERVAL, THERE IS A SET of *selection constants* $m_k(i)$

$$\text{for } d \in [d_i, d_{i+1}), \quad q_{j+1} = k \;\; \text{if} \;\; m_k(i) \leq rw[j] \leq m_{k+1}(i) - ulp$$

DEFINITION OF $m_k(i)$

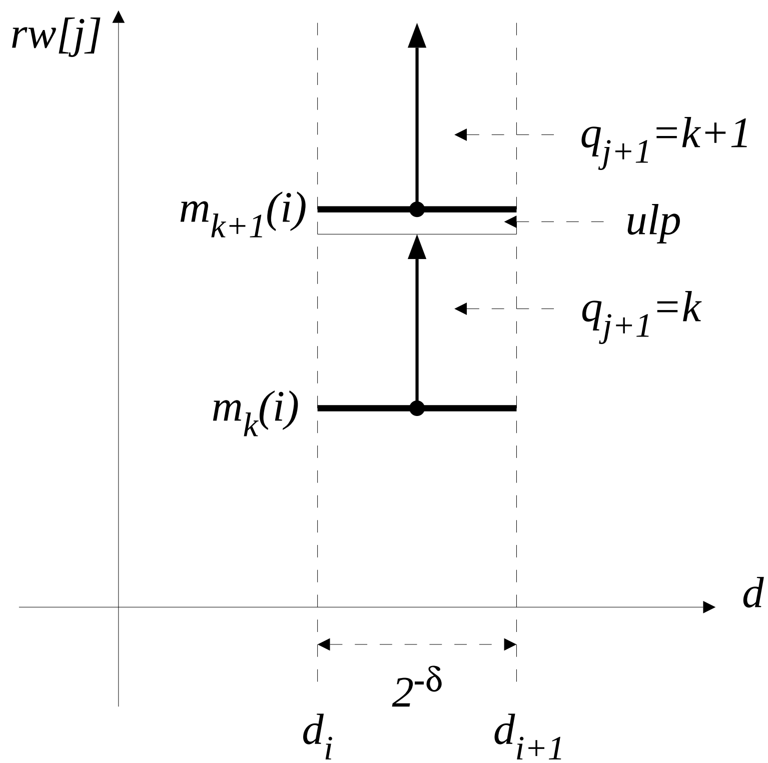


Figure 5.20: DEFINITION OF $m_k(i)$.

RANGE CONDITION FOR SELECTION CONSTANTS

$$max(L_k(d_i), L_k(d_{i+1})) \leq m_k(i) \leq min(U_{k-1}(d_i), U_{k-1}(d_{i+1})) + ulp$$

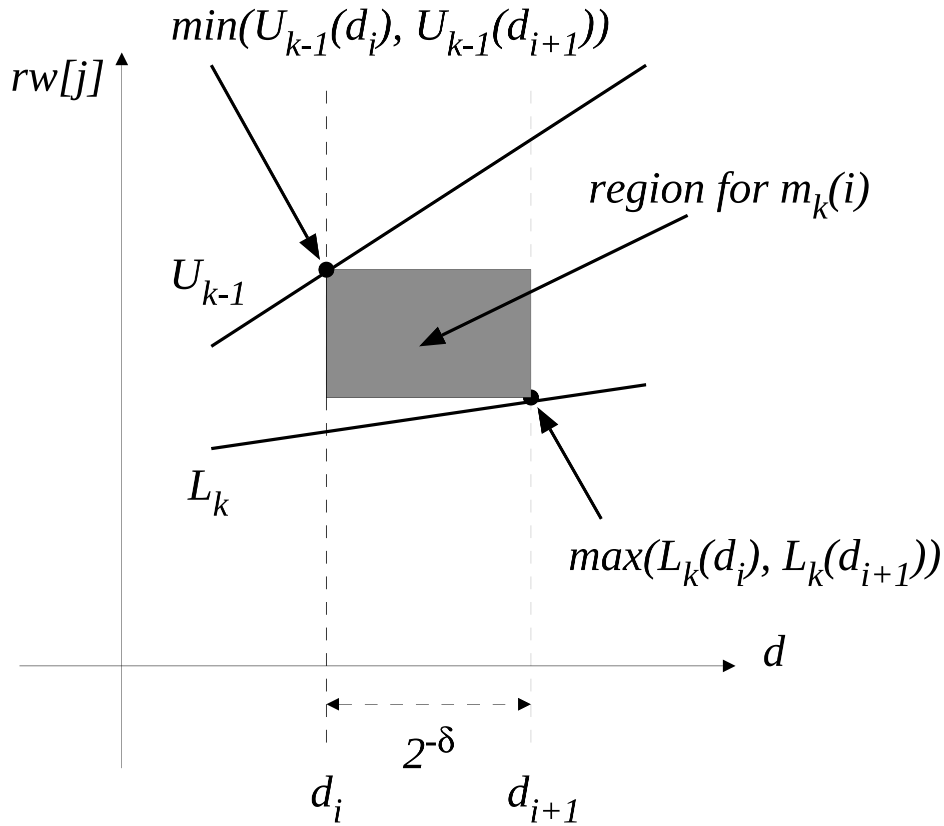


Figure 5.21: SELECTION CONSTANT REGION.

LOW-PRECISION SELECTION CONSTANTS

$$m_k(i) = A_k(i)2^{-c}$$

WHERE $A_k(i)$ IS INTEGER

- CAN USE TRUNCATED RESIDUAL IN COMPARISONS WITH SELECTION CONSTANTS
- RESIDUAL MUST BE IN 2's COMPLEMENT
- SELECTION CONDITIONS

$$\begin{array}{ll} \text{for } k > 0 & L_k(d_i + 2^{-\delta}) \leq A_k(i)2^{-c} \leq U_{k-1}(d_i) \\ \text{for } k \leq 0 & L_k(d_i) \leq A_k(i)2^{-c} \leq U_{k-1}(d_i + 2^{-\delta}) + ulp \end{array} \tag{5.1}$$

