

BIOMEDICAL SIGNALS

LAB SESSION 3 WITH MATLAB SPECTRAL ESTIMATION IN STATIONARY SIGNALS

FIRST PART: PRE-LABORATORY WITH NON-BIOMEDICAL SIGNALS

The same techniques used in system identification (with input and output signals) are used with a single time series or signal (including bioelectric or biomedical). The application of parametric and non-parametric methods (the latter are based on the DFT) allow the calculation of the Power Spectral Density, thus allowing obtaining information on the frequency components of these signals. This allows, in the case of biomedical signals, to extract characteristics of the signal that otherwise could not be detected and that will facilitate the support to clinical diagnosis.

The System Identification toolbox can therefore be used for the signals spectral estimation.

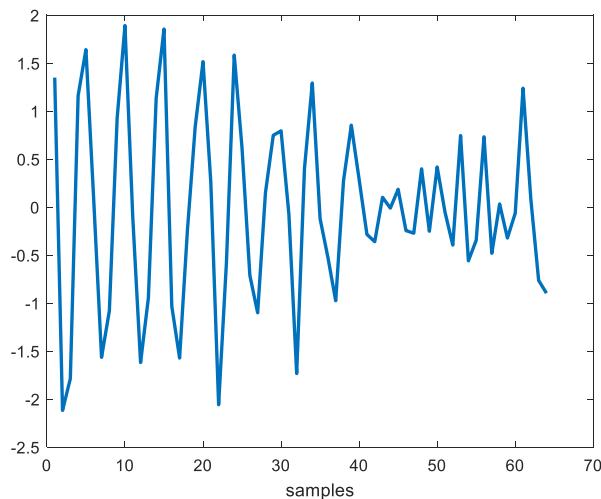
1. SYNTHETIC TEST SIGNAL

The test signal is obtained from the book: The complex data in L. Marple: S.L. Marple, Jr, Digital Spectral Analysis with Applications, Prentice-Hall, Englewood Cliffs, NJ 1987, and it is uploaded as follows:

```
>> load marple
```

The real part of the signal is selected for the spectral estimation and, then, it is displayed:

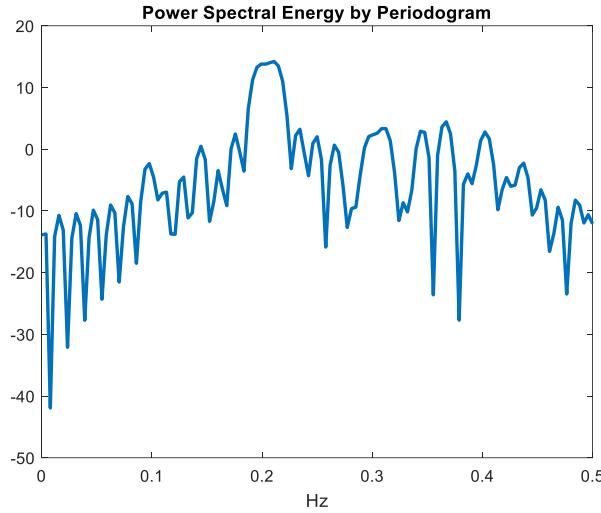
```
>> marple =real(marple);  
>> plot(marple),  
>> xlabel('samples')
```



1.1 Periodogram.

Firstly, the periodogram is calculated (or 'sample spectrum'), and the corresponding Power Spectral Density function estimated from this non-parametric method is displayed:

```
>> [Pxx,f]=periodogram(marple,[],256,1,'onesided');
>> plot(f,10*log10((Pxx/2)/(256/64)));
>> xlabel('Hz');
>> title('Power Spectral Energy by Periodogram');
```

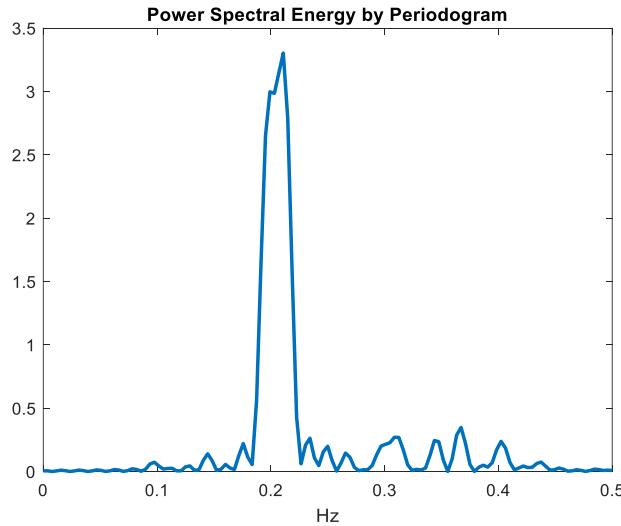


The periodogram is scaled firstly by 2 because the option 'onesided' multiplies by 2 the PSD in order to have the whole signal energy in only half frequency range. In addition, 256 points have been considered in the calculation of FFT inside the periodogram function when the signal record has only 64 samples. According with the calculations done by this function, the resulting PSD must be later scaled dividing by (N/L) being N the number of FFT points and L the number of signal samples.

As no specific sampling frequency has been considered at the beginning, we have introduced the unit value which is considered by default. Therefore, the frequency range up to which the spectral density is displayed is 0.5, corresponding to the Nyquist frequency, since it is known that the spectrum is symmetric and periodic.

