Time Frequency Analysis and Characterization of Phonocardiography Signals

Marc Palomer, Cinthia Krämer, Universitat de Barcelona (January 2023)

***Abstract*— *Objective:* This paper presents a comprehensive evaluation of various estimation techniques used in the analysis of phonocardiographic (PCG) signals, which are recordings of heart sounds obtained during auscultation. Techniques evaluated include Power Spectral Density (PSD) estimations as Welch or Burg periodograms, time frequency representations (TFR) as spectrogram, scalogram (continuous wavelet transform) and quadratic time-frequency representations. Empirical mode decomposition method also has been evaluated. The main objective of this work is to evaluate the effectiveness of these techniques and their optimal parameters to distinguishing and analyse normal heartbeats and 3 different cardiac pathologies. A secondary goal is to develop an algorithm that automatically classifies the type of heartbeat to extract time and frequency characteristics of each PCG component using the optimal parameters.**

***Significance:* Accurately diagnosing heart disease is critical to the proper treatment and management of patients. The work considers a set of estimation techniques that allow for a comprehensive assessment of the strengths and limitations of each technique when analysing PCG signals.**

***Index Terms*— Phonocardiography, Frequency estimation, Time-frequency analysis, Continuous wavelet transform, Quadratic time-frequency representations, Empirical decomposition, Algorithm development, Automatic Signal processing.**

INTRODUCTION

Phonocardiography (PCG) is a non-invasive technique used to record the sounds of the heart obtained during auscultation. The PCG signal is a valuable tool in the diagnosis of various cardiac diseases, including valvular disorders and congenital heart diseases. For a healthy heart the PCG signal contains the first and second heart sounds (S1 and S2). S1 is the murmur produced by the atrioventricular (AV) valves closing during ventricular systole, while S2 is the murmur produced by the semilunar (SL) valves closing during ventricular diastole. These sounds can be heard through a stethoscope and are used to assess heart valve function and to detect certain heart conditions.

However, the interpretation of PCG signals can be challenging due to the presence of multiple components with different temporal and frequency characteristics. The accurate and efficient analysis of PCG signals is critical for the proper treatment and management of patients.

In this work, we evaluate different frequency and time-frequency estimation techniques, including the spectrogram, continuous wavelet transforms (CWT), quadratic time-frequency representations (QTFR), and empirical decomposition into oscillation modes, applied to PCG signals of normal beats and various pathologies. The goal of this evaluation is to identify the most suitable techniques for the analysis of PCG signals and to identify the characteristics that differentiate between normal and pathological signals.

Additionally, we propose an algorithm that automatically obtains the PSD representation and characteristics of each PCG component, as well as the TFR. The algorithm uses parameters extracted from the PSD that best represent each signal component to characterize and differentiate between signals corresponding to a normal heartbeat and those associated with different pathologies, with this information, the algorithm delivers the best PSD and TFR representations tunning their parameters.

MATERIALS AND METHODS

Cardiac signals

The material used in practice consists of 4 different phonocardiography signals (one normal and one for each of the three pathologies). For each type of signal, a 0.7s window containing only 1 heartbeat was extracted for the analysis (Figure 1).

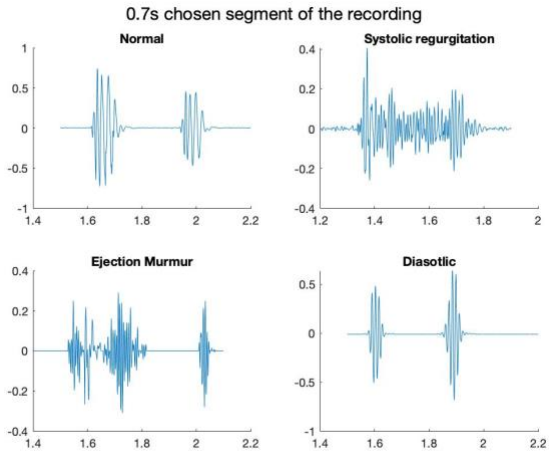


Figure 1: Segments (heartbeats) used for analysis

Definition of each type of heartbeat:

- Normal cardiac signal: Standard heartbeat with its typically separated S1 and S2 segments.

- Diastolic fixed S2 split (DFS2) (ORPHA:99106, Human Phenotype Ontology): Lack of variation in the splitting between the two components of the second heart sound with respiration. Normally, the aortic valve closure (A2) is followed by the pulmonic valve closure (P2) but the A2-P2 interval increases with inspiration and decreases with expiration

- Systolic Mitral Regurgitation (SMR) (HP:0001653, Human Phenotype Ontology): An abnormality of the mitral valve characterized by its insufficiency or incompetence, resulting in retrograde leaking of blood through the mitral valve upon ventricular contraction. Normally the mitral valve, which is located between the left atrium and left ventricle, opens to allow unidirectional blood flow from the left atrium to the left ventricle, and then closes to prevent blood from flowing back into the left atrium. In SMR, the mitral valve does not close properly and some blood flows back into the left atrium during systole. This can lead to an increased workload on the left ventricle of the heart, which over time can lead to heart failure.

- Heart ejection murmur (EM) of grade 1(HP:0030148, Human Phenotype Ontology): An extra or unusual sound heard during a heartbeat caused vibrations resulting from the flow of blood through the heart. It can be caused by turbulent blood flow through the heart valves. It is typically heard during systole, which is the period of contraction of the heart, and is caused by the rapid flow of blood from the left ventricle into the aorta or from the right ventricle into the pulmonary artery. Ejection murmurs are usually caused by aortic stenosis, which is a condition in which the aortic valve narrows, usually produced by calcification of the valve, and that produces reduced blood flow and increases turbulence. This turbulent blood flow and produces a characteristic "whooshing" sound that can be heard with a stethoscope.
  Analogously, mitral stenosis causes increased resistance to blood flow into the left ventricle, leading to an ejection murmur. Ejection murmurs are typically heard over the aortic or pulmonic valve areas, depending on where the stenosis is located. They are graded on a scale of 1 to 6, with 1 being the quietest and 6 being the loudest. The grade of the murmur is also related to the severity of stenosis. [1][5]

All the PCG signals used in the study are recorded with a sampling frequency of 11025 Hz. Since most of the frequency content of the PCG signals is below 500 Hz and to reduce the computational complexity of time-frequency based methods, it is advisable to perform a resampling of the signal at 1000 Hz, this has been done using the *leerwav.m* function of the virtual classroom of Abel Torres, to which the PCG signal is introduced and returns the variables: pcg (corresponding to the signal PCG resampled to 1000 Hz), t (the time vector), and fs (corresponding to the sample rate).

PSD estimations

In order to characterize the Power Spectral Density (PSD) of the signals two approaches have been used, the Welch periodogram and the Burg periodogram as the autoregression method.

The Welch periodogram of a signal is the average the multiple periodograms of smaller segments of the same signal, with overlap between the mentioned segments. This averaging can improve the estimate of the power spectral density, but it also increases the variance of the estimate. Therefore, the Welch method is better suited for stationary signal like ECG signals than for non-stationary signals. For the Welch periodogram the matlab function *pwelch()* from the signal processing toolbox is used, and different sizes of windows or overlapping proportions are evaluated.

The second PSD estimation proposed is the auto-regressive Burg method. The *pburg()* matlab function from the signal processing toolbox is used for this purpose. When estimating the Burg model, the values of the samples outside the observation window are not assumed to be zero which allows the model to predict future samples outside the observation window and in accordance with the nature of the process. To find the model, the algorithm is usually based on minimizing the optimization condition, commonly the prediction error ρ = sigma^2. The parameter *order* from the *pburg()* has been tuned to optimise PSD estimation for each of the four cardiac signals.

When working with sound signals, the Burg algorithm or the Yule-Walker method are the most common choices for estimating the AR coefficients. The Burg algorithm is an extension of the Yule-

Walker method that is particularly well-suited for signals with highly non-stationary characteristics, like sound signals. [2][3]

After identifying the best PSD estimation, the function *psdpar()* will be applied to that PSD. Which is used to extract parameters like the Shannon entropy, skewness, and statistical frequency characteristics from the PSD. As the welch method is not well suited for the PSD estimation of sound signals, the PSD estimations generated with this method will not be analysed further using the *psdpar()* function.

The parameters extracted from the PSD estimation are useful for signal characterisation, and more specifically will be used to classify the signal given as input to the algorithm developed at the end of the project.

Time Frequency Representations

In order to characterize the frequency behaviour of the 4 types of signals across time different time-frequency representations (TFR) have been proposed. In this way it can be observed the frequency content of the signal in the parts containing pathological behaviour specifically.

The first proposed TFR are spectrograms on which different window sizes have been tested. In addition, the parameters window type (rectangular vs. hamming) and reassigned vs. non-reassigned were compared. Only these two types of windows have been tested, as well as those with more different behaviours. Due to spectrograms' strict compromise between time and frequency resolution, which is far apart from the theoretical maximum resolution $\Lambda f \cdot \Lambda t \geq \frac{1}{4\pi}$, other approaches have been explored.

Secondly, a scalogram for each of the signal was generated. A scalogram is a representation of the continuous wavelet transform (CWT) of a signal, more detailed it plots the squared magnitude of the CWT coefficients as a function of both time and frequency. The CWT is a mathematical technique that decomposes a signal into a series of scaled and translated versions of a chosen wavelet function. This is a powerful tool to analyse non-stationary signals, as it can detect changes in frequency and amplitude over time with higher resolution than spectrogram.

The parameters of a scalogram are essentially the envelope, the wavelet, and the number of oscillations of the wavelet k. The envelope type and the number of oscillations is changed to obtain the best scalogram result for each of the signals. The Morlet wavelet is used for our case.

The Morlet wavelet is a complex sinusoid modulated by a Gaussian function, which gives it a trade-off between time and frequency resolution. It has good time resolution because the Gaussian function decays rapidly, meaning that the wavelet has most of its energy concentrated in a small-time interval. It has good frequency localization because the complex sinusoid has a definite frequency. [4]

On third place, different quadratic TFRs have been tested, as the Wigner-Ville (WV), the Bohn Jordan (BJ), the Choi-Williams (CW) and the Generalised Exponential (GE) distributions. These distributions are based on the representation of the ambiguity function of the signal. Also, the ambiguity function is filtered analogously as by a 2-D low-pass filter with different parameters to tune depending on the given distribution in order to reduce the interferences that typically appear in QTFR. For instance, WV has no filter applied, and generated the TFR directly from the ambiguity distribution. On the other hand, BJ automatically tunes the parameters to deliver the most optimal filter to apply, while CW has the sigma parameter to tune, and the GE has sigma, N and M.

