SOLUTIONS: SYSTEMS IDENTIFICATION

1. Solution

a) Model $\mathcal{M}_1$ can be written as

$$\mathcal{M}_1 : \widehat{\varphi}_1(v, \theta_1, \theta_2) = \Psi_1(v)^\top \theta$$

where vectors $\Psi_1(v)$ and θ are defined as follows:

$$\Psi_1(v) := \begin{bmatrix} e^{-4(v-11/5)^2} \\ e^{-4(v-12/5)^2} \end{bmatrix}, \qquad \theta := \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}$$

Analogously, Model $\mathcal{M}_2$ can be written as

$$\mathcal{M}_2 : \widehat{\varphi}_2(v, \alpha_1, \alpha_2) = \Psi_2(v)^\top \alpha$$

where vectors $\Psi_2(v)$ and α are defined as follows:

$$\Psi_2(v) := \begin{bmatrix} v \\ 1 \end{bmatrix}, \qquad \alpha := \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$$

Consider the generic set

$$\Theta = \{(v(j), m_c(j)),\ j = 1, \ldots, M\}.$$

and the above-defined function

$$\widehat{\varphi}_1(v, \theta_1, \theta_2) = \Psi_1(v)^\top \theta$$

Introduce the error variable given by

$$e_1(j) = m_c(j) - \widehat{\varphi}_1[v(j), \theta_1, \theta_2] = m_c(j) - \Psi_1[v(j)]^\top \theta,\ j = 1, \ldots, M.$$

Moreover, consider the following cost function:

$$J(\theta) = \frac{1}{M} \sum_{j=1}^{M} [e_1(j)]^2 = \frac{1}{M} \sum_{j=1}^{M} \{m_c(j) - \Psi_1[v(j)]^\top \theta\}^2.$$

The minimisation of $J(\theta)$ with respect to the unknown vector θ yields the optimal (in the least-squares sense) set of parameters $\theta_1^\circ, \theta_2^\circ$ of model $\mathcal{M}_1$. Specifically:

$$\theta^\circ = [\theta_1^\circ, \theta_2^\circ]^\top = \arg\min_\theta J(\theta)$$

provided that a unique minimum of $J(\theta)$ does exist.

Since model $\mathcal{M}_2$ takes on the same structure as $\mathcal{M}_1$, that is

$$\mathcal{M}_2 : \widehat{\varphi}_2(v, \alpha_1, \alpha_2) = \Psi_2(v)^\top \alpha,$$

by introducing the error variable given by

$$e_2(j) = m_c(j) - \widehat{\varphi}_2[v(j), \alpha_1, \alpha_2] = m_c(j) - \Psi_2[v(j)]^\top \alpha,\ j = 1, \ldots, M,$$

and the cost function

$$J(\theta) = \frac{1}{M}\sum_{j=1}^{M}[e_2(j)]^2 = \frac{1}{M}\sum_{j=1}^{M}\{m_c(j) - \Psi_2[v(j)]^\top \alpha\}^2,$$

the same solution approach gives the optimal (in the least-squares sense) set of parameters $\alpha_1^\circ, \alpha_2^\circ$ of model $\mathscr{M}_2$.

[8 Marks]

b) According to the answer to Question 1-a), the same solution approach yields the optimal (in the least-squares sense) models $\mathscr{M}_1$ and $\mathscr{M}_2$. One first considers a generic model of the form

$$\widehat{\varphi}(v, \sigma_1, \sigma_2) = \Psi(v)^\top \sigma$$

a generic error variable

$$e(j) = m_c(j) - \widehat{\varphi}(v, \sigma_1, \sigma_2) = m_c(j) - \Psi[v(j)]^\top \sigma\,,\ j = 1, \ldots, M,$$

and a generic cost function

$$J(\theta) = \frac{1}{M}\sum_{j=1}^{M}[e(j)]^2 = \frac{1}{M}\sum_{j=1}^{M}\{m_c(j) - \Psi[v(j)]^\top \sigma\}^2.$$

To compute the optimal solution $\sigma^\circ = [\sigma_1^\circ, \sigma_1^\circ]^\top$, the gradient of $J(\theta)$ with respect to the vector σ is determined:

$$\begin{aligned}
\frac{\partial}{\partial\sigma}J(\sigma) &= \frac{1}{M}\sum_{j=1}^{M}\frac{\partial}{\partial\sigma}\left\{[m_c(j) - \Psi(v(j))^\top\sigma]^2\right\} = \\
&= \frac{2}{M}\sum_{j=1}^{M}[m_c(j) - \Psi(v(j))^\top\sigma]\frac{\partial}{\partial\sigma}\left[m_c(j) - \Psi(v(j))^\top\sigma\right] = \\
&= -\frac{2}{M}\sum_{j=1}^{M}[m_c(j) - \Psi(v(j))^\top\sigma]\Psi(v(j))^\top
\end{aligned}$$

where $\Psi(v)^\top$ is a 2-dimensional row vector. Now, by imposing

$$\frac{\partial}{\partial\sigma}J(\sigma) = [0, 0],$$

after some algebra, the following linear problem is obtained:

$$\left\{\sum_{j=1}^{M}\Psi(v(j))\Psi(v(j))^\top\right\}\sigma = \sum_{j=1}^{M}m_c(j)\Psi(v(j))$$

that has a unique solution σ° if the matrix

$$\sum_{j=1}^{M}\Psi(v(j))\Psi(v(j))^\top$$

is non-singular.

