MODULE #6: CORRELATION FUNCTIONS

BMDE:519 ANALYSIS OF BIOMEDICAL SIGNALS AND SYSTEMS

Correlation Functions

- Characterize the sequential structure of signals in the time domain
- Useful to detect:
  - repeated patterns within a signal
  - similarities between signals
- Auto-Correlation
  - correlation of a signal with itself
  - used mainly in signal analysis
- Cross-Correlation
  - correlation of a signal with another signal
  - used in both signal and system analysis.

Auto-Correlation

$$R_{xx}(\tau) = E\left[x(t)x\left(t+\tau\right)\right]$$

- Characterizes the sequential structure of a signal, *x(t).*
- Measures the correlation of a signal, *x(t),* with itself at a lag, τ , time units previously.
- Changes with the lag
- Maximum occurs at zero lag
- Changes in correlation with lag characterize a signals time domain sequential structure.
- The power spectrum of a signal also describes its sequential structure - but in the frequency domain.

Auto-Correlation

- There are three variants of the auto-correlation:
  - the auto-correlation function
  - the auto-covariance function
  - the auto-correlation-coefficient function
- It common, but confusing, for all three forms to be referred to simply as the auto-correlation.

Auto-Correlation Function

$$R_{xx}(\tau) = E\left[x(t)x\left(t+\tau\right)\right]$$

- Value depends on both mean and variance of the signal
- Rarely used

Auto-Correlation Maxima

- Auto-correlation functions have their maxima at zero lag

$$\left|R_{xx}(\tau)\right| \leq R_{xx}(0)$$

Auto-Covariance Function

$$C_{xx}(\tau) = E\left[(x(t) - \mu_x)(x(t+\tau) - \mu_x)\right]$$

- Eliminates dependence on signal mean
- Variance as a function of lag
- Value still depends on signal amplitude

Auto-Covariance Function

Related to the auto-correlation function by

$$C_{xx}(\tau) = R_{xx}(\tau) - {\mu_x}^2$$

Zero lag covariance is equal to the signal variance

$$\mathrm{C}_{\mathrm{xx}}\left(0\right) = \sigma_x^2$$

Auto-Correlation Coefficient Function

$$r_{xx}(\tau) = \frac{C_{xx}(\tau)}{C_{xx}(0)}$$

$$= \frac{C_{xx}(\tau)}{\sigma^2}$$

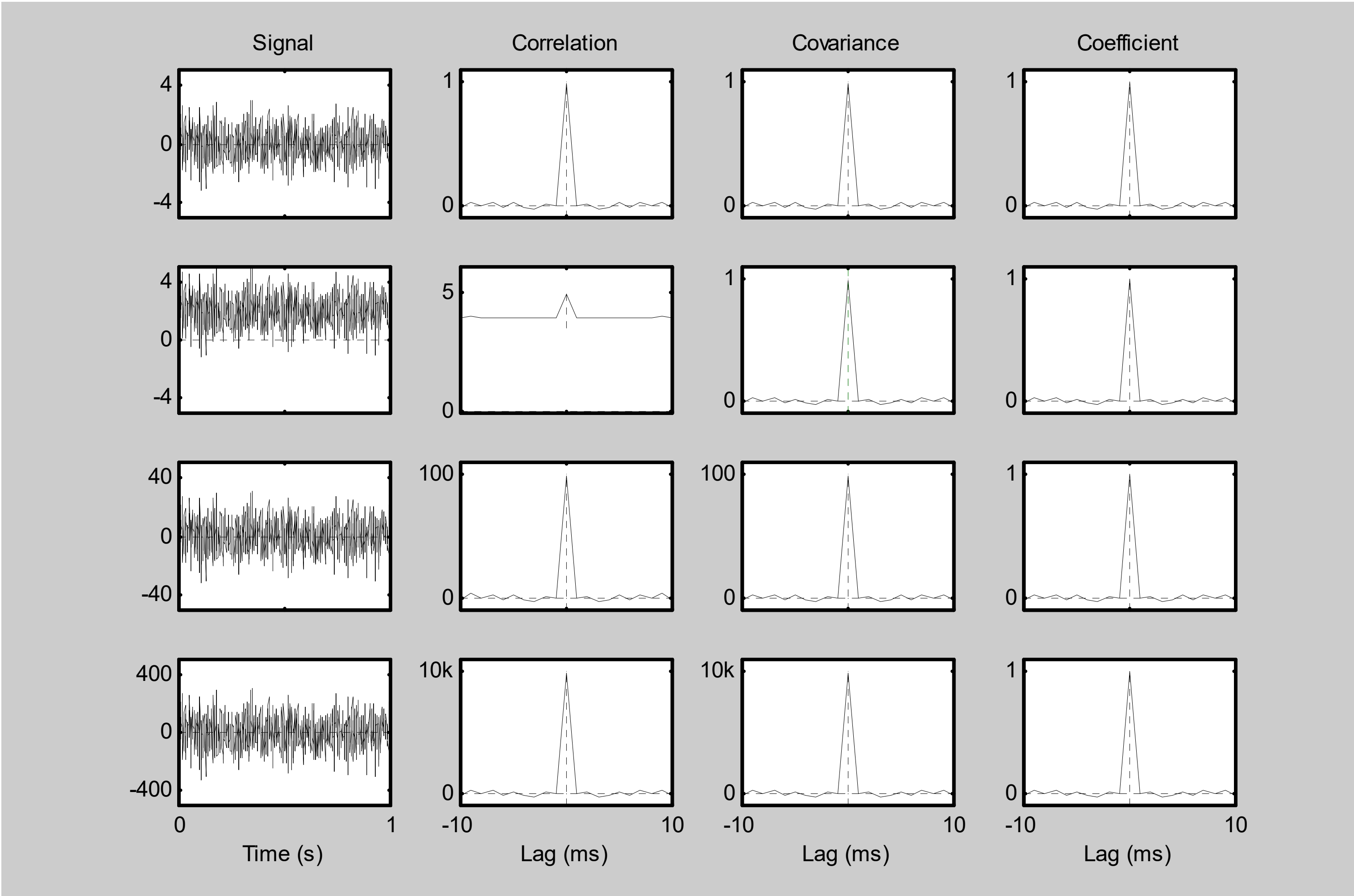
$$\left|C_{xx}(\tau)\right| \leq C_{xx}(0)$$

$$\Rightarrow \quad -1 \leq r_{xx}(\tau) \leq 1$$

$$\Rightarrow r_{xx}(0) = 1$$

- Does not depend on signal amplitude or mean
- Simplifies comparing the autocorrelation properties of signals with different amplitudes

Auto-Correlation Functions



Auto-Correlation Properties

- The auto-correlation coefficient function values are statistical correlations of the signal with itself:
  - +1 = completely positive correlation
  - 0 = no correlation
  - -1 = completely negative correlation.

Auto-Correlation Properties

- Auto-correlation functions are even functions

$$R_{xx}(\tau) = R_{xx}(-\tau)$$

- Auto-correlations are symmetric about 0

Auto-Correlation Properties

- The auto-correlation of a periodic signal is periodic with the same period

If: $x(t) = x(t + nT)$ for $n = 1, 2, \cdots n$

Then $R_{xx}(\tau) = R_{xx}(\tau + nT)$ for $n = 1, 2, \cdots n$

Auto-Correlation Properties

- The auto-correlation and power spectral density are Fourier transforms of each other

$$R_{xx}(\tau) = \int_{-\infty}^{\infty} \Phi_{xx}(f) e^{j2\pi f \tau} df$$

$$\Phi_{xx}(f) = \int_{-\infty}^{\infty} R_{xx}(\tau) e^{-j2\pi f \tau} d\tau$$

Auto-Correlation Properties

- The autocorrelation of the sum of two completely uncorrelated signals is the sum of the two autocorrelations:

If: x, y are uncorrelated and

$$z(t) = x(t) + y(t)$$

Then:

$$R_{zz}(\tau) = R_{xx}(\tau) + R_{yy}(\tau)$$

Auto-Correlation Properties

- The autocorrelation of a white noise signal has a peak at lag 0 and is 0 at all other lags.

$$\text{If: } x(t) \text{ is white}$$

$$\text{Then: } \quad \mathrm{R}_{xx}(\tau) = \delta(\tau)$$

$$\textit{where}$$

$$\delta(0) = 1$$

$$\delta(\tau) = 0 \quad \forall \text{ other values of } \tau$$

Cross-Correlation Function

$$R_{xy}(\tau) = E\left[x(t)\,y(t+\tau)\right]$$

- Rarely used because its values depend on both the means and variances of *x(t)* and *y(t)*.

