ME 233 Spring 2012
Solution to Homework #9

1. (a) With the given conditions, we have

$$\begin{aligned}
\sum_{j=0}^{k} y^T(j)u(j) &= \sum_{j=0}^{k} \left(y_1^T(j) + y_2^T(j)\right) u(j) \\
&= \sum_{j=0}^{k} y_1^T(j)u(j) + \sum_{j=0}^{k} y_2^T(j)u(j) \\
&\geq -\gamma_1^2 - \gamma_2^2 \triangleq -\gamma^2
\end{aligned}$$

Thus, the parallel combination satisfies the Popov inequality.

(b) With the given conditions, we have

$$\begin{aligned}
-\gamma_1^2 &\leqslant \sum_{j=0}^{k} y^T(j)\left(u(j) - y_1(j)\right) = \sum_{j=0}^{k} y^T(j)u(j) - \sum_{j=0}^{k} y^T(j)y_1(j) \\
\Rightarrow \sum_{j=0}^{k} y^T(j)u(j) &\geq -\gamma_1^2 + \sum_{j=0}^{k} y^T(j)y_1(j) \geq -\gamma_1^2 - \gamma_2^2 \triangleq -\gamma^2
\end{aligned}$$

Thus, the feedback combination satisfies the Popov inequality.

2. (a)

$$\begin{aligned}
G(e^{j\omega}) &= \frac{e^{j\omega} - b}{e^{j\omega} - a} = \frac{\left(e^{j\omega} - b\right)\left(e^{-j\omega} - a\right)}{|e^{j\omega} - a|^2} = \frac{1 + ab - \left(ae^{j\omega} + be^{-j\omega}\right)}{|e^{j\omega} - a|^2} \\
&= \frac{1 + ab - (a+b)\cos\omega - j(a-b)\sin\omega}{|e^{j\omega} - a|^2}
\end{aligned}$$

In order to make $G(e^{j\omega})$ touches the imaginary axis, we must have

$$0 = 1 + ab - (a+b)\cos\omega \qquad \Leftrightarrow \qquad 1 + ab = (a+b)\cos\omega\ .$$

Since $\cos\omega \in [-1, 1]$, we have

$$\begin{aligned}
\exists\omega \text{ s.t. } 0 = 1 + ab - (a+b)\cos\omega &\Leftrightarrow (1+ab)^2 \leq (a+b)^2 \\
&\Leftrightarrow 1 + a^2b^2 \leq a^2 + b^2 \\
&\Leftrightarrow (1-a^2)(1-b^2) \leq 0 \\
&\Leftrightarrow \frac{1-b^2}{1-a^2} \leq 0 \\
&\Leftrightarrow \frac{b^2-1}{1-a^2} \geq 0\ .
\end{aligned}$$

(b) Obviously,

$$\text{All poles of } G(z) \text{ are asymptotically stable} \quad \Leftrightarrow \quad a^2 < 1.$$

From the results in the previous part, we know

$$\begin{aligned}
\mathcal{R}e\left\{G(e^{j\omega})\right\} > 0 \text{ for } \forall\omega \in [0, 2\pi] \quad &\Leftrightarrow \quad 1 + ab - (a+b)\cos\omega > 0 \text{ for } \forall\omega \in [0, 2\pi] \\
&\Leftrightarrow \quad 1 + ab > |a + b| \\
&\Leftrightarrow \quad 1 + a^2b^2 > a^2 + b^2 \text{ and } 1 + ab > 0 \\
&\Leftrightarrow \quad (1 - a^2)(1 - b^2) > 0 \text{ and } 1 + ab > 0 \; .
\end{aligned}$$

With $a^2 < 1$ for asymptotically stable poles, we have

$$\begin{aligned}
G(z) \text{ is SPR} \quad &\Leftrightarrow \quad 1 - b^2 > 0,\, 1 - a^2 > 0, \text{ and } 1 + ab > 0 \\
&\Leftrightarrow \quad 1 - b^2 > 0 \,\text{and}\, 1 - a^2 > 0 \\
&\Leftrightarrow \quad b^2 < 1 \,\text{and}\, a^2 < 1 \; .
\end{aligned}$$

3. These are the key points:

   (a) You should notice that, although the system converges very slowly, it does converge for constant gains with a white input and no measurement noise. Furthermore, the rate of convergence will be proportional to the gain times the parameter value, so a balanced matrix (or one slightly favoring the b terms) will work best. The constant gain case exhibits more oscillation in the parameter estimates than the least squares and least squares with forgetting factor. The convergence for the constant gain is also slower than for the least squares algorithms. The least squares with forgetting factor had the fastest convergence by several orders of magnitude.

