

Automatic heart sounds analysis using Empirical Mode Decomposition and Higher-Order Spectra

Foteini Savvidou, Stavros Tsimpoukis, Stavros Filosidis, Chrysoula Psyrouki, Konstantinos Barmpounakis and Leontios Hadjileontiadis

Abstract—Cardiovascular diseases (CVD) are one of the leading causes of morbidity and mortality in the world. The diagnosis of heart disorders is a useful but usually difficult task. Phonocardiogram (PCG) signals contain valuable information about the mechanical function and state of the human heart that is useful in the diagnosis of such diseases. The present work proposes a computer-aided technique based on the analysis of the fundamental heart sounds (FHS), i.e., the first heart sound (S1) and the second heart sound (S2), using Empirical Mode Decomposition (EMD) and Higher-Order Spectra (HOS). EMD allows the decomposition of the PCG signals into several oscillatory components (Intrinsic Mode Functions, IMF) increasing thus the observability of the FHS. By employing a kurtosis-based criterion, we located the IMFs that contain relevant information regarding S1 and S2. On those IMFs, we estimated the bispectrum of FHS and explored features derived from the magnitude of the bispectrum for the quantification of heart conditions (normal and abnormal). Statistical tests were used to facilitate feature selection and the derived features were fed to a decision tree-based ensemble model classifier (XGBoost) for the distinction of normal and abnormal heart sounds. Experimental results have shown that, overall, the XGBoost classifier determines the heart condition in a percentage of 82% balanced accuracy. Moreover, cepstral analysis was carried out for the identification of the periodicity of a heart cycle and the acquisition of the impulse response that is affiliated with the heart sound signals (normal and abnormal). The results suggest that the bispectral and cepstral analysis of the PCG signals can facilitate the diagnosis of heart disorders in everyday clinical practice.

Index Terms—Empirical mode decomposition, Bispectrum, Cepstrum, Phonocardiogram, Heart sounds.

I. INTRODUCTION

CARDIOVASCULAR DISEASES (CVDs) are the leading cause of death globally, taking an estimated 17.9 million lives each year [1]. Medical inspections should be performed on a regular basis for the timely detection and diagnosis of heart abnormalities. As a first step, a process automated to the greatest extent, which provides preliminary information, i.e., whether examined heart condition is normal or abnormal, would give to the healthcare systems worldwide a valuable asset for the fight against CVDs. The ultimate aim of this process is to refer patients to further assessment early enough.

Although there are several medical tests to diagnose heart conditions, there are a number of factors, such as cost-effectiveness, simplicity and portability, that should be taken

The authors are with the Signal Processing and Biomedical Technology Unit, Telecommunications Laboratory, Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, GR 54124 Thessaloniki, Greece (e-mail: sfoteini@ece.auth.gr; tsimpous@ece.auth.gr; stavrosf@ece.auth.gr; chrysyr@ece.auth.gr; kmparmpou@ece.auth.gr; leontios@auth.gr)

into account in search of the appropriate physical examination. Electrocardiogram (ECG) and Phonocardiogram (PCG) recordings constitute two of the most commonly used techniques for identifying heart issues. ECG signals are produced by the electrical activities of the heart, while PCG signals are produced by mechanical activities of the heart. In particular, a PCG is a graphic record in the form of a wave in which you can see the heart sounds obtained with a stethoscope. The PCG allows to provide data on the timing, relative intensity, frequency, quality, tone, timbre and precise location of the different components of the cardiac sound, in an objective and repeatable manner [2].

The cardiac cycle is the performance of the human heart from the beginning of one heartbeat to the beginning of the next. It consists of two periods: one during which the heart muscle relaxes and refills with blood, called diastole, following a period of robust contraction and pumping of blood, called systole. When a stethoscope is placed on the chest over different regions of the heart, there are four basic heart sounds that can be heard in a cardiac cycle. The most fundamental heart sounds are the first and second sounds, usually abbreviated as S1 and S2. S1 occurs at the beginning of isovolumetric ventricular contraction, when already closed mitral and tricuspid valves suddenly reach their elastic limit due to the rapid increase in pressure within the ventricles. S2 occurs at the beginning of diastole with the closure of the aortic and pulmonic valves [3]. Heart sound segmentation (HSS) is a necessary process before PCG analysis starts.

The proposed method of the current study for the PCG analysis involves Empirical Mode Decomposition (EMD). EMD algorithm is a fully adaptive, data-driven, and time-frequency technique for analysis and decomposition of non-stationary signals [4]. The signal is decomposed into a finite number of components, called Intrinsic Mode Functions (IMFs), corresponding to different frequencies and a residue. One of the tasks in which IMF extraction can prove to be useful is the denoising of the heart sound signals. Following these efforts, in this paper, the heart phonocardiogram is analyzed by employing EMD combined with kurtosis statistics to locate the occurrences of S1, S2 and maintain the IMFs where fundamental heart sounds are illustrated better. Furthermore, in a time window, we computed the bispectrum of S1 and S2 and extracted features which fed to a XGBoost classifier. Several statistical hypothesis tests were also carried out to find significant differences in the parameters of interest between conditions (normal vs. abnormal), and between heart sounds (S1 vs. S2).

The current approach employs the PhysioNet Computing in Cardiology Challenge 2016 (CinC 2016) database [5] and its main contribution lies in the development of a machine learning algorithm for the binary classification of heart condition into normal and abnormal.

The rest of the paper is structured as follows. Section II introduces the PCG dataset analyzed in this study and some experimental issues, while the mathematical background and the proposed methodology are described in Section III. Section IV presents and discusses the simulation, statistical analysis, and classification results obtained. Finally, the conclusions are provided in Section V.

II. EXPERIMENTAL DATASET

The database analyzed in the present study is composed of 3153 short-duration heart sound recordings that were sourced from seven contributing research groups and divided into six databases (with file names prefixed alphabetically, a through f). All recordings were resampled to 2000Hz and 16-bit resolution and provided in .wav format. Heart sound recordings were divided into two types: normal (healthy subjects) and abnormal (patients with a confirmed cardiac diagnosis) [5]. The patients were noted to suffer from a variety of illnesses, including mitral valve prolapse (MVP), aortic disease (AD), coronary artery disease (CAD), mitral regurgitation (MR), and aortic stenosis (AS). Specific classification for the abnormal PCG recordings was not provided on a case-by-case basis. The recordings were labeled as MPC (miscellaneous pathological conditions) when lacking detailed information about the pathological condition. In addition, reference annotations of the location of S1, systole, S2, and diastole states for each heartbeat were provided in the database. Reference annotations were generated by an algorithm based on a hidden semi-Markov model (HSMM) [7] and thereafter manually reviewed and corrected [5].

In this study, we considered 2867 heart sound recordings from the totality of the PCG recordings provided in the database; 286 PCG recordings were discarded due to poor signal quality and/or missing annotations. It should be noted that the PCG recordings database is unbalanced, i.e., the number of normal recordings is not equal to that of abnormal ones. Specifically, 2295 recordings were labeled as normal and 572 recordings were labeled as abnormal (AD: 13, AS: 8, Benign: 114, CAD: 219, MPC: 80, MR: 12, MVP: 126). Anthropometric characteristics of the patients were not provided in every PCG recording, and thus statistical analysis of anthropometric characteristics of the participants was not carried out.