Cross-Covariance Function

$$C_{xy}(\tau) = E\left[\left(x(t) - \mu_x\right)\left(y(t+\tau) - \mu_y\right)\right]$$

- Related to the cross-correlation function by

$$C_{xy}(\tau) = R_{xy}(\tau) - \mu_x \mu_y$$

- Does not depend on signal means
- Does depend on variances

Cross-Correlation Coefficient Function

$$r_{xy}(\tau) = \frac{C_{xy}(\tau)}{\sqrt{\left[C_{xx}(0)C_{yy}(0)\right]}}$$

- Normalized cross-correlation function
- Does not depend on
  - mean values
  - variances
- Useful for examining correlations between sets of signals with different amplitudes and/or mean values

Cross-Correlation Coefficient Function

$$r_{xy}(\tau) = \frac{C_{xy}(\tau)}{\sqrt{\left[C_{xx}(0)C_{yy}(0)\right]}}$$

If $\quad r_{xy}(0) = 1$

then $\quad x(t) = ky(t)$

Cross-Correlation Properties

- Cross-correlation coefficient function measures the statistical correlation between two signals as a function of the lag, τ, between them:
  - +1 = completely positive correlation
  - 0 = no correlation
  - -1 = completely negative correlation

Cross-Correlation Properties

- Cross-correlations are neither even nor odd, but

$$R_{xy}(\tau) = R_{yx}(-\tau)$$

- The magnitude is limited by that of the auto-correlations

$$\left|R_{xy}(\tau)\right| \leq \sqrt{R_{xx}(0)\, R_{yy}(0)}$$

Cross-Correlation Properties

- The cross-correlation and cross--power spectral density are Fourier transforms of each other

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} \Phi_{xy}(f) e^{j2\pi f\tau} df$$

$$\Phi_{xy}(f) = \int_{-\infty}^{\infty} R_{xy}(\tau) e^{-j2\pi f\tau} d\tau$$

Cross-Correlation Properties: Delays

- If y is a delayed, scaled version of x(t) with added noise

$$y(t) = \alpha\, x(t - \tau_0) + n(t)$$

- The, the cross-correlation function is the input auto-correlation scaled and displaced by the delay

$$R_{xy}(\tau) = \alpha\, R_{xx}(\tau - \tau_0)$$

- Maximum value occurs at lag equal to delay.

Cross-Correlation Estimation

- Sample signals at times

$$t = 0 \text{ to } t=(N-1)\Delta t$$

- Giving samples

$$x(i) \text{ and } y(i) \text{ for } i = 1, 2, \cdots N$$

- Unbiased estimate at discrete lag k is

$$R_{xy}(k) = \frac{1}{N-k} \sum_{i=1}^{N-k} x(i)\, y(i+k)$$

- Biased Estimate

$$R_{xy}(k) = \frac{1}{N} \sum_{i=1}^{N-k} x(i)\, y(i+k)$$

Effects of Noise

Practical measurements are noisy:

$$w(t) = x(t) + n(t)$$

$$z(t) = y(t) + v(t)$$

where:

$x(t), y(t)$ are the noise free signals of interest

$n(t), v(t)$ are independent noise signals

Noise & Auto-Correlation Estimates

$$w(t) = x(t) + n(t)$$

$$R_{ww}(\tau) = E\Big[\big(x(t)+n(t)\big)\big(x(t+\tau)+n(t+\tau)\big)\Big]$$

$$= R_{xx}(\tau) + \cancel{R_{xn}}(\tau) + \cancel{R_{nx}}(\tau) + R_{nn}(\tau)$$

$$R_{xn} \equiv R_{nx} = 0 \quad \text{but } R_{nn} \neq 0$$

$$R_{ww}(\tau) = R_{xx}(\tau) + R_{nn}(\tau)$$

Noise & Auto-Correlation Estimates

Bias Term

$$R_{ww}(\tau) = R_{xx}(\tau) + R_{nn}(\tau)$$

- Autocorrelation estimates are biased by noise

Noise & Cross-Correlation Estimates

$$w(t) = x(t) + n(t)$$

$$z(t) = y(t) + v(t)$$

$$R_{wz}(\tau) = E\left[w(t)z(t+\tau)\right]$$

$$= E\left[\left(x(t)+n(t)\right)\left(y(t+\tau)+v(t+\tau)\right)\right]$$

$$= R_{xy}(\tau) + \cancel{R_{xv}}(\tau) + \cancel{R_{ny}}(\tau) + \cancel{R_{nv}}(\tau)$$

$$R_{xv} \equiv R_{ny} \equiv R_{nv} \equiv 0$$

$$R_{wz}(\tau) = R_{xy}(\tau)$$

Noise & Cross-Correlation Estimates

$$R_{wz}(\tau) = R_{xy}(\tau)$$

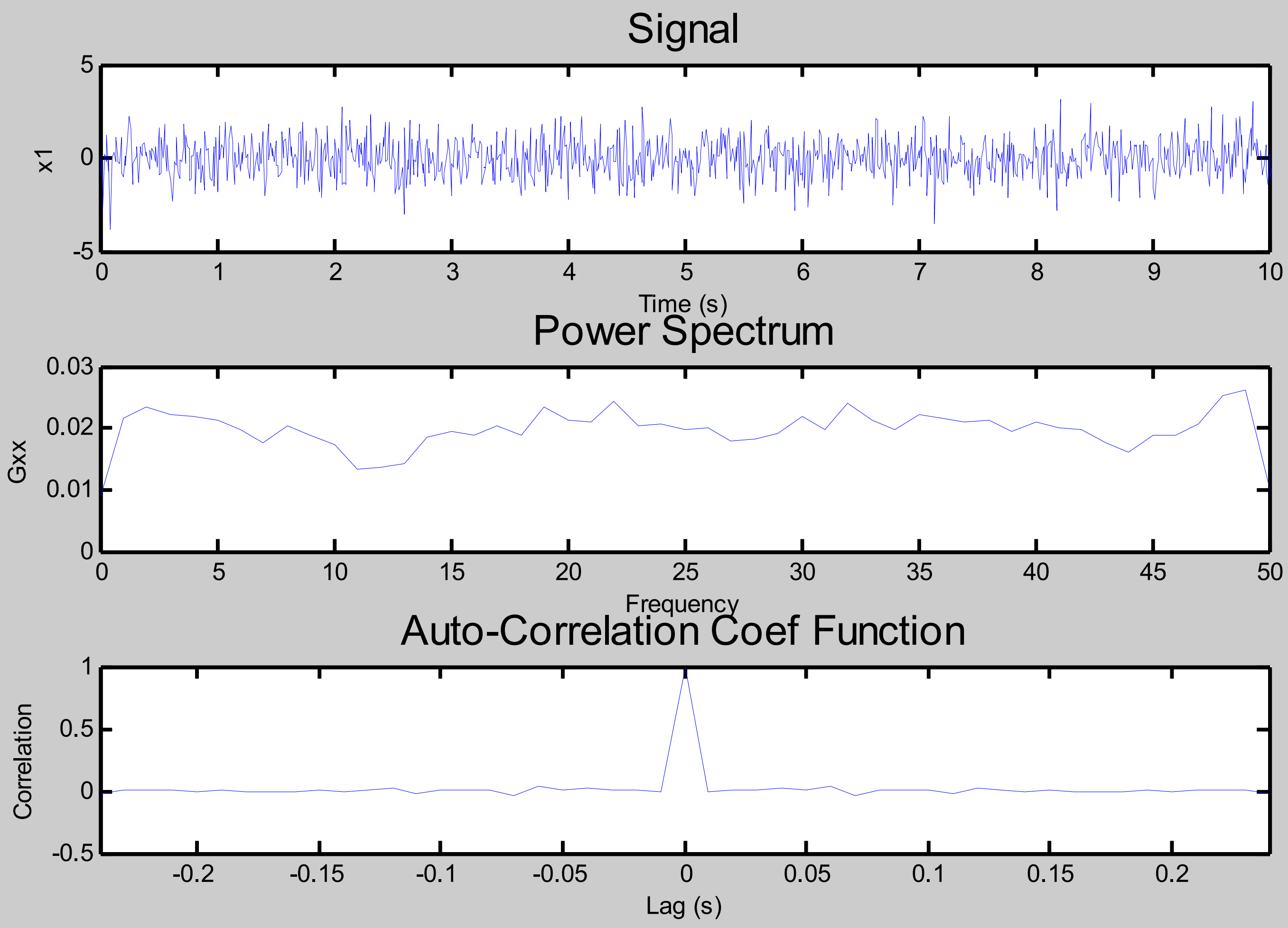
- Cross-correlation function estimates are not biased by noise.