The point is that this Quadratic TFRs have interferences, which have nothing to do with the signal, and that should be filtered for best performance. These interferences appear side by side with the central part of the ambiguity function. The aim of good filter would be to have a wide vertical area of pass filter and a narrow horizontal line, to filter all the interferences while leaving the component of interest. The responsible for filter shape are the parameters mentioned above. Similarly in all the distributions, a rising sigma value increases the filter vertical and horizontal lines, avoiding filtering the signal but also the interferences. In the

specific case for the GE, an increasing M decreases filtering, while an increasing N increases filtering.

Finally, empirical mode decomposition was applied to the four signals. Empirical mode decomposition (EMD) is a method of signal decomposition that is based on the local maxima and minima of the signal. The basic idea behind EMD is to decompose a non-stationary signal into a set of intrinsic mode functions (IMFs), which are defined as the local maxima and minima of the signal. The components are oscillatory and localized in time and frequency. Each IMF represents a different frequency component of the signal, and together they can be used to reconstruct the original signal.

Algorithm

Our algorithm takes in four inputs: a signal of one heartbeat, measuring 0.7 seconds and sampled at a frequency of 1000 Hz, referred to as 'signal'; a time vector for the signal, referred to as 't'; an analysis type with options of PSD or TFR, referred to as ‘analysistype’; and a type of signal pathology, referred to as 'type', with 4 different options including normal and pathological signals.

The algorithm first checks if there is a 'type' variable present. If not, it uses a *k-nearest neighbour* classification algorithm to determine the type of signal pathology. After this the type of signal is determined and the algorithm proceeds with 4 cases of if statements, each corresponding to a specific signal pathology. If the ‘type’ variable is present only one of these if statements will be executed, depending on type of signal. These if statements adjust the parameters for PSD and TFR estimations to the optimal parameters identified prior in the results for each signal pathology.

The next block of the algorithm then checks the ‘annotationtype’ input and returns the appropriate plot. If the PSD option is chosen, the output is a plot that includes the signal, the PSD estimation made with the Burg method (pburg) and optimized order, and the PSD estimation including PSD parameters using psdpar.m. Finally, a variable that is a structure containing all extracted parameters of the PSD is returned.

If TFR is chosen, the output is a plot that includes the signal, the scalogram with optimized k, the quadratic representations based on the generalized exponential distribution with optimized sigma, N and M, and again a variable that is a struct containing all characterization parameters of the signal.

RESULTS

The following estimations from the results chapter are performed with one heartbeat only.

Power Spectral Density Estimation

The criteria to decide which PSD parameters perform best are as follows:

I. The PSD has clear and distinct peaks, indicating the presence of a dominant frequency in the data
II. The PSD should have a smooth and continuous shape, with minimal noise or distortion

Depending on the signals characteristics the PSD shape can have different appearances. For the different PCG signal the PSD was generated using the Welch and Burg Method.

(1) Welch

PSD calculation using *pwelch()* matlab function has been done and identification of the optimal performance of the welch periodogram under changing window size D and overlap S has also been assessed

Small values of D should theoretically deliver high temporal resolution and low frequency resolution. As the welch periodogram is an averaging frequential representation, spatial resolution is not directly visible, but frequential is. What can be derived from this is that as welch periodogram is an averaging method, the smaller the window (D) and the higher the overlap (S) the more averaged will be our PSD, and less variability of the possibly non-stationary components of the signal will not be seen.

In reality, the overlap (S) parameter tunning plots have not been shown in this result section because compared to the window size (W) its effect on PSD estimation shape and quality was not significant.

Normal Signal

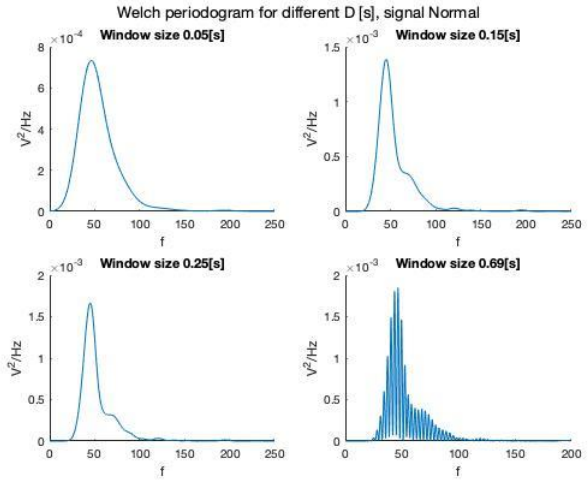


Figure 2: Welch Periodogram of Normal heartbeat over D

As mentioned in the previous paragraph, the secondary "peak" around 70Hz is averaged when the window is decreased, meaning that its energy probably comes from a component of the signal which is non-stationary. Furthermore, this right peak could be useful for heartbeat classification.

DFS2

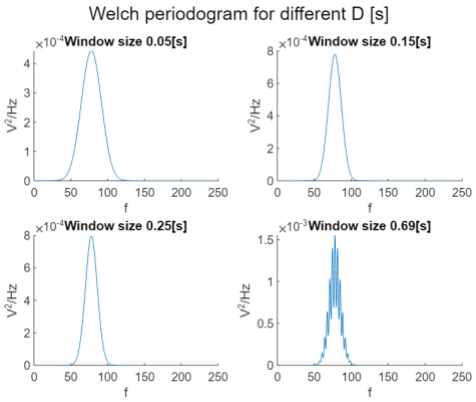


Figure 3: Welch periodogram of DFS2 signal with changing window size and overlap of XXX.

The PSD plot illustrates a single prominent peak for smaller window sizes. As the window size increases, additional peaks become visible. This is because a larger window size improves the frequency resolution. This DFS2 PSD representation shows that this pathology has a frequency peak around 80Hz and a symetrical shape, which could later be used for pathology identification.

Ejection murmur

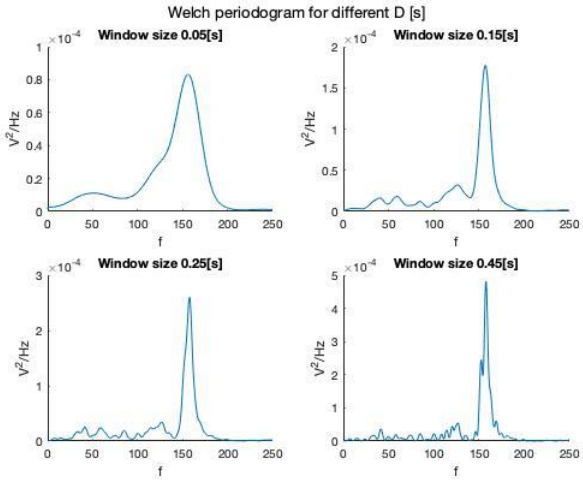


Figure 4: Welch Periodogram of Ejection murmur over D

As stated before, increasing window size increases the frequency resolution which is why most peaks are identified using the biggest window. The PSD shows a clear frequency peak around 160 Hz and a characteristic tail to the left of the peak, which will later be useful for pathology classification.

The EM signal has higher frequency components which are also visible in the PSD when comparing it to any of the other heartbeat labels.

SMR signal

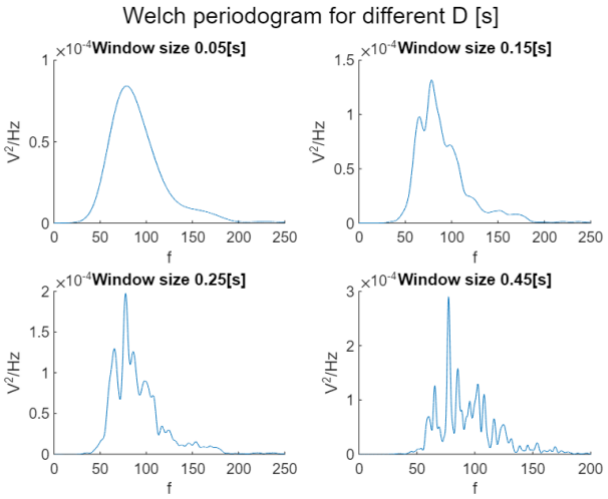


Figure 5: Welch Periodogram of Systolic Mitral Regurgitation heartbeat over D

The SMR signal, like the EM signal, contains higher frequencies than the normal and DFS2 signals. The Welch periodogram method may not provide an accurate representation of these signals as it has lower frequency resolution and may show a high degree of variability, resulting in overlapping peaks.

(2) Burg Method

The pburg function takes a sequence of samples of a signal as input and returns an estimate of the PSD as output. It has the order of the model and the number of samples used for the fast Fourier transform (NFFT) as parameters. A NFFT of 4096 samples was set and the order of the estimation changed to see which order performs best for the different signals. The order of the AR model refers to the number of previous time steps used to predict the current time step. When the order is increased, the AR model becomes more complex and can better approximate the underlying PSD of the signal. Meaning that the higher the order the higher the frequency resolution and the estimation confidence of the resulting PSD. As the signal is very short several high orders between 200 and 320 are tested to see if the PSD quality is sufficient. The Results for the normal signal are displayed in figure 6.

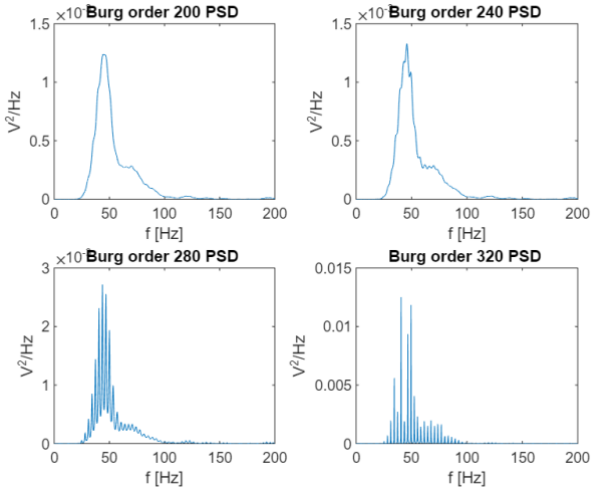


*Figure 6: PSD estimation of the **normal signal** with Burg method for different model orders.*

All estimations show a smooth trajectory, but the best result was obtained with an order of 320 for the normal signal. This PSD shows all peaks very clearly and shows no signs of noise or distortion. Therefore, this PSD will be used for the *psdpar()* function, which generated the result shown in figure 7.