III. METHODOLOGY

A. Pre-processing

The sampling frequency of the recordings in the database is 2000 Hz. To enhance the quality of the acquired PCG signals and reduce the high-frequency noise, a 3rd-order median filter and a 10th-order low-pass Butterworth filter were initially applied to the signals. In the present study, we were mainly interested in the analysis of the properties of the fundamental

heart sounds S1 and S2, whose major energy concentrations lie below 150 Hz [5]. Thus, the cutoff frequency of the low pass filter was set to 150 Hz.

B. Empirical Mode Decomposition (EMD)

The Empirical mode decomposition (EMD) is a time-frequency signal analysis technique that is applicable to non-stationary and non-linear signals. EMD does not require a priori basis functions for signal decomposition, as it is an adaptive and data-dependent method [4]. Thus, it is useful in the processing of signals acquired from biological systems, including the analysis of cardiac sounds [8]-[11]. EMD decomposes a signal $x(t)$ into a finite (and often small) number (M) of amplitude and frequency modulated (AM-FM) components called Intrinsic Mode Functions (IMFs), $IMF_1(t), IMF_2(t), \dots, IMF_M(t)$ and a residual, $r(t)$, via an iterative process. After decomposition, the original signal can be expressed as:

$$x(t) = \sum_{m=1}^M IMF_m(t) + r(t) \quad (1)$$

where $IMF_m(t)$ denotes the m^{th} IMF, $r(t)$ is the residue and M is the total number of IMFs extracted.

An IMF must fulfill two conditions: (i) the number of extrema and the number of zero crossings must be equal or differ at most by one within the whole data set, and (ii) the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero at any point [4]. The iterative procedure of extracting M-oscillatory modes (IMFs) from a given signal is called sifting. The sifting process is summarized as follows [4]:

- 1) Identify all the local maxima and minima from the signal $x(t)$.
- 2) Interpolate with cubic splines all local maxima and minima to produce an upper and a lower envelope, respectively.
- 3) Compute the mean, $m(t)$, of the upper and lower envelopes.
- 4) Calculate the residue between the signal $x(t)$ and the mean $m(t)$:

$$h(t) = x(t) - m(t)$$

- 5) Ideally, $h(t)$ satisfies the above-mentioned conditions of an IMF, if not, repeat the previous steps, but by treating the calculated residue $h(t)$ as input data. Repeat this sifting procedure k times, until $h_k(t)$ is an IMF.
- 6) The first IMF is $IMF_1(t) = h_k(t)$
- 7) Calculate the residue between the original signal $x(t)$ and the first IMF component of the data:

$$R_1(t) = x(t) - IMF_1(t)$$

- 8) The above-described sifting process is repeated but by taking as input the residue $R_1(t)$. The rest of the IMFs can be obtained from subsequent residuals $R_j(t)$ by repeating the same procedure.
- 9) The sifting process stops when no more IMFs can be extracted from the final residue ($R_M(t)$ becomes a monotonic function).

In the present study, all the IMFs were first extracted from the filtered PCG signals. Since EMD is a data-driven technique, each signal was decomposed into a different number of IMFs. The number of extracted IMF components ranges from 6 to 13. So, for uniformity in signal processing and comparability of results, we decomposed all the PCG recordings into six ($M=6$) IMF components, $IMF_1(t), IMF_2(t), \dots, IMF_6(t)$. The extracted IMFs are shown in Fig. 1 for both normal and abnormal heart sounds.

C. Higher-order Statistics (HOS)

Higher-order statistics and spectra analysis is an advanced signal processing technique that is applicable in the analysis of non-linear signals. In contrast to first- and second-order statistics (e.g., mean, variance, autocorrelation), HOS, i.e., moments and cumulants of third order and beyond, allow detecting deviations from Gaussianity and quantifying the non-linear characteristics in the signal. Biomedical signals are essentially non-linear, non-stationary, and non-Gaussian, and therefore HOS offer advantages in the analysis of such signals [12]-[14]. In general, there are three main motivations for using HOS in signal processing: (i) to extract information due to deviation from Gaussianity, as HOS of Gaussian signals are zero, (ii) to recover the phase as well as the magnitude response of a system, and (iii) to detect non-linear systems and quantify the non-linear properties of the signals [12].

The second-, third- and fourth-order cumulants of a fourth-order stationary zero-mean random process for zero-lag are defined as:

$$\gamma_2^x = c_2^x(0) = E\{x^2(k)\}(\text{variance}) \quad (2)$$

$$\gamma_3^x = c_3^x(0,0) = E\{x^3(k)\}(\text{skewness}) \quad (3)$$

$$\gamma_4^x = c_4^x(0,0,0) = \frac{E\{x^4(k)\}}{(\gamma_2^x)^2} - 3(\text{kurtosis}) \quad (4)$$

1) *Kurtosis*: In the absence of heart sounds, a PCG signal follows a Gaussian distribution, while when a heart sound occurs it does not [10]. If $x(k)$ is a random process following a Gaussian distribution, then kurtosis will approach zero. Since the aim of this study is to maintain the IMFs where fundamental heart sounds are illustrated better, a kurtosis-based criterion was implemented to detect deviations from Gaussianity.

The kurtosis of a stationary zero-mean process is defined by 4. Since kurtosis is estimated using a finite-length window, its value depends on the length of the signal. Thus, the kurtosis estimate exists in a confidence interval. Hence, a lower and an upper bound for the estimator has to be defined. Given a random variable X of mean μ and standard deviation σ , the kurtosis estimator can be framed between two values by utilizing the Chebyshev inequality. The Chebyshev inequality can be expressed as [10]:

$$P\left\{X \in \left(\mu - \frac{\sigma}{\sqrt{1-q}}, \mu + \frac{\sigma}{\sqrt{1-q}}\right)\right\} > q, \quad (5)$$

where $q = 1 - (\sigma^2/\epsilon^2)$, $\epsilon > 0$ and σ^2/ϵ^2 are sufficiently small. For a zero-mean, N-sample observation $x(k)$ from a random

variable X , the kurtosis estimate $\hat{\gamma}^4$ of $x(k)$ is given by 4. If X is assumed to be Gaussian, then it is proved that:

$$\hat{m} \approx \frac{6}{N} \text{ and } \hat{s}^2 \approx \frac{24}{N} \quad (6)$$

where \hat{m} and \hat{s} are the mean and the standard deviation estimates of $\hat{\gamma}^4$, respectively. Based on 5 and 6, it is easy to show that:

$$P\{\hat{\gamma}^4 \in I(N, q)\} > q, \quad (7)$$

where

$$I(N, q) = \left(\frac{6}{N} - \sqrt{\frac{24}{N(1-q)}}, \frac{6}{N} + \sqrt{\frac{24}{N(1-q)}} \right) \quad (8)$$

concluding that if the kurtosis estimate of an N-sample observation $x(k)$ lies outside $I(N, q)$, then it is natural to assume that $x(k)$ does not follow a Gaussian distribution. Probability q was initially set equal to 0.999. The higher q value, the stricter Gaussianity test is suggested [10].

D. Bispectrum

Bispectrum has been utilized extensively in the analysis of signals acquired from biological systems. Just as the power spectrum corresponds to the frequency domain representation of the second-order cumulants, the third-order spectrum or bispectrum is defined as the two-dimensional discrete Fourier transform of the third-order cumulants. In contrast to the power spectrum which suppresses the phase information of a process, bispectrum can detect and quantify quadratic phase coupling (QPC) between a signal's oscillatory components due to their non-linear interaction [13], [16]. Two frequency components are said to be quadratically phased coupled, if a third frequency component, whose frequency and phase are the sum of the frequencies and phases of the first two components, exists [15], [17]. Thus, Bispectrum represents the interactions between each pair of the signal's frequency components.

