Problem Set 2

6.435 System Identification
Ali J. Azarbayejani
9 October 1991

1. Nonparametric time- and frequency-domain methods.

Ljung Problem 6G.3

$$
\begin{aligned}
E\left|\hat{G}_N(e^{i\omega})\right|^2 &= E\hat{G}_N(e^{i\omega})\overline{\hat{G}_N(e^{i\omega})} \\
&= E\hat{G}_N(e^{i\omega})\hat{G}_N(e^{-i\omega}) \\
&= E\left(G_o(e^{i\omega}) + \frac{R_N(\omega)}{U_N(\omega)} + \frac{V_N(\omega)}{U_N(\omega)}\right)\left(G_o(e^{-i\omega}) + \frac{R_N(-\omega)}{U_N(-\omega)} + \frac{V_N(-\omega)}{U_N(-\omega)}\right) \\
&= \left|G_o(e^{i\omega})\right|^2 + \frac{G_o(e^{i\omega})R_N(-\omega)}{U_N(-\omega)} + \frac{G_o(e^{i\omega})}{U_N(-\omega)}EV_N(-\omega) \\
&\quad + \frac{G_o(e^{-i\omega})R_N(\omega)}{U_N(\omega)} + \frac{|R_N(\omega)|^2}{|U_N(\omega)|^2} + \frac{R_N(\omega)}{|U_N(\omega)|^2}EV_N(-\omega) \\
&\quad + \frac{G_o(e^{-i\omega})}{U_N(\omega)}EV_N(\omega) + \frac{R_N(-\omega)}{|U_N(\omega)|^2}EV_N(\omega) + \frac{E\left|V_N(\omega)\right|^2}{|U_N(\omega)|^2}
\end{aligned}
$$

Since $v(t)$ is zero mean, $EV_N(\omega) = EV_N(-\omega) = 0$. Also, asymptotically $R_N(\omega) \to 0$ and $E\left|V_N(\omega)\right|^2 \to \Phi_v(\omega)$, thus

$$
E\left|\hat{G}_N(e^{i\omega})\right|^2 \quad \to \quad \left|G_o(e^{i\omega})\right|^2 + \frac{\Phi_v(\omega)}{|U_N(\omega)|^2}
$$

asymptotically as N gets large.

Ljung Problem 6G.4

The Trooly-Cukey spectral estimate is

$$
\hat{\Phi}_u^N(\omega) = \frac{1}{M}\sum_{k=1}^{M}\left|U_R^{(k)}(\omega)\right|^2.
$$

Since Ljung does not state the cross-spectral estimate explicitly, we must figger out what it might be. Using an approach identical to the proof of Lemma 2.1, it can be shown that $EX_N(\omega)\overline{Y_N}(\omega) \xrightarrow{weak} \Phi_{xy}(\omega)$. Thus the Trooly-Cukey cross-spectral estimate is

$$
\hat{\Phi}_{yu}^N(\omega) = \frac{1}{M}\sum_{k=1}^{M} Y_R^{(k)}(\omega)\overline{U_R^{(k)}}(\omega)
$$

Equation (6.70) is

$$\begin{aligned}
\hat{G}_N(e^{i\omega}) &= \frac{\sum_{k=1}^{M} \left|U_R^{(k)}(\omega)\right|^2 \hat{G}_R^{(k)}(e^{i\omega})}{\sum_{k=1}^{M} \left|U_R^{(k)}(\omega)\right|^2} \\
&= \frac{\frac{1}{M}\sum_{k=1}^{M} U_R^{(k)}(\omega)\overline{U_R^{(k)}}(\omega)Y_R^{(k)}(\omega)/U_R^{(k)}(\omega)}{\frac{1}{M}\sum_{k=1}^{M} \left|U_R^{(k)}(\omega)\right|^2} \\
&= \frac{\hat{\Phi}_{yu}^N(\omega)}{\hat{\Phi}_u^N(\omega)}
\end{aligned}$$

Ljung Problem 6E.1

For the impulse response method, we took

$$\hat{g}(t) = y(t)/\alpha$$

for $t = 1, \ldots, N$. An estimate for $G_o(e^{i\omega})$ might be the Fourier transform of this estimated impulse response:

$$\begin{aligned}
\hat{G}_o(e^{i\omega}) &= \sum_{k=1}^{N} \hat{g}(t)e^{-i\omega k} \\
&= \frac{\sum_{k=1}^{N} y(k)e^{-i\omega k}}{\alpha} \\
&= \frac{\sqrt{N}Y_N(\omega)}{\sqrt{N}U_N(\omega)}
\end{aligned}$$

which is the ETFE after cancelling the common constant.