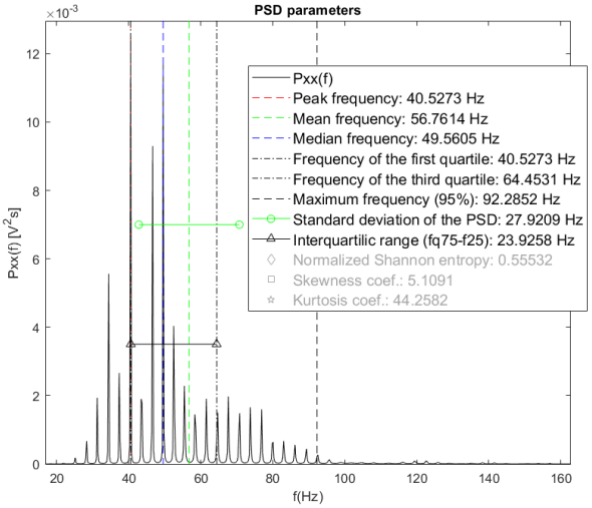


*Figure 7: The characterization of the PSD estimation using the burg method of the **non-pathological PCG**.*

The analysis of the signal's power spectral density (PSD) using Shannon entropy and skewness coefficient shows that the signal has a high level of uncertainty, with an estimation uncertainty of approximately 56%. The skewness coefficient also indicates that the distribution of the signal's PSD is not symmetrical, with the tail of the distribution leaning towards the right side of the maximal frequency. The maximum frequency of the signal is found to be 57 Hz, with the mean frequency lying around 36 Hz. This suggests that most of the signal's energy is concentrated in the lower frequency range of the signal, with less energy present at higher frequencies.

The PSD of the diastolic signal shows only two main peaks which are most visible using order 280, even though also order 320 allows a clear peak identification (see figure 8). No new peaks can be identified when further increasing the model order.

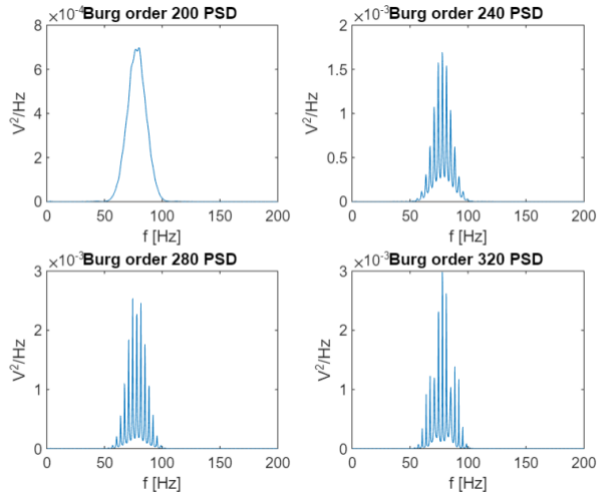


*Figure 8: PSD estimation of the **DFS2 signal** with Burg method for different model orders.*

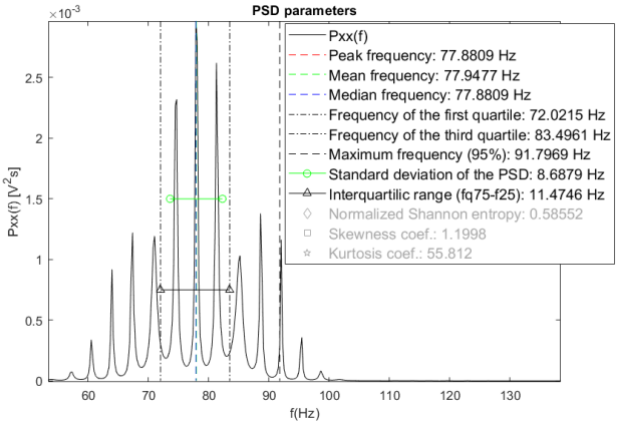


*Figure 9: The characterization of the PSD estimation using the burg method with order 280 of the **DFS2 PCG**.*

The analysis of the diastolic fixed S2 split signal's power spectral density (PSD) using Shannon entropy and skewness coefficient reveals that the signal has a slightly higher level of uncertainty than the normal signal, with an estimation uncertainty of around 58%. The mean frequency of the signal is found to be around 49 Hz. The PSD is also found to be relatively symmetrical, as reflected by the low skewness coefficient. However, it still shows a tail on the right side of the signal. This suggests that the signal's energy is distributed relatively evenly across a wide range of frequencies, with a slight concentration towards higher frequencies.

The PSD estimate of the ejection murmur signal is shown in figure 10. Since the signal contains higher frequency components than the normal and DFS2 signals, the order of the Burg method should also be

increased. The identified peaks increase with order, so the PSD for orders 360 to 480 was calculated with a step size of 40, so the PSD with order 440 performed best and was further analyzed using the psdpar() function shown in figure 11.

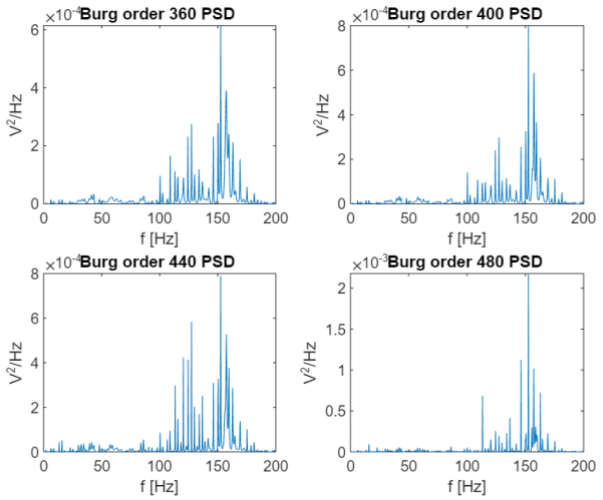


*Figure 10: PSD estimation of the **ejection murmur signal** with Burg method for model order 440.*

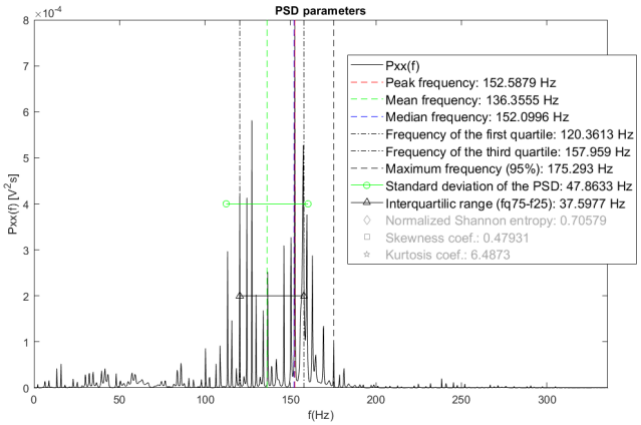


*Figure 11: The characterization of the PSD estimation using the burg method of the **ejection murmur PCG**.*

Even with a higher order the Shannon entropy for the EM signal is higher than for the normal and the DFS2 signal, which is probably due to the many different frequency components present in the EM signal compared to the other two respectively. The skewness coefficient is close to zero meaning that the PSD is quite symmetrical. The mean frequency is 86 Hz, and the maximal frequency is 110 Hz. Both are higher than for the DFS2 and normal signal which is concordant with the signal's appearance.

For the SMR signal the PSD estimation also requires higher orders than the normal and the DFS2 signal. The orders 360 to 480 were tested again, and the estimation with an overlap of 440 looks best to characterize the signal, as it has the highest frequency resolution (see figure 12). The SMR signal is the only one showing a certain white noise component around 0 in the signal between the heart sounds, where the other signal has no visible noise component. This noise might be represented in the PSD on the right sided tail.

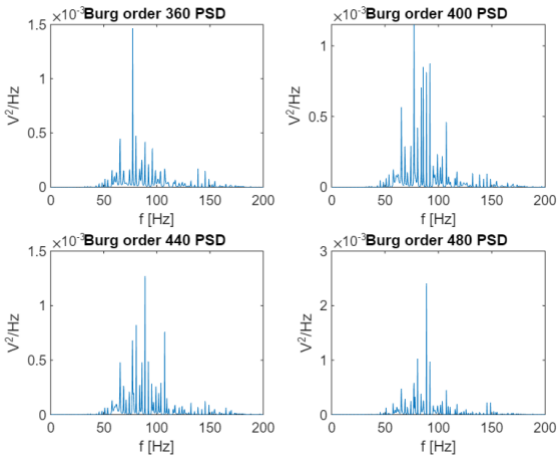


*Figure 12: PSD estimation of the **SMR signal** with Burg method with order 440.*

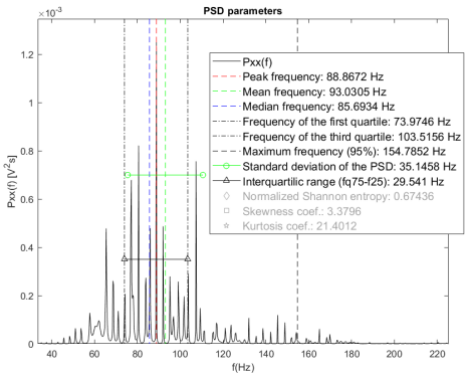


*Figure 13: The characterization of the PSD estimation using the burg method with order 440 of the **SMR PCG**.*

The Power Spectral Density (PSD) of the SMR signal exhibits a similar shape to that of the EM signal, as it also contains higher frequency components (see figure 13). Furthermore, its shape has a right-skewed tail, as evidenced by its skewness coefficient of approximately 3.4 and its Shannon entropy of 67%, which is higher than that of the normal and DFS2 signals.

When analysing the mean frequency and maximal frequency of the PSD, it can be noted that the mean

frequency is 58 Hz, and the maximal frequency is 97 Hz. This is lower than that of the EM signal, but higher than that of the normal and DFS2 signals.

Using the Burg method for PSD estimation, it is observed that smooth trajectories and clean peaks can be obtained for high enough orders. The best trade-off between quality and computational effort was found to be between 320 and 440 samples, depending on the characteristics of the signal.
These results suggest that it may be possible to differentiate with high confidence between the normal signal and the SMR or EM signals, as they exhibit distinct characteristics. To distinguish between the normal and DFS2 signals, the skewness coefficient seems to be the most promising feature, as the PSD shape of the normal signal has a right-skewed tail, whereas the DFS2 signal is highly symmetrical.
The same applies to the SMR and EM signals, but they also have more distinct peak frequencies, which may be a better method of differentiation in this case. However, it should be noted that the confidence of the estimations is not very high, which could lead to false classifications of pathologies.

