1.

a. The coherence spectra related to H_{11} and H_{22} are rather good up to 1 Hz. Hence, the estimated models cannot be trusted up to more than 1 Hz. However, also in this frequency interval, we notice that the accuracy of the models H_{12} and H_{21} are quite poor, and, in particular, much lower than the one for H_{11} and H_{22}.

b. From the process set-up, we see that u_1 has its largest influence on y_1 and u_2 on y_2. Hence, the best pairing is $u_1 = C_1 y_1$ and $u_2 = C_2 y_2$.

2.

$$y_k = \underbrace{b_1 c_0}_{\theta_1} + \underbrace{b_1 c_1}_{\theta_2} u_{k-1} + \underbrace{b_1 c_2}_{\theta_3} u_{k-1}^2 - \underbrace{a_1}_{\theta_4} y_{k-1} - \underbrace{a_2}_{\theta_5} y_{k-2}$$

$$\begin{pmatrix} y_3 \\ y_4 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} 1 & u_2 & u_2^2 & y_2 & y_1 \\ 1 & u_3 & u_3^2 & y_3 & y_3 \\ \vdots & \vdots & \vdots & \vdots & \\ 1 & u_{N-1} & u_{N-1}^2 & y_{N-1} & y_{N-2} \end{pmatrix} \theta$$

Using the LMS approach, we find

$$\hat{\theta} = (\Phi^T \Phi)^{-1} \Phi^T Y.$$

From $\hat{\theta}$, $\hat{a}_1$ and $\hat{a}_2$ are readily calculated. To obtain other parameters, we make use of the assumption on the static gain, i.e,

$$\frac{b_1}{1 + a_1 + a_2} = 1 \Rightarrow \hat{b}_1 = 1 + \hat{a}_1 + \hat{a}_2$$

So, we can find all parameters as below

$$\hat{b}_1 = 1 + \hat{\theta}_4 + \hat{\theta}_5, \qquad \hat{a}_1 = \hat{\theta}_4, \qquad \hat{a}_2 = \hat{\theta}_5,$$
$$\hat{c}_0 = \frac{\hat{\theta}_1}{1 + \hat{\theta}_4 + \hat{\theta}_5}, \qquad \hat{c}_1 = \frac{\hat{\theta}_2}{1 + \hat{\theta}_4 + \hat{\theta}_5}, \qquad \hat{c}_2 = \frac{\hat{\theta}_3}{1 + \hat{\theta}_4 + \hat{\theta}_5}$$

3.

a. The recursive least-squares algorithm is

$$\begin{aligned} \hat{\theta}_k &= \hat{\theta}_{k-1} + P_k \phi_k \epsilon_k \\ \epsilon_k &= y_k - \phi_k^T \hat{\theta}_{k-1} \\ P_k &= P_{k-1} - \frac{P_{k-1} \phi_k \phi_k^T P_{k-1}}{1 + \phi_k^T P_{k-1} \phi_k} \end{aligned}$$

Introduce

$$\begin{aligned} \phi_k &= 1 \\ P_k &= (\sum_{i=1}^{k} \phi_i \phi_i^T)^{-1} = \frac{1}{k} \end{aligned}$$

where, the parameter variance has been calculated using the standard definition. Alternatively, $P_1 = 1$ can be calculated and used in the recursive equation to obtain the same result.

Finally,

$$\hat{\theta}_k = \hat{\theta}_{k-1} + \frac{1}{k}(y_k - \hat{\theta}_{k-1}) = \left(1 - \frac{1}{k}\right)\hat{\theta}_{k-1} + \frac{1}{k}y_k$$

b. The regression model is

$$\begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \theta + \begin{pmatrix} e_1 \\ \vdots \\ e_N \end{pmatrix}$$

with the least-squares estimate

$$\hat{\theta}_k = (\Phi^T\Phi)^{-1}\Phi^T Y \Rightarrow \hat{\theta}_k = \frac{1}{k}\sum_{k=1}^{N} y_k.$$

