KULEUVEN

BIOMEDICAL ENGINEERING

Assignment 5: ICA/Wavelets

Jan Everaert and Alexander Meulemans

Prof. S. Van Huffel
Prof A. Bertrand

December 22, 2017

Contents

Chapter 1

Exercise 5.1.1

1.1 Independent Component Analysis on signal X

In this exercise 4 different sources are used. These (normalized) sources are plotted in figure 1.1 on the left side. Source 1 is 10s of an ECG recording at sample fequency of 500Hz. The other 3 sources simulate noise: white gaussian noise, powerline interference and a sawtooth to represent e.g. instrument noise. To simulate measurements of different mixtures (x), the sources (s) are multiplied by a mixing matrix (A). This gives the signals showed in figure 1.2. Since A is a square (4 by 4) matrix, the number of mixture signals (which would correspond with the number of measurements) equals the number of sources.

Now we can perform Independent Component Analysis (ICA) on these four mixture signals (x). In this analysis the following equation is solved: $y = W * x$, with both y and W unknown. The solution is a set of independent component signals (y) and a demixing matrix (W). These independent component signals are an estimation of the different (independent) source signals, which are usually unknown, present in the mixture signals (these are usually measured). The result of this analysis is presented in figure1.1, where the normalized independent components (on the right side) are linked with the original source signals. The normalization is needed because the energies of the sources cannot be determined: both y and W are unknown, so any scalar multiplier in one of the sources y_i could always be canceled by multiplying the the corresponding row of W with the same scalar. The sign of the signals must be matched since the sign of the independent components is random (same reason as above, with scalar = -1). To link the independent components with the original sources, the cross correlation between each component and each source is calculated and the pairs with the maximum correlation are taken. The RMSE between the independent signals (estimated sources) and the corresponding sources can be calculated and the results are displayed in table1.1

1.1.1 Artifact removal using ICA

We can now use these source estimations to eliminate a certain component from the mixed signals (that would be measurements). Since the mixed signals (x)

Table 1.1: RMSE estimated sources

Signals	RMSE
Source 1 - IC 1	0.062807
Source 2 - IC 2	0.072028
Source 3 - IC 3	0.1075
Source 4 - IC 4	0.020138

can be recomposed based on the source estimation (y) and the (pseudo)inverse of the demixing matrix W: $x = y * W^+$, a certain estimated source can be left out by setting the corresponding column of W^+ to zero. In this way, for example the sawtooth signal can be eliminated from the mixed signals. The effect of removing the sawtooth signal is given in figure1.3 where on top, one of the original mixed signals (the third) and the same mixed signal with the sawtooth removed is plotted, and on the bottom the difference between those two signal. If we now perform ICA again on the newly formed mixed signals, only three source estimations will be found as can be seen in figure1.4

1.1.2 Artifact removal using single channel filters

Although this is a multichannel signal, the processing methods of single channel signals can be still applied. Each channel of the multichannel signal is than treated as an individual channel. And so the filter techniques for single channel signals can be used to filter the two noise sources that are left in the mixture signals. A quick look at the frequency spectra of the four sources already gives a good indication on how well frequency domain based filters would work. In figure1.5, the frequency responses of the sources are plotted. The frequency response of the ecg source (useful source 1) is the largest in the $0-50$ Hz region. The white noise source (source 2) has a frequency response magnitude that is almost constant over the whole frequency range. The frequency response of the sawtooth source (source 3) is also larger in the low frequency region and therefor it would be not possible to filter this source from the mixed signals by use of frequency domain based filters. Luckily for us, this signal was already removed earlier. The last noise source (source 4) is a perfect 50Hz sine wave. This can also clearly be observed in the frequency response.

The noise from source 4, 50Hz sine wave, can easily be filtered out by the use of a notch filter. The transfer function for this filter is given in equation 1.1, where $z_{1,2} = cos(\theta_0) \pm i * sin(\theta_0)$. The filter has zeros at $\theta_0 = \pm\frac{f_0}{f_s}$ (f_s is the sample freq and equals 500Hz, f_0 is the freq of the sin wave that has to be filtered and equals 50Hz). There are also four poles present of which two are placed at zero, and two are placed close to the zeros of the transfer function.The latter has as effect that the frequency range that is filtered out by the zeros narrows. And since the noise source is exact 50Hz this is desirable. The downside is that a small nonlinearity in the phase response is introduced around 50Hz as ca be seen on figure1.6. However, sine the non linearity is small and is situated on the edge of the useful frequency spectrum of the ecg, the value of p can be chosen relatively high (is this example $p = 0.9$).

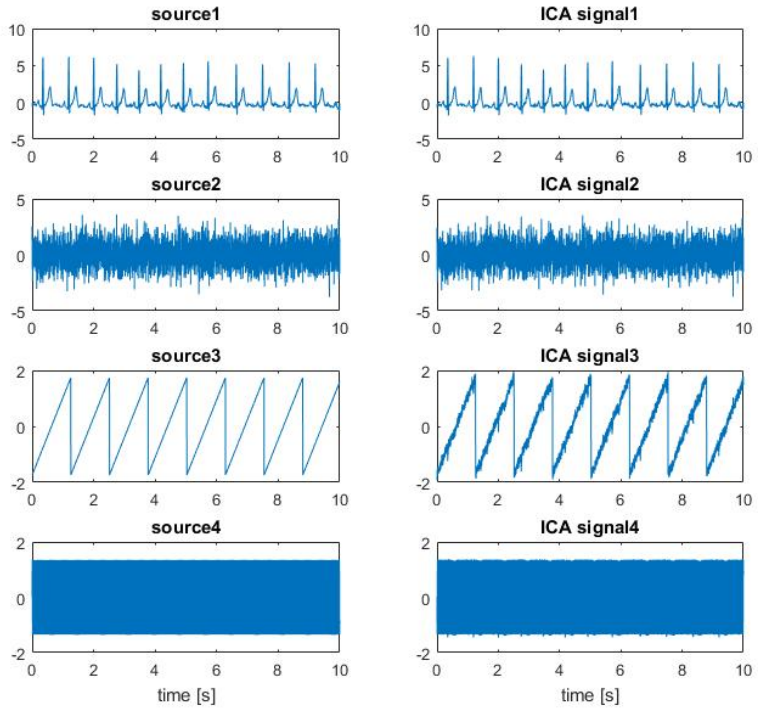

Figure 1.1: The original sources (left) and the IC from he ICA of x1 (right)

$$H(z) = \frac{(z - z_1) * (z - z_2)}{(z - p * z_1) * (z - p * z_2)} = \frac{z^2 - 2 * cos(\theta_0) + 1}{z^2 - 2 * p * cos(\theta_0) + p^2} \tag{1.1}$$

The result of this notch filter is presented in figure1.7. Only one of the four mixture signals is displayed in the figures, because, as the difference between the different mixture signals is rather small (especially for the frequency spectrum), the effect on the other mixture signals will almost be the same. From this figure, as well in the time as in the frequency domain, can be seen that the 50Hz sin wave is removed without disturbing the ECG signal too much.
For filtering of the random white noise, a low pass filter can be used. This will only filter the higher frequency content of the noise signal out, but since the low frequency spectrum overlaps with the useful frequency spectrum, this can not be removed by a frequency domain based filter. A Butterworth filter with a cutoff frequency of 50Hz (the characteristic frequency range of the ECG signal) and n=3 is used. Figure1.8 shows the result of this filter, again for one of the four signals (but the effect on the three other is the same).

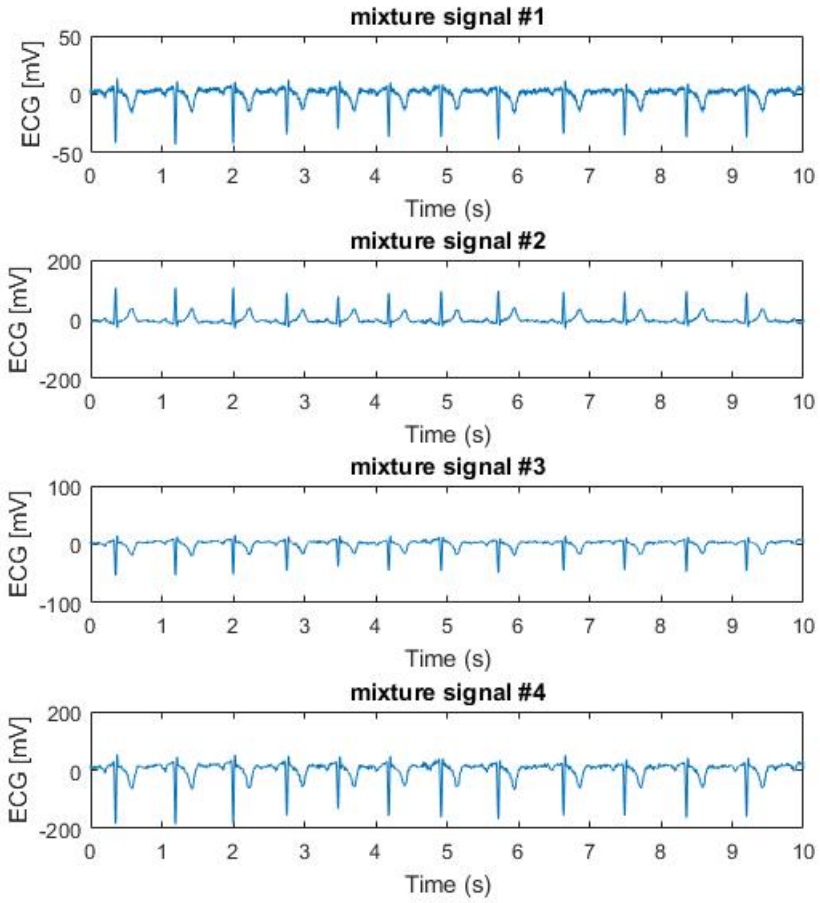

Figure 1.2: The mixed signals

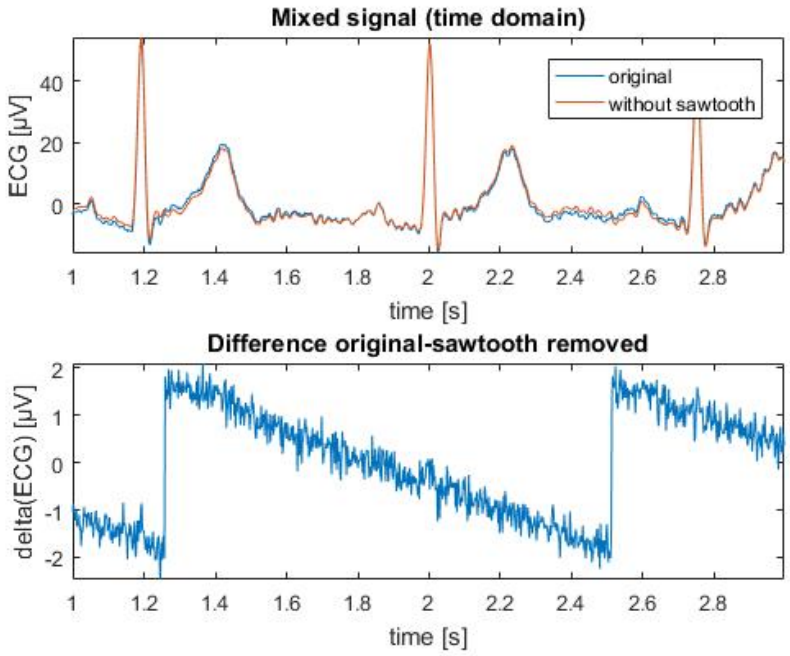

Figure 1.3: The original mixed channel 1 and the mixed channel 1 with the sawtooth signal removed (top) The difference between the two signals on top (bottom)