USE OF SELECTION CONSTANTS

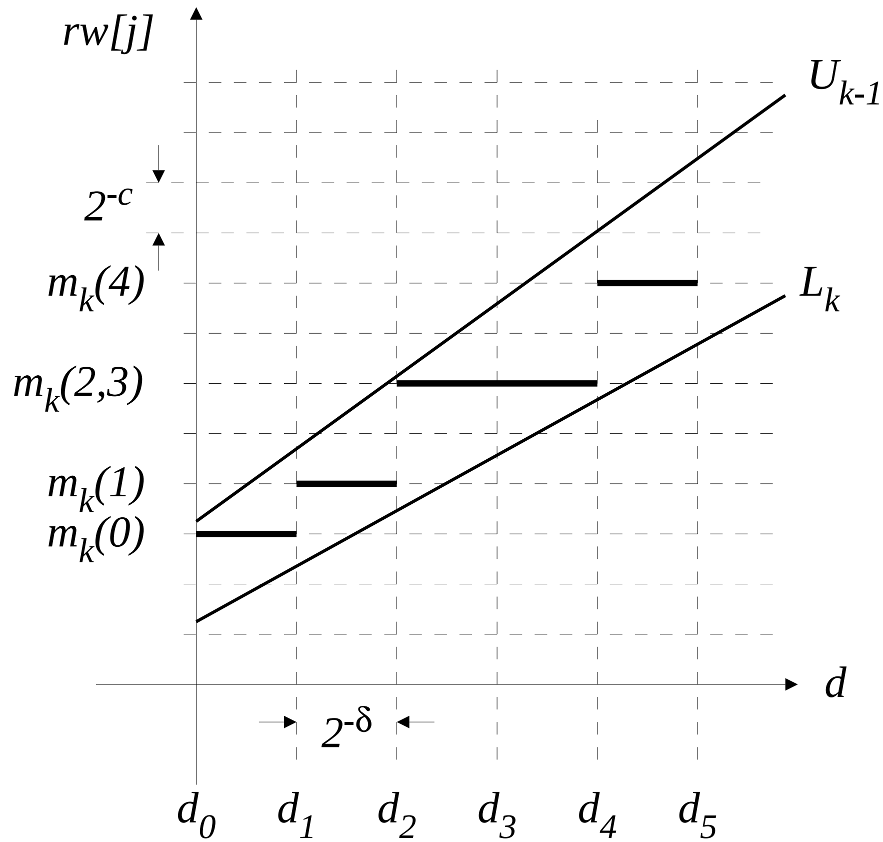


Figure 5.22: QUOTIENT-DIGIT SELECTION WITH SELECTION CONSTANTS.

SELECTION WITH TRUNCATED RESIDUAL AND DIVISOR

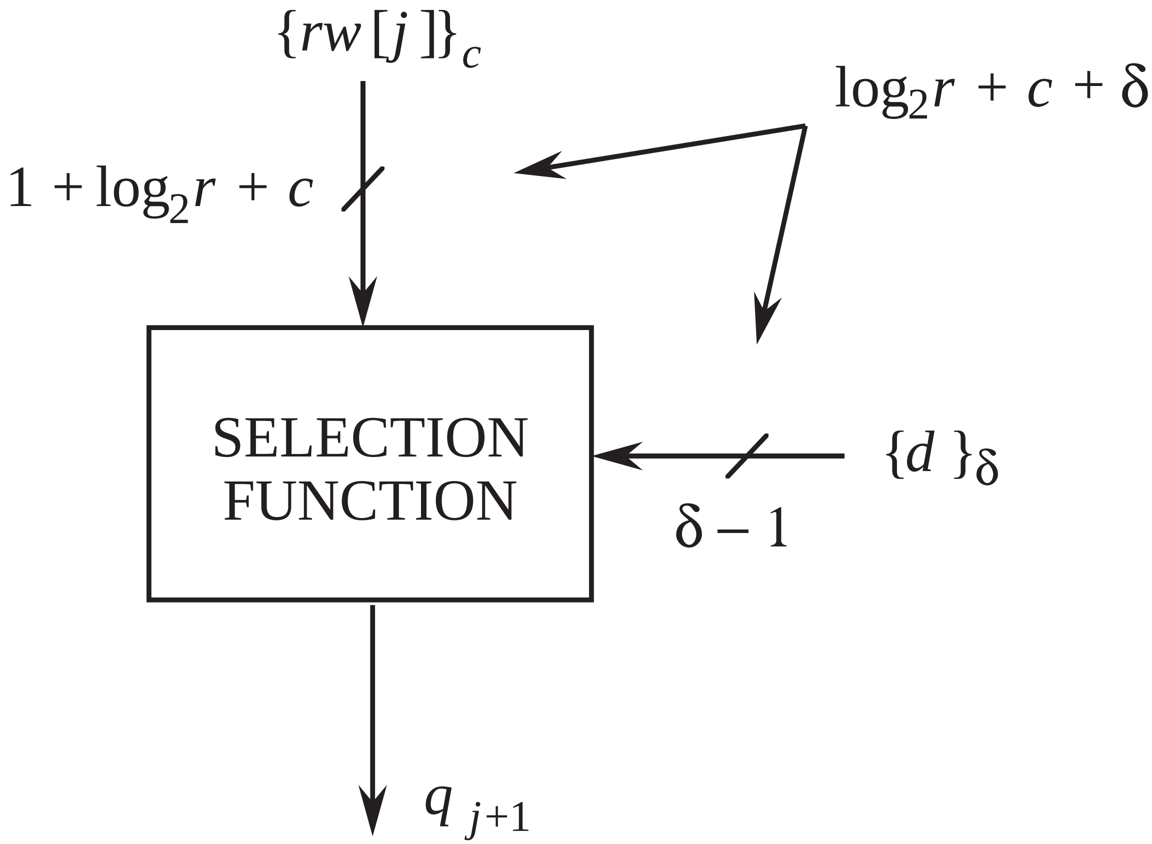


Figure 5.23: SELECTION WITH TRUNCATED RESIDUAL AND DIVISOR.