We have plotted the power spectrum in decibels (dB), however, the most common way to visualize the PSD in biomedical applications is with linear units both on the frequency axis and on the power axis (y axis). In addition, we have considered 256 points for the FFT in the periodogram as an input parameter. As the signal has only 64 samples, zeros are added up to the number of points. Thus, the first 128 values (points) of the FFT correspond to the frequency range from 0 Hz to the Nyquist frequency ($1 / 2T$ Hz or π / T rad/s).

```
>> plot(f,(Pxx/2)/(256/64));
>> xlabel('Hz');
>> title('Power Spectral Energy by Periodogram');
```



Alternatively, the periodogram can be calculated directly with the equation using the FFT as follows:

```
>> fmarple=fft(marple,256);
>> Pxxf=abs(fmarple(1:129)).^2/256;
>> plot(f, (Pxx/2)/(256/64),f,Pxxf);
>> xlabel('Hz');
>> title('Comparison of Power Spectral Energies by Periodogram');
```

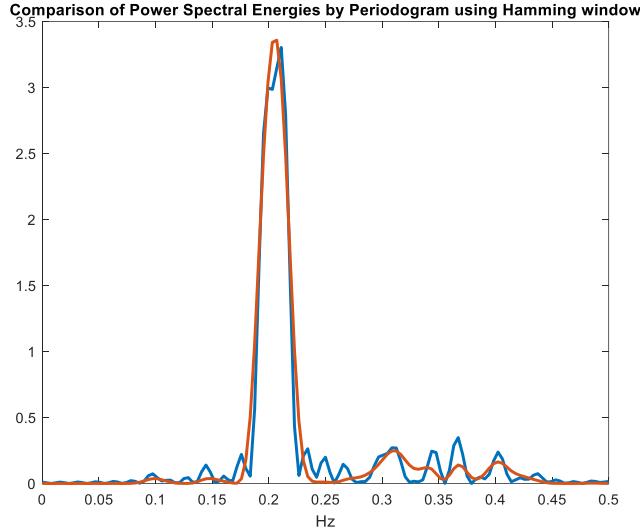
Curves are totally overlapped because PSD are the same.

To observe the time window effect, we can consider a Hamming window:

```
>> wmarple=hamming(length(marple)).*marple;
>> fwmarple=fft(wmarple,256);
>> Pxxfw=abs(fwmarple(1:129)).^2/256;
```

The Hamming window reduces the initial signal energy, thus there is a factor to be considered to rescale the PSD:

```
>> factor= sum(marple.^2)/sum(wmarple.^2);
>> Pxxfw=factor.*Pxxfw;
>> plot(f, Pxxf, f, Pxxfw);
>> xlabel('Hz');
>> title('Comparison of Power Spectral Energies by Periodogram using Hamming window');
```

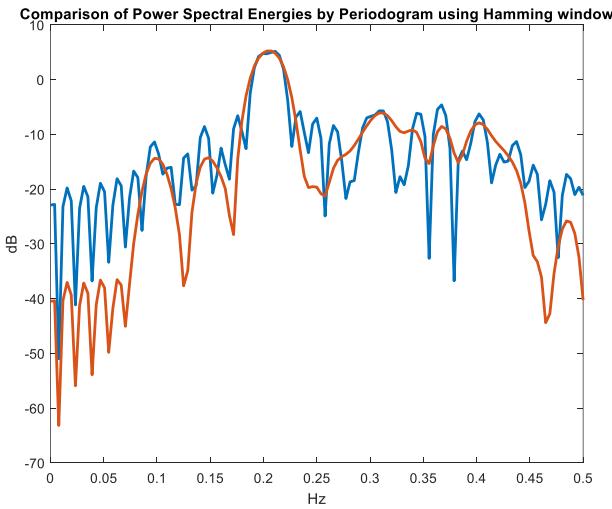


Using the periodogram command with the Hamming window:

```
>> [PxxH,f]=periodogram(marple,hamming(length(marple)),256,1,'onesided');
```

The last PSD obtained must be rescaled according to the initial signal energy:

```
>> factor2=(sum(Pxx)+sum(Pxx(2:end-1)))/ (sum(PxxH)+sum(PxxH(2:end-1)));
>> plot(f, 10*log10((Pxx/2)/(256/64)),f,10*log10((factor2*PxxH/2)/(256/64)));
>> xlabel('Hz');
>> title('Comparison of Power Spectral Energies by Periodogram using Hamming window');
```



The resolution with Hamming window is worse (in red color) but the leakage has been reduced a lot.

1.2 Welch periodogram

As the signal recording corresponds to 64 samples, we can observe a relatively narrow frequency resolution in the periodogram. An averaged periogram (Welch periodogram) can be performed where each segment is 32 samples,

which will provide a spectral resolution of 1/32 Hz. This estimate, therefore, will have lower resolution, but its variance will also be lower.

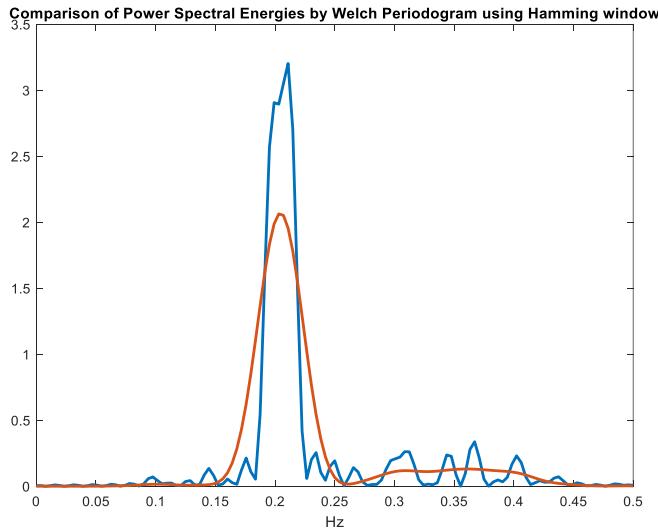
```
>> number_segments = 3;
>> overlap = 0.5;
>> samples_segment = floor(length(marple)/(number_segments-(number_segments-1)*overlap));
>> [PxxW,f]=pwelch (marple,samples_segment,floor(samples_segment*overlap),256,1,'onesided');
```