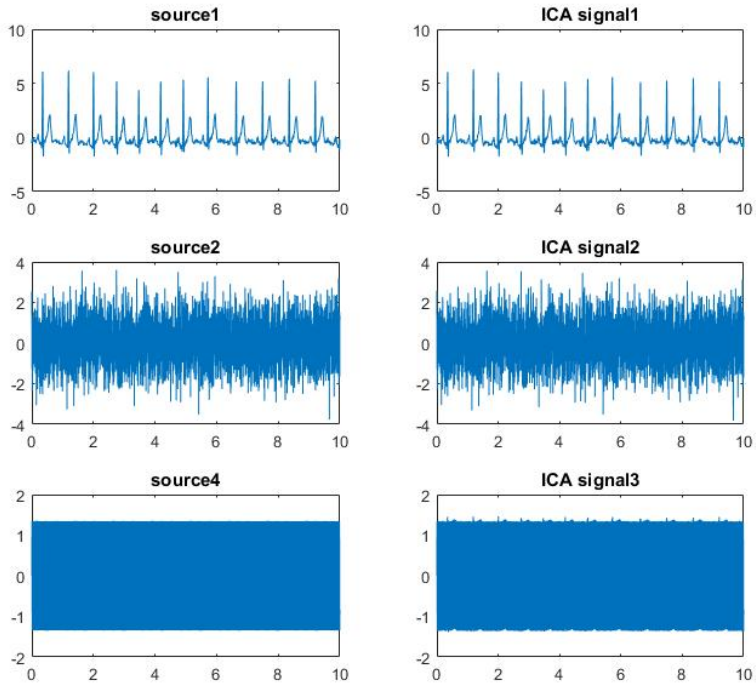

Figure 1.4: The original sources (left) and the IC from he ICA of x1 (right)

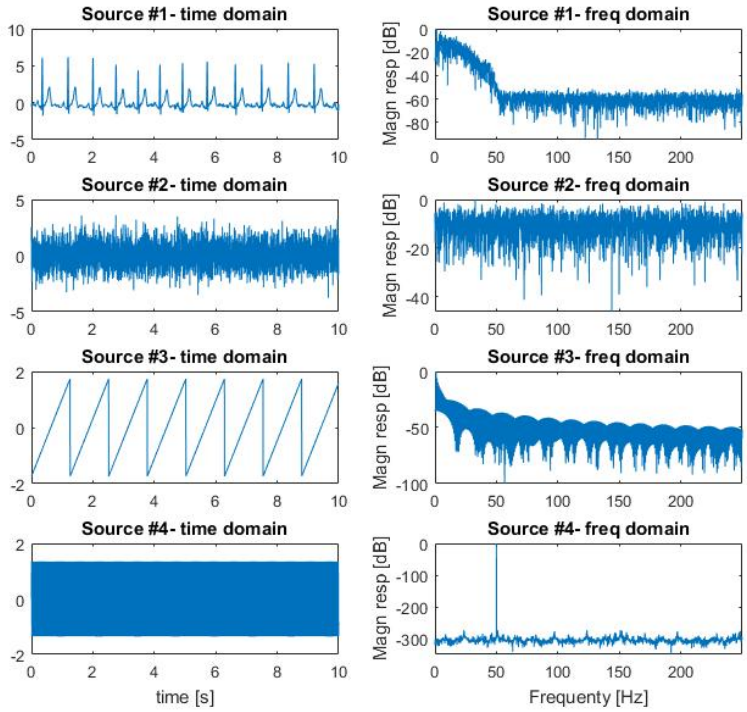

Figure 1.5: The original sources in time domain (left) and original sources in freq domain (right)

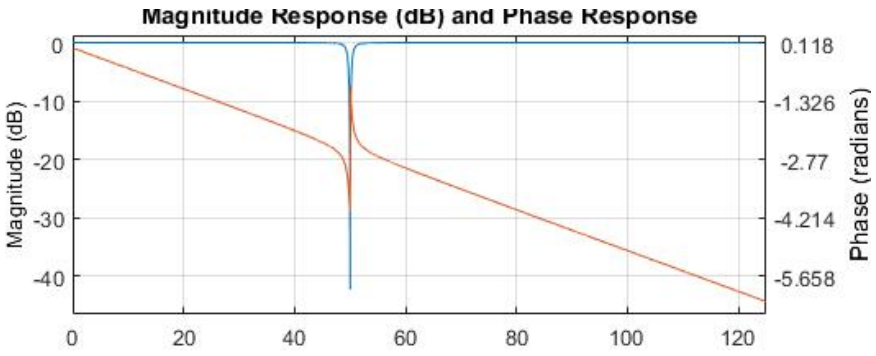

Figure 1.6: Frequency response of the notch filter (magnitude and phase

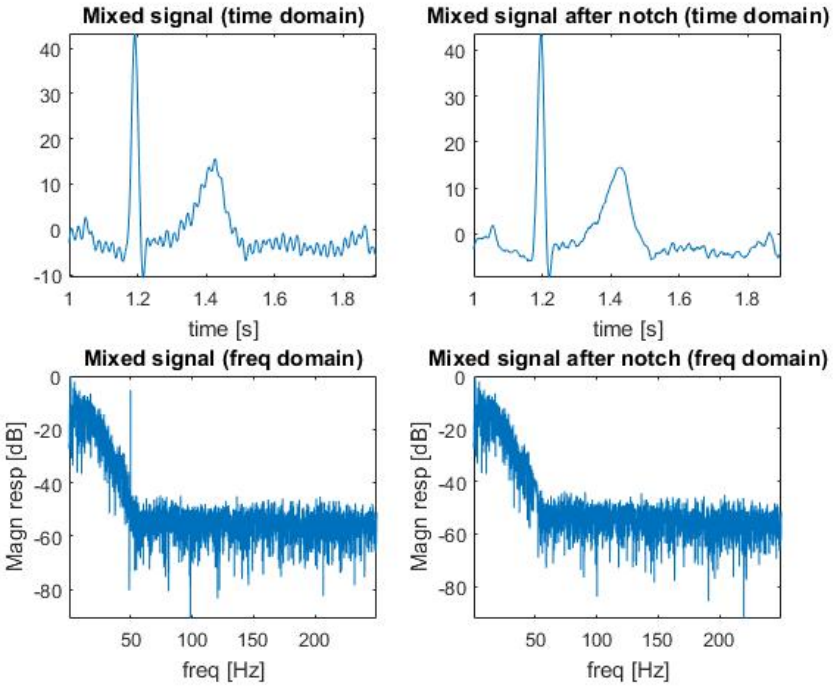

Figure 1.7: The effect of the notch filter (top: time domain, bottom: frequency domain, left: one of the mixture signals, right: the same mixture signal after notch

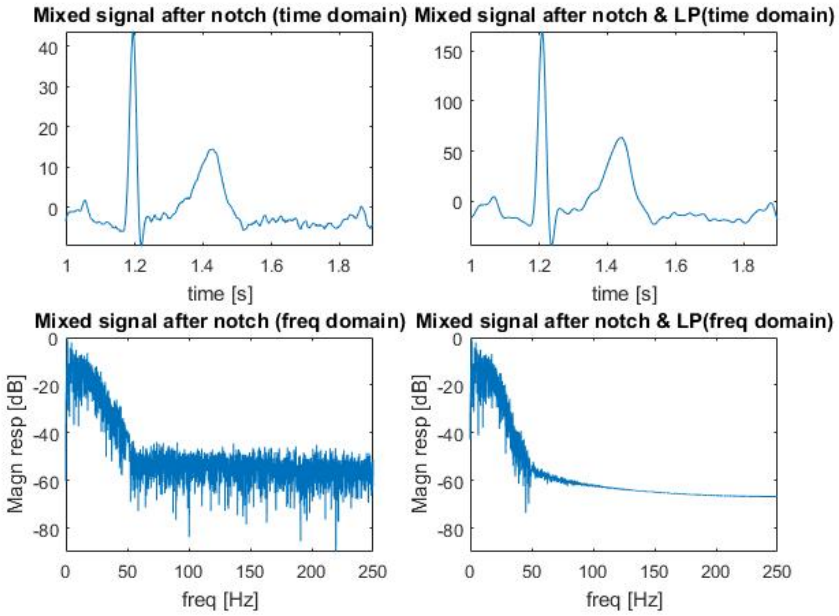

Figure 1.8: The effect of the LP filter (top: time domain, bottom: frequency domain, left: one of the mixture signals after notch, right: the same mixture signal after notch and LP

1.2 ICA on signal X2 and X3

In the next part the same analysis is done, but on different mixture signals. This would correspond with different measurements of the same sources. First the mixing matrix A is changed, so now $x2 = A2 * s$. A2 is a 6 by 4 matrix and therefore there will be six mixture signals (fig1.9). This would correspond with 6 different measurement signals of 4 independent sources. The result of an ICA of these mixed signals is shown in figure1.10, again normalized and combined with the corresponding source. The RMSE of these source estimations is listed in table1.2. Note that the output of the fastICA algorithm consists out of four sources, while the input were six mixture signals. This is can be explained by looking at the mixture signals. They are a noise free linear combination of 4 independent sources. Thus, the covariance matrix of these mixture signals will only have four non-zero eigenvalues. By looking at the amount of non-zero eigenvalues, the fastICA algorithm knows that it can reduce the dimension of the problem from six to four.