   (b) This system has 4 unknown parameters (because there is no b_0 term), so a persistence of excitation of level four (rather than the usual 5 for $m = 2$, $n = 3$) will be needed to see convergence. With no measurement noise, you should see convergence for white noise and 2 or 3 sine inputs, but not for a step input (level 1) or single sine (level 2).

   (c) Least squares should reject white noise because it effectively averages it out over all time. (A little error will remain.) However, colored noise can be correlated with past values, which violates the assumptions of the least squares design, so the parameter estimates will converge to incorrect values. Finally, forgetting factor prevents noise from being completely averaged, and new bursts of noise can enter into the system gain, causing the parameters to deviate occasionally from their true values.

4. (a)

$$\begin{aligned}
e(k+1) &= y(k+1) - \hat{y}(k+1) = y(k+1) - \hat{\theta}^T(k+1)\hat{\phi}(k) \\
&= y(k+1) - \left[\hat{\theta}^T(k) + \frac{1}{\lambda_1(k)} v(k+1)\hat{\phi}^T(k)F(k)\right]\hat{\phi}(k) \\
&= y(k+1) - \hat{y}^o(k+1) + v(k+1) - v(k+1)\left[1 + \frac{1}{\lambda_1(k)}\hat{\phi}^T(k)F(k)\hat{\phi}(k)\right] \\
&= y(k+1) - \hat{y}^o(k+1) + v(k+1) - v(k+1)\left[\frac{\lambda_1(k) + \hat{\phi}^T(k)F(k)\hat{\phi}(k)}{\lambda_1(k)}\right] \\
&= y(k+1) - \hat{y}^o(k+1) + v(k+1) - v^o(k+1) \\
&= e^o(k+1) + v(k+1) - e^o(k+1) - c_1 e(k) - c_2 e(k-1) \\
&= v(k+1) - c_1 e(k) - c_2 e(k-1) \\
\Rightarrow v(k+1) &= e(k+1) + c_1 e(k) + c_2 e(k-1)
\end{aligned}$$

(b)

$$
\begin{aligned}
e(k+1) &= y(k+1) - \hat{y}(k+1) = \theta^T \phi(k) - \hat{\theta}^T(k+1)\hat{\phi}(k) \\
&= \theta^T \hat{\phi}(k) - \hat{\theta}^T(k+1)\hat{\phi}(k) + \theta^T \phi(k) - \theta^T \hat{\phi}(k) \\
&= \tilde{\theta}^T(k+1)\hat{\phi}(k) + \theta^T \left[\phi(k) - \hat{\phi}(k)\right] \\
&= m(k+1) + a_1 \left[\hat{y}(k) - y(k)\right] + a_2 \left[\hat{y}(k-1) - y(k-1)\right] \\
&= m(k+1) - a_1 e(k) - a_2 e(k-1) \\
\Rightarrow E(z) &= \frac{1}{1 + a_1 z^{-1} + a_2 z^{-2}} M(z)
\end{aligned}
$$

Since

$$V(z) = \left[1 + c_1 z^{-1} + c_2 z^{-2}\right] E(z)$$

we can say that

$$
\begin{aligned}
V(z) &= \frac{1 + c_1 z^{-1} + c_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}} M(z) \\
\Rightarrow G(z^{-1}) &= \frac{1 + c_1 z^{-1} + c_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}}
\end{aligned}
$$

(c) Applying the same block diagram operations to the feedback loop as were applied in lecture results in the block diagram shown in Figure 1. Note that with the exception of the $v(k+1)$ in the parameter update equation, the PAA is the same as was presented in lecture. However, since NL has $v(k+1)$ as its input, NL is the same as was presented in lecture. Thus, we know that NL_1 is P-class for $\lambda_2(k) \geq 0$ and $0 < \lambda_1(k) \leq 1$. Now note that NL_2 is made up of NL_1 in negative feedback with a time-varying gain. For now, we will assume that this time-varying gain is P-class so that we can conclude that NL_2 is P-class. (We will justify this assumption in the next part.) Thus, under this assumption, the feedback loop is made up of an LTI block, $G(z^{-1}) - \lambda/2$, in feedback with the P-class nonlinearity, NL_2.