Three segments have been considered to average with an overlap of 50%. The length of each segment (samples_segment) is calculated following the formula above. In this case, the length is 32 sample. A Hamming window is applied by default. The resulting PSD must be rescaled as follows:

```
>> factor3=(sum(Pxx)+sum(Pxx(2:end-1)))/ (sum(PxxW)+sum(PxxW(2:end-1)));
```

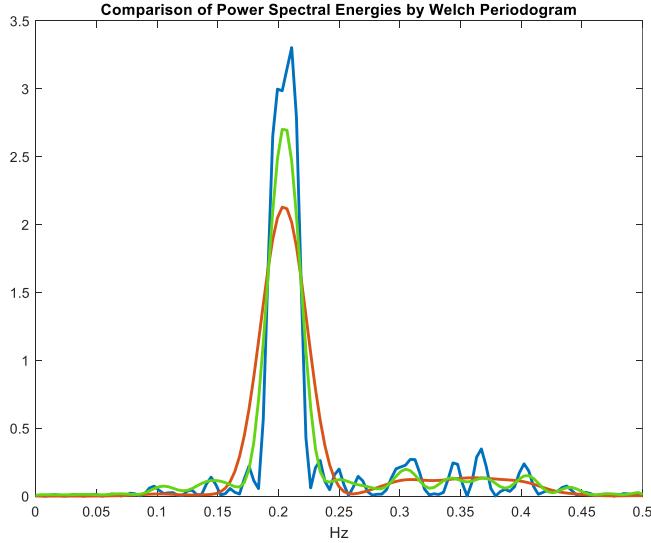
After calculating said averaged periodogram, both power spectral densities are plotted (averaged in red and not averaged without time window in blue) for comparison:

```
>> plot(f,(Pxx/2)/(264/64),f,factor3*(PxxW/2)/(264/64));
>> xlabel('Hz');
>> title('Comparison of Power Spectral Energies by Welch Periodogram using Hamming window');
```



It is possible to calculate the Welch Periodogram considering another time window, for example, a rectangular time window as follows:

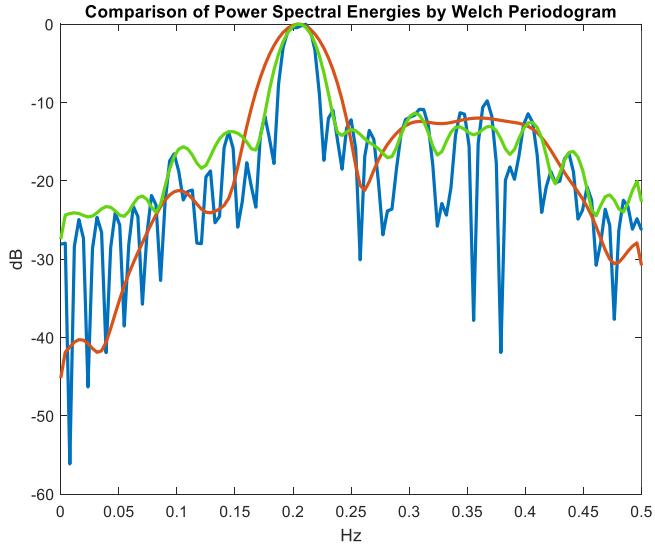
```
>> [PxxWR,f]=pwelch (marple,ones(samples_segment,1),floor(samples_segment*overlap),256,1,'onesided');
>> factor4=(sum(Pxx)+sum(Pxx(2:end-1)))/ (sum(PxxWR)+sum(PxxWR(2:end-1)));
>> figure
>> plot(f,(Pxx/2)/(264/64),f,factor3*(PxxW/2)/(264/64),f,factor4*(PxxWR/2)/(264/64));
>> xlabel('Hz');
>> title('Comparison of Power Spectral Energies by Welch Periodogram using Hamming window');
```



Periodogram is in blue color, Welch Periodogram with rectangular and Hamming window are in green and red colors, respectively

In decibels (dB) and comparing with respect to the maximum PSD value to better compare the resolution and leakage between them:

```
>> figure
>> plot(f,10*log10(Pxx/max(Pxx)),f, 10*log10(PxxW/max(PxxW)),f, 10*log10(PxxWR/max(PxxWR)));
>> xlabel('Hz');
>> title('Comparison of Power Spectral Energies by Welch Periodogram using Hamming window');
```



Periodogram in blue color, Welch Periodogram with rectangular and Hamming window in green and red colors, respectively. It is clear the loss of resolution by averaging with Welch periodogram. In addition, using a Hamming window the frequency resolution was worse but the leakage has been reduced (lower variance) smoothing the PSD. All functions have been normalized to their maximum values for a easier comparison.