The bispectrum $B(f_1, f_2)$ of a stationary random process can be obtained directly from the discrete Fourier Transform of the third-order cumulants as follows:

$$B(f_1, f_2) = E[X(f_1)X(f_2)X^*(f_1 + f_2)] \quad (9)$$

Bispectrum presents symmetries that could further facilitate its calculation. Specifically, the bispectrum of a signal exhibits 12 symmetric regions in the plane (f_1, f_2) , while the calculation of the bispectrum in one of these regions is enough for the description of the bispectrum in the whole bi-frequency domain. Bispectrum is calculated in the triangular region $\Omega = \{\omega_2 \geq 0, \omega_1 \geq \omega_2, \omega_1 + \omega_2 \leq \pi\}$, which is named non-redundant region [16].

In the current study, we computed the bispectrum in an observation window in each IMF component to explore the properties of the fundamental heart sounds (S1 and S2). Specifically, from the annotations provided in the database on the location of S1 and S2 in time, we identified the location of the fundamental heart sounds and computed the bispectrum

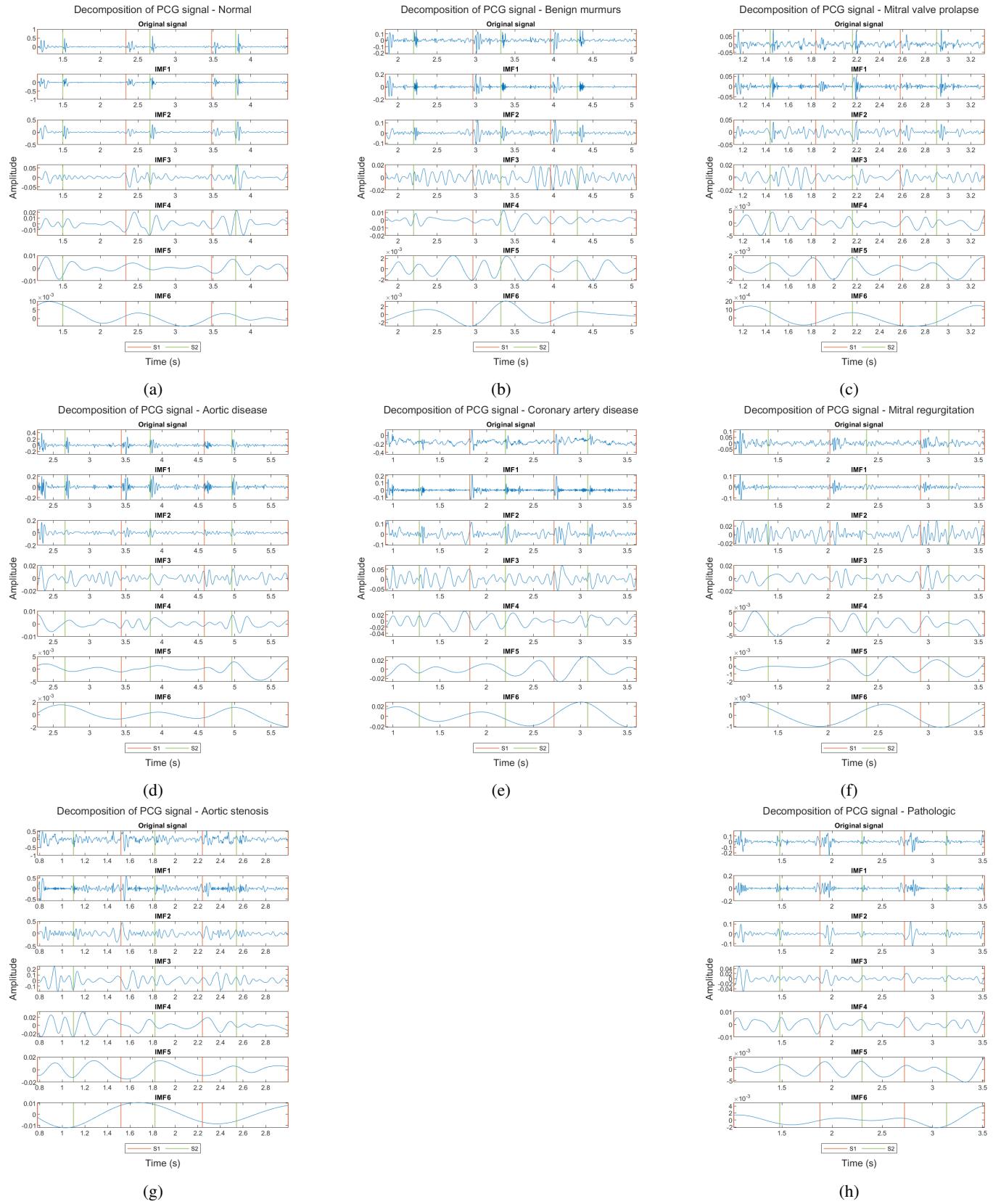


Fig. 1: Decomposition of PCG signals into IMFs using EMD technique. The location of S1 and S2 are shown as reference.

of S1 and S2 in segments of length 140ms (280 samples) and 110ms (220 samples) respectively, which agrees with the average duration of S1 and S2. The direct method was used to estimate the bispectrum using the bispecd function of the MATLAB Higher Order Statistics Toolbox (HOSA). The number of points used for each fast Fourier transform (NFFT) was 512. The bispectrum of each frame was extracted and averaged over all frames. Bispectra for signals from various systolic murmurs are presented in Fig. 4. Bispectral estimations were used to detect quadratic phase coupling and to extract quantitative features.

1) *Bispectral features extraction:* The plot of the bispectrum often enables us to differentiate pathological conditions and reveal a lot of information about the heart sound signal. However, it is difficult to utilize these plots for automatic classification of heart murmurs by computers. Thus, we need to quantify the characteristics of the bispectrum. Due to the symmetry of the bispectrum, it is enough to investigate the non-redundant region. In this paper, quantitative features based on the magnitude of the bispectrum were calculated to facilitate automatic machine learning. These bispectrum-based features are: mean (H1), standard deviation (H2), maximum (H3), and minimum (H4) value of the bispectrum, skewness (H5), kurtosis (H6), sum of logarithmic amplitudes of diagonal elements (H7), normalized bispectral entropy (BE1), and normalized bispectral squared entropy (BE2). The above-mentioned features were calculated in the triangular region $\Omega = \{f_2 \geq 0, f_1 \geq f_2, f_1 + f_2 \leq 150\}$.

The mean, standard deviation, maximum, and minimum values of the bispectrum, and sum of logarithmic amplitudes of diagonal elements are computed as follows:

$$H_1 = \frac{1}{L} \sum_{\Omega} |B(f_1, f_2)| \quad (10a)$$

$$H_2 = \sqrt{\frac{1}{L-1} \sum_{\Omega} (|B(f_1, f_2)| - H_1)^2} \quad (10b)$$

$$H_3 = \max \sum_{\Omega} |B(f_1, f_2)| \quad (10c)$$

$$H_4 = \min \sum_{\Omega} |B(f_1, f_2)| \quad (10d)$$

$$H_7 = \sum_{\Omega} \log(|B(f_k, f_k)|) \quad (10e)$$

where L is the total number of sampling points in the bispectral matrix, and Ω refers to the region of interest as described above.