Data Lengths

- Correlation functions should be calculated for lengths that are much shorter than the data length
- It is desirable to calculate correlation functions at lags no greater than ¼ of the data length.
- It is not acceptable to calculate correlation functions at lags greater than ½ the data length

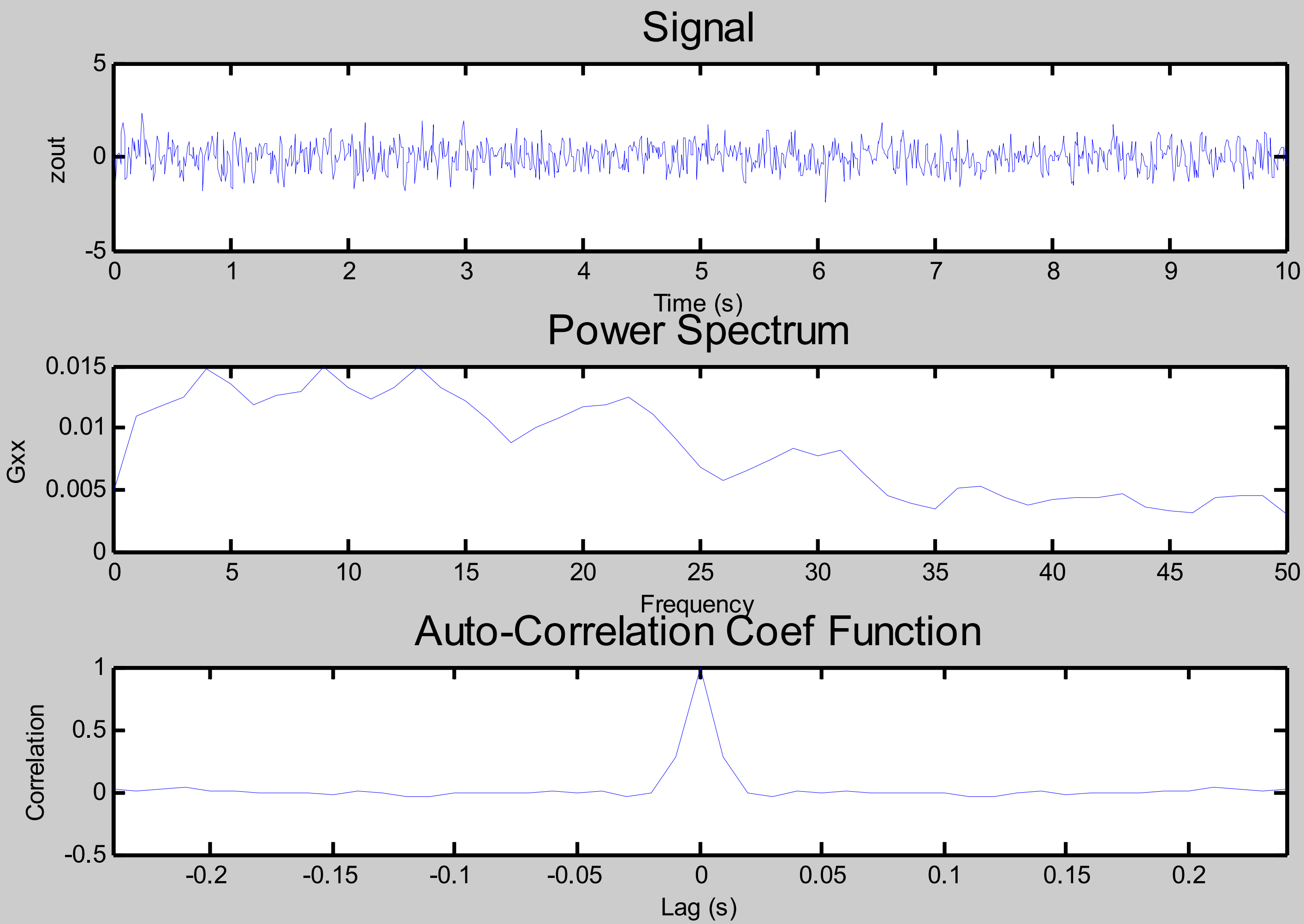
Auto-Correlation Examples

White Noise



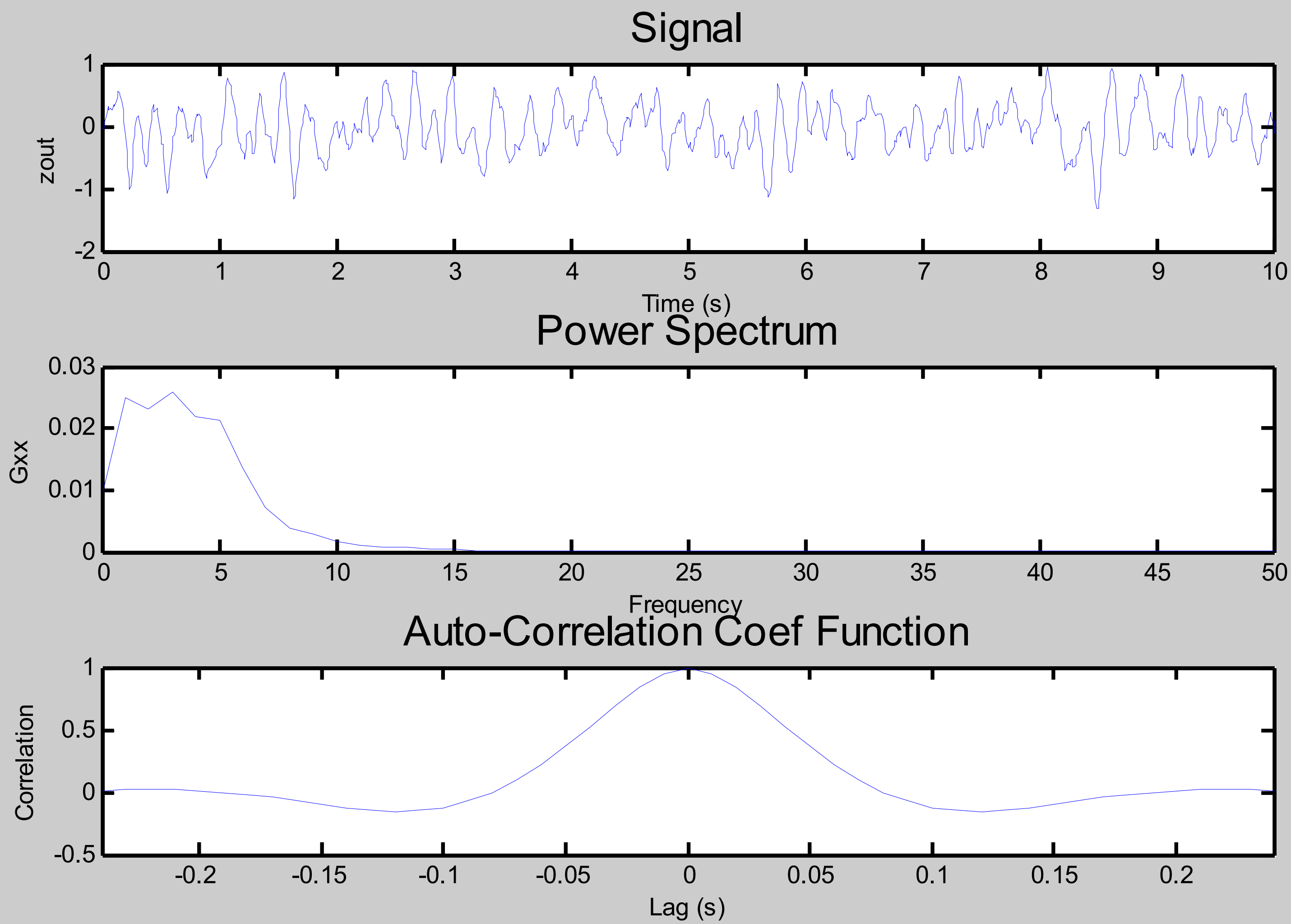
White Noise

Lightly Filtered Noise



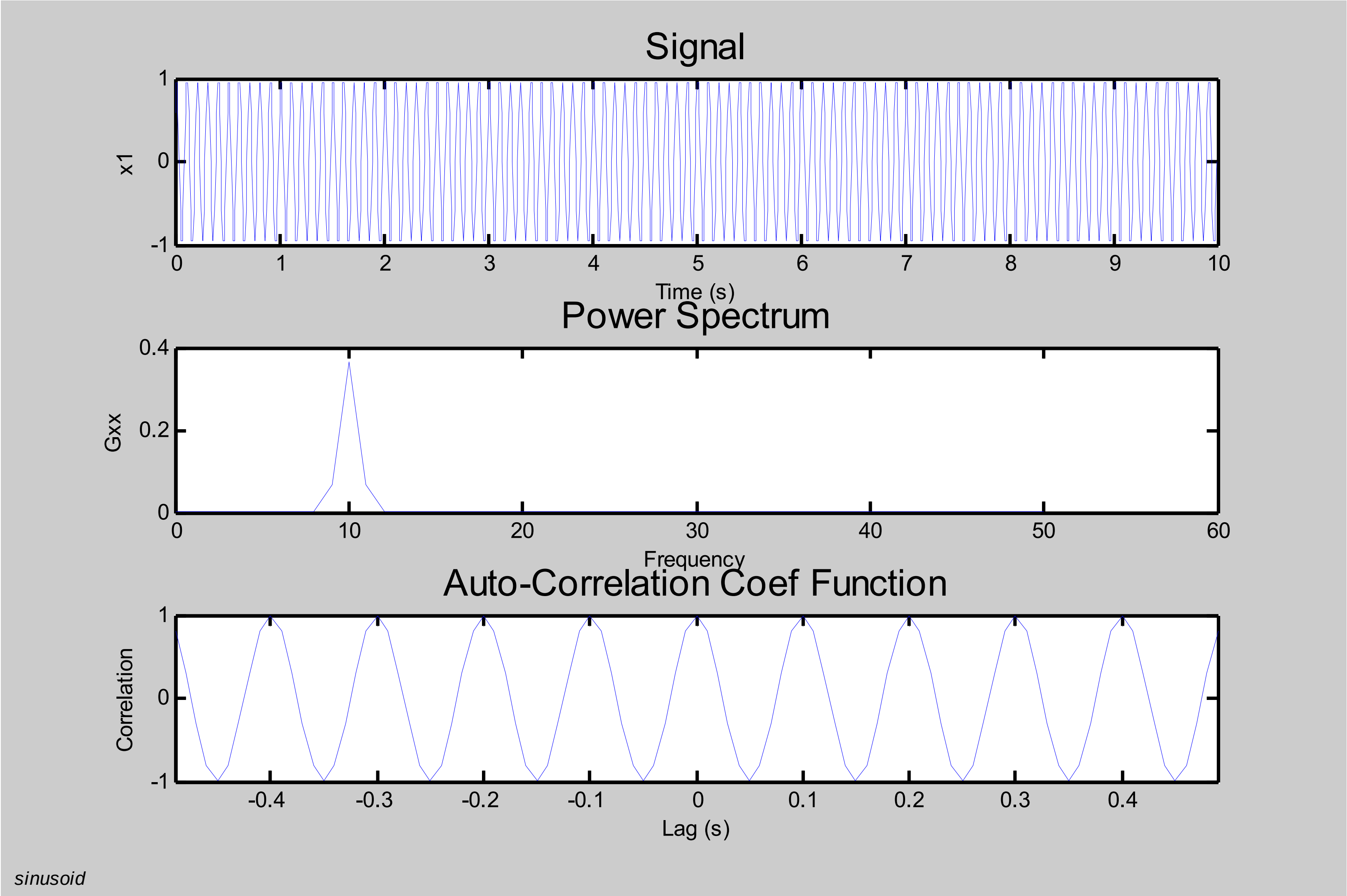
Filtered Noise

Heavily Filtered Noise

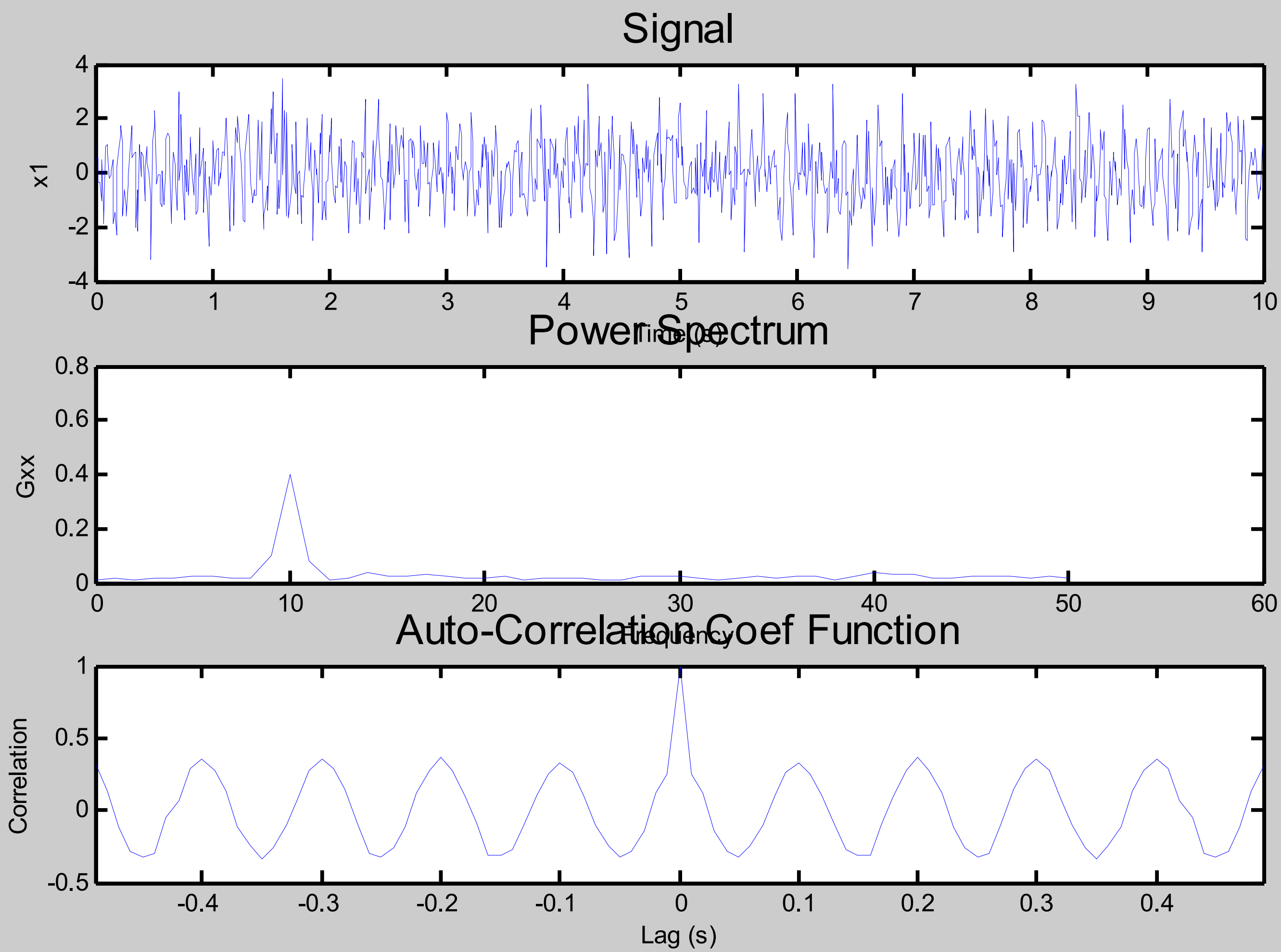


Filtered Noise

Sinusoid

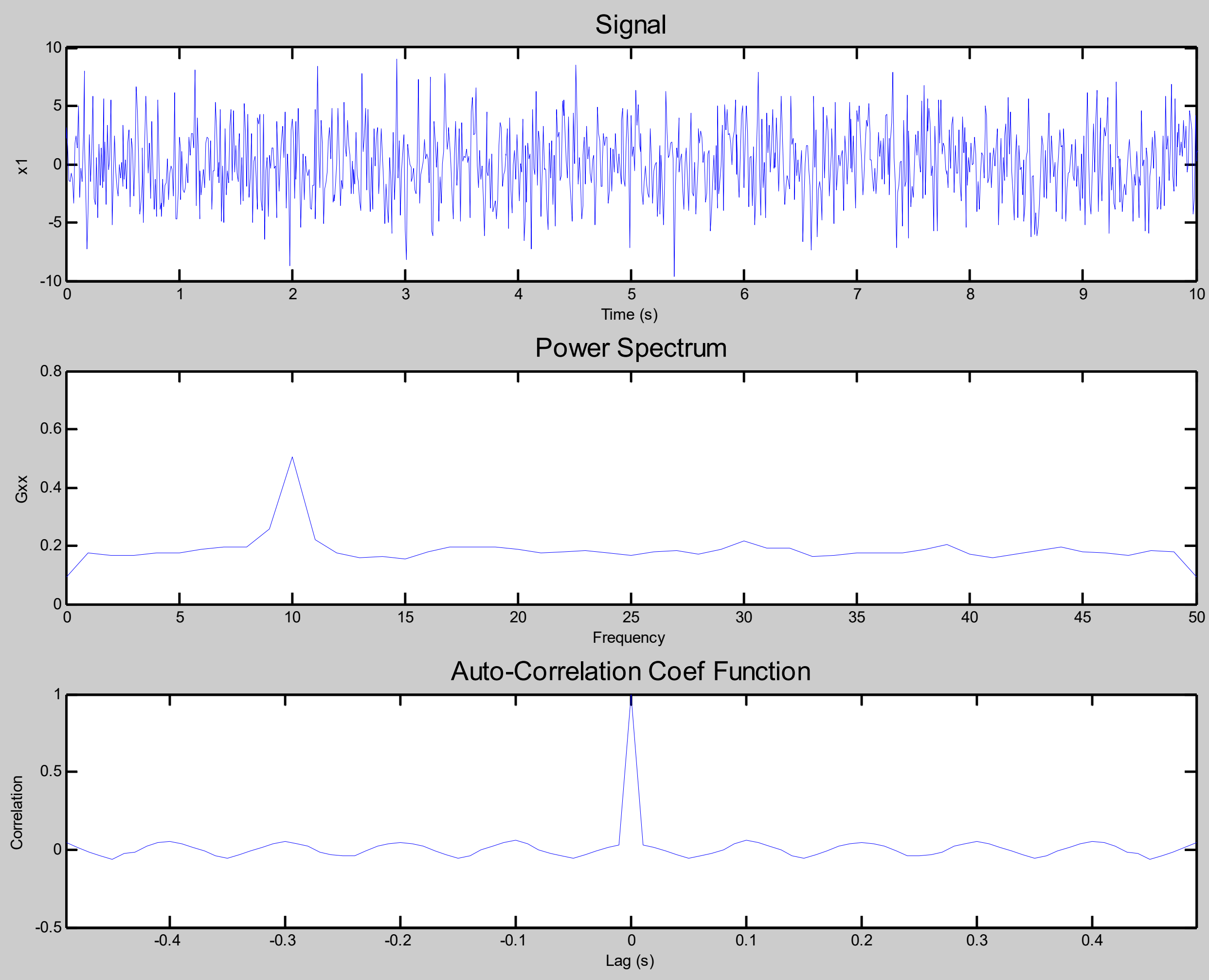

Signal
x1
1
0
-1
0 1 2 3 4 5 6 7 8 9 10
Time (s)
Power Spectrum
Gxx
0.4
0.2
0
0 10 20 30 40 50 60
Frequency
Auto-Correlation Coef Function
Correlation
1
0
-1
-0.4 -0.3 -0.2 -0.1 0 0.1 0.2 0.3 0.4
Lag (s)
sinusoid

Sinusoid + low noise



Sinusoid + low noise

Sinusoid + large noise



