1 Spectrum Estimation

1.1 Discrete Fourier Transform Basics

a. The ideal Fourier (magnitude) spectrum of a 20 Hz sine wave and the theoretical continuous frequency DTFT (magnitude) spectrum for a windowed sine wave at 20 Hz are shown below. Since a sine wave is a real periodic signal, we expect the ideal magnitude spectrum to be an even discrete function consisting of only dirac delta functions. The magnitude spectrum of the rectangular window used is a *sinc* function. The ideal magnitude spectrum is convoluted with the *sinc* function to produce the theoretical continuous frequency DTFT (magnitude) spectrum; it is clear that the window has smearing and leakage of the ideal spectrum.
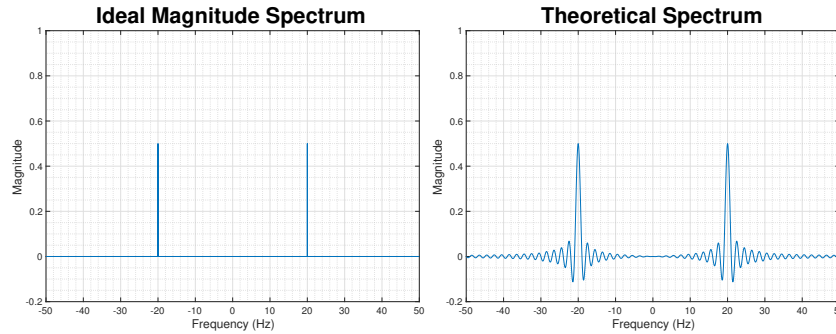

Figure 1: Ideal magnitude spectrum and Theoretical Continuous DTFT of 20 Hz sine wave

b. At first sight, it looks like the DFT spectrum on the left in Figure 1.1 is exactly the same as ideal spectrum of a 20 Hz sine wave. However, we know that this cannot be the case. The ideal spectrum is derived from a continuous infinite duration sine wave whereas the spectra presented below are both derived from finite duration discrete-time sine waves. In fact, both the spectra below are sampled versions of the theoretical continuous frequency DTFT (magnitude) spectrum shown in Figure 1.1. The resemblance described above comes from the fact that **coherent sampling** was performed. The rectangular window used to limit is signal is 0.1 s long and thus the *sinc* function will have intercepts at $f = \frac{1}{0.1} = 10Hz$. 100 samples are evenly divided across an axis that is 1000 Hz long and thus, each sample is separated by 10 Hz and exactly corresponds to frequencies at which the *sinc* function is 0, expect at ± 20 Hz. The spectral smearing and leakage that come about due to the rectangular window become evident when more points from the continuous DTFT spectrum are sampled.
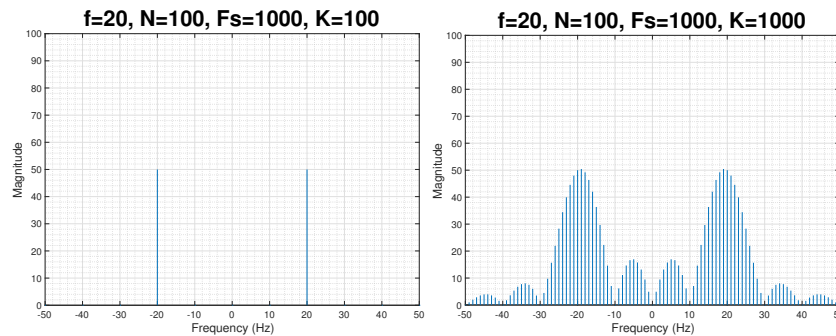

Figure 2: 100 and 100 Point DFT Spectra of 20 Hz Sine Wave

c. The graphs below show the K-point DFT (magnitude) spectrum for $K = 100$ and $K = 2^{16}$; the graph on the right was generated using the `plot` function rather than the `stem` function since the theoretical continuous frequency DTFT has been sampled so densely that it is effectively continuous. The graphs below illustrate the effects of sampling the continuous frequency DTFT sparsely; the true peaks are not correctly identified. In fact, the peaks occur at bin $\frac{f_R}{f_s} = \frac{fxN}{f_s} = \frac{24*100}{1000} = 2.4$; since this bin does not actually exist in the 100-point DFT spectrum, there is no clear peak at 24 Hz. Increasing the value of K can be used to solve the incoherent sampling problem. This is easily obtained by zero-padding the signal.
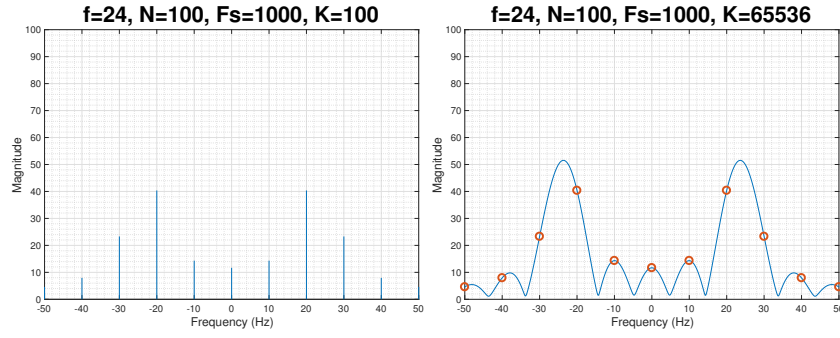
Figure 3: DFT Spectra Illustrating the Effects of Incoherent Sampling

