9.8 a) The only values that need to be computed explicitly are the values to be stored in the delay elements and the outputs. Eliminating all intermediate node values in the recurrence equations, which were derived in Problem 6.7, we get the difference equations in computable order:

$$v_2(n+1) := v_1(n)$$
$$v_1(n+1) := (\alpha_1 - 1)\, x(n) - \alpha_1\, v_2(n)$$
$$v_4(n+1) := v_3(n)$$
$$v_3(n+1) := (\alpha_3 + 1)\, v_0(n) - \alpha_3\, v_4(n)$$
$$v_6(n+1) := v_5(n)$$
$$v_5(n+1) := \alpha_1(1+\alpha_5)\, x(n) + [\alpha_5 - \alpha_1(1+\alpha_5)]\, v_2(n) - \alpha_5\, v_6(n)$$
$$v_0(n+1) := x(n)$$
$$y_1(n) := \alpha_1\alpha_5\, x(n) + \alpha_3\, v_0(n) - \alpha_5(1-\alpha_1)\, v_2(n) + (1-\alpha_3)\, v_4(n) + (1+\alpha_5)\, v_6(n)$$
$$y_2(n) := \alpha_1\alpha_5\, x(n) - \alpha_3\, v_0(n) - \alpha_5(1-\alpha_1)\, v_2(n) - (1-\alpha_3)\, v_4(n) + (1+\alpha_5)\, v_6(n)$$

b) Inserting the quantized adaptor coefficient values we get:

$$v_1(n+1) := [1123\, x(n) + 99\, v_2(n)]\, 2^{-10}$$
$$v_3(n+1) := [1405\, v_0(n) + 381\, v_4(n)]2^{-10}$$
$$v_5(n+1) := [-53361\, x(n) + 1598577\, v_2(n) + 1545216\, v_6(n)]\, 2^{-21}$$
$$v_0(n+1) := x(n)$$
$$y_1(n) := [149391\, x(n) + 780288\, v_0(n) + 1694607\, v_2(n) + 2877440\, v_4(n) + 551936\, v_6(n)]\, 2^{-21}$$
$$y_2(n) := [149391\, x(n) - 780288\, v_0(n) + 1694607\, v_2(n) - 2877440\, v_4(n) + 551936\, v_6(n)]\, 2^{-21}$$

In practice, this new set of equations should be scaled in order to optimize the dynamic range. The scaled coefficients tend to be of the same magnitude. The word length of these new coefficients is rather long, since they have not been optimized for this application. Instead, the adaptor coefficients have been optimized to have a favorable representation with few non-zero digits.

Five vector-multipliers are needed for the fully parallel implementation showed in Fig. P9.8.

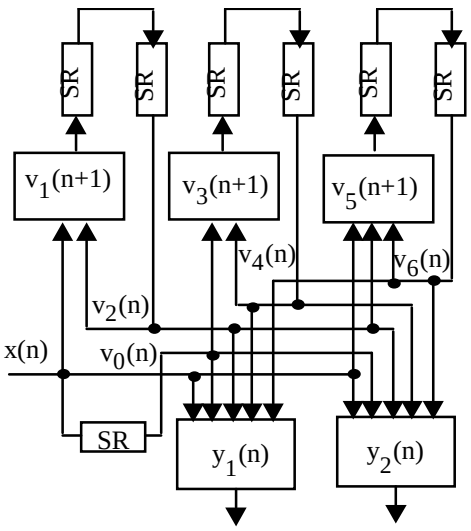
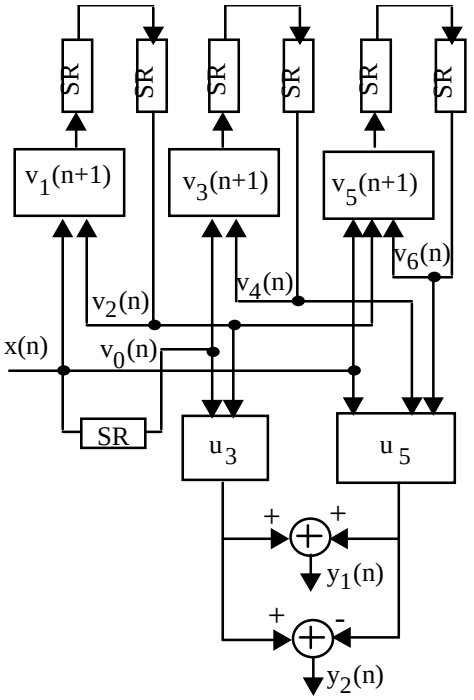


Fig. P9.8. Vector-multiplier based realization of lattice wave digital filter.