Sinusoid + high noise

Auto-Correlation Applications

- Auto-correlation functions are often used to assess periodicity in a signal.
  - A periodic component hidden in a noisy signal will appear in the auto-correlation function with the same period.
- Auto-correlation functions of stochastic signals tend to “die out" as the lag increases.
  - The lag at which the auto-covariance function drops to zero can be is a measure of the process “memory”

Cross-Correlation Applications

- Cross-correlation functions measure the sequential relation between two signals.
  - Template matching
    - i.e. spike detection
- Two signals may each have considerable sequential structure but not be correlated

Cross Correlation Applications

- Often used use to determine the value of the delay between two signals.
- The delay is given by the lag at which the maximum value of the cross-covariance function occurs.
  - Comparing emitted and reflected signals can be used for source location
    - Bagpipes in an apartment building
    - 2D cross correlation often used for image registration
    - Submarine echo location

Cross Correlation Applications

- Cross correlation of the same signal measured at different locations can be used to estimate velocity.
- Applications include:
  - Bladder function
  - Peripheral nerve measurements
  - Fish migration

Cross Correlation Applications

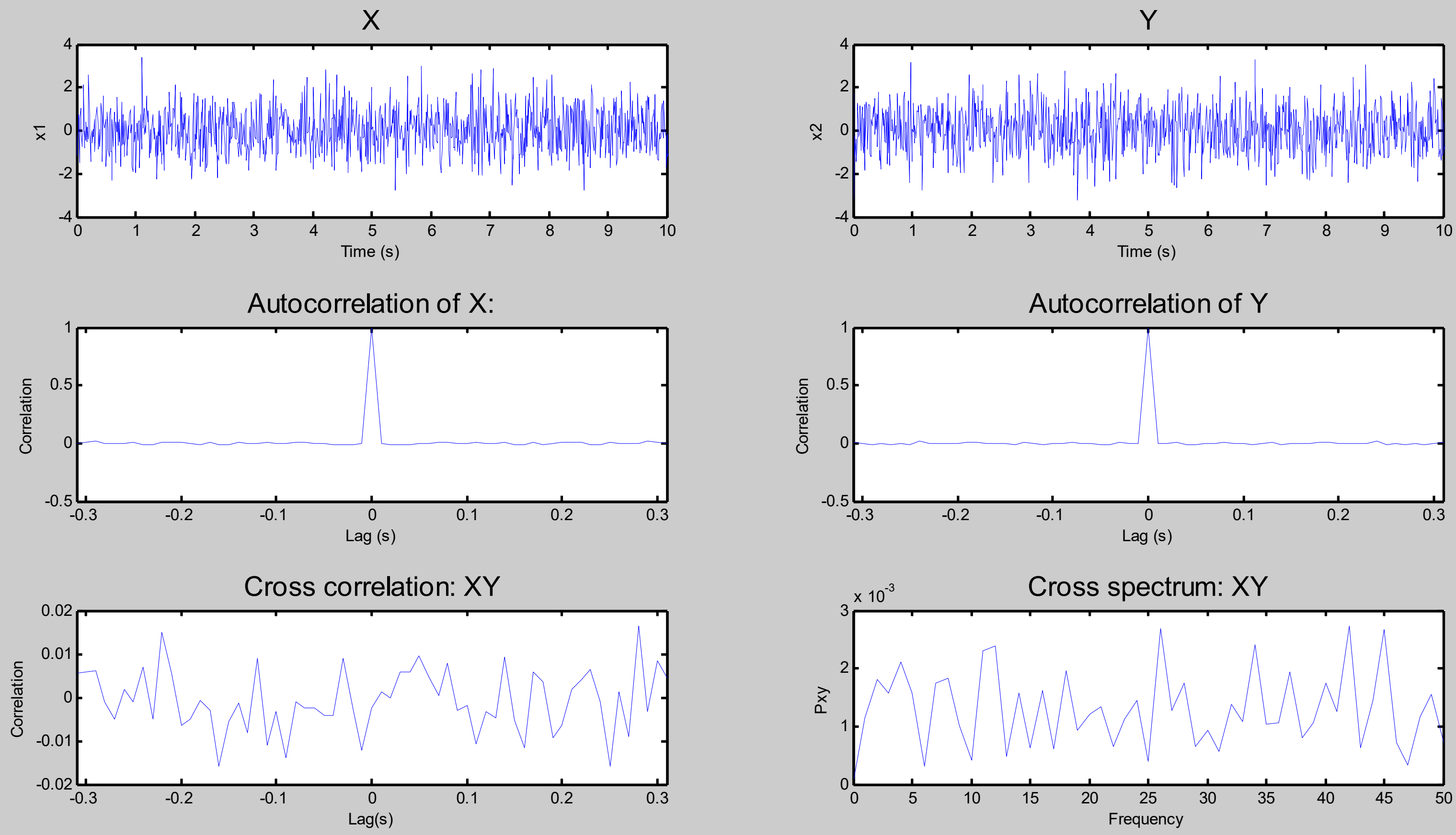
- Not all cross-correlation peaks are due to conduction delay
- The dynamic relation between input and output may result in delayed peaks in the cross-correlation function

Correlation is not Causation

- The presence of a correlation between two signals does not imply they are causally related
- Many signals that are highly correlated but not related
  - Common inputs
    - Correlation between IQ of post- World War II children and the number of teeth
  - Happen chance: http://www.tylervigen.com/