When evaluating the performance of the Welch and Burg methods for analysing this type of signal, it was apparent that the Burg method would yield superior results. This was confirmed by the outcome of our analysis, as the Power Spectral Density (PSD) estimation using the Welch method yielded inferior results compared to the Burg method. As a result, the Burg method will be used in the final algorithm for determining the characteristics of the input signals.

Time Frequency Representation

(1) Short Fourier Transform (Spectrogram)

The spectrogram is a TFR that, as mentioned in the *"Materials and methods"* section, has a good energy fidelity, but a very low combined frequency and time resolution. For what concerns us in this work, which is the characterisation of the four different cardiac signals, we have not used the components' energy from TFRs.

The theoretical maximum frequency and time resolution is shown in the *"Materials and methods"* section, and it presents a compromise between the two, which is modulated by the size of the window of the spectrogram. The smaller the window in time, the higher temporal resolution and the lower the spatial resolution of the spectrogram. This can be seen in the following figures and has been taken into consideration to decide which value for window size (D) is most suitable for the TFR of our cardiac heartbeats. Furthermore, the type (shape) of the window applied (rectangular, hanning, hamming etc.) algo modulates this resolution compromise. The rectangular and hamming window shape have been chosen to evaluate their suitability for TFR too.

Normal signal

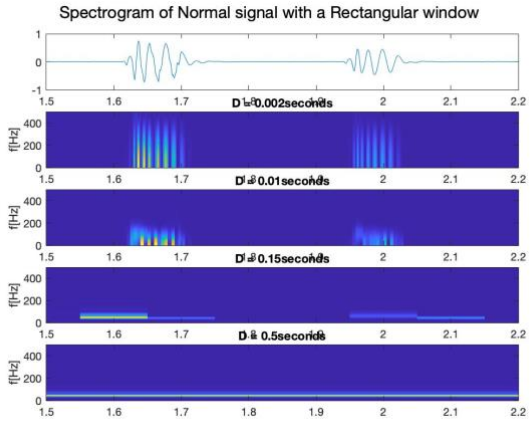


Figure 14: Spectrogram of ***Normal PCG*** *with hamming window and different window sizes (D).*

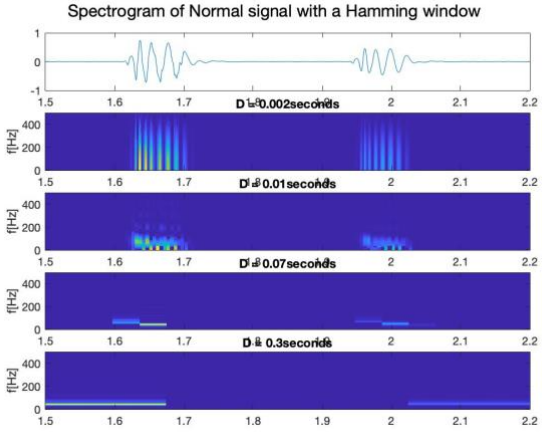


Figure 15: Spectrogram of ***Normal PCG*** *with hamming window and different window sizes (D).*

With the 0.07- and 0.15-seconds windows it is visible that there is at least two different components in the signal. Also, the hamming window presents better time and frequency

resolution than the rectangular one (as compared in D=0.01).

SMR signal

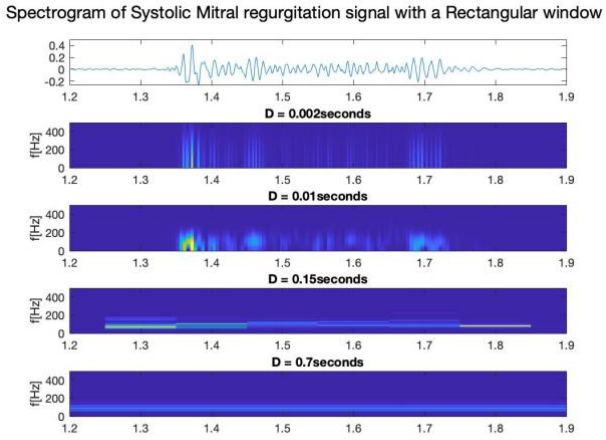


*Figure 16: Spectrogram of **SMR PCG** with rectangular window different window sizes (D).*

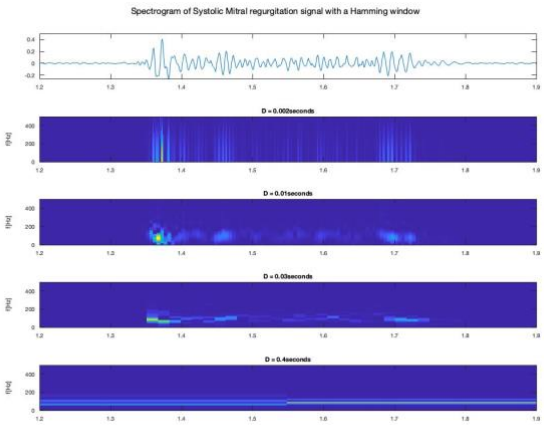


*Figure 17: Spectrogram of **SMR PCG** with hamming window different window sizes (D).*

Both spectrograms show information of the pathological are between cardiac sounds S1 and S2, being best observed in the rectangular window of D=0.15, where it is visible that the inter S segment presents higher frequencies than the S1 and S2 segments, being so the regurgitation.

Ejection murmur

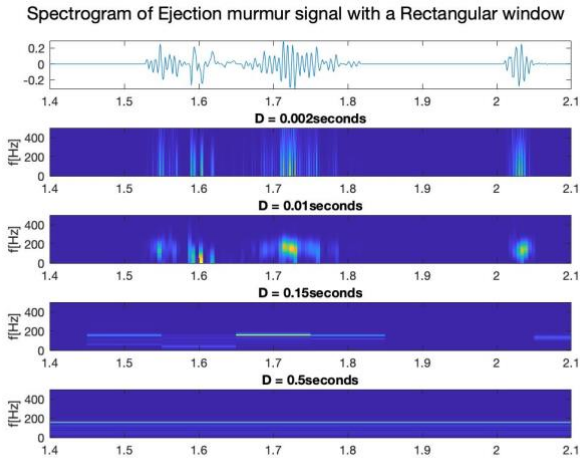


*Figure 18: Spectrogram of **ER PCG** with rectangular window different window sizes (D).*

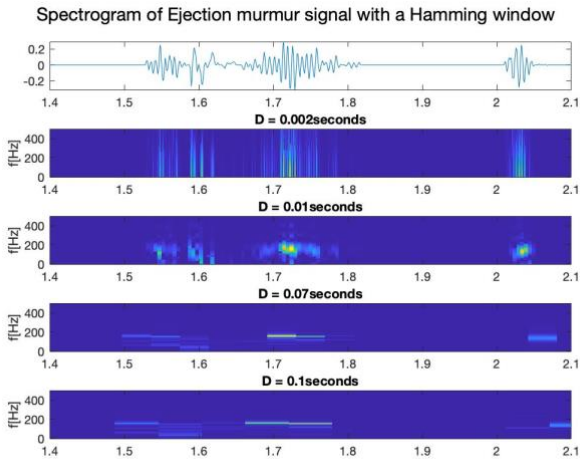


*Figure 19: Spectrogram of **ER PCG** with hamming window different window sizes (D).*

The heart ejection murmur, which is visible as the intense wave right after the S1, presents frequencies of almost 200Hz, far beyond the ones expected for S1. The hamming window with width of 0.07s has been evaluated to be the best one for frequency characterisation, while maintaining acceptable spatial resolution.

DFS2

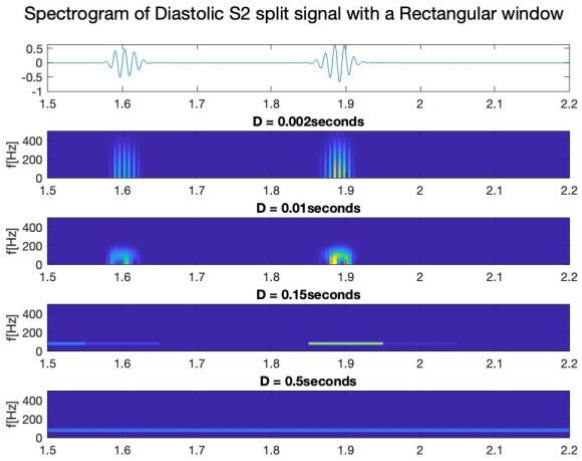


*Figure 20: Spectrogram of **DFS2 PCG** with hamming rectangular different window sizes (D).*

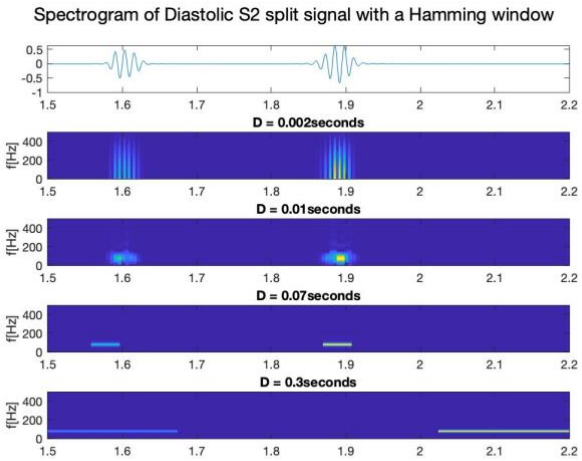


*Figure 21: Spectrogram of **DFS2 PCG** with hamming window different window sizes (D).*

At first sight, the results from DFS2 appear really similar to those of the normal heartbeats, being the mean difference the duration of both S1 and S2. A D=0.07 seconds with hamming seems to deliver the best resolution for DFS2.

In general terms the spectrogram is useful as a TFR first insight, but for visual characterisation of the signal other TFR are due to be used to evaluate their performance, as the scalogram, the QTFR or the EMD.

(2) Continuous Wavelet Transform (Scalogram)

The Hanning and exponential envelopes were both tested on each of the signals, with the exponential envelope always showing much better time and frequency resolution. Since the sampling rate is already decimated but still very high (1000 Hz), and the signal between the two heart sounds does not contain time-frequency information, the signals that have a significant pause between S1 and S2 are split into these two sections, respectively.