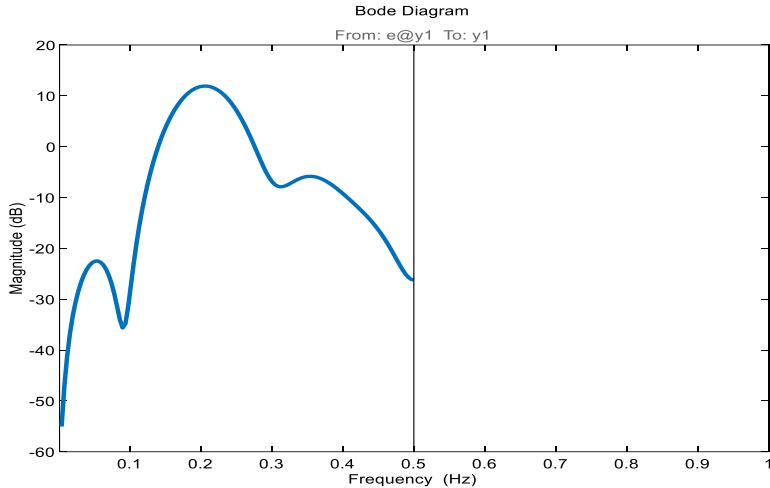
1.3 Correlogram

The correlogram, also called the Blackman-Tukey estimation, allows the calculation of the Spectral Power Density by a non-parametric method (applying the FFT on the estimated autocorrelation function):

```
>> corr1 = spa(marple);
```

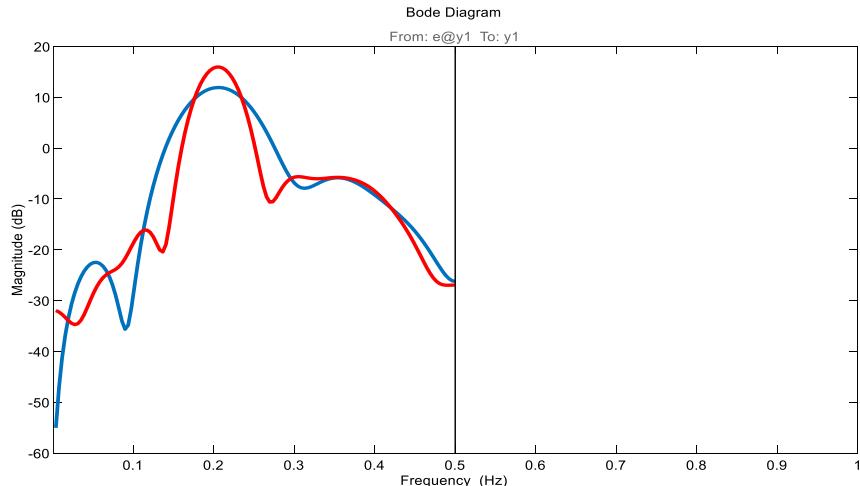
Said spectral estimate is drawn:

```
>> ffplot(corr1);
```



The PSD is too smooth, the duration, by default, of the window used by the command is very small for this amount of signal data: one tenth (10%), which means only 6 samples for the autocorrelation function. A longer window duration is chosen (20), and the new correlogram is displayed, in a green line, comparing it with the averaged blue periods:

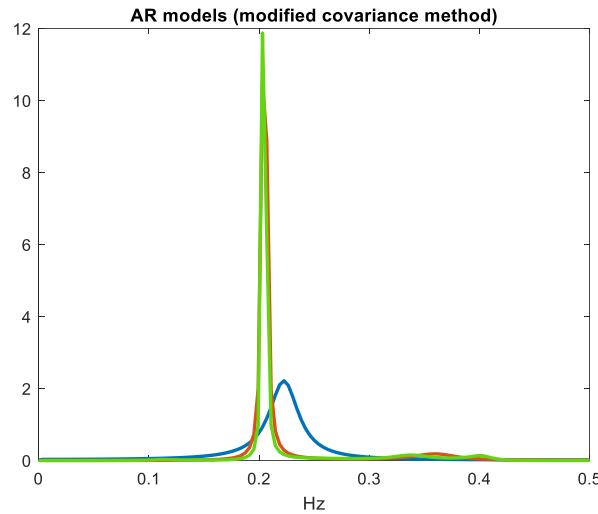
```
>> corr2 = spa(marple,20);
>> hold on;
>> ffplot(corr2,'r');
```



1.5 AR(5) Model: modified covariance method.

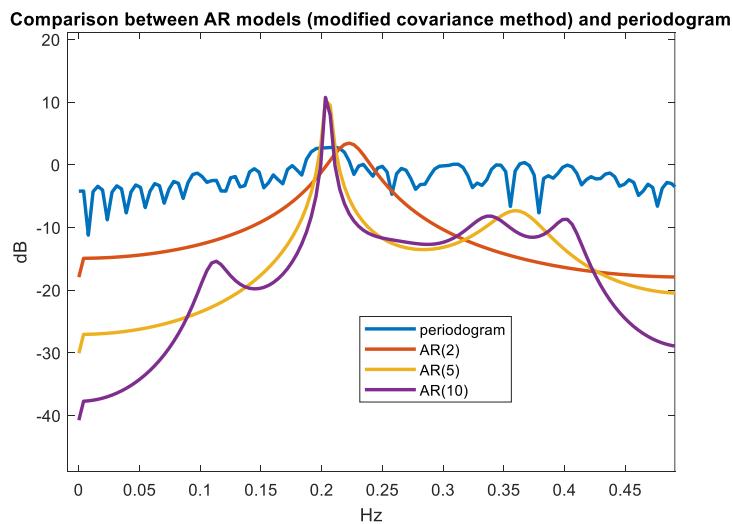
Signal data is adjusted to an AR model of order 2, 5 and 10 and the PSD is calculated:

```
>> [Parmcov2,f]= pmcov(marple,2,256,1,'onesided');
>> [Parmcov5,f]= pmcov(marple,5,256,1,'onesided');
>> [Parmcov10,f]= pmcov(marple,10,256,1,'onesided');
>> plot(f,(Parmcov2/2)/(256/64), f,(Parmcov5/2)/(256/64), f,(Parmcov10/2)/(256/64));
>> xlabel('Hz');
>> title('AR models (modified covariance method)');
```