Normalized bispectral entropy and normalized bispectral squared entropy are defined as [16]:

$$BE_1 = - \sum_n p_n \log p_n \quad (11a)$$

$$BE_2 = - \sum_n q_n \log q_n \quad (11b)$$

where

$$p_n = \frac{|B(f_1, f_2)|}{\sum_{\Omega} |B(f_1, f_2)|} \text{ and } q_n = \frac{|B(f_1, f_2)|^2}{\sum_{\Omega} |B(f_1, f_2)|^2}$$

and Ω is the region of interest. The above entropies were used to characterize signal variability and irregularity from

bispectrum. They have been previously used for automatic analysis of bio-signals from bispectral plots [18], [19], [20].

For each condition of the PCG recordings (normal and abnormal) and for each pathology, we computed the mean and standard deviation of the aforementioned features. The results are presented in Tables II-V. Two non-parametric statistical tests were also carried out to select the significant features and IMFs. The Wilcoxon rank-sum test (Mann-Whitney U test) was used to find the significant differences in the extracted features between normal and abnormal PCG recordings (independent samples) for each fundamental heart sound (S1 and S2), while the Wilcoxon signed-rank test was used to check the significance of the bispectral features between the fundamental heart sounds S1 and S2 (paired samples) for each class (normal and abnormal). Both statistical tests make no assumption about the distribution of data and are non-parametric alternatives to the Student's t-test. Wilcoxon rank-sum is a statistical test for two populations when samples are independent, while Wilcoxon signed-rank test is used to compare two populations when the observations are paired. Features with a p-value less than or equal to 0.05 were considered to be statistically significant.

E. Cepstral analysis

The Cepstral Analysis is a tool for detecting periodicity in the frequency spectrum and it is mainly used in pitch detection, radar and sonar applications, speech analysis, and diagnostics [21].

Cepstrum is an anagram of spectrum. The first paper on cepstrum analysis defined it as "the power spectrum of the logarithm of the power spectrum". The real cepstrum has a focus on periodic effects in the amplitudes of the spectrum. It is written as:

$$C_r = \mathcal{F}^{-1}\{\log(|\mathcal{F}\{f(t)\}|)\} \quad (12)$$

and the complex cepstrum is written as:

$$C_r = \mathcal{F}^{-1}\{\log(\mathcal{F}\{f(t)\})\} \quad (13)$$

The real cepstrum is used to export the pitch of the signals. The peak with the maximum value contains the information of the periodicity.

A very powerful processing tool is an exponential "lifter" (window) applied to the cepstrum, which is shown to extract the modal part of the response (with a small extra damping of each mode corresponding to the window). This can then be used to repress or enhance the modal information in the response according to the application [21].

It is assumed that the signal segments are described by:

$$y(t) = h(t) * x(t)$$

where $y(t)$ is the signal segment, $h(t)$ is the impulse response and $x(t)$ is the input. The "lifter" is used to export and remove the impulse response from the cepstrum of the signal. The result from subtraction constitutes the input $x(t)$. The Normalized Root Mean Squared Error (NRMSE) is used in

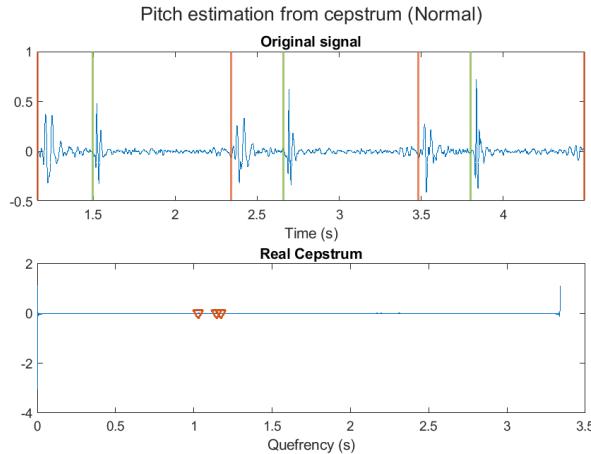


Fig. 2: Pitch estimation from Real Cepstrum for a normal PCG recording.

order to find the part of cepstrum that contains the impulse response. The NRMSE is described as:

$$NRMSE = \frac{RMSE}{\max(y(k)) - \min(y(k))},$$

where $RMSE = \sqrt{\frac{\sum_{k=1}^N (x(k)-y(k))^2}{N}}$. The minimum value of NRMSE ensures that the approach of the impulse response is acceptable. The reconstructed signal is described as the convolution of the input and the impulse response. The results are shown in Fig. 7.

F. Classification

To deal with the tabular data after the feature extraction performed in previous subsections, we opted for XGBoost to perform the binary classification task. XGBoost is a decision-tree-based ensemble Machine Learning algorithm that uses gradient boosting.

The given dataset is quite unbalanced, with most sound recordings being of properly functioning hearts, which raises a problem in model evaluation, and overfitting. This is because the frequently used accuracy metric is misleading, since the model can get away with classifying every sample as normal and still perform acceptably. To mitigate this problem we used metrics such as recall, precision, f1 score and balanced accuracy to evaluate our model predictions in the validation dataset. Intuitively, precision is the ability of the classifier not to label as positive a sample that is negative, and recall is the ability of the classifier to find all the positive samples.

To create a train and validation dataset and avoid data leakage, stratified dataset splitting was used to evenly distribute classes across the two, as well as during cross validation.

Furthermore, the dataset itself has to be balanced to avoid overfitting. This can be achieved with oversampling or undersampling techniques. In our experiments we saw that SMO-TEENN, a method combining SMOTE (Synthetic Minority Over-sampling Technique) to perform oversampling and then uses edited nearest-neighbors to clean the remaining sampled space. The second step helps because simple SMOTE can

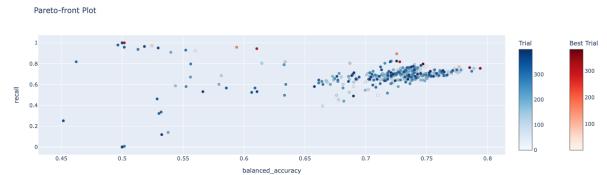


Fig. 3: Hyperparameter tuning

generate noisy samples by interpolating new points between marginal outliers and inliers. This issue can be solved by cleaning the space resulting from over-sampling.

To further improve model performance, thorough hyperparameter tuning was conducted over a variety of model parameters such as boosting methods, subsampling, model depth, breadth or depth prioritization. An image of the tuning results based on two selected metrics can be seen in Fig. 3

IV. RESULTS AND DISCUSSION

A. Kurtosis-based IMFs selection

Excerpts from the experimental results derived from the application of EMD on the dataset are depicted in Fig. 1, both for normal cases (Fig. 1a) and for cases where cardiac murmurs are present (Fig. 1b-1h). The locations of S1 and S2 provided in the database are shown in the signals to facilitate the visual evaluation of the results. We can clearly observe that the signal's information is handed out in different ways in the IMFs. Specifically, relevant information regarding S1, and S2 can be seen in the first three or four IMFs, while the last two IMFs do not depict useful information about the fundamental heart sounds in the majority of the signals. Moreover, more oscillations are present in the first two IMFs in the pathologic cases than in the normal cases, while the amplitude of the first IMF of the normal signal (Fig. 1a) remains almost constant between the peaks of S1 and S2.

The results obtained from the visual inspection of the figures are confirmed by the kurtosis-based Gaussianity test applied. The percentage of signals in which each IMF component depicts useful information about S1 and S2 are presented in Table I. The first four IMFs that contain the S1 and S2 important components were selected and thus, further analysis was performed only on them.

