Imperial College London

# EE4-13 Adaptive Signal Processing and Machine Intelligence (2018-2019)

Imperial College London

Department of Electrical and Electronic Engineering

# Coursework

*Authors:* CID

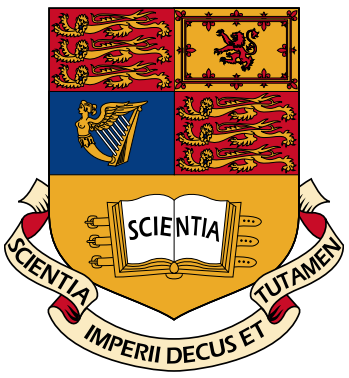
Adel Haddad 01060023

*Supervisor:*

Prof. Danilo P. Mandic

Date: April 2019

Word Count: 0



# Contents

# 1 Classical and Modern Spectrum Estimation

## 1.1 Properties of Power Spectral Density (PSD)

### 1.1.a Approximation in the Definition of PSD

Starting with the equation provided, (1), we: use the relationship between modulus and complex conjugate for complex numbers, we move out the summation terms from the expectation operator, we factor out the exponential terms - as they are independent of the random variable $x$ and we finally use the following substitution: $g(\tau) = r_{xx}(\tau)e^{-j\omega\tau}$.

$$
\begin{aligned}
P(\omega) &= \lim_{N\to\infty} \mathbb{E}\left\{\frac{1}{N}\left|\sum_{n=0}^{N-1} x(n)e^{-j\omega n}\right|^2\right\} \qquad (1)\\
&= \lim_{N\to\infty} \mathbb{E}\left\{\frac{1}{N}\sum_{m=0}^{N-1} x(m)e^{-j\omega m}\sum_{k=0}^{N-1} x^*(k)e^{j\omega k}\right\}\\
&= \lim_{N\to\infty} \frac{1}{N}\sum_{m=0}^{N-1}\sum_{k=0}^{N-1} \mathbb{E}\left\{x(m)e^{-j\omega m}x^*(k)e^{j\omega k}\right\}\\
&= \lim_{N\to\infty} \frac{1}{N}\sum_{m=0}^{N-1}\sum_{k=0}^{N-1} \mathbb{E}\left\{x(m)x^*(k)\right\}e^{-j\omega(m-k)}\\
&= \lim_{N\to\infty} \frac{1}{N}\sum_{m=0}^{N-1}\sum_{k=0}^{N-1} r_{xx}(m-k)e^{-j\omega(m-k)} = \lim_{N\to\infty} \frac{1}{N}\sum_{m=0}^{N-1}\sum_{k=0}^{N-1} g(m-k) \qquad (2)
\end{aligned}
$$

We can convert the double summation into a single summation using Equation 3, and recalling the earlier substitution:

$$
\sum_{m=-N}^{N}\sum_{k=-N}^{N} g(m-k) = \sum_{\tau=-2N}^{2N} (2N+1-|\tau|)g(\tau) \qquad (3)
$$

(2) can then be written as:

$$
\begin{aligned}
P(\omega) &= \lim_{N\to\infty} \frac{1}{N}\sum_{\tau=-(N-1)}^{N-1} (N-|\tau|)r_{xx}(\tau)e^{-j\omega\tau}\\
&= \lim_{N\to\infty} \sum_{\tau=-(N-1)}^{N-1} r_{xx}(\tau)e^{-j\omega\tau} - \lim_{N\to\infty} \frac{1}{N}|\tau| \sum_{\tau=-(N-1)}^{N-1} r_{xx}(\tau)e^{-j\omega\tau}\\
&\approx \sum_{\tau=-\infty}^{\infty} r_{xx}(\tau)e^{-j\omega\tau} \qquad (4)
\end{aligned}
$$

### 1.1.b Simulation of the Limiting Case

An unbiased estimator of the autocorrelation will provide a case where the correlation does not rapidly decay. This would violate the derivation shown in Section 1.1.a, thereby hindering equivalence. An example is shown in Figure 1.1.1, here we see a sine function with an unbiased estimator:

