EE 422G - Signals and Systems Laboratory

Lab 5 Filter Applications

Kevin D. Donohue
Department of Electrical and Computer Engineering
University of Kentucky
Lexington, KY 40506
October 3, 2015

Objectives:

- Apply knowledge of signal and noise properties to design filters to enhance signal detection in noise.
- Apply analysis tools and experimental techniques to verify filter performance.

1. Background

An ultrasonic A-scan is a signal created by sending a pulse of high frequency sound into an object and recording back-scattered energy. This pulse-echo principle is similar to what is done in radar or what a bat does to navigate. The timing of the return echoes are used to image or locate the scatterers in the insonified field. In non-destructive evaluation (NDE) this pulse-echo ultrasonic technique is used to scan materials for internal flaws, such as cracks and other defects without having to cut (destroy) the material for inspection. Figure 1 shows examples of A-scans from a stainless rod. Stainless steel is composed of grain-like structures or crystals on the order of 0.1 mm with random distributions of sizes, shapes, and orientations throughout the material. These microstructures scatter portions of insonifying energy from the propagating ultrasonic pulse back to the receiving element (the piezoelectric properties of ultrasound transducer element allow for the same element to be used to generate and receive ultrasonic energy). The grain-like structures throughout the material result in backscattered energy appearing over the duration of the entire A-scan. Since this *noise* results from a process in the scanned environment it is sometimes referred to as *process noise*. This is in contrast to artifacts and variations from the instrumentation, which is referred to as *system noise*).

Figures 1a and 1b show A-scans from insonified flaws simulated by flat-bottom holes 4mm in diameter. Figure 1a shows a case where the flaw echo is clearly stronger than the surrounding scattered grain echoes. Figure 1b shows another A-scan where surrounding scattered grain echoes are stronger than the flaw echo. While the individual grains echoes are much smaller and weaker than the flaw echo, the large number of grain structures distributed throughout the volume sometimes result in an in-phase (coherent) addition to create strong echoes. The coherent addition of grain scattering results in amplitude variations dependent on the relative positions of the grain scatterers and the wavelength of the illuminating energy. Since the factors contributing to this complex process cannot be practically known, the grain scattering is modeled statistically as a random noise process.
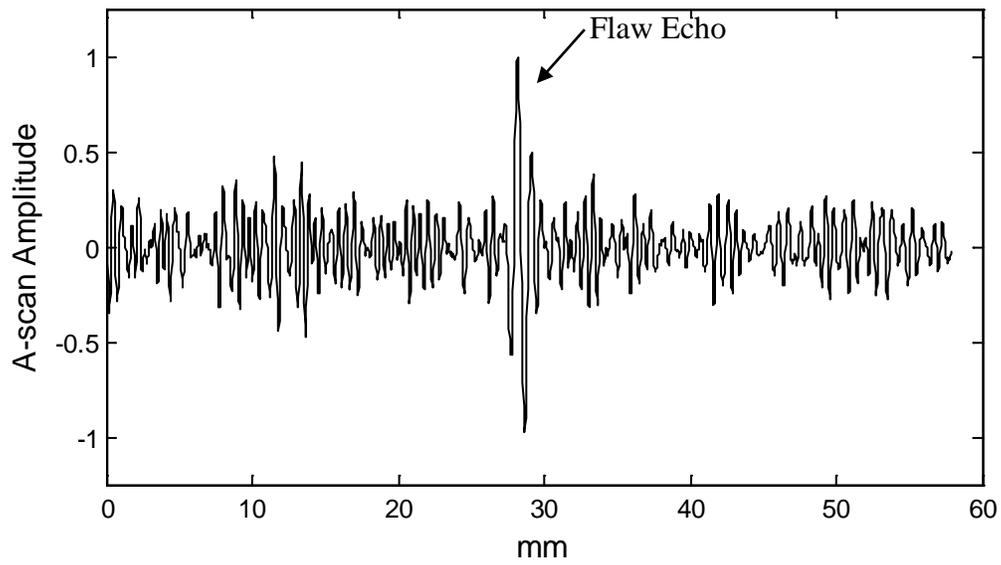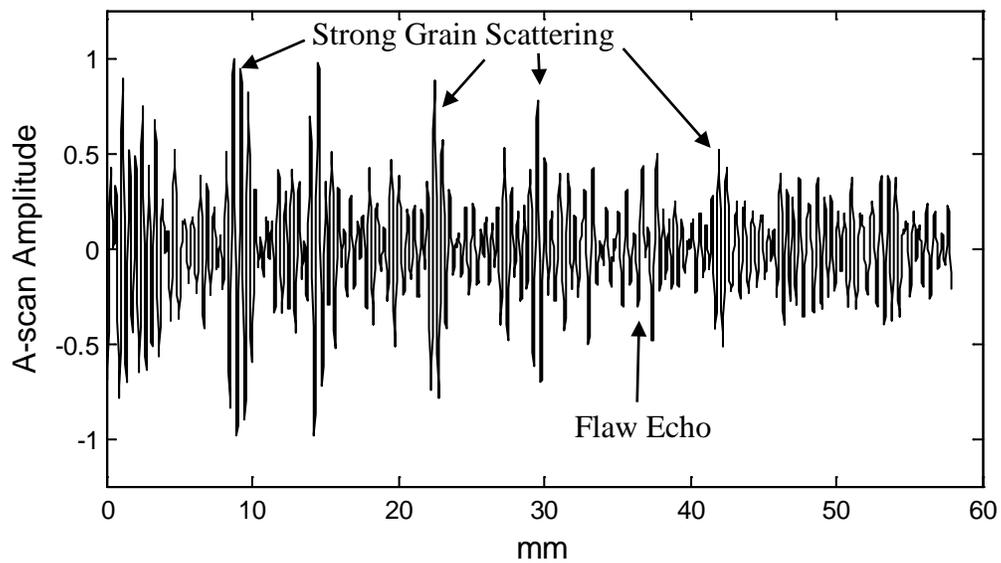
Figure 1. Examples of A-scans from stainless-steel samples with flaws simulated by a drilled 4mm flat-bottom hole. A 5 MHz transducer was used to create the insonifying pulse and receive the scattered energy. (a) Case where flaw echo is stronger than surrounding backscattered energy from the grain structures. (b) Case where flaw echo is similar to or weaker than surrounding backscattered energy from grain structures.

