A Method for the Estimation of the Timing of Heart Sound Components Through Blind Source Separation in Multi-Source Phonocardiography

Noemi Giordano
Department of Electronics and Telecommunications
Politecnico di Torino
Torino, Italy
noemi.giordano@polito.it

Marco Knaflitz
Department of Electronics and Telecommunications
Politecnico di Torino
Torino, Italy
marco.knaflitz@polito.it

***Abstract*—Recently, phonocardiography (PCG) has gained importance as a diagnostic tool for cardiovascular diseases. In particular, the measurement of the time of occurrence of heart sounds may be of interest, in the clinical context, for the analysis of the electromechanical coupling of the heart. To date, though, there is no standardization concerning the positioning of the microphone probe over the chest, and this causes low accuracy and consistency in the measured timing values. Multi-source phonocardiography is a promising approach to face the stated issue. In this work, we present a methodology to estimate the latency of the components of the two main heart sounds towards the corresponding R-wave peak based on the Blind Source Separation (BSS) of the contributions of the left and right side of the heart. We tested our algorithm on a sample population of 12 subjects over 10-minute long recordings of three simultaneous PCG signals and one electrocardiographic (ECG) signal for reference. Results show that the approach is robust with respect to the usage of different algorithms to perform BSS (FastICA, JADE). The measured timing values are consistent with what measured by means of a single-source algorithm we previously developed. This methodology looks promising in terms of obtaining accurate measurements of the time of occurrence of heart sound components and may have an impact in the clinical context.**

Keywords—heart sounds, multi-source phonocardiography, Blind Source Separation, Independent Component Analysis

I. INTRODUCTION

Phonocardiography (PCG) is the electronic recording, by means of a microphone transducer, of the acoustic waves generated by the closing of cardiac valves. These sounds are expression of the mechanical behavior of the heart. Even though auscultation is a routine screening tool in the clinical practice, it relies on an experienced examiner. This leads to a lack of objectivity and to the impossibility of performing accurate measurements of some heart features, such as the latency of heart sounds with respect to the corresponding ventricular depolarization. Digital phonocardiography aims at overcoming the stated limitations and at providing a quantitative analysis of heart sounds.

The measurement of the timing of occurrence of heart sounds is expected to have a clinical impact. Indeed, the latency of heart sounds with respect to the corresponding R-wave, in a simultaneous electrocardiographic (ECG) recording, is correlated to the electromechanical coupling of the heart. A quantitative insight in the latter may provide valuable indications for a range of pathologies, such as heart valve diseases and heart failure, with a potential clinical impact in the near future [1].

Nevertheless, the relative novelty of this topic introduces a number of critical issues that need to be dealt with. One concerns the localization of the microphone on the chest of the subject. It was proved that the localization of the microphone affects the measure of heart sounds latency with respect to the corresponding R-wave [2]. At this time, no standardization exists on the microphone positioning: this makes it difficult to give a clinical interpretation of the measured latency values.

This work deals with the unavailability of a reference for the microphone localization in PCG recordings, which leads to uncertainties in the measure of the timing of heart sounds latencies. In particular, there is a clinical interest in evaluating the latency of heart sounds components, namely mitral and tricuspid components for the first, and aortic and pulmonary components for the second heart sound.

Phonocardiography is a noninvasive technique. Therefore, the PCG signal is not a trace of the actual acoustic waves as generated by the cardiac valves, but a trace of the cardiac acoustic waves when they reach the chest wall, after propagation across the biological tissues. The measurement of the timing of heart sounds should be the closest possible to what we would measure if we could record the signal as it is generated by its source.

In this context, a Blind Source Separation (BSS) approach looks promising. Blind Source Separation is commonly used in acoustics to identify and separate a number of source signals from a set of mixtures acquired at different spatial locations. The same approach could be applied in multi-source phonocardiography to separate the acoustic signals generated by different cardiac structures. On the separated source signals, a measurement of heart sounds features, such as their latency towards corresponding R-wave peaks, can be accurately performed.