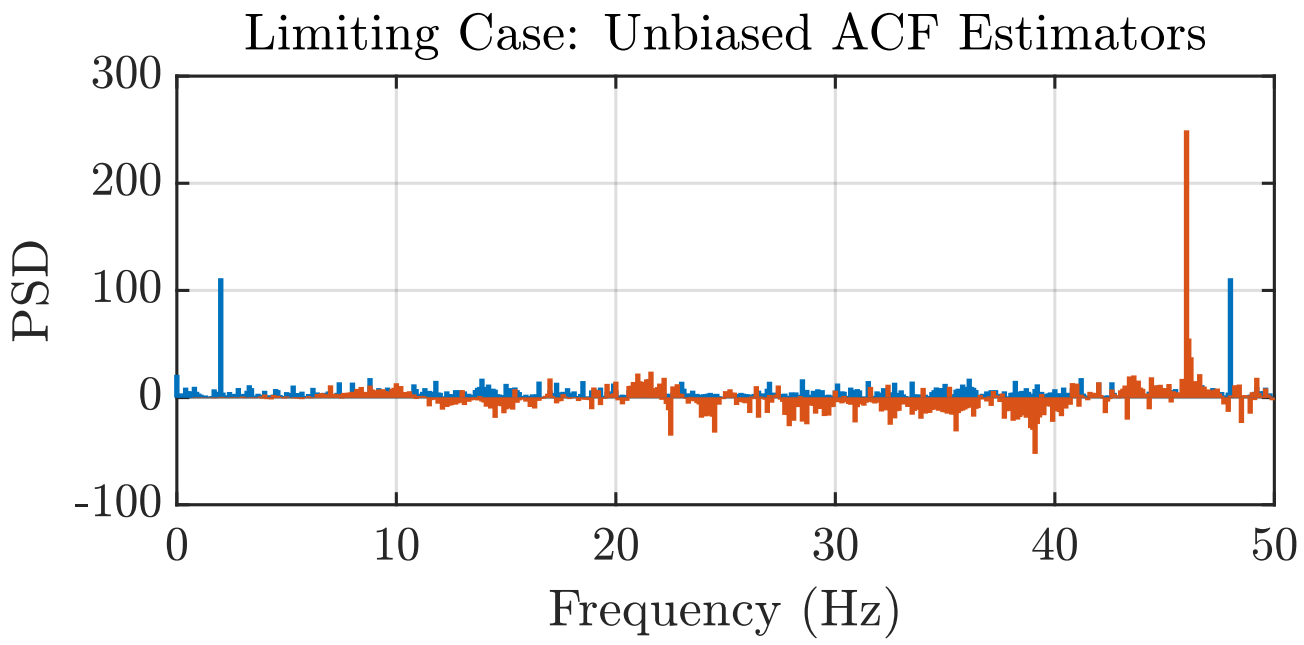


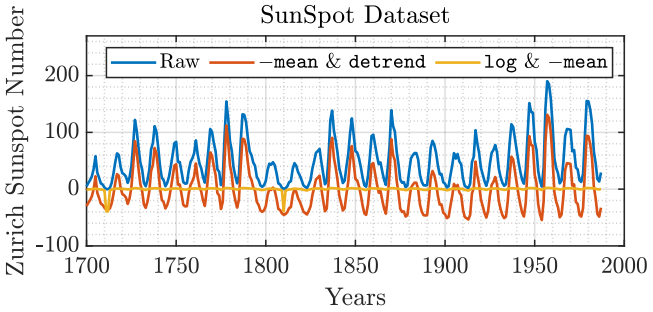

Figure 1.1.1: An example of the Limiting Case of the Periodogram Definition

## 1.2 Periodogram-based Methods Applied to Real–World Data

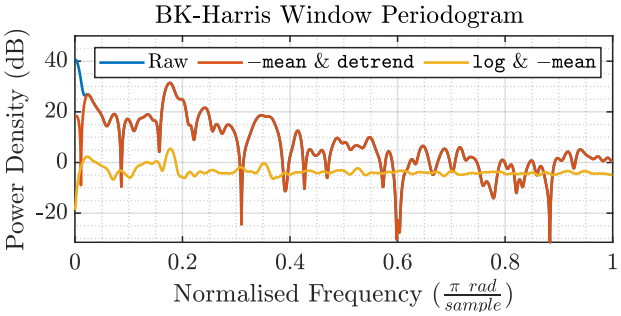
### 1.2.a The SunSpot Dataset

Figure 1.2.1. The mean's influence on data is the offset DC bias, captured in the $f = 0$ component of the periodogram. Hence as we would expect, subtracting the `mean` reduces its magnitude in the periodogram. `detrend` removes linear trends, it seems in the case of this data set that most linear trends are captured at $f \lessapprox 0.02 rad/sample$.

The natural logarithm was taken using: `log`. As the logarithm has a compression effect on magnitude we see that the magnitude of both raw and periodogram is greatly attenuated. We note that frequencies of interest and its harmonics appear as more distinct when compared to the rest of the periodogram.



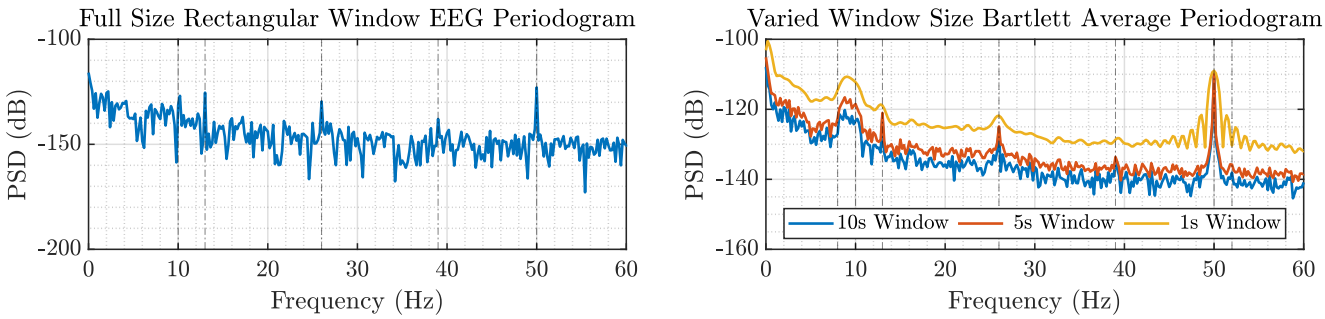

(a) Raw and its preprocessed datas




(b) Periodograms

Figure 1.2.1

### 1.2.b The EEG Dataset




(a) Standard Periodogram

(b) Bartlett Average Periodograms

Figure 1.2.2

| **Response** | **Expected Range** (Hz) | **Observed Range** (Hz) |
|---|---|---|
| Alpha Rhythm | $8-10$ | $8-10$ |
| SSVEP | $\texttt{range}[11-20]$ | $13n$ |
| Power-Line | 50 | 50 |

Table 1.2.1: EEG Frequency Peaks of the Periodogram. $n$ refers to harmonics

The standard periodogram has identifiable peaks, showing in Figure 1.2.2, quantified in Table 1.2.1. We are unable to identify the 3rd harmonic of the SSVEP, at 52Hzs it is too close to the power-line interference at 50Hz. The main difference in the 10s window averaged periodogram is clearer peak isolation compared to the surrounding periodogram and emphasis on the range of frequencies of the alpha-rhythm, rather than a single discrete peak.