1.2 Properties of Power Spectral Density (PSD)

Approximation in the definition of PSD

To show the equivalence of the two equations we shall start with (1).

$$P(\omega) = \lim_{N\to\infty} E\left\{\frac{1}{N}\left|\sum_{n=0}^{N-1} x(n)e^{-jn\omega}\right|^2\right\} \tag{1}$$

$$= \lim_{N\to\infty} E\left\{\frac{1}{N}\sum_{m=0}^{N-1} x(m)e^{-jm\omega}\sum_{n=0}^{N-1} x^*(n)e^{jn\omega}\right\}$$

$$= \lim_{N\to\infty}\frac{1}{N}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} E\left\{x(m)x^*(n)\right\}e^{-jm\omega}e^{jn\omega}$$

$$= \lim_{N\to\infty}\frac{1}{N}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} r_{xx}(m-n)e^{-j(m-n)\omega}$$

Substituting $\tau = m - n$ and expressing the incorporating the double summation into one summation, we get:

$$P(\omega) = \lim_{N\to\infty}\frac{1}{N}\sum_{\tau=-N+1}^{N-1}(N-|\tau|)r_{xx}(\tau)e^{-j\tau\omega}$$

$$= \sum_{\tau=-\infty}^{\infty} r(\tau)e^{-j\tau\omega} - \lim_{N\to\infty}\frac{1}{N}\sum_{\tau=-N+1}^{N-1}|\tau|r(\tau)e^{j\tau\omega}$$

And thus under the mild assumption that the covariance sequence $r(k)$ decays rapidly we have shown that (1) is equivalent to (2), which is simply the continuous DTFT of the Autocorrelation Function (ACF).

$$P(\omega) = \sum_{\tau=-\infty}^{\infty} r(\tau)e^{-j\tau\omega} \tag{2}$$

a. The Fourier Transform of a real symmetrical signal will also be real and symmetrical; the symmetry in the FT is due to the signal being real whereas the fact that the imaginary component of the FT is zero is due to the symmetrical nature of the signal. Matlab's FFT algorithm assumes that the incoming signal is wrapped and thus the specific structure of **x**, described in the coursework, is necessary. In contrast to the zero-padding performed in Section 1.1, here we pad zeros to the middle of the vector **x** to preserve symmetry. This is important to note that in this specific scenario, zero-padding is actually the same as extending the duration for which we observe the signal; this is because for $|k| \geq M$, the autocorrelation function is 0. As such, zero-padding is increasing the frequency resolution of $P(\omega_k)$. In contrast, zero-padding is usually used to increase the number of points for the which the continue DTFT is sampled and the frequency resolution depends solely on the size of the rectangular window used to restrict the signal from an infinite duration to a finite duration. The graphs below clearly illustrate this point. It is interesting to note that the autocorrelation function described in the coursework is actually the ACF of the rectangular window. As M is increased from 10 to 128, the width of the mainlobe decreases; this is exactly what we expect. Smearing of the spectra is inversely proportional to the width of the rectangular window.
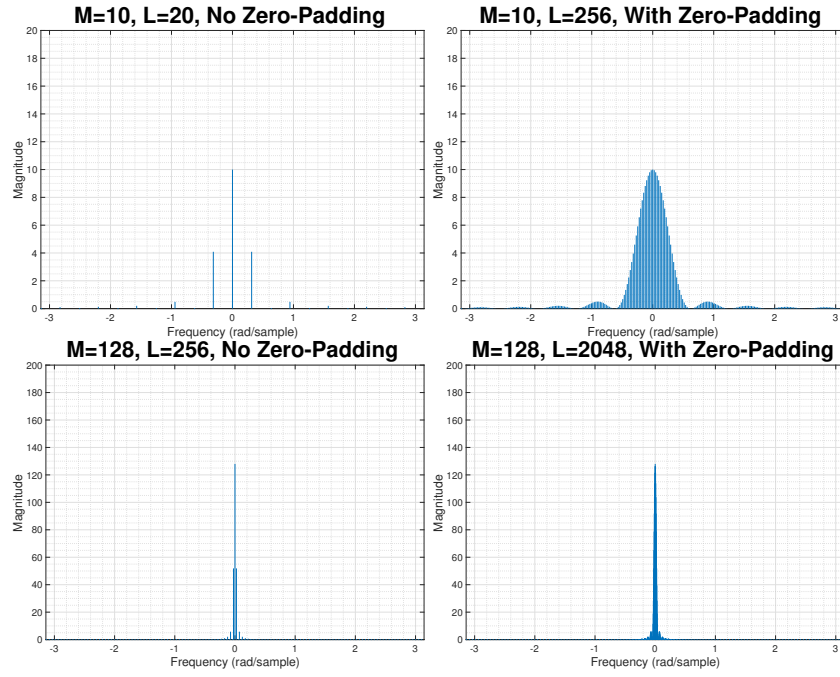
Figure 4: $P(\omega_k)$ of $r(k)$ Described in the Coursework with and without Zero-Padding for $M = 10$ and $M = 100$