Figure 1 illustrates that the echo strength alone is not sufficient to detect a material defect. There are, however, spectral differences between flaw and grain scattered energy that can be exploited through filtering. While scattering strength is directly proportional to difference in density and/or elasticity at material boundaries, there is also a frequency sensitivity related to scatterer size. If the scatterer boundary is large with respect to the wavelength (sometimes referred to as optical scattering), energy at all wavelengths reflect with significant strength resulting in a strong echo. Alternatively, if the scatterer boundary is small with respect to the wavelength, weak scattering occurs (sometimes referred to as Rayleigh scattering). In this case, the long wavelengths (low frequencies) tend to pass through the small scatterers with little energy loss. So low-frequency signal components will exhibit weaker scattering than higher frequency components for smaller scatterers. (Similar reasoning applies for optics in explaining why the sky is blue and sometimes red at sun rise/set).

In the case of the ultrasound scans in Fig. 1, the insonifying pulse with center frequency of 5 MHz corresponds to a wavelength of 1.2 mm (assuming a sound speed of 5790 m/s in steel), which is smaller than the 4.11mm flaw scatterer. The grain structures, however, are on the order of 0.1 mm, which is an order of magnitude smaller than the center frequency wavelength. This wavelength to scatterer size relationship is in the Rayleigh scattering region, which exhibits significant frequency sensitivity to scatterer size. The insonifying pulse has a bandwidth of about 4MHz, which corresponds to a frequency range from 3 MHz to 7 MHz. This frequency range corresponds to wavelengths ranging from 2mm to 0.8mm. So it is expected that grain echoes will scatter energy from the upper end of the transducer spectrum with greater strength than from frequencies at the lower end. The flaw, one the other hand, scatters energy from the full spectrum of the transducer. Another factor that impacts the frequency distribution of the received energy is the grain scatterers cause the propagating pulse to progressively lose higher-frequency energy at a greater rate than the lower frequencies. Therefore, high frequencies in the propagating pulse are attenuated more so than the lower frequencies. So it is expected that the received signal from scatterers at greater depth have a low-pass emphasis due to propagation effects.