## 1.3 Correlation Estimation

### 1.3.a Unbiased and Biased ACF Estimates

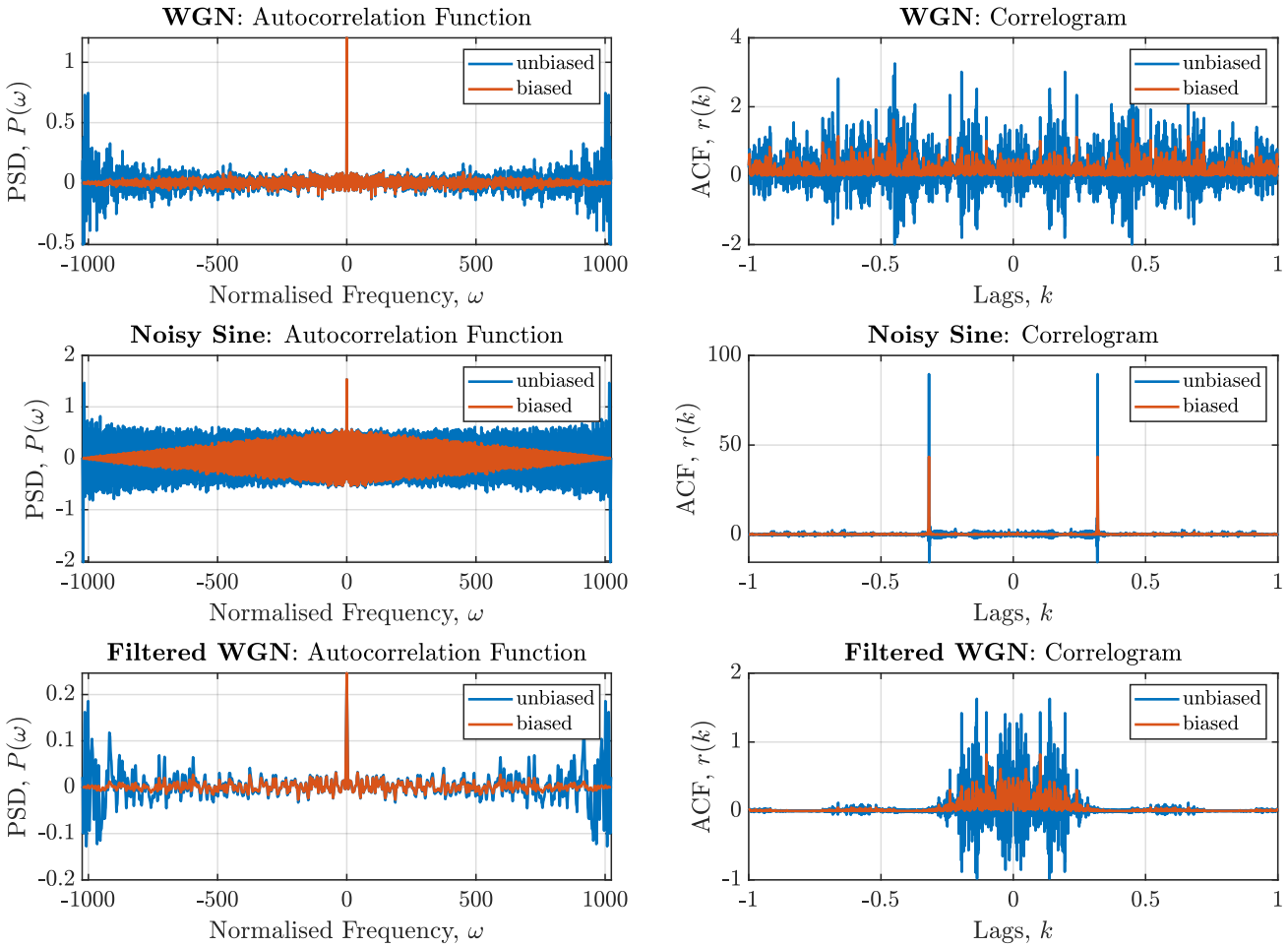


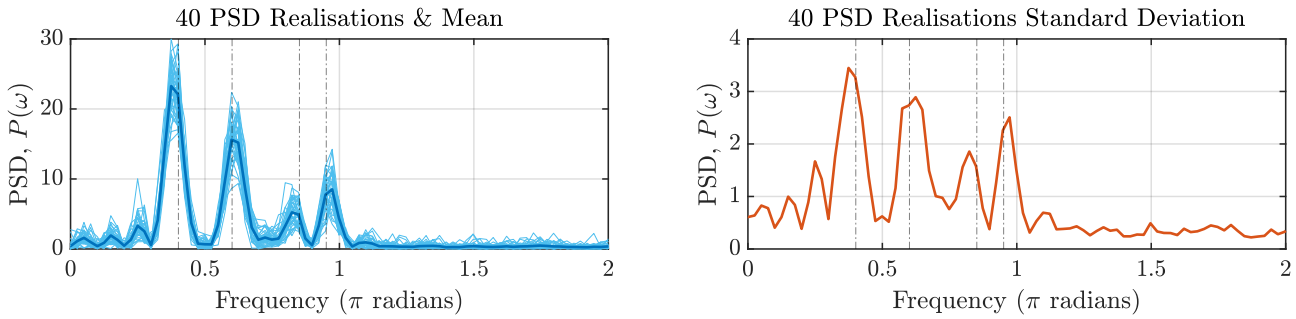

Figure 1.3.1: Set of Auto-Correlation Functions (ACFs) and their Correlograms

Figure 1.3.1. For the autocorrelation functions: we can see that the biased estimator tends to 0 for increasing lag magnitude, whereas the unbiased estimator remains somewhat constant, although at the extremes it begins to increase to approximately double the constant value.
For the correlograms: we observe that the biased estimator does not contain negative values.