b. Due to the symmetrical nature of the signal for which we are taking the FFT, we do not expect any imaginary components. The graph below shows that the imaginary components are extremely small and are about 15 orders of magnitude smaller than the real components. Above, the `abs` function was used to eliminate the imaginary components. Using the `real` function had no significant difference due to the extremely small values for the imaginary components.
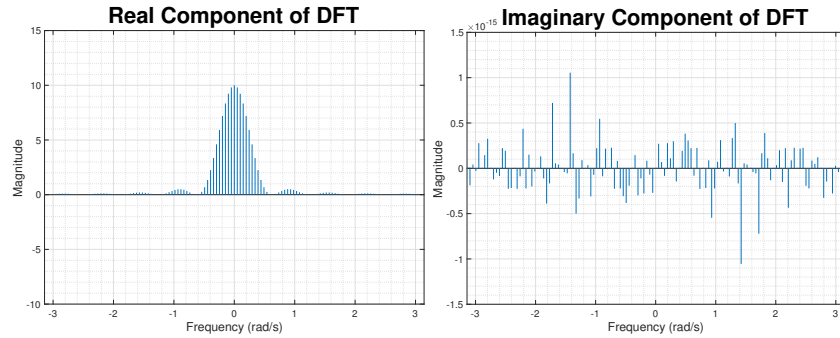

Figure 5: Comparing the magnitudes of the Real and Imaginary Components of $P(\omega_k)$

c. The graph below shows that in this case, the real and imaginary components of the DFT are of the same magnitude and thus using the `real` is not suitable. The correct Power Spectral Density can be obtained using the `abs` function in a similar fashion to part (a).
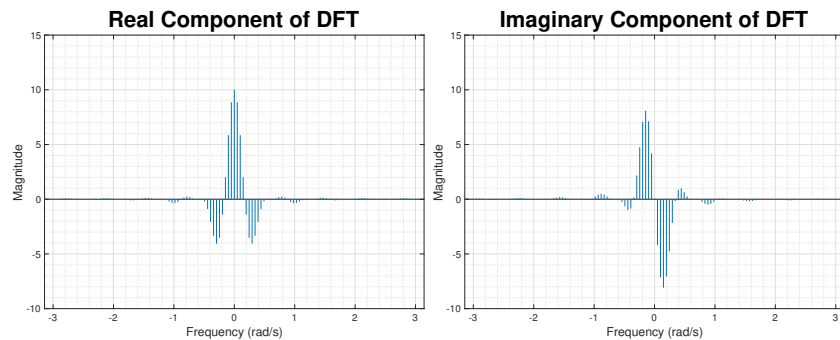

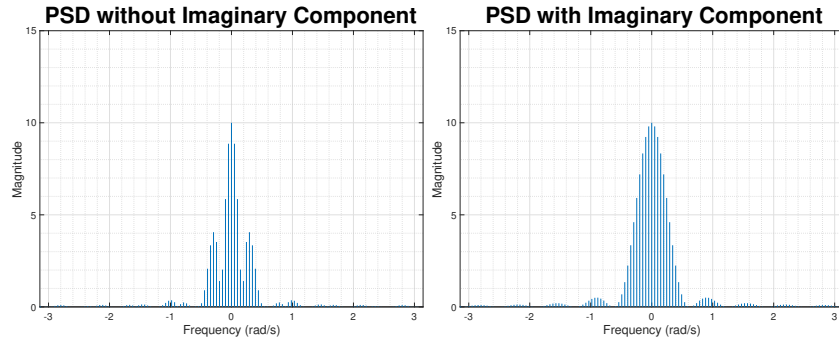Figure 6: Magnitudes of the Real and Imaginary Components are of the Same Order of Magnitude

4

Figure 7: Erroneous Spectral Values because Imaginary and Real Components are Same Order of Magnitude

d. The vectors **w** and **n** do indeed depend on if the sequence is odd or even. The formula are described in the table below.