Concerning model $\mathscr{M}_1$, one has:

$$\sum_{j=1}^{M}\Psi_1(v(j))\Psi_1(v(j))^\top = \begin{bmatrix} \sum_{j=1}^{M} e^{-8(v(j)-11/5)^2} & \sum_{j=1}^{M} e^{-4(v(j)-11/5)^2-4(v(j)-12/5)^2} \\ \sum_{j=1}^{M} e^{-4(v(j)-11/5)^2-4(v(j)-12/5)^2} & \sum_{j=1}^{M} e^{-8(v(j)-12/5)^2} \end{bmatrix}$$

and

$$\sum_{j=1}^{M} m_c(j)\Psi_1(v(j)) = \begin{bmatrix} \sum_{j=1}^{M} m_c(j)e^{-4(v(j)-11/5)^2} \\ \sum_{j=1}^{M} m_c(j)e^{-4(v(j)-12/5)^2} \end{bmatrix}$$

Finally, concerning model $\mathcal{M}_2$, one has:

$$\sum_{j=1}^{M} \Psi_2(v(j))\Psi_2(v(j))^\top = \begin{bmatrix} \sum_{j=1}^{M} v(j)^2 & \sum_{j=1}^{M} v(j) \\ \sum_{j=1}^{M} v(j) & M \end{bmatrix}$$

and

$$\sum_{j=1}^{M} m_c(j)\Psi_2(v(j)) = \begin{bmatrix} \sum_{j=1}^{M} m_c(j)v(j) \\ \sum_{j=1}^{M} m_c(j) \end{bmatrix}$$

[8 Marks]

c) Consider the set Θ given in the text of Question 1-c) of the exam paper:

$$\Theta = \{(2,0.03),(2.1,0.77),(2.2,0.79),(2.3,0.9),(2.4,0.84)(2.5,0.91)\}$$

and substitute the numerical values of $v(j)$ and $m_c(j)$ into the general solution obtained in the answer to Question 1-b).

Concerning model $\mathcal{M}_1$, one gets:

$$\begin{bmatrix} 4.78 & 4.42 \\ 4.42 & 4.34 \end{bmatrix} \cdot \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} 3.77 \\ 3.8 \end{bmatrix} \Longrightarrow \theta^\circ = \begin{bmatrix} \theta_1^\circ \\ \theta_2^\circ \end{bmatrix} = \begin{bmatrix} -0.36 \\ 1.24 \end{bmatrix}$$

Accordingly, the computation of the cost

$$J(\theta^\circ) = \frac{1}{6}\sum_{j=1}^{6}[e_1(j)]^2$$

defined in the answer to Question 1-a) gives $J(\theta^\circ) \simeq 0.03$.

Concerning model $\mathcal{M}_2$, one gets:

$$\begin{bmatrix} 30.55 & 13.5 \\ 13.5 & 6 \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 9.78 \\ 4.24 \end{bmatrix} \Longrightarrow \alpha^\circ = \begin{bmatrix} \alpha_1^\circ \\ \alpha_2^\circ \end{bmatrix} = \begin{bmatrix} 1.37 \\ -2.38 \end{bmatrix}$$

and the computation of the cost

$$J(\alpha^\circ) = \frac{1}{6}\sum_{j=1}^{6}[e_2(j)]^2$$

defined in the answer to Question 1-a) gives $J(\alpha^\circ) \simeq 0.04$.

Since

$$J(\alpha^\circ) > J(\theta^\circ)$$

one can conclude that the nonlinear $\mathcal{M}_1$ shows a better (in the least-squares sense) agreement with the measured data set than the linear model $\mathcal{M}_2$.

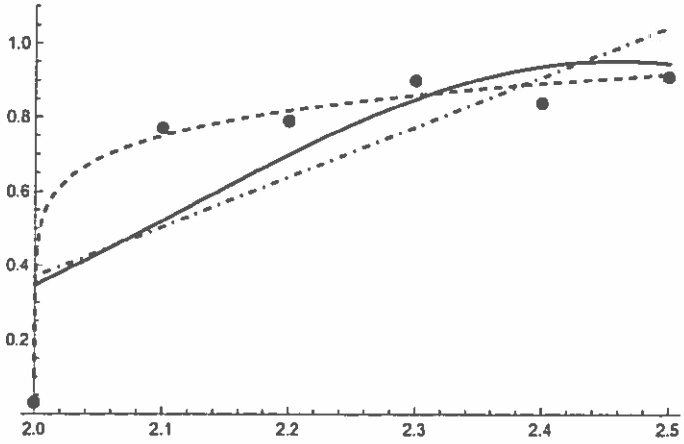

Figure 1.1

Remark for the students (not part of the solution of the exam paper).

Fig. 1.1 shows the diagrams of functions $\widehat{\varphi}_1(v,\theta_1^\circ,\theta_2^\circ)$ (continuous line) and $\widehat{\varphi}_2(v,\alpha_1^\circ,\alpha_2^\circ)$ (dash-dotted line), of the unknown model $\varphi(v)$ (dashed line) of the six noisy measurements $m_c(j), j=1,\ldots,6$ (bullet points).