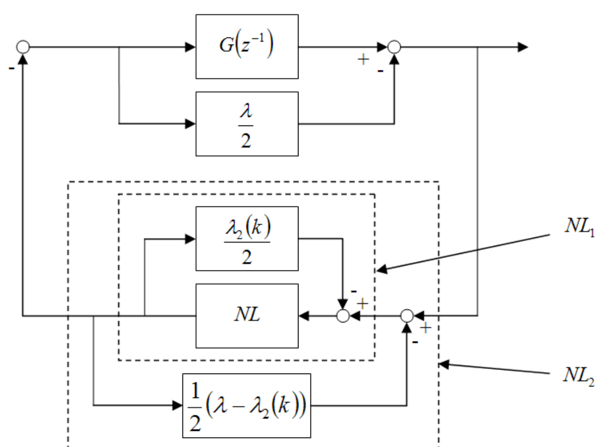

Figure 1: Block diagram of feedback loop after performing block diagram operations

(d) Based on the results in the previous section, we know that the only condition we need to impose on $G(z^{-1})$ is that

$$G(z^{-1}) - \frac{\lambda}{2} \quad - \quad \text{SPR}$$

where λ is chosen so that

$$0 \leq \lambda_2(k) \leq \lambda < 2, \qquad \forall k$$

Note in particular that this assumption implies that the polynomial $1 + a_1 q^{-1} + a_2 q^{-2}$ must be anti-Schur.

(e) First note that, as in the series-parallel system analyzed in lecture, this choice of λ will guarantee that the time-varying gain in NL_2 and, hence, NL_2 itself are P-class nonlinearities. Now, for convenience, we will define the transfer function of the LTI block and its output respectively as

$$\begin{aligned}\overline{G}(z^{-1}) &= G(z^{-1}) - \frac{\lambda}{2} \\ \overline{v}(k) &= v(k) - \frac{\lambda}{2} m(k)\end{aligned}$$

Since $\overline{G}(z^{-1})$ is SPR by assumption, the system will be asymptotically hyperstable. Therefore, we know that

$$\lim_{k\to\infty} \overline{v}(k) = 0 \qquad\qquad \lim_{k\to\infty} m(k) = 0 \;.$$

Since $v(k) = \overline{v}(k) + \lambda m(k)/2$, we therefore see that

$$\lim_{k\to\infty} v(k) = 0 \;.$$

Using the result of the hint for part (b), we see that

$$[1 + a_1 q^{-1} + a_2 q^{-2}]e(k) = m(k) \;.$$

Since $m(k) \to 0$ and the polynomial $1 + a_1 q^{-1} + a_2 q^{-2}$ is anti-Schur, we therefore see that

$$\lim_{k\to\infty} e(k) = 0 \;.$$

It now only remains to show that $e^o(k)$ and $v^o(k)$ converge to zero. Since the polynomial $1 + a_1 q^{-1} + a_2 q^{-2}$ is anti-Schur, the plant dynamics must be asymptotically stable. Thus, if $u(k)$ is bounded, then $y(k)$ must also be bounded. Since $e(k)$ is bounded, this also implies that $\hat{y}(k) = y(k) - e(k)$ must also be bounded. Therefore, $\hat{\phi}(k)$ must be bounded. Since $F(k)$ is assumed to be bounded, this implies that $\lambda_1(k) + \hat{\phi}^T(k)F(k)\hat{\phi}(k)$ remains bounded, which in turn implies that

$$\lim_{k\to\infty} v^o(k+1) = \lim_{k\to\infty} \left\{ \frac{\lambda_1(k) + \hat{\phi}^T(k)F(k)\hat{\phi}(k)}{\lambda_1(k)} v(k+1) \right\} = 0$$

provided that $\lambda_1(k)$ does not converge to 0. Therefore

$$\lim_{k\to\infty} e^o(k) = \lim_{k\to\infty} \{v^o(k) - c_1 e(k-1) - c_2 e(k-2)\} = 0 \;.$$

5. Figure 2 shows the plot of $\overline{G}(z^{-1})$ for $[c_1, c_2, \lambda] = [1.3, 0.42, 1]$. Note that since $\overline{G}(z^{-1})$ is asymptotically stable and its Nyquist plot lies in the open right half-plane, $\overline{G}(z^{-1})$ is SPR. Thus, we will use these values to conduct our simulations for different parameter adaptation algorithms.

   In the following simulation, we choose the initial $F = \text{diag}\{10^2, 10^2, 10^2, 10^2\}$. For the least square gain with forgetting factor, $\lambda_1 = 0.95$.