Sequence Type	Formula for w	Formula for n
Odd	$-\pi : \frac{2\pi}{L} : \pi - \frac{2\pi}{L}$	$-\frac{L}{2} : 1 : \frac{L}{2} - 1$
Even	$-\pi + \frac{\pi}{L} : \frac{2\pi}{L-1} : \pi - \frac{\pi}{L}$	$-\frac{L-1}{2} : 1 : \frac{L-1}{2}$

Table 1: Formule for vectors **w** and **n**

1.3   Resolution and Leakage of Periodogram-based Methods

Figure 8 shows the Bartlett Window for $N = 128$ and $N = 512$ on both linear and logarithmic scales. The red line is graphed to show the 3 dB frequency of each window.
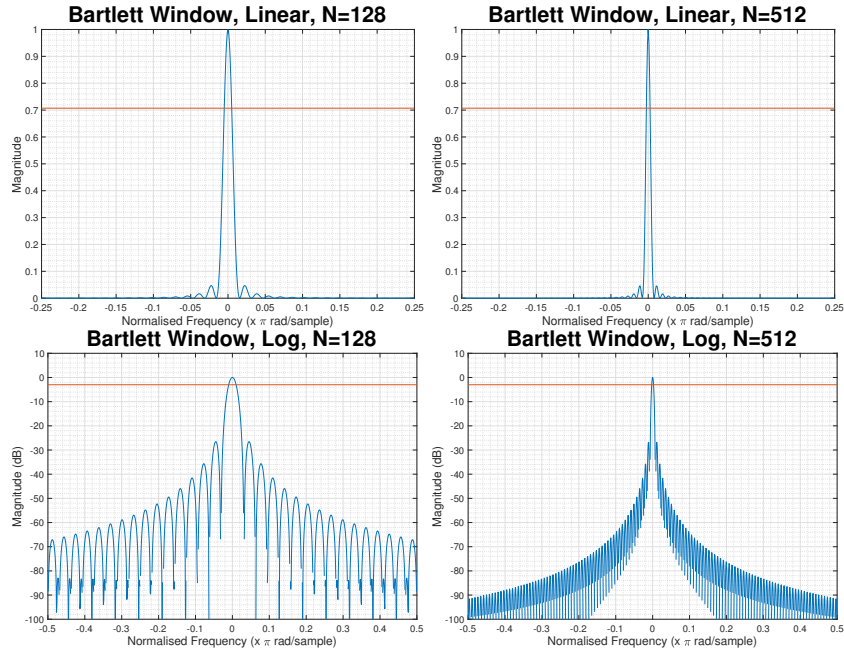

Figure 8: Bartlett Window with Linear and Log Scales for $N = 128$ and $N = 512$

To find the relationship between the width of the mainlobe and the length of the window, the 3 dB frequency is evaluated for a range of window lengths[1]. The figure below shows the relationship between the normalized 3 dB width of the main lobe of the Bartlett Window,ω_c, and both N and $\frac{1}{N}$.

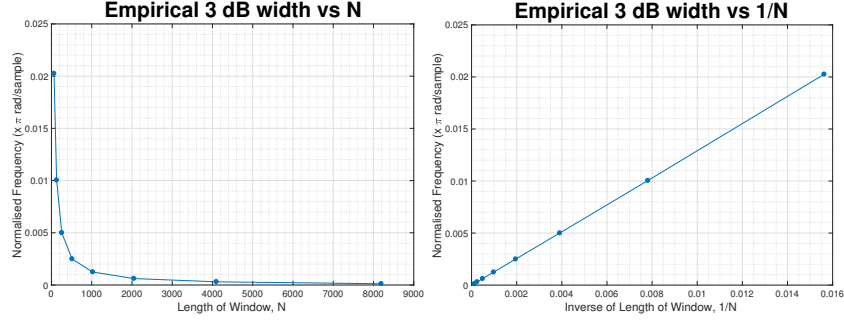[1]The intersection points were found using the function `InterX` found on matlab file exchange.