Figure 2 illustrates the spectral differences between the grain and the flaw echo. The average spectra or power spectral densities (PSDs) for the A-scans of Figure 1 are plotted. The PSD is computed with the PWELCH function in Matlab, which takes the squared FFT magnitude of small overlapping segments of data from the whole segment and averages them together. This is sometimes called the *hopping-window* approach or Welch's methods for spectral estimation from random processes. The A-scans consist of 2000 samples each, sampled at 100MHz. For the spectra in Fig. 2, a hopping window size of 128 samples was used, with 64 points of overlap between adjacent windows. A tapering window was used (a hamming window) and the FFT length is increased through zero padding to obtain a 256 point FFT (double the actual number of data points). Spectra for the flaw were more tedious to obtain since the flaw echoes occur only at a limited location in the scans. The beginning and ending sample points around the flaw echo were obtained by looking at the plot and extracting out the flaw section only for the FFT, and the magnitudes were averaged together over several A-scans (padding with

zeros helped to keep the FFT length the same to the vectors could be averaged together point by point from the different scans).

Sample Matlab code can be used to compute and plot spectra for the grain. The Matlab variable *ac1* is a vector containing the A-scan points.

```
fs = 100e6;   % Sampling frequency
wl = 128;     %   Hopping Window Length
nfft = 2*wl;  % Number of FFT points
wolap = fix(wl/2);   %   Number of overlapping points in hopping window
%   Apply the hopping window method to estimate spectrum
[p,f] = pwelch(ac1,hamming(wl),wolap,nfft,fs);
figure(1) %   Plot the resulting PSD, but scale to get power in each bin
 %   Divide Frequency axis by 1e6 to get Units in MHz
plot(f/1e6,(p*fs/wl),'k')
xlabel('MHz')
ylabel('PSD Magnitude')
set(gca,'Xlim', [0 10]) %   Zoom in on 0 to 10 MHz on X-axis.
```

For the flaw spectra, segments dominated by flaw echo energy were too small to use *pwelch()*, so the FFT magnitudes from segments containing the flaw were tediously extracted (with knowledge of where the flaw was located and graphic inspection to get the beginning and ending sample points) and averaged together.
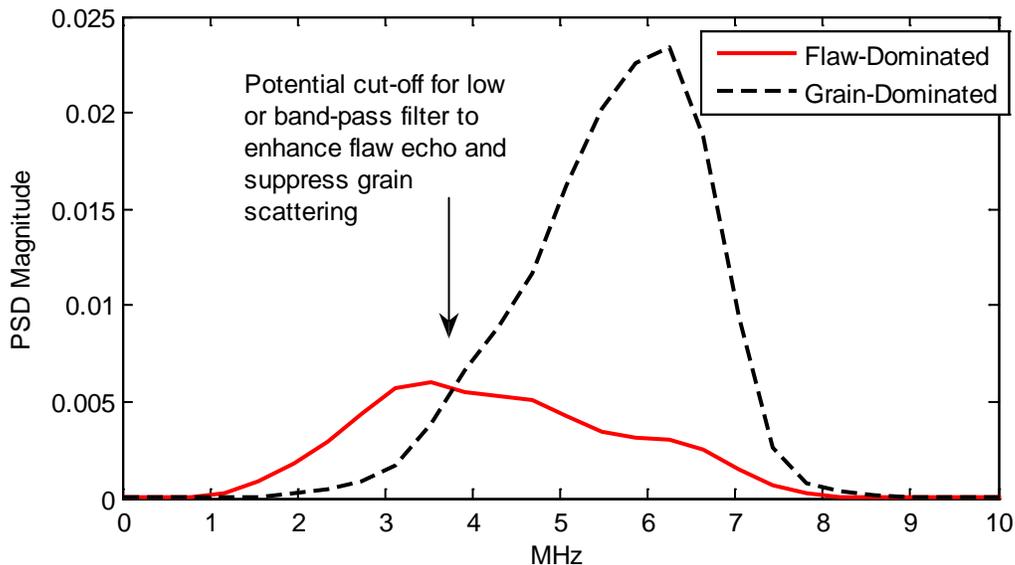

Figure 2. Spectral comparison between flaw and grain dominated A-scans of Fig. 1.

