**Assignment #1: estimation of correlation and PSD functions for LTI systems inputs and outputs**

Note: this document uses the same notation as in the course slides for Section 2, e.g., $m_x$, $\hat{m}_x$, $\phi_{yx}(l)$, $\gamma_{yx}(l)$, $\hat{\phi}_{yx}(l)$, $p_{yx}(l)$, $\Phi_{yx}(e^{j\omega})$, $P_{yx}(e^{j\omega})$, and $P^W_{yx}(e^{j\omega})$.

Throughout the assignment, we use the following relationships describing two LTI systems:

$$d(n) = x(n) * g(n) = \sum_{k=-\infty}^{\infty} x(k)g(n-k), \text{ with impulse response } g(n) = 2\delta(n-1) + \delta(n-2)$$

$$y(n) = x(n) * h(n) = \sum_{k=-\infty}^{\infty} x(k)h(n-k), \text{ with impulse response } h(n) = \delta(n-2) + 3\delta(n-3).$$

Each signal $x(n)$, $d(n)$, $y(n)$ is an observation (sequence, realization) over $L$ samples from wide sense stationary, mean ergodic and correlation ergodic random processes.

The signal $x(n)$ has an autocorrelation function of:

$$\phi_{xx}(l) = \sigma_x^2\delta(l) + |m_x|^2$$

with $\sigma_x^2 = 1$ and $m_x = 5$.

Question 1

a) Find the theoretical expressions for $\phi_{yd}(l)$, $\phi_{yx}(l)$, $\phi_{yy}(l)$.

b) Use the following code to generate an instance of the sequences $x(n)$, $d(n)$, $y(n)$:

```
L=2^16;
g=[0,2,1];
h=[0,0,1,3];
extra_samples_to_remove_transients = max(length(g),length(h))-1;

sigma_x_sq=1;
mx=5;
x=sqrt(sigma_x_sq)*randn(1,L+extra_samples_to_remove_transients)+mx;
d=filter(g,1,x);
y=filter(h,1,x);

% remove transients
x=x(extra_samples_to_remove_transients+1:end);
d=d(extra_samples_to_remove_transients+1:end);
y=y(extra_samples_to_remove_transients+1:end);
```

Compute the auto-correlation estimate $\hat{\phi}_{yx}(l)$ and plot it on the same figure as the theoretical $\phi_{yx}(l)$ (e.g., two subplots). Repeat for $\hat{\phi}_{yd}(l)/\phi_{yd}(l)$, and $\hat{\phi}_{yy}(l)/\phi_{yy}(l)$. If you use functions such as *xcorr*() instead of programming by yourself the details of the estimation, make sure that you use it with parameters such that it gives the same result as our definition.

Note: for all the plots, you must include meaningful titles, labels for vertical and horizontal axis, and a grid.

c) Compute the auto-covariance estimate $p_{yx}(l)$ and plot it on the same figure as the theoretical $\gamma_{yx}(l)$ (two subplots). Repeat for $p_{yd}(l)/\gamma_{yd}(l)$, and $p_{yy}(l)/\gamma_{yy}(l)$. Again, if you use functions such as *xcorr*() instead of programming by yourself the details of the estimation, make sure that you use it with parameters such that it gives the same result as our definition.

d) Generating $N = 100$ sequences of $y(n)$, you can experimentally approximate the variance of the estimate with the following expression:

$$\operatorname{var}\left\{\hat{\phi}_{yy}(l)\right\} = E\left[\left(\hat{\phi}_{yy}(l) - E\left[\hat{\phi}_{yy}(l)\right]\right)^2\right] \approx \frac{1}{N}\sum_{n=1}^{N}\left(\hat{\phi}_{yy}^{n}(l) - \hat{m}_{\phi}\right)^2$$

where $\hat{m}_{\phi} = \frac{1}{N}\sum_{n=1}^{N}\hat{\phi}_{yy}^{n}(l)$ and $\hat{\phi}_{yy}^{n}(l)$ represents the estimate measured from the $n^{\text{th}}$ sequence of $y(n)$. Likewise, $\operatorname{var}\left\{p_{yy}(l)\right\}$ can be experimentally approximated.