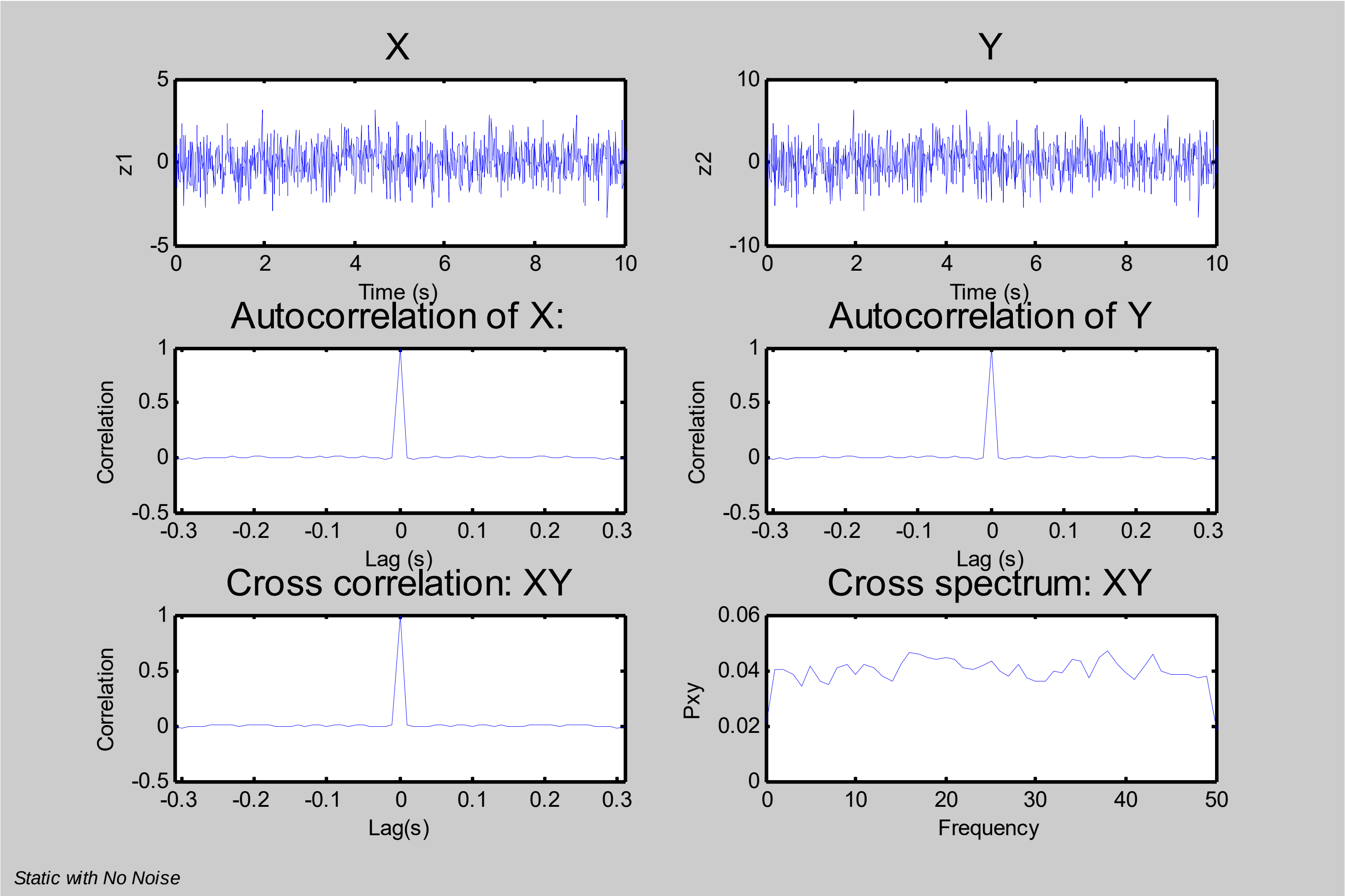
Cross Correlation Examples

No Correlation

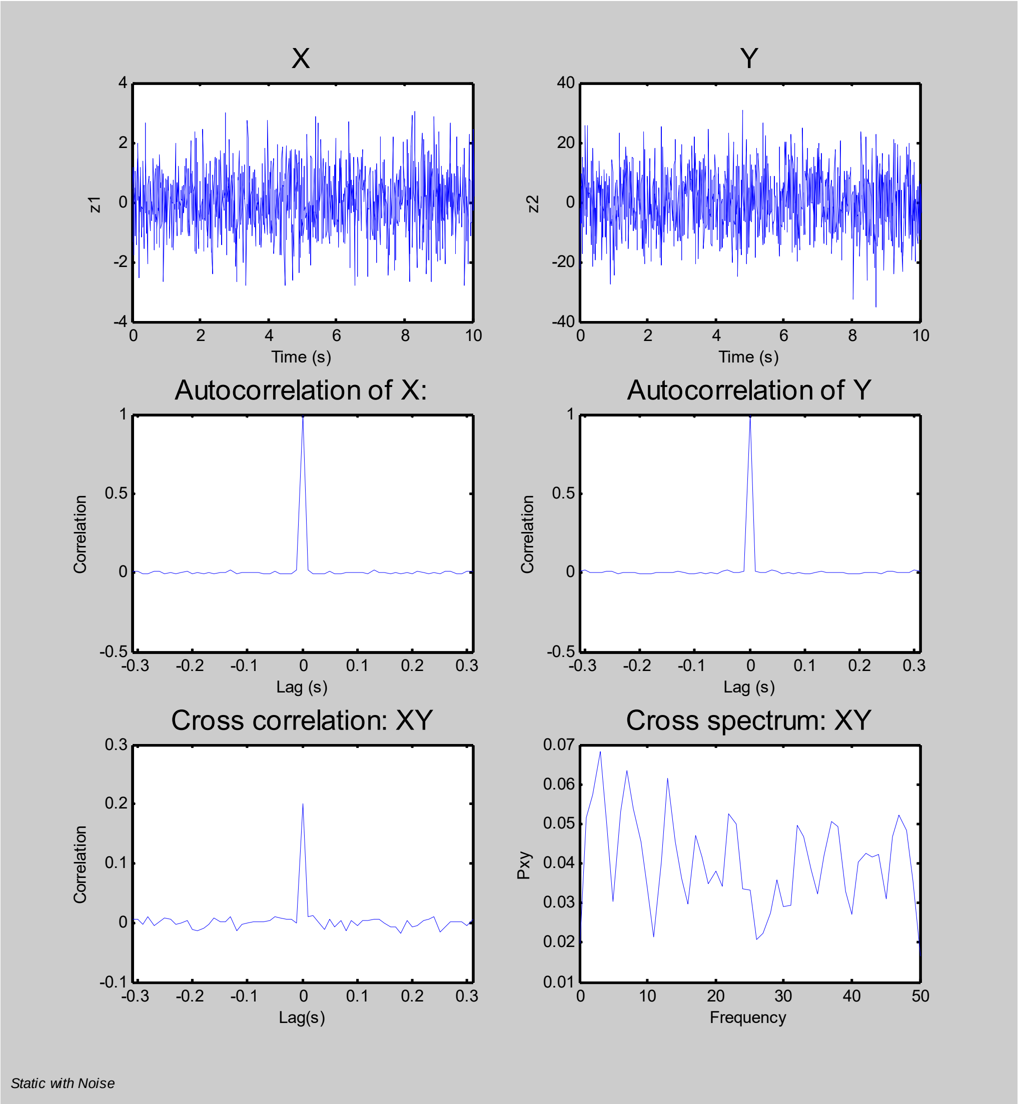


No Correlation

