PhonoTrack

Main Objective

1. Building a Reference Dataset.
2. Giving a Baseline Solution

Data Type: Phonocardiogram Signal

The Phonocardiogram (PCG) is a powerful tool in the field of medicine, providing a noninvasive and valuable means of monitoring vital signals related to heart health. It is presented as a plot of high-fidelity recordings of the sounds and murmurs generated by the heart.

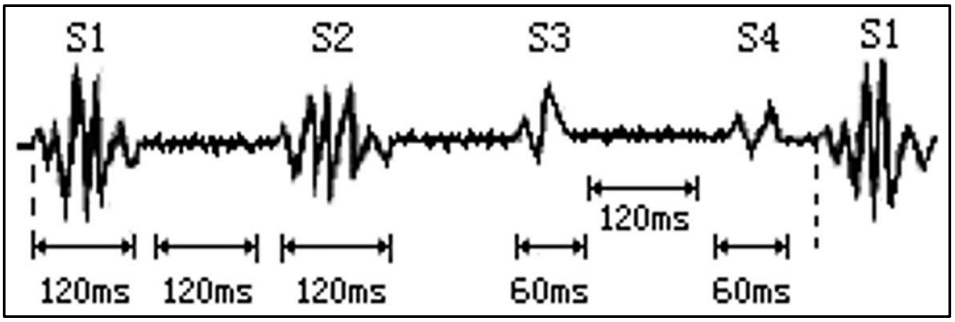


Figure 01: Different components of PCG Signal [1]

It is one of the most traditional biological signals and its measurement requires a vibration or sound converter into an electrical signal; microphones, pressure transducers or accelerometers can be placed for this purpose on the surface of the patient's chest. The individual heart sounds are best heard at certain locations on the chest, so the sensors should be placed in several auscultator areas: mitral, aortic, pulmonary and tricuspid. A frequency range of the sensor can be from 0.05 to 1000 Hz but in fact, over 95 % of the acoustic energy of a PCG signal lies under 75 Hz [10].

Thus, this signal captures the acoustic vibrations resulting from heart auscultation, representing the unique sounds produced during the cardiac cycle. These sounds are influenced by various physiological factors, including aortic pressure, atrial pressure, ventricular pressure, ventricular volume, and the electrocardiogram. Of particular interest are two distinct sounds known as S1 and S2:

1. S1: This sound is produced by the closure of the mitral and tricuspid valves at the beginning of systole, the phase of the cardiac cycle when the heart contracts to pump blood.
2. S2: This sound is generated by the closure of the aortic and pulmonary valves at the end of the systole, marking the completion of the heart's contraction.

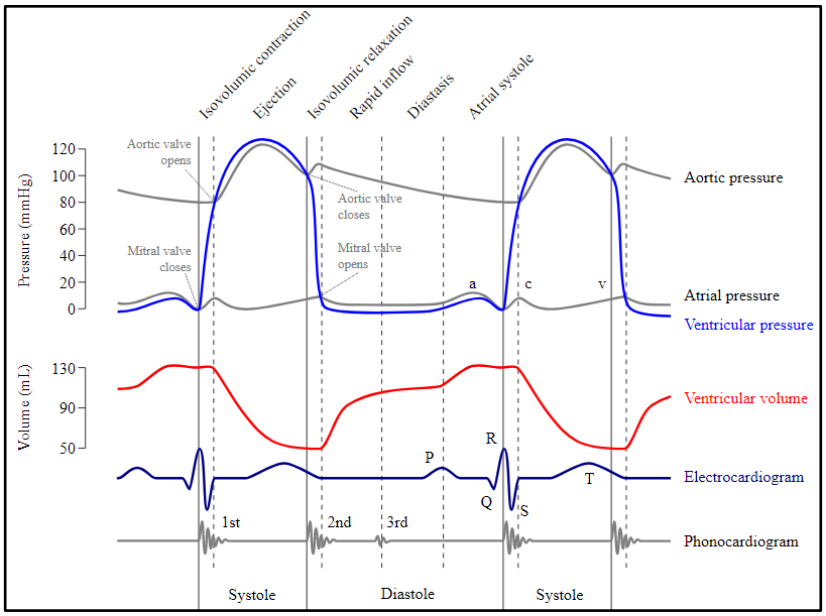


Figure 02: Origin of phonocardiogram Signal

The timing and quality of these sounds hold crucial information about the heart's function and can serve as indicators of potential issues such as murmurs, valve disorders, congenital heart defects, and other structural abnormalities. Additionally, the PCG can be utilized as a sign of heart failure.

Beyond its diagnostic capabilities, the PCG enables the extraction of essential vital signs in a reliable manner. Notably, heart rate and blood pressure can be accurately derived from the PCG data, as this signal is relatively cleaner and less prone to artifacts.

The noninvasive nature of phonocardiography makes it a highly valuable tool in healthcare settings, facilitating the continuous monitoring of heart health and early detection of potential cardiac abnormalities. By analyzing the PCG signal, medical professionals can gain valuable insights into a patient's cardiovascular status, enabling timely interventions and personalized treatment plans to improve overall cardiac well-being. As medical technology advances, the phonocardiogram continues to play a crucial role in enhancing patient care and promoting heart health worldwide.

Relation between PCG and other Non-invasive signals

Different Non-invasive signals are used to measure the cardiac condition of the human body. Some of the most used signals are ECG (Electrocardiogram), PCG (Phonocardiogram), and PPG (Photo plethysmography). As all of these signals monitor cardiac activity there are many relations in between these signals.

a) ECG: ECG measures the electrical activity of the heart and records the depolarization and repolarization of the heart chambers. It is typically measured using electrodes placed on

the skin, most commonly on the chest, limbs, or 12-lead ECG systems. ECG provides information about heart rate, heart rhythm, and can help diagnose various cardiac conditions.

b) PPG: PPG measures changes in blood volume in peripheral blood vessels, such as the fingertip, earlobe, or wrist. It is commonly measured using pulse oximeters and wearable fitness trackers. PPG provides information about heart rate, oxygen saturation (SpO2), and peripheral blood flow.

c) PCG: PCG records the acoustic vibrations produced during heart auscultation. It is measured using specialized microphones or piezoelectric sensors placed on the chest. PCG provides information about heart sounds (S1 and S2), and murmurs, and can assist in diagnosing heart valve disorders or other cardiac abnormalities.

As PCG and ECG both are collected directly from the chest, they very closely resemble cardiac events. On the other hand PPG is collected from other body parts, that's why it's synchronized but there is a delay element in between PPG and other 2 signals.

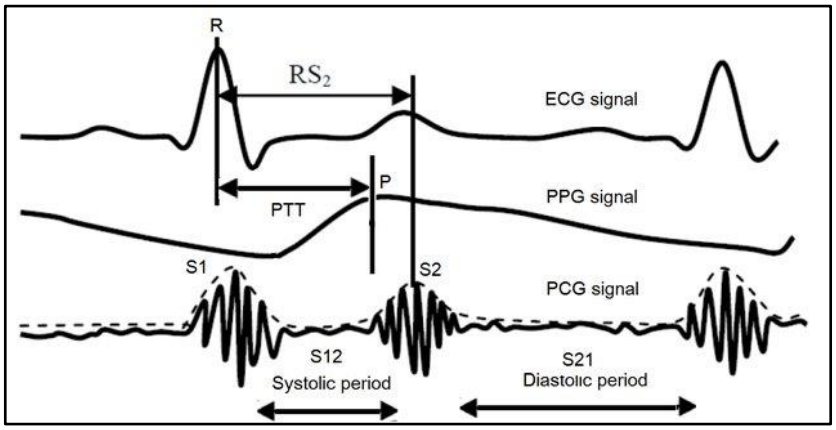


Figure 03: Simultaneous presentation of the three physiological signals ECG PPG and PCG [2]

There is also a slight delay in PCG from ECG due to electrical signals propagating faster than sound signals. S1 starts at 0.02~0.04 seconds lagging behind the beginning of the QRS wave in the electrocardiogram, lasting for about 0.08-0.15 seconds. It is caused by the flowing of blood into the aorta, when both the mitral valve and the tricuspid valves close, and ventricular systole occurs.S2 happens at the beginning of diastole. Its frequency is comparatively high and its duration (about 0.07~0.12 seconds) is shorter than S1. S3 shows low-frequency and small- amplitude. It is 0.12~0.20 seconds away from S2 behind the T wave in the electrocardiogram and lasts for 0.05~0.06 seconds. S3 is caused by the vibration of the ventricular wall. S4 lags behind P wave for 0.15~0.18 seconds, and S4 generally has small- amplitude and is caused by the casthe e that blood stream flows rapidly into the ventricle and vibrates the ventricular wall [1].

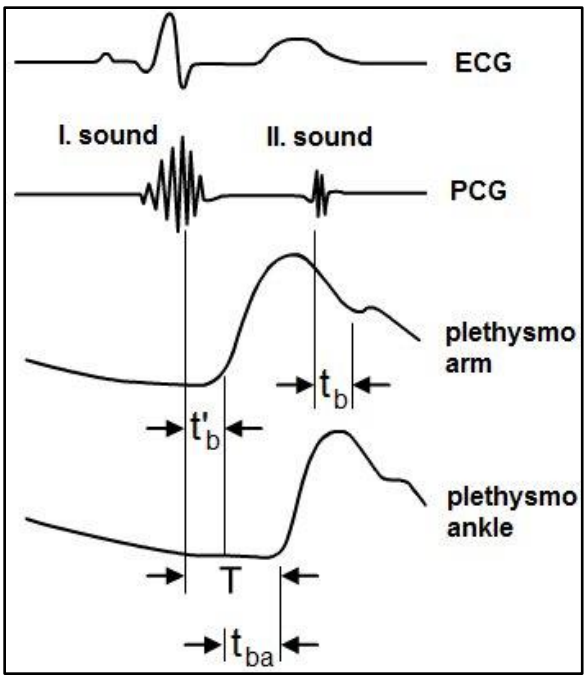


Figure 04: Signal Synchronization [3].

Vital signs from PCG

1. Heart Rate:

a. Definition:

Heart rate from Phonocardiogram (PCG) refers to the measurement of the number of heart beats per minute obtained from the analysis of the acoustic signals recorded during heart auscultation. It is interchangeably used with Pulse Rate as both terms represent the frequency at which the heart contracts and pumps blood throughout the body. Heart Rate is an important factor for measuring overall health because it reflects the efficiency and effectiveness of the cardiovascular system, which plays a central role in supplying oxygen and nutrients to the body's tissues and organs. Regular monitoring of heart rate allows for early detection of potential issues and enables healthcare professionals to tailor interventions and treatments to promote overall well-being and reduce the risk of cardiovascular complications.

b. Theory:

As S1 peak corresponds to r peaks of ECG signal, we can calculate Heart Rate from the S1-S1 interval just like the R-R interval.