AR models are compared with the periodogram using a rectangular window:

```
>>
plot(f,(10*log10(Pxx/2)/(256/64)),f,10*log10((Parmcov2/2)/(256/64)),f,10*log10((Parmcov5/2)/(256/64)),f,10*log10((Parmcov10/2)/(256/64)));
>> xlabel('Hz');
>> ylabel('dB')
>> title('Comparison between AR models (modified covariance method) and periodogram');
```

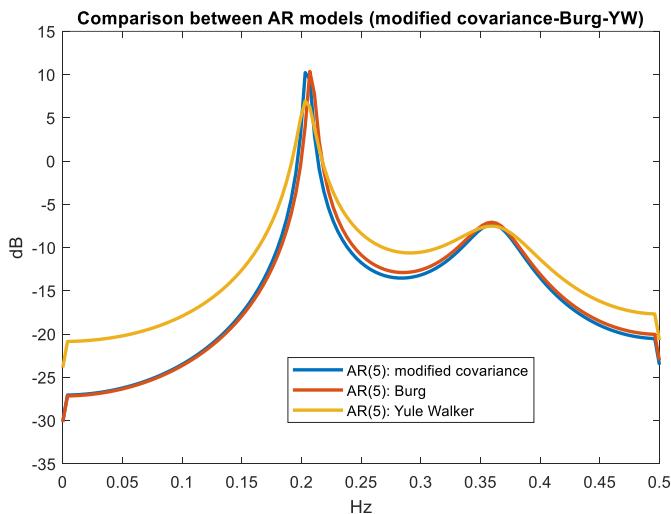


Resolution with AR models is higher.

1.6 AR(5) model: Burg and Yule Walker methods.

Other AR models can be estimated by the methods of Burg and Yule Walker as follows:

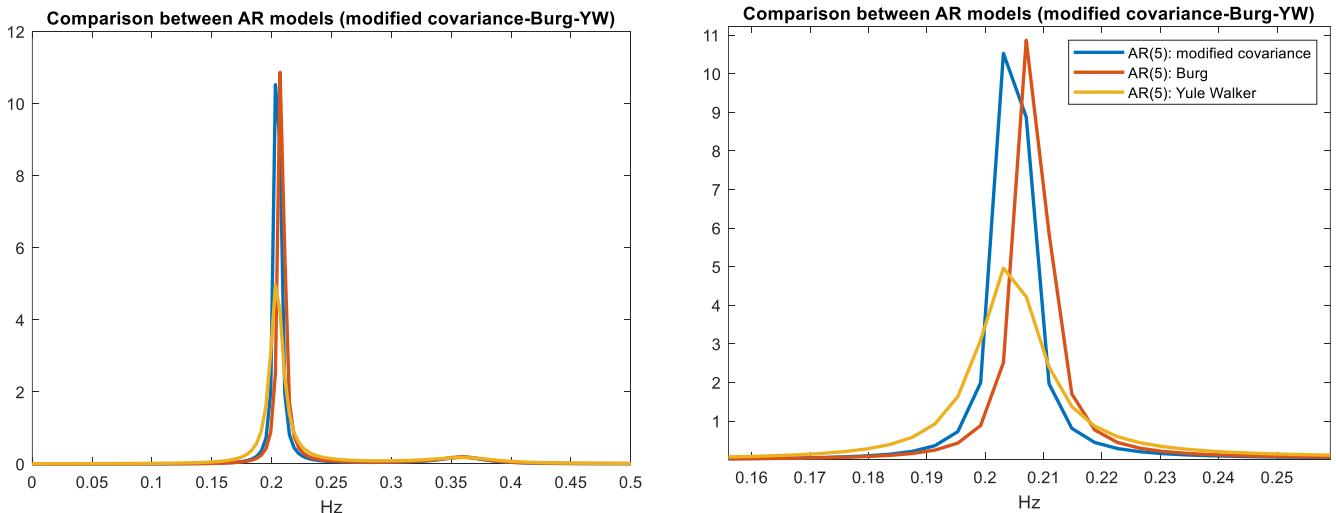
```
>> [Parb5,f]= pbburg(marple,5,256,1,'onesided');
>> [Paryw5,f]= pyulear(marple,5,256,1,'onesided');
>> plot(f,10*log10((Parmcov5/2)/(256/64)),f,10*log10((Parb5/2)/(256/64)),f,10*log10((Paryw5/2)/(256/64)));
>> xlabel('Hz');
>> ylabel('dB')
>> title('Comparison between AR models (modified covariance-Burg-YW)');
>> legend('AR(5): modified covariance','AR(5): Burg','AR(5): Yule Walker')
```



The three methods (particularly the first two) find very similar spectral estimation: Yule Walker equations get an poorer estimation of PSD (spectral peak smaller and wider).

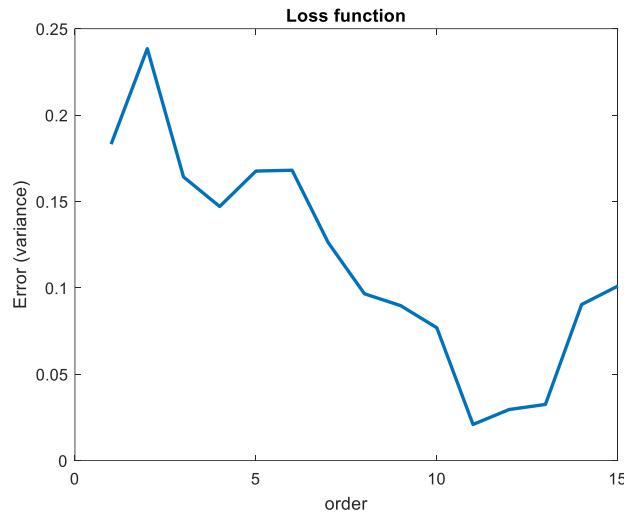
Finally, PSDs can be plotted not using dB but linear scale as follows:

```
>> plot(f,(Parmcov5/2)/(256/64),f,(Parb5/2)/(256/64),f,(Paryw5/2)/(256/64));
>> xlabel('Hz');
>> title('Comparison between AR models (modified covariance-Burg-YW)');
>> legend('AR(5): modified covariance','AR(5): Burg','AR(5): Yule Walker')
```