To find a suitable number of oscillation the k was reduced and increased iteratively, which revealed that low oscillation numbers performed better, which is why the k parameter was changed starting from $4/\pi$ up to $12/\pi$ for each of the signals. To visualize the power distribution the main ridges of the signal are plotted using the ftridge() function using just one main ridge. The results are presented in the following figures 22 to 25.

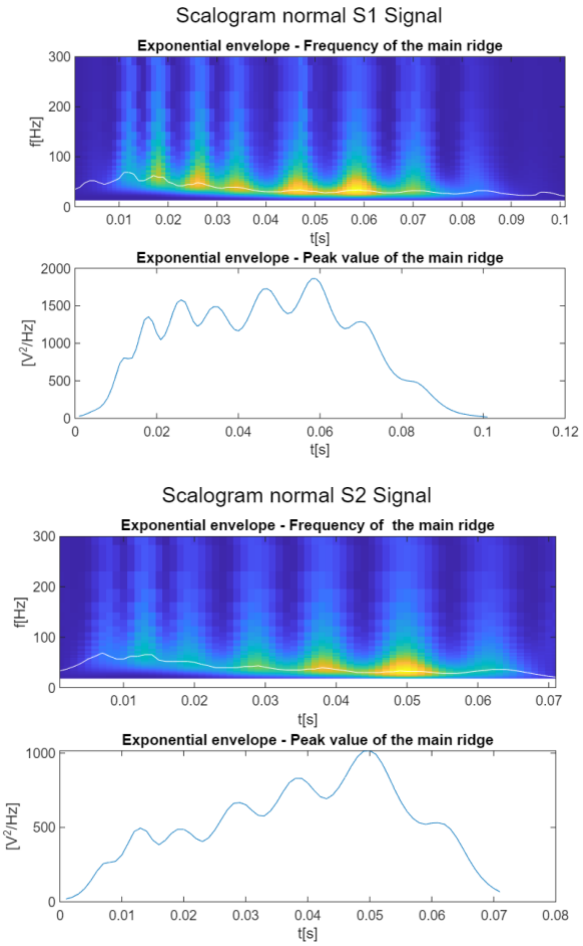


Figure 22: Scalogram of normal S1 signal (top) and S2 signal (bottom).

Peak output occurs 0.06 seconds after the first heart sound and 0.05 seconds after the second heart sound begins. The total output of the first heart sound (~ 1800 V^2/Hz) is almost twice that of the second (~ 1000 V^2/Hz), which makes physiological sense since

the first heart sound is produced by ventricular systole while the second heart sound is produced by ventricular diastole.
S1 is the sound produced by the atrioventricular (AV) valves closing during ventricular systole, while S2 is the sound produced by the semilunar (SL) valves closing during ventricular diastole.

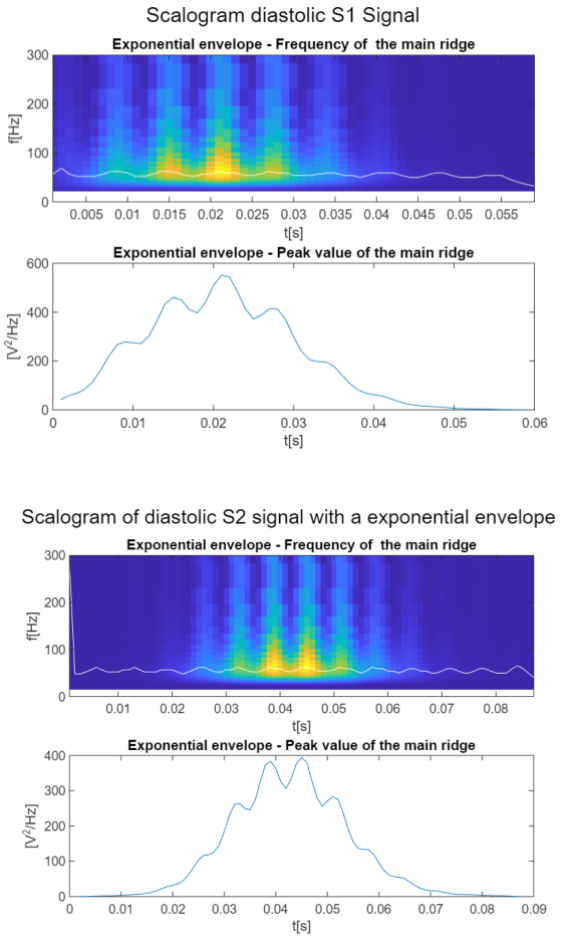


Figure 23: Scalogram of DFS2 S1 signal (top) and S2 signal (bottom).

The DFS2 signal shows a similar overall energy in the first and second heart sound. Both shapes of the heart sounds are also similar. From a pathological point of view this makes sense as the split between S1 and S2 is always present in a person with DFS2, even during diastole, which is the period of relaxation between heartbeats. This can be caused by several factors, such as a delay in the closure of the aortic valve due to aortic stenosis (narrowing of the aortic valve). This can lead to less energy in the heart sounds because the valve is not closing as quickly or as efficiently as it should. [5]

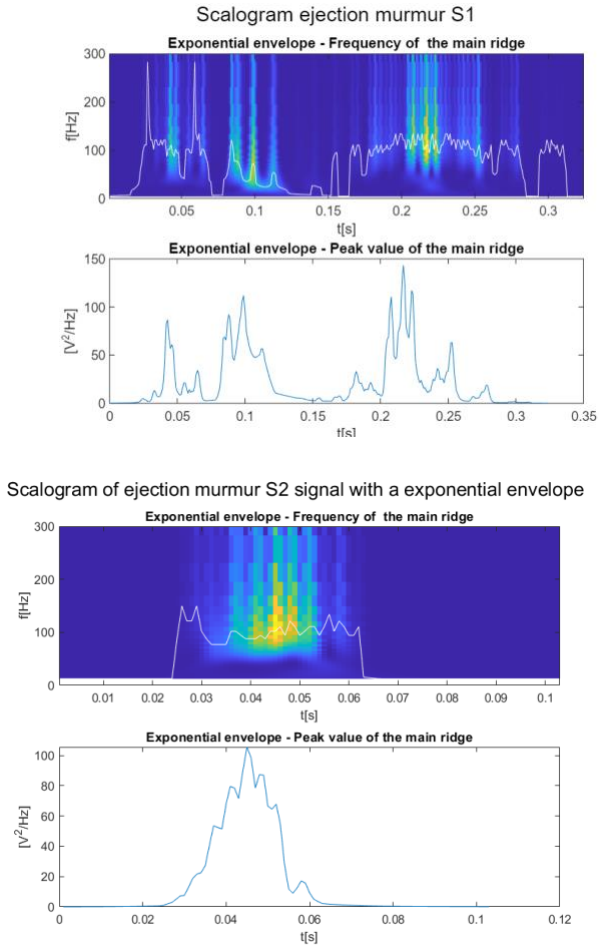


Figure 24: Scalogram of EM S1 signal (top) and S2 signal (bottom).

In the scalogram of the EM signal, we observe a distinct lack of the smooth progression of energy over time that is typically present in normal heart signals. Additionally, the overall power of the EM signal is much lower than that of the normal or DFS2 S1 signal, with maximum power reaching around 100 V^2/Hz. This is likely since EM sounds are typically heard during systole, which is the period of contraction of the heart. It is reasonable to expect that the S1 sound would be quite different from that of a healthy heart, as the pathological sound is produced by the rapid flow of blood through a narrowed heart valve, such as from the left ventricle into the aorta or from the right ventricle into the pulmonary artery during systole.
The S2 is less powerful and shorter in duration than the S2 of a healthy heart and has a similar shape and duration as the DFS2 S2 signal.

In the case of the SMR signal there is no clear separation between S1 and S2 which is why this signal is not divided into two sections.

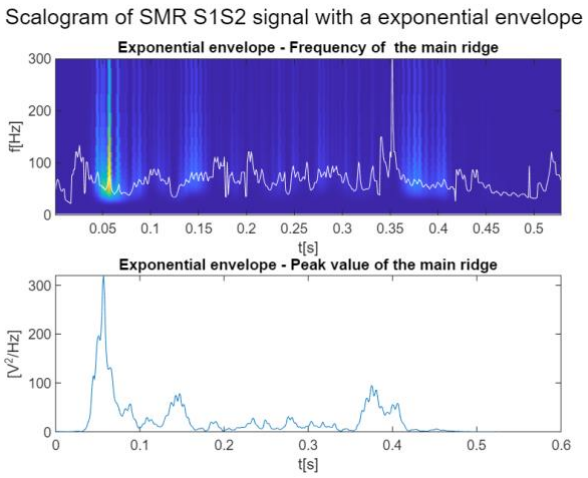


Figure 25: Scalogram of SMR S1 and S2 signal.

In the signal of a patient with SMR, the temporal energy development of the heart sound signal deviates from the normal, smooth pattern, also probably due to the k value. The SMR signal has a small interval of about 0.2 seconds, which may be the pause between S1 and S2. Within this interval, there are low energy components present, typically around 25 to 30 V^2/Hz, that indicate the backflow of blood through the mitral valve, which is not closing completely in people with SMR.
The backflow of blood results in the heart sounds being less loud and less well-directed than those of a healthy heart. This is reflected in the results, where the maximal power of the SMR signal is around 300 V^2/Hz, which is weaker than the maximal power of the normal heart signal. Additionally, the first heart sound has more energy than the second, as is typical for a healthy heart, but at 300 V^2/Hz, it is still much weaker than the healthy or diastolic fixed S2 split sound. The second heart sound reaches around 100 V^2/Hz, like the ejection murmur sound.

(3) Quadratic Time Frequency Representation

The theory about quadratic time frequency representation has already been mentioned in the "Materials and methods" section, and how their approach is to try to filter the ambiguity function to take out the interferences seen in the WV distribution. It was also commented that the different distributions where different in the senso of the filter they applied and the parameters to tune this filter.
As a matter of fact, we have seen that the quadratic distribution which presented better results for all the four signals was the generalised exponential (GE) distribution, which also was the one with more parameters to tune, three in this case (sigma, M and N), whose effect on the filter has also been discussed in the "Materials and methods" section.

Furthermore, only the GE will be shown in this results section, and the other distributions can be seen in the matlab code.