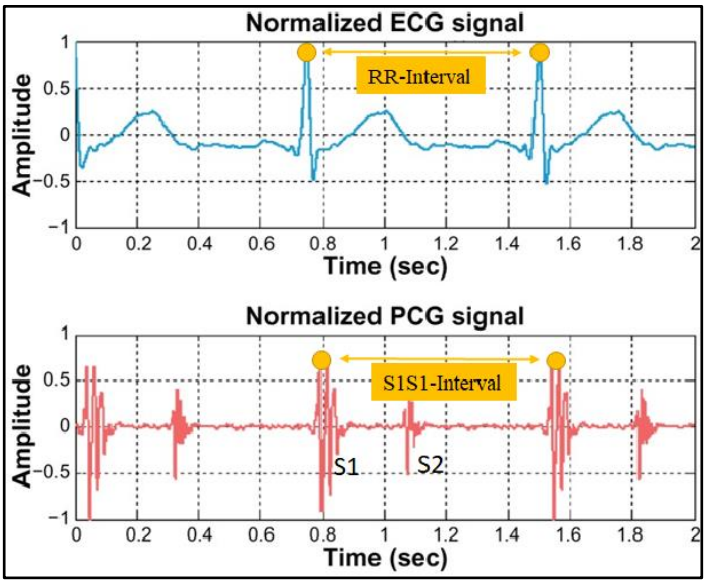


Figure 05: Relation between R-R interval and S1-S1 interval

- First, we must find the number of S1 peaks in the given frames. Number of intervals will be 1 less than the number of peaks.
  - N(interval) = N(peaks) - 1
- Get the average interval duration in seconds by identifying 1st and last peaks and dividing their differences by Number of intervals.
- Then Similar to HR calculation from ECG:
  - HR (BPM) = 60 / S1-S1 interval (in seconds)

The heart rate calculated from both ECG and PCG should be similar, as they represent the same physiological parameter, i.e., the number of heart beats per minute. However, the heart rate derived from PCG might have slight variations from ECG due to the inherent differences between the two signals (Electrical vs Sound)

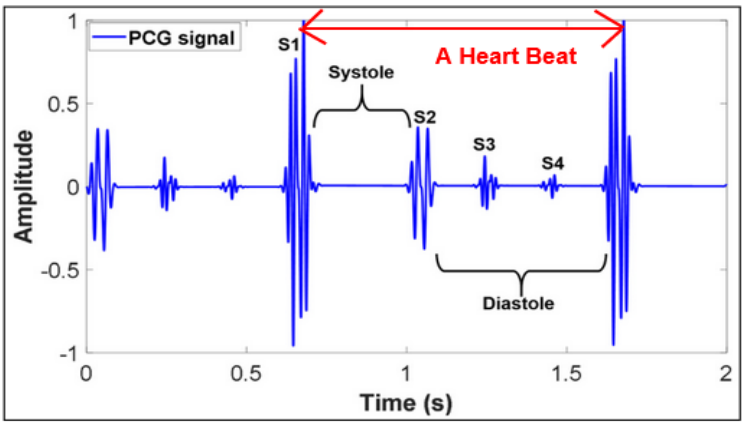


Figure 06: Calculating a Heartbeat from S1-S1 peak

2. Heart Rate Variability:

a. Definition:

Heart rate variability (HRV) is a measure of the variation in time intervals between successive heartbeats, also known as R-R intervals. In other words, it is the fluctuation in the time duration between each heartbeat. Higher HRV is generally considered a sign of good cardiovascular health and adaptability of the autonomic nervous system. It suggests that the body can quickly adjust to various stressors and demands, promoting better heart function and overall well-being.

HRV provides information about the autonomic nervous system's activity, which regulates heart rate in response to various internal and external stimuli which isn't accessible through Heart Rate. Although acquisition of these two parameters are similar, while heart rate is a fundamental measure of heart activity, HRV provides a more nuanced and comprehensive view of cardiovascular health and autonomic nervous system function. It's expected to have some variability in between heart beats in healthy hearts.

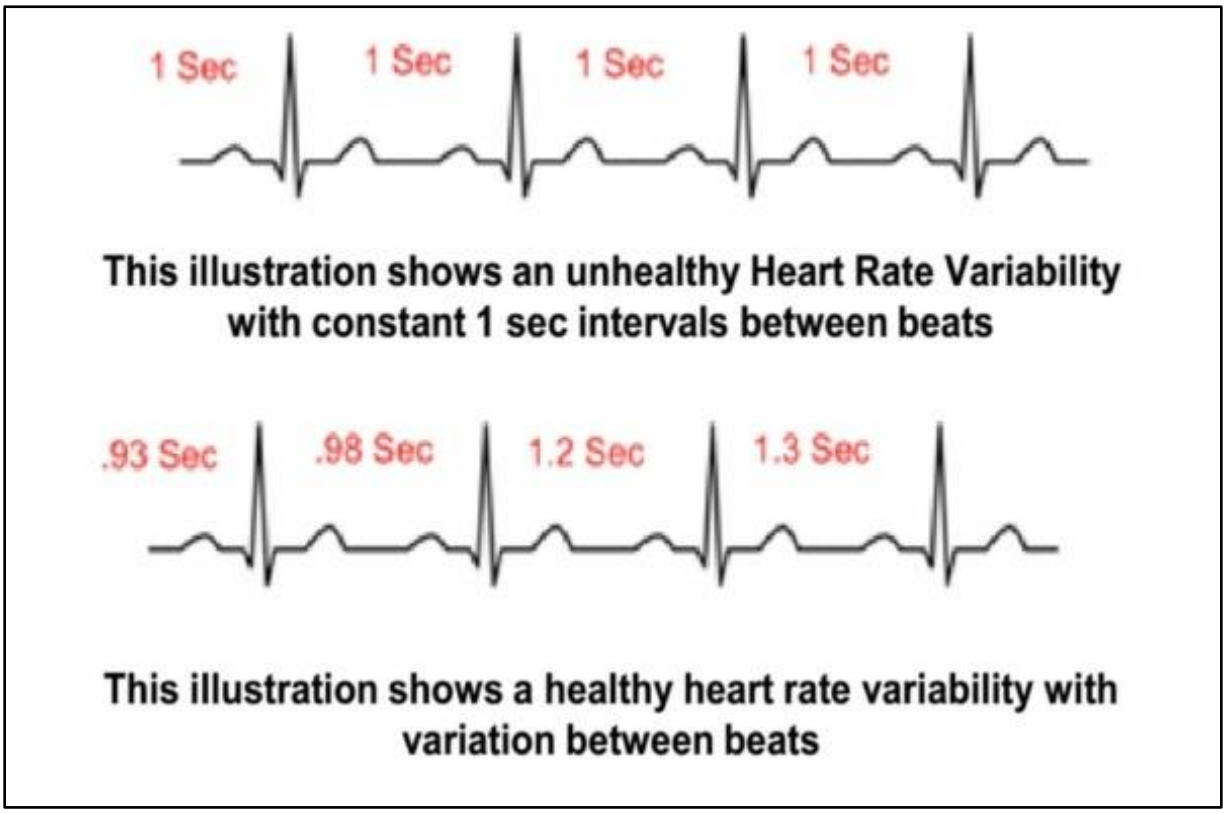


Figure 07: Healthy vs Unhealthy HRV [4].

b. Theory:

Heart rate variability can be measured in many different ways from Heart Rate. Among them measuring HRV by calculating Standard Deviation of Heart Rate in a given frame among each R-R or S1-S1 interval is one of the simplest methods.

Table 1

HRV time-domain measures.

Parameter	Unit	Description
SDNN	ms	Standard deviation of NN intervals
SDRR	ms	Standard deviation of RR intervals
SDANN	ms	Standard deviation of the average NN intervals for each 5 min segment of a 24 h HRV recording
SDNN index (SDNNI)	ms	Mean of the standard deviations of all the NN intervals for each 5 min segment of a 24 h HRV recording
pNN50	%	Percentage of successive RR intervals that differ by more than 50 ms
HR Max – HR Min	bpm	Average difference between the highest and lowest heart rates during each respiratory cycle
RMSSD	ms	Root mean square of successive RR interval differences
HRV triangular index		Integral of the density of the RR interval histogram divided by its height
TINN	ms	Baseline width of the RR interval histogram

Open in a separate window

Interbeat interval, time interval between successive heartbeats; NN intervals, interbeat intervals from which artifacts have been removed; RR intervals, interbeat intervals between all successive heartbeats.

Table 01: Time domain methods of calculating HRV [5].

Table 2

HRV frequency-domain measures.

Parameter	Unit	Description
ULF power	ms^2	Absolute power of the ultra-low-frequency band (≤0.003 Hz)
VLF power	ms^2	Absolute power of the very-low-frequency band (0.0033–0.04 Hz)
LF peak	Hz	Peak frequency of the low-frequency band (0.04–0.15 Hz)
LF power	ms^2	Absolute power of the low-frequency band (0.04–0.15 Hz)
LF power	nu	Relative power of the low-frequency band (0.04–0.15 Hz) in normal units
LF power	%	Relative power of the low-frequency band (0.04–0.15 Hz)
HF peak	Hz	Peak frequency of the high-frequency band (0.15–0.4 Hz)
HF power	ms^2	Absolute power of the high-frequency band (0.15–0.4 Hz)
HF power	nu	Relative power of the high-frequency band (0.15–0.4 Hz) in normal units
HF power	%	Relative power of the high-frequency band (0.15–0.4 Hz)
LF/HF	%	Ratio of LF-to-HF power

Table 02: Frequency domain methods of calculating HRV [5].

Normal Values of Standard Measures of HRV[85]

Time Domain Analysis			Spectral Analysis		
Variable	Units	Normal Values (mean±SD)	Variable	Units	Normal Values (mean±SD)
IBI	ms	926±90	LF	ms^2	519±291
SDNN	ms	50±16	HF	ms^2	657±777
RMSSD	ms	42±15	LF	nu	52±10
			HF	nu	40±10
			LF/HF ratio		2.8±2.6

Table 03: Normal Values of Standard Measures of HRV [6].