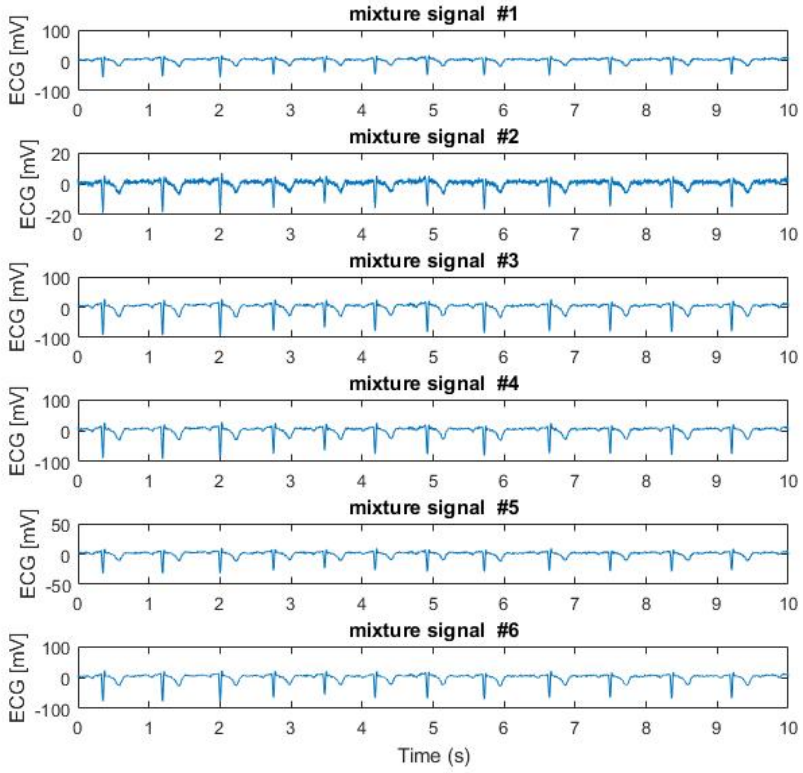

Figure 1.9: The mixed signals X2

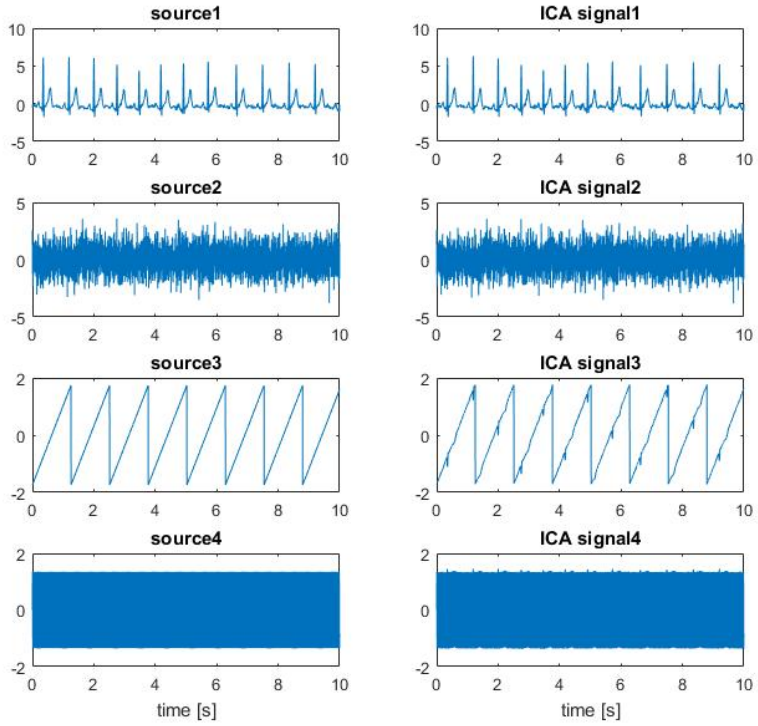

Figure 1.10: The original sources (left) and the IC from the ICA of x2 (right)

To obtain the mixed signals x3, the same mixing matrix A2 is used but on top of that, additional white noise (uniformly distributed) is added in the form of a random vector: x3 = A2*s + R. So again, six mixture signals (figure 1.11) are obtained on which ICA can be performed. As source estimations of the ICA we find six independent signals, four of them corresponding to the four exact sources and two random components (that are the most corresponding to the white noise source) as can be seen in figure1.12. From the RMSE of these source estimations (table1.3) however, it is clear that the apart from the fact that they are all random, the IC3 and IC4 have no real correspondence with the white noise source. Note that, in contrary with the previous example, the fastICA algorithm now estimates six independent components instead of four. This can again be explained by the covariance matrix of the mixture signals. Due to the extra added white uniformly distributed noise, the covariance matrix has six non-zero eigenvalues (although the two extra eigenvalues are close to zero). Thus, the fastICA algorithm cannot reduce the dimension of the problem and searches for six independent components. In essence, there are six new independent sources introduced by the random vector R (one on each mixture channel). However, there are only six channels available, so the ICA algorithm can at most separate 6 independent components in total (so two extra noise sources instead of 6). Important to note is that the added noise is uniformly distributed. Therefore, the ICA algorithm is able to separate those extra noise sources from the white Gaussian noise source.

Table 1.2: RMSE estimated sources of X2

Signals	RMSE
Source 1 - IC 1	0.062809
Source 2 - IC 2	0.040596
Source 3 - IC 3	0.056781
Source 4 - IC 4	0.020139

If the added noise would also be Gaussian, the ICA algorithm would not be able to separate those extra noise sources from the existing noise source, because ICA can separate only one Gaussian distributed source. This is due to the fact that an orthogonal transformation of Gaussian independent variables still results in an independent set of Gaussian variables. The ICA algorithm thus cannot find a direction of separation which leads to less Gaussiannity of those sources, and will not be able to separate the sources.

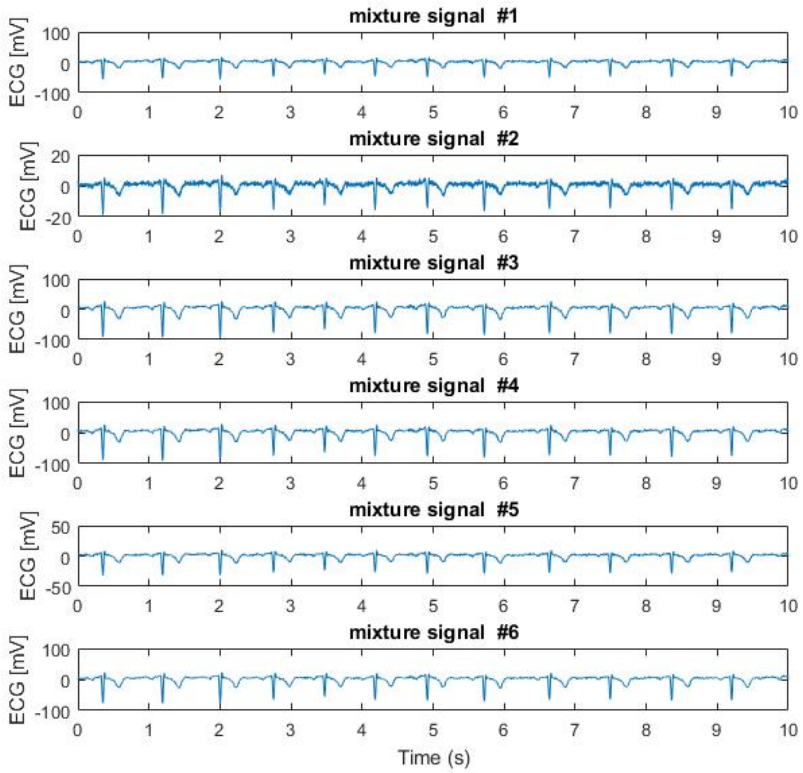

Figure 1.11: The mixed signals X3

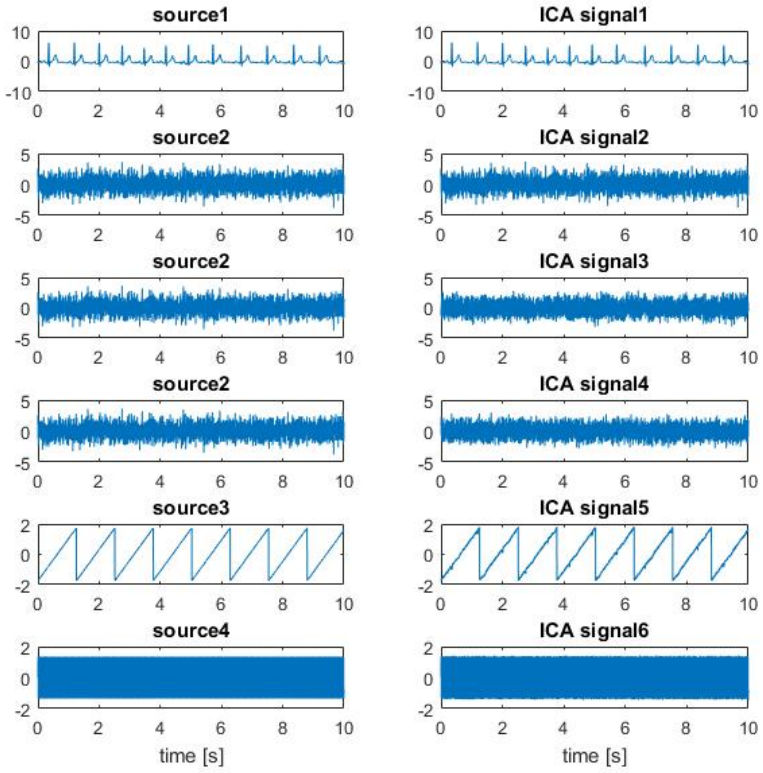

Figure 1.12: The original sources (left) and the IC from the ICA of x3 (right)

Table 1.3: RMSE estimated sources of X3

Signals	RMSE
Source 1 - IC 1	0.056451
Source 2 - IC 2	0.073476
Source 2 - IC 3	1.3818
Source 2 - IC 4	1.3874
Source 3 - IC 5	0.05555
Source 4 - IC 6	0.033

Chapter 2

Exercise 5.1.2

This exercise uses EEG data, gathered from an auditory attention experiment. The objective is to remove the blink artifacts from the data.

2.1 Visualisation of the EEG data

The EEG multichannel data is visualised in figure 2.1. The eye blink artifacts are indicated in red. The blinks are best noticeable in the Fpz, F3 and F4 channels, which are located at the front of the head. Because this exercise focuses on Independent Component Analysis and not on event detection, the eye blinks were marked by visual inspection. If the reader would want to automatize the event detection process, methods such as template matching or matched filter could be used (as was explored in exercise session 2).

2.2 Independent Component Analysis

The EEG multichannel signal contains noise artifacts from the eye blinking of the patient. As already showed in the previous exercise, independent component analysis can be used to remove these artifacts from the eeg signal. The independent components (source estimations) resulting from the ICA are given in figure 2.2 and the eye blink source is clearly recognizable (component one). The blinks are marked in red on the figure. For diagnostic and other reasons, it can be interesting to investigate the spatial occurrence of the sources. This can be done by correlating the sources with the EEG channels and visualizing the results. In figure 2.3, this spatial distribution is plotted for the 12 independent components. It shows that the first component, corresponding with the eye blink source, originates from the front side of the head (region of the Fp3, Fp4, Fz electrodes). The symmetry (left/right) implies that the patient blinked with both eyes simultaneously.