Figure 9: Empirical Relationship between Normalized 3 dB width of Mainlobe for Bartlett Window and the Length of the Window

In matlab, a linear relationship between the normalized frequency and $1/N$ is obtained using `pinv`. The linear relation is:

$$\omega_c = \frac{1.2965(2\pi)}{N}$$

The figure below shows the independence of the sidelobe peak to the length of the window. The sidelobe peak is at a constant value of $-26.5dB$ relative to the height of the mainlobe.
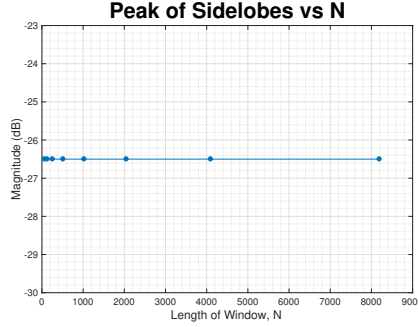

Figure 10: Empirical Relationship between Sidelobe Peak for Bartlett Window and the Length of the Window

b. The signal $x(n)$ has been very carefully designed to study the frequency resolutions of different windows. This is because the frequency of one sinewave different from the frequency of the other by a factor of $\frac{\alpha 2\pi}{N}$; The resolution of a spectral estimation method has the following general form described in (3). Thus, the signal $x(n)$ can be used to understand the scaling factor.

$$\text{Res}\left[\hat{P}_{per}\left(e^{j\omega}\right)\right] = c\frac{2\pi}{N} \tag{3}$$

The following graphs show the periodograms for $\alpha \in 1, 0.65, 0.60$. It is clear that for values of $\alpha \leq 0.60$, the two peak are not distinguishable. This value of alpha is significantly smaller than the value quoted in the lecture notes and in [?]. The value quoted in both pieces of literature take into account that the periodogram has some noise however we have removed noise and thus observe greater resolution.
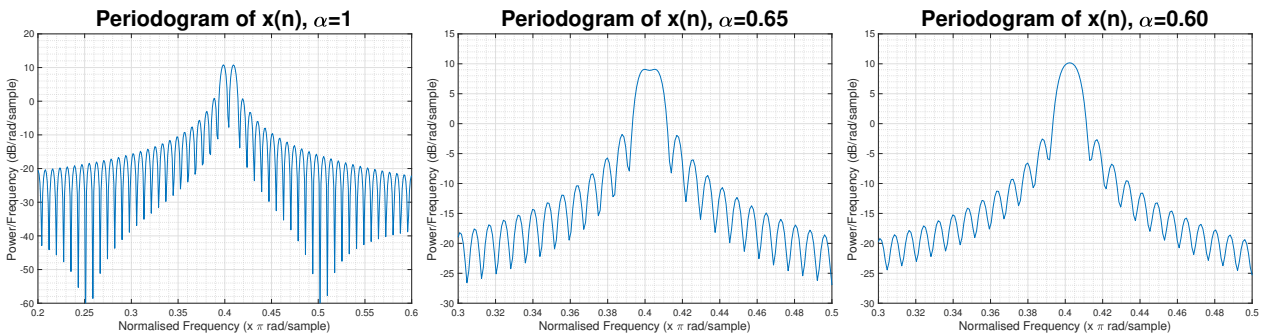

Figure 11: Periodogram of Signal $x(n)$ for Different Values of α

6

c. The hamming window has a wider mainlobe than the rectangular window. As such, there is greater spectral leakage and thus the two frequencies are indistinguishable beyond $\alpha = 0.70$. Although hamming window has a wider mainlobe, its sidelobes are much lower and this is evident in the figure below.
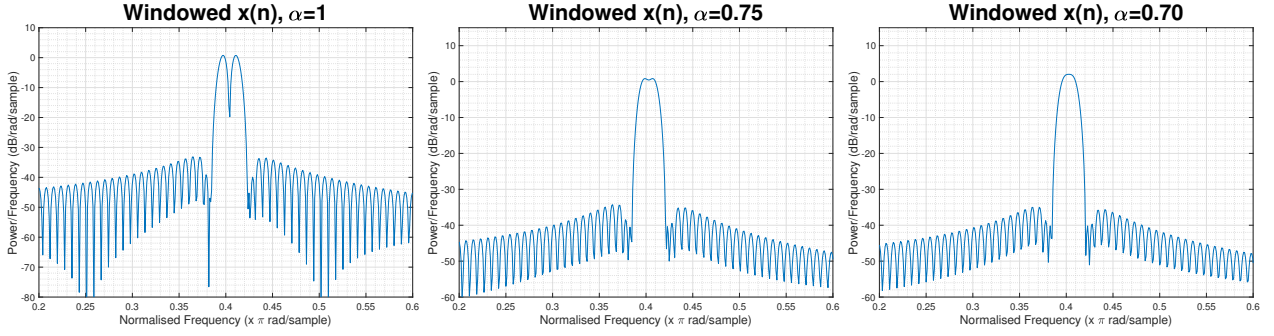