TABLE I: Percentage of signals in which each IMF depicts useful information about S1 and S2

IMF	Percentage (%)
1	100.00
2	98.22
3	88.70
4	77.64
5	57.10
6	38.47

B. Analysis of PCG signals using bispectral features

Fig. 4 shows examples of bispectrum for signals from various systolic murmurs. The colors indicate the relative

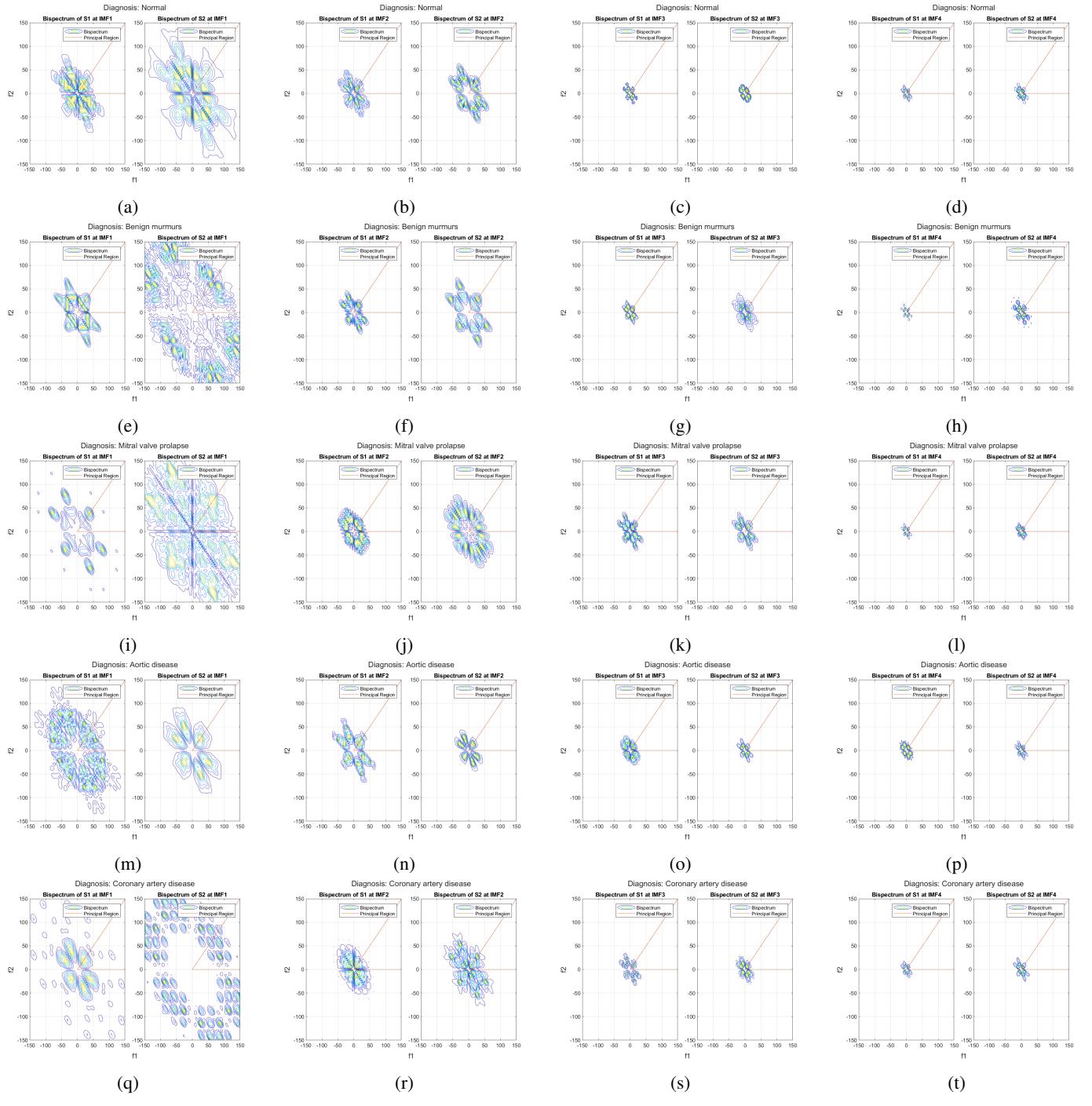


Fig. 4: Bispectra of S1 and S2 computed at each IMF.

changes in the bispectrum amplitude. Blue and yellow colors indicate the highest and lowest amplitude, respectively. The bispectrum represents the degree of interaction between the frequencies f_1 and f_2 . This interaction can be related to the non-linearities present in the signal and indicate the presence of QPC. According to Fig. 4, we can observe that, in all the cases, as the order of the IMF component increases, the bispectrum plot is concentrated in a narrower bi-frequency range as expected. Differences in the bispectrum patterns and the location of the pronounced peaks are presented both

between S1 and S2 and among different pathologies. Though, the bispectral plot computed on the fourth IMF does not provide useful information for the visual discrimination of heart sounds. According to Fig. 4a, the bispectrum of S1 exhibits a prominent peak at around 15 to 20 Hz on the diagonal $f_1 = f_2$, while the bispectrum of S2 exhibits a prominent peak at around $(f_1, f_2) = (35, 10)$ Hz. Using the normal PCG signal as a reference, it can be observed that the bispectral peaks are shifted towards the high frequencies in the case of Benign murmurs (Fig. 4e), Mitral valve prolapse

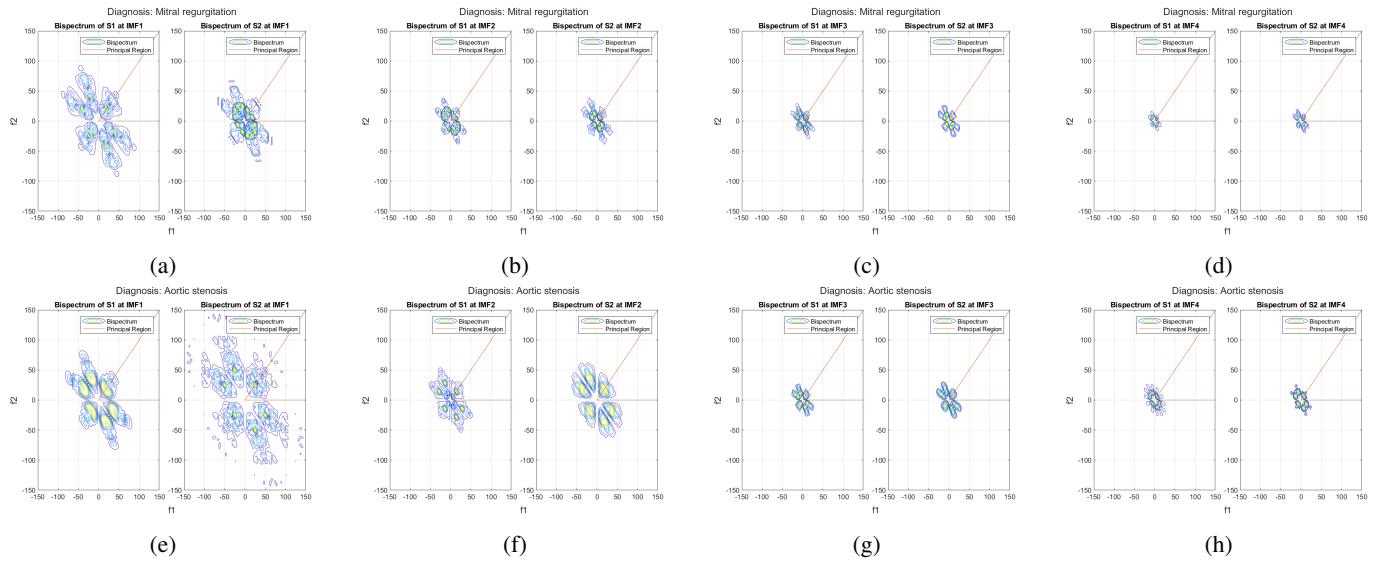


Fig. 5: Bispectra of S1 and S2 computed at each IMF. (Continued from Fig. 4).