By dividing the summation into two parts, we get

$$\hat{\theta}_k = \frac{1}{k}\left(\sum_{k=1}^{N-1} y_k + y_k\right) = \frac{1}{k}\left((k-1)\hat{\theta}_{k-1} + y_k\right) = \left(1 - \frac{1}{k}\right)\hat{\theta}_{k-1} + \frac{1}{k}y_k.$$

4. The transfer function is

a.

$$H(z) = \frac{z-1}{z^2 - 1.79z + 0.792}$$

The controllable canonical form is given by

$$x_{k+1} = \begin{pmatrix} 1.79 & -0.792 \\ 1 & 0 \end{pmatrix} x_k + \begin{pmatrix} 1 \\ 0 \end{pmatrix} u_k$$

$$y_k = \begin{pmatrix} 1 & -1 \end{pmatrix} x_k$$

b. First, we calculate the observability Gramian Q by solving the Lyapunov equation.

$$\Phi^T Q \Phi - Q + C^T C = 0$$

$$2.204Q_{11} + 3.58Q_{12} + Q_{22} + 1.0 = 0$$
$$-1.418Q_{11} - 1.792Q_{12} - 1.0 = 0$$
$$0.627Q_{11} - Q_{22} + 1.0 = 0$$

with the solution

$$Q_{11} = 2.6844, Q_{12} = -2.6817, Q_{22} = 2.6838$$

$$\Sigma = Q_z = T^{-T}QT^{-1} = \begin{pmatrix} 2.6837 & 0 \\ 0 & 2.4045 \end{pmatrix}$$

Since the elements of matrix Σ have the same order of magnitude, it is not advisable to reduce the order of the model.

c.

$$z_{k+1} = \begin{pmatrix} 0.791 & 0.0423 \\ -0.0423 & 0.999 \end{pmatrix} z_k + \begin{pmatrix} -1 \\ 0.0118 \end{pmatrix} u_k$$
$$y_k = \begin{pmatrix} -1 & -0.0118 \end{pmatrix} z(k)$$

And the reduced model is

$$z^1_{k+1} \approx (0.791 + \frac{0.0423}{1-0.999}(-0.0423))z^1_k + (-1 + \frac{0.0423}{1-0.999}0.0118)u_k$$
$$= -0.998 z^1_k - 0.5u_k$$
$$y_k \approx (-1 + \frac{-0.0118}{1-0.999}(-0.0423))z^1_k + \frac{-0.0118}{1-0.999}(-1)u_k = -0.5z^1_k - 0.139u_k$$

Therefore,

$$H_r(z) = \frac{0.1113z - 0.139}{z + 0.998}$$

Accordingly, the reduced system is not simply resulted from cancelling the pole and the zero.

5.

a. Since e_k is white noise we know that

$$f_e(e_2, e_3, \ldots, e_N) = f_e(e_2) f_e(e_3) \cdots f_e(e_N)$$

Writing the residuals ϵ_k as a function of $\bar{\theta}$

$$\epsilon(\bar{\theta}) = y_k - \varphi_k^T \bar{\theta}$$

we get the likelihood function

$$L(\bar{\theta}) = \prod_{k=2}^{N} f_e(\epsilon(\bar{\theta})) = \alpha^N e^{-\beta \sum_{k=2}^{N} |y_k - \varphi_k^T \bar{\theta}|^p}$$

The optimization problem to be solved to obtain the ML estimate, $\hat{\theta}$ is

$$\hat{\theta} = \arg\max_{\bar{\theta}} L(\bar{\theta})$$

This problem might be simplified by taking the logarithm of $L(\bar{\theta})$.

b. An outlier is a point in the regression data, r for which the numerical value of $|y_k - \varphi_k^T \bar{\theta}|$ is drastically different compared with the rest. If p is large this one term might dominate the rest and in the limiting case when p goes to infinity, the optimization problem solved is

$$\max_{\bar{\theta}} |y_r - \varphi_r^T \bar{\theta}|$$

since the other terms disappear.