3. Blood Pressure:

a. Definition:

Blood pressure refers to the force exerted by the blood against the walls of the arteries as the heart pumps blood throughout the body. It is measured using two values: systolic pressure (the higher value when the heart contracts) and diastolic pressure (the lower value when the heart relaxes). It can fluctuate throughout the day due to various factors, including physical activity, stress, emotions, and the body's autonomic nervous system responses.

It’s a very important parameter to measure in order to get idea about health condition because-

a) Detection of Hypertension: Measuring blood pressure regularly allows for the early detection of hypertension, which is crucial as it is often asymptomatic. Untreated hypertension can lead to serious health complications, such as heart disease, stroke, kidney damage, and vision problems.
b) Monitoring Cardiovascular Health: Regular blood pressure measurements can help assess the overall condition of the heart and blood vessels, allowing healthcare professionals to identify potential risks and implement preventive measures.
c) Assessment of Treatment Effectiveness: Blood pressure measurements help healthcare professionals adjust medications, lifestyle modifications, or other interventions to achieve optimal blood pressure control.
d) Risk Assessment: It is a significant risk factor for heart disease, stroke, and other cardiovascular disorders. People with high blood pressure may need more comprehensive assessments and preventive measures to reduce their risk of developing these conditions.
e) Diagnostic Tool: specific guidelines use blood pressure thresholds to diagnose hypertension or to classify the severity of hypertension.

Thus, monitoring blood pressure plays a crucial role in preventive healthcare and is a key component of routine health screenings and clinical assessments.

b. Theory:

Blood pressure is typically not directly measured from the Phonocardiogram (PCG) signal. While the PCG signal can provide valuable information about the heart's mechanical activity and any potential abnormalities in heart sounds, it does not directly provide information about blood pressure levels. Other noninvasive methods, such as PPG, ECG and other oscillometric devices, are used to estimate blood pressure based on the characteristics of the arterial pulse wave, but these methods are slightly different from direct blood pressure measurements obtained through sphygmomanometer.

However, Blood pressure has relation with other health parameters like Heart Rate, Heart Rate Variability, Pulse Transit Time (PTT), Pulse Wave Velocity (PWV) etc. Also, different features obtained from PCG signals like systole and diastole durations, systole and diastole rise and decays, systole and diastole peaks etc have some indirect relationship with the Blood Pressure.

We can make use of these indirect relationships to calculate blood pressure using PCG. Various signal processing and Deep Learning based methods are proposed to do so.

Among them a simple solution can be semi empirical linear models such as this equations [7]

$$P_{\text{sys}} = 289 - 0.6t_{s1} - 0.01t_s - 0.9\alpha_{\text{ins1}} - 13.3\varphi_{\text{ins1}} + 1.44\alpha_{\text{des1}} - 9.02\varphi_{\text{des1}} - \alpha_{\text{des2}} + 5.1\varphi_{\text{des2}}, \quad (3)$$

$$P_{\text{dias}} = -34.735 - 0.576\text{HR} - 0.305t_{s2} - 0.193t_s + 9.064\varphi_{\text{des1}} - 0.584_{\alpha_{\text{ins2}}} - 8.162\varphi_{\text{ins2}}, \quad (4)$$

4. Respiration Rate:

a. Definition:

Respiration rate, also known as breathing rate, refers to the number of breaths a person takes per minute. It is a vital physiological parameter that measures the frequency of breathing and is an essential indicator of respiratory health.

Respiration rate is typically expressed in breaths per minute (BPM). In healthy adults at rest, the normal respiration rate ranges between 12 to 20 breaths per minute. However, respiration rate can vary depending on several factors, such as age, physical activity, emotional state, and overall health.

b. Theory:

Respiration rate can be measured directly by monitoring the rising and lowering of Diaphragm. The following figure measures respiratory pressure signal to identify respiration rate:

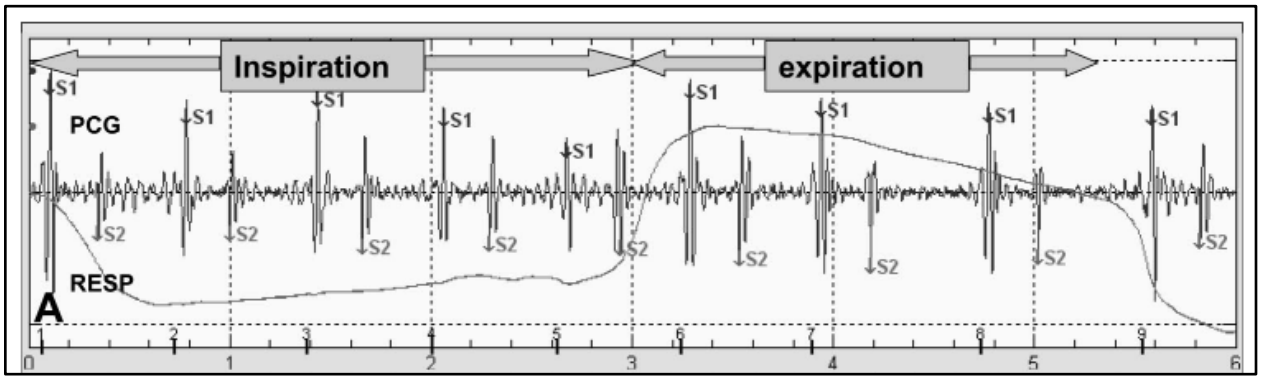


Fig 08: Morphological variability of heart sounds. A: during a normal respiratory cycle, first (S1) and second (S2) heart sounds exhibit considerable changes in morphology. [22]

In the PCG sound as its intensity is lower than the S1 and S2 sound it resides in the background sound which can be easily mixed up with other noises.

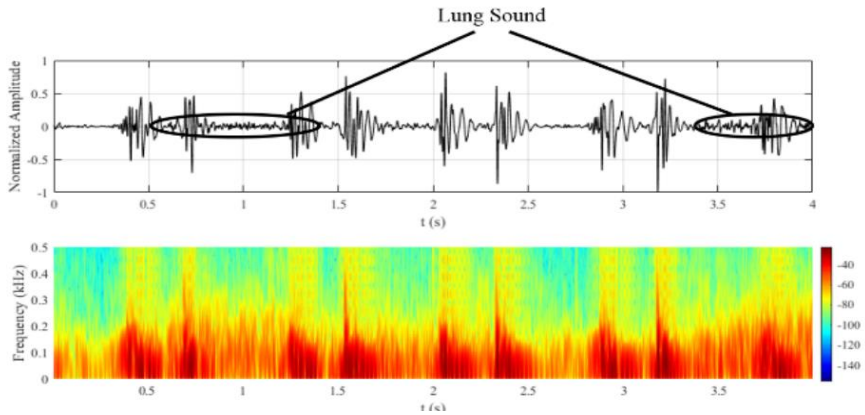


Fig 09: PCG and its spectrogram from Mitral Area [7]

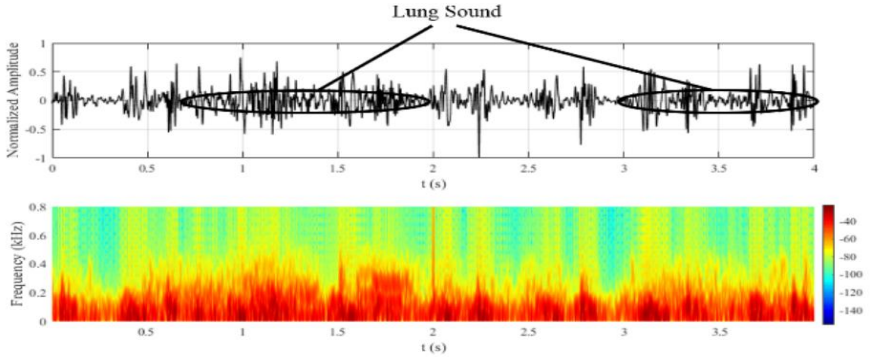


Fig 10: PCG and its spectrogram from aortic Area [7]

So, extracting respiratory sound and directly extracting respiratory rate from it is quite challenging. But as different papers highlighted the relationship between HR, HRV, BP and RR and also the relationship of RR with different features extracted from PCG [23]. The following table summarizes some of the features extracted from PCG to estimate RR. Additional features can replace the ECG extracted features in this table.

Feature Class	Source Signals	Feature Symbol	Description
Timing	ECG	HR	Heart rate
Timing	PCG	$LVET$	Left ventricular ejection time
Timing	ECG, PCG	PEP	Pre-ejection period
Area	ECG	QRS_{area}	QRS complex area
Area	PCG	$S1_{area}$	S1 area
Area	PCG	$S2_{area}$	S2 area
Amplitude	ECG	QRS_{amp}	QRS complex amplitude
Amplitude	PCG	$S1_{amp}$	S1 amplitude
Amplitude	PCG	$S2_{amp}$	S2 amplitude
Morphology	ECG	QRS_{PCA}	Morphological variations of the QRS complex
Morphology	PCG	$S1_{PCA}$	Morphological variations of the S1 peak
Morphology	PCG	$S2_{PCA}$	Morphological variations of the S2 peak

Table 04: ECG- and PCG-derived respiration features. [23]

A semi-empirical model like we have used in blood pressure or ML/DL based model can be used to make use of these features to extract correct respiratory rate. Previously extracted labels (HR, HRV, and BP) can be used as features as they are related.

5. Oxygen Saturation(SpO_2)

a. Definition:

Oxygen saturation, abbreviated as SpO2, is a crucial measure in assessing respiratory and circulatory health. It represents the percentage of hemoglobin in the blood that is bound to oxygen at a specific moment. Hemoglobin is responsible for transporting oxygen from the lungs to tissues. A higher SpO2 percentage indicates better oxygenation, as more hemoglobin is carrying oxygen. Typically expressed as a percentage, normal levels range from 95% to 100% in healthy individuals. Monitoring SpO2 helps healthcare professionals evaluate oxygen delivery and guide treatment decisions, especially in individuals with respiratory or cardiovascular conditions.

b. Theory:

Oxygen saturation can't be directly estimated from PCG signals as sound signals like PCG have no information about oxygen capacity in blood. But different parameters in PCG are influenced by SpO2 and hereby it can be measured by a semi empirical model, or ML/DL algorithm like in the previous manner.