Figure 2 shows the grain echoes are characterized by an emphasis on the higher frequencies of the transducer bandwidth, while the flaw echoes are characterized by an emphasis on the lower frequencies. The grain and flaw structures effectively filter the backscattered energy based on the size and wavelength relations. Figure 2 suggests that either a band-pass or low-pass filter with an upper cut-off frequency around 3.6 MHz would help suppress grain scatterer energy relative to the flaw echo energy.

Based on the spectra of Fig. 2, a band-pass filter was designed with upper and lower cut-off frequencies of 1.5MHz and 3.2MHz, and applied to the A-scan of Fig. 1b. The filtered result is plotted in Fig. 3. The filtered output in Fig. 3a shows the flaw echo dominating the A-scan making it detectable (with no false detections) using a simple threshold. There is a slight delay or shifting toward the right of the flaw's original position due to the filter delay. Figure 3b shows the absolute value of the filtered output. This makes the amplitude variations easier to observe in the graph (reduces the required dynamic range by a factor of 2) and emphasizes the echo peak amplitude, which is independent of whether the oscillating pressure wave is positive or negative. The filter performance can be characterized by the ratio of the flaw peak to the maximum grain peak. This ratio is referred to as the peak-signal-to-noise ratio and is given by:

$$P_{SNR} = \frac{\max(|y[n]|)}{\max(|y[k]|)} \quad \text{for } n \in \{\text{Flaw echo region}\} \text{ and } k \in \{\text{Grain only echo region}\} \quad (1)$$

A larger P_{SNR} metric implies a better filter performance for enhancing flaw echoes amplitudes over grain scattering amplitudes. There are other performance metrics that can be used such those involving power ratios or RMS values. The selection of the performance metric often involves matching a numeric quantity derived from the data to heuristic assessments of performance. For the purpose of this lab, it is assumed that simple threshold detection will be applied after processing so the goal is to maximize the distance between the flaw echo peak and the highest (worst-case) grain echo peak. A high P_{SNR} value implies a low false positive and high flaw detection rate for a broad range of thresholds.