Ljung Problem 6E.2

Clearly, $u(t)$ and $y(t)$ are highly correlated, in fact they are scaled versions of one another and

$$U_N(\omega) = -KY_N(\omega).$$

Thus,

$$\hat{G}_N(e^{i\omega}) = -1/K.$$

The reason Lemma 6.1 does not apply here is because the feedback causes $u(t)$ to be highly correlated with $v(t)$. In fact,

$$u(t) = \frac{-K}{1+KG_o}v(t)$$

which implies that Euv is a nonzero function of Ev^2. Lemma 6.1 requires uncorrelated $\{u(t)\}$ and $\{v(t)\}$.

Ljung Problem 6E.3

Take a slightly more general case. Let w_k, $k = 1, \ldots, M$ be independent random variables with mean m and variances $E(w_k - m)^2 = \lambda_k$. Now, for

$$w = \sum_{k=1}^{M} \alpha_k w_k$$

we wish to find $\{\alpha_k\}$ to satisfy (a) $Ew = m$, where I have changed the condition appropriately for my slightly more general case. Since

$$\begin{aligned} Ew &= \sum_{k=1}^{M} \alpha_k E w_k \\ &= m \sum_{k=1}^{M} \alpha_k \end{aligned}$$

the constraint imposed by (a) is that the coefficients sum to 1:

$$\sum_{k=1}^{M} \alpha_k = 1.$$

Now we must consider (b), minimize $E(w-m)^2$. First,

$$\begin{aligned} E(w-m)^2 &= E\left(\sum_k \alpha_k w_k \quad - \quad m\right)^2 \\ &= E\left(\left(\sum_k \alpha_k w_k\right)^2 - 2m\sum_k \alpha_k w_k + m^2\right) \\ &= \sum_k \alpha_k^2 E w_k^2 + 2\sum_{k\neq j} \alpha_j \alpha_k \underbrace{E w_j w_k}_{zero} - 2m \sum_k \alpha_k \underbrace{E w_k}_{m} + m^2 \\ &= \sum_k \alpha_k^2(\lambda_k + m^2) - 2m^2 + m^2 \end{aligned}$$

where the constraint (a) has been enforced to obtain the second term of the last equality.

Since m is constant, minimizing $\sum_k \alpha_k^2(\lambda_k + m^2)$ will result in the desired minimization. Thus we have a constrained minimization problem and can employ a Lagrange multiplier method. In this method we would like to minimize

$$Q = \sum_{k=1}^{M} \alpha_k^2(\lambda_k + m^2) + \mu\left(\sum_{k=1}^{M} \alpha_k - 1\right)$$

The minimum can be found by where the partial derivatives w.r.t α_k and μ are zero. Thus,

$$\frac{\partial Q}{\partial \alpha_k} = 2\alpha_k(\lambda_k + m^2) + \mu = 0$$

$$\frac{\partial Q}{\partial \mu} = \sum_{k=1}^{M} \alpha_k - 1 = 0$$

must be satisfied. These are $M+1$ independent linear equations in the unknowns $\{\alpha_k\}$ and μ, hence there is a unique solution. Considering that

$$2\alpha_j(\lambda_j + m^2) = 2\alpha_k(\lambda_k + m^2)$$

must be satisfied for all j, k, the solution must be of the form $\alpha_k = c_k/(\lambda_k + m^2)$. The first M partial derivatives give $c_k = -\mu/2$. Thus, from the last partial derivative,

$$\mu = \frac{-2}{\sum_k 1/(\lambda_k + m^2)}$$

resulting in

$$c_k = \frac{1}{\sum_k 1/(\lambda_k + m^2)}$$

Now, defining the *inverse mean square* as $\eta_k = 1/(\lambda_k + m^2)$ and the sum of inverse mean squares as $S_M = \sum_{k=1}^{M} \eta_k$, we have

$$\alpha_k = \frac{\eta_k}{S_M}$$

For the special case $m = 1$, $\eta_k = 1/(\lambda_k + 1)$.

The minimal variance of w, call it λ^* is

$$\begin{aligned}
\lambda^* &= \sum_k \left(\frac{\eta_k}{S_M}\right)^2 \eta_k^{-1} \quad - \quad m^2 \\
&= \frac{1}{S_M}\sum_k \alpha_k \quad - \quad m^2 \\
&= \frac{1 - m^2 S_M}{S_M}
\end{aligned}$$

Application of this technology: Ljung equation (6.69) gives an estimate of $w = \hat{G}_N(e^{i\omega})$ as the weighted sum of independent measurements, where $\alpha_k = 1/M$ for all k. Since the sum of alphas is unity, the mean of the estimate is the same as the means of the measurements (assuming they all have the same mean), but this simple set of alphas does not guarantee the minimum variance on the estimate.