The model order can be selected by several criteria using half signal for the training and the second half of the signal for the validation. The error evaluation permits to select the most appropriate order. The Akaike Information Criteria (AIC) can be used firstly:

```
>> data1=iddata(marple(1:32),[],1);
>> data1=iddata(marple(33:64),[],1);
>> Vm=arxstruc(data1,data2,[1:15]');
>> plot(Vm(2,1:end-1),Vm(1,1:end-1));
```



The order according to the AIC:

```
>> naic=selstruc(Vm,'aic')
```

```
naic =
11
```

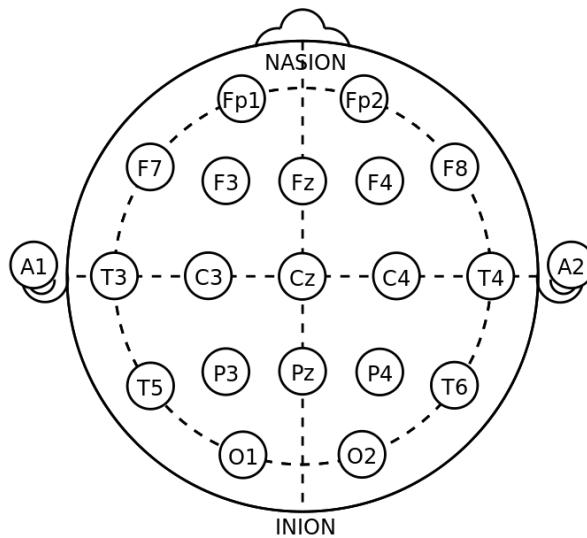
```
>> nmdl=selstruc(Vm,'mdl')
```

```
nmdl =
11
```

2. SPECTRAL ESTIMATION OF EEG SIGNALS DURING SLEEP.

In this section, EEG signals will be considered to test different methods for spectral estimation. EEG signals were recorded during sleep. The available data is the following:

- Data files associated with different sleep stages determined by a neurophysiologist: Awake, Stage S1, Stage S2, Slow Wave Activity/Sleep (SWS) and REM.
- In each file there are 21 biological channels recorded during 30 seconds with a sampling frequency of 100 Hz: The first two are EOG and the other 19 channels are EEG signals where the electrodes were located in the following places on the scalp following the international system 10-20: 'Fp1'; 'Fp2'; 'F7'; 'F3'; 'Fz'; 'F4'; 'F8'; 'T3'; 'C3'; 'Cz'; 'C4'; 'T4'; 'T5'; 'P3'; 'Pz'; 'P4'; 'T6'; 'O1' and 'O2' (with this order).



It is well known that power of frequency bands from EEG changes during sleep when the person moves to a new stage, for example deeper sleep. The frequency bands are **delta**, **theta**, **alfa**, **sigma** and **beta** among others with the associated frequency ranges shown in the slide 18 of the presentation of Chapter 1. An extra band will be also considered in this Computer Session: **low delta** between 0.5 Hz and 2 Hz. Sleep scoring was updated by the American Academy of Sleep Medicine (AASM) in 2007 as explained in the document “Chapter 3. Sleep stages and scoring techniques” available at ATENEA. The sections “Electroencephalographic Activity During Wakefulness and Sleep” and “Stages of Sleep Stages” are specially interesting to learn more about the different activities produced by the brain at specific stages and their cerebral location. In addition, the official document published by the AASM in 2007 about the rules to score the sleep in different stages is available at ATENEA (document named “The AASM-Manual for Scoring-of Sleep and Associated Events”)

These powers for each frequency band are calculated from the Power Spectral Density (o being more strict from the Power Spectral Energy) and can be absolute powers or relative ones when they are normalized by the total power of the signal. This function can be also estimated using different spectral techniques as presented with synthetic signals in the previous Section.

The following estimators will be tested in EEG signals:

- non parametric methods such as periodogram and modified periodogram with a Hamming window) and Welch periodogram;
- parametric methods such as AR models with different orders.

The objective is to evaluate the impact of the estimator in:

- The PSD functions obtained from EEG signals.
- The relative power indexes: few bands can be more affected than the others of selecting different spectral estimators.
- The brain location of predominant rhythm/band can be also affected or different depending on the spectral estimator.
- The analysis of the relative powers during sleep at different sleep stages where it is expected more sigma activity in the Stage S2, tending to lower frequencies with deeper sleep, more alpha activity during awake etc.

It is usually used to consider EEG signal stationary during 5-seconds epochs. Thus, the PSD should be estimated each 5 seconds interval and the powers are average from the 6 epochs available in the MATLAB files for each Stage.