LOWER BOUND ON δ

- CONSIDER CASE $k > 0$ (similar argument for $k \leq 0$);
- FROM CONTINUITY CONDITION

$$U_{k-1}(d_i) - L_k(d_i + 2^{-\delta}) \geq 0$$

$$(\rho + k - 1)d_i - (-\rho + k)(d_i + 2^{-\delta}) \geq 0$$

$\Longrightarrow$

$$(2\rho - 1)d_i \geq (k - \rho)2^{-\delta}$$

- Worst case: $d \geq 1/2$ and $k \leq a$:

$$2^{-\delta} \leq \frac{2\rho - 1}{2(a - \rho)} = \frac{2\rho - 1}{2\rho(r - 2)}$$

- MINIMUM VALUE OF δ CAN RESULT IN A LARGE VALUE OF c
- OPTIMIZE THE VALUES OF δ AND c TOGETHER

RADIX-4 DIVISION WITH NONREDUNDANT RESIDUAL

- Known as Robertson's division
- $q_j \in \{-2, -1, 0, 1, 2\}$
- $U_k = (\frac{2}{3} + k)d \qquad L_k = (-\frac{2}{3} + k)d$
- BOUND ON δ

$$2^{-\delta} \leq \frac{2\rho - 1}{2(a - \rho)} = \frac{1}{8}$$

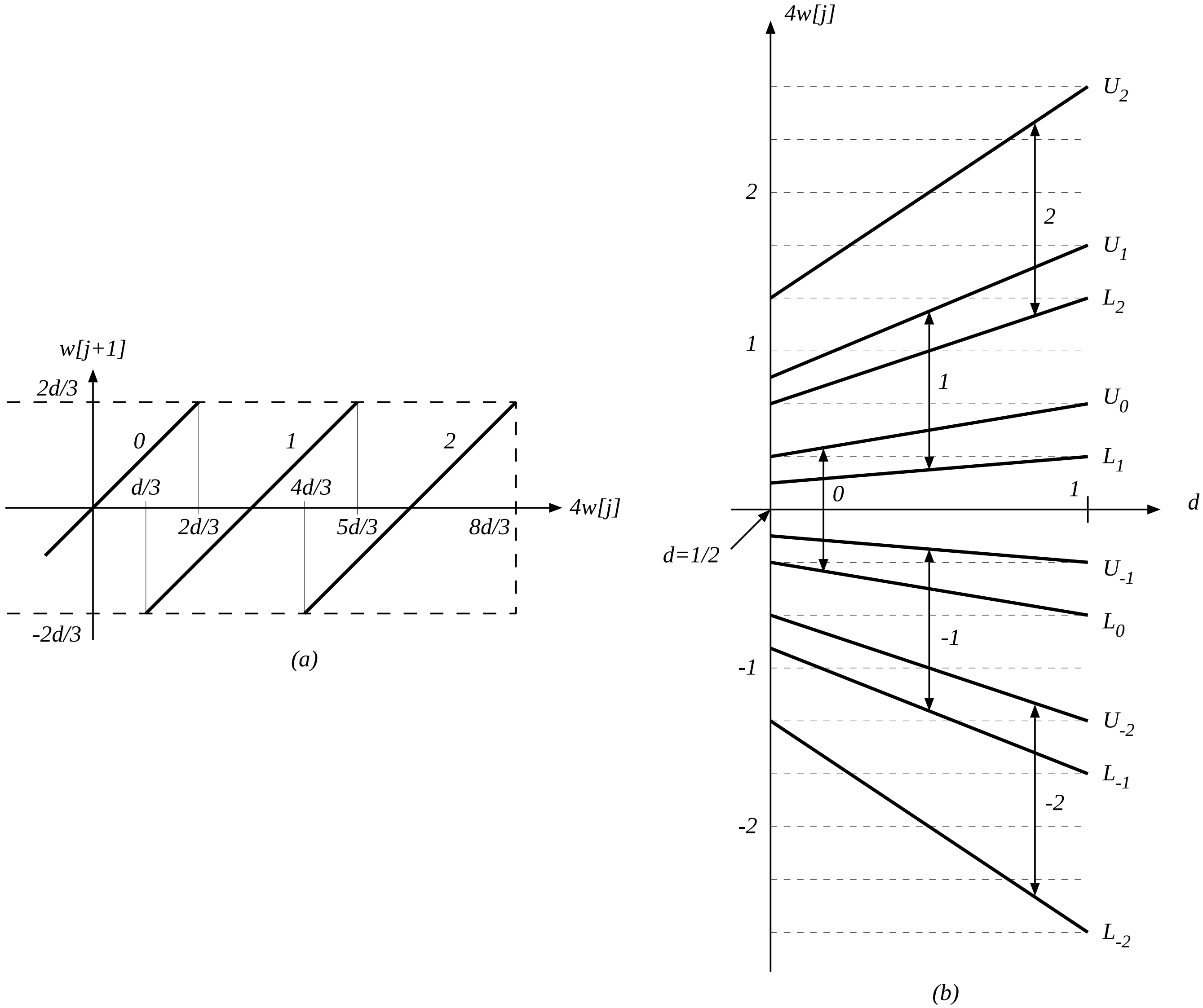


Figure 5.24: ROBERTSON'S AND PD DIAGRAMS FOR RADIX-4 AND $a = 2$.

$[d_i, d_{i+1})^+$	[8, 9)	[9, 10)	[10, 11)	[11, 12)
$L_2(d_{i+1}), U_1(d_i)^\#$	36, 40	40, 45	44, 50	48, 55
$m_2(i)^*$	**6**	**7**	**8**	**8**
$L_1(d_{i+1}), U_0(d_i)^\#$	9,16	10, 18	11, 20	12, 22
$m_1(i)$	**2**	**2**	**2**	**2**
$L_0(d_i), U_{-1}(d_{i+1})^\#$	-16, -9	-18, -10	-20, -11	-22, -12
$m_0(i)$	**-2**	**-2**	**-2**	**-2**
$L_{-1}(d_i), U_{-2}(d_{i+1})^\#$	-40, -36	-45, -40	-50, -44	-55, -48
$m_{-1}(i)$	**-6**	**-7**	**-8**	**-8**
$[d_i, d_{i+1})^+$	[12, 13)	[13, 14)	[14, 15)	[15, 16)
$L_2(d_{i+1}), U_1(d_i)^\#$	52, 60	56, 65	60, 70	64, 75
$m_2(i)$	**10**	**10**	**10**	**12**
$L_1(d_{i+1}), U_0(d_i)^\#$	13, 24	14, 26	15, 28	16, 30
$m_1(i)$	**4**	**4**	**4**	**4**
$L_0(d_i), U_{-1}(d_{i+1})^\#$	-24, -13	-26, -14	-28, -15	-30, -16
$m_0(i)$	**-4**	**-4**	**-4**	**-4**
$L_{-1}(d_i), U_{-2}(d_{i+1})^\#$	-60, -52	-65, -56	-70, -60	-75, -64
$m_{-1}(i)$	**-10**	**-10**	**-10**	**-12**