### 1.3.b Biased ACF Estimator PSDs

The process simulated was the following:

$$x(n) = 2sin(2\pi 0.4n) + 1.75sin(2\pi 0.6n) + 0.85sin(2\pi 0.85n) + 1.2sin(2\pi 0.95n) + \eta(n) \quad \eta \sim \mathcal{N}(0,1) \tag{5}$$

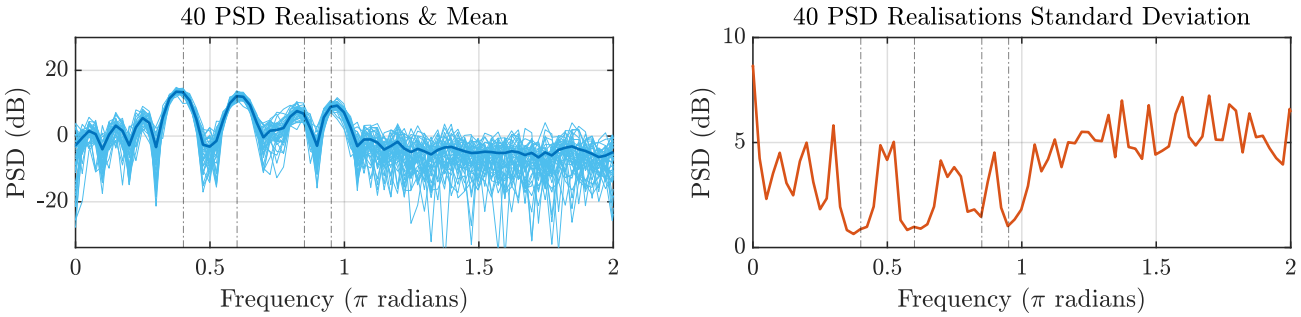



(a) Periodogram (b) Standard Deviation

Figure 1.3.2: The total number of data points used was 512, black vertical lines indicate the model defined frequencies

It is interesting to see the low frequency resolution influences the accuracy of the peak with respect to the actual frequencies used, Figure 1.3.2.

### 1.3.c Biased ACF Estimator PSDs on the dB Scale




(a) Periodogram (b) Standard Deviation

Figure 1.3.3: The total number of data points used was 512, black vertical lines indicate the model defined frequencies

Figure 1.3.3. It is advantageous that the standard deviation decreases around our frequencies of interest instead of increases.

### 1.3.d Influence of Data Samples on the PSD

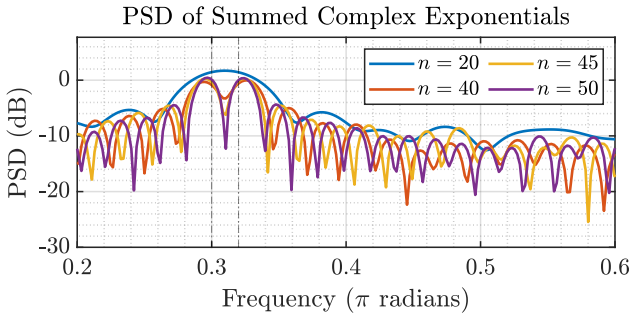



Figure 1.3.4: PSD while varying $n$, the number of Data Samples used

In Figure 1.3.4 we can clearly see that the frequency resolution is insufficient at lower sample numbers, resulting in aliasing of the desired frequency peaks.