Consequently, the application of Blind Source Separation techniques to multi-source PCG recordings is expected to overcome the main critical issue of single-source phonocardiography, i.e. the positioning of the microphone over the subject chest. In fact, the estimation of the latency of heart sound components over the separated source signals generated by different cardiac structures is expected to be more robust and accurate with respect to its equivalent measured in single-source PCG. This would allow for a more reliable interpretation of the latency values in the clinical context.

The aim of this work is to propose an automated algorithm to separate source signals in multi-source phonocardiography through Blind Source Separation techniques and evaluate the

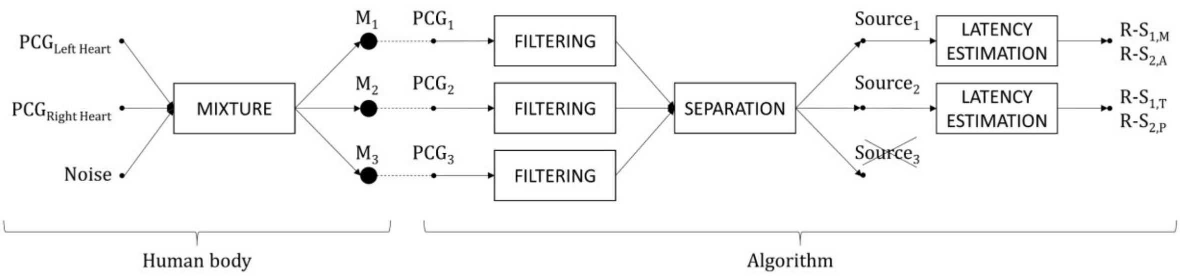

Fig. 1. Block diagram of the mixture process and of the developed algorithm for source separation and latency estimation.

time of occurrence of the two components of the two main heart sounds within the source signals.

II. Materials and Methods

A. Rationale

The described methodology is applied to multi-source phonocardiography recordings including three PCG signals and one simultaneous electrocardiogram (ECG) used as reference.

Given the availability of three mixtures, recorded at different spatial positions with respect to the acoustic sources, Blind Source Separation is capable of separating up to three independent source signals.

Here, we apply BSS with the scope of separating the following three signals:

1) Contribution of the right side of the heart.
2) Contribution of the left side of the heart.
3) Contribution of noise.

The reason why we believe we can separate the latter sources resides in the physiology of the cardiac cycle and, specifically, in the mechanical structure and behavior of the heart. Indeed, right and left side of the heart can be considered as independent from a mechanical point of view, since the blood flowing in them derives from different circulation circuits. Because the opening and closing of the valves is passive and only depends on the pressure gradients between cardiac chambers on the same side of the heart, the contributions of the two sides on the recorded PCG signal can be considered as independent [3].

After having identified three different source signals by means of a Blind Source Separation technique, we select the two sources corresponding to the cardiac contributions through a Signal-to Noise Ratio assessment.

In the end, we use an adaptation of an algorithm we previously developed [4] for the segmentation and classification of heart sounds within the source signals and the estimation of their latency with respect to the corresponding R-wave peak.

Fig. 1 presents the block diagram of the mixture process of the contributions of interest inside the human body and of the developed algorithm for the separation of the sources and the estimation of the timing of heart sounds.

B. State of the Art

The application of Blind Source Separation to physiological multi-source acoustic signals has been proposed in the literature mainly with two objectives:

1) the separation of lung and heart sounds.
2) the separation of fetal and maternal heart sounds.

Few works are available concerning the usage of BSS methods in multi-source phonocardiography with the scope of distinguishing sounds generated by different cardiac structures.