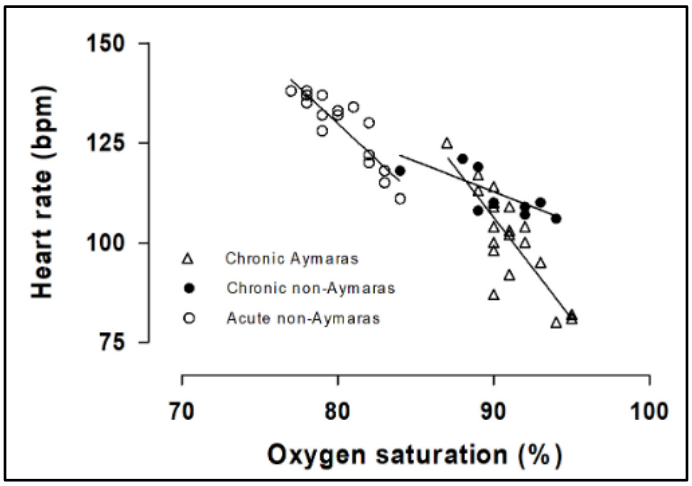


Fig 11: Inverse relationship between Heart Rate and SpO_2 [24].

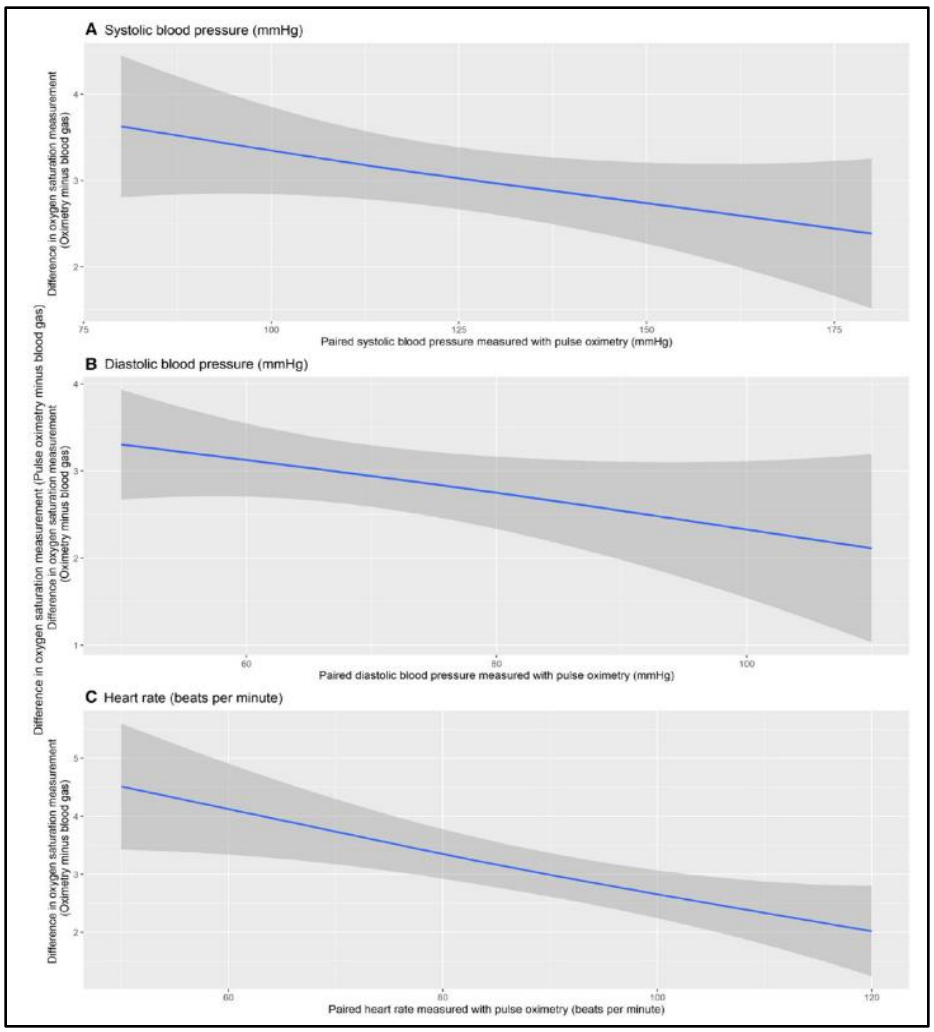


Fig 12: Inverse relationship of Heart Rate and BP withSpO2 [25].

Noise and distortions in PCG

As PCG is a sound signal, various noises and distortions can occur while collecting PCG data. Normal frequency distribution of PCG signal is shown below-

Sound	S1	S2	S3	S4
Frequency	30-100 Hz	Above 100 Hz	20-25 Hz	Below 30 Hz

Table 05: The four basic heart sounds and their frequency range [8].

Also the frequency distribution of PCG signal with respect to its two main components S1 and S2 are shown below:

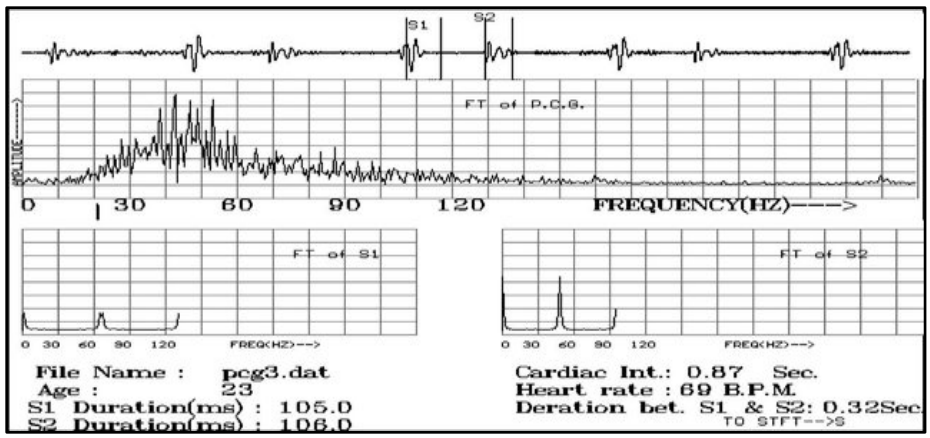


Figure 13: Frequency domain analysis (FT) of the PCG signal [9].

These Frequency distributions vary from person to person. A comparison of dominant frequency regions among persons to persons are shown below:

Patient ID	Dominant frequency range in whole signal (Hz)	Peak frequency of S1 (Hz)	Peak frequency of S2 (Hz)
PCG1	50 – 110	38	34
PCG2	35 – 90	38	30
PCG3	28 – 90	62	52
PCG4	35 – 80	48	45
PCG5	40 – 110	38	35
PCG6	25 – 110	75	55
PCG7	27 – 110	50	45
PCG8	15 – 118	30	45
PCG9	30 – 90	35	32

Table 06: Results obtained from frequency domain analysis (FT) of PCG records [9].

Various noises and artifacts that overlap in frequency regions corrupts PCG signals. The following figure shows the effect of artifacts on the frequency spectrum of PCG signals. This is why cleaning PCG signals is a hard problem.

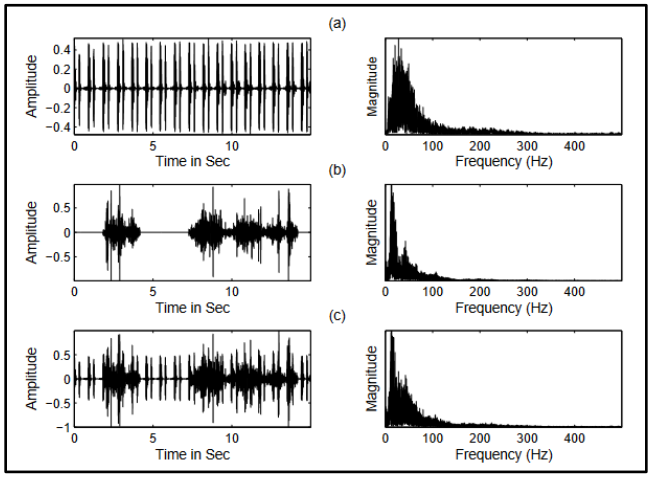


Figure 14: Frequency domain analysis (FT) of the PCG signal [9].

Some of these artifacts are [12] –

- Type I: High intensity and short duration artifacts
    - Coughing
    - Hiccups
    - Stethoscope movement
- Type II: Low intensity and long duration artifacts
    - Fast breathing
    - Deep breathing
    - Body movement

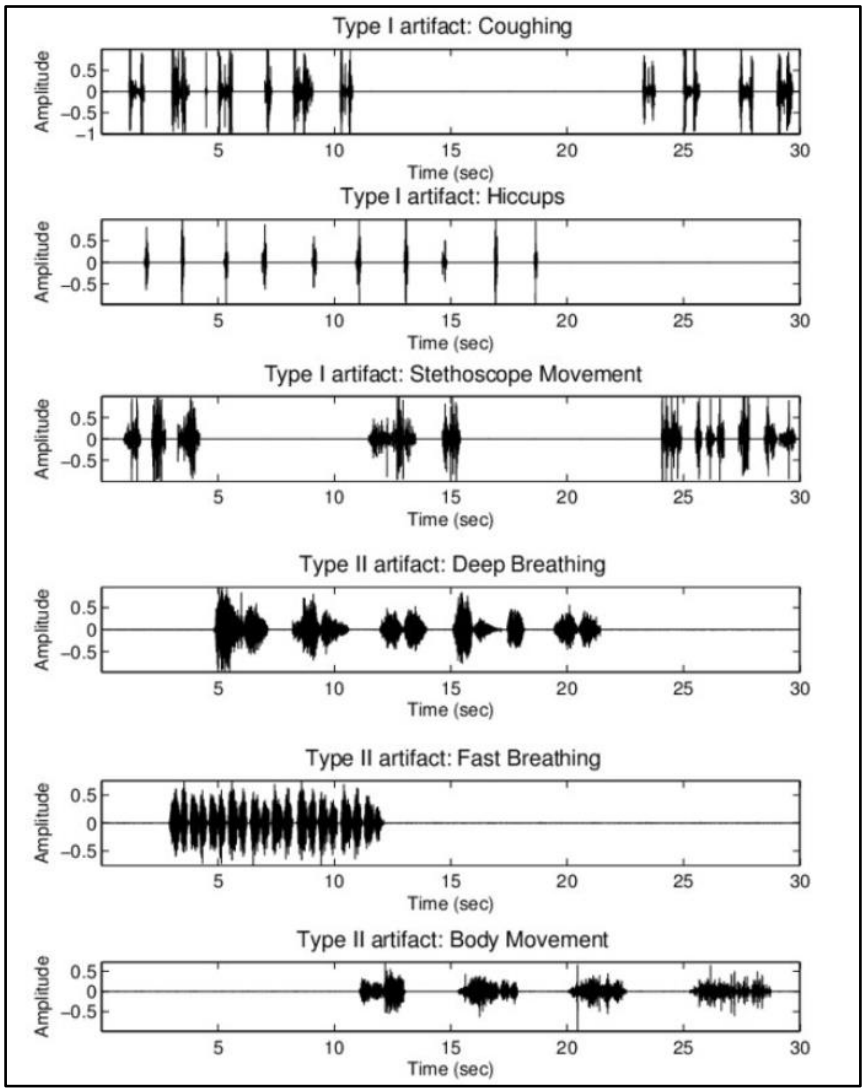


Figure 15: Six different types of artifacts [12].