### 1.3.e MUltiple SIgnal Classification (MUSIC) Estimator

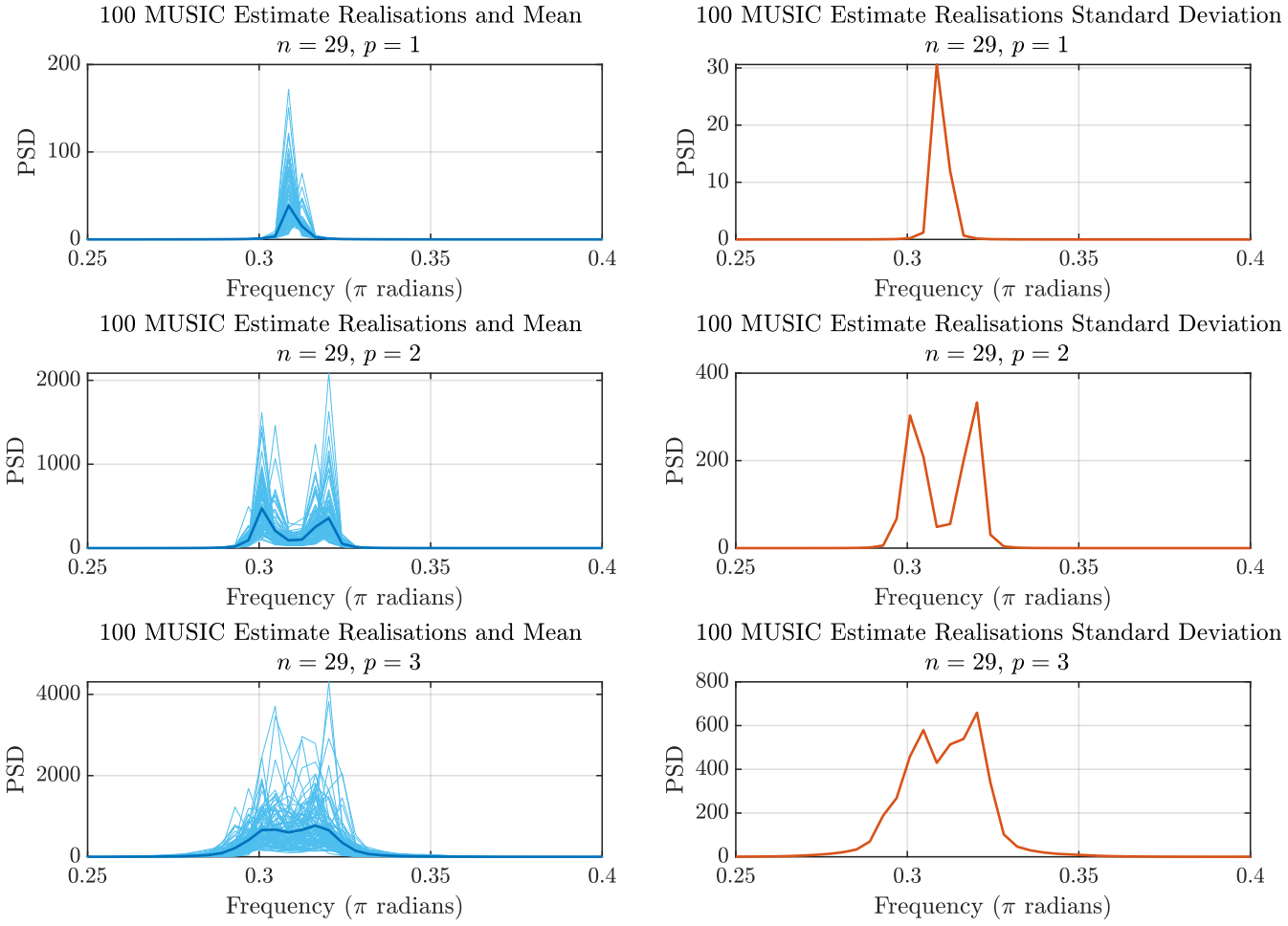



Figure 1.3.5: Set of Auto-Correlation Functions (ACFs) and their Correlograms.
$n$ is the number of samples used, $p$ is the Signal Space Dimensionality

The MUSIC estimate identifies eigenvalues of an autocorrelation matrix in descending magnitude order - which indicate directions of largest variability in the subspace.
The first line creates the 14 dimensional modified correlation matrix $R_{xx}$ for the dataset, which is subsequently used by the MUSIC algorithm with $p$ setting the Signal Subspace Dimensionality in the second line.
The third line plots the Pseudospectrum (PSD Estimate) against the frequency axis.

Figure 1.3.5 shows the MUSIC estimate plots for the given equation. The spectrum is not very detailed, peaks are not all sharp, but are clearly distinguishable from the surrounding periodogram. In the general case the MUSIC estimator is a good choice if the signal space dimensionality is known, especially as it works on such a limited set of samples.

The periodogram is equally suitable for resolving peaks, but requires more samples for a meaningful resolution.

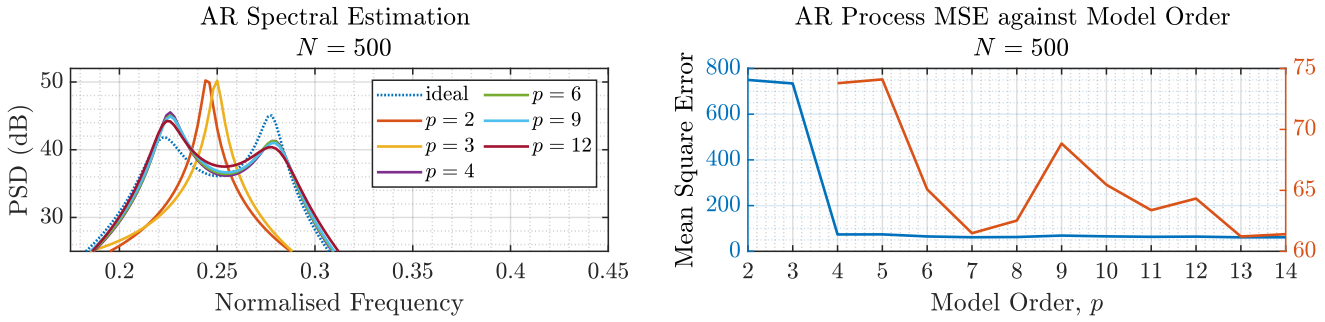
In both periodogram and MUSIC estimators we see that the standard deviation, hence variance increase around the peaks, except when using the log scale of the periodogram, where it decreases.

## 1.4 Spectrum of Autoregressive (AR) Processes

### 1.4.a Shortcomings of the Unbiased ACF in finding AR Parameters

As the unbiased estimator allows for negative values, at a computational level it will require more bits to store, especially for larger values.

### 1.4.b Error of the AR PSD Estimate



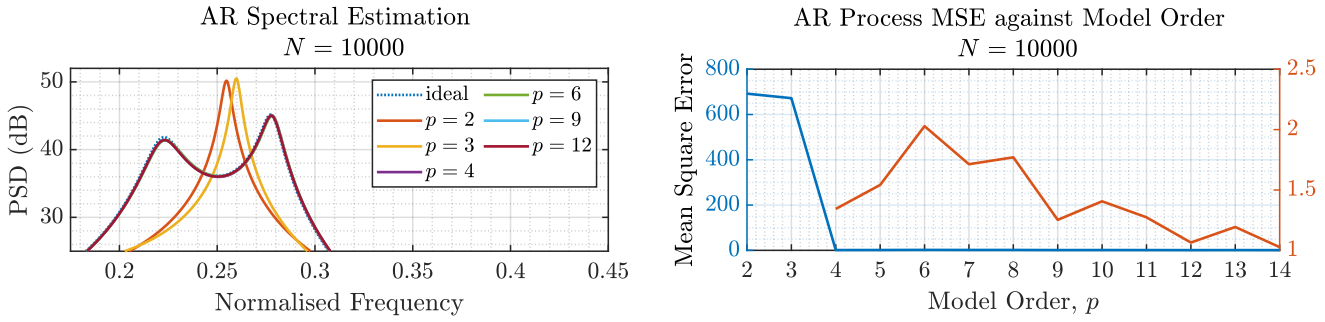

(a) AR Periodogram and its $p$ Order Estimates

(b) Mean Squared Error

Figure 1.4.1

We can see in (a) of 1.4.1, that increasing the order tends towards a better solution. But (b) notes that the most drastic difference is at the model order, 4, which matches the order of the process defined, subsequent increases of the model order do increase its likeliness to the true response, but changes are not so drastic.