Ljung equation (6.70) again estimates $\hat{G}_N(e^{i\omega})$ as a weighted sum of measurements, but this time with

$$\alpha_k = \frac{1/E\left|\hat{\hat{G}}_R^{(k)}(e^{i\omega}) - G_o(e^{i\omega})\right|^2}{\sum_k 1/E\left|\hat{\hat{G}}_R^{(k)}(e^{i\omega}) - G_o(e^{i\omega})\right|^2}$$

$$= \frac{\left|U_R^{(k)}(\omega)\right|^2 / \Phi_v(\omega)}{\sum_k \left|U_R^{(k)}(\omega)\right|^2 / \Phi_v(\omega)}$$

$$= \frac{\left|U_R^{(k)}(\omega)\right|^2}{\sum_k \left|U_R^{(k)}(\omega)\right|^2}$$

This estimate of the transfer function is minimum variance based on the estimates of variance for the ETFE, which come from a finite sample of data. As R gets large, this estimate will asymptotically have the minimum variance.

Ljung Problem 6T.2

$$\begin{aligned}
E\left|S_N\right|^2 &= \frac{1}{N^2} E\left|\sum_{t=1}^{N}\sum_{s=1}^{N} \alpha_t \alpha_s v(t) v(s)\right| \\
&\leq \frac{1}{N^2} E\sum_{t=1}^{N}\sum_{s=1}^{N} \left|\alpha_t \alpha_s\right| v(t) v(s) \\
&\leq \frac{C_1^2}{N^2}\sum_{t=1}^{N}\sum_{s=1}^{N} R_v(t-s) \\
&= \frac{C_1^2}{N^2}\sum_{t=1}^{N}\sum_{\tau=t-N}^{t-1} R_v(\tau) \\
&= \frac{C_1^2}{N^2}\sum_{t=1}^{N}\left(\Phi_v(0) - \sum_{\tau=-\infty}^{t-N-1} R_v(\tau) - \sum_{\tau=t}^{\infty} R_v(\tau)\right)
\end{aligned}$$

Now, consider

$$\begin{aligned}
\left|\sum_{t=1}^{N}\sum_{\tau=-\infty}^{t-N} R_v(\tau)\right| &\leq \sum_{t=1}^{N}\sum_{\tau=-\infty}^{t-N} \left|R_v(\tau)\right| \\
&= \sum_{\tau=-\infty}^{-N}\sum_{t=1}^{N}\left|R_v(\tau)\right| + \sum_{\tau=-N+1}^{-1}\sum_{t=1}^{-\tau}\left|R_v(\tau)\right| \\
&= \sum_{\tau=-\infty}^{-N} N\left|R_v(\tau)\right| + \sum_{\tau=-N+1}^{-1}\left|\tau\right|\left|R_v(\tau)\right| \\
&\leq \sum_{\tau=-\infty}^{-1}\left|\tau\right|\left|R_v(\tau)\right|.
\end{aligned}$$

Similarly,

$$\left|\sum_{t=1}^{N}\sum_{\tau=t}^{\infty} R_v(\tau)\right| \leq \sum_{\tau=1}^{\infty} \tau\left|R_v(\tau)\right|$$

Thus,

$$\begin{aligned} E\left|S_N\right|^2 &\leq \frac{C_1^2}{N^2}\left(N\Phi_v(0) + N(2C_3)\right) \\ &= \frac{C_1^2(\Phi_v(0)+2C_3)}{N} = \frac{C_2}{N} \end{aligned}$$

where

$$C_3 \geq \sum_{-\infty}^{\infty} |\tau R_v(\tau)|$$

In relation to Lemma 6.1, relate α_t to $e^{-i\omega t}$, and C_1 to unity. Then, $V_N(\omega) = \sqrt{N} S_N$ and

$$\begin{aligned} E\left|V_N(\omega)\right|^2 &= NE\left|S_N\right|^2 \\ &\leq NC_2/N = C_2 \\ &= \Phi_v(0) + 2C_3 \end{aligned}$$

with C_3 defined above. This is a somewhat weaker result than in Lemma 6.2 where it is shown that the variance of the noise transform approaches the spectrum asymptotically for all values of ω.