QUOTIENT-DIGIT SELECTION FOR RADIX-4 DIVISION; NON-REDUNDANT RESIDUAL

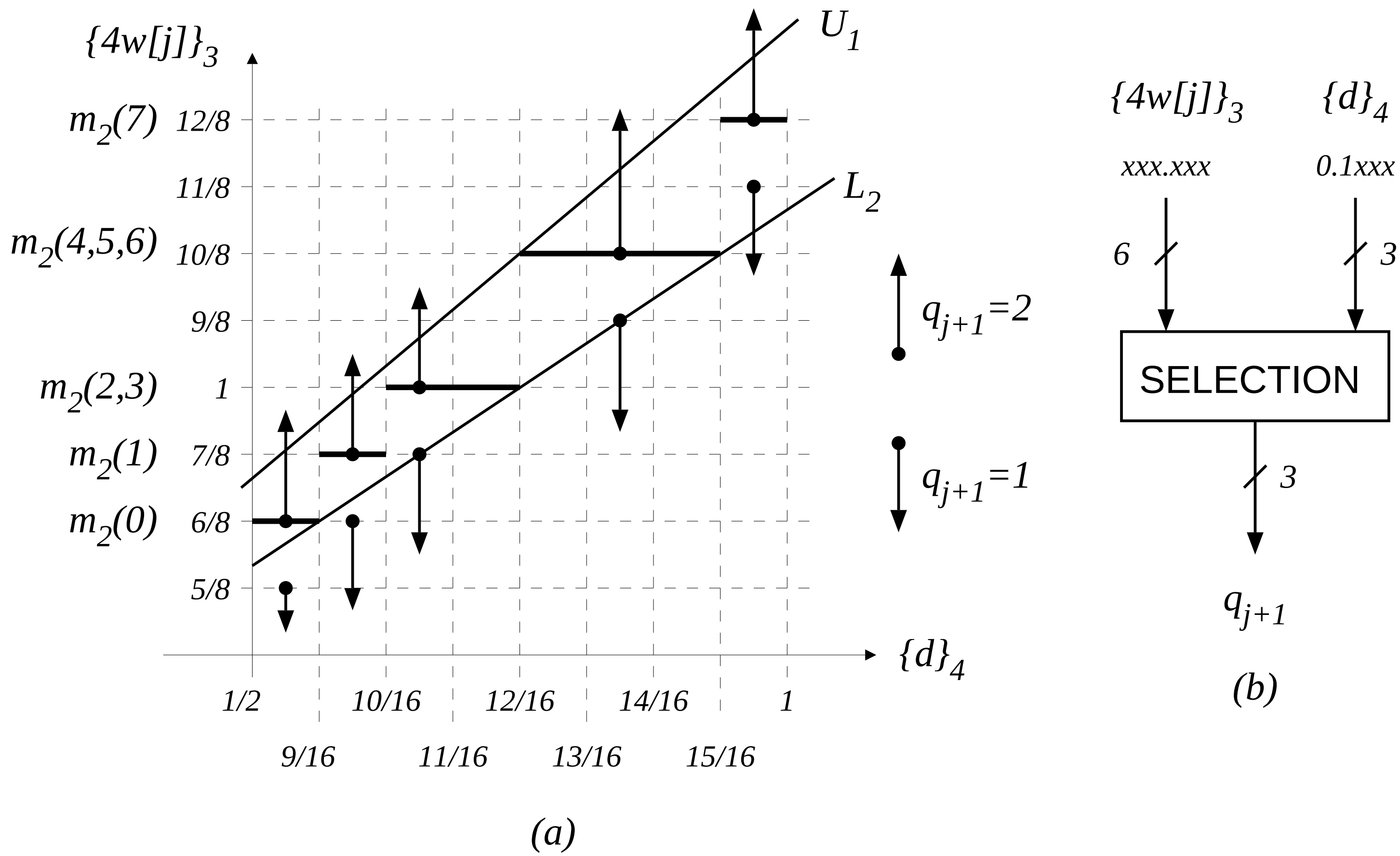


Figure 5.25: QUOTIENT-DIGIT SELECTION: (a) FRAGMENT OF THE P-D DIAGRAM. (b) IMPLEMENTATION.

USE OF REDUNDANT ADDER

- SO FAR: COMPUTE $rw[j]$ IN FULL PRECISION, TRUNCATE, AND COMPARE WITH LOW-PRECISION CONSTANTS
- FULL-PRECISION ADDITION: significant portion of the cycle time
- OVERLAP BETWEEN SELECTION INTERVALS $\Longrightarrow$COULD USE AN ESTIMATE OF $rw[j]$
- ERROR IN ESTIMATE:

$$\epsilon_{min} \leq y - \widehat{y} \leq \epsilon_{max}$$

- BASIC CONSTRAINT: if we choose $q_{j+1} = k$ for an estimate $\widehat{y}$ then

$$y \in [\widehat{y} + \epsilon_{min},\ \widehat{y} + \epsilon_{max}]$$

$$L_k^* = L_k - \epsilon_{min}$$

$$U_k^* = U_k - \epsilon_{max}$$

CONSTRAINTS ON SELECTION CONSTANTS

$$max(L_k^*(d_i), L_k^*(d_{i+1})) \leq m_k(i) \leq min(U_{k-1}^*(d_i), U_{k-1}^*(d_{i+1}))$$