Figure 12: Periodogram of Signal $x(n)$, Windowed using a Hamming Window, for Different Values of α

d. The following figures show the periodogram obtained using the rectangular window. All windows have to trade-off the width of the mainlobe and the relative heights of the sidelobes. The rectangular window lies at an extreme end of this spectrum in that it has the smallest mainlobe however it also has the highest sidelobes. This causes the least amount of smearing and bias but results in the most spectral leakage. As such, the ability to distinguish the second peak deteriorates significantly as a_2 decreases. The amplitude threshold identification of the second sinusoidal term was slightly easier for $\alpha = 12$ however it did not change significantly. The reason for this will be explained in the next part.
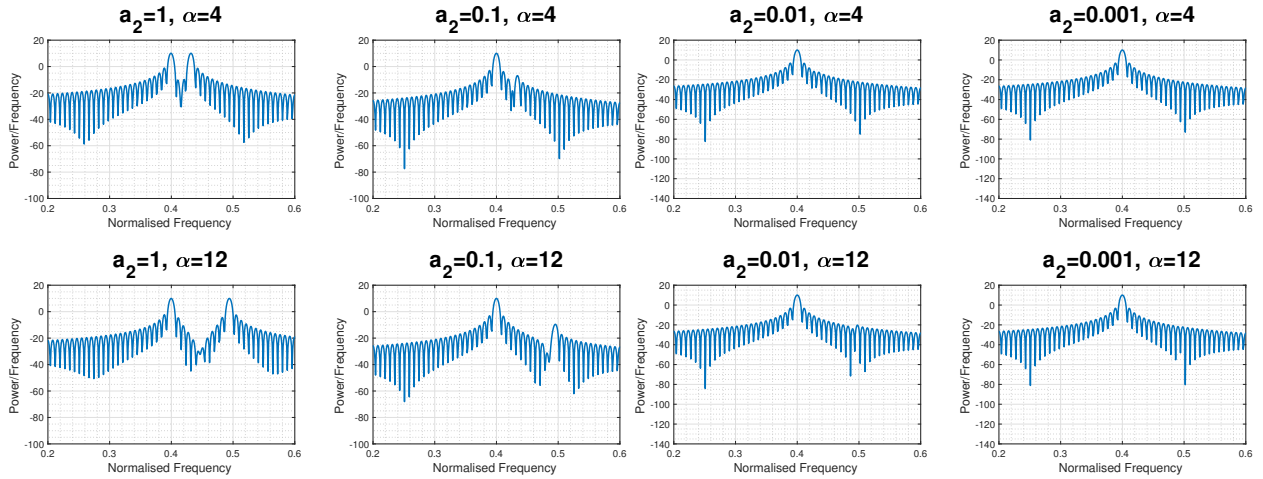

Figure 13: Identification of Sinusoids using Rectangular Window

e. A rectangular window with $N = 256$ can be thought of as a digital filter with 256 zeros; as such, it is able to fix the gain to be 0 at 256 points. At frequencies, $f = 4/N$ and $f = 12/N$ the gain of the filter is 0. However, the sidelobes of the rectangular window are not constant and they decrease as distance from the mainlobe increases. As such, the amplitude threshold identification is slightly easier at $\alpha = 12$.
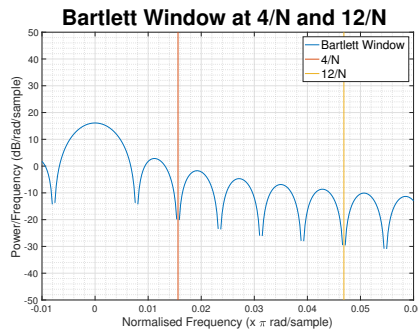

Figure 14: Power Spectral Density of the Bartlett Window at $f = \frac{4}{N}$ and $f = \frac{12}{N}$