[4 Marks]

2. a) For the values $a = 0$, $b \neq 0$ set in Question 2-a), by inspecting the block-scheme depicted in Fig. 2.1, we get

$$w(t) = \xi(t) + b\xi(t-1) + \eta(t) \qquad (\star)$$

The process $w(\cdot)$ is stationary because its realizations $w(t)$ are generated as a linear combination of realizations $\eta(t)$ of the stationary stochastic process $\eta(\cdot)$, and realizations $\xi(t)$ and $\xi(t-1)$ of the stationary stochastic process $\xi(\cdot)$ (recall from the text of Question 2 that $\xi(\cdot)$ and $\eta(\cdot)$ are supposed to be independent).

Moreover, from Equation $(\star)$ above, one gets:

$$\mathbb{E}[w(t)] = \mathbb{E}[\xi(t)] + b\mathbb{E}[\xi(t-1)] + \mathbb{E}[\eta(t)] = 0$$

Since $\mathbb{E}(w) = 0$, the variance $\text{var}(w)$ of process $w(\cdot)$ can be determined as follows:

$$\text{var}(w) = \gamma_w(0) = \mathbb{E}[w(t)^2] = \mathbb{E}\{[\xi(t) + b\xi(t-1) + \eta(t)]^2\} = \sigma_\eta^2 + (1+b^2)\sigma_\xi^2$$

Finally, the correlation function $\gamma_w(\tau)$ for all $\tau \geq 0$ is determined as follows:

$$\begin{aligned}\gamma_w(\tau) &= \mathbb{E}\{[\xi(t) + b\xi(t-1) + \eta(t)][\xi(t-\tau) + b\xi(t-\tau-1) + \eta(t-\tau)]\} \\ &= \mathbb{E}[\xi(t)\xi(t-\tau)] + \mathbb{E}[\eta(t)\eta(t-\tau)] + b\mathbb{E}[\xi(t)\xi(t-\tau-1)] \\ &\qquad + b\mathbb{E}[\xi(t-1)\xi(t-\tau)] + b^2\mathbb{E}[\xi(t-1)\xi(t-\tau-1)]\end{aligned}$$

Then:

$$\begin{aligned}\gamma_w(0) &= \sigma_\eta^2 + (1+b^2)\sigma_\xi^2 = \text{var}(w) \\ \gamma_w(1) &= \gamma_w(-1) = b\sigma_\xi^2 \\ \gamma_w(\tau) &= 0, \quad \forall \tau: |\tau| \geq 2\end{aligned}$$

[4 Marks]

b) Since the stochastic process $v(\cdot)$ is a moving-average one, it is stationary. Hence, in order the stationary stochastic processes $w(\cdot)$ and $v(\cdot)$ to have the same spectra, they should have the same complex spectra, that is:

$$\Gamma_w(\omega) = \Gamma_v(\omega), \forall \omega \in [-\pi, \pi] \iff \Phi_w(z) = \Phi_v(z), \forall z \in \mathbb{C}$$

Owing to Equation $(\star)$ in the answer to Question 2-a), to the independence of the stochastic processes $\xi(\cdot)$ and $\eta(\cdot)$, and using the values $a = 0$, $b = -1/2$, $\sigma_\eta^2 = 1$, $\sigma_\xi^2 = 4$ set in Question 2-b), one gets:

$$\Phi_w(z) = 4\left(1 - \frac{1}{2}z^{-1}\right)\left(1 - \frac{1}{2}z\right) + 1 = \frac{-2z^2 + 6z - 2}{z}$$

Analogously, concerning the moving-average stochastic process $v(\cdot)$, one immediately obtains

$$\Phi_v(z) = (1 + cz^{-1})(1 + cz)\sigma_e^2 = [1 + c^2 + c(z^{-1} + z)]\sigma_e^2 = \frac{cz^2 + (1+c^2)z + c}{z}$$

Hence, from the identity

$$\Phi_w(z) = \Phi_v(z), \forall z \in \mathbb{C}$$

one gets the following conditions on the parameters c and σ_e^2:

$$\begin{cases} (1+c^2)\sigma_e^2 = 6 \\ c\sigma_e^2 = -2 \end{cases}$$

Two solutions for c are obtained:

$$c_{(1)} = -\frac{1}{2}(3+\sqrt{5}) \simeq -2.62 \quad \text{and} \quad c_{(2)} = -\frac{1}{2}(3-\sqrt{5}) \simeq -0.38$$

To satisfy the condition $|c| < 1$ given in Question 2-b), one has

$$c = c_{(2)} = -\frac{1}{2}(3-\sqrt{5})$$

and consequently

$$\sigma_e^2 = -\frac{2}{c} = \frac{4}{3-\sqrt{5}} \simeq 5.24$$

The spectrum is given by

$$\Gamma_v(\omega) = \left| e^{j\omega} - \frac{1}{2}(3-\sqrt{5}) \right|^2 \frac{4}{3-\sqrt{5}}.$$

After a little algebra, we obtain

$$\Gamma(\omega) = 6 - 4\cos(\omega).$$

The behaviour of $\Gamma(\omega)$ in the interval $\omega \in [-\pi, \pi]$ is shown in Figs. 2.1.