Verify that:

- the estimation variance $\operatorname{var}\left\{\hat{\phi}_{yy}(l)\right\}$ is larger for large lag values $|l|$
- the estimation variance $\operatorname{var}\left\{p_{yy}(l)\right\}$ is smaller for large lag values $|l|$
- Note: a direct comparison of $\operatorname{var}\left\{\hat{\phi}_{yy}(l)\right\}$ and $\operatorname{var}\left\{p_{yy}(l)\right\}$ would be misleading, because their levels are also proportional to the values of $\phi_{yy}(l)$ and $\gamma_{yy}(l)$, respectively (and $\phi_{yy}(l)$ and $\gamma_{yy}(l)$ are very different).

Question 2

Note: for this question, do not use Matlab/Octave functions such as *pwelch*(), *periodogram*(), or *cpsd*() to compute the PSD quantities. Use the basic *fft*() function and the equations for the PSD estimates.

Note: frequency domain signals in this question are written with the discrete time Fourier transform (DTFT) notation (e.g. $\Phi_{xx}(e^{j\omega})$, $-\pi \le \omega < \pi$), but it is understood that in practice they are discretized with fast Fourier transforms (FFTs) $\Phi_{xx}(k) \quad 0 \le k \le N_{fft} - 1$, where $N_{fft}$ is the size of the FFT (after zero-padding, if zero-padding is used). FFT results like $\Phi_{xx}(k)$ can be plotted to approximate DTFTs, with proper scaling $\omega_k = 2\pi / N_{fft}$ and "wrap-around" or "fftshift" processing, so that the second half of the $\Phi_{xx}(k)$ values appears as negative frequencies in the displayed $\Phi_{xx}(e^{j\omega})$ approximation.

a) Find theoretical expressions for $\Phi_{yx}(e^{j\omega})$, $\Phi_{yd}(e^{j\omega})$, $\Phi_{yy}(e^{j\omega})$.

b) Generating $N = 100$ sequences of $y(n)$, you can experimentally measure the variance of the periodogram estimate $P_{yy}(e^{j\omega})$ with the following expression: $\operatorname{var}\left\{P_{yy}(e^{j\omega})\right\} = E\left[\left(P_{yy}(e^{j\omega}) - E\left[P_{yy}(e^{j\omega})\right]\right)^2\right] \approx \frac{1}{N}\sum_{n=1}^{N}\left(P_{yy}^{n}(e^{j\omega}) - \hat{m}_P\right)^2$, where $\hat{m}_P = \frac{1}{N}\sum_{n=1}^{N} P_{yy}^{n}(e^{j\omega})$ and $P_{yy}^{n}(e^{j\omega})$ represents the estimate measured from the $n^{\text{th}}$ sequence of $y(n)$. Using different values of $L$ (ex. $L = 2^{12}$ and $L = 2^{16}$), verify that the variance of the periodogram estimate $P_{yy}(e^{j\omega})$ does not significantly decay as the number of observed samples $L$ is increased.

c) With $L = 2^{16}$, $M = 128$, $K = 1 + 2(L - M)/M = 1023$ (case with 50% overlap between windows), and using a (normalized) Hamming window, find the Welch estimates $P_{yx}^{W}(e^{j\omega})$, $P_{yd}^{W}(e^{j\omega})$ and $P_{yy}^{W}(e^{j\omega})$. Plot $\left|P_{yx}^{W}(e^{j\omega})\right|$ and the theoretical $\left|\Phi_{yx}(e^{j\omega})\right|$ on the same figure. Repeat with $\left|P_{yd}^{W}(e^{j\omega})\right|$ and $\left|\Phi_{yd}(e^{j\omega})\right|$, and $\left|P_{yy}^{W}(e^{j\omega})\right|$ and $\left|\Phi_{yy}(e^{j\omega})\right|$.

d) With $L = 2^{16}$, $M = 128$, $K = 1 + 2(L - M)/M = 1023$ (case with 50% overlap between windows), and using a (normalized) Hamming window, compute an estimate of the coherence between $x(n)$ and $y(n)$ with $\hat{\Psi}_{yx}(e^{j\omega}) = P_{yx}^{W}(e^{j\omega}) / P_{xx}^{W}(e^{j\omega})$, and compare $\left|\hat{\Psi}_{yx}(e^{j\omega})\right|$ with the theoretical coherence magnitude $\left|\Psi_{yx}(e^{j\omega})\right|$.