f. As shown in the figure below, it is possible to resolve the two components even when $a_2 = 0.001$. Notice that when the Chebyshev window is used, the overlapping of the mainlobes causes trouble when trying to distinguish between two frequcies; this is clear when $\alpha = 4$. This represents the tradeoff that windows have to make between the width of the mainlobe as the height of the sidelobes. The rectangular window has a small mainlobe and thus the trouble in identification comes about because of the leakage effects whereas the Chebyshev window cause more smearing and less leakage.
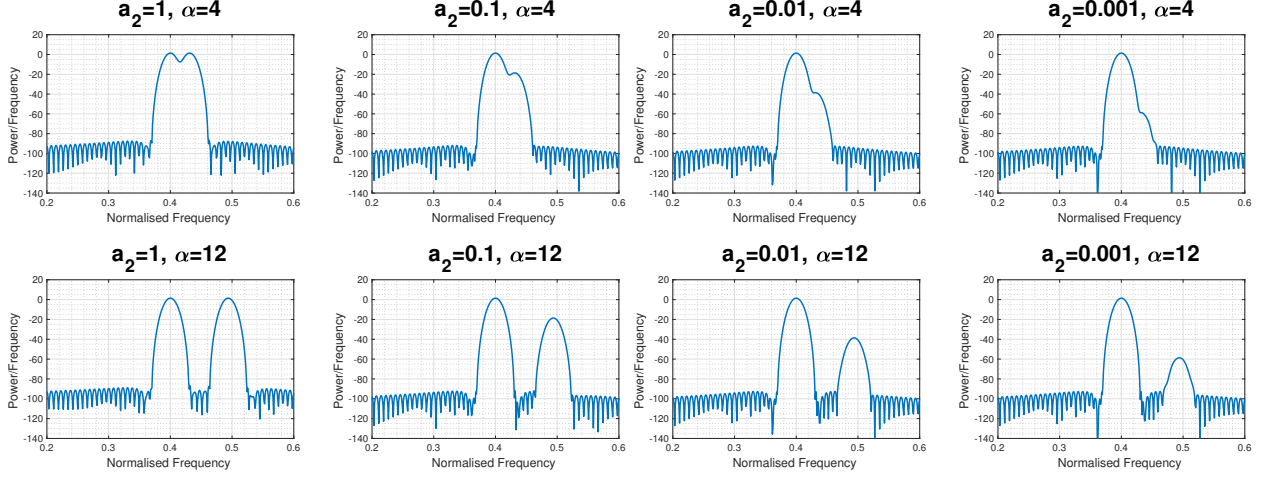

Figure 15: Identification of Sinusoids using Chebyshev Window

The Blackman-Tukey method was used to find the periodograms and the results are shown below. The autocorrelation function is calculated but only 64 values are used; the tail end of the autocorrelation is not reliable as they were calculated by averaging very few points and thus they are excluded when the Power Spectral Density is calculated. The results are not promising and the Blackman-Tukey method performs poorly compared to the Chebyshev window. However, this is completely expected. The Blackman-Tukey method is used to reduce the variance of the periodogram estimate by reducing the variance in the autocorrelation estimate. A reduction is variance is traded-off for resolution; notice the number of troughs between the two peaks has reduced by a factor of 4 when the estimates are compared to the estimate using the rectangular window. This is expected since we have used $\frac{1}{4}$ the number of points to calculate the DFT.
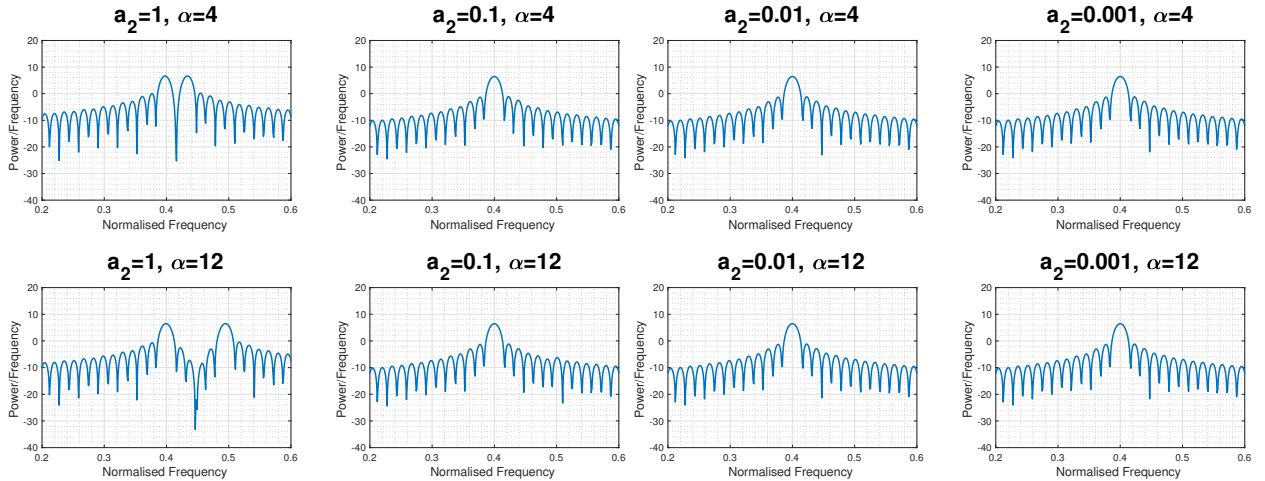