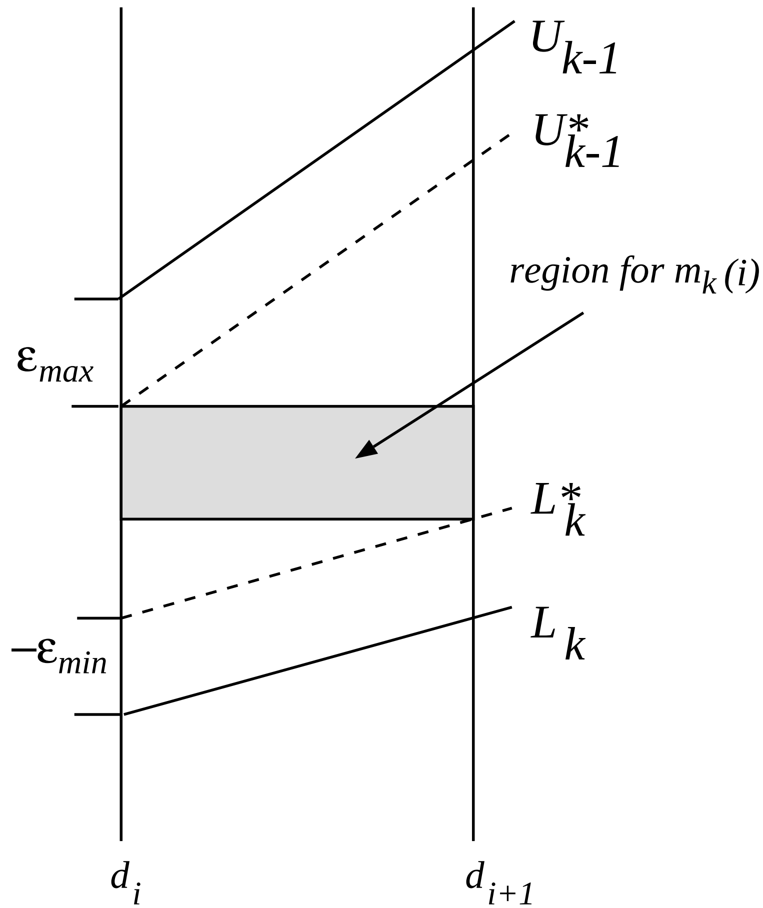


Figure 5.26: CONSTRAINTS FOR SELECTION BASED ON ESTIMATES.

MINIMUM OVERLAP AND RANGE OF ESTIMATE

- OVERLAP

$$min(U^*_{k-1}(d_i), U^*_{k-1}(d_{i+1})) - max(L^*_k(d_i), L^*_k(d_{i+1})) \geq 0$$

- RANGE

$$|rw[j]| \leq r\rho d < r\rho \quad (\text{for } d < 1)$$

$$-r\rho - \epsilon_{max} < \widehat{y} < r\rho - \epsilon_{min}$$

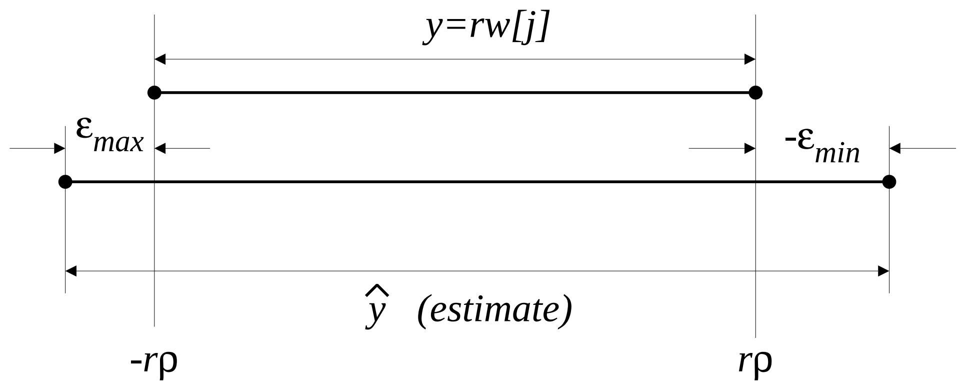


Figure 5.27: RANGE OF ESTIMATE

REDUNDANT ADDER

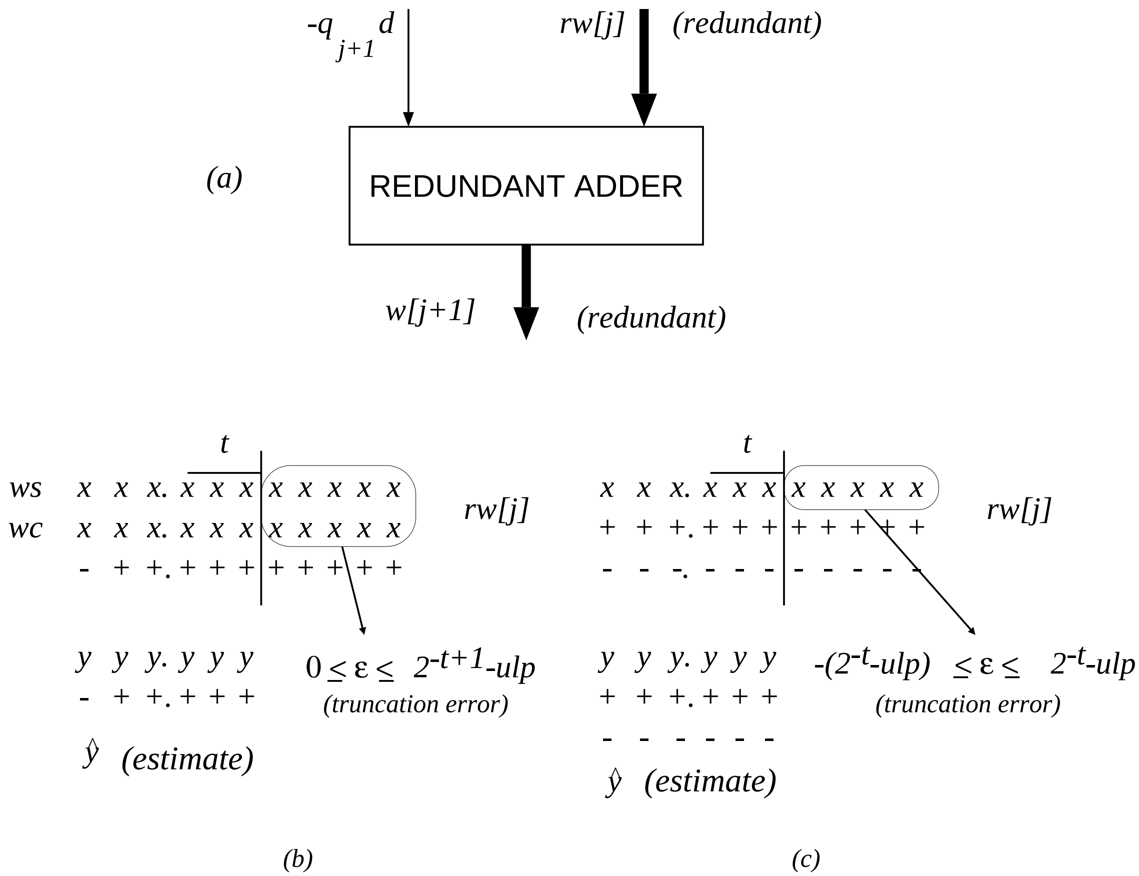


Figure 5.28: USE OF REDUNDANT ADDER: (a) Redundant adder. (b) Carry-save case. (c) Signed-digit case.