(Fig. 4i), Aortic disease (Fig. 4m), and Coronary artery disease (Fig. 4q (S2)). Moreover, for the pathological signals in the case of Mitral regurgitation and Aortic stenosis, the couplings tended to be at lower frequencies compared to normal case.

The spreads of the bispectrum magnitude of the heart sounds signals present also differences. In the case of Benign murmurs (Fig. 4e (S1)), Aortic disease (Fig. 4m (S2)), Mitral regurgitation (Fig. 5a), and Aortic stenosis (Fig. 5e (S2)) the bispectral spread is relatively lower than that of the normal case, while in the case of Benign murmurs (Fig. 4e (S2)), Mitral valve prolapse (Fig. 4i (S2)), Aortic disease (Fig. 4m (S1)), and Coronary artery disease (Fig. 4q) the bispectral spread is higher than that of the normal case.

Similar conclusions can be drawn for the bispectral graphs of the remaining IMF components. Generally, as the order of the IMF increases, the prominent peaks of the bispectrum tend to cluster in the same bi-frequency region.

Tables II, III, IV and V show the results of the metrics evaluated on the bispectrum of each IMF and heart sound for normal and abnormal PCG recordings, respectively. Concerning the HOS (skewness and kurtosis) of the magnitude of the bispectrum, we can observe that these metrics were greater in the case of abnormal signals than that of the normal signals for both S1 and S2 heart sounds for all IMFs and the HOS of the bispectral magnitude increase as the order of the IMF increases. Concerning the bispectral entropies, in the case of the normal signals the bispectral entropy was higher than that of the abnormal signals for both S1 and S2 for the first three IMFs and the bispectral squared entropy was higher in the normal signals than that of the abnormal signals for both S1 and S2 for all IMFs. The entropy quantifies the regularity and complexity of time series data and thus, it should have a higher value in the case of normal cases and a lower value for abnormal cases, indicating smaller variability (or inherent periodicity) in the beat-to-beat interval [22]. The lower entropy was in the fourth IMF in all the cases. Concerning the sum of logarithmic amplitudes of diagonal elements, this value is

lower in the pathologic signals than that of the normal signals in all the cases.

The statistical analysis carried out showed that: (i) statistically significant differences were found in S1 for normal vs. abnormal PCG recordings in each IMF, (ii) statistically significant differences were found in S2 for normal vs. abnormal PCG recordings in each IMF, except in minimum value in IMF 2-4, (iii) regardless of the condition of the PCG recording (normal, abnormal), statistically significant differences were found in S1 vs. S2 in each IMF, except in sum of logarithmic amplitudes of diagonal element in IMF 3 and 4 of the abnormal signals, where no statistically significant difference was found in S1 vs. S2.

C. Analysis of Cepstrum

The mean value of the pitch for each case and the standard deviation are shown in Table VI. In all heart conditions, the mean value of pitch lies between 0.7 and 0.9 seconds and the standard deviation lies in the interval [0.16, 0.29] seconds. It turns out that, on average, signals with Aortic Disease and Mitral Valve Prolapse have a higher pitch (0.93865 seconds and 0.90225 seconds) and signals with Aortic Stenosis have a lower pitch (0.72275 seconds). Normal signals have a mean value of pitch of almost 0.8 seconds.

As shown in Fig. 7, the reconstructed signal approaches the original signal. The impulse response does not present similarities for signals with the same disease.

D. Classification Results

The classifier performed acceptably in different metrics when tested in the evaluation dataset. Notably as expected, the normal signal detection outperforms the abnormal one, but taking recall into account, the classifier performs well. Specifically, recall is an important metric, since it shows that we do not label abnormal signals as normal as often while sacrificing some precision.

TABLE II: Mean \pm standard deviation of the bispectral features of S1, computed at each IMF for normal PCG recordings

Feature Name	IMF1	IMF2	IMF3	IMF4
Mean	4.91e-06 \pm 8.51e-06	1.32e-06 \pm 3.79e-06	1.55e-07 \pm 1.04e-06	9.59e-09 \pm 5.42e-08
Standard Deviation	1.00e-05 \pm 1.98e-05	5.99e-06 \pm 1.90e-05	1.38e-06 \pm 8.94e-06	9.50e-08 \pm 5.37e-07
Minimum	1.27e-52 \pm 8.82e-52	1.97e-52 \pm 1.70e-51	1.24e-52 \pm 9.99e-52	1.04e-52 \pm 8.73e-52
Maximum	7.18e-05 \pm 1.64e-04	6.06e-05 \pm 2.09e-04	2.19e-05 \pm 1.44e-04	1.74e-06 \pm 9.96e-06
Skewness	2.69 \pm 1.26	5.37 \pm 1.74	9.65 \pm 2.59	15.18 \pm 2.64
Kurtosis	10.25 \pm 12.47	36.26 \pm 27.53	108.85 \pm 60.84	249.8 \pm 77.34
Bispectral Entropy	5.01 \pm 0.44	3.72 \pm 0.42	2.71 \pm 0.43	2.28 \pm 0.47
Bispectral Squared Entropy	4.15 \pm 0.62	2.81 \pm 0.52	1.62 \pm 0.51	0.69 \pm 0.39
Sum of log amplitudes	-371.13 \pm 40.66	-426.25 \pm 40.99	-500.53 \pm 48.74	-551.21 \pm 49.83

TABLE III: Mean \pm standard deviation of the bispectral features of S2, computed at each IMF for normal PCG recordings

Feature Name	IMF1	IMF2	IMF3	IMF4
Mean	3.36e-06 \pm 8.47e-06	6.49e-07 \pm 1.61e-06	1.49e-07 \pm 6.54e-07	1.05e-08 \pm 8.99e-08
Standard Deviation	4.93e-06 \pm 1.29e-05	2.59e-06 \pm 8.18e-06	1.15e-06 \pm 5.33e-06	8.31e-08 \pm 7.45e-07
Minimum	8.36e-53 \pm 9.35e-52	7.77e-53 \pm 6.38e-52	1.24e-52 \pm 1.00e-51	8.34e-53 \pm 1.03e-51
Maximum	2.97e-06 \pm 8.13e-05	2.51e-05 \pm 9.68e-05	1.72e-05 \pm 8.33e-05	1.39e-06 \pm 1.27e-05
Skewness	2.06 \pm 1.25	4.25 \pm 1.73	8.34 \pm 2.10	12.10 \pm 2.07
Kurtosis	6.36 \pm 10.61	23.15 \pm 24.47	81.09 \pm 42.52	162.98 \pm 51.18
Bispectral Entropy	5.25 \pm 0.40	4.07 \pm 0.42	3.05 \pm 0.38	2.83 \pm 0.43
Bispectral Squared Entropy	4.51 \pm 0.62	3.23 \pm 0.56	1.91 \pm 0.47	1.17 \pm 0.38
Sum of log amplitudes	-386.34 \pm 48.26	-433.51 \pm 44.53	-508.49 \pm 54.24	-554.24 \pm 52.77

TABLE IV: Mean \pm standard deviation of the bispectral features of S1, computed at each IMF for abnormal PCG recordings