The technical assumption on the correlation function is not terribly restrictive because typical noise correlations drop off rapidly. For example, a white noise satisfies the criterion and any noise with a finite length correlation function satisfies the criterion.

2. Hum.

We have

$$y = Gu + v$$

$$u = u^* + w$$

$$v = v^* + w$$

where u^* and v^* are the uncorrupted input and noise processes, respectively, and w is a 60Hz hum. By assuming that u and v have discrete frequency profiles that have no common frequency except that of w, we can be assured that u^*, v^*, and w are uncorrelated.

Thus,

$$\begin{aligned} \Phi_u(\omega) &= \Phi_{u^*}(\omega) + \underbrace{\Phi_{u^*w}(\omega)}_{0} + \underbrace{\Phi_{wu^*}(\omega)}_{0} + \Phi_w(\omega) \\ &= \Phi_{u^*}(\omega) + \Phi_w(\omega) \end{aligned}$$

(the cross terms are zero) and

$$\begin{aligned} \Phi_{vu}(\omega) &= \underbrace{\Phi_{v^*u^*}(\omega)}_{0} + \underbrace{\Phi_{v^*w}(\omega)}_{0} + \underbrace{\Phi_{wu^*}(\omega)}_{0} + \Phi_w(\omega) \\ &= \Phi_w(\omega) \end{aligned}$$

By transforming $Ey(t)u(t-\tau)$, we get

$$\Phi_{yu}(\omega) = G(e^{i\omega})\Phi_u(\omega) + \Phi_{vu}(\omega)$$

Thus,

$$\begin{aligned}\hat{G}(e^{i\omega}) &= \frac{\hat{\Phi}_{yu}(\omega)}{\hat{\Phi}_u(\omega)} \\ &= G(e^{i\omega}) + \frac{\Phi_{vu}(\omega)}{\hat{\Phi}_u(\omega)} \\ &= G(e^{i\omega}) + \frac{\Phi_w(\omega)}{\Phi_{u^*}(\omega) + \Phi_w(\omega)}\end{aligned}$$

The estimate is biased. The decision to use this approach is questionable and warrants further review.

3. Properties and smoothing of ETFEs.

(a) Generate data.

Ten values of N were used for generating data (M=10). For $k = 1, \ldots, M$ I used $N_k = 10k^2$. Thus

$$\{N_k\} = \{10, 40, 90, 160, 250, 360, 490, 640, 810, 1000\}.$$

For each value of N_k, forty simulations were run (P=40). This is done with the following MATLAB script for the ETFE estimates in part (b). A similar script was used to generate estimates using SPA in part (c).

```
% 6.435 Problem Set 2
% Problem 3:

% This is the actual system for Problem 3.
a = [1];
b = [0 1];
c = [1 0.5];
d = [1 -0.2];
f = [1 -0.6];
TH_ue_to_y = mktheta(a,b,c,d,f);

% This is for M=10 different data lengths
M = 10;
for k=1:M;
k,
% ...this generates the input sequence
N = 10*k^2,
N_array(k) = N;
```

```
for j=1:N;
u_j = 0;
for i=1:4;
u_j = u_j + cos(pi*i*j/2);
end
u(j) = u_j;
end
% ...and this generates the noise sequence.
rand('normal');
% This here simulates and estimates several times for this N
P = 40;
for j = 1:P;
e = rand(N,1);
y = idsim([u' e],TH_ue_to_y);
G_est = etfe([y u']);
g_pi_2(j) = G_est(65,2);
g_pi(j) = G_est(129,2);
end
mx(k) = mean(g_pi_2);
my(k) = mean(g_pi);
var_pi_2(k) = mean(diag(g_pi_2'*g_pi_2)) - (mean(g_pi_2))^2;
corr(k) = mean(diag(g_pi_2'*g_pi)) - mean(g_pi_2)*mean(g_pi);
end

plot(N_array,var_pi_2);
xlabel('number of steps, N');
ylabel('Variance at PI/2');
text(0.3,0.8,'M=10 values of N','sc');
text(0.3,0.75,'P=40 runs at each N','sc');
text(0.3,0.7,'ETFE Method','sc');
meta graph3b_var;

pause
plot(N_array,corr);
xlabel('number of steps, N');
ylabel('Correlation between PI/2 and PI');
text(0.3,0.5,'M=10 values of N','sc');
text(0.3,0.45,'P=40 runs at each N','sc');
text(0.3,0.4,'ETFE Method','sc');
meta graph3b_corr;
```

(b) ETFE estimates of G.