In this document we'll start working with the Awake signals as an example and this should be reproduced for the other stages:

Load AWAKE

We select the first 5 seconds epoch and the last 19EEG channels:

```
eegi=signals(1:500,3:21);
```

Frequency bands come from 0.5 Hz to 35 Hz, thus it is convenient to remove the very low frequency components associated mainly with the movement and other brain activities different to delta rhythms. This action will permit to estimate better the PSD. You can check the estimation twice: high frequency filtering or without filtering:

```
[b,a]=ellip(6,0.5,40,.4/50,'high');  
eeg=filtfilt(b,a,eegi);
```

% Method 1

Firstly, the periodogram is estimated considering a FFT with 1000 points:

```
[Pxx,f]=periodogram(eeg,[],1000,100,'onesided'); %Method 1
```

% Method 2

Then, the periodogram using a Hamming window is calculated:

```
[PxxH,f]=periodogram(eeg,hamming(500),1000,100,'onesided'); %Method 2
```

% Method 3

Finally, the averaged Welch periodogram is obtained:

```

number_segments = 4;
overlap = 0.5;
samples_segment = floor(length(eeg)/(number_segments-(number_segments-1)*overlap));
[PxxW,f]=pwelch(eeg,samples_segment,floor(samples_segment*overlap),1000,100,'onesided');

% Method 4

```

Then the PSD is calculated by an AR model using the Burg method: the first selection is a low order of 2:

```
[Parb2,f]=pburg(eeg,2,1000,100,'onesided');% Method 4
```

```
% Method 5
```

and then, a high order of 50:

```
[Parb50,f]=pburg(eeg,50,1000,100,'onesided'); % Method 5
```

```
% Methods 6 and 7
```

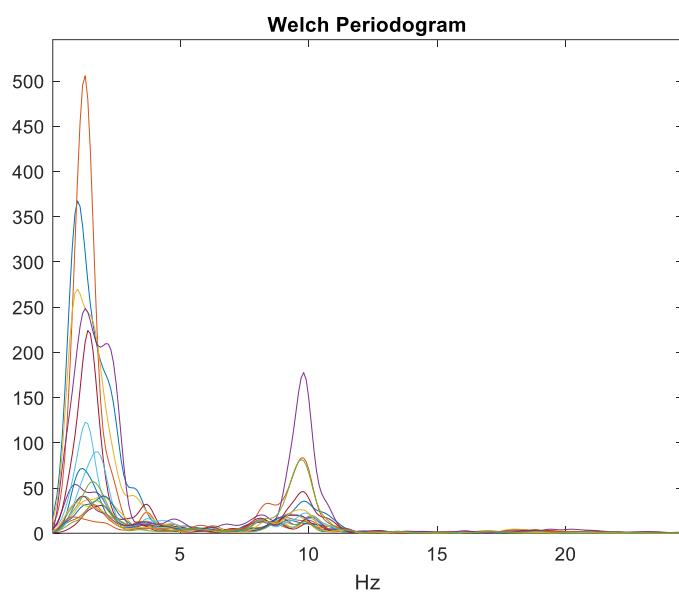
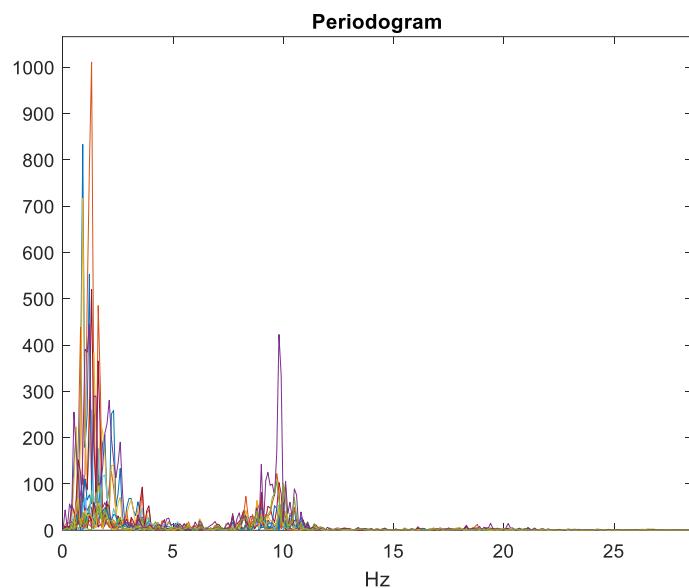
Finally, orders proposed following the AIC and MDL criteria were selected. As each EEG channel/signal in the first epoch can result a different order, the average order between the 19 channels proposed by each criterion is selected and applied in all EEG channels:

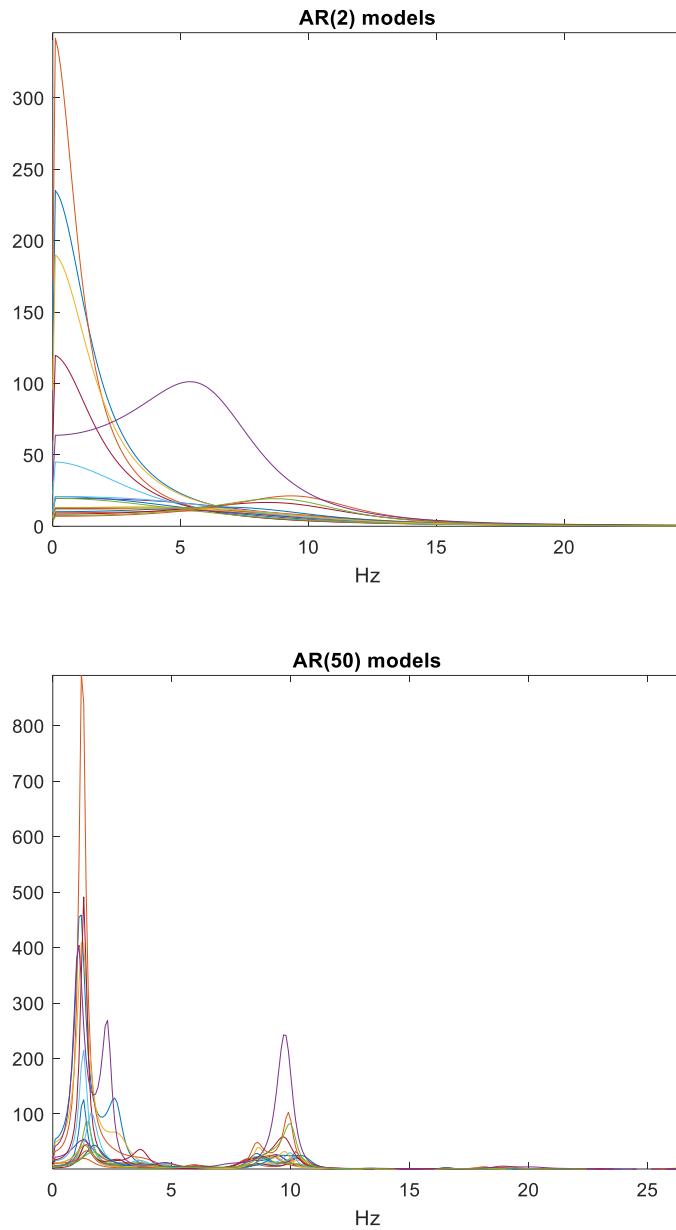
```

for n=1:19,
    data1= iddata(eeg(1:250,n),[],100);
    data2= iddata(eeg(251:500,n),[],100);
    V1=arxstruc(data1,data2,[1:60]');
    nnaic(n)=selstruc(V1,'aic');
    nnmdl(n)=selstruc(V1,'mdl');
end
naic=round(mean(nnaic))
nnmdl=round(mean(nnmdl))
[Parbaic,f]=pburg(eeg,naic,1000,100,'onesided'); % Method 6
[Parbmdl,f]=pburg(eeg,nnmdl,1000,100,'onesided'); % method 7