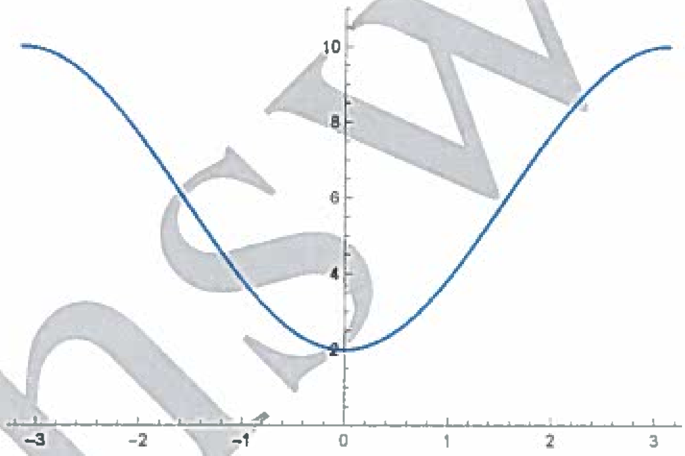

Figure 2.1 Plot of the spectrum $\Gamma(\omega) = 6 - 4\cos(\omega)$.

[7 Marks]

c) For the values $a = 1/3$, $b = -1/2$, $\sigma_\eta^2 = 1$, $\sigma_\xi^2 = 4$ set in Question 2-c), by inspecting the block-scheme depicted in Fig. 2.1, one immediately obtains

$$w(t) = \frac{1 - \frac{1}{2}z^{-1}}{1 + \frac{1}{3}z^{-1}}\xi(t) + \frac{1}{1 + \frac{1}{3}z^{-1}}\eta(t)$$

Owing to the independence of the stochastic processes $\xi(\cdot)$ and $\eta(\cdot)$ and to the linearity of the dynamic systems considered in the block-scheme depicted in Fig. 2.1, one immediately gets

$$\Gamma_w'(\omega) = \Gamma_{w;\xi}'(\omega) + \Gamma_{w;\eta}'(\omega)$$

where $\Gamma'_{w,\xi}(\omega)$ denotes the spectrum of the process $w(\cdot)$ when $\eta(\cdot) = 0$ and $\Gamma'_{w,\eta}(\omega)$ denotes the spectrum of the process $w(\cdot)$ when $\xi(\cdot) = 0$.

Hence, the spectrum $\Gamma'_{w,\xi}(\omega)$ is given by

$$\Gamma'_{w,\xi}(\omega) = \frac{\left|e^{j\omega} - \frac{1}{2}\right|^2}{\left|e^{j\omega} + \frac{1}{3}\right|^2} 4 = \cdots = \frac{45 - 36\cos(\omega)}{10 + 6\cos(\omega)}$$

whereas the spectrum $\Gamma'_{w,\eta}(\omega)$ is given by

$$\Gamma'_{w,\eta}(\omega) = \frac{1}{\left|e^{j\omega} + \frac{1}{3}\right|^2} = \cdots = \frac{9}{10 + 6\cos(\omega)}$$

which finally gives

$$\Gamma'_{w}(\omega) = \frac{54 - 36\cos(\omega)}{10 + 6\cos(\omega)}$$

The behaviour of $\Gamma'_{w}(\omega)$ in the interval $\omega \in [-\pi, \pi]$ is shown in Figs. 2.2.

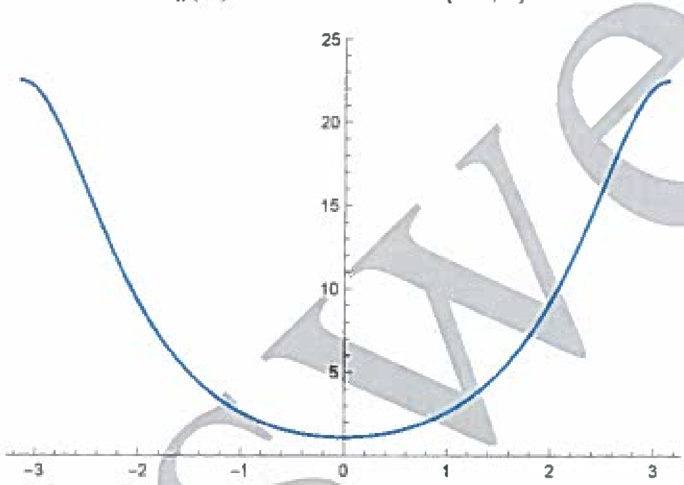

Figure 2.2 Plot of the spectrum $\Gamma'_{w}(\omega) = \frac{54-36\cos(\omega)}{10+6\cos(\omega)}$.

[6 Marks]

d) In Fig. 2.3, the diagram of $\Gamma(\omega) = 6 - 4\cos(\omega)$ obtained in the answer to Question 2-b) and the diagram of $\Gamma_w'(\omega) = \frac{54-36\cos(\omega)}{10+6\cos(\omega)}$ obtained in the answer to Question 2-c) are compared.

As can be clearly seen, the spectrum $\Gamma_w'(\omega)$ shows more power at high frequencies than the spectrum $\Gamma(\omega)$. This is due to the presence of the auto-regressive part of the model of the process considered in Question 2-c) which is characterised by a transfer function with a pole located in $-\frac{1}{3}$.