To remove the blink artifacts, the column of W^+ (the pseudo inverse of the demixing matrix) corresponding to the eye blink component is set to zero, exactly as in the previous exercise. The new, blink free, eeg signal is then obtained by $x = y * \tilde{W}^+$ (with $\tilde{W}^+$ the pseudo inverse of the demixing matrix, with the replaced column of zeros) . This resulting, blink free, signal is given in figure2.4 and the

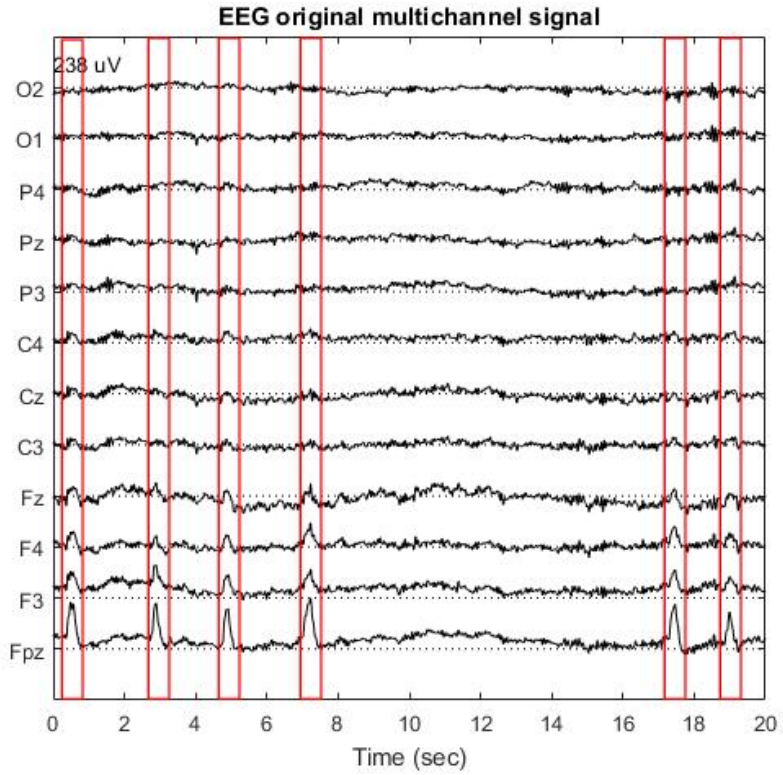

Figure 2.1: The original eeg signal with the presence of eye blink artifacts (in red)

eye blink artifact is successfully removed. As a check, ICA is now applied again on the filtered signal and as can be seen in figure 2.5 the eye blink component is not present anymore.

In this exercise, the independent component corresponding to the eye blinking was visually detected and hard coded in the software. In real life applications it is desirable to automatize this process of blink removal. The blink independent component identification can be done by applying matched filters or template matching to the normalized independent components, and measuring which resulting signal displays clear (periodic) peaks (which correspond to the blinks). A draft of this possible method was implemented in the code, but was not finished because of a lack of time. Another possible method to automatize the selection of the artifact independent components is to investigate the spatial distribution of the independent components. As can be seen in figure 2.3, the blinking artifact is only present at the front of the head, while the other components are distributed over larger regions. Component and 4 and 12 are also primary located at the front of the head, but their distribution is less symmetric than the blinking component. So the symmetry can also be taken into account.

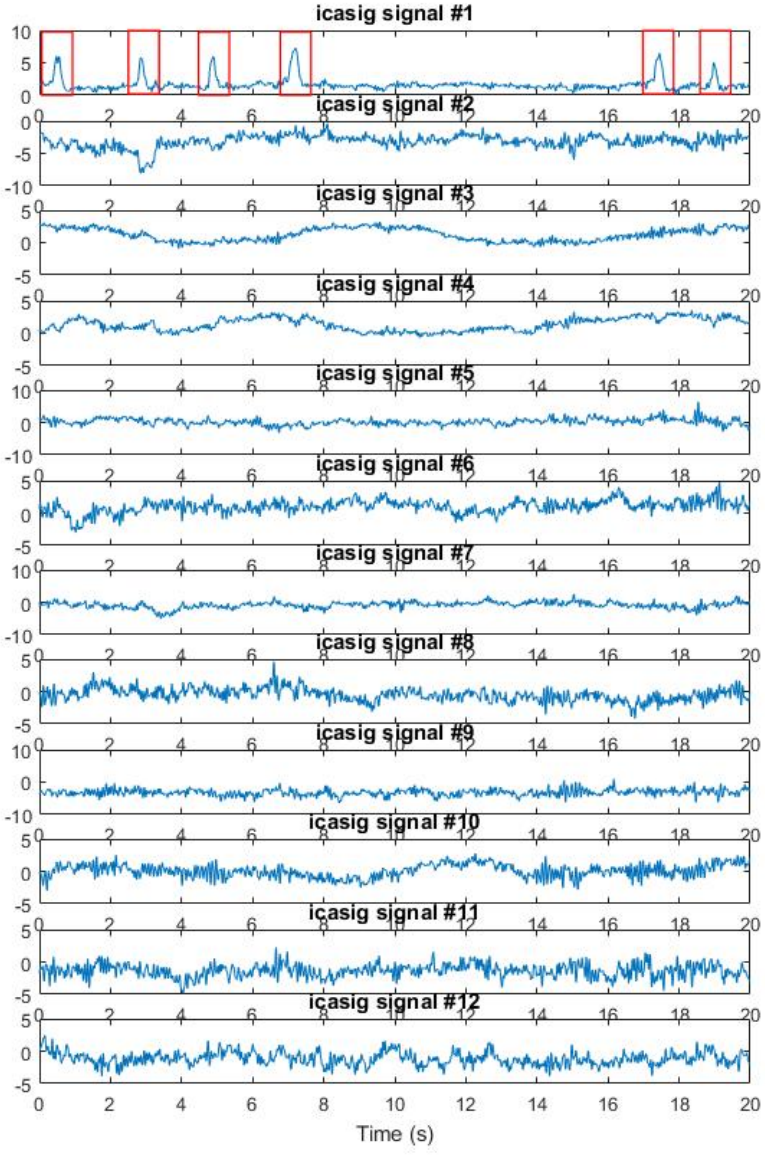

Figure 2.2: The IC with the eye blink source as IC 1 (first signal)

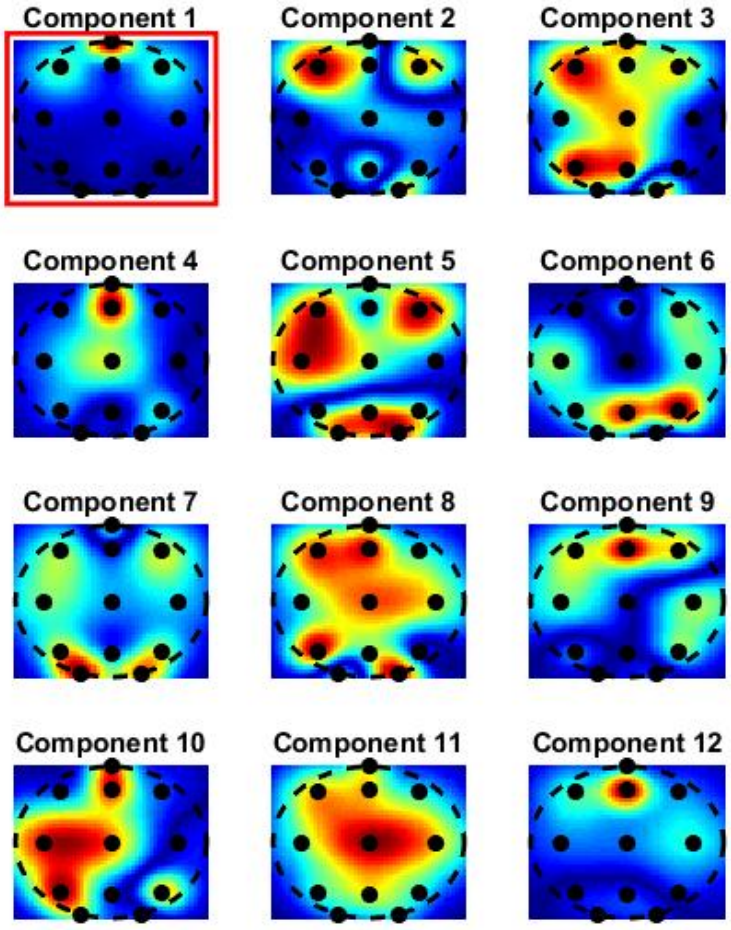

Figure 2.3: The spatial distribution of the IC with the eye blink source as IC 1 (first signal)

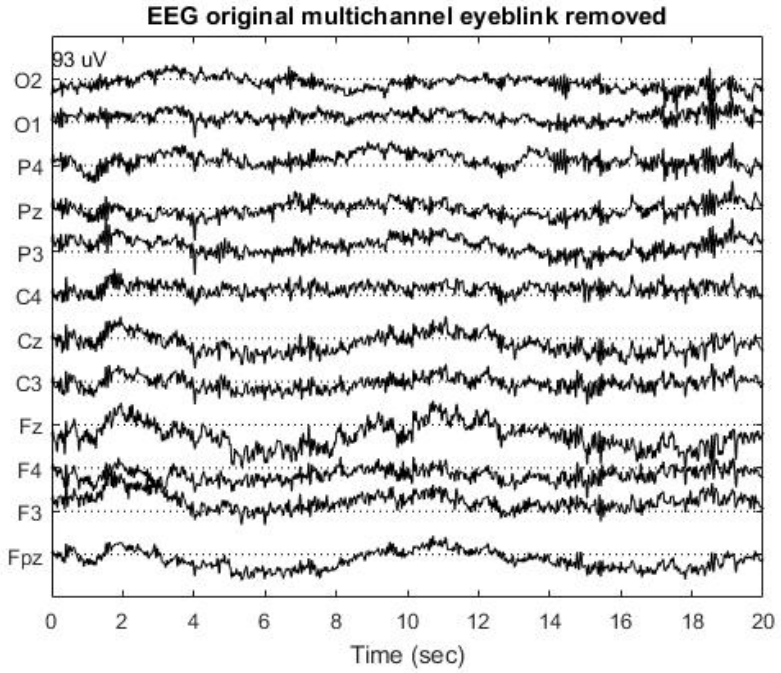

Figure 2.4: The filtered EEG signal, without the eye blink artifacts)