```

PSD from some of these methods are presented in the following figures:

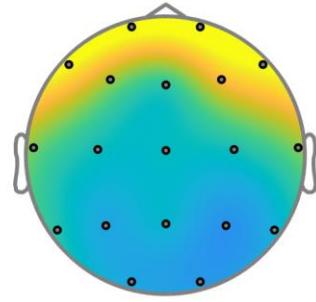




All PSD from the 19 EEG channels (epoch 1) are plotted in each figure

Relative powers must be calculated for each PSD function as the sum of energy within the frequency range, that is absolute powers, divided by the total power. For each band (low delta, delta, theta, alpha, sigma, beta) the average value in the 6 epochs can be also calculated. Then 19 values for delta band for example are obtained in a specific sleep stage. This can be plotted in topographic maps like this:

```
draw_topogram(delta1, [0 .7]);
```



Each value is associated with a color from blue (minimum value in this case 0) to yellow (maximum value in this case 0.7 which means 70% for delta relative power). We can observe that most of the delta activity is focused in the frontal area. The plot is interpolated between channels for a better visualization. This can be obtained for each band and each method and finally every sleep stage. We should compare the plots according to the objective mentioned above.

This function is available as a P-file in the Atenea. Thus, only the running version can be used. The help text is the following:

```
function draw_topogram(data, limits, ax, cmap)

% DRAW_TOPOGRAM Draws a 19-channel topographic map of 10/20 EEG signals
%
% Example1:
%
% figure; draw_topogram([0 0 3 4 5 4 3 2 2 9 2 2 7 7 15 7 7 12 12]');
%
% Example2:
%
% cmap = [0.1 0.1 0.4;
%      0.4 0.4 0.6;
%      0.7 0.7 0.8;
%      1.0 1.0 1.0;
%      0.8 0.7 0.7;
%      0.6 0.4 0.4;
%      0.4 0.1 0.1];
% figure;
% ax1 = subplot(211);
% ax2 = subplot(212); plot(rand(5));
% draw_topogram([0 0 1 1 3 1 1 0 2 3 2 0 -1 -1 -3 -1 -1 -2 -2]', [-3 3], ax1, cmap);
%
% This last example is a statistical probability map where values 0, 1,
% 2 and 3 represent different p-values (no significance, tendence, significant
% difference and very significant difference, respectively). Sign
% indicates direction of the detected change.
```

Thus, it is possible to select the minimum and maximum values associated with the extreme colors depending on our interest of visualizing. For example, if we are interested in comparing

the 7 spectral methods in a frequency band a figure with 7 topograms/subplots would be useful to compare the brain distribution/location of a specific power. Then, for a good visual comparison the color scales should be the same in all topograms, that is, **the same limits [min max] should be considered in the figure**. The *min* must be 0 and the *max* should be the maximum value for all channels and all methods considered in the figure.

Summarizing, you should explore the spectral estimators proposed in this Section. To present and explain differences of the PSD functions using different methods on the EEG. Can we observe the same tendency in all frequency bands during deeper sleep stages? Is there any band more affected than others by the estimator selection? which model orders are proposed in the AIC and MDL criteria? You can also try other orders to explore changes in the PSD, etc.

For this purpose, the following results are requested (not limited):

1. For the sleep stage S1 only (in order not to do an exhaustive work with all stages):
 - a. To compare the PSD functions for the channels Fz, Cz and Pz obtained in all 7 spectral methods in a figure (consider only the epoch 1).
 - b. To compare the topograms of all 7 spectral methods in a figure for each relative power (indicate the *max* value considered in each of both figures).
- We can conclude that if we obtain different results with a specific method, they are not feasible. However, similar results with most of the methods means feasibility. Discuss this with the figures obtained.
2. To observe changes of each relative power in frequency band during sleep (comparing the five sleep stages in a figure with the same *max* value for the five topograms which must be indicated). You can select only two spectral methods: selecting one of them with similar performance in most methods and another different to observe the influence of the estimator in checking these changes (each method in separated figures). Select the Welch periodogram and AR with order selected by MDL criteria.

Description, smart comments and conclusions are very welcome.