Normal Signal

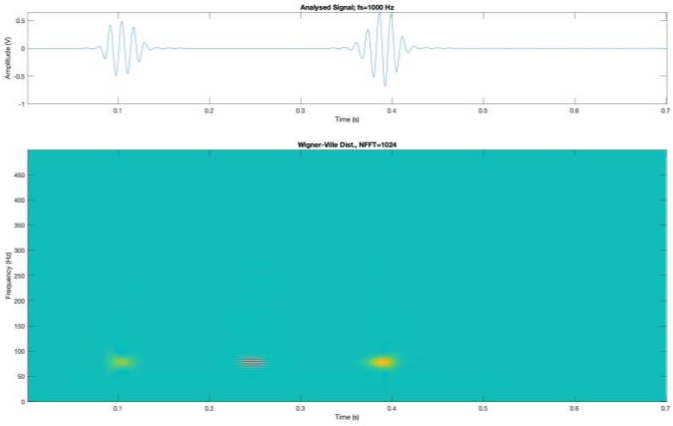


Figure 26: WV QTFR analytical with visible interferences between S1 and S2. Normal Signal

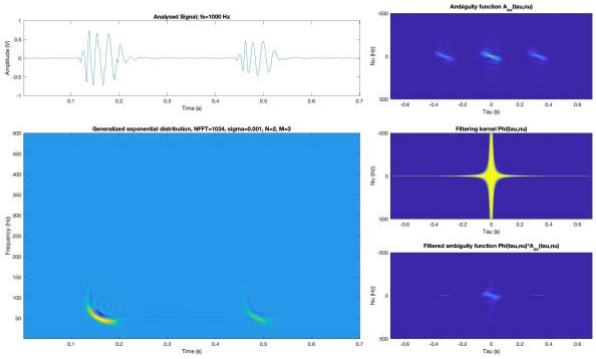


Figure 26:GE QTFR with visibly filtered interferences between S1 and S2. Normal signal

DFS2 Signal

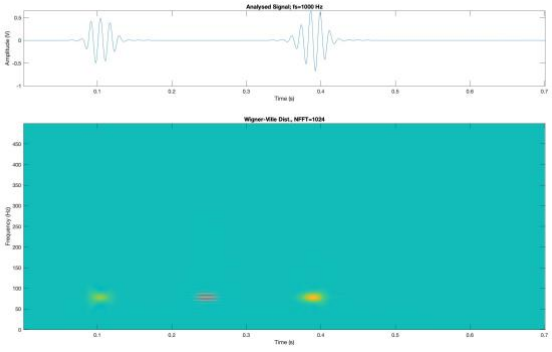


Figure 27: WV QTFR analytical with visible interferences between S1 and S2. DFS2 signal.

GE parameters: sigma=0.0001, M=3, N=2

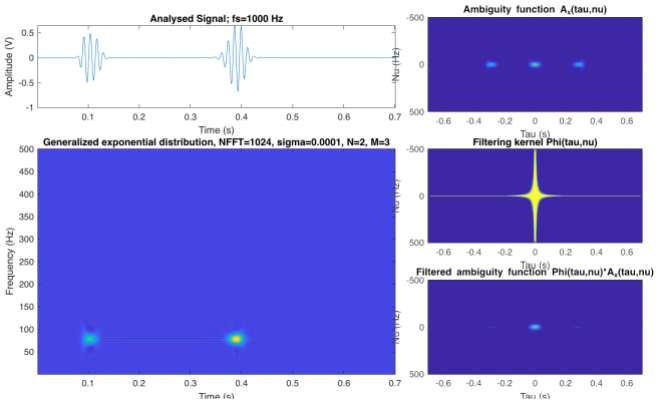


Figure 28: GE QTFR with visibly filtered interferences between S1 and S2. DFS2 signal

SMR Signal

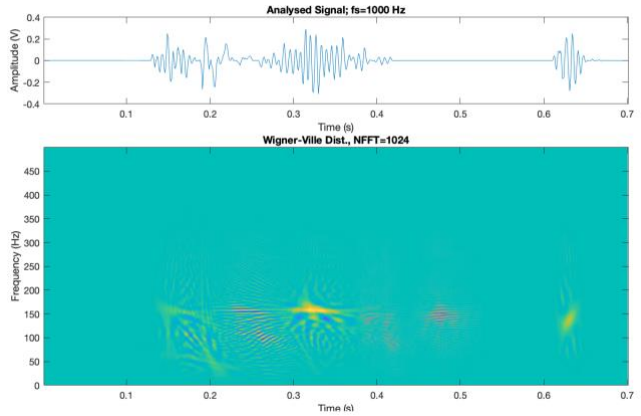


Figure 29: WV QTFR analytical with visible interferences between S1 and S2. SMR signal.

GE parameters: sigma=0.00005, M=3, N=2

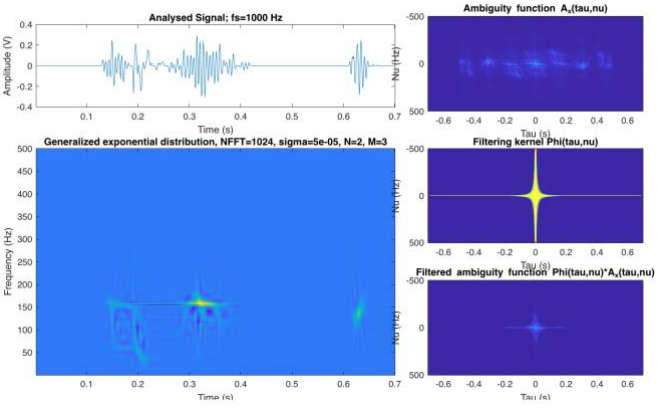


Figure 30: GE QTFR with visibly filtered interferences between S1 and S2. SMR signal

ER Signal

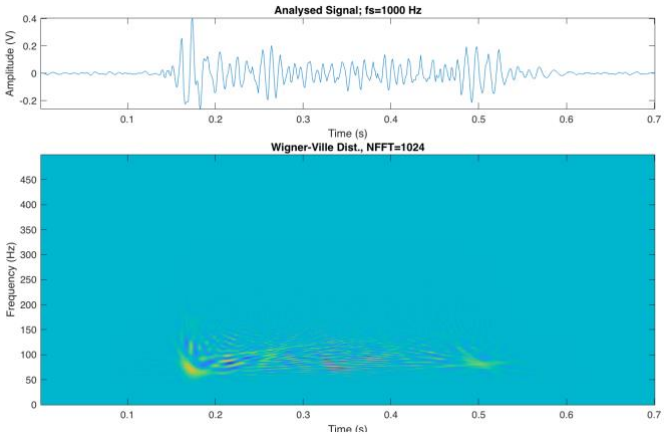


Figure 31: WV QTFR analytical with visible interferences between S1 and S2. ER signal.

GE parameters: sigma=0.5, M=3, N=2

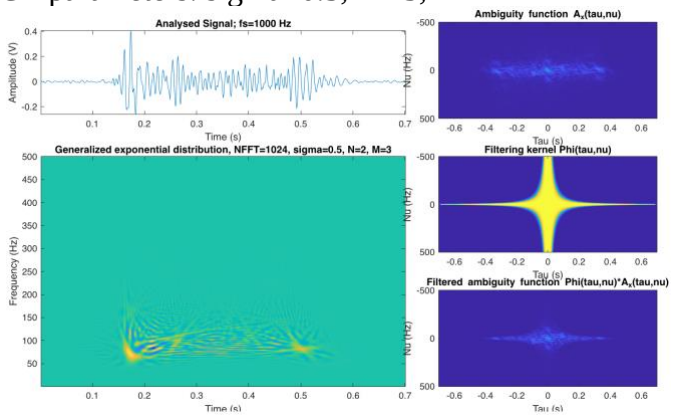


Figure 32: GE QTFR with visibly filtered interferences between S1 and S2. ER signal

As a summary of all the plots from this section, it is visible for each type of pathology that the interferences present in the WV TFR where totally or partially filtered by the GE distribution using different parameters.

The GE was the distribution showing better results above the CW, BJ and GE distributions, being therefore the chosen one for the processing of the heartbeats in the algorithm that will be explained in the sections below.

(4) Empirical Estimation

The EMD algorithm begins by identifying the signal's local maxima and minima, and then fits a line through these points to create the first IMF. This

process is repeated for the residual signal until a specific termination criterion is met, for example the number of IMFs or the energy of the residual signal. Each IMF represents a specific frequency of the original signal, and the sum of all the IMFs reconstructs the original signal. Using EMD, it is possible to eliminate specific components such as noise and reconstruct the signal. You can also extract relevant components for further analysis. [6]

The results of the EMD for the different signals is shown in figure 31 to figure 33.

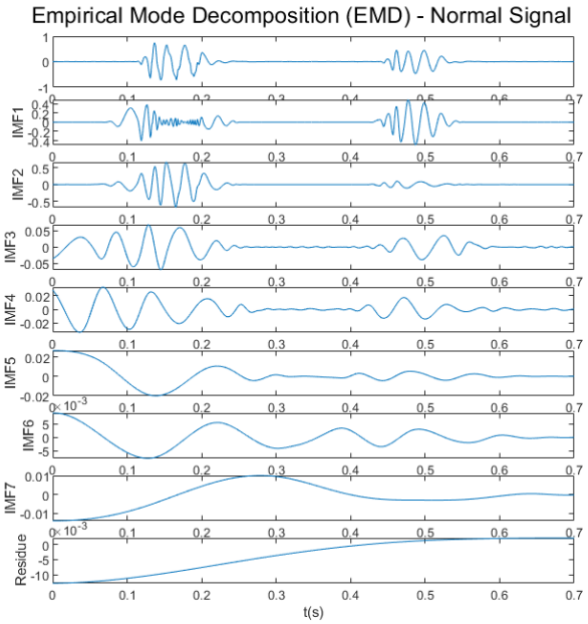


Figure 31: EMD of the normal signal, with original signal (top), IMFs and residue.