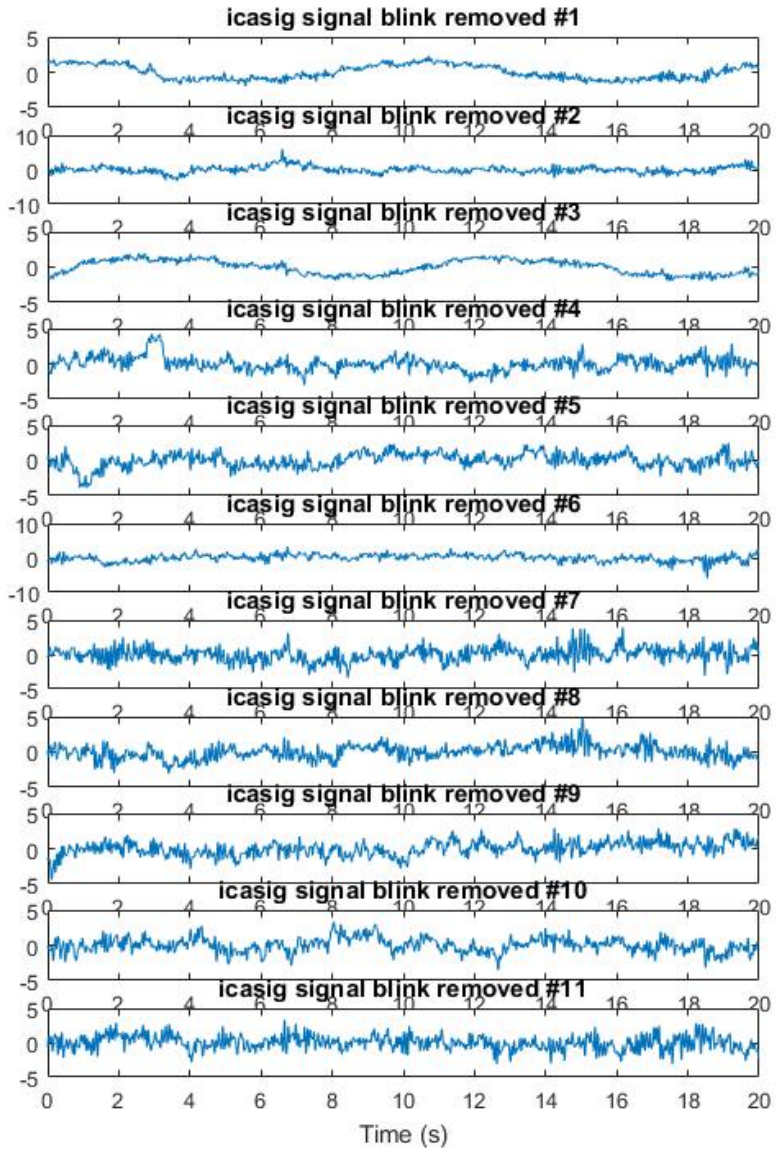

Figure 2.5: The resulting independent components of the ICA on the EEG data with eye blinks removed

Chapter 3

Exercise 5.2.1

In this exercise an ECG signal of 15 seconds sampled at 300Hz is considered. The aim of this exercise is to compress the ECG data. This is done by performing a wavelet decomposition and selecting the most representable coefficients.

3.1 Wavelet decomposition

The ECG signal is decomposed with Daubechies wavelet 4 (db4), using 5 levels of decomposition. Each of the the different approximation (A1) and detail (D1-D5) coefficients covers a certain frequency band as given in table3.1. As can be seen, the higher detail coefficients have a better frequency resolution, but due to the down sampling that is done during the wavelet decomposition, the time resolution will be poorer too. This can be observed in figure 3.2 where the original ECG signal is plotted on top and the approximation and detail coefficients in the following order (A5, D5, D4, D3, D2, D1). The decreasing time resolution (number of dots) is clearly visible.

Figure 3.1: Daubechies wavelet 4: Scaling and wavelet function

3.2 Compression of the signal using wavelets

In order to reduce the number of coefficients the M largest wavelet coefficients (approx and detail) are selected and based on these, the original signal is reconstructed (to a certain extend) In figure 3.3, the magnitudes (absolute value) of

Table 3.1: The frequency bands of the different coefficients

Coefficient	frequency band	time resolution
A5	[0-5 Hz]	$1.07e-1s$
D5	[5-10 Hz]	$1.07e-1s$
D4	[10-19 Hz]	$5.33e-2s$
D3	[19-38 Hz]	$2.67e-2s$
D2	[38-75 Hz]	$1.33e-2s$
D1	[75-150 Hz]	$6.66e-3s$

all the wavelet coefficients are plotted. From this figure it is clear that the coefficients A5, D5, D4 and D3 will be the most important for reconstructing the ECG signal. This makes sense since the frequency bands of these coefficients (see table3.1) overlap with the characteristic part of the frequency spectrum of an ECG signal (range 0.5-50 HZ). The reconstruction of the ECG signal with a selection of M coefficients is done for M = 10, 25, 50, 100, 500 and 1000. The results for M = 10, 25, 50 and 100 can be found respectively in figure 3.4, 3.5, 3.6, and 3.7, where the original and reconstructed signal are plotted on top, with the used coefficients (M largest) underneath. It is clear that the reconstructed signal improves as the number of used coefficients M increases. This can be also concluded from the RMSE of the reconstruction (with the original as reference) listed in table 3.2 and from the close up of the signal for different M in figure3.8. The optimal amount M of used coefficients for the reconstruction depends on the application. If for example only the RR intervals are needed, a reconstruction with $M = 10$ or $M = 25$ could be used, because the QRS complex is clearly visible there (figure 3.8. If more detailed information is needed, M should be chosen bigger. With $M \geq 50$, the P wave becomes visible and with $M \geq 100$, the T wave becomes visible (but the offset is biased). With $M = 500$, almost all the information of the shape of the ECG signal is incorporated, while having the data compressed to 11% of the original data size.

When we now look at the complete reconstructed signal for one of the M, for example M=50, the shape of the different QRS complexes differ throughout the signal. Depending on which coefficients the reconstruction is made, a more ϕ or ψ (the scaling and wavelet function from figure 3.1) shape can be recognized. This is clarified by a close up or the reconstruction for M=50 from sample 1750 to 2250. The first wave clearly tends more to the ψ function where the seconds one looks more like the ϕ function. The explanation for this can be found on figure 3.6. At the position of the first QRS complex (around sample 1850) there is no A5 coefficient present. This means that the reconstruction at this position is made based only on the details coefficients (D5 and D4) that as already mentioned before are computed by a high pass filter that is related to the ψ wavelet function. At the position of the second QRS complex (around sample 2050), there is a A coefficient present and therefore here it is possible to obtain a ϕ function shape. So depending of the frequency content of the section of the signal, the shape can differ. Low frequencies (within the frequency band of A5) will induce larger A5 coefficients and thus the reconstruction of that section will have a large part corresponding to ϕ wave. Higher frequencies trigger ψ like waves in the reconstruction.

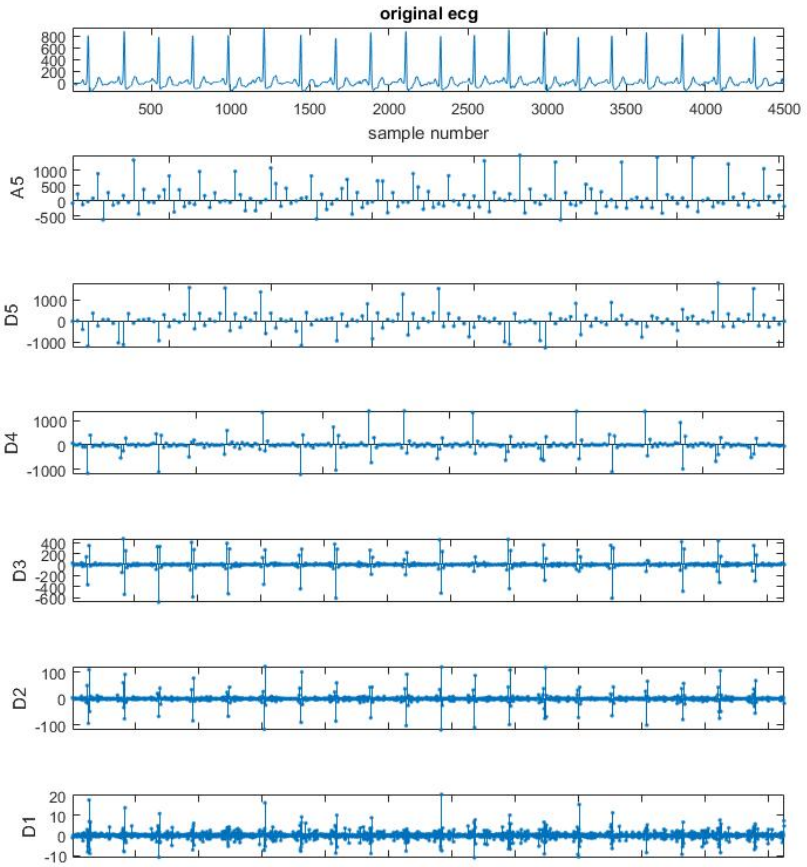

Figure 3.2: wavelet decomposition of the ECG signal

Note on selection of wavelets For this exercise, db4 wavelets were used for the wavelet decomposition. In general, there is no clear 'best' wavelet. The 'best' wavelet depends on the application (used dataset). Thus to chose the optimal wavelet, it has to be matched to the used dataset. In this case, ECG data is used. The most characteristic part of this data is the QRS complex. The db4 wavelets have already a shape that closely resembles the shape of the QRS complex, which indicates that it is a good choice of wavelet for this application. Other desirable properties of the db4 wavelets are their orthogonality and their simplicity (only 4 parameters). Because of their computational simplicity , they are also useful for online applications.