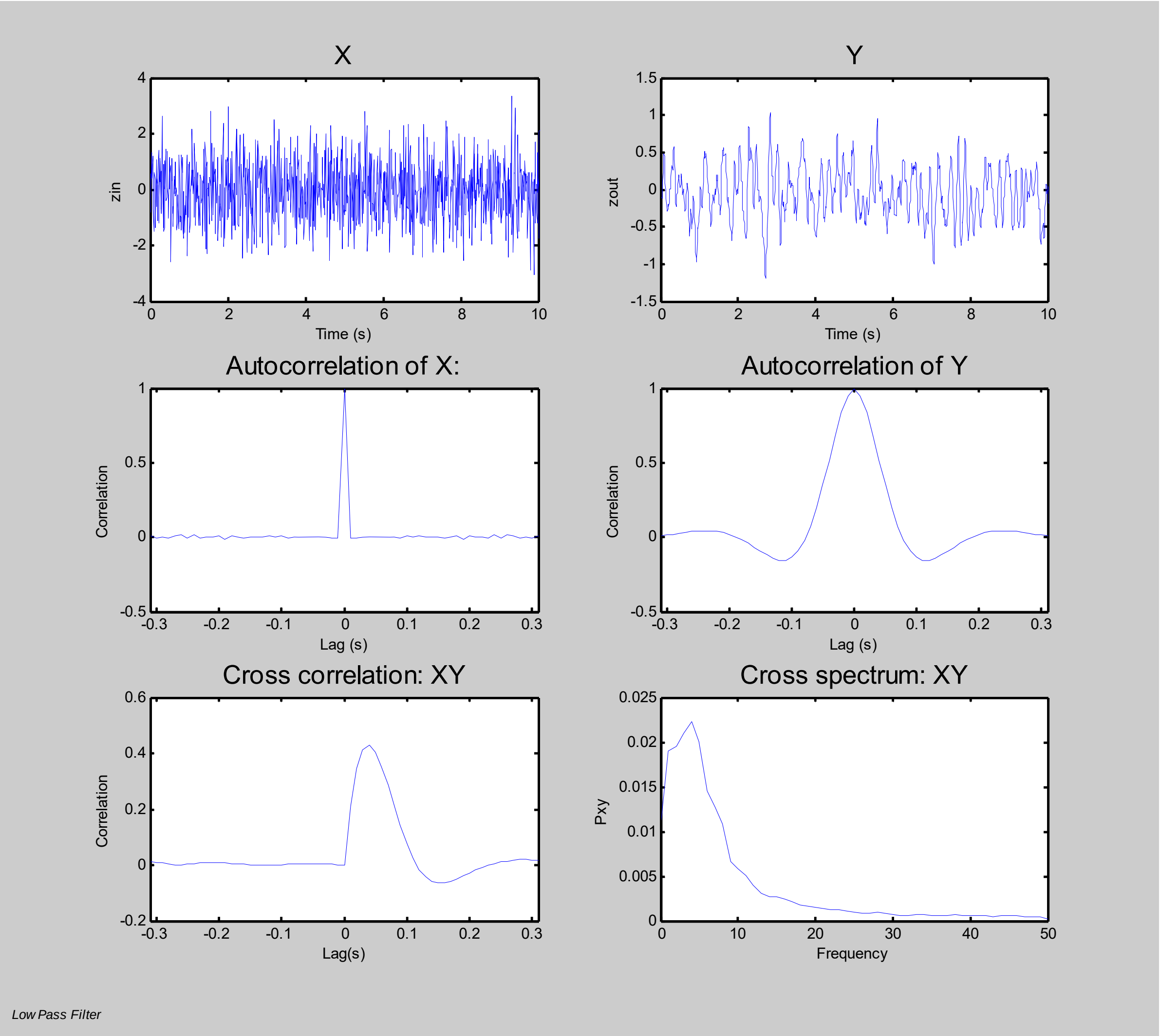
Static Relation: $y = k * x$



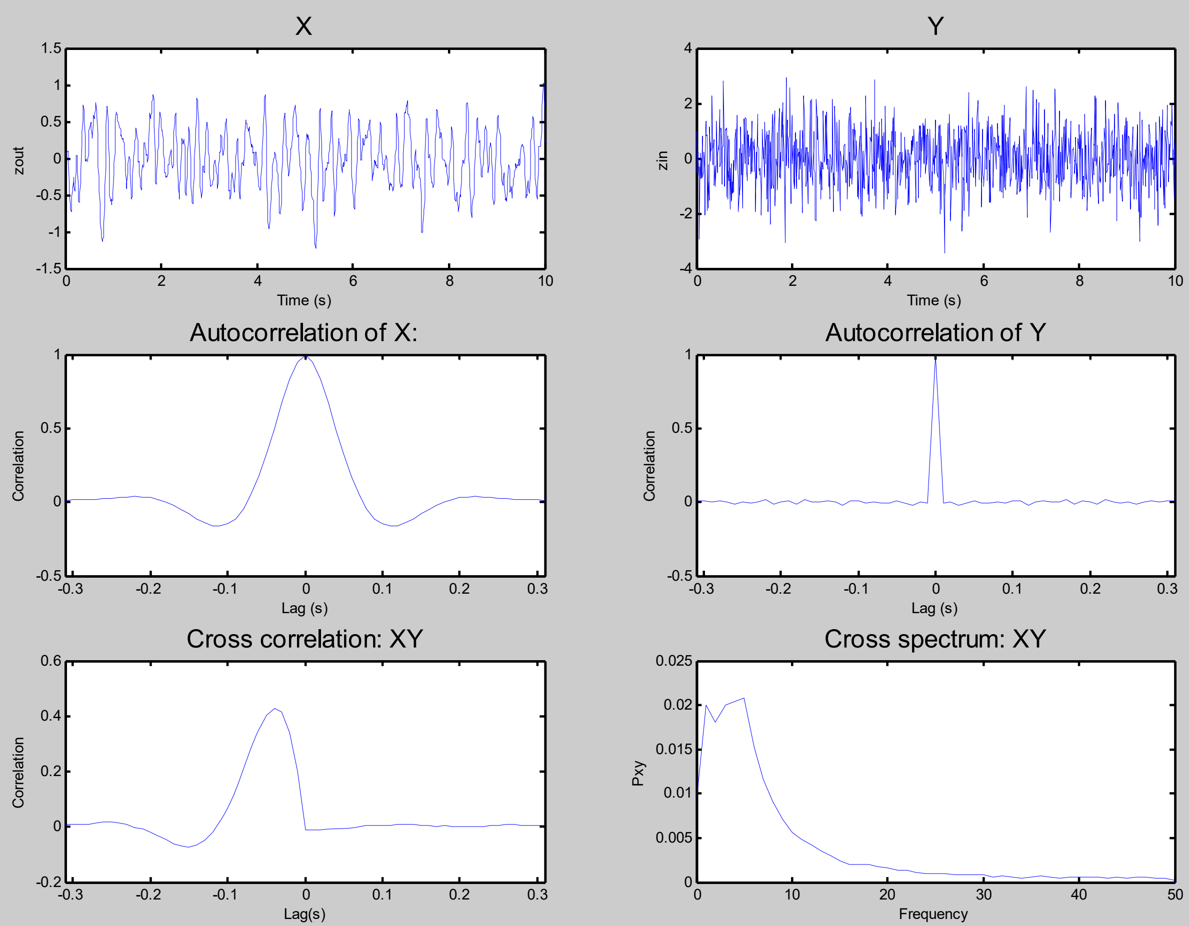
Static Relation with Noise $y = k * x + e$