6. Using standard trigonometric formulas rewrite

$$y(t) = c\sin(\omega t + \phi) + e(t) = \sin(\omega t)c\cos(\phi) + \sin(\phi)c\cos(\omega t) + e(t).$$

Put

$$\theta = \begin{pmatrix} c\cos(\phi) & c\sin(\phi) \end{pmatrix}$$

and form the equation system

$$\underbrace{\begin{pmatrix} y(t_1) \\ \vdots \\ y(t_N) \end{pmatrix}}_{Y} = \underbrace{\begin{pmatrix} \sin(\omega t_1) & \cos(\omega t_1) \\ \vdots & \vdots \\ \sin(\omega t_N) & \cos(\omega t_N) \end{pmatrix}}_{A} \theta + \underbrace{\begin{pmatrix} e(t_1) \\ \vdots \\ e(t_N) \end{pmatrix}}_{E}.$$

The least-squares estimate of θ is given by

$$\begin{pmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{pmatrix} = \left(A^T A\right)^{-1} A^T Y.$$

Estimates of c and ϕ can then be found by

$$\hat{c} = (\hat{\theta}_1)^2 + (\hat{\theta}_2)^2$$

and

$$\hat{\phi} = \begin{cases} \arctan(\frac{\hat{\theta}_2}{\hat{\theta}_1}), & \hat{\theta}_1 \geq 0,\ \hat{\theta}_2 \geq 0 \\ \arctan(\frac{\hat{\theta}_2}{\hat{\theta}_1}) + \pi, & \hat{\theta}_1 \leq 0,\ \hat{\theta}_2 \geq 0 \\ \arctan(\frac{\hat{\theta}_2}{\hat{\theta}_1}) - \pi, & \hat{\theta}_1 \leq 0,\ \hat{\theta}_2 \leq 0 \end{cases}$$

7. The best unbiased linear estimator is given by the least-squares estimate. Ergodicity gives

$$\hat{b} = \frac{\sum_{i=0}^{N} u_i y_{i+1}}{\sum_{i=0}^{N} u_i^2} \rightarrow \frac{R_{yu}(1)}{R_{uu}(0)} = \frac{R_{uu}(0) + R_{uu}(1)}{R_{uu}(0)} = 1 + \frac{R_{uu}(1)}{R_{uu}(0)}.$$

Thus we get the following estimates of b for the different input signals:

a. $R_u(k) = c^2$ so $\hat{b} = 2$

b. $R_u(k) = (-1)^k$ so $\hat{b} = 0$

c. $R_u(k) = \delta(k)\sigma^2$ so $\hat{b} = 1$

8. First we note that $\sigma_r^2 = 0$ implies $u_k = -Ky_k$ and then note that the task can be solved in several ways. Here we present two alternatives, using spectrum analysis and least-squares estimation.

- Spectrum analysis: we note that the system can be written on the form

$$y_k = H(z)u_{k-1} + e_k$$

where $H(z)$ is the pulse-transfer function of the system. As $u_k = -Ky_k$ this gives us the following closed-loop dynamics from e_k to y_k:

$$\begin{aligned} y_k &= \frac{1}{1+H(z)K}e_k \\ u_k &= -\frac{K}{1+H(z)K}e_k \end{aligned}$$

Calculating the cross-spectrum $S_{yu}(i\omega)$ and the autospectrum $S_{uu}(i\omega)$ gives

$$\begin{aligned} S_{yu}(i\omega) &= -\frac{K}{|1+HK|^2}S_{ee}(i\omega) \\ S_{uu}(i\omega) &= \frac{K^2}{|1+HK|^2}S_{ee}(i\omega) \end{aligned}$$

The transfer function estimate $\widehat{H} = S_{yu}(i\omega)/S_{uu}(i\omega)$ then gives

$$\widehat{H} = S_{yu}(i\omega)/S_{uu}(i\omega) = -\frac{1}{K}$$

which shows that in lack of a persistently exciting r_k, the estimation of the process model fails.