Table 3.2: RMSE of the reconstructed signals for different M

Number of coeff	RMSE
$M = 10$	135.08
$M = 25$	112.56
$M = 50$	85.853
$M = 100$	60.803
$M = 500$	7.5612
$M = 1000$	1.9723

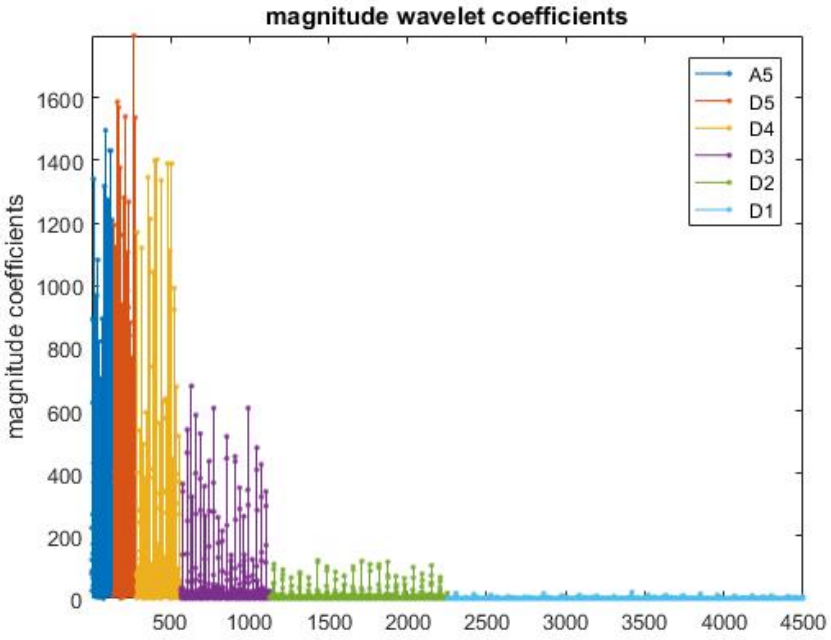

Figure 3.3: The magnitude (absolute value) of the wavelet coefficients

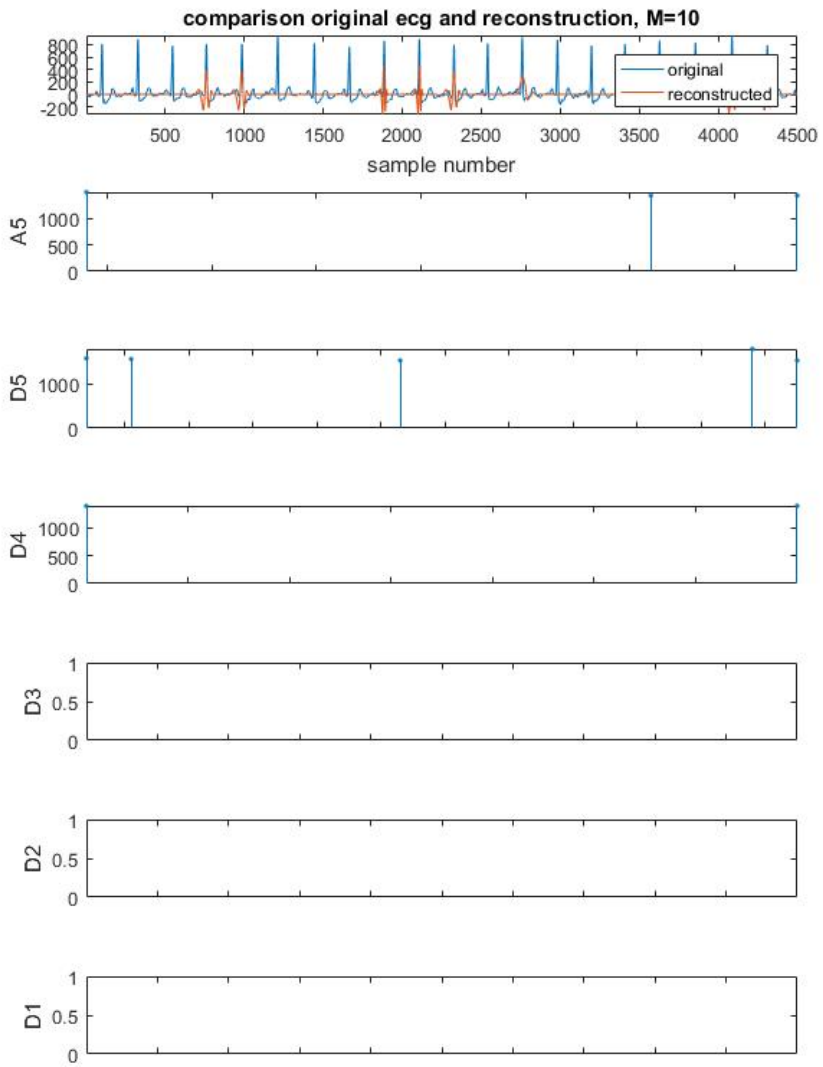

Figure 3.4: Reconstruction of the ECG using M=10 coefficients

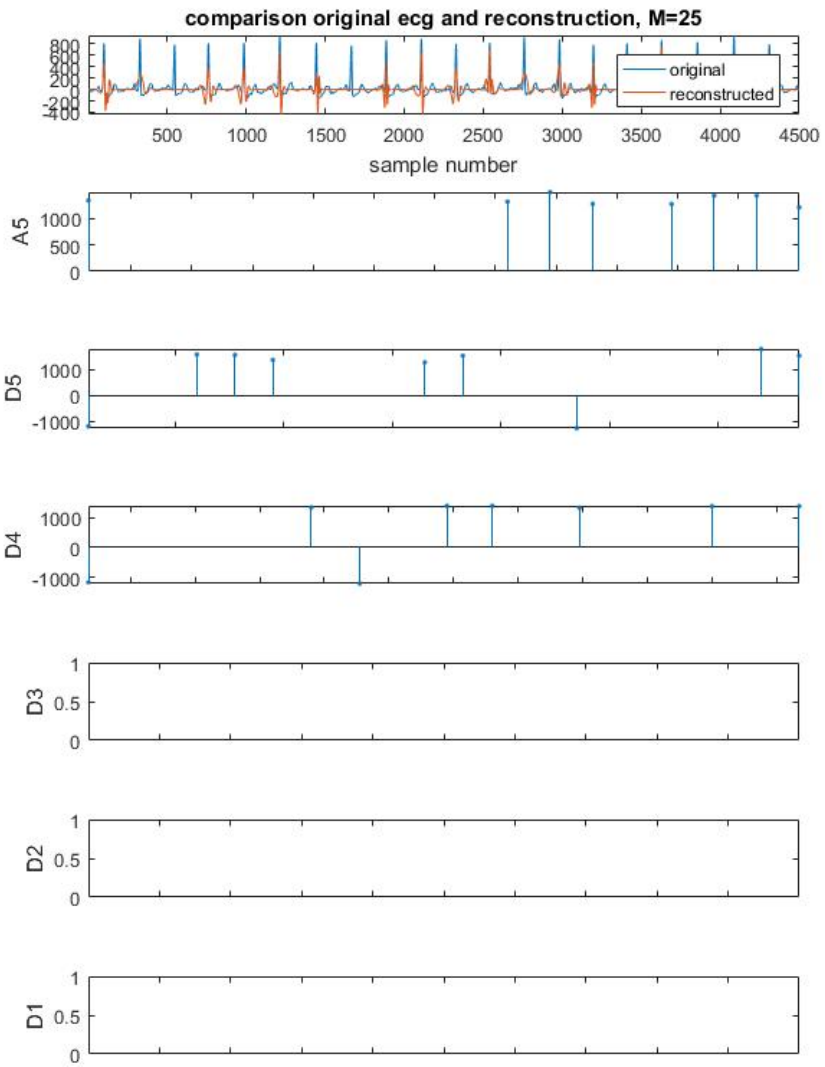

Figure 3.5: Reconstruction of the ECG using M=25 coefficients

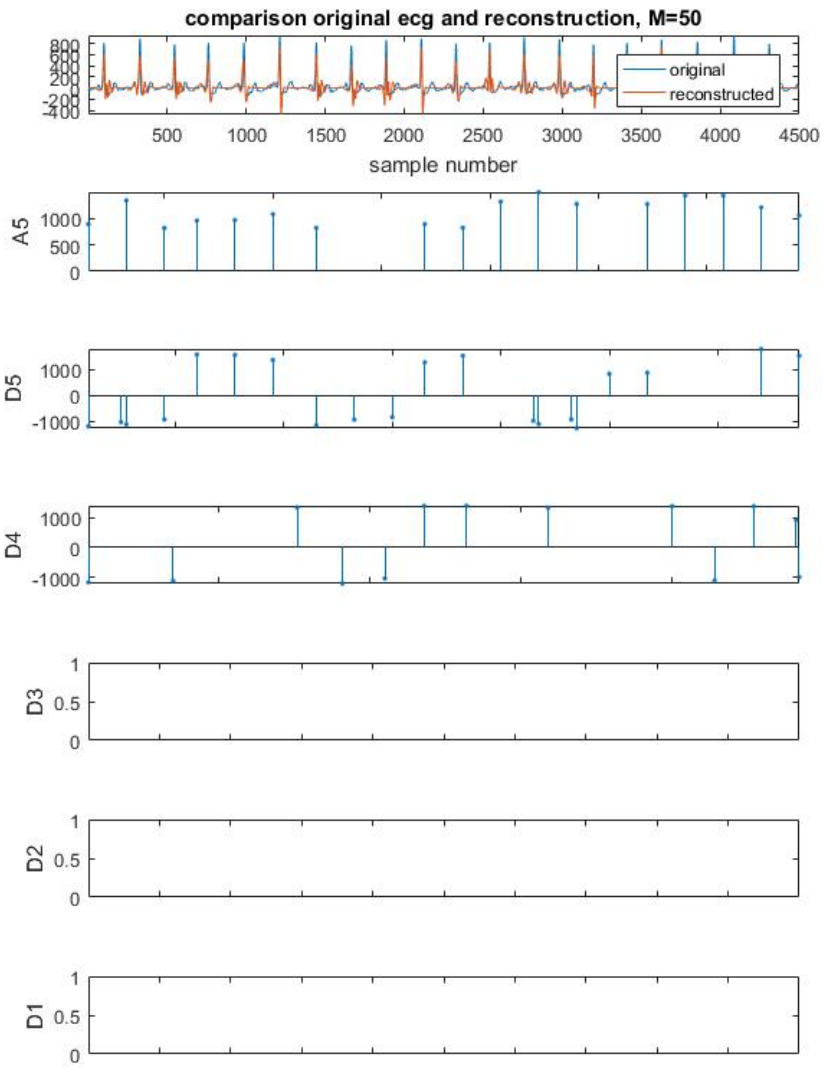

Figure 3.6: Reconstruction of the ECG using M=50 coefficients

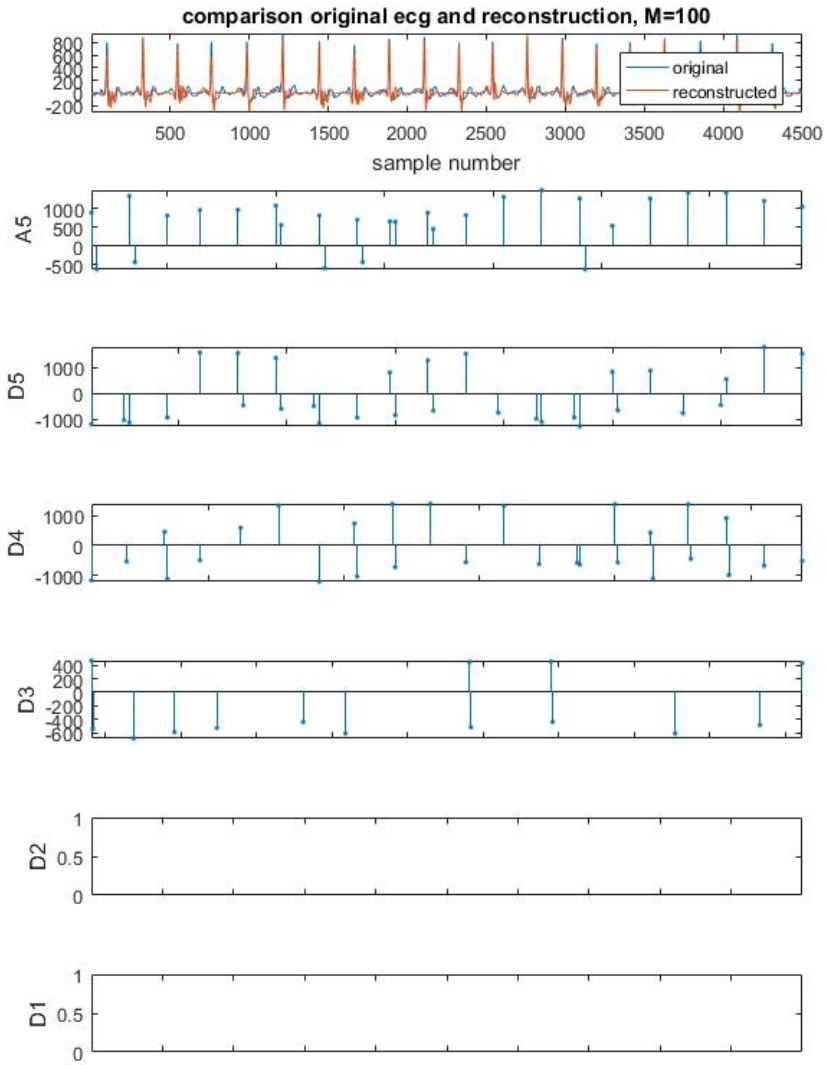

Figure 3.7: Reconstruction of the ECG using M=100 coefficients

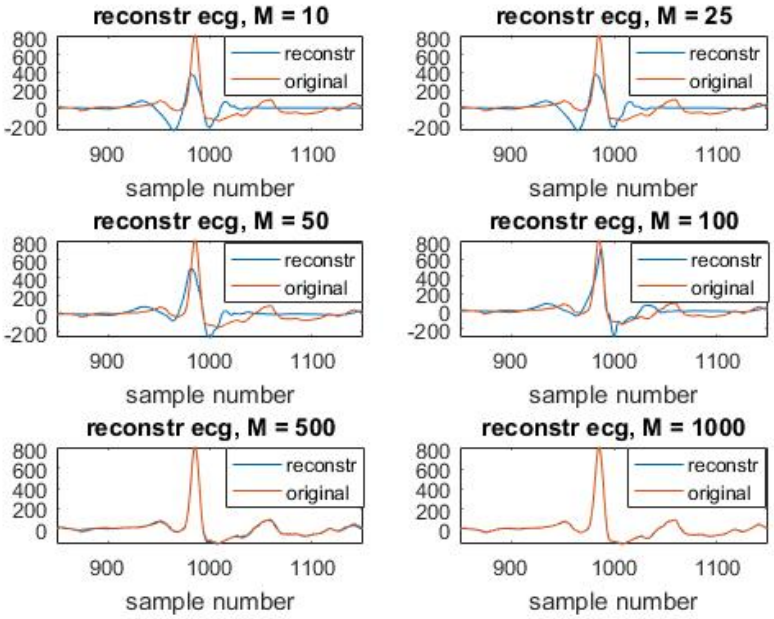

Figure 3.8: Overview of the reconstructions for different M compared to the original

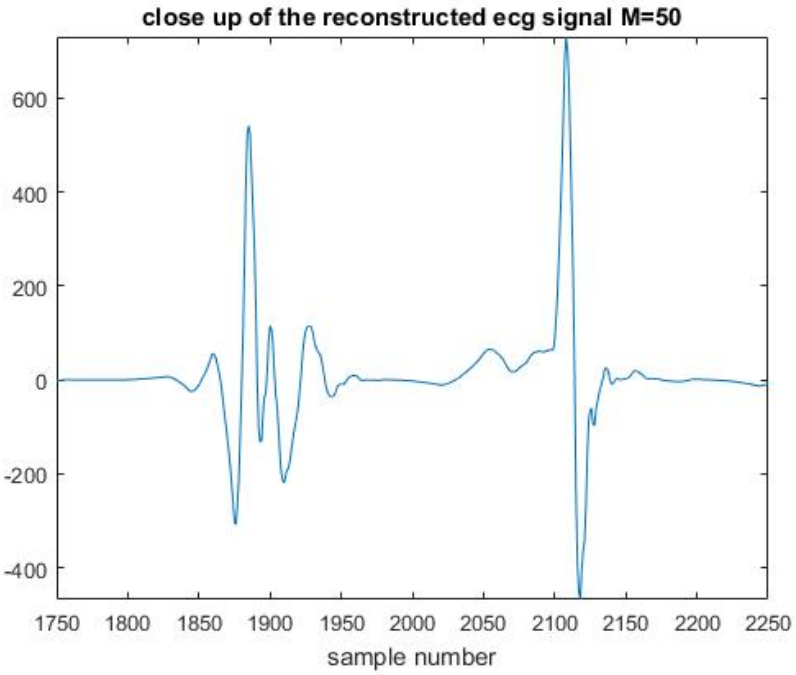

Figure 3.9: Close up of the reconstructed ECG using M=150 coefficients

Chapter 4

Exercise 5.2.2

For the classification of EEG segments the discrete wavelet transform can be particularly useful, because it separates the signals in different frequency bands. This will be illustrated in this exercise where a test set of EEG segments has to be classified as being seizure or non seizure based on the information of a training set.

4.1 Visualisation of the data

The raw data consists of segments of 2s of the EEG data, sampled at 256Hz (140 train segments and 106 test segments). The segments of the training set are plotted, all after each other, in figure4.1. On this figure, it can be seen that the seizure samples have a higher average absolute amplitude and will thus have a higher average total energy content. This total energy can be extracted as a possible feature. For classification, it is sometimes useful to standardize the data (classifiers are trained more easily with standardized data). However, this sometimes leads to a loss of information. In this case, the standardization of each segment would mean that each segment has the same total energy content. This will thus lead to a loss of information. Not all information is lost, because the relative differences in energy content between different frequency bands still remain, but as mentioned above, the seizure signals have on average a higher total energy, thus that information will be lost. This was also seen in the classification results (table 4.2 and table 4.3): with standardized data, the accuracies of the classification were lower. The