CARRY-SAVE ADDER

- ERRORS

$$\epsilon_{min} = 0 \qquad \epsilon_{max} = 2^{-t+1} - ulp$$

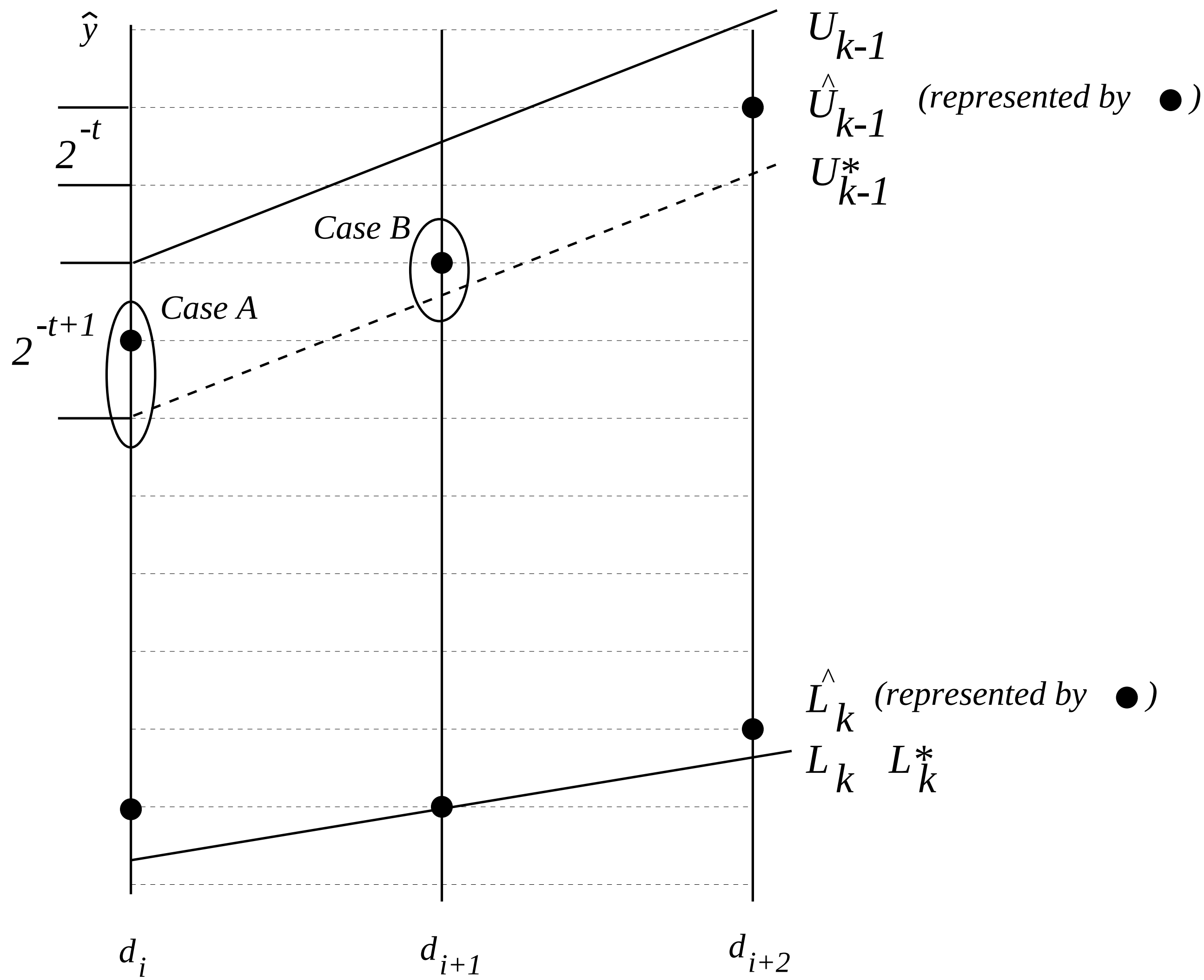
- RESTRICTED SELECTION INTERVAL

$$U_k^* = U_k - 2^{-t+1} + ulp$$

$$L_k^* = L_k$$

$$\widehat{U}_{k-1} = \lfloor U_{k-1}^* + 2^{-t} \rfloor_t = \lfloor U_{k-1} - 2^{-t} \rfloor_t$$

$$\widehat{L}_k = \lceil L_k^* \rceil_t = \lceil L_k \rceil_t$$



Case A: U^*_{k-1} *is on the grid;*

$\hat{U}_{k-1} = U^*_{k-1} + 2^{-t}$ *on the grid*

Case B: U^*_{k-1} *is off the grid;*

$\hat{U}_{k-1} > U^*_{k-1}$ *on the grid*

Figure 5.29: $\hat{U}$ and $\hat{L}$ for residual in carry-save form.

LOWER BOUND FOR t and δ

(for positive k)

$$\widehat{U}_{k-1}(d_i) - \widehat{L}_k(d_{i+1}) \geq 0$$

$$U_{k-1}(d_i) - 2^{-t} - L_k(d_{i+1}) \geq 0$$

$$\frac{2\rho - 1}{2} - (a - \rho)2^{-\delta} \geq 2^{-t}$$

- RANGE

$$\lfloor -r\rho - 2^{-t} \rfloor_t \leq \widehat{y} \leq \lfloor r\rho - ulp \rfloor_t$$