Feature Name	IMF1	IMF2	IMF3	IMF4
Mean	4.24e-06 \pm 1.14e-05	1.28e-06 \pm 3.46e-06	1.64e-07 \pm 6.75e-07	1.19e-08 \pm 9.88e-08
Standard Deviation	1.02e-05 \pm 2.51e-05	6.59e-06 \pm 1.94e-05	1.43e-06 \pm 6.44e-06	1.17e-07 \pm 9.64e-07
Minimum	1.39e-052 \pm 6.45e-52	1.27e-52 \pm 7.70e-52	3.63e-52 \pm 5.11e-51	1.42e-52 \pm 1.23e-51
Maximum	8.10e-05 \pm 2.25e-04	7.02e-05 \pm 2.14e-04	2.18e-05 \pm 1.05e-04	2.15e-05 \pm 1.78e-05
Skewness	3.21 \pm 1.27	5.93 \pm 1.74	10.63 \pm 2.74	15.50 \pm 2.79
Kurtosis	13.80 \pm 13.68	42.93 \pm 27.59	129.80 \pm 66.89	262.13 \pm 80.99
Bispectral Entropy	4.79 \pm 0.48	3.54 \pm 0.46	2.59 \pm 0.47	2.42 \pm 0.47
Bispectral Squared Entropy	3.85 \pm 0.62	2.61 \pm 0.52	1.42 \pm 0.52	0.65 \pm 0.42
Sum of log amplitudes	-392.88 \pm 51.49	-448.67 \pm 54.63	-516.80 \pm 61.60	-571.26 \pm 64.12

TABLE V: Mean \pm standard deviation of the bispectral features of S2, computed at each IMF for abnormal PCG recordings

Feature Name	IMF1	IMF2	IMF3	IMF4
Mean	1.91e-06 \pm 7.12e-06	5.21e-07 \pm 2.29e-06	1.63e-07 \pm 8.60e-07	7.90e-09 \pm 4.82e-08
Standard Deviation	3.73e-06 \pm 1.20e-05	2.30e-06 \pm 8.79e-06	1.24e-06 \pm 7.22e-06	5.75e-08 \pm 3.56e-07
Minimum	1.79e-52 \pm 3.02e-51	8.12e-53 \pm 5.56e-52	1.08e-52 \pm 9.67e-52	1.13e-52 \pm 1.27e-51
Maximum	2.40e-05 \pm 7.86e-05	2.44e-05 \pm 9.58e-05	1.82e-05 \pm 1.16e-04	9.11e-07 \pm 5.77e-06
Skewness	2.53 \pm 1.27	5.16 \pm 1.99	9.33 \pm 1.97	12.58 \pm 1.99
Kurtosis	8.91 \pm 10.12	33.63 \pm 29.96	98.91 \pm 40.03	178.17 \pm 51.08
Bispectral Entropy	5.04 \pm 0.47	3.82 \pm 0.48	2.91 \pm 0.42	2.92 \pm 0.38
Bispectral Squared Entropy	4.22 \pm 0.68	2.89 \pm 0.62	1.67 \pm 0.44	1.11 \pm 0.35
Sum of log amplitudes	-407.16 \pm 48.44	-453.23 \pm 49.21	-515.57 \pm 59.55	-572.05 \pm 61.78

TABLE VI: Mean \pm standard deviation of pitch and inter-beat interval (IBI) computed from cepstrum and annotations, respectively

Class	Pitch (s)	IBI (s)
Normal	0.80084 \pm 0.2084	0.80626 \pm 0.18082
Abnormal	0.85382 \pm 0.27785	0.87419 \pm 0.23014
Diagnosis	Pitch (s)	IBI (s)
AD	0.93865 \pm 0.26944	0.96087 \pm 0.19764
AS	0.72275 \pm 0.18746	0.78946 \pm 0.13973
Benign	0.8787 \pm 0.20599	0.90261 \pm 0.15115
CAD	0.84299 \pm 0.28514	0.85833 \pm 0.21648
MR	0.7855 \pm 0.1651	0.76654 \pm 0.14431
MVP	0.90225 \pm 0.22387	0.92924 \pm 0.16285

V. CONCLUSION

The importance of heart sounds assessment arises from the need to seek information on the severity of cardiovascular disorders. Therefore, an accurate cardiac state assessment approach is required to identify early symptoms of the heart diseases. In the present work, the analysis of the fundamental heart sounds (S1 and S2) from PCG recordings through EMD, a time-frequency decomposition technique, and Higher-Order Statistics is proposed. The method was tested on an experimental database, recorded in multiple clinical environments, including normal heart sound samples, samples of aortic stenosis, samples of mitral regurgitation and other diseases. The EMD-Bispectrum-XGBoost scheme employed in this study proved to be a promising tool for the classification of PCG

TABLE VII: Classification Results

	precision	recall	f1-score	support	balanced accuracy
Normal (0)	0.96	0.79	0.86	574	
Abnormal (1)	0.52	0.85	0.63	143	
				717	0.82
macro avg	0.73	0.82	0.75	717	
weighted avg	0.86	0.80	0.82	717	

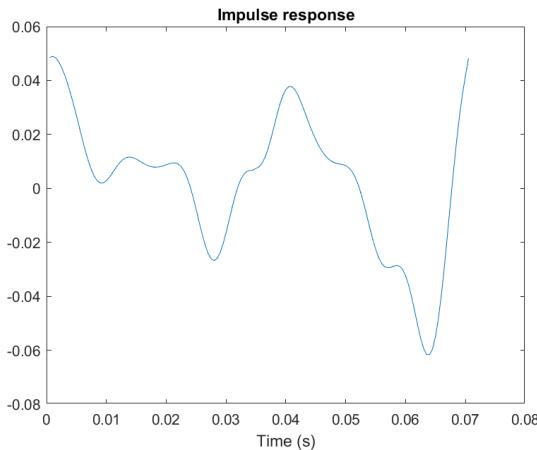


Fig. 6: Impulse Response estimated from Complex Cepstrum.

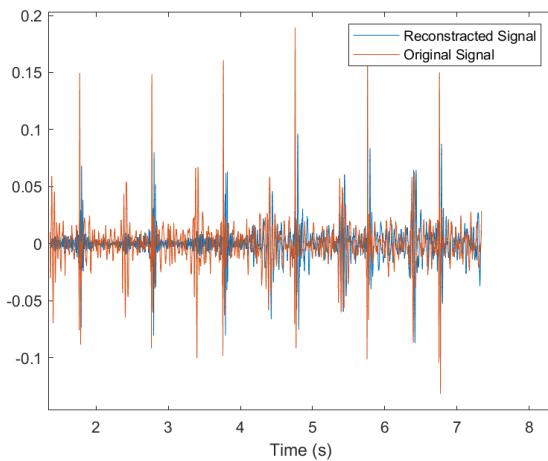


Fig. 7: Original and Reconstructed signal.

signals into normal and abnormal. The easier acquisition of PCG signals as compared to ECG signals makes the proposed method suitable for rural areas, developing economies, home-care, and non-clinical environments.

In this work, the EMD technique and kurtosis features were used to locate the IMFs that contain relevant information about the fundamental heart sounds. The statistical results showed that FHS were depicted stronger in the first four IMFs. The results were confirmed by visual inspection of the IMFs. Moreover, the bispectrum of S1 and S2 revealed the presence of QPC in the PCG signals and the different QPC relations in each heart condition and FHS. The proposed method of heart condition (normal and abnormal) classification by using features derived from the magnitude of the bispectrum and XGBoost classifier has shown reliable performance and shed

the light upon further investigation of the bispectral features as an effective tool for heart disorders detection. A detailed investigation and analysis of the physiology of each heart disease may enable us to decode the hidden patterns of QPC relations in the bispectrum.