These laboratory exercises involve working with data obtained from an ultrasonic scanner. The **technical details** concerning data collection are as follows. The A-scans were obtained from three 2-in diameter stainless-steel rods that were heat treated to obtain various grain sizes. A flaw was simulated in each specimen by drilling a flat-bottom hole of 4.22-mm diameter. The stainless-steel samples were placed in a water bath and scanned with a U2-h KB-Aerotech Alpha transducer with a center frequency of 5 MHz and a Gaussian-shaped spectrum with a 4 MHz bandwidth. The received echoes were digitized at a sampling rate of 100 MHz, and each scan was averaged 200 times in a LeCroy 9400 digital oscilloscope to reduce time varying noise (i.e. 200 scans were taken at the same position and average together). Average grain sizes for the different stainless-steel rods scanned were 86, 106, and 160 µm. These average grain sizes were determined from micrographs using a linear intercept method.
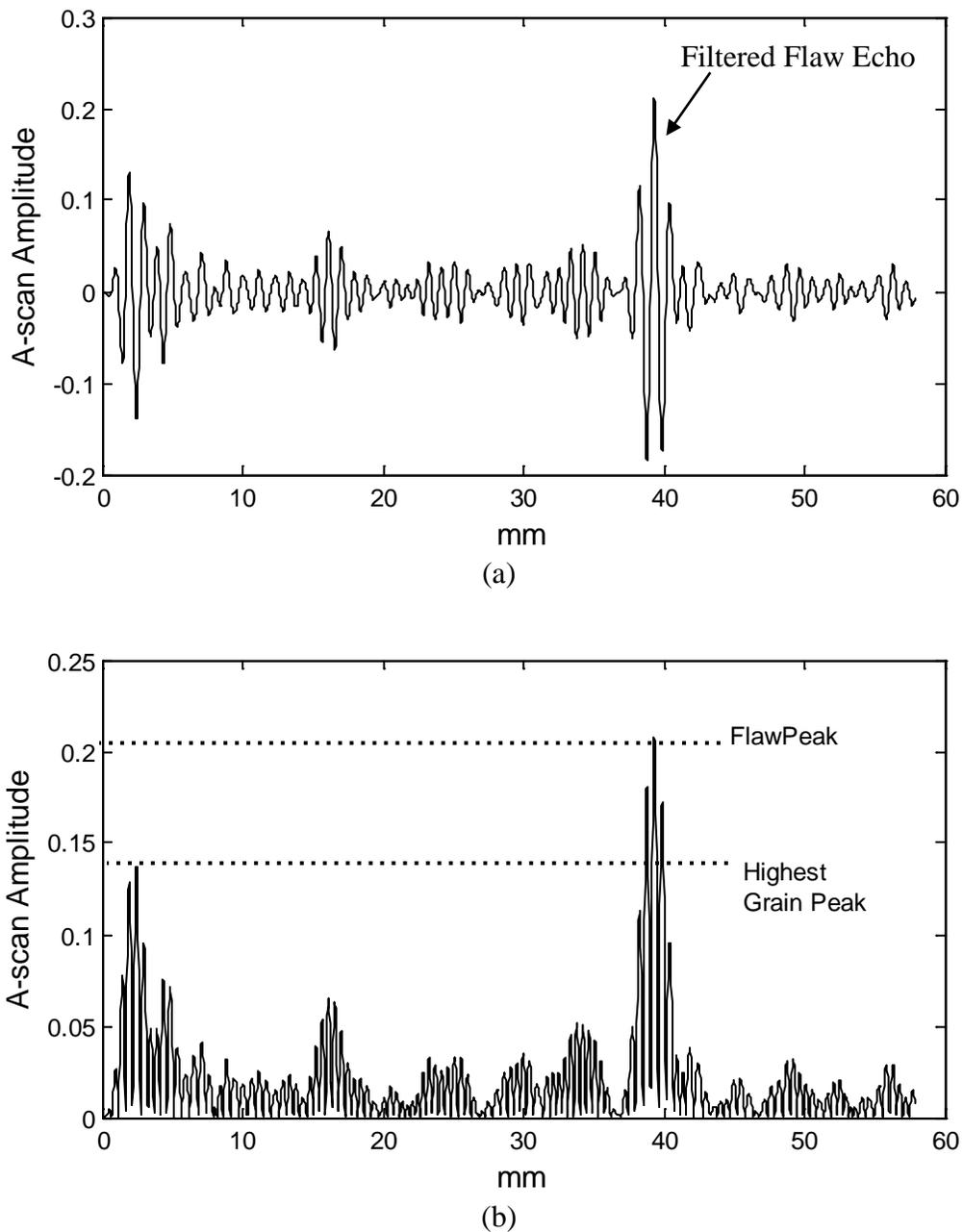
Figure 3. Filtered A-scan of Fig 1b. (a) Direct filtered output. (b) The absolute value of filtered output for better peak comparisons. The resulting P_{SNR} is about 0.21/0.14 = 1.5 or 3.5 dB.

Grain noise models:

When echo amplitude envelopes vary from multiple scattering centers, resulting in random levels of constructive and destructive interference, they can be modeled with a

Rayleigh distribution. Let **x** be the random variable denoting the envelope amplitude in this case. The Rayleigh probability density function (pdf) is given by:

$$p_{\mathbf{x}}(x) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad \text{for} \quad x \geq 0, \tag{2}$$

where scaling parameter, σ, is related to the mean by:

$$E[x] = \sigma\sqrt{\frac{\pi}{2}}, \tag{3}$$

and the cumulative distribution function (cdf) is given by:

$$P_{\mathbf{x}}(x) = 1 - \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad \text{for} \quad x \geq 0 \tag{4}$$

If the sample mean is estimated from the noise/data, Eq. (3) can be used to estimate σ, and Eq. (4) can be manipulated to find a threshold required to achieve a particular false positive rate.