Thus, some of the noises and artifacts overlap with frequency range on this spectrum. They can be divided into 3 categories frequency range wise [10]-

- High-frequency (HF) noise with the frequency range above 20 Hz, representing the sound artifacts from the environment and patient's body, optionally some electromagnetic interference.
- Low-frequency (LF) noise with the frequency range up to 100 Hz , simulating the motion artifacts which occur mainly due to the motion of the sensor on the patient's chest;
- Whole-frequency (WF) noise containing a combination of voice signals, reproducing the speech noise from the environment

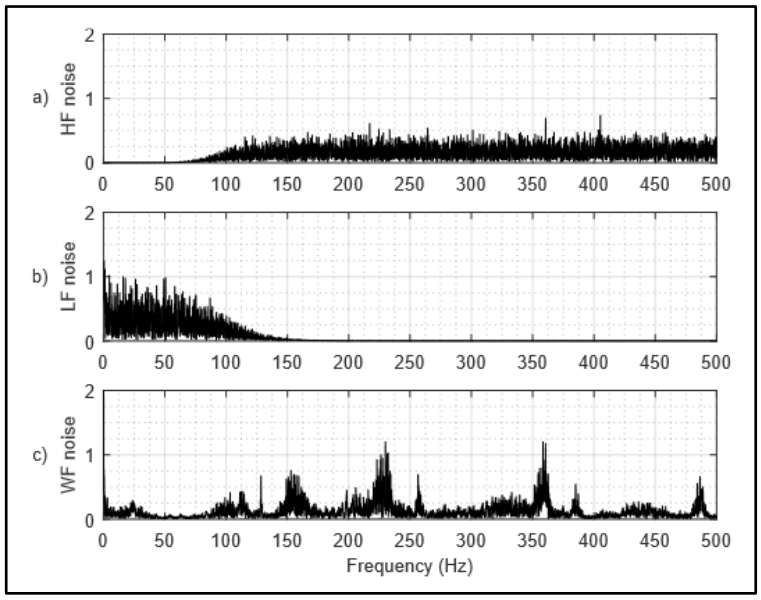


Figure 16: Frequency spectra of the used types of noise, from above: HF, LF and WF noise. [11].

These noises come from various sources. Some of the common of these are-

1. Ambient Noise: Ambient noise refers to background noise present in the environment during the Phonocardiogram (PCG) recording. It includes various sounds from the surroundings, such as electronic equipment, ventilation systems, room acoustics, or other external factors. Ambient noise can be considered as random noise and may contain components that follow a Gaussian distribution.
2. Gaussian Noise: Gaussian noise is random noise that follows a Gaussian (normal) distribution. It is a type of statistical noise with equal probability of positive and negative values, and it can be present in PCG signals due to various sources, including ambient noise, sensor noise, or recording artifacts.
3. Movement Artifact: Movement artifacts in PCG signals occur when there is motion or movement during the recording process. Patient movement, sensor movement, or other motion-related factors can introduce artifacts into the PCG signal. Movement artifacts can affect the accuracy of heart sound analysis and may appear as random disturbances with a Gaussian distribution.

4. Ambient Music or Speech: In some cases, ambient music or speech from the environment can be inadvertently recorded in PCG signals. These sounds can be considered as background noise and may add to the overall ambient noise present in the recording.
5. Doors closing and Other Environmental Noise: Environmental sounds like doors closing, footsteps, or other noises from the surroundings can be considered as ambient noise that may be present in PCG recordings.
6. Breathing Sounds: Breathing sounds from the patient can introduce respiratory artifacts into the PCG signal. These sounds may vary in frequency and amplitude depending on the patient's breathing pattern and can be considered as random noise in the PCG recording.
7. Chest Movements: Chest movements, especially during breathing or patient repositioning, can cause artifacts in the PCG signal. These movements can introduce random disturbances that may affect the clarity of the heart sounds.
8. Cough, Speech, Intestinal and Stomach Growls, or Acoustic Damping of the Bones: Other sounds produced by the patient's body, such as coughing, speech, intestinal and stomach growls, or acoustic damping of the bones, can also contribute to noise in the PCG signal. These sounds may have a random nature and add to the overall noise present in the recording.

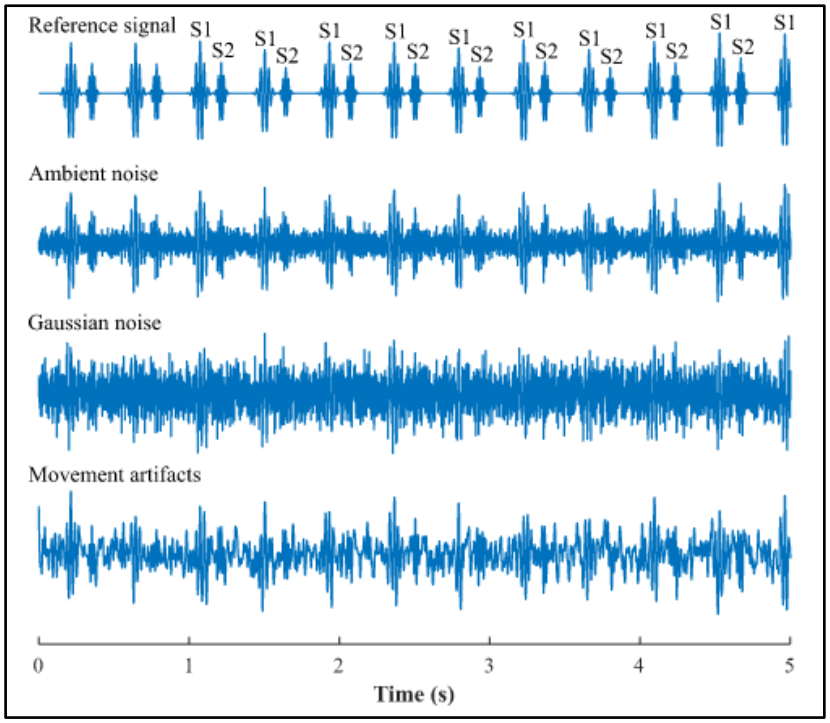


Figure 17: Frequency spectra of the used types of noise, from above: HF, LF and WF noise [17].

It's essential to address and minimize these sources of noise in PCG recordings to obtain clean and reliable signals for accurate heart sound analysis. Signal processing techniques, noise reduction

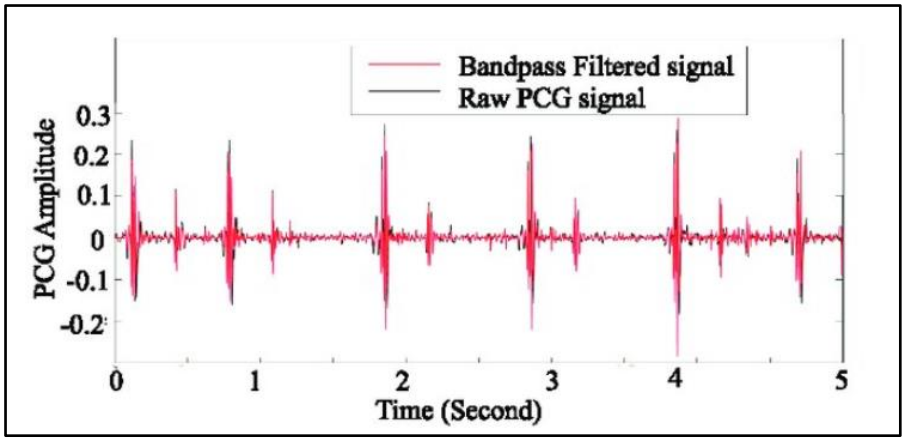
algorithms, and proper patient positioning are some of the strategies used to improve the quality of PCG recordings and enhance the accuracy of heart sound analysis.

Denoising Techniques

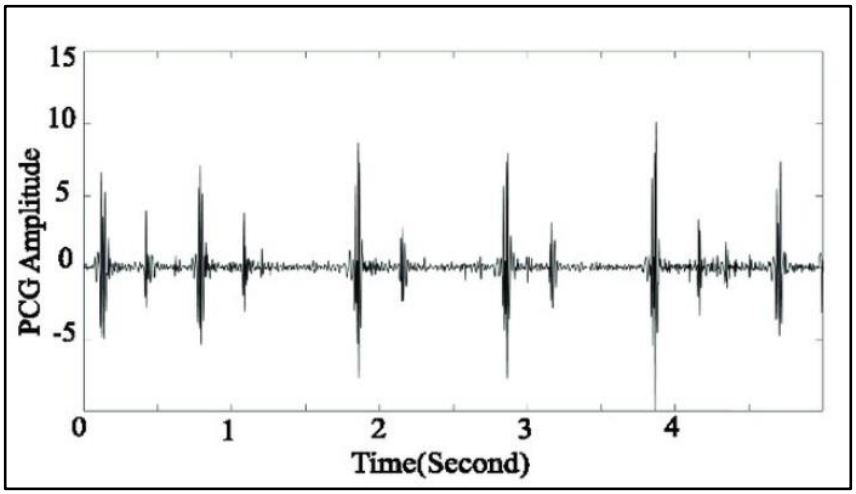
As shown before, just bandpass filtering or simple techniques can't clean PCG signals as the existing noises overlap in both time and frequency domain with the signal. Various signal processing methods have been used to optimize the denoising process of PCG signals. Some of the best performing and commonly used methods are presented below:

1. **Classical Butterworth band-pass filter:**

   The classical Butterworth band-pass filter is a type of analog or digital filter with a flat frequency response in the passband and a gradual roll-off in the stopbands. The High frequency noise can be reduced using these Bandpass filters with cut-off frequencies from 25 to 400 Hz. (Mubarak et al., 2018)



(a)



(b)

Figure 18: (a) preprocessed PCG signal after applying bandpass filter, (b) normalization of filtered PCG signal. [12]

2. Spectral subtraction

Spectral subtraction is a noise reduction technique used in signal processing to remove background noise from a noisy signal. It works by estimating the spectral profile of the noise from the noisy signal and then subtracting this estimated noise profile from the original signal, leaving behind the desired clean signal.