   Figure 5 and Figure 3 show the least squares parameter estimates vs. time with and without measurement noise respectively. Note that in this particular set of simulations, the measurement noise actually improved the speed of convergence for small k. Figure 4 shows the least squares parameter estimates with forgetting factor for small k. For large k, the parameter estimates display burst and go unstable, because the gain matrix $F(k)$ becomes unbounded for these choices of λ_1 and λ_2. From the simulation results, we see the forgetting factor can improve the convergence speed of the parameter estimates. Figure 6 shows the results with colored measurement noise. Compared to the case of white measurement noise, the parameter estimates under color measurement noise converge more slowly. However, in this parallel adaptation scheme, the parameter estimates converge to correct values even in the face of colored noise. This is the main advantage of using parallel adaptation instead of series-parallel adaptation.

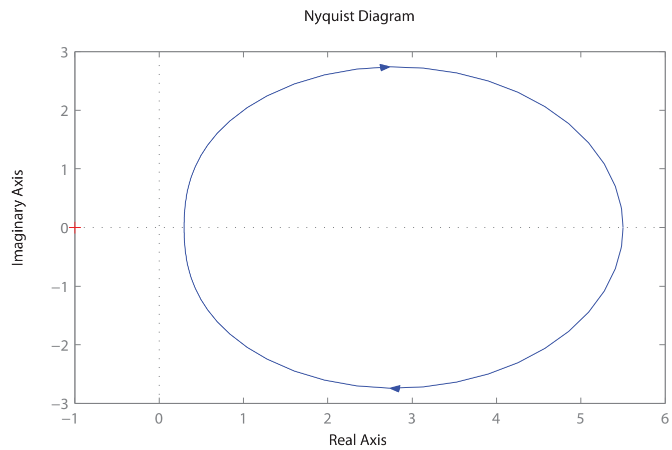

Figure 2: Nyquist plot of $\overline{G}(z^{-1})$ for $[c_1, c_2, \lambda] = [1.3, 0.42, 1]$