### 1.4.c Error of the AR PSD Estimate with more Samples




(a) AR Periodogram and its $p$ Order Estimates

(b) Mean Squared Error

Figure 1.4.2

The same trend is seen in Figure 1.4.2, as we saw with $N = 500$, although the estimate now matches the underlying model much better, reflected in the drastically lower MSE.

It is noted that for a more valid comparison of model order's influence on the estimate's error the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC) are more suitable quantifiers of model order suitability than MSE. But the accuracy trend reflected in the MSE is still valid for our comparisons.

## 1.5 Real World Signals: Respiratory Sinus Arrhythmia from RR-Intervals

### 1.5.a Standard & Average PSDs of the RRI Dataset

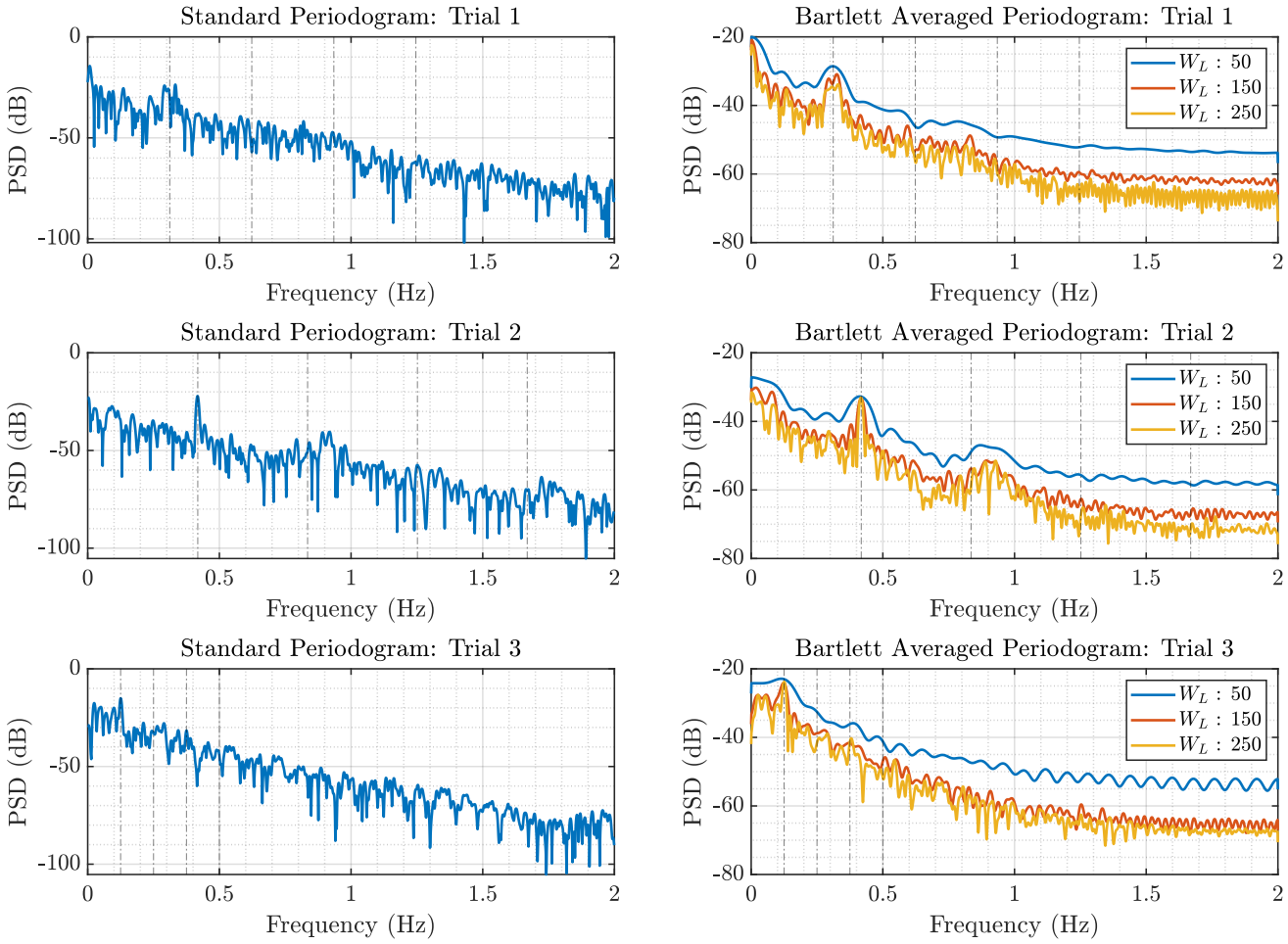



Figure 1.5.1: Standard and Bartlett Average Periodograms.
$W_L$ is the Window Length used. First 4 harmonics denoted by vertical black lines. Zero Padded Signal Length: 4096.