The above MATLAB script generates ETFE estimates. At each value of N_k, the mean and variance of $\hat{G}_{N_k}(e^{i\pi/2})$ is computed from the forty data runs at that N_k. Also, the correlation between

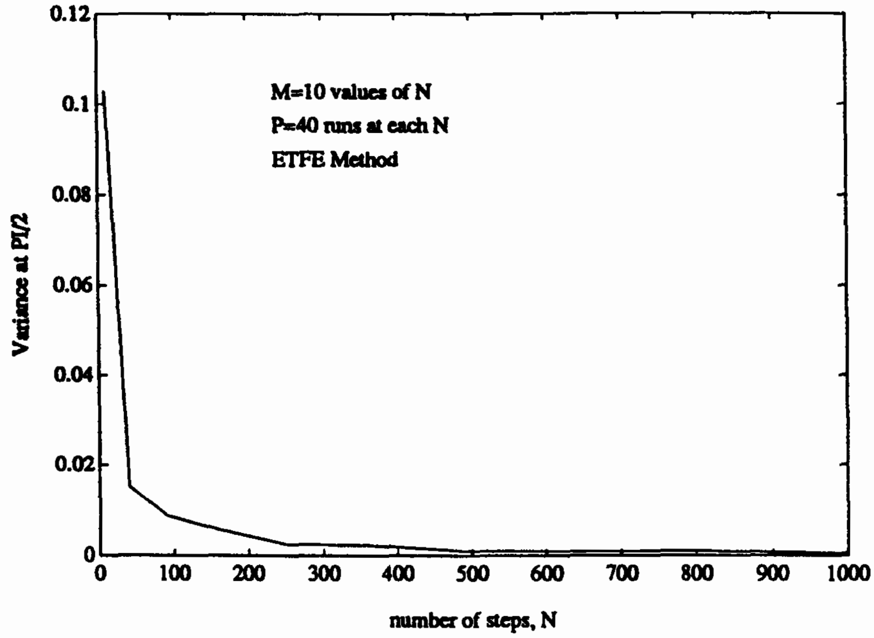

Figure 1: ETFE estimate of variance versus data length (P3b).

$\hat{\hat{G}}_{N_k}(e^{i\pi/2})$ and $\hat{\hat{G}}_{N_k}(e^{i\pi})$ is also computed and collected for each value of N_k.

The results are shown in Figures 1 and 2 and do confirm the theoretical properties of ETFE. Both should asymptotically decrease, the variance to a constant number and the correlation to zero, as described in Ljung Chapter 6.

(c) Spectral analysis.

The results are shown in Figures 3 and 4. The differences in my plots are not very marked. Both plots exhibit the same behavior, qualitatively and, to a large extent, quantitatively.

I could comment on the tradeoff between window length and data length, but I did not include this tradeoff in my simulations. In general, however, long windows will improve the variance but at the expense of introducing a bias, because the windowed frequencies will have some variance among them. For short data sequences, a long window might be the only way of achieving a low-variance estimate. My data shows, however, that with a data length over 500 samples, a window of 1 is as good as a window of 20. Thus for long data sequences, the window should be kept small to avoid biasing the estimate.

4. Nonparametric identification in feedback loop.

(a) Generate data.