2. Pre-Laboratory Assignment

1. Download file lab5nde.mat from:

   http://www.engr.uky.edu/~donohue/ee422/data/lab5nde.mat

   and load it into your Matlab workspace with the *load* command (i.e. if lab5nde is in current directory, simply type **load lab5nde.mat,** or you can use the command *uiload* to open a directory navigator window to search for the file) Once loaded type **whos** and the workspace should contain the following vectors and parameters:

Name	Size	Bytes	Class	Attributes
a1	2000x1	16000	double	
a1_posmm	1x1	8	double	
a2	2000x1	16000	double	
a2_posmm	1x1	8	double	
c	1x1	8	double	
fs	1x1	8	double	
nfa1	2000x1	16000	double	
nfa1_posmm	1x1	8	double	
nfa2	2000x1	16000	double	
nfa2_posmm	1x1	8	double	
nfa3	2000x1	16000	double	
nfa3_posmm	1x1	8	double	

   The vectors *a1* and *a2* are the sample A-scans that do not need filtering to make the flaw echo stronger than the grain. The vectors *nfa1*, *nfa2*, and *nfa3* are A-scans that need filtering in order for the flaw to be detectable (i.e. become larger

than the grain echoes). The data acquisition and material parameters for this data set are included in the workspace and are given as follows:

c   => is the speed of sound in stainless steel.
fs   => is the sampling rate.
a1_posmm   => is position of the flaw in millimeters for a1
a2_posmm   => is position of the flaw in millimeters for a2
nfa1_posmm => is position of the flaw in millimeters for naf1
nfa2_posmm => is position of the flaw in millimeters for naf2
nfa3_posmm => is position of the flaw in millimeters for naf3

Write a script to plot all 5 scans similar to those in Fig. 1, and label the X-axis in millimeters. The Y-axis is a voltage value from the digitizer/quantizer and is proportional to the acoustic pressure of the echo on the receiving transducer. Since the quantizer scale was not calibrated, there are no meaningful units for these values, so they can be simply label as "amplitude." This is typical in applications where shape and relative amplitudes of the signal are more important than the actual value. (**Hint**: the challenging part of this problem is coming up with the x-axis. The samples are based on time, with $fs=100 \times 10^6$ samples per second, so the velocity of sound multiplied by time is distance. However, you must divide by 2 to compensate for the roundtrip time for the pulse to travel to the scatter position and back again. You can check if your axis is correct by plotting the A-scans that do not need filtering and comparing their peak location with the given flaw position numbers).

2. For the five A-scans in the data set, use the *pwelch* function to plot the magnitude spectrum for each. Show only the frequency axis over 0 to 10 MHz (can use the *xlim*() command or zoom option on figure). Put each spectrum plot in its own figure. Make sure axes are properly labeled and use figure labels and captions to clearly identify the original A-scan along with its corresponding spectrum.

3. For a given set of grain-only A-scans (no flaw present), the envelope peaks were collected at a specific depth range where the average peak value was computed to be 0.62. Assume the peak values are Rayleigh distributed. Determine the threshold for flaw detection such that a false-positive error probability of 1/10,000 $= 1 \times 10^{-4}$ can be expected.

3. Laboratory Assignment

1. Design two filters (one FIR and one IIR) to optimize P_{SNR} for the 3 A-scans that need filtering (naf1, naf2 and naf3) in the *lab5nde* mat file. The optimized filter is the one that maximizes the average P_{SNR} for the 3 samples over all possible filter parameters. Write a commented script that opens the data files from the current directory and performs the optimization with an iterative loop (See suggestion below). From the script it should be clear how the best order

and cut-off frequencies were determined. **Indicate to the TA which was the best filter (type, order, and cut-offs), and present the following results:**