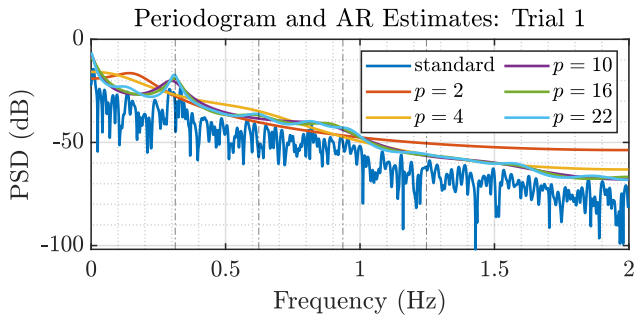
| **Breathing Type** | **Expected** (BPM) | **Observed Peak** (BPM) |
|---|---|---|
| Normal (Trial 1) | $10 - 15$ | $0.31 \times 60 \approx 18.7$ |
| Fast (Trial 2) | 25 | $0.41 \times 60 \approx 25$ |
| Slow (Trial 3) | 7.5 | $0.125 \times 60 = 7.5$ |

Table 1.5.1: Breaths Per Minute (BPM), i.e. Observed Frequency $\times 60$, for all trials.

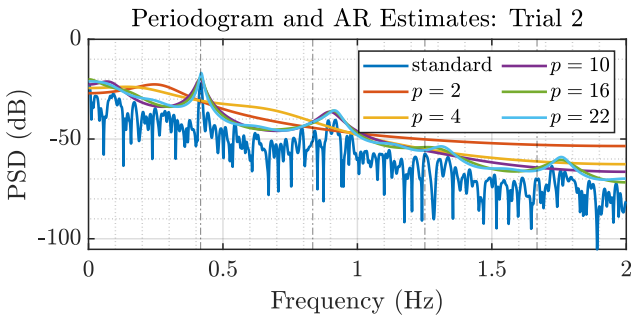
### 1.5.b Analysis of the RRI PSD Estimates

Figure 1.5.1 shows us distinct peaks for each trial that somewhat agrees with the breathing expected. We note that harmonics are difficult to distinguish, despite the large zero padding of the signal. Table 1.5.1 highlights how the observed peaks align with the expected peaks.

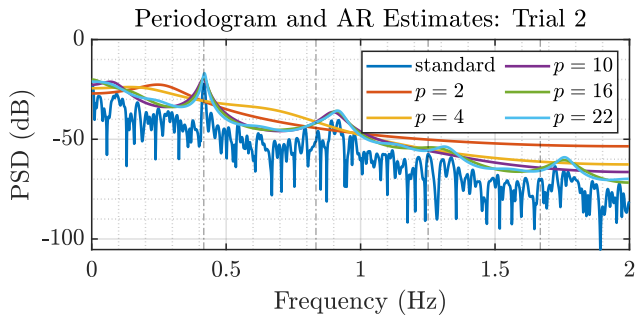
### 1.5.c AR PSD Estimate for the RRI Dataset




(a)




(b)



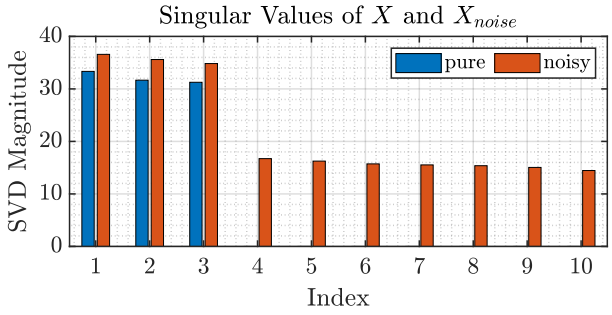

(c)

Figure 1.5.3: AR Estimate Periodograms. $p$ is the model order.

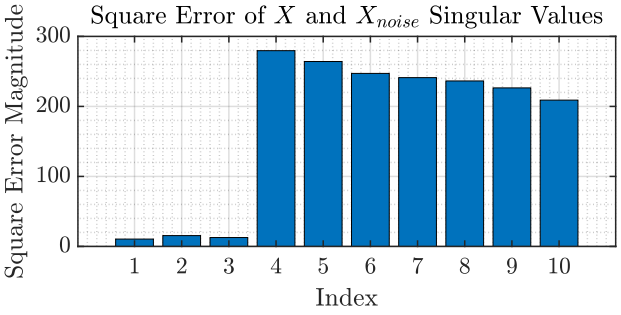
The AR spectral estimate in Figure 1.5.3 correctly identified the peak from models of order $p \gtrapprox 10$. We note that the estimate resembles a smooth envelope over our PSD and correctly identifies clear peaks in the standard PSD at $p = 10$. We also can see for higher order models that there are move fluctuations as the model starts to overfit for the input data's natural noisiness.

## 1.6 Robust Regression

### 1.6.a Single Value Decomposition (SVD)




(a) SVD




(b) Square Error

Figure 1.6.1

The rank of the input data would be indicated when the SVD Error increases or the non-zero SVD values end (for pure data) - hence the data is of rank 3.
Noise makes the SVD magnitude non-zero where we expect it to be zero for the pure input. Hence, in the real world where we would not have the pure dataset we would need to purely look at Figure 1.6.1 (a), if the SVD magnitude of the noise was comparable to that of the signal we would have difficulty establishing a criteria to distinguish signal and noise subspaces from the SVD method.

### 1.6.b Low Rank Approximation Error

By looking at the mean square error between the pure signal and the low-rank approximation of the noise, Figure 1.6.2, we note that the minimum error is when the low-rank approximation matches that of the pure data.