Future work will focus on the usage of cepstral analysis in heart sound classification, utilizing the pitch and the impulse response of each signal. Deep learning models like convolutional neural networks or attention-based networks should offer superior performance and could be experimented with.

AVAILABILITY OF DATA AND MATERIALS

The dataset analyzed during this study is available in the PhysioNet repository: <https://physionet.org/content/challenge-2016>. The source code of the implementations used to compute the presented results can be obtained from the public repository: <https://github.com/sfoteini/phonocardiogram-heart-sound-analysis>.

REFERENCES

- [1] World Health Organization, [https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))
- [2] Dalcame, <http://www.dalcame.com/fono.html#Yz8EvS81tQK>
- [3] Ashraf, Hassan & Waris, Muhammad & Gilani, Syed & Tariq, Umair & Alquhayz, Hani. (2021). Threshold Parameters Selection for Empirical Mode Decomposition-Based EMG Signal Denoising.
- [4] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih et al., *The empirical mode decomposition and the hilbert spectrum for nonlinear and nonstationary time series analysis*, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, vol. 454, no. 2, pp. 903–995, 1998. 4
- [5] C. Liu, D. Springer, Q. Li, B. Moody, RA. Juan, FJ. Chorro, F. Castells, JM. Roig, I. Silva, AE. Johnson, Z. Syed, SE. Schmidt, CD. Papadaniil, L. Hadjileontiadis, H. Naseri, A. Moukadem, A. Dieterlen, C. Brandt, H. Tang, M. Samieinasab, MR. Samieinasab, R. Sameni, RG. Mark and GD. Clifford, *An open access database for the evaluation of heart sound algorithms*. Physiol Meas. 2016 Dec;37(12):2181-2213
- [6] Goldberger, A., L. Amaral, L. Glass, J. Hausdorff, P. C. Ivanov, R. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley (2000). *PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals*. Circulation [Online]. 101 (23), pp. e215–e220.
- [7] D. B. Springer, L. Tarassenko and G. D. Clifford, *Logistic Regression-HSMM-Based Heart Sound Segmentation*, in IEEE Transactions on Biomedical Engineering, vol. 63, no. 4, pp. 822-832, April 2016, doi: 10.1109/TBME.2015.2475278.
- [8] A. Cheema, and M. Singh, *Psychological stress detection using phonocardiography signal: An empirical mode decomposition approach*, Biomedical Signal Processing and Control, Volume 49, 2019, Pages 493-505, ISSN 1746-8094, <https://doi.org/10.1016/j.bspc.2018.12.028>.
- [9] A. Cheema, M. Singh, *An application of phonocardiography signals for psychological stress detection using non-linear entropy based features in empirical mode decomposition domain*, Applied Soft Computing, Volume 77, 2019, Pages 24-33, ISSN 1568-4946, <https://doi.org/10.1016/j.asoc.2019.01.006>.
- [10] C. Papadaniil, and L. Hadjileontiadis (2014). *Efficient Heart Sound Segmentation and Extraction Using Ensemble Empirical Mode Decomposition and Kurtosis Features*. IEEE journal of biomedical and health informatics. 18. 1138-52. 10.1109/JBHI.2013.2294399.

- [11] M. Altuve, L. Suárez Mantilla, and J. Ardila (2020). *Fundamental heart sounds analysis using improved complete ensemble EMD with adaptive noise*. Biocybernetics and Biomedical Engineering. 40. 10.1016/j.bbe.2019.12.007.
- [12] C. L. Nikias, *Higher-order spectral analysis*, Proceedings of the 15th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 1993, pp. 319-319, doi: 10.1109/IEMBS.1993.978564.
- [13] C. L. Nikias, and A. P. Petropulu, *Higher-Order Spectral Analysis: A Nonlinear Signal Processing Framework*, Prentice-Hall, Inc., 1993
- [14] C. L. Nikias and J. M. Mendel, *Signal processing with higher-order spectra*, IEEE Signal processing magazine, vol. 10, no. 3, pp. 10-37, 1993.
- [15] Kiciński, W., Szczepański, A. *Quadratic Phase Coupling Phenomenon and Its Properties*. Hydroacoustics 2004, 7, 97–106.
- [16] Chua K.C., Chandran V., Acharya U.R., Lim C.M. *Application of higher order statistics/spectra in biomedical signals—A review*. Med. Eng. Phys. 2010;32:679–689. doi: 10.1016/j.medengphy.2010.04.009.
- [17] P. Venkatakrishnan, R. Sukarnesh, and S. Sangeetha, *Detection of quadratic phase coupling from human eeg signals using higher order statistics and spectra*, Signal, Image and Video Processing, vol. 5, no. 2, pp. 217–229, 2011.
- [18] Berriah, Sid Ahmed et al. *Severity cardiac analysis using the Higher-order spectra*. Appl. Math. Comput. 409 (2021): 126389.
- [19] Chua, K. C., Chandran, V., Rajendra Acharya, U. & Lim, C. M. *Analysis of epileptic EEG signals using higher order spectra*, Journal of Medical Engineering Technology 33, 42–50, <https://doi.org/10.1080/03091900701559408> (2009).
- [20] Bou Assi E, Gagliano L, Rihana S, Nguyen D. K., Sawan M., *Bispectrum Features and Multilayer Perceptron Classifier to Enhance Seizure Prediction*, Sci Rep. 2018;8:15491. <https://doi.org/10.1038/s41598-018-33969-9>
- [21] Robert B. Randall, *A history of cepstrum analysis and its application to mechanical problems*. December 2017
- [22] Chua, Kuang & Chandran, V. & Acharya, U Rajendra & Lim, Choo. (2009). *Cardiac state diagnosis using higher order spectra of heart rate variability*. Journal of medical engineering & technology. 32. 145-55. 10.1080/03091900601050862



Stavros Filosidis is currently an undergraduate student in the Department of Electrical and Computer Engineering at the Aristotle University of Thessaloniki (AUTH). His research interests include Deep Learning, Computer Vision, Distributed Computing and Optimization. Has previously worked for Amazon (AWS) and Bloomberg.



Chrysoula Psyrouki is currently an undergraduate student in the Department of Electrical and Computer Engineering at the Aristotle University of Thessaloniki (AUTH), Thessaloniki, Greece, affiliated with the Telecommunications Laboratory. Her research interests lie in the area of Digital Image Processing, Data Analysis and Robotics.



Foteini Savvidou is currently an undergraduate student in the Department of Electrical and Computer Engineering at the Aristotle University of Thessaloniki (AUTH), Thessaloniki, Greece, affiliated with the Telecommunications Laboratory. Her research interests lie in the area of wireless communications, Internet of Things, cloud and edge computing, and signal processing techniques for biomedical applications. In addition to her academic activities, she is actively involved in the Microsoft developer communities, co-authored a Microsoft course, published a number of technical articles, and participated in various national and international technology conferences.



Konstantinos Barmpounakis is currently an undergraduate student in the Department of Electrical and Computer Engineering at the Aristotle University of Thessaloniki (AUTH), Thessaloniki, Greece, affiliated with the Electronics and Computer Laboratories. His research interests lie in the area of Data Analysis, Machine Learning and Signal Processing.



Stavros Tsimpoukis is currently an undergraduate student in the Department of Electrical and Computer Engineering at the Aristotle University of Thessaloniki (AUTH), Thessaloniki, Greece, affiliated with the Telecommunications Laboratory. His research interests lie in the area of wireless communications, Machine Learning and Signal Processing.