We start with

$$y = Gu + v$$

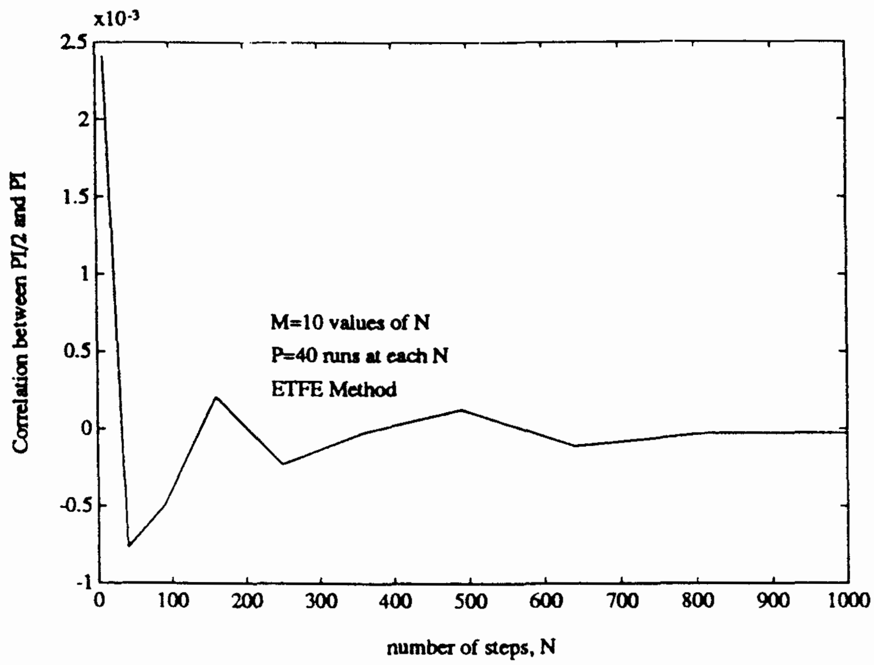

Figure 2: ETFE estimate of correlation versus data length (P3b).

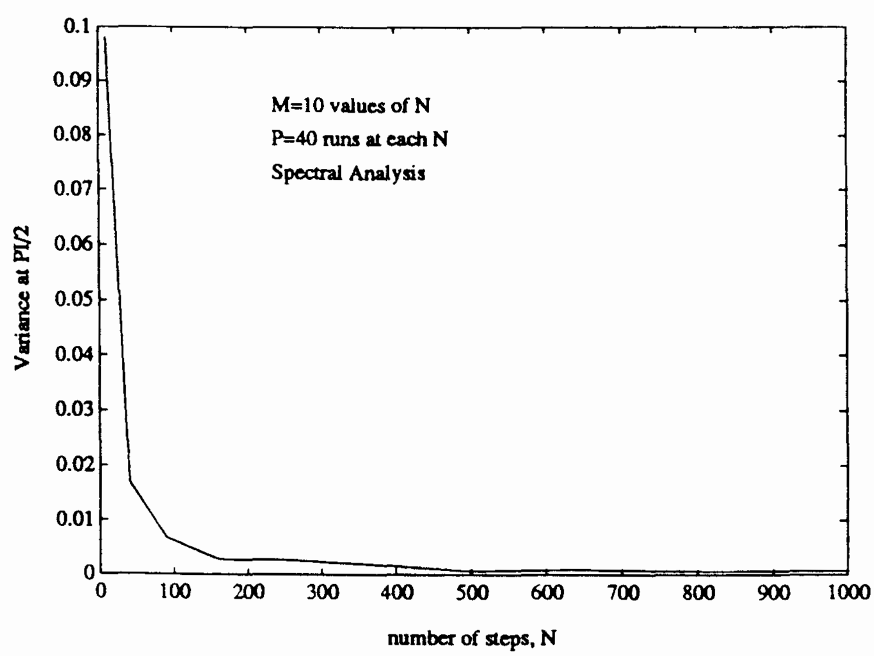

Figure 3: Spectral estimate of variance versus data length (P3c).

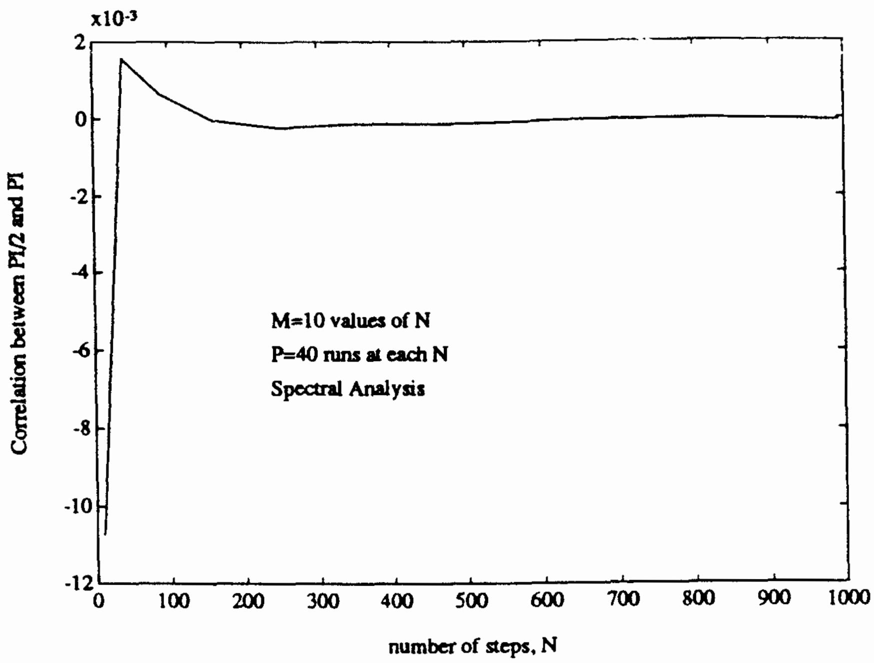

Figure 4: Spectral estimate of correlation versus data length (P3c).