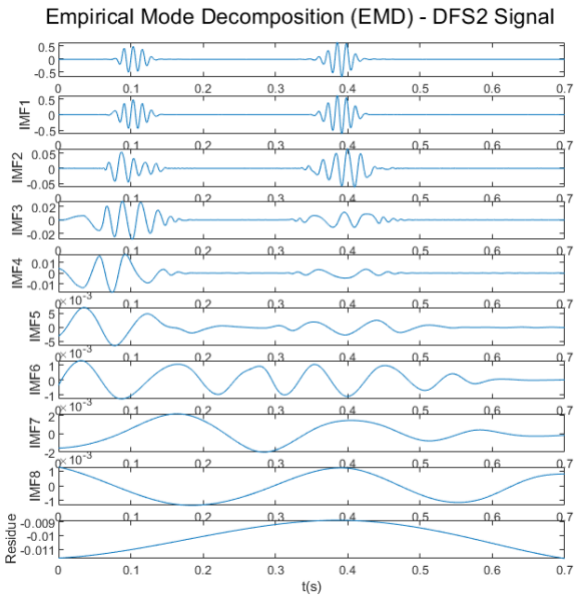


Figure 32: EMD of the DFS2 signal, with original signal (top), IMFs and residue.

When comparing the EMD of the normal and the DFS2 signal, it is visible that most IMFs are very similar besides the first one. But there is no specific IMF that seems promising for pathology classification. The same applies for the EM and the SMR signal. There are visible differences but the characterizations of the signals via EMD does not reveal much valuable information that can be used in a classifier at first sight.

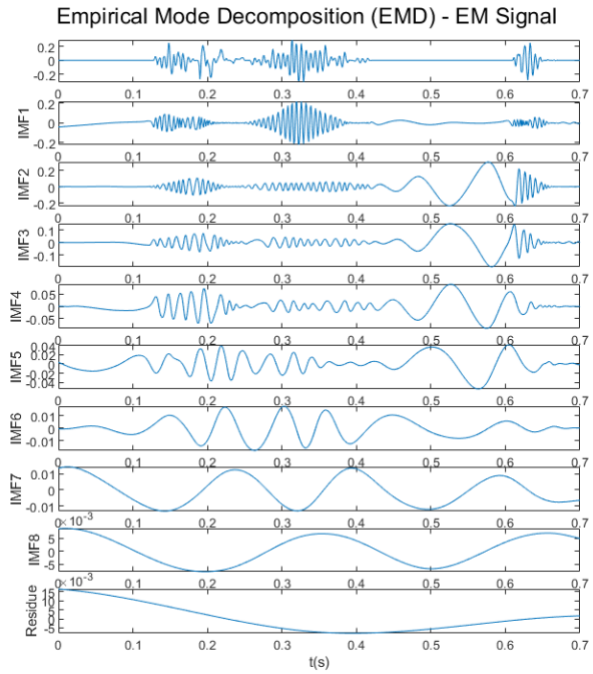


Figure 33: EMD of the EM signal, with original signal (top), IMFs and residue.

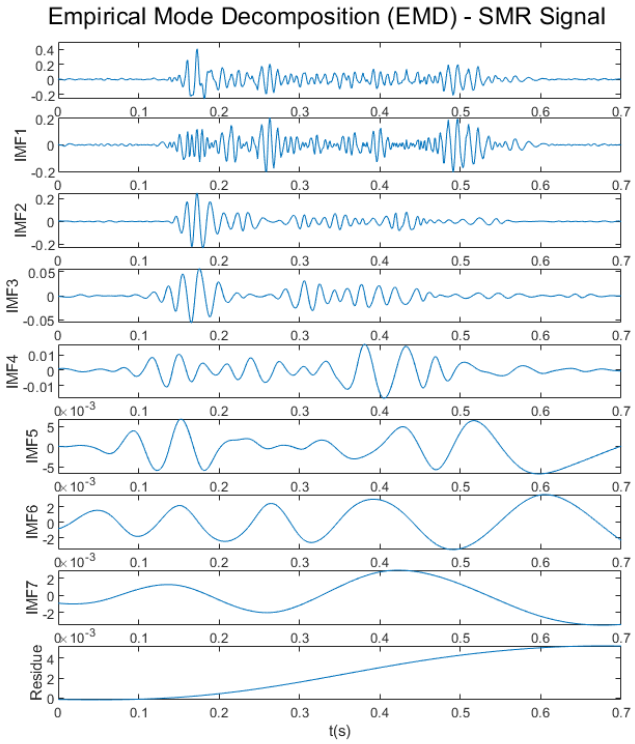


Figure 34: EMD of the SMR signal, with original signal (top), IMFs and residue.

With the EMD the frequency components of the signals were extracted allowing a better understanding of the signal's components against time. However, for the purpose of classifying the signals, EMD doesn't appear as the best choice as its interpretation is not as straight forward as the one from PSD of from the other TFR estimation.
Furthermore, data that contains breaks (like the EM and SMR signal) might detect peaks of the data belonging to different components on different time scales. As a result, the disassembled IMFs will consist of intermittent oscillating components spanning a wide range of different scales. This is known as the mode-mixing phenomenon which not only causes strong aliasing in the time-frequency distribution, but also obscures the physical meaning of individual IMFs.[7]

The EMD has other advantages that might be more relevant for other purposes. As data driven approach, EMD does not rely on assumptions about the signal which is useful when analysing an unknown signal.[7]

(5) Algorithm

The algorithm developed, as explained in the *"Materials and methods"* section, is based on two main blocks. The first one looks if the signal type (pathology) is introduced to the algorithm or not, and if it is not, the algorithm classifies the analysed signal into one of the four classes possible (normal, diastolic fixes S2 split etc.) using the parameters it extracts from the PSD estimation of the signal automatically.

Once the type of signal is already defined, the algorithm either delivers a PSD related plot or a TFR related plot depending on what we want. The best methods and the optimized parameters for each method are applied to each type of signal depending on their class and the knowledge acquired in this study.

The algorithm always returns the PSD parameters of the given signal, as well as the resulting plot with the optimized methods and parameters as shown in the figures below (the PSD title says the order used, in the Burg periodogram, changing and being always the optimal one) .

Even though further performance tests should be done, the algorithm seems to perfectly identify the type of pathology and correctly delivers the given parameters and figures.

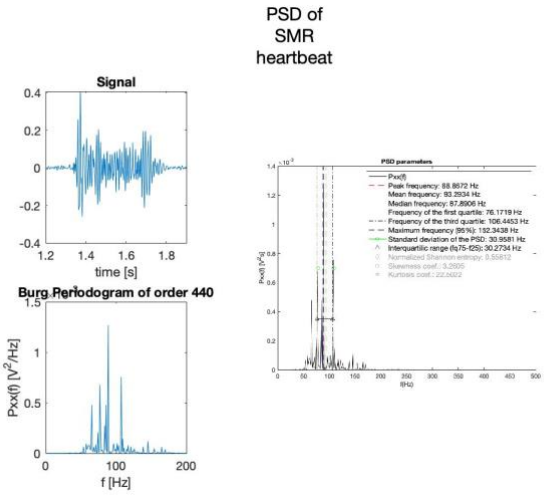


Figure 35: Algorithm Output for "PSD" of a SMR signal

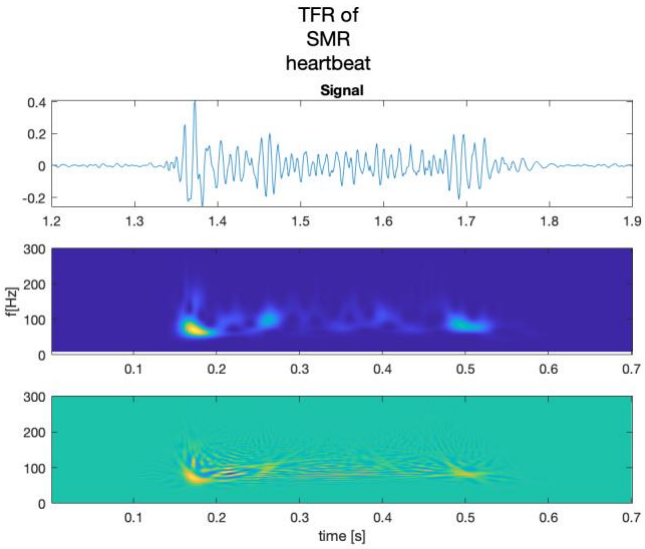


Figure 36: Algorithm Output for "TFR" of a SMR signal

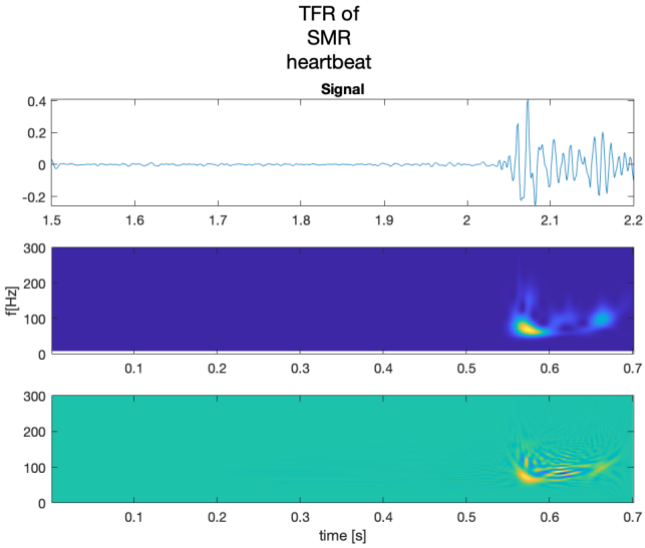


Figure 37: Algorithm Output for "TFR" of a random segment of SMR signal

As can be seen in the figures above, the algorithm detects the type of signal using the parameters extracted from the PSD even when the segment is not a heartbeat itself, but a random segment of the signal (the last part of the figure 37 signal is the beginning of the figure 36 signal).

CONCLUSION

In conclusion, an algorithm has been developed to detect pathological phonocardiography signals by utilizing power spectral density (PSD) parameters. The Burg method was found to be the most effective method for estimating the PSD (probably because the signal is non-stationary), providing accurate results for distinguishing between healthy, EM, DSF2 and SMR PCG recordings. While the scalogram also revealed different characteristics of the signals over time, it was not chosen as the primary method for classification due to its higher computational effort and complexity when trying to extract the signals characteristics. However, the characteristics from the scalogram are still provided as a visual output, offering a comprehensive characterization of the input signal over frequency and time.

ACKNOWLEDGMENT

We thank Professor Abel Torres for the material provided.

REFERENCES

[1] - Joseph J, Naqvi SY, Giri J, Goldberg S. *"Aortic Stenosis: Pathophysiology, Diagnosis, and Therapy."* Am J Med. 2017 Mar;130(3):253-263. doi: 10.1016/j.amjmed.2016.10.005. Epub 2016 Nov 1. PMID: 27810479.