Nigam et al. [5] applied BSS to second heart sound segments to separate aortic and pulmonary components and measure the split. Geethu et al. [6] proposed an algorithm to separate through BSS the contribute of six generic cardiac structures. Tong et al. [7] applied BSS on delayed versions of the same PCG signals with the scope of separating contributes from different cardiac structures. Yang et al. [8] used Independent Component Analysis, a typical BSS approach, to separate cardiac murmurs from the two main heart sounds.

To our best knowledge, no state-of-the-art method allows for separating the contribution of the right and left side of the heart to measure the timing of the components of each heart sound.

C. Acquisition System

The methodology described in this work was applied to simultaneous recordings of one ECG signal, used as a time reference, and three PCG signals.

The recording relied on the commercial 4-channel acquisition system for biomedical signals ReMotus® (ItMeD, Italy). This system simultaneously samples the four channels with a sampling frequency of 1 kSa/s and converts them through a 24-bit A/D converter. The system guarantees a 3-dB bandwidth from DC to 262 Hz, which is suitable for the recording of ECG and PCG. ReMotus® is fully isolated for patient safety (CF type) and connected via USB to a computer for the visualization, processing and storage of the acquired signals.

The first channel of the acquisition system records ECG signal, by using three disposable Ag/AgCl electrodes connected to the acquisition system through an active probe with 1 GΩ input impedance and 0 dB gain. The electrodes are located on the subject chest as to recreate the I standard lead, as shown by Fig. 2.

The remaining three channels are connected to custom made wired microphone probes to perform PCG recording. Each probe is composed of a condenser microphone, the

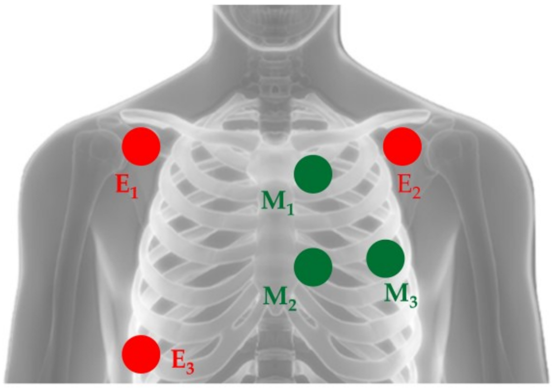

Fig. 2. Positioning of eletctrodes (E_1, E_2 and E_3) and microphone probes (M_1, M_2 and M_3) over the subject chest.

signal conditioning circuit, and a 3D-printed stethoscope-shaped plastic container. The frequency response of the probes spans from 2 Hz to 250 Hz, suitable for the recording of the two main heart sounds.

The location of the microphone probes over the subject chest is fixed referring to anatomical references and is inspired by the traditional auscultation areas. Fig. 2 shows that two microphones are located along the sternum, respectively over the second and fifth intercostal spaces, and the third is located over the fifth intercostal space 70 mm from the sternum.

D. Pre-Processing

First, R-wave peaks need to be located within the ECG recording. R-wave peaks represent the electrical depolarization of the ventricles and serve as a reference of the beginning of a cardiac cycle.

To this purpose, we applied to the ECG signal a modified version of Pan-Tompkins algorithm [9]. First, we applied the cascade of a median filter for baseline wandering removal with 100-ms moving window and a low-pass FIR filter with cut-off frequency at 35 Hz. Then the filtered signal is subjected to a five-point derivative filter, point-wise squaring, and moving average integration with a 150-ms window. R-peaks are located within the final signal through a two-step thresholding process.

PCG recordings are prepared for Blind Source Separation by means of a pre-filtering stage. The scope is to improve the quality of the input signals, and thus the capability of BSS techniques of distinguishing the signal sources of interest. In fact, a wide range of artifacts, generated by the environment, the subject or their interactions, usually corrupts the PCG signal. Filtering is carried out by a digital band-pass filter (IIR Chebyshev) with cut-off frequencies at 20 Hz and 100 Hz.

It should be highlighted that lung sounds, whose frequency content spans up to 2 Hz [10], are strongly attenuated in the pre-filtering step.