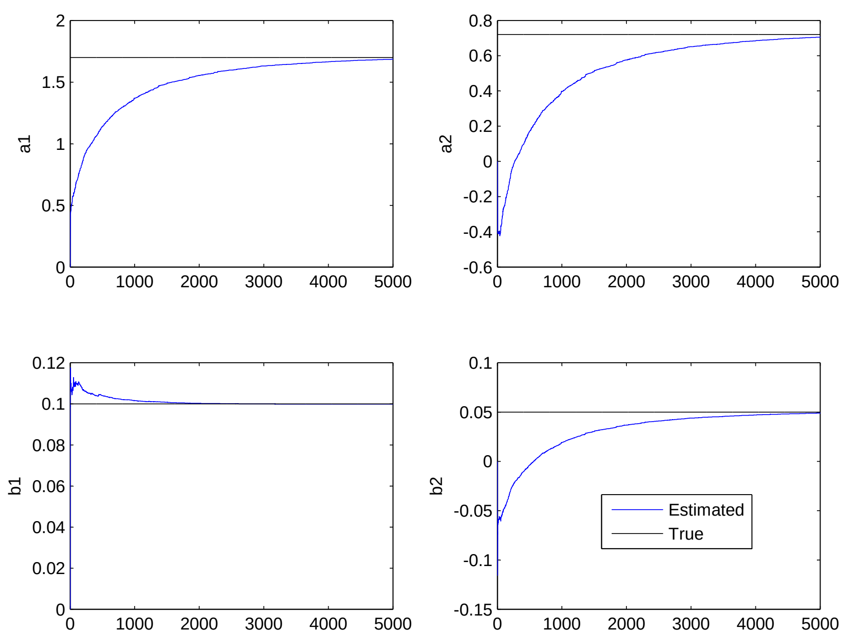

Figure 3: Random input, no measurement noise, and no forgetting factor

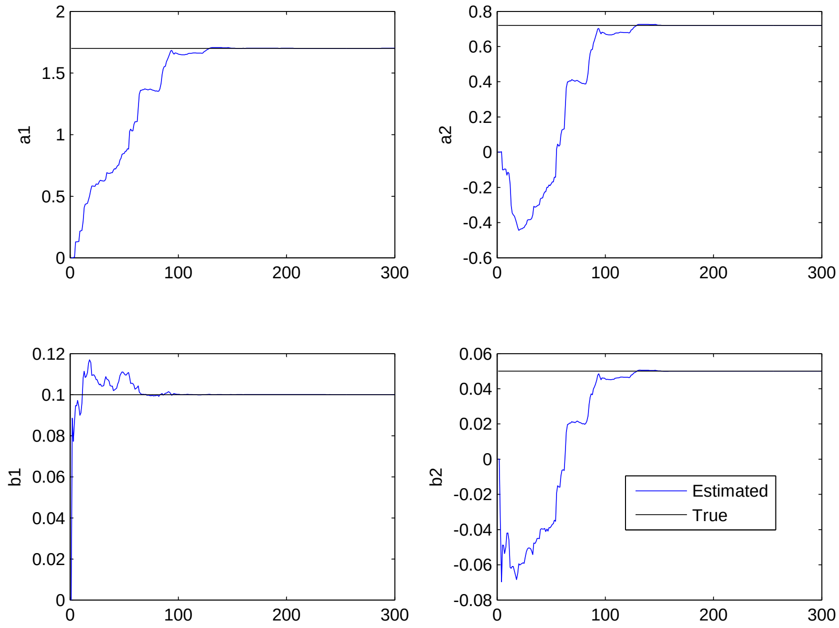

Figure 4: Random input, no measurement noise, and least square gain with forgetting factor

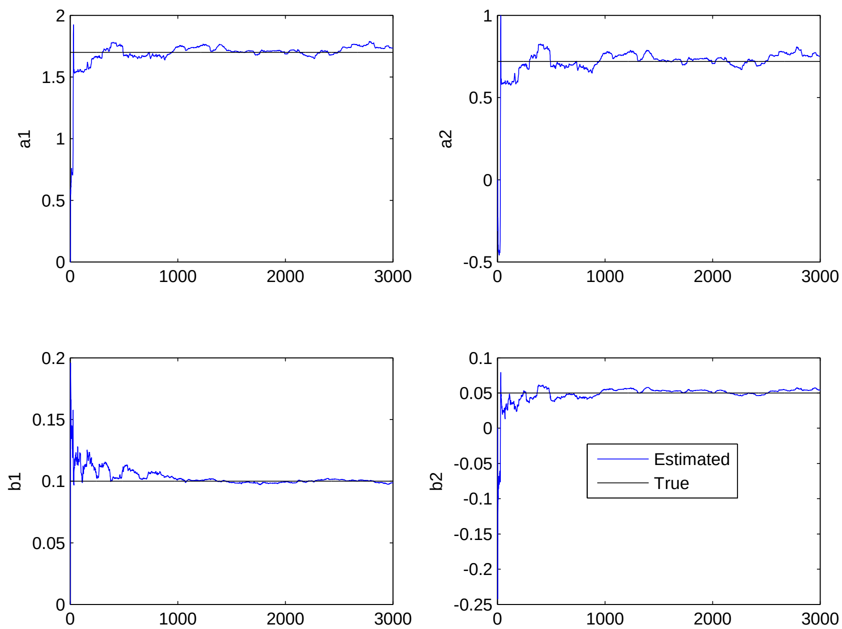

Figure 5: Random input, white measurement noise, and no forgetting factor

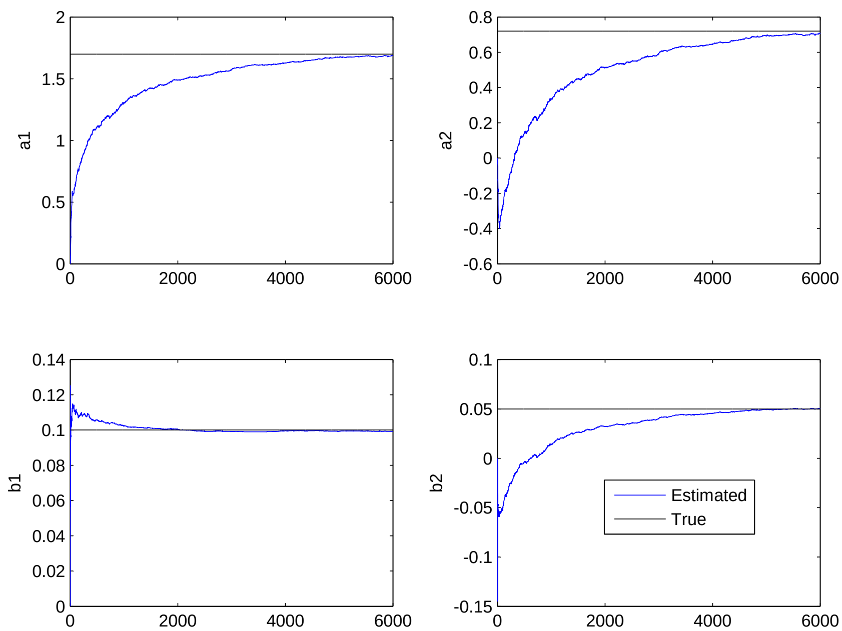

Figure 6: Random input, colored measurement noise, and no forgetting factor