accuracy of the classification based on total energy was 50% (random) as expected, because the total energy of each segment is equal due to the standardization. One could also standardize the data in the other dimension (standardize over all k^{th} samples of the segments), but this dimension has no meaning in this case. For the results of this chapter, the original raw data was used (no standardization).

4.2 Feature extraction

One possible feature to use in the classification could be the total signal energy. Figure4.2 gives the energy present over the whole frequency spectrum of each of the segments of the training set. The energy level of the seizure sample is on

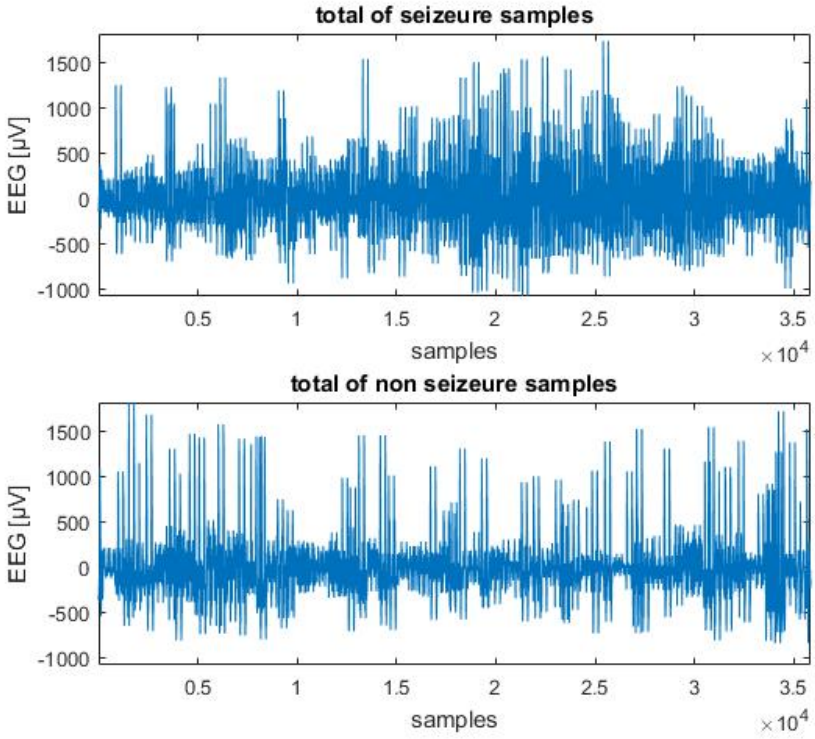

Figure 4.1: All the seizure segments of the trainings set plotted after each other (top), all the non seizure samples of the trainings set plotted after each other (bottom)

average a bit higher, but it is clear that there is no good separation between the two classes possible based on this one feature.

Another method to extract features is to compute the energy of the segments in different frequency bands. The wavelet decomposition with db4 wavelets in 5 levels results in the frequency bands displayed in table 4.1. Each frequency band corresponds with a specific EEG rhythm. The energies contained in the approximation and detail coefficients for all the segment of the training set can be now calculated and the result is given in figure4.3 under the form of boxplots. The 0 represents all the non seizure segments, the 1 all the seizure segments. It can be seen that the classification based on the total energy (whole frequency range), as already concluded before, is poor. The same is true for the energies of the approximation coefficient and the detail coefficient D3. The four other on the other hand, show better (not complete) separation and would therefore be more suited to use as an individual feature for classification. For the lower frequency bands (given by D5 and D4), the energies of the seizure segments are higher. This is according to the expectations since seizure eeg is characterized by recruiting theta waves (around 5 Hz) [1]. For the higher frequency bands (given by D1 and D2) exactly the opposite can be observed. It is possible to standardize the features after the feature extraction step (because the features have high absolute values), but

because there is also information in the difference in energy between frequency bands, this step is omitted.