$\lfloor z \rfloor_t = 2^{-t} \lfloor 2^t z \rfloor$

RADIX-2 DIVISION WITH CARRY-SAVE ADDER

$$\frac{1}{2} - 0 \times 2^{-\delta} \geq 2^{-t}$$

$$max(\widehat{L}_k(d_i), \widehat{L}_k(d_{i+1})) \leq m_k(i) \leq min(\widehat{U}_{k-1}(d_i), \widehat{U}_{k-1}(d_{i+1}))$$

$$\begin{aligned} \widehat{L}_1(1) &= 0 \\ \widehat{U}_0(1/2) &= 0 \\ \widehat{L}_0(1/2) &= -1/2 \\ \widehat{U}_{-1}(1) &= -1/2 \end{aligned}$$

$$(\widehat{L}_1(1) = 0) \leq m_1 \leq (\widehat{U}_0(1/2) = 0)$$

$$(\widehat{L}_0(1/2) = -1/2) \leq m_0 \leq (\widehat{U}_{-1}(1) = -1/2)$$

This results in the selection constants $m_1 = 0$ and $m_0 = -1/2$

SELECTION FUNCTION

$$\lfloor -2 - 2^{-1} \rfloor_1 \leq \widehat{y} \leq \lfloor 2 - ulp \rfloor_1$$

$$-\frac{5}{2} \leq \widehat{y} \leq 3/2$$

$$q_{j+1} = \begin{cases} 1 & \text{if } 0 \leq \widehat{y} \leq 3/2 \\ 0 & \text{if } \widehat{y} = -1/2 \\ -1 & \text{if } -5/2 \leq \widehat{y} \leq -1 \end{cases}$$

RADIX-2 WITH CSA: PD PLOT

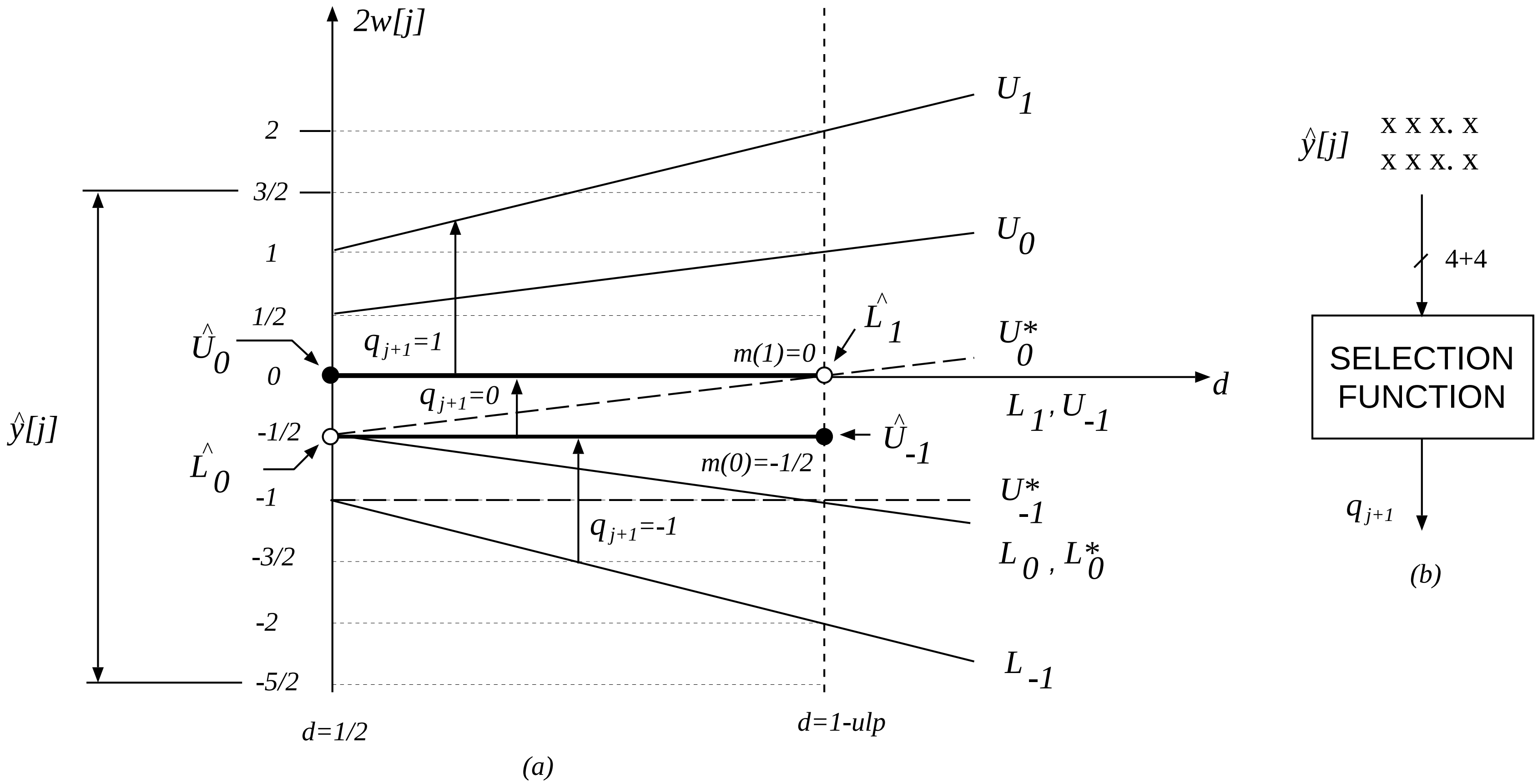


Figure 5.30: RADIX-2 DIVISION WITH CARRY-SAVE ADDER: (a) P-D PLOT. (b) SELECTION FUNCTION.

IMPLEMENTATION

$q_{j+1} = (q_s, q_m)$

$$q_m = (p_{-1}p_0p_1)' \tag{5.2}$$
$$q_s = p_{-2} \oplus (g_{-1} + p_{-1}g_0 + p_{-1}p_0g_1)$$