- **plot of magnitude response of best filter**
- **the average P_{SNR} value from the filtered A-scans**
- **plots of the 3 A-scans (naf1, naf2 and naf3) before and after filtering (clearly label the figures).**
- **plots of the 2 A-scans (*a1* and *a2*) before and after filtering (clearly label the figures) to ensure it did not lower their already-good P_{SNR} values.**

Be sure to record/save the values for the peak grain noise echoes after filtering in each of the 5 scans. For all 5 A-scans determine the filter delay. Since you are given the actual flaw location, compare these to the (peak) location in the filtered data and compute the average filter delay in millimeters. Refilter these 5 A-scan using the same filter coefficients, but use the *filtfilt* command instead of *filter*. Did the results turn out as you expected?

Suggestions: This filter design approach involves a training set. In many cases a physical model is not available for deriving the best filters analytically. So data are collected with known targets to form a testing and training set. Filter parameters are then optimized over all training set samples. I recommend writing scripts to automate the optimization process (as much as possible), such as including loops to iterate through possible filter parameters. The script should be included in your procedure section along with narratives describing your intentions/objectives for each procedure. Also remember that filters have delays. So depending on the filter order, you may have a noticeable shifting to the right of the true flaw peak position. For FIR filters, this shift will be equal to half the filter order in samples, and in some cases you may not be able to apply high-order FIR filters directly because the segment size is too small. Since the A-scans consist of 2000 sample points, the *filter* command in Matlab will truncate the output after 2000. If the flaw signal is toward the end of the segment, you may wind up pushing it beyond the truncation point. It is up to you to try different filters (the various types of FIR and IIR filters). The most critical parameters will likely be the cut-off frequencies, and then maybe the filter order.

2. From the grain echo portion of the signals, select the 5 highest envelope peaks after filtering with the best filter of Exercise 1, assume a Rayleigh distribution for the noise peaks and compute the threshold needed to result in a false positive error probability of 1×10^{-4}. Estimate the detection probability using this threshold (*i.e.* what percentage of the flaw echoes exceed this threshold). Determine the smallest false positive error probability that will result in a 100% detection probability (**Hint:** You can only base your judgment on the data you have, so find threshold that will detected the weakest flaw echo and use the Rayleigh distribution to compute the false positive error probability from that threshold).

3. From the class data file postings, download the mat file lab5test.mat and load this into your workspace.

   http://www.engr.uky.edu/~donohue/ee422/data/lab5test.mat

   This mat file contains 3 A-scans in vectors ta1, ta2, ta3, acquired with the same scanner used to collect the training data. Use the filter developed in the previous problems to find the flaw locations for each A-scan. Indicate the estimated flaw position in millimeters for each file, as well as the plots of filtered A-scans used to locate the flaws. In estimating the location, the filter delay will cause a shift between the peak in the filtered signal and the actual position. Use the estimated filter delay computed in the previous exercise to correct for this delay in your location estimates. **Show the TA your plots indicating the locations of the detected flaws. Also compare the heights of the peaks to the 100% detection threshold found in Part 2. Indicate to the TA the resulting detection and false-positive rate computed over all 3 signals.**

4. For the optimized FIR and IIR filters determined in Part 1, compute the group delays. This can be done with Matlab's command:

    >> [G,f] = grpdelay(b,a,N,fs);

   See the help files on *grpdelay* for details on the input/output arguments. This works similar to the *freqz* command but it shows the delay in samples that an input signal frequency component will undergo in passing through the filter. For FIR filter the group delay will be constant over all frequencies resulting in an undistorted time shift of the filtered flaw echo. For IIR filters the delay will be variable over the filter frequency range, but should be relatively constant over the passband. In determining the delay for the IIR filter, you should focus on the spectral region where the input signal contains the most energy. **Show the TA plots of the group delays for the IIR and FIR optimized filters used in Part 1, and describe whether the delays were consistent with the delays you actually found in Part 1.**

5. **Hand in (upload mfile) the script used to optimize for the best performing filter from Part 1. (Due before the next lab period, so you can clean up the comments and structure the code so it is more readable and usable by someone other than the author. The instructor/TA should be able to run this and get the same results you did in the lab).**