Figure 16: Blackman-Tukey Method for Identification of Two Sinusoidal Tones

1.4 Periodogram-based Methods Applied to Real-World Data

a. The figure below shows the periodogram, with and without pre-processing, for the sunspot time series. It is evident that the periodograms produced by removing the mean and removing the trend are very similar. These two periodograms are also extremely similar to the original periodogram beyond $f_n = 0.2$.
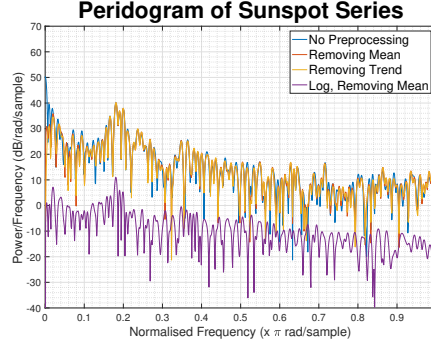
Figure 17: Standard Periodogram, with and without Preprocessing, for Sunspot Time Series

The original periodogram has a DC component that causes the periodogram to have a spike at $f = 0$ of approximately 50 dB. The removal of the mean removes the the spectral components close to $f = 0$. The mean can be visualised as a rectangular signal that has been added to a zero-mean signal. Through the linearity of the Fourier Transform and the fact that the spectrum of a rectangular signal is a *sinc*, the spectrum of the mean-removed signal only differs from the original signal at low frequencies. The `detrend` function removes linear trends. The Fourier Transform of linear functions is a spike at $f = 0$ and a magnitude proportional to $\frac{1}{f^2}$ for $f \neq 0$. As such, using `detrend` prior to formulating the periodogram affects low frequency components.

Note that before we are able to take logarithms of the time-series, a small offset of 0.001 was added to the entire series. This allowed the logarithms to be taken because the sunspot series has values of 0 at certain points. Since this offset is added to all the terms, it too was removed when the mean of the series was subtracted. The shape of the periodogram obtained is similar to the periodogram obtained solely by removing the mean and tread. The main difference is in the fact that spikes have been greatly accentuated.

b. The standard periodogram of the EEG obtained from an electrode located at the POz region of the head is graphed in the figure below. The peaks in the spectrum corresponding to the SSVEP are clearly observed at $f = 13, 26, 39, 52$. As expected, the fundamental frequency response peak is the largest whereas the height of the harmonics is monotonically decreasing.
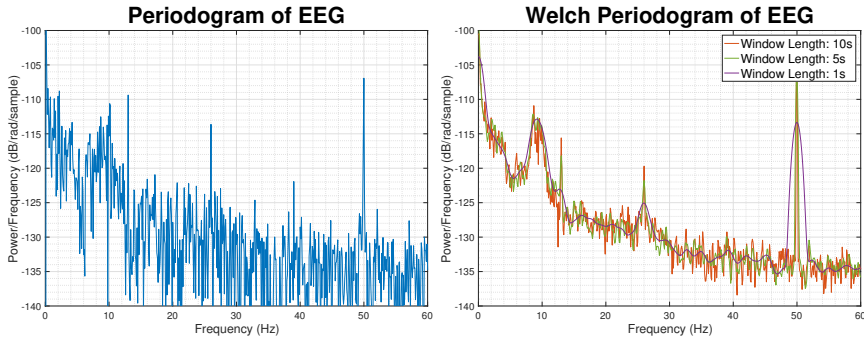

Figure 18: Standard Periodogram and Welch Periodogram with Hamming Window for EEG samples Obtained from POz Location on the Scalp

The figure also shows the Welch periodograms obtained by averaging the signal over different window lengths (10 s, 5 s, 1 s). As the duration of the observation window becomes smaller, the variance of the periodogram decreases. A smaller observation window means that the number of periodograms obtained increases; there are a greater number of periodograms to average over and thus the variance decreases. However, using a shorter obervation window means that the frequency resolution of the periodogram has decreased. The damning effect of decreasing the frequency resolution is evident when the periodogram obtained from by averaging over 1 s windows is studied. The large alpha-rhythm present in the EEG have masked the distinct peak that occurs at $13Hz$. This makes identifying the fundamental frequency response peak much harder. Notice that the first harmonic at $f_n = 26$ is still visible. This is due to the fact that there are no large frequency components around $f = 26$ that can cause spectral leakage and thus mask the harmonic. This shows the clear trade off between frequency resolution and variance reduction that has to be made when averaging periodograms.