9.9 The outputs of the last two adaptors are computed using vector-multipliers. The outputs of the filter are then obtained by a bit-serial addition and subtraction of these values. The simplified implementation is shown below. This scheme reduces the number of terms in the two inner products, u_{34} and u_{54}, at the small expense of two adders.

The new set of difference equations is:

$$v_1(n+1) := [1123\, x(n) + 99\, v_2(n)]\, 2^{-10}$$
$$v_3(n+1) := [1405\, v_0(n) + 381\, v_4(n)]2^{-10}$$
$$v_5(n+1) := [-53361\, x(n) + 1598577\, v_2(n) + 1545216\, v_6(n)]\, 2^{-21}$$
$$v_0(n+1) := x(n)$$
$$u_{34}(n) := [381\, v_0(n) + 1405\, v_4(n)]2^{-10}$$
$$u_{54}(n) := [-149391\, x(n) + 1694607\, v_2(n) + 551936\, v_6(n)]2^{-21}$$
$$y_1(n) := u_{34}(n) + u_{54}(n)$$
$$y_2(n) := u_{34}(n) - u_{54}(n)$$



9.10 a) A shared-memory architecture is selected.

b) The number of bits/s to and from the processors are: $\dfrac{2N_{PE}}{T_{PE}}$

The number of bits/s to and from the memories are:

$$\frac{W_{RAM}N_{RAM}}{T_{RAM}}$$

These transmission rates should be equal for an ideal architecture. Hence, we have:

$$N_{RAM} = \frac{2N_{PE}T_{PE}}{W_{RAM}T_{PE}}$$

9.11 The total number of operations, N_{op}, is:

$$N_{op} = \frac{N}{2}\log_2(N) = \frac{1024}{2}\log_2(1024) = 5120$$

These are performed in T_{FFT} = 1 ms which means that the number of operations per second is:

$$\frac{N_{op}}{T_{FFT}} = \frac{5120}{1\times10^{-3}} = 5.12 \text{ MOp/s}$$

The clock frequency on the bit-serial lines through A-A' are:

$$f_{CL} = \frac{N_{op}}{T_{FFT}}W_d = 5.12\times10^6\times21 \approx 108 \text{ MHz}$$

The total bit rate through A-A' is: $(4 + 4)\times107\times10^6$ = 860 Mbit/s

The bit rate through B-B': $\frac{W_m}{T_m} = f_m\ W_m = 21\ f_m$

The bit rates through the cuts must be equal:

$860\times10^6 = 21\ f_m \Rightarrow f_m = 860\times10^6 / 21 = 41$ MHz

The memory clock rate will be halved if the word length is doubled. A single fast RAM (41 MHz) is enough to support the butterfly PE.

9.12 First we estimate the computational workload. The total workload is:

N_{op} = 352.8 + 352.8 + 352.8 + 705.6 = 1.764 MOp/s

The clock rate in cut A-A' is: $f_{CL} = N_{op}\ W_d = 1.764 \approx 35.3$ MHz

The bit rate in A-A' is: $(2 + 2)f_{CL} = 4 \times 35.3 = 141.2$ Mbit/s

The bit rate in B-B' is $f_{mem}\ W_{mem}$

Stage	1	2	3	4
Order	17	9	5	5
no. adaptors	8	4	2	2
frequency	44.1	88.2	176.4	352.8
kOPS	352.8	352.8	352.8	705.6

If we choose $W_{mem} = 20$ bit we will have: $f_{mem} = \frac{141.2}{20} \approx 7.1$ MHz

To get $f_{mem} < 20$ MHz we must select $W_{mem} > \frac{141.2}{20} \approx 7.1$ bit

9.13 a) The number of 1-D DCT that have to be computed per second is:

$$N_{tot} = (16 + 16 + 2) \times 240000$$

That is: 16 + 16 1-D DCT. We looses two cycles between the row and column computations because of the pipelining of the processors. The execution time for one processor is:

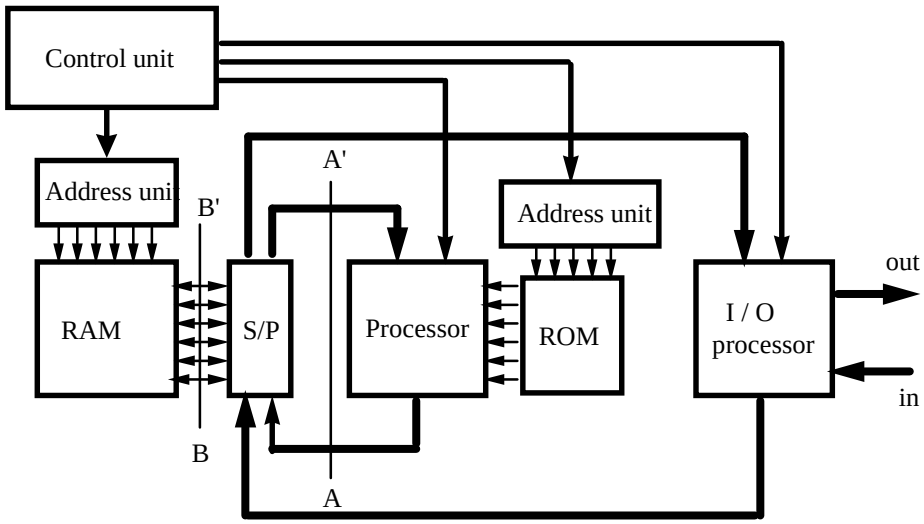
$\frac{14}{120 \times 10^6}$ at maximal clock frequency, i.e., we can compute:

