Problem 1

Part 1 We work with the model

$$y(k) = \frac{B(z)}{A(z)}u(k) + \frac{H_1(z)}{A(z)}e(k).$$

(a) $h_1 = 0$: this is a standard ARX model with one-step ahead predictor

$$\hat{y}(k|k-1) = (1-A)y + Bu.$$

(b) Here it is convenient to work with the scaled problem $y_C(k) = C^{-1}y(k)$ where the one-step ahead predictor is

$$\hat{y}_C(k|k-1) = (1-A)y_C + Bu_C.$$

$C(z) = 1 + z^{-1}$ is not stably invertible, so I am not sure this is really justified.

Part 2 The model is

$$y(k) = \underbrace{\frac{F(z)}{D(z)}}_{\doteq G_2(z)} u(k) + \frac{H_1(z)}{D(z)}e(k) + \frac{H_2(z)}{D(z)}w(k). \tag{1}$$

The estimate $\hat{\theta}$ can be found using the one-step ahead predictor for the modified variables $y_H \doteq H_1^{-1}y$ and can be computed in the usual way: multiply by D and rearrange it as

$$y = (1-D)y + Fu + H_1 e + H_2 w \tag{2}$$

multiply by H_1^{-1} (provided H_1 is stably invertible, which $H_1 = 1 + z^{-1}$ is not)

$$y_H = (1-D)y_H + Fu_H + e + H_2 w_H.$$

The one-step ahead predictor for y_H is

$$\begin{aligned}\hat{y}_H(k|k-1) &= (1-D)y_H + Fu_H + H_2 w_H \\ &= \underbrace{\left[(1-D)y_H + Fu_H + z^{-1}w_H\right]}_{\Phi_H \theta_0} + w_H(k).\end{aligned}$$

Since H_2 has a pass-through term, the regressor becomes affine.

Multiplying $\hat{y}_H(k|k-1)$ by $H_1(z)$

$$H_1(z)\hat{y}_H(k|k-1) = (1-D)y + Fu + H_2 w$$

does *not* give the one step-ahead predictor for $y(k)$, because from the definition, we expect $y(k) - \hat{y}(k|k-1) = e(k)$; here instead we have

$$y(k) - H_1(z)\hat{y}_H(k|k-1) = H_1(z)e(k)$$

using eq. (2). To compute the correct one-step ahead predictor for $y(k)$, we use the formula derived in class

$$\hat{y} = H^{-1}Gx + (1 - H^{-1})y, \quad Gx = \frac{F}{D}u + \frac{H_2}{D}w, \quad H = \frac{H_1}{D}.$$

The formula holds also when $\frac{H_2}{D}$ has a pass-through term (at least I do not see a reason why it should not). Straightforward application gives

$$\begin{aligned}\hat{y}(k|k-1) &= Fu_H + H_2 w_H + (H_1 - D)y_H \\ &= (1-D)y_H + Fu_H + H_2 w_H + (H_1 - 1)y_H\end{aligned}$$

where the second form can be used for the parametrization and the first to compute the prediction once the parameters $\hat{\theta}$ are known.

Problem 2

Given the two pulse responses $\{U_1, Y_1\}$ and $\{U_2, Y_2\}$ from the same system $G(z)$, how do we estimate $\hat{G}(z)$?

One way is to estimate the two separately and combine the estimates:

$$\hat{G}_1 = (\Phi_1^\top \Phi_1)^{-1}\Phi_1^\top Y_1, \qquad \hat{G}_2 = (\Phi_2^\top \Phi_2)^{-1}\Phi_2^\top Y_2, \qquad \hat{G} = \frac{\hat{G}_1 + \hat{G}_2}{2}$$

Another way is by stacking the measurements:

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} \Phi_1 \\ \Phi_2 \end{bmatrix} \hat{g}, \qquad \hat{g} = \frac{1}{\Phi_1^\top \Phi_1 + \Phi_2^\top \Phi_2}(\Phi_1^\top Y_1 + \Phi_2^\top Y_2)$$

Clearly the two results are not the same, even when the regressors are noise-free (*e.g.* when one applies two different sets of control inputs) but they are both unbiased. Numerically, the standard deviation is also equal, see implementation `test3.m`.

The decisive argument in favour of the second approach is that the measurement with the stronger signal should count more. This is encoded in Φ: a stronger control input results in larger Φ and Y, since $Y = \Phi G_0 + e$.

Problem 3

The systems model is

$$y(k) = \frac{B(z)}{A(z)} u(k) + \frac{1}{A(z)} e(k)$$

After computing $u(k) = r(k) - C(z)y(k)$, the problem is a standard ARX model.