Effect of spectral subtraction on speech signal:

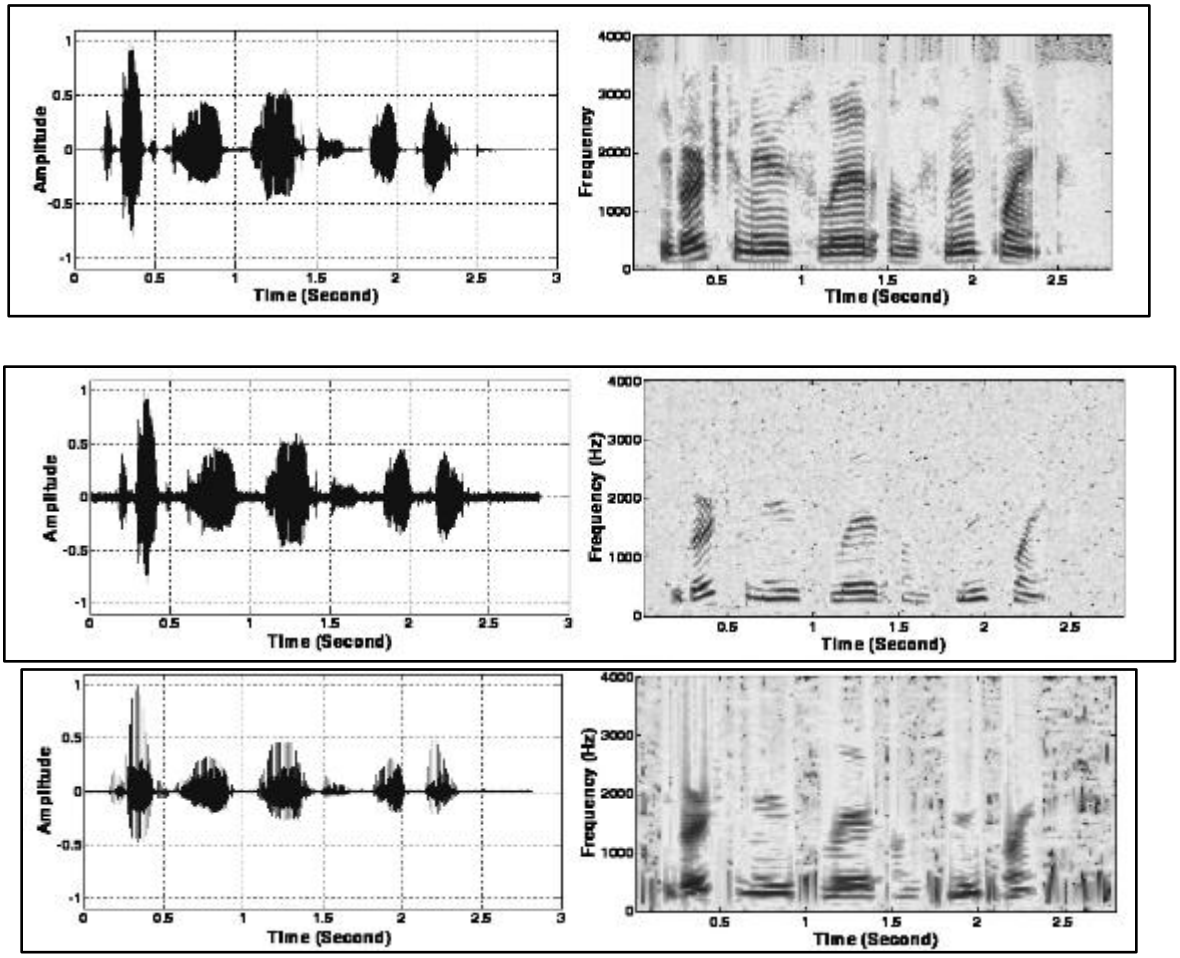


Figure19: Waveforms and spectrograms (From top to bottom): (i) Clean speech; (ii) Noisy speech (white noise at 15 dB); (iii) Speech enhanced by different subtractive-type algorithms [12].

Multiple implementation of spectral subtraction on speech signal:

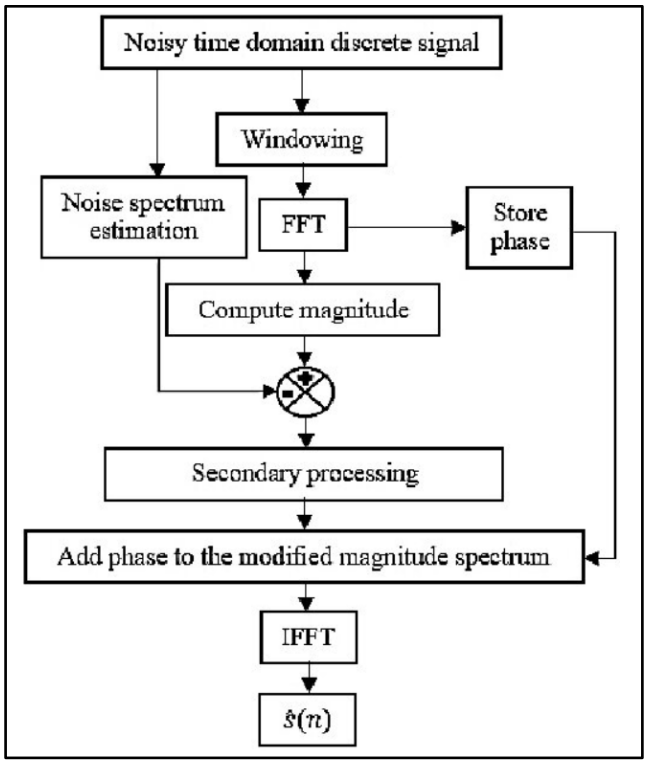


Figure 20: Flow Chart Showing the Spectral Subtraction Followed by Secondary Processing Needed to Reduce Auditory Effects due to Spectral Error. [13]

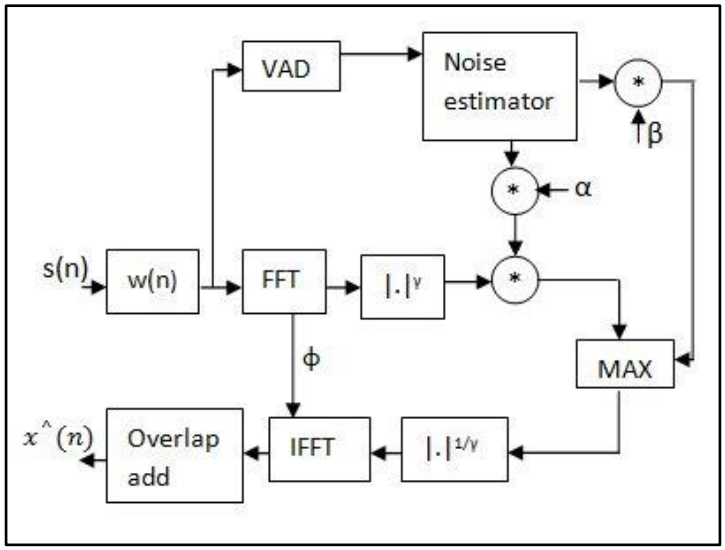


Figure 21: Spectral Subtraction Method Block Diagram. [14]

3. **Wavelet Transform**
    a. continuous wavelet transforms (CWT) (time frequency resolution)
    b. Morlet wavelet

c. wavelet energy spectrum (WES) and the wavelet frequency spectrum (WFS) used to determine which one works better [mamun-2]
d. CWT, WES, and WFS
e. Metrics : AR, ARFS, NRMSE
f. adaptive wavelet transform (AWT) [imtiaz-1]

Wavelet transform (WT) is a frequently tested and very effective method able to suppress noise when its setting is optimized for a given purpose. Contrary to the FT, WT treats frequency logarithmically which corresponds to the acoustic perception of the human body and therefore is more suitable for the analysis of the sound-related signals. It decomposes the input signal into a set of wavelets – a wave-like oscillation with two basic properties: scale, which relates to the signal's frequency, and location, which defines its position in time. In the WT-based filtering, one should pay attention to selecting suitable set of the parameters, such as the wavelet base (e.g. Daubechiens or Coiflets) or level of decomposition. Moreover, the effectiveness of WT-based filters can be further increased by combining this algorithm with other filtration methods.

CWT vs DWT:
Wavelet transform can be implemented in a continuous and discrete manner. The CWT is continuous and provides a more continuous time-frequency representation of the signal, but it is computationally more demanding and redundant. On the other hand, the DWT is discrete and computationally efficient, but its time-frequency representation is more discrete and may not be as smooth as the CWT. The continuous wavelet transform (CWT) is especially suitable for the analysis of non-stationary signals and to obtain the TFR, due to its high resolution, both in time and in frequency.

The CWT is more appropriate than the DWT for the spectral analysis of non-stationary signals because the DWT does not use all the scales which contract and expand the mother wavelet. If all the scales in the range related to the frequency band of the signal are not used, loss of some information is unavoidable and this will lead to difficulties in the extraction of the features of the original signal (Ergen and Tatar 2001).

Morelet Wavelet:
In some literature they have highlighted that Morelet is the most reliable wavelet for the time– frequency analysis of PCG signals.