$\frac{120 \times 10^6}{14} = 8.57 \times 10^6$ 1-D DCT/s. The required number of processors is:

$$N_p = \frac{34 \times 240000 \times 14}{120 \times 10^6} \approx 0.952 < 1$$

Thus, one processor is enough. The processor schedule becomes sequential, i.e., any feasible schedule is acceptable.

b)



c) The number of bits per second passing to and back from the processor is: $34 \times 240000 \times 14 = 114.24$ Mbit/s. The processor must operate higher than 114 MHz.

d) Through the cut A-A' : (16 + 16) $34 \times 240000 \times 14 = 3.66$ Gbit/s.
Through B-B' :

$$\frac{N_m \cdot W_m}{T_{RAM}} = \frac{N_m \cdot W_m}{17 \times 10^{-9}} \Rightarrow N_m = \frac{3.66 \times 10^9 \times 17 \times 10^{-9}}{W_m}$$

We select $W_m = 12 \Rightarrow N_m \approx 5.18$

e) The memory must hold a whole 16×16 data matrix, i.e.,
$16 \times 16 \times 12 = 3072$ bits.

f) We select 8 RAMs each with 32×12-bit words. This selection yields a simple organization of the memories.

g) We have already selected $W_m = 12$. The required access frequency for the memories is: $3.65568 \times 10^9 = 8 \times 12 \times f_m$
$\Rightarrow f_m = 38$ MHz

9.14 a) The FFT has N_{tot} butterflies:

$$N_{tot} = \frac{N}{2}\log_2(N) = \frac{1024}{2}\log_2(1024) = 5120 \text{ butterflies}$$

The number of processor is: $N_p = \frac{5120 \times 1000 \times 23}{120 \times 10^6} \approx 0.981 < 1$

The processor schedule is sequential since only one processor is used.

b) See Problem 9.13.

c) In average we execute $\frac{5120}{10^{-3}}$ butterflies per second. The bit rate to the processor is: $\frac{5120}{10^{-3}} \times 23 \approx 118$ MHz

d) The number of bits per second through the cut A-A' :

$$(4 + 4) \times \frac{5120}{10^{-3}} \times 23 = 942.08 \text{ Mbit/s}$$

and through the cut B-B' :

$$\frac{N_m \cdot W_m}{T_{RAM}} = \frac{N_m \cdot W_m}{17 \times 10^{-9}} \Rightarrow N_m = \frac{942.08 \times 10^6 \times 17 \times 10^{-9}}{W_m}$$

We select $W_m = 21 \Rightarrow N_m = \frac{942.08 \times 10^6 \times 17 \times 10^{-9}}{21} \approx 0.76 < 1$

e) We need 2×1024×21-bit words ⇒ 43008 bits.

f) We select 8 RAMs each with 128×42-bit words. This selection yield a reasonable length/width ratio for the memories.

g) The memory access rate is: $942.08 \times 10^6 = 8 \times 42 \times f_m \Rightarrow$
$\Rightarrow f_m = 2.80$ MHz

9.15 a) The number of adaptors per second is:

$$N_{op} = 2 \times (8 + 4 \times 2 + 2 \times 2^2 + 2 \times 2^3) = 80$$

The sample rate is 44.1 kHz. The number of processors is:
$N_p = \frac{80 \times 44.1 \times 10^3 \times 24}{120 \times 10^6} \approx 0.706 < 1$

The processor schedule is sequential since only one processor is used.

b) See Problem 9.13.

c) We execute $80 \times 44.1 \times 10^3$ adaptor operations per second. The I/O bit rate for the processor is: $80 \times 44.1 \times 10^3 \times 24 \approx 85$ MHz

d) The number of bit/s through the cut A-A' is:

$$(2 + 2) \times 80 \times 44.1 \times 10^3 \times 24 = 338.7 \text{ Mbit/s}$$

and through the cut B-B' :

$$\frac{N_m \cdot W_m}{T_{RAM}} = \frac{N_m \cdot W_m}{17 \times 10^{-9}} \Rightarrow N_m = \frac{338.7 \times 10^6 \times 17 \times 10^{-9}}{W_m}$$

We select $W_m = 22 \Rightarrow N_m \approx 0.26 < 1$

e) The memory requirement is less than the number of outputs, i.e., less than 64 × 22-bit words or 1408 bits. An lower estimate of the memory requirement is equal to the number of delay elements, i.e., 32 × 22-bit words or 704 bits.

f) We select one RAM with at least 32 × 22-bit or jet better 64 × 22-bit words.