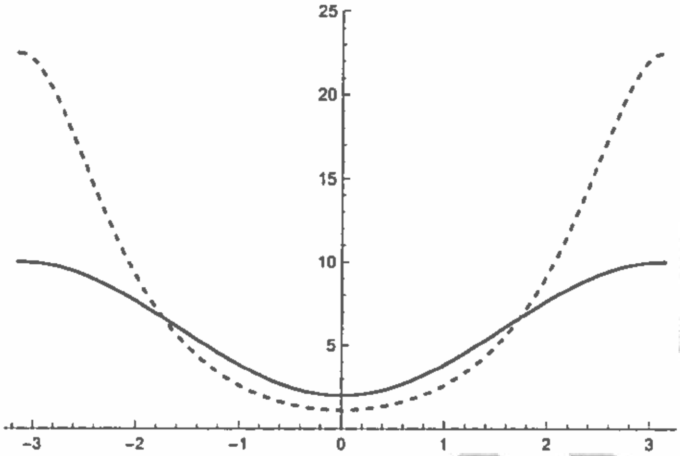

Figure 2.3 Comparison of the spectra $\Gamma(\omega) = 6 - 4\cos(\omega)$ (continuous-line) and $\Gamma_w'(\omega) = \frac{54-36\cos(\omega)}{10+6\cos(\omega)}$ (dashed line).

[3 Marks]

3. a) By inspecting the block-scheme shown in Fig. 3.1 in Question 3 of the text of the exam paper, and owing to the linearity of the system, one gets

$$v(t) = bz^{-1}\left(1 - \frac{1}{2}z^{-1}\right)\xi(t)$$

$$w(t) = (1 + cz^{-1})e(t)$$

and hence

$$y(t) = \frac{1}{1+az^{-1}}[v(t) + w(t)] = \frac{1+cz^{-1}}{1+az^{-1}}e(t) + \frac{b\left(1-\frac{1}{2}z^{-1}\right)}{1+az^{-1}}\xi(t-1)$$

thus obtaining the following difference equation

$$y(t) = -ay(t-1) + e(t) + ce(t-1) + b\xi(t-1) - \frac{b}{2}\xi(t-2)$$

Moreover, since $|a| < 1$ and $e(\cdot)$ and $\xi(\cdot)$ are stationary stochastic processes, the process $y(\cdot)$ is stationary as well.

[3 Marks]

b) From the difference equation determined in the answer to Question 3-a), one gets:

$$\begin{aligned}\gamma_{yy}(1) &= \mathbb{E}[y(t)y(t-1)] \\ &= \mathbb{E}\{[-ay(t-1) + e(t) + ce(t-1) + b\xi(t-1) - \frac{b}{2}\xi(t-2)]y(t-1)\}\end{aligned}$$

and hence, owing to the stationarity of the stochastic process $y(\cdot)$, it follows that

$$\begin{aligned}\gamma_{yy}(1) &= -a\gamma_{yy}(0) + c\mathbb{E}[e(t-1)y(t-1)] - \frac{b}{2}\mathbb{E}[\xi(t-2)y(t-1)] \\ &= -a\gamma_{yy}(0) + c\sigma_e^2 - \frac{b^2}{2}\sigma_\xi^2\end{aligned}$$

The calculation of the variance $\gamma_{yy}(0)$ of the process $y(\cdot)$ is thus needed. First, one has:

$$\mathbb{E}[y(t)] = -a\mathbb{E}[y(t-1)] + \mathbb{E}[e(t)] + c\mathbb{E}[e(t-1)] + b\mathbb{E}[\xi(t-1)] - \frac{b}{2}\mathbb{E}[\xi(t-2)]$$

and thus $\mathbb{E}[y(t)] = -a\mathbb{E}[y(t)] \Longrightarrow \mathbb{E}[y(t)] = 0$. Then:

$$\begin{aligned}\gamma_{yy}(0) = \mathbb{E}[y(t)^2] &= \mathbb{E}\{[-ay(t-1) + e(t) + ce(t-1) + b\xi(t-1) - \frac{b}{2}\xi(t-2)]^2\} \\ &= a^2\gamma_{yy}(0) + (1 + c^2 - 2ac)\sigma_e^2 + \left(\frac{5}{4} + a\right)b^2\sigma_\xi^2\end{aligned}$$

thus getting

$$\gamma_{yy}(0) = \frac{1 + c^2 - 2ac}{1-a^2}\sigma_e^2 + \frac{\frac{5}{4} + a}{1-a^2}b^2\sigma_\xi^2$$

and hence, after some algebra:

$$\gamma_{yy}(1) = \frac{(c-a)(1-ac)}{1-a^2}\sigma_e^2 - \frac{\frac{1}{2}(a+2)(a+\frac{1}{2})b^2}{1-a^2}\sigma_\xi^2$$

Inspecting the block-scheme shown in Fig. 3.1 in Question 3 of the text of the exam paper, one gets

$$u(t) = b\xi(t) - \frac{b}{2}\xi(t-1)$$

Hence

$$\begin{aligned}
&\gamma_{yu}(0) = \mathbb{E}[y(t)u(t)] \\
&= \mathbb{E}\{[-ay(t-1)+e(t)+ce(t-1)+b\xi(t-1)-\frac{b}{2}\xi(t-2)][b\xi(t)-\frac{b}{2}\xi(t-1)]\} \\
&\qquad = -\frac{b^2}{2}\sigma_\xi^2
\end{aligned}$$

Finally,

$$\begin{aligned}
&\gamma_{yu}(1) = \mathbb{E}[y(t)u(t-1)] \\
&= \mathbb{E}\{[-ay(t-1)+e(t)+ce(t-1)+b\xi(t-1)-\frac{b}{2}\xi(t-2)][b\xi(t-1)-\frac{b}{2}\xi(t-2)]\} \\
&\qquad = \left(\frac{5}{4}+\frac{a}{2}\right)b^2\sigma_\xi^2
\end{aligned}$$