- Least-squares: one natural way of trying to solve this estimation problem is to write it on the form

$$y_k = \begin{pmatrix} -y_{k-1} & u_{k-1} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} + e_k$$

Forming the regressor matrix based on N observations, however, gives the results

$$\Phi = \begin{pmatrix} -y_1 & u_1 \\ \vdots & \vdots \\ -y_{N-1} & u_{N-1} \end{pmatrix} = \begin{pmatrix} -y_1 & -Ky_1 \\ \vdots & \vdots \\ -y_{N-1} & -Ky_{N-1} \end{pmatrix}$$

which in turn gives

$$\Phi^T\Phi = \begin{pmatrix} \sum_{k=2}^{N-1} y_{k-1}^2 & K\sum_{k=2}^{N-1} y_{k-1}^2 \\ K\sum_{k=2}^{N-1} y_{k-1}^2 & K^2\sum_{k=2}^{N-1} y_{k-1}^2 \end{pmatrix}$$

As rank $(\Phi^T\Phi) = 1$ there will be no unique solution of the least-squares problem, and hence we can not estimate the process model.

9.

a. One way is to verify the factorization property by direct substitution of Markov parameters $h_k = CA^{k-1}B$ into the Hankel matrix.

$$
\begin{aligned}
\mathcal{H}_{r,s}^{(k)} &= \begin{pmatrix} h_{k+1} & h_{k+2} & \cdots & h_{k+s} \\ h_{k+2} & h_{k+3} & \cdots & h_{k+s+1} \\ \vdots & \vdots & \ddots & \vdots \\ h_{k+r} & h_{k+r+1} & \cdots & h_{k+r+s-1} \end{pmatrix} \\
&= \begin{pmatrix} CA^kB & CA^{k+1}B & \cdots & CA^{k+s-1}B \\ CA^{k+1}B & CA^{k+2}B & \cdots & CA^{k+s}B \\ \vdots & \vdots & \ddots & \vdots \\ CA^{k+r-1}B & CA^{k+r}B & \cdots & CA^{k+r+s+2}B \end{pmatrix} \\
&= \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{r-1} \end{pmatrix} A^k \begin{pmatrix} B & AB & \dots & A^{s-1}B \end{pmatrix}
\end{aligned}
$$

b. Using a numerical factorization such as the singular value decomposition it is possible to find estimates of the extended observability and controllability matrices. In turn, this information can be used to determine a state-space realization $\{A, B, C\}$. In the factorization above, the matrix

$$
\mathcal{O}_r = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{r-1} \end{pmatrix}
$$

is the extended observability matrix and

$$
\mathcal{C}_s = \begin{pmatrix} B & AB & \dots & A^{s-1}B \end{pmatrix}
$$

the extended controllability matrix.

For $k = 0$, the factorization is then:

$$
\begin{aligned}
H_{r,s}^{(0)} &= \mathcal{O}_r \cdot \mathcal{C}_s \\
&= U \Sigma V^T \\
&= U \Sigma^{1/2} \Sigma^{1/2} V^T
\end{aligned}
$$

Where the second inequality is obtained through singular value decomposition. We then have:

$$
\begin{aligned}
\mathcal{O}_r &= U \cdot \Sigma^{1/2} \\
\Rightarrow \mathcal{O}_r^\dagger &= \Sigma^{-1/2} U^T
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{C}_s &= \Sigma^{1/2} \cdot V^T \\
\Rightarrow \mathcal{C}_s^\dagger &= V^T \Sigma^{-1/2}
\end{aligned}
$$

The dagger sign on e.g. $\mathscr{O}_r^\dagger$ denotes the pseudo inverse.

The state space matrices are then, inserting the expressions for the pseudoinverse for the extended observability and controllability in for example $A = \mathscr{O}_r^\dagger H_{r,s}^{(1)} \mathscr{C}_s^\dagger$ and using the expressions above:

$$\begin{aligned}
\hat{A}_n &= \mathscr{O}_r^\dagger H_{r,s}^{(1)} \mathscr{C}_s^\dagger \\
&= \Sigma_n^{-1/2} U_n^T H_{r,s}^{(1)} V_n \Sigma_n^{-1/2} \\
\hat{B}_n &= \mathscr{C}_s \cdot \begin{bmatrix} I_{m\times m} & 0_{m\times(s-1)m} \end{bmatrix}^T \\
\hat{C}_n &= \left[\begin{pmatrix} I_{p\times p} \\ 0^T_{p\times(s-1)p} \end{pmatrix} \right]^T \mathscr{O}_r
\end{aligned}$$