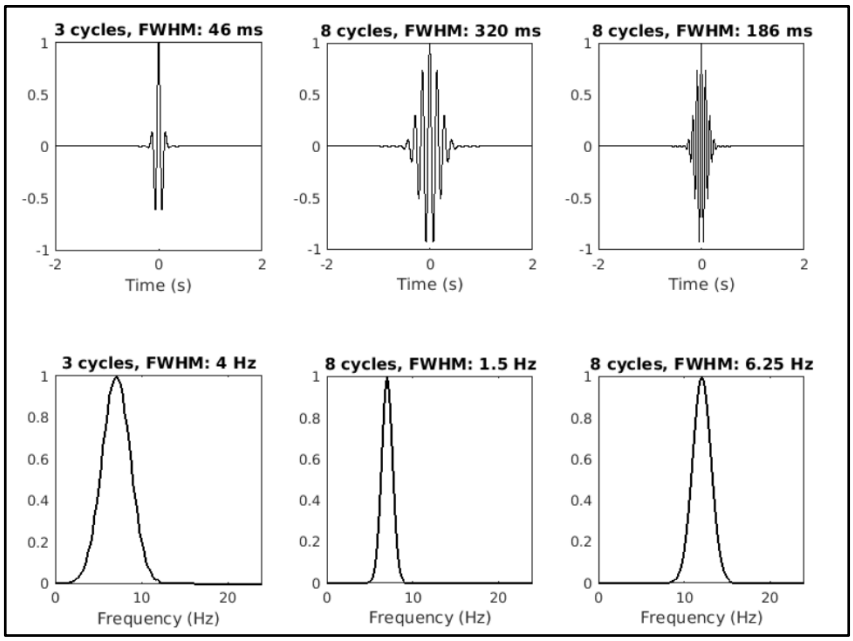


Figure 22: Three Morlet wavelets in the time domain (upper row) and in the frequency domain (lower row) with two frequencies and different time-frequency trade-off parameters. [16]

A complex Morlet wavelet w can be defined as the product of a complex sine wave and a Gaussian window,

$$w = e^{2i\pi ft} e^{\frac{-t^2}{2\sigma^2}} \quad (1)$$

where i is the imaginary operator, f is frequency in Hz, and t is time in seconds. To avoid introducing a phase shift, t should be centered at t=0, for example, by defining t from -2 to +2 seconds is the width of the Gaussian, which is defined as

$$\sigma = \frac{n}{2\pi f} \quad (2)$$

The parameter that defines the time-frequency precision trade-off is n, which is often referred to as the "number of cycles." For neurophysiology data such as EEG, MEG, and LFP, typical values of n range from 2 to 15 over frequencies between 2 Hz and 80 Hz. Other applications may require a different number of cycles.

A better definition:

$$w = e^{2i\pi ft} e^{\frac{-4ln(2)t^2}{h^2}} \quad (3)$$

The new parameter here is h, which is the FWHM in seconds.

Adaptive Wavelet Transform (AWT):

Adaptive thresholding is actually a DWT method where each wavelet coefficient was assigned with a certain threshold value. A moving window with a specified length is used, whereby the threshold responds to changes of the noise power in the signal waveform. Thresholding can be further divided into hard and soft. In the case of hard thresholding, the coefficients below the threshold are replaced by zeros, while, in the case of soft thresholding, the remaining non-zero coefficients are shifted towards zero by the size of the threshold. Finally, an inverse discrete wavelet transform (IDWT) can be used to reconstruct the filtered PCG signal.

Choosing correct CWT type [18]:

Although there are many studies to determine which wavelet type is optimal for decomposing the signal in DWT application such as denoising, there is not any noteworthy study proposing which wavelet is suitable for CWT analysis to obtain TFR of a signal, particularly PCG signals. Nevertheless, it has not yet been studied which wavelet is suitable for the CWT analysis of PCG signals. Here, we used two spectral methods, the **wavelet energy spectrum (WES)** and the **wavelet frequency spectrum (WFS)**, derived from the CWT, in order to determine the best wavelet function. The determination was made by comparing these wavelet spectra to the energy and the frequency spectra of the PCG signal. While the energy spectrum (ES) is directly obtained from the interested PCG signal, the frequency spectrum (FS) is obtained from the AR modeling of the signal.

The definition of WES:

$$\mathrm{WES}(n) = \frac{1}{N}\sum_{k=1}^{M} |\mathrm{CWT}(n,k)|^2,$$

where M is the number of scales and N is the length of signal.

The definition of WFS:

$$\mathrm{WFS}(f) = \sum_{n=1}^{N} \left|\mathrm{CWT}\left(n, \frac{f_s f_0}{s}\right)\right|^2,$$

where, f is the local frequency.

4. **Empirical mode decomposition (EMD)[17]:**

Empirical mode decomposition (EMD) is a signal processing technique that is able to decompose any non-stationary and non-linear signal into oscillating components. These bandlimited components are also called intrinsic mode functions (IMFs). Each IMF must comply with two basic conditions. First, the number of extrema and the number of zero

crossings must be the same or may differ by one at most. Second, at any point, the mean value of the envelopes defined by the local maxima and the envelopes defined by the local minima is zero. The algorithm is described below:-

1) The upper and lower envelopes, respectively of the local maxima and minima, are estimated using a cubic spline interpolation.
2) The mean m(t) of the two envelopes is subtracted from the original signal s(t).

$$e1(t) = s(t) - m(t).$$

3) If e1(t) meets both of the aforementioned conditions, it is denoted as IMF1. If not, steps 1. and 2. are repeated until the first IMF is obtained.
4) Subsequently, a residue ri(t) is defined by subtracting e1(t) from the input signal s(t).

$$ri(t) = s(t) - e1(t).$$

5) To obtain subsequent IMFs, the entire process is repeated, but, instead of the input signal s(t), the residue ri(t) is used.

The entire process is repeated until the final residue is a monotonic function or a constant. The original signal can be reconstructed by summing all extracted IMFs and the last residue.

5. Ensemble empirical mode decomposition (EEMD)[17]:

The ensemble empirical mode decomposition (EEMD) method was devised as an improved version of the empirical mode decomposition (EMD) procedure. The EMD is very effective, but it is burdened by limitations in the form of so-called mode mixing problems. This is a phenomenon in which the EMD is not able to correctly decompose the signal into individual IMFs, resulting in one IMF containing several components of very different frequencies. This problem usually occurs when the original low-frequency signal contains discontinuous and isolated high-frequency oscillations. To prevent this phenomenon, an extended version of this method, termed EEMD, was proposed. With the EEMD, individual IMFs are obtained by averaging the results of several EMD cycles, wherein random white noise with a predefined amplitude is added to the input signal. The EEMD method algorithm can be described in the following step:-

1) Set the number of ensemble trials N and the standard deviation of the added noise Nstd.
2) Add a random white noise to the input signal.
3) Decompose the resulting signal into individual IMFs using the EMD algorithm.
4) Repeat step 1 and 2 N -times, but with different white noise series each time.
5) Determine the final IMF j(t) of the EEMD by averaging all IMFs related to N trials.

$$IMF_j(t) = \frac{1}{N}\sum_{i=1}^{N} IMF_{i,j}(t),$$

where j is a number of an IMF scale.

The added white noise leads to the correct decomposition of the signal into IMFs using the EMD method. The final result is the average of a large number of iterations, in whic a different white noise is used. By averaging, the white noise is eliminated, and only the resulting signal remains.

6. **Autoregressive models**

An autoregressive model is a type of time-series model where the current value of the signal is modeled as a linear combination of its past values and a white noise term. The AR model of order 'p' can be represented as follows:

$$y(t) = c + \sum(ai * y(t-i)) + \varepsilon(t)$$

where:

a. y(t) is the current value of the signal at time 't',
b. ai are the coefficients of the model,
c. y(t-i) are past values of the signal, typically from t-1 to t-p,
d. c is a constant term,
e. ε(t) is the white noise term at time 't'.

7. **Short-time FT (STFT and Wigner distribution == to capture time frequency resolution) [wavelet solves this resolution problem]**

Short-Time Fourier Transform (STFT) is a commonly used technique for analyzing and denoising signals, including phonocardiogram (PCG) signals. The STFT allows us to observe the frequency content of a signal over time by computing the Fourier Transform of small overlapping segments of the signal. The equation for STFT of a signal x(t) at time t and frequency f is as follows:

$$STFT(t, f) = \int[x(t) * w(t - \tau) * e^{(-2\pi i f \tau)}] \, d\tau$$

Where:

a. x(t) is the original PCG signal.
b. w(t - τ) is the window function applied to the signal at time t with a time-shift of τ. The window function is used to control the trade-off between time and frequency resolution and to mitigate artifacts at the edges of the windowed segment.
c. f is the frequency variable.
d. e^(-2πifτ) represents the complex exponential used in the Fourier Transform.
e. STFT(t, f) is the resulting representation of the signal in the time-frequency domain.

Similarly, the Wigner distribution is another time-frequency analysis method that can be used for PCG signal denoising. It provides a representation of the signal's time-varying frequency content and allows for joint time and frequency analysis. However, the Wigner distribution is not a true distribution since it may exhibit negative values in certain cases.

The equation for the continuous Wigner distribution of a signal x(t) at time t and frequency f is given by:

$$W(t, f) = \int[x(t + \tau/2) * x^*(t - \tau/2) * e^{(-2\pi i f \tau)}] \, d\tau$$

Where:

- x(t) is the original PCG signal.
- x*(t) denotes the complex conjugate of the signal x(t).
- f is the frequency variable.
- τ is the time-lag variable.
- W(t, f) is the resulting Wigner distribution representing the time-frequency content of the signal.

Time–frequency analysis by means of TFR methods based on traditional FT such as short-time FT(STFT) or Wigner distribution is a useful and basic tool to obtain the TFR of such signals.

If the analysis window is made short enough to capture rapid changes in the signal, it becomes impossible to resolve the frequency components of the signal, which are close in frequency within the analysis windows. On the other hand, if the time window is made long enough to permit good frequency resolution, it becomes difficult to determine where, in time, the various frequency components act.

An alternative method to overcome the resolution problem is to use the wavelet transform (WT) to get a better TFR of non-stationary signals.

8. Independent Component Analysis(ICA) and Principal Component Analysis (PCA):

ICA is a blind source separation technique that assumes the observed data is a linear combination of statistically independent sources. In the context of PCG denoising, it can be applied to separate the underlying independent cardiac sound sources from the noise.

The equation for ICA-based denoising can be represented as follows:

$$X = AS + N$$

Where:

- X is the observed PCG signal (mixtures) as a matrix of size M x N, where M is the number of sensor channels (recorded mixtures) and N is the number of samples.
- A is the mixing matrix of size M x M, which represents the linear mixing process.
- S is the source matrix of size M x N, which represents the independent source signals (cardiac sounds) to be extracted.
- N is the noise matrix of size M x N, which represents the noise components.

The goal of ICA is to estimate the unmixing matrix W (inverse of A) such that:

$$S = W * X$$

In this way, we can obtain the estimated independent source signals S, which ideally contain the denoised PCG components.