$$v = He$$

$$u = r - Ky$$

$$G = \frac{q^{-1}}{1 - 0.5q^{-1}} \qquad K = 0.5 \qquad H = \frac{1 + 0.5q^{-1}}{1 - 0.2q^{-1}}$$

After algebraic manipulation, we have

$$y = q^{-1} r + \frac{1 - 0.25q^{-2}}{1 - 0.2q^{-1}} e$$

The above system is used to generate $y(t)$ from white sequences $r(t)$ and $e(t)$. Computationally, the input $u(t)$ is computed from the sequences $r(t)$ and $y(t)$. In a real experiment, u and y are available directly. The MATLAB script is

```
% This is the closed loop system with input r and noise e
a = [1];
b = [0 1];
c = [1 0 -0.25];
d = [1 -0.2];
TH_g_cl = mktheta(a,b,c,d);

% This is a simulation to generate y; (u is then computed)
```

```
rand('normal');
e = rand(1000,1);
r = rand(1000,1);
y = idsim([r e],TH_g_cl);
u = r - 0.5*y;
```

(b) Estimate G, estimate $\hat{\Phi}_v^N$

The correlation method for estimating the transfer function and the noise spectrum is implemented in the MATLAB function *spa*. The associated formulas are (6.51) and (6.78) in Ljung. The MATLAB script follows, the actual versus estimated transfer function is shown in Figure 5 and the actual versus estimated noise spectrum is shown in Figure 6.

```
% This is the open-loop system G with input u and output y
a = [1 -0.5];
b = [0 1];
c = 0;
TH_g = mktheta(a,b,c);
G_g = trf(TH_g);

% This is the noise generator with input e and output v
a = [1 -0.2];
b = [1 0.5];
c = 0;
TH_h = mktheta(a,b,c);
G_h= trf(TH_h);
NSP_v_act(1,:) = [100 0];
for k=2:length(G_h)
NSP_v_act(k,:) = [G_h(k,1) G_h(k,2)^2];
end

% This is the spectral estimate (based on correlation) and
%     the noise spectrum estimate for the data.
[G_g_est NSP_g_est] = spa([y u]);
bodeplot([G_g G_g_est]);
text(0.2,0.85,'G,G_est for y=Gu+He,u=r-Ky','sc');
meta graph4_1;

% This is a plot of the spectrum.
bodeplot([NSP_v_act NSP_g_est]);
text(0.2,0.85,'NSP,NSP_est for y=Gu+He,u=r-Ky','sc');
meta graph4_2;
```

(c) Theoretical explanation.

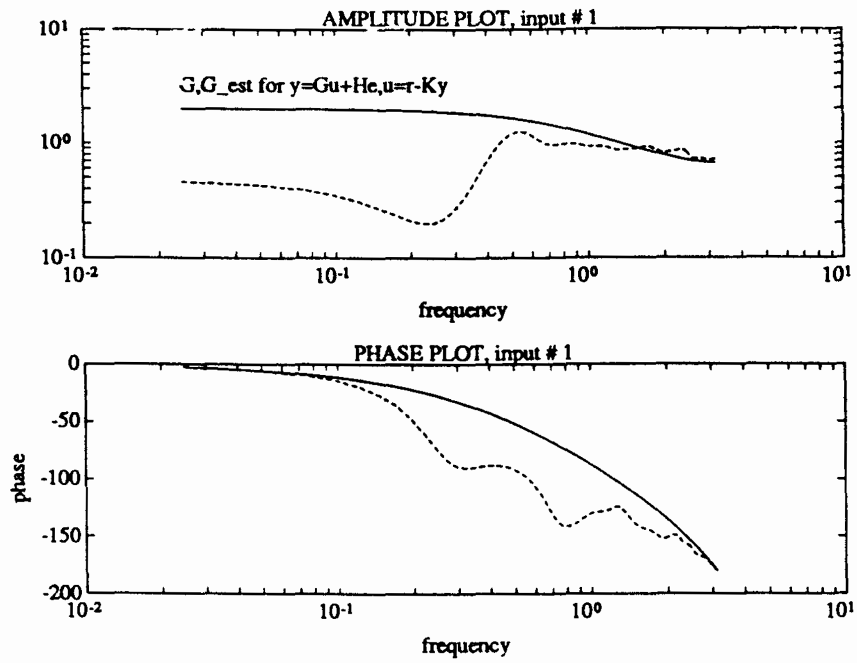

Figure 5: Spectral analysis vs. actual plant in feedback loop (P4b).