where

$$p_i = c_i \oplus s_i \qquad g_i = c_i \cdot s_i$$

and

$$(c_{-2}, c_{-1}, c_0, c_1)$$

$$(s_{-2}, s_{-1}, s_0, s_1)$$

$$q_{j+1} = 0 \qquad (q_s, q_m) = (1, 0)$$

RADIX-4 ALGORITHM WITH CARRY-SAVE ADDER

$$\frac{1}{6} - \frac{4}{3}2^{-\delta} \geq 2^{-t}$$

$$2^{-t} \leq \frac{1}{6} - \frac{1}{12} = \frac{1}{12}$$

$$\lfloor -8/3 - 1/16 \rfloor_4 \leq \hat{y} \leq \lfloor 8/3 - ulp \rfloor_4$$

$$-\frac{44}{16} \leq \hat{y} \leq \frac{42}{16}$$

$[d_i, d_{i+1})^+$	[8, 9)	[9, 10)	[10, 11)	[11, 12)
$\widehat{L}_2(d_{i+1}), \widehat{U}_1(d_i)^+$	12, 12	14, 14	15, 15	16, 17
$m_2(i)^+$	**12**	**14**	**15**	**16**
$\widehat{L}_1(d_{i+1}), \widehat{U}_0(d_i)^+$	3, 4	4, 5	4, 5	4, 6
$m_1(i)$	**4**	**4**	**4**	**4**
$\widehat{L}_0(d_i), \widehat{U}_{-1}(d_{i+1})^+$	-5, -4	-6, -5	-6, -5	-7, -5
$m_0(i)$	**-4**	**-6**	**-6**	**-6**
$\widehat{L}_{-1}(d_i), \widehat{U}_{-2}(d_{i+1})^+$	-13, -13	-15, -15	-16, -1 6	-18, -17
$m_{-1}(i)$	**-13**	**-15**	**-16**	**-18**
$[d_i, d_{i+1})^+$	[12, 13)	[13, 14)	[14, 15)	[15, 16)
$\widehat{L}_2(d_{i+1}), \widehat{U}_1(d_i)^+$	18, 19	19, 20	20, 22	22, 24
$m_2(i)$	**18**	**20**	**20**	**24**
$\widehat{L}_1(d_{i+1}), \widehat{U}_0(d_i)^+$	4, 7	5, 7	5, 8	6, 9
$m_1(i)$	**6**	**6**	**8**	**8**
$\widehat{L}_0(d_i), \widehat{U}_{-1}(d_{i+1})^+$	-8, -6	-8, -6	-9, -6	-10, -7
$m_0(i)$	**-8**	**-8**	**-8**	**-8**
$\widehat{L}_{-1}(d_i), \widehat{U}_{-2}(d_{i+1})^+$	-20, -19	-21, -20	-23, -2 1	-25, -23
$m_{-1}(i)$	**-20**	**-20**	**-22**	**-24**

+: real value = shown value/16; $\widehat{L}_k = \lceil L_k \rceil_4$, $\widehat{U}_k = \lfloor U_k - \frac{1}{16} \rfloor_4$.

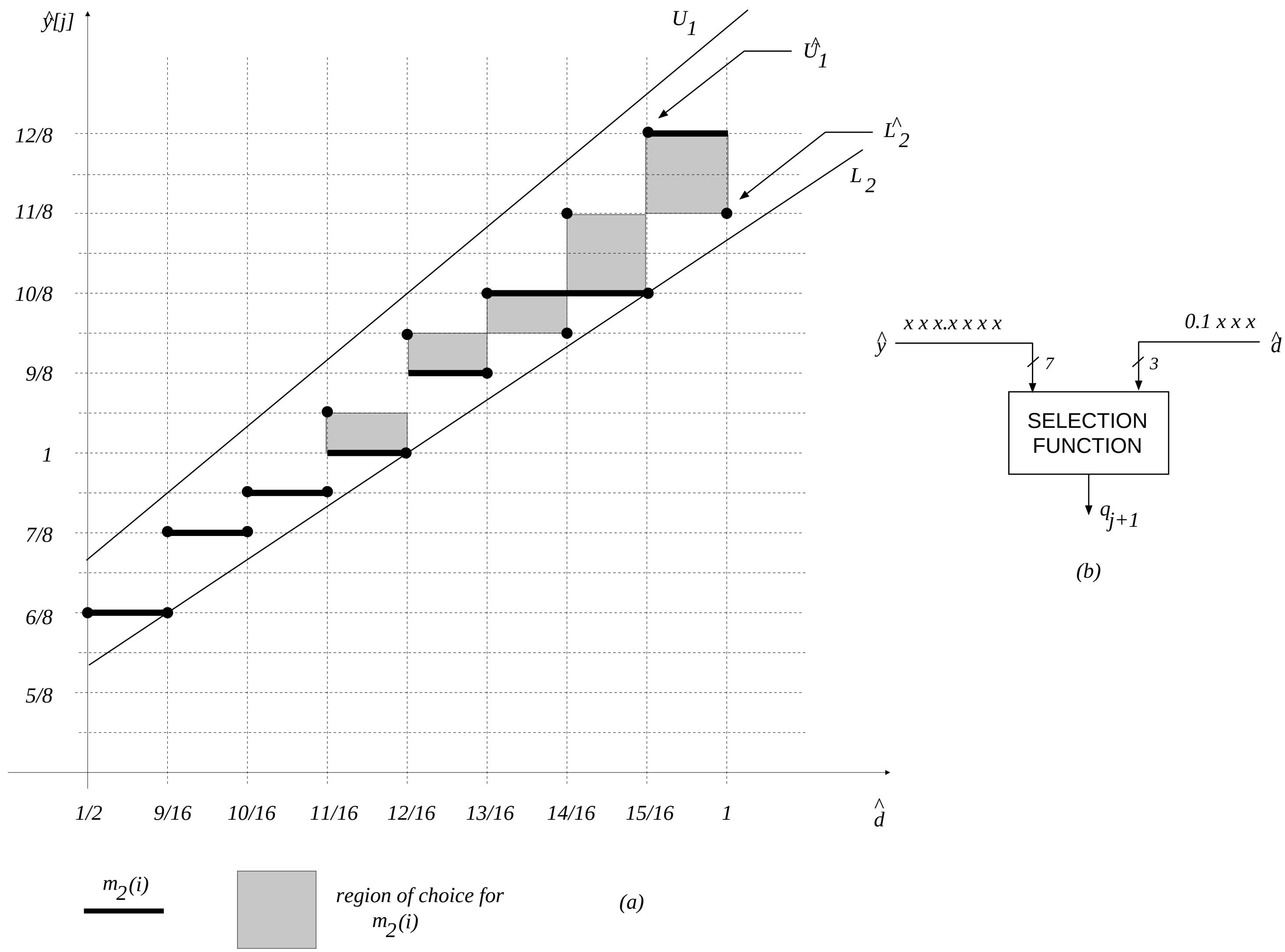


Figure 5.31: SELECTION FUNCTION FOR RADIX-4 SCHEME WITH CARRY-SAVE ADDER: (a) Fragment of P-D diagram. (b) Implementation.