[7 Marks]

c) Since the stationary $u(\cdot)$ is supposed to be measurable, the model in prediction form

$$\widehat{\mathcal{M}}(\theta): \hat{y}(t|t-1) = \alpha y(t-1) + \beta u(t-1), \text{ where } \theta := [\alpha, \beta]^\top$$

is a legitimate predictor because to compute the prediction $\hat{y}(t|t-1)$ of $y(t)$ it uses the values of $y(t-1)$ and $u(t-1)$ that are available. The vector $\overline{\theta}$ can be obtained by minimising the cost function

$$\begin{aligned}
J(\theta) &= \mathbb{E}\{[y(t)-\hat{y}(t|t-1)]^2\} \\
&= \gamma_{yy}(0) + \mathbb{E}\{[\alpha y(t-1)+\beta u(t-1)]^2\} - 2\mathbb{E}\{y(t)[\alpha y(t-1)+\beta u(t-1)]\} \\
&\qquad = (1+\alpha^2)\gamma_{yy}(0) + \beta^2\gamma_{uu}(0) + 2\alpha\beta\gamma_{yu}(0) - 2\alpha\gamma_{yy}(1) - 2\beta\gamma_{yu}(1)
\end{aligned}$$

Now

$$\frac{\partial J}{\partial \theta} = [0,0] \iff \begin{cases} \dfrac{\partial J}{\partial \alpha} = 0 \\[2ex] \dfrac{\partial J}{\partial \beta} = 0 \end{cases}$$

Hence, the following linear system of equations has to be solved with respect to α and β:

$$\begin{cases} \alpha\gamma_{yy}(0) + \beta\gamma_{yu}(0) = \gamma_{yy}(1) \\[2ex] \alpha\gamma_{yu}(0) + \beta\gamma_{uu}(0) = \gamma_{yu}(1) \end{cases}$$

where $\gamma_{yy}(0)$, $\gamma_{yu}(0)$, $\gamma_{yy}(1)$, and $\gamma_{yu}(1)$ have been computed in the answer to Question 3-b), whereas $\gamma_{uu}(0)$ is given by

$$\gamma_{uu}(0) = \mathbb{E}[u(t)^2] = \mathbb{E}\{[b\xi(t) - \frac{b}{2}\xi(t-1)]^2\} = \frac{5}{4}b^2\sigma_\xi^2$$

Thus, the solution $\overline{\theta} = [\overline{\alpha}, \overline{\beta}]^\top$ is given by

$$\overline{\theta} = \begin{bmatrix} \overline{\alpha} \\ \overline{\beta} \end{bmatrix} = \frac{1}{\gamma_{yy}(0)\gamma_{uu}(0) - \gamma_{yu}(0)^2} \begin{bmatrix} \gamma_{uu}(0) & -\gamma_{yu}(0) \\ -\gamma_{yu}(0) & \gamma_{yy}(0) \end{bmatrix} \begin{bmatrix} \gamma_{yy}(1) \\ \gamma_{yu}(1) \end{bmatrix}$$

[7 Marks]

d) Substituting the values $a = 1/2$, $b = 1/3$, $c = 1/4$, $\sigma_e^2 = 1$, $\sigma_\xi^2 = 9$ into the expressions for $\gamma_{yy}(0)$, $\gamma_{yu}(0)$, $\gamma_{yy}(1)$, $\gamma_{yu}(1)$, and $\gamma_{uu}(0)$, one gets:

$$\gamma_{yy}(0) = \frac{41}{12};\ \gamma_{yu}(0) = -\frac{1}{2};\ \gamma_{yy}(1) = -\frac{59}{24};\ \gamma_{yu}(1) = \frac{3}{2};\ \gamma_{uu}(0) = \frac{5}{4};$$

Moreover, using the above numerical values in the general solution for $\overline{\theta}$ provided in the answer to Question 3-c), one obtains

$$\overline{\theta} = \begin{bmatrix} \overline{\alpha} \\ \overline{\beta} \end{bmatrix} = \begin{bmatrix} -\frac{223}{386} \\ \frac{187}{193} \end{bmatrix} \simeq \begin{bmatrix} -0.58 \\ 0.97 \end{bmatrix}$$

Finally, the variance of the prediction error can be obtained by substituting the value of $\overline{\theta}$ and the numerical values of $\gamma_{yy}(0)$, $\gamma_{yu}(0)$, $\gamma_{yy}(1)$, $\gamma_{yu}(1)$, and $\gamma_{uu}(0)$ computed above into the expression of

$$\begin{aligned} J(\theta) &= \mathbb{E}\{[y(t) - \hat{y}(t|t-1)]^2 \\ &= \gamma_{yy}(0) + \mathbb{E}\{[\alpha y(t-1) + \beta u(t-1)]^2\} - 2\mathbb{E}\{y(t)[\alpha y(t-1) + \beta u(t-1)]\} \\ &= (1+\alpha^2)\gamma_{yy}(0) + \beta^2\gamma_{uu}(0) + 2\alpha\beta\gamma_{yu}(0) - 2\alpha\gamma_{yy}(1) - 2\beta\gamma_{yu}(1) \end{aligned}$$

hence obtaining

$$\text{var}[\varepsilon(t)] = \text{var}[y(t) - \hat{y}(t|t-1)] = \frac{1677}{3088} \simeq 0.54$$

As should be expected, the variance of the prediction error is smaller than the variance of the output process.