Whereas, PCA is a dimensionality reduction technique that finds the principal components in the data, which are orthogonal to each other and represent the directions of maximum variance. In the context of PCG denoising, PCA can be used to reduce the dimensionality of the data while preserving the most important information.

The equation for PCA-based denoising involves computing the principal components and projecting the data onto the first few principal components:

- Compute the mean-centered data matrix X' of size M x N.
- Calculate the covariance matrix C = X' * X'^T.
- Compute the eigenvectors and eigenvalues of C.
- Sort the eigenvectors in descending order of their corresponding eigenvalues.
- Select the first 'k' eigenvectors to form the projection matrix P of size M x k, where 'k' is the desired reduced dimensionality.
- Project the data onto the reduced subspace: Y = P^T * X'.

The resulting matrix Y contains the denoised PCG signals in the reduced dimensional space.

Both ICA and PCA offer effective ways to denoise PCG signals, but the choice between the two techniques depends on the specific characteristics of the noise and the signal. ICA is more suitable when the sources are assumed to be statistically independent, while PCA

is appropriate for dimensionality reduction when the noise is considered to be more correlated across different sensor channels. ICA and PCA, achieved very good results for SNR improvement and fHR determination, but higher accuracy was achieved using the ICA method. [19]

9. Adaptive filters - LMS, NLMS, RLS

Adaptive filters, which are self-learning filters changing their parameters depending on the change in the parameters of the input signal, are also very popular. These types of filters allow filtering interference from the useful signal if it changes its parameters over time or its parameters are not known in advance. The NLMS algorithm achieved better results in determining the HR according to the sensitivity and positive predictive value parameters, while the LMS algorithm achieved better results according to the SNR and root mean square error RMSE) parameters.

a. Least Mean Squares (LMS) Filter:

The LMS algorithm minimizes the mean square error between the desired output and the filter's output by adjusting its coefficients in the negative gradient direction of the error.

The weight update equation for the LMS filter is given by:

$$w(k+1) = w(k) + \mu * e(k) * x(k)$$

Where:

- w(k) is the vector of filter coefficients at iteration 'k'.
- μ is the step size, also known as the learning rate, which controls the convergence speed and stability of the algorithm.
- e(k) = d(k) - y(k) is the error signal, where d(k) is the desired output, and y(k) is the output of the filter.
- x(k) is the input vector at iteration 'k', usually a window of past input samples.

b. Normalized LMS (NLMS) Filter:

The NLMS algorithm is a variation of the LMS algorithm that normalizes the weight update to make it less sensitive to the scale of the input vector.

The weight update equation for the NLMS filter is given by:

$$w(k+1) = w(k) + (\mu / (\alpha + ||x(k)||^2)) * e(k) * x(k)$$

Where:

- w(k), μ, e(k), and x(k) are the same as in the LMS filter.
- α is a small positive constant added to the denominator to prevent division by zero.

c. Recursive Least Squares (RLS) Filter:

The RLS algorithm computes the filter coefficients recursively by minimizing the weighted sum of the past squared errors.

The weight update equation for the RLS filter is given by:

$$w(k+1) = w(k) + K(k) * e(k)$$

Where:

- w(k) is the vector of filter coefficients at iteration 'k'.
- K(k) is the Kalman gain matrix at iteration 'k', calculated as:
- $K(k) = P(k) * x(k) / (\lambda + x^T(k) * P(k) * x(k))$

where P(k) is the inverse of the autocorrelation matrix of the input vector up to iteration 'k', and λ is a small positive constant added to the denominator to ensure numerical stability.

- e(k) = d(k) - y(k) is the error signal, where d(k) is the desired output, and y(k) is the output of the filter.
- x(k) is the input vector at iteration 'k', usually a window of past input samples.

10. Efficient version of Fast Independent Component Analysis(EFICA)[20]:

It's a method based on single-channel separation consisting of three steps-

a. EMD
b. SVD
c. ICA

Very good results were obtained in filtering using this method for the PCG signals.

11. Variational Mode Decomposition (VMD)[21]:

Variational Mode Decomposition (VMD) is a data-driven technique used for adaptively decomposing a signal into a set of modes or components with varying frequency content. It is particularly useful for analyzing signals with non-stationary and nonlinear behavior, such as phonocardiogram (PCG) signals.

The VMD algorithm decomposes a given signal x(t) into K modes, each represented by a mode function mk(t) and an associated frequency ωk, such that:

$$x(t) \approx \sum(Ak * mk(t) * \cos(2\pi * \omega k * t + \varphi k))$$

Where:

- x(t) is the original signal.
- Ak is the amplitude of mode k.
- mk(t) is the mode function, representing the oscillatory behavior of mode k.
- ωk is the instantaneous frequency of mode k at time t.
- φk is the phase of mode k at time t.

The goal of VMD is to find the mode functions mk(t), frequencies ωk, and amplitudes Ak that best represent the signal x(t) and allow for an approximate reconstruction of the original signal.

VMD is formulated as a variational optimization problem. The decomposition is obtained by minimizing the following cost function with respect to the mode functions mk(t) and frequencies ωk:

$$J = \Sigma(||x(t) - \sum(Ak * mk(t) * \cos(2\pi * \omega k * t + \varphi k))||^2) + \Sigma(\mu * ||\partial t^p mk(t)||^2)$$

Where:

- The first term represents the fidelity to the original signal, ensuring that the sum of the reconstructed modes closely approximates x(t).
- The second term is a regularization term that enforces smoothness and ensures that each mode is locally band-limited with p being the order of the derivative and μ a regularization parameter.

The optimization problem can be solved using various numerical methods like gradient descent or alternating least squares. By iteratively updating the mode functions and frequencies, the VMD algorithm efficiently decomposes the signal into different modes

with varying frequency content, which can be useful for denoising, feature extraction, and time-frequency analysis.

In this [21] paper, 7 modes are decomposed. Among them, 3rd mode had the least error. So 3rd mode was used to calculate the Shannon energy.

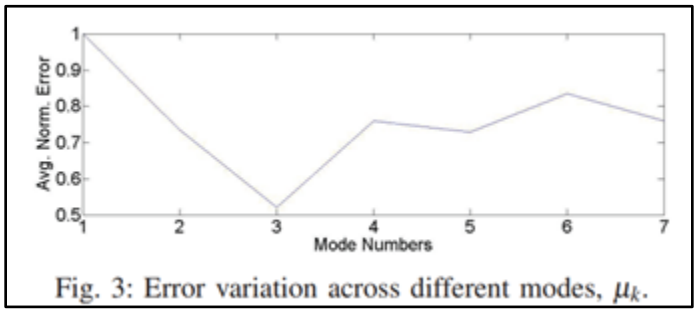


Fig. 3: Error variation across different modes, μ_k.

Figure 23: Error variation across different modes. [21]

Detection of Heart Sound

1. Autocorrelation + another model (WT, matching pursuit, or model-based individual correlation)
2. 4-step algorithm[48]
   - All sounds were localized using optimized S-transform
   - Detecting S1 and S2 using Shannon transform
   - Feature extraction using SVD
   - Classification using ANN
3. Duration-dependent hidden Markov model (DHMM) [49] - [50]
4. Gaussian regression[51]
5. Multiple steps[52] = WT + Hilbert phase envelope determination (to determine the boundaries)
6. EEMD with kurtosis feature [53]
7. Envelope detection algorithm using Shannon energy
8. combination of frequency filtering, energy detection, and interval regulation[56]
9. Envelope detector using Hilbert transform and interrelated supplementary processes[57]
10.

Metrics Used

The following metrics are used to judge the quality of the solution for denoising the PCG signal-

1. Correlation

2. Bland Altman Analysis
3. SNR,
4. ACC,
5. SE, and
6. PPV
7. mean error of heart interval measurement (|1Ti|), and
8. Harmonic mean between SE and PPV (F1).
9. Detection of S1 heart sounds:
   The detection was performed using the Pan-Tompkins algorithm
   - Noise cancellation
   - Signal is derived by a derivative filter, followed by squaring, and moving window integration.
   - The Pan-Tompkins decision rule, using adaptive thresholding, is applied.
   - Using TP, FP and FN values, it was possible to further determine the accuracy (ACC), the sensitivity (SE), the positive predictive value (PPV), and F1 score

References:

1. Recognition of heart sound based on distribution of Choi-Williams
2. New Parameter Available in Phonocardiogram for Blood Pressure Estimation
3. Pulse transition time in heamodynamic measurement: Methods, problems, hardware equipment and first results
4. What is Heart Rate Variability & How Can You Improve It with Velocity
5. An Overview of Heart Rate Variability Metrics and Norms
6. Heart rate variability
7. Calculating Blood Pressure Based on Measured Heart Sounds
8. Time-Frequency Analysis, Denoising, Compression, Segmentation, and Classification of PCG Signals
9. Computer aided analysis of phonocardiogram
10. Elimination of Interference in Phonocardiogram Signal Based on Wavelet Transform and Empirical Mode Decomposition
11. Improved Computerized Cardiac Auscultation by Discarding Artifact Contaminated PCG Signal Sub-sequence
12. Short unsegmented PCG classification based on ensemble classifier
13. Comparison of Speech Enhancement Algorithms
14. Spectral Subtraction Method Block Diagram
15. Speech Enhancement using Spectral Subtraction-type Algorithms: A Comparison and Simulation Study
16. A better way to define and describe Morlet wavelets for time-frequency analysis
17. Passive Fetal Monitoring by Advanced Signal Processing Methods in Fetal Phonocardiography
18. Time–frequency analysis of phonocardiogram signals using wavelet transform: a comparative study
19. Fetal phonocardiography signal processing from abdominal records by non-adaptive method
20. A Single channel Phonocardiograph Processing using EMD, SVD, and EFICA
21. Segmentation and detection of first and second heart sounds (Si and S2) using variational mode decomposition
22. Respiratory modulation of heart sound morphology
23. Wearable Cardiorespiratory Monitoring Employing a Multimodal Digital Patch Stethoscope: Estimation of ECG, PEP, LVET and Respiration Using a 55 mm Single-Lead ECG and Phonocardiogram
24. Oxygen saturation and heart rate in children at high altitude. A different response of Aymaras and non-Aymaras with chronic exposure at 3500 m
25. Inverse association between blood pressure and pulse oximetry accuracy: an observational study in patients with suspected or confirmed COVID-19infection