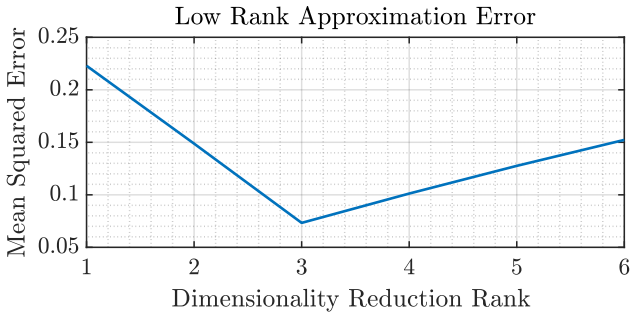
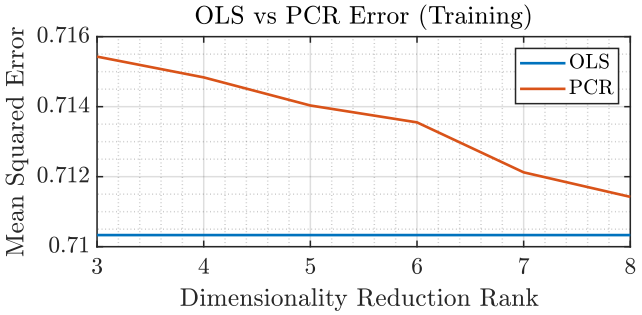
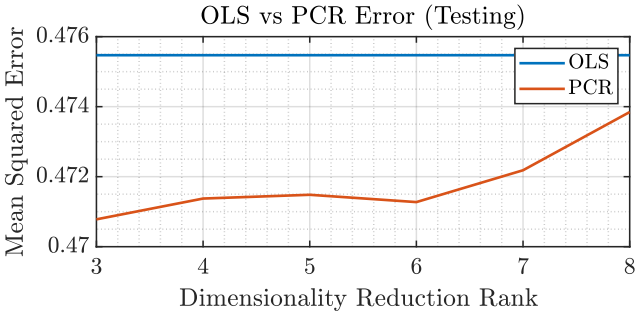



Figure 1.6.2: Effect of Changing Rank on the Approximation Error

### 1.6.c Ordinary Least Squares (OLS) & Principle Component Regression (PCR) Estimate Errors




(a) Training Dataset Error




(b) Testing Dataset Error

Figure 1.6.3

In both datasets we see that at the true model order the differences are tiny, in the training dataset we see that PCR underperforms by 0.85%, while in the testing dataset PCR outperforms by 1.05%. We can also note that decreasing the PCR model order further than the true model order, improves performance in the Training dataset but reduces performance in the Testing dataset.

### 1.6.d Ordinary Least Squares (OLS) & Principle Component Regression (PCR) Estimate Errors - Part 2

Again if we look at the true model order, Figure 1.6.4, PCR outperforms OLS by 0.79%. We note how further reduction in dimensionality decreases the performance of PCR. In regard to effectiveness of each scheme, if the model order is unknown and cannot be approximated accurately, OLS is a safer choice, however best performance - although marginal - can be achieved if the model order is known and PCR is implemented.

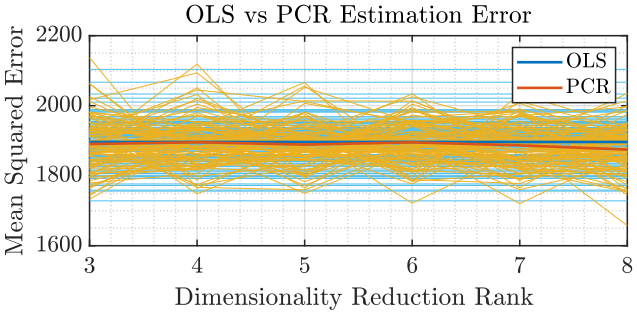



Figure 1.6.4: Mean Square Error over several Realisations

# 2 Adaptive Signal Processing

## 2.1 The Least Mean Square (LMS) Algorithm

### 2.1.a Example Correlation Matrix

### 2.1.b Example Learning Curve for an AR(2) Process

### 2.1.c Misadjustment

### 2.1.d Steady State Estimation

### 2.1.e Leaky LMS Derivation

### 2.1.f Example Leaky LMS for an AR(2) Process

## 2.2 Adaptive Step Sizes

### 2.2.a Gradient Adaptive Step Sizes (GASS) Comparisons for a MA(1) System

### 2.2.b NLMS Update Equation Equivalence

### 2.2.c Generalised Normalised Gradient Descent (GNGD) Example for a MA(1) System

## 2.3 Adaptive Noise Cancellation

### 2.3.a Adaptive Line Enhancer (ALE) Delay Investigation

### 2.3.b Adaptive Line Enhancer (ALE) Filter Order Investigation

### 2.3.c Adaptive Noise Cancellation (ANC) Comparison

### 2.3.d Adaptive Noise Cancellation (ANC) on real EEG Data

# 3 Widely Linear Filtering and Adaptive Spectrum Estimation

## 3.1 Complex LMS and Widely Linear Modelling

### 3.1.a Widely Linear Moving Average (WLMA) Process using Complex LMS (CLMS) & Augmented CLMS (ACLMS)

### 3.1.b CLMS and ACLMS Suitability with real (wind) data

### 3.1.c Balanced and UnBalanced Voltages - Investigating Circularity

### 3.1.d Derivation for estimating Frequency from Filter Weights

### 3.1.e Investigating the Accuracy of the Filter Estimates

## 3.2 Adaptive AR Model Based Time-Frequency Estimation

### 3.2.a Frequency Modulated (FM) Signal Investigated with an AR(1) Process

### 3.2.b Frequency Modulated (FM) Signal Investigated with CLMS

## 3.3 A Real Time Spectrum Analyser Using Least Mean Square

### 3.3.a The Discrete Fourier Transform (DFT) Formula and the Least Square Solution

### 3.3.b Fourier Transform in terms of Changes of Basis & Projections

### 3.3.c The DFT-CLMS vs. the AR(1) at Spectrum Analysis

### 3.3.d The DFT-CLMS on real (EEG) Data