[3 Marks]

4. Solution

a) By inspecting the block-scheme depicted in Fig. 4.1 in the text of the exam paper, it follows that

$$w(t) = \frac{1}{2}\xi(t-3) - 5\xi(t-4)$$

and hence the process $w(\cdot)$ is stationary because it is a linear combination of stationary stochastic processes.

Moreover

$$v(t) = \frac{1}{1 - z^{-1} + \frac{1}{2}z^{-2}} w(t) = \frac{z^2}{z^2 - z + \frac{1}{2}} w(t)$$

The roots of the denominator $z^2 - z + \frac{1}{2}$ are located inside the unit circle, hence the stochastic process $v(\cdot)$ is stationary as well.

Since $\mathbb{E}(w) = 0$, one immediately gets $\mathbb{E}(v) = 0$. Therefore:

$$\text{var}(v) = \mathbb{E}[v(t)^2] = \gamma_v(0)$$

where $\gamma_v(\tau)$ is the correlation function of process $v(\cdot)$. It is convenient to set

$$\eta(t) := \xi(t-3), \quad \eta(\cdot) \sim WN(0,4)$$

Then:

$$\mathbb{E}[v(t)^2] = \mathbb{E}\left\{[v(t-1) - \frac{1}{2}v(t-2) + \frac{1}{2}\eta(t) - 5\eta(t-1)]^2\right\}$$

After some algebra, it follows that

$$\gamma_v(0) = \frac{5}{4}\gamma_v(0) - \gamma_v(1) + 81 \Longrightarrow \gamma_v(0) = 4\gamma_v(1) - 324$$

Moreover

$$\gamma_v(1) = \mathbb{E}\left\{[v(t-1) - \frac{1}{2}v(t-2) + \frac{1}{2}\eta(t) - 5\eta(t-1)]v(t-1)\right\} = \gamma_v(0) - \frac{1}{2}\gamma_v(1) - 10$$

Thus

$$\gamma_v(1) = \frac{2}{3}\gamma_v(0) - \frac{20}{3}$$

hence obtaining:

$$\text{var}(v) = \gamma_v(0) = \frac{1052}{5} = 210.4$$

[5 Marks]

b) By inspecting the block-scheme depicted in Fig. 4.1 in the text of the exam paper, it follows that

$$v(t) = G(z)\xi(t) \quad \text{with} \quad G(z) = \frac{\frac{1}{2}z^{-3} - 5z^{-4}}{1 - z^{-1} + \frac{1}{2}z^{-2}}$$

Clearly $G(z)$ is not expressed in canonical form.

To obtain the equivalent canonical form $\widehat{G}(z)$, it is convenient to resort again to the definition already introduced in the answer to Question 4-a), that is:

$$\eta(t) := \xi(t-3), \quad \eta(\cdot) \sim WN(0,4)$$

Then:

$$v(t) = \frac{\frac{1}{2} - 5z^{-1}}{1 - z^{-1} + \frac{1}{2}z^{-2}}\eta(t) = \frac{1}{2}\frac{z^2 - 10z}{z^2 - z + \frac{1}{2}}\eta(t)$$

By letting

$$e(t) := \xi(t-3), \quad e(\cdot) \sim WN(0,1)$$

it follows that

$$v(t) = \frac{z^2 - 10z}{z^2 - z + \frac{1}{2}}e(t)$$

However, the transfer function $\frac{z^2-10}{z^2-z+\frac{1}{2}}$ is still not expressed in canonical form since one of the zeros lies outside the unit circle. The spectral canonical form is then given by:

$$v(t) = \widehat{G}(z)r(t) = \frac{z^2 - \frac{1}{10}z}{z^2 - z + \frac{1}{2}}r(t) \quad \text{with} \quad r \sim WN(0,100)$$

In fact, it is easy to see that

$$\frac{z^2 - 10z}{z^2 - z + \frac{1}{2}}e(t) \quad \text{and} \quad \frac{z^2 - \frac{1}{10}z}{z^2 - z + \frac{1}{2}}r(t)$$

have the same complex spectra.

[4 Marks]

c) Consider the model of the process $v(\cdot)$ in canonical form determined in the answer to Question 4-b), that is:

$$v(t) = \frac{z^2 - \frac{1}{10}z}{z^2 - z + \frac{1}{2}}r(t) = \frac{1 - \frac{1}{10}z^{-1}}{1 - z^{-1} + \frac{1}{2}z^{-2}}r(t) \quad \text{with} \quad r \sim WN(0,100)$$

Thus:

$$A(z)v(t) = C(z)r(t)$$

where

$$A(z) = 1 - z^{-1} + \frac{1}{2}z^{-2}, \quad C(z) = 1 - \frac{1}{10}z^{-1}.$$

By carrying out one iteration of polynomial division of $C(z)$ by $A(z)$ one gets:

$$\begin{array}{llllll}
1 & -\frac{1}{10}z^{-1} & & 1 & -z^{-1} & +\frac{1}{2}z^{-2} \\
-1 & z^{-1} & -\frac{1}{2}z^{-2} & 1 & & \\
// & \frac{9}{10}z^{-1} & -\frac{1}{2}z^{-2} & & &
\end{array}$$