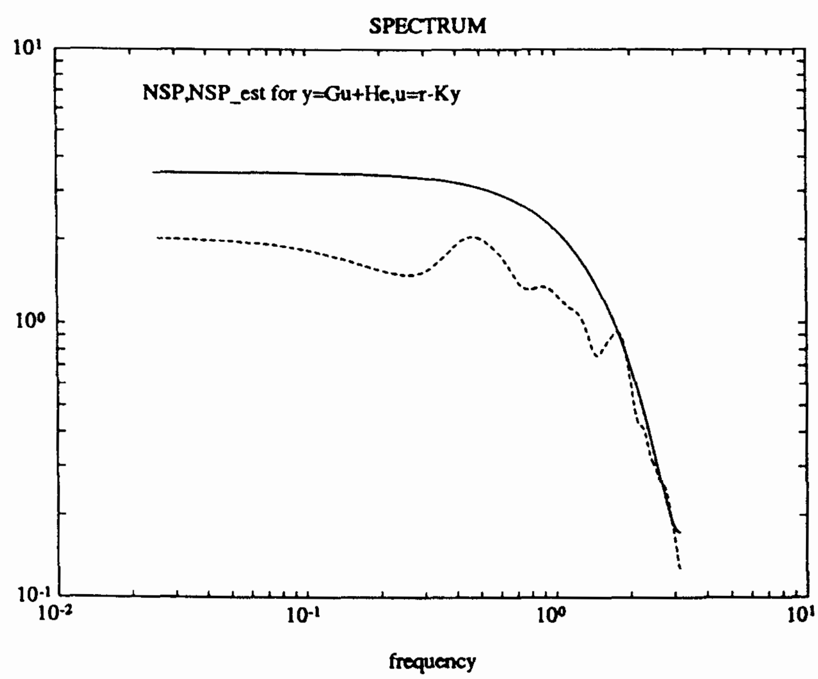

Figure 6: Spectral analysis vs. actual plant in feedback loop (P4b).

One should not expect the correlation method to work very well because the estimate is only unbiased if the input and the noise are uncorrelated, which is not the case for our feedback loop. The noise actual noise spectrum, as shown in Figure 6, has a relatively high gain at DC and falls off at high frequency. Thus, we should expect a good transfer function estimate and good noise spectrum estimate only at high frequency, which we see is the case.

To see the details of the feedback system, consider

$$y = Gu + He$$

Taking the expectation with respect to u yields

$$\begin{aligned} R_{yu}(\tau) &= Ey(t)u(t-\tau) \\ &= \sum_k g(k)Eu(t-k)u(t-\tau) + \sum_k h(k)Ee(t-k)u(t-\tau) \\ &= \sum_k g(k)R_u(\tau-k) + \sum_k h(k)R_{eu}(\tau-k) \end{aligned}$$

Fourier transforming yields

$$\Phi_{yu}(\omega) = G(e^{i\omega})\Phi_u(\omega) + H(e^{i\omega})\Phi_{eu}(\omega)$$

Similarly, starting with

$$u = r - Ky = r - KGu - KHe$$

and taking the expectation with respect to e yields

$$R_{ue}(\tau) = \underbrace{R_{re}(\tau)}_{0} - K\sum_k g(k)R_{ue}(\tau-k) - K\sum_k h(k)R_e(\tau-k)$$

which transforms to

$$\Phi_{ue}(\omega) = -KG(e^{i\omega})\Phi_{ue}(\omega) - KH(e^{i\omega})\Phi_e(\omega)$$

or

$$\Phi_{eu}(\omega) = -K\overline{G(e^{i\omega})}\Phi_{eu}(\omega) - K\overline{H(e^{i\omega})\Phi_e(\omega)}$$

In abbreviated notation,

$$\begin{aligned} \Phi_{yu} &= G\Phi_u + H\Phi_{eu} \\ &= G\Phi_u + H\frac{-K\overline{H\Phi_e}}{1+K\overline{G}} \end{aligned}$$

Thus the spectral estimate will yield

$$\begin{aligned} \hat{G} &= G + \frac{-K\,|H|^2\,\overline{\Phi_e}}{(1+K\overline{G})\Phi_u} \\ &= G + \frac{-K\overline{\Phi_v}}{(1+K\overline{G})_u} \end{aligned}$$

and the estimate is clearly biased.