-Oj, ni har helt rätt Sir Dwayne. När jag knackar här börjar han surra och han är ARG!

g) We have already selected W_m = 22. The access rate for the RAM is selected so that a balanced architecture is obtained.

$338.7 \times 10^6 = 1 \times 22 \times f_m \Rightarrow f_m \approx 15$ MHz

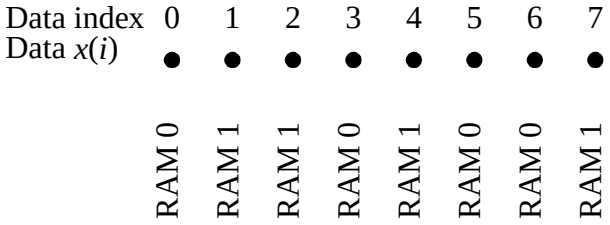
9.18 To verify the FFT architecture alternatives, we use an 8-point FFT.

Architecture 1:

•A butterfly PE shall always use data from two different RAMs.

•Butterflies from 0 to $N/4 - 1$ are assigned to PE0 and rest to PE1.

The memory assignment can be done with the exclusion graph method (See Chapter 7.11).



Memory assignment

The assignment for the PEs are done by the requirement 2.

We can verify that all data come from different RAMs, for example, at the first stage, data pairs $\{x(0), x(4)\}$, $\{x(1), x(5)\}$, $\{x(2), x(6)\}$, and $\{x(3), x(7)\}$ come from both RAMs. The verification for the second and the last stage are left to the readers.

Architecture 2:
•The variables with indices 0 to $N/2-1$ are assigned to RAM0 and the rest to RAM1.
•Butterflies 0 to $N/4-1$ are assigned to PE0 and the rest to PE1.

At the first stage, data pairs $\{x(0), x(4)\}$, $\{x(1), x(5)\}$, $\{x(2), x(6)\}$, and $\{x(3), x(7)\}$ come from both RAMs. At the second stage, data pairs $\{x(0), x(2)\}$ and $\{x(1), x(3)\}$ come from RAM0 and the rest from RAM1. At the last stage, data pairs $\{x(0), x(1)\}$ and $\{x(2), x(3)\}$ come from RAM0 and the rest from RAM1, i.e., after the first stage, we have two separated 4-point FFTs on two different PEs.

Architecture 3:
•A butterfly PE shall always use data from two different RAMs.
•A RAM is not connected to both inputs of a butterfly PE.
This architecture has less restriction than that of Architecture 1, where PEs are assigned to butterflies.

Architecture 4:
•The variables with indices 0 to $N/2-1$ are assigned to RAM0 and the rest to RAM1.
•A RAM is not connected to both inputs of a butterfly PE.
It is impossible to implement FFT with this architecture. For example, at the last stage, the data pairs $\{x(0), x(1)\}$ should be processed in a butterfly PE. According to the first restriction, $\{x(0), x(1)\}$ should be stored in RAM0, which is not allowed according to the second restiction.

9.20 The delay time and supply voltage is given by

$$t = \frac{C_{load} V_{supply}}{k(V_{supply} - V_T)^2} \quad (1)$$

where k is a factor depends on the technology and size of transistors and it is a constant for voltage scaling, the same for C_{load}.

The power consumption is given by

$$P = C_{load} {V_{supply}}^2 f_{clk} \quad (2)$$

The sample rate is 10 MHz corresponds to a clock frequency of 30 MHz.

$$T_{clk} = \frac{1}{30 \times 10^6} s = 33{,}3ns$$

Compare it with the case of 35 MHz, or clock frequency of 105 MHz

$$T_{oldclk} = \frac{1}{105 \times 10^6} s = 9{,}52ns$$

From equation (1), we can scale down the voltage from 5 V to 2.24 V, so the power consumption can be reduced to

$P_{new} = \frac{2{,}24^2 \times 30 \times 10^6}{5^2 \times 105 \times 10^6} \cdot P_{old} = 0{,}0572 \times 30 mW = 1{,}72 mW$ according to equation (2).