Then:

$$\widehat{W}(z) = \frac{C(z)}{A(z)} = 1 + z^{-1}\frac{\frac{9}{10} - \frac{1}{2}z^{-1}}{1 - z^{-1} + \frac{1}{2}z^{-2}}.$$

and hence the transfer function of the one-step ahead predictor of $v(t+1)$ from the white noise process $r(t)$ is given by

$$\widehat{W}_1(z) = \frac{\frac{9}{10} - \frac{1}{2}z^{-1}}{1 - z^{-1} + \frac{1}{2}z^{-2}}$$

Moreover, the transfer function of the one-step ahead predictor of $v(t+1)$ from the past data $v(t)$ is

$$W_1(z) = \frac{\frac{9}{10} - \frac{1}{2}z^{-1}}{1 - \frac{1}{10}z^{-1}}\,.$$

and, accordingly, the difference equation implementing the one-step ahead predictor of $v(t+1)$ from the data $v(t)$ is

$$\hat{v}(t+1|t) = \frac{1}{10}\hat{v}(t|t-1) + \frac{9}{10}v(t) - \frac{1}{2}v(t-1)\,.$$

To compute $\text{var}[\varepsilon_1(t)]$, from

$$\widehat{W}(z) = \frac{C(z)}{A(z)} = 1 + z^{-1}\frac{\frac{9}{10} - \frac{1}{2}z^{-1}}{1 - z^{-1} + \frac{1}{2}z^{-2}}\,,$$

it follows that

$$\text{var}[\varepsilon_1(t)] = \text{var}[v(t+1) - \hat{v}(t+1|t)] = 1 \cdot \text{var}[r(t+1)] = 100\,.$$

[4 Marks]

d) Using the expression

$$A(z)v(t) = C(z)r(t)$$

with

$$A(z) = 1 - z^{-1} + \frac{1}{2}z^{-2}, \quad C(z) = 1 - \frac{1}{10}z^{-1}\,.$$

that has been used in the answer to Question 4-c), another iteration of polynomial division of $C(z)$ by $A(z)$ is carried out in the following:

$$\begin{array}{llll|lll}
1 & -\frac{1}{10}z^{-1} & & & 1 & -z^{-1} & +\frac{1}{2}z^{-2} \\
-1 & z^{-1} & -\frac{1}{2}z^{-2} & & 1 & \frac{9}{10}z^{-1} & \\
// & \frac{9}{10}z^{-1} & -\frac{1}{2}z^{-2} & & & & \\
// & -\frac{9}{10}z^{-1} & \frac{9}{10}z^{-2} & -\frac{9}{20}z^{-3} & & & \\
// & // & \frac{2}{5}z^{-2} & -\frac{9}{20}z^{-3} & & &
\end{array}$$

and thus

$$\widehat{W}(z) = \frac{C(z)}{A(z)} = 1 + \frac{9}{10}z^{-1} + z^{-2}\frac{\frac{2}{5} - \frac{9}{20}z^{-1}}{1 - z^{-1} + \frac{1}{2}z^{-2}}$$

Therefore, the transfer function of the two-steps ahead predictor of $v(t+2)$ from the white noise process $r(t)$ is given by

$$\widehat{W}_2(z) = \frac{\frac{2}{5} - \frac{9}{20}z^{-1}}{1 - z^{-1} + \frac{1}{2}z^{-2}}$$

and thus the transfer function of the two-steps ahead predictor of $v(t+2)$ from the past data $v(t)$ is

$$W_2(z) = \frac{\frac{2}{5} - \frac{9}{20}z^{-1}}{1 - \frac{1}{10}z^{-1}}$$

Finally, the difference equation implementing the two-step ahead predictor of $v(t+2)$ from the data $v(t)$ is

$$\hat{v}(t+2|t) = \frac{1}{10}\hat{v}(t+1|t-1) + \frac{2}{5}v(t) - \frac{9}{20}v(t-1).$$

To compute $\mathrm{var}[\varepsilon_2(t)]$, from

$$\widehat{W}(z) = \frac{C(z)}{A(z)} = 1 + \frac{9}{10}z^{-1} + z^{-2}\frac{\frac{2}{5} - \frac{9}{20}z^{-1}}{1 - z^{-1} + \frac{1}{2}z^{-2}}$$

it follows that

$$\mathrm{var}[\varepsilon_2(t)] = \mathrm{var}[v(t+2) - \hat{v}(t+2|t)] =$$
$$= 1 \cdot \mathrm{var}[r(t+2)] + \left(\frac{9}{10}\right)^2 \cdot \mathrm{var}[r(t+1)] = \left(1 + \frac{81}{100}\right) \cdot 100 = 181.$$

[4 Marks]

e) The comparison between $\mathrm{var}[\varepsilon_1(t)]$ and $\mathrm{var}[\varepsilon_2(t)]$ gives

$$\mathrm{var}[\varepsilon_2(t)] = 181 > 100 = \mathrm{var}[\varepsilon_1(t)]$$

This confirms that the variance of the prediction error $\mathrm{var}[\varepsilon_r(t)]$ increases with the number r of steps-ahead of the prediction that is computed. Moreover (see the answer to Question 4-a))

$$\mathrm{var}[\varepsilon_r(t)] < \mathrm{var}[v(t)] = 210.4, \forall r \geq 1$$

[3 Marks]