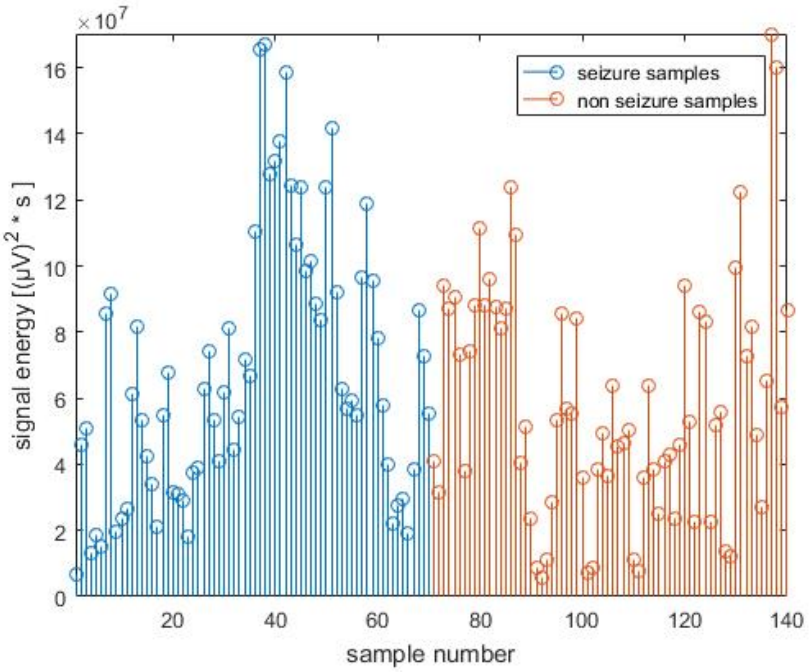

Figure 4.2: The signal energy (over the whole frequency spectrum) of all the segments of the training set

Table 4.1: The frequency bands of the different coefficients and the corresponding (approx) EEG rhythms

Coefficient	frequency band	eeg rhythm
A5	[0-4 Hz]	δ waves
D5	[4-8 Hz]	θ waves
D4	[8-16 Hz]	α waves
D3	[16-32 Hz]	β waves
D2	[32-64 Hz]	γ waves
D1	[64-128 Hz]	

4.3 Classification

The next step is to classify the test data based on the training data. This classification is done with different feature sets. First, the total signal energy is used, afterwards the energy of each frequency band is used separately and finally, all frequency band energies are used together in the classification. Linear Discriminant Analysis (LDA) is used for the classification. Table 4.2 summarizes the

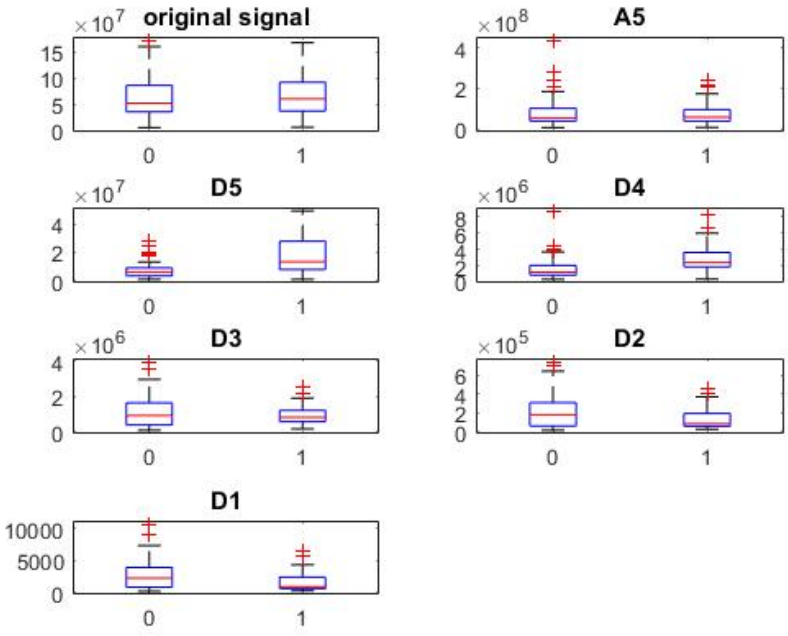

Figure 4.3: Boxplot comparing the energies of the seizure(1)/non seizure(0) segments in the different frequency bands (related to the coefficients)

results.

Figure 4.4 visualizes the classification of the test set based on the total signal energy. Because there is only one feature, a simple threshold is used to separate the data. As expected, the samples with a lower energy value are classified as non seizure and visa versa. This results in a low sensitivity, because a lot of seizure segments have a total energy content below the threshold.

Figure 4.5 visualizes the classification of the test set based on the individual sub-band energies. Here it can be observed that in the lower frequency bands (θ and α waves), the energies in the seizure segments is (on average) higher, where for the higher frequency bands (D1 and D2) it is the other way around. Note that the sensitivity of those high frequency bands is very high, but the specificity very low. Thus almost all seizure segments will have a low high-frequency energy content, but also a lot of non-seizure segments will have a low high-frequency energy content. This can clearly be seen on figure 4.5. Also clear from this figure is that the energies from the frequency bands of D5 and D4 would be the best as individual features. The exact results of these individual features can be found in table4.2.
The combination of all the sub-band energies leads to the best classification results, because the different discriminating aspects of the features can be combined. This leads to the highest accuracy of 92% (Table 4.2). A last option could be to combine the sub-band energy features with the total energy feature. How-

ever, the total energy of the segment is equal to a weighted sum (linear combination) of the sub-band energies, so it does not introduce extra information in the linear classifier. This can be seen in Table 4.2: the accuracy of both classifications stays the same, which indicates that no extra classification power is introduced by the extra feature. There is a small trade-off between sensitivity and specificity, because the LDA algorithm has an extra feature to deal with, and thus will have a slightly different solution.

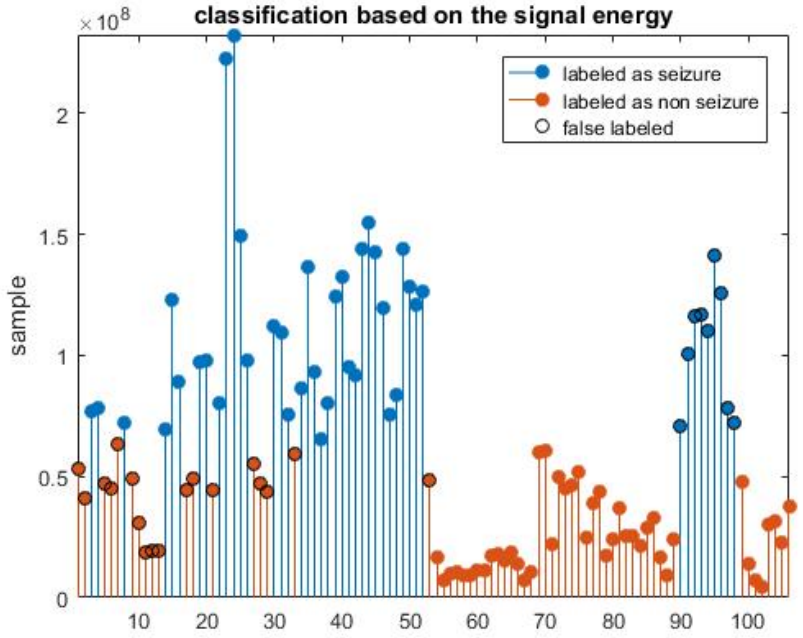

Figure 4.4: Classification of the training set based on the signal energy (total freq spectrum)

Features	accuracy	sensitivity	specificity
'total signal'	0.74528	0.66038	0.83019
'A5'	0.29245	0.39623	0.18868
'D5'	0.86792	0.84906	0.88679
'D4'	0.81132	0.77358	0.84906
'D3'	0.66038	0.75472	0.56604
'D2'	0.74528	0.90566	0.58491
'D1'	0.73585	0.90566	0.56604
'D1, D2, D3, D4, D5, A5'	0.92453	0.90566	0.9434
'total signal + coeff'	0.92453	0.96226	0.88679

Table 4.2: The accuracy, sensitivity and specificity of LDA classification using different feature vectors

Features	accuracy	sensitivity	specificity
'total signal'	0.5	0	1
'A5'	0.67925	0.71698	0.64151
'D5'	0.76415	0.71698	0.81132
'D4'	0.59434	0.56604	0.62264
'D3'	0.78302	0.90566	0.66038
'D2'	0.73585	0.86792	0.60377
'D1'	0.51887	0.77358	0.26415
'D1, D2, D3, D4, D5, A5'	0.81132	0.92453	0.69811

Table 4.3: The accuracy, sensitivity and specificity of LDA classification using different feature vectors after first have standardized the samples

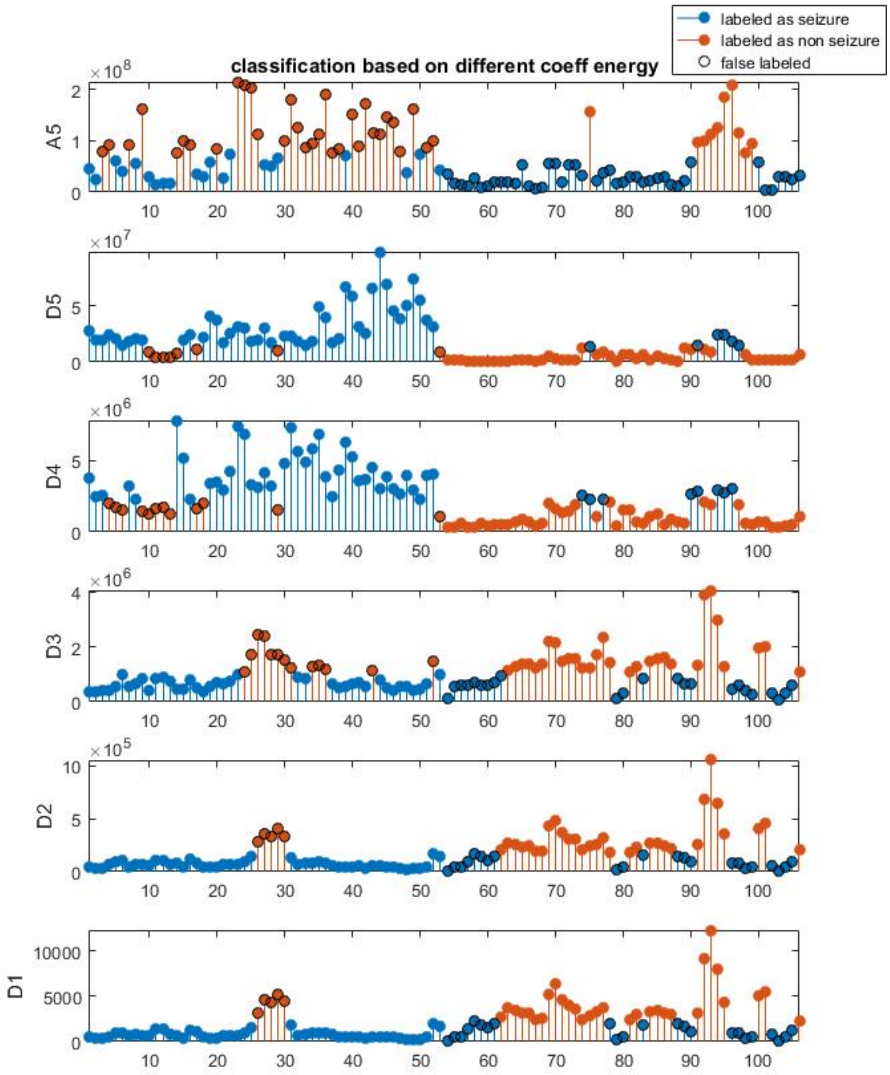

Figure 4.5: Classification of the training set based on the different frequency band signal energy

Bibliography

[1] Rangaraj M. Rangayyan, *Biomedical Signal Analysis, Second Edition*, 2015

[2] S. Van Huffel, *SESSION 5: ICA/Wavelets*, 2017