Low Pass Filter

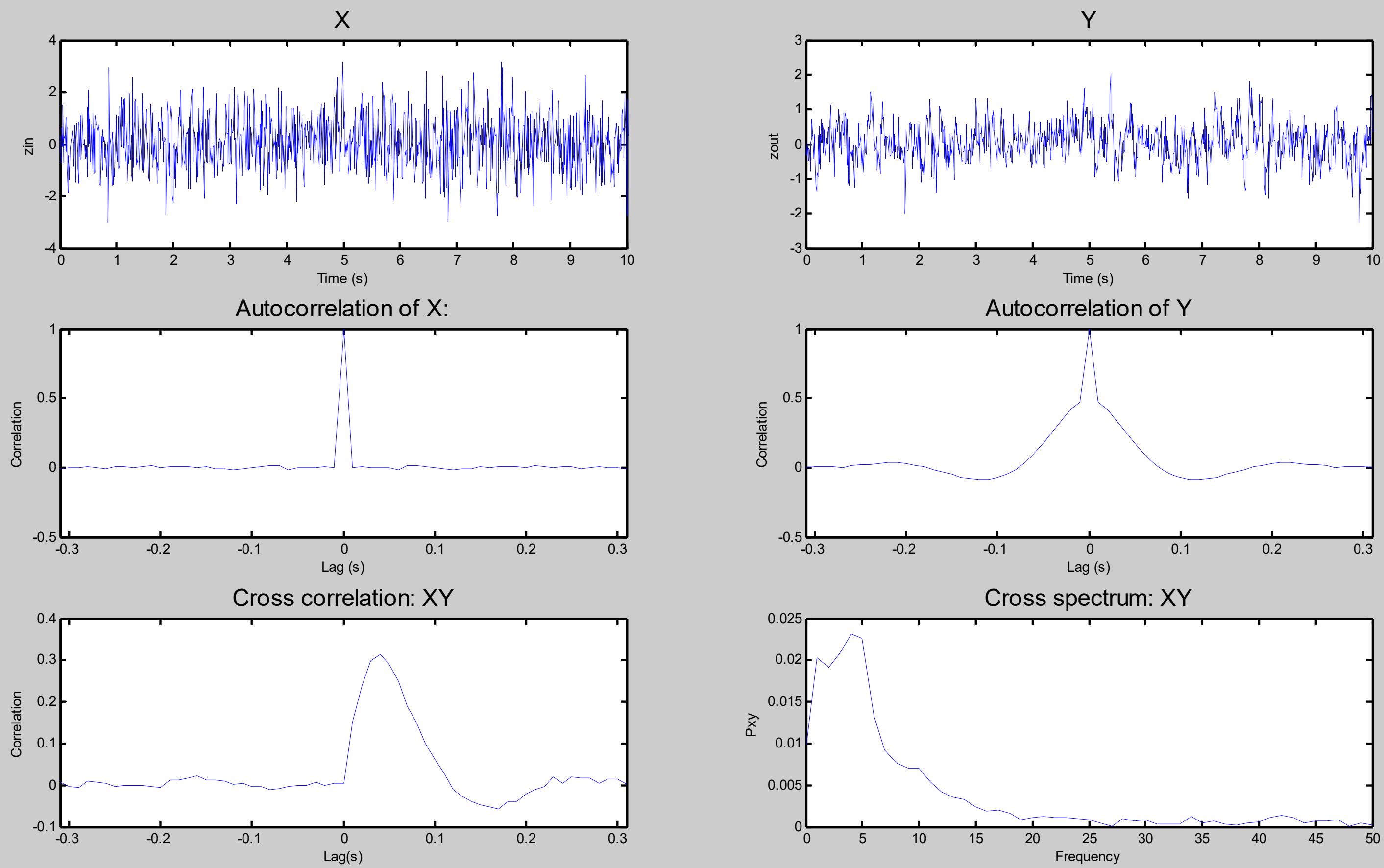


Low pass filter: input & output reversed



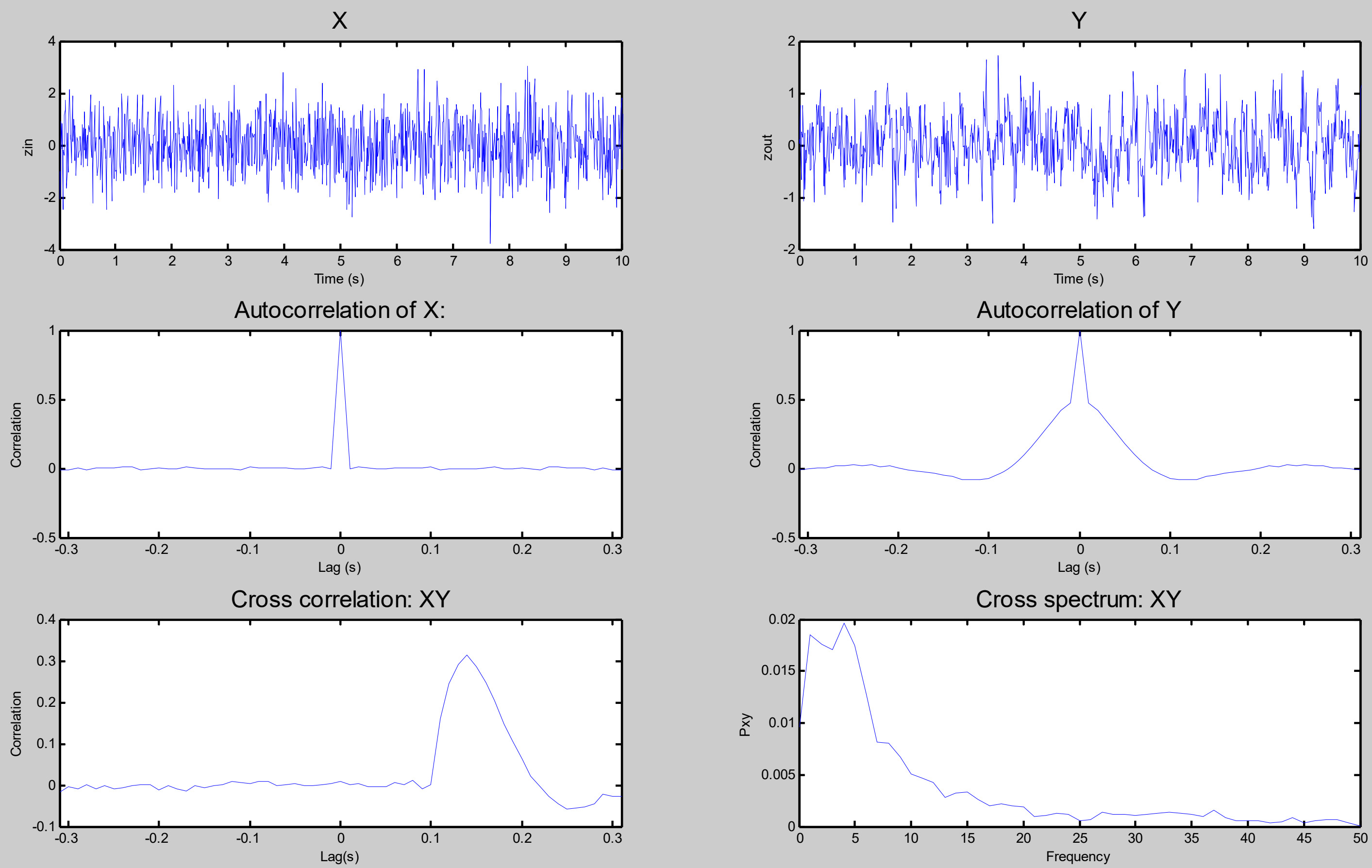
Low Pass Filter with input and output reversed

Low Pass Filter with Noise



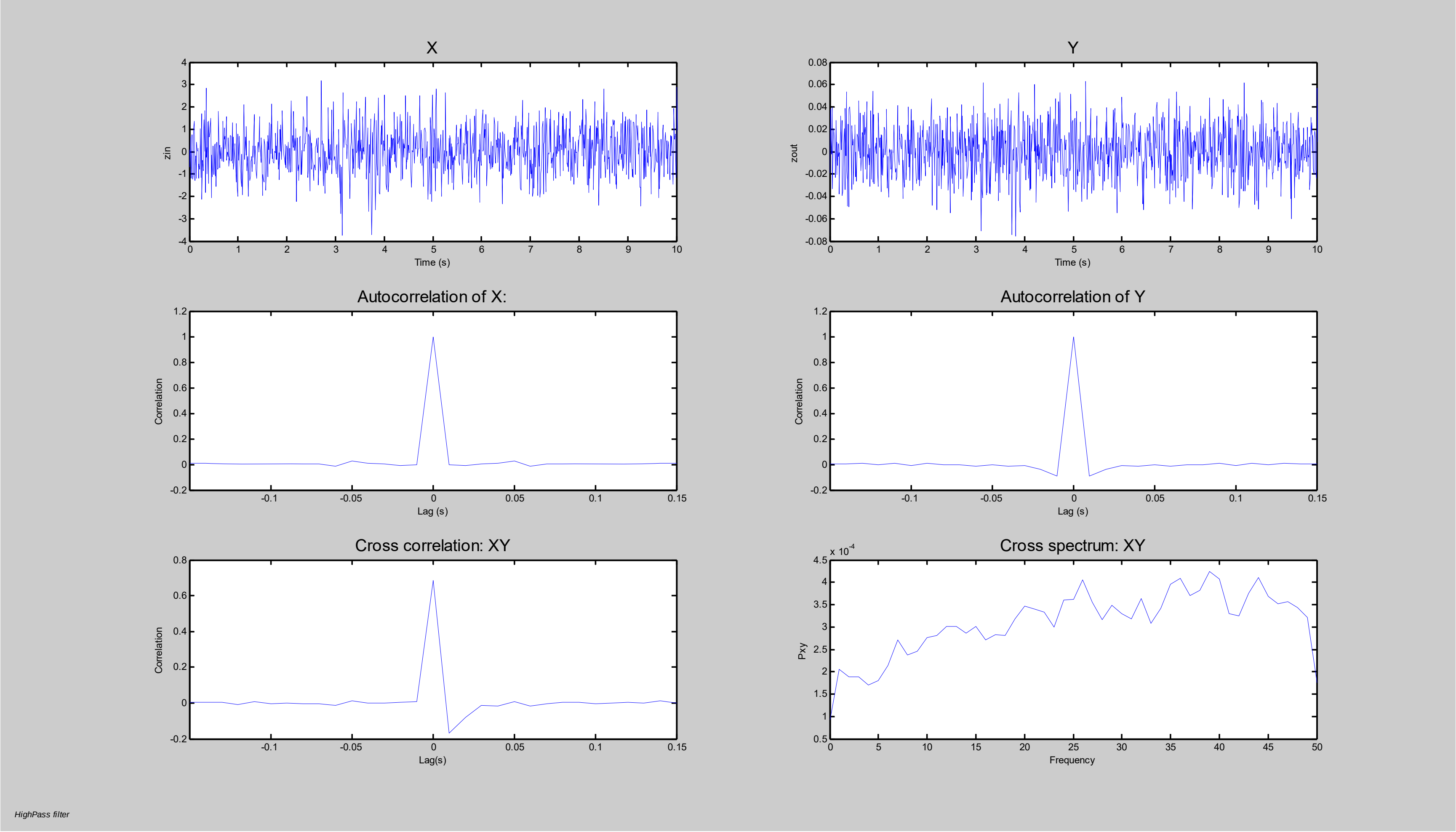
Low Pass Filter with Noise

Low pass filter with noise + delay



Low Pass Filter with Noise & Delay

High Pass Filter



Assignment-1

- Load the data set edu*519m6* from MyCourses.
- This data set contains the test signals x1,x2, …., x10.
- Signals x1,…x5 have two columns with the domain and range information.
- Signals x6,…x10 have three columns containing domain and range values for input and output signals from some unknown system.

Assignment-2

1. Write a matlab function called *mycorel* which calculates the biased auto-correlation-coefficient function using a relation similar to that given in the notes for the cross-correlation function.

- Demonstrate its use with a sample of normal, white data having a mean of 2 and standard deviation of 5.
- (2.5/10)

Assignment-3

- Signals x1 -> x5.
  - Compute and plot the auto-correlation-coefficient function for each signal.
  - What can you deduce about each signal from your results?

  - (3.75/10)

Assignment-4

3. Signals x6 -> x10.
   - Compute and plot
     - the auto-correlation-coefficient functions for the input and output signals
     - the input-output cross-correlation coefficient function
     - Plot both sides of the correlation functions

- What can you deduce about each of these signal pairs from your results.

- (3.75/10).

Assignment-5

- Present your results in the form of a concise report of no more than 10 pages.
- Append a copy of your correlation function to your report.
- Make use of the matlab function xcov or xcorr. Take care to interpret your results in terms of how the lags are used in this function.
- Choose the length of your correlation functions carefully.