[2] - Kari Roth, Ismo Kauppinen, Paulo A. A. Esquef, Vesa Valimaki *"FREQUENCY WARPED BURG'S METHOD FOR AR-MODELING"* in IEEE Explore, 2003

[3] - Abel Torres, *"Frequency-Domain Characterization"*, oral presentation WS22/23, Universitat de Barcelona/Universitat Polytecnica de Barcelona

[4] - M. V. Subbarao and P. Samundiswary, *"Time-frequency analysis of non-stationary signals using frequency slice wavelet transform,"* 2016 10th International Conference on Intelligent Systems and Control (ISCO), Coimbatore, India, 2016, pp. 1-6, doi: 10.1109/ISCO.2016.7726999

[5] - El-Segaier M, Lilja O, Lukkarinen S, Sörnmo L, Sepponen R, Pesonen E. *"Computer-based detection and analysis of heart sound and murmur."* Ann Biomed Eng. 2005 Jul;33(7):937-42. doi: 10.1007/s10439-005-4053-3. PMID: 16060534.

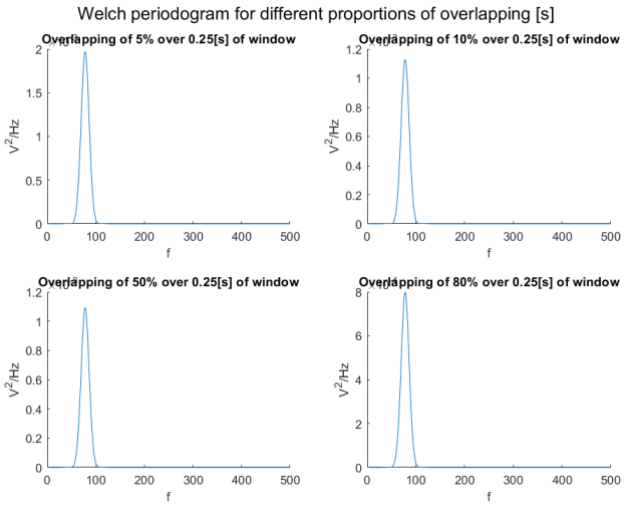
[6] - https://www.mathworks.com/discovery/empirical-mode-decomposition.html (Accessed 18.01.2023)

[7] - Wang YH, Hu K, Lo MT. Uniform Phase Empirical Mode Decomposition: An Optimal Hybridization of Masking Signal and Ensemble Approaches. IEEE Access. 2018;6:34819-34833. doi: 10.1109/ACCESS.2018.2847634. Epub 2018 Jun 15. PMID: 31106103; PMCID: PMC6521981.

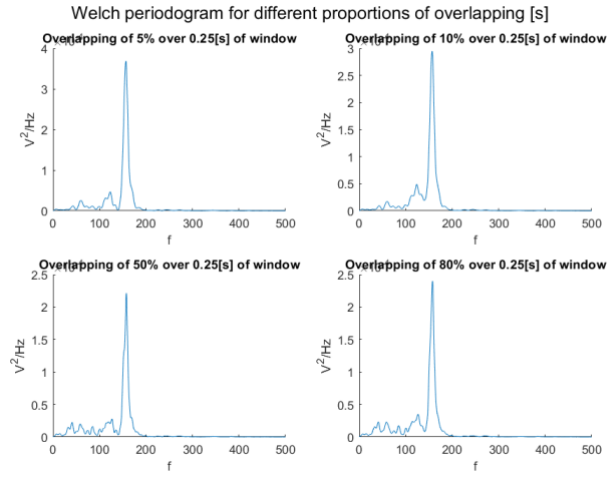
APPENDIX

WELCH PERIODOGRAMS

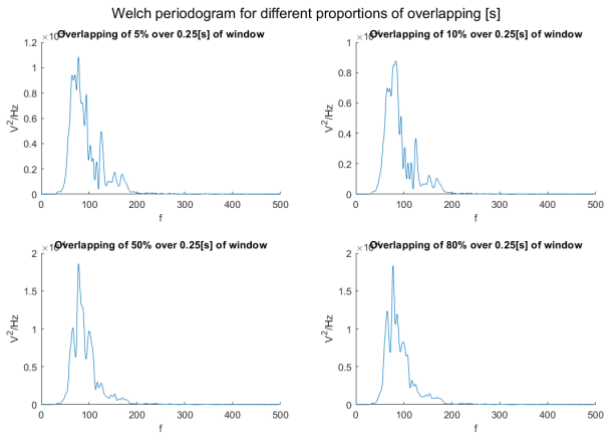
(1) DFS2 SIGNAL



(2) EM SIGNAL

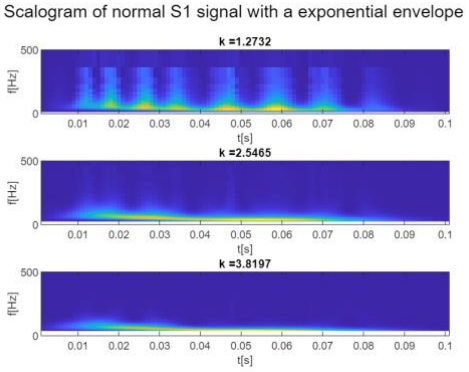


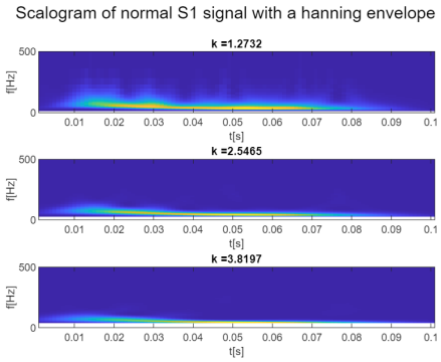
(3) SMR SIGNAL

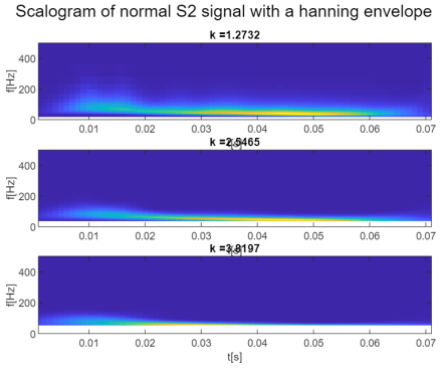


SCALOGRAMS

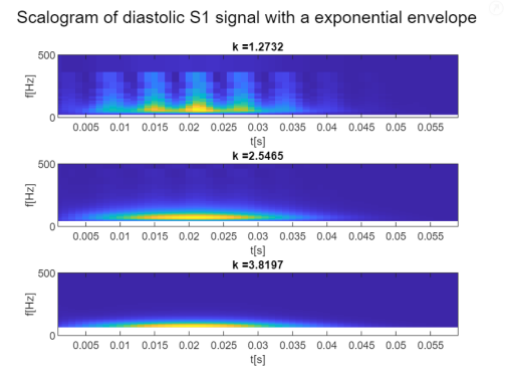
(1) NORMAL SIGNAL

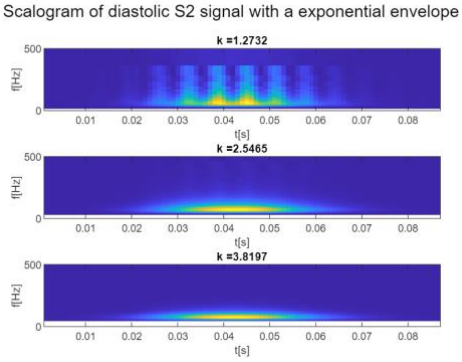




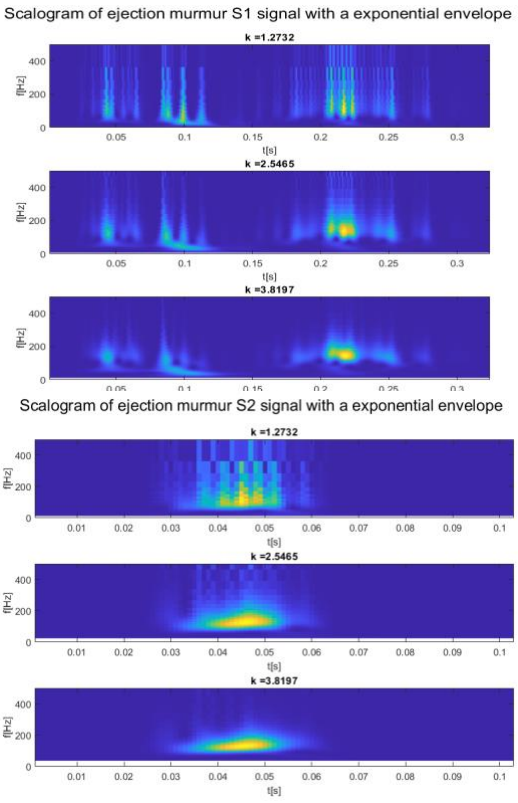


(2) DFS2 SIGNAL

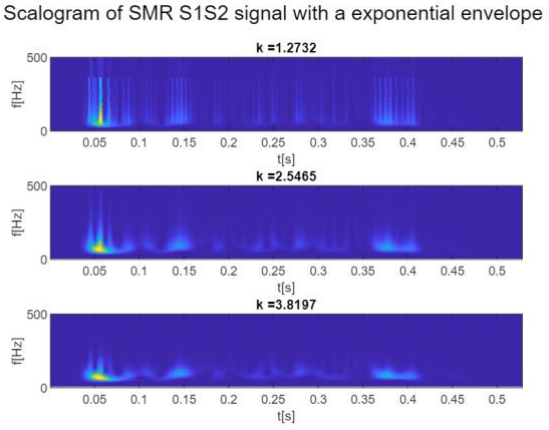




(3) EM SIGNAL


Scalogram of ejection murmur S1 signal with a exponential envelope
k =1.2732
k =2.5465
k =3.8197
f[Hz]
t[s]
Scalogram of ejection murmur S2 signal with a exponential envelope
k =1.2732
k =2.5465
k =3.8197
f[Hz]
t[s]

(4) SMR SIGNAL


Scalogram of SMR S1S2 signal with a exponential envelope
k =1.2732
k =2.5465
k =3.8197
f[Hz]
t[s]

QUADRATIC TIME REPRESENTATIONS

(1) NORMAL SIGNAL

(2) DFS2 SIGNAL

(3) EM SIGNAL

(4) SMR SIGNAL