E. Blind Source Separation

Blind Source Separation (BSS) techniques aim at identifying the impinging contributions present in a multi-source recorded signal without any information about the sources, nor the mixture mechanism, nor the geometry of the recording system [11]. The sole assumption of the BSS approach is that the source signals are independent: no further information is needed. The issue is well known in the acoustics field under the name of "cocktail party problem" [11].

From a mathematical point of view, given the signal matrix $S = [s_1(t), \ldots, s_N(t)]^T$, where $s_i(t)$ is the signal emitted by the i-th source, and a noise contribution $n(t)$, the mixture matrix $X = [x_1(t), \ldots, x_M(t)]^T$, where $x_j(t)$ is the signal recorded by the j-th microphone, can be expressed as in:

$$X(t) = AS(t) + n(t) \tag{1}$$

where A is an M-by-N matrix having the role of a spatial transfer function between sources and microphones. It is usually referred to as *mixing matrix*.

Assuming that the noise contribution can be removed in a pre-filtering stage, the goal of BSS techniques is to estimate a separation N-by-M matrix W, such as:

$$\hat{S}(t) = WX(t) \tag{2}$$

where $\hat{S}$ is an estimation of the signal matrix. W is usually referred to as *unmixing matrix* and approximates the inverse of the mixing matrix A.

Among the range of families of available techniques, Independent Component Analysis is one of the most widely used. It is a statistical technique aimed at transforming the available signals into a suitable statistical domain, where components are as independent as possible from each other. It works under the assumption of non-gaussianity and mutual independence of the source signals of interest.

Different ICA algorithms mainly differ on the measure of independence which is maximized to try and decorrelate the signals. Fourth-order statistical moments (kurtosis) are most often chosen.

In this work, we tested two algorithms for Independent Component Analysis, namely:

1) FastICA. It is the most popular algorithm for ICA because of its efficiency. It has an iterative scheme that searches for an orthogonal rotation of the components that maximizes a measure of non-gaussianity (kurtosis in our work) [12].

2) Joint Approximation of Diagonalization of Eigen-matrices (JADE). It is another ICA algorithm which proved to have good performances in the latest years. It is based on the construction and diagonalization of fourth-order cumulant tensor from data [13], [14].

The rationale for testing two ICA approaches is to evaluate the robustness of the overall algorithm with respect to the BSS methodology.

It should be highlighted that, before applying the ICA algorithms, data pre-whitening is performed to make the signals compliant with the requirements of the algorithms.

F. Detection and Timing of Heart Sounds

As anticipated, the output of the BSS phase are three source signals, which we expect to represent respectively the contribution of the left heart, the contribution of the right heart and the contribution of noise.

The contribution of noise is the source signal with lowest average Signal-to-Noise Ratio (SNR). We define SNR of a PCG signal as:

$$SNR = 20 \ \log \frac{A_S}{4 \ \sigma_N} \quad (3)$$

where A_S is a measure of the signal amplitude and is defined as the peak-to-peak value of the heart sound of interest, whereas $4 \ \sigma_N$ is a measure of the noise amplitude and is defined as the width of the 95% band of the noise time series.

Since the noise contribution is of no use for further processing, the source signal with the lowest average SNR is discarded.

For the segmentation and classification of heart sounds within the source signals we used an adaptation of an algorithm we previously developed [4]. This algorithm relies on an envelope-based approach, where the envelope of the PCG signal is computed as the normalized point-wise second order Shannon energy (SE) of the signal, defined as:

$$SE = -\frac{1}{N} \sum_{i=1}^{N} x^2(i) \cdot \log x^2(i) \quad (4)$$

where N is the length of the moving integration window, set to 20 samples at a 1 kHz sampling frequency. It should be highlighted that the moving integration window slides one sample at a time, with the scope of maintaining a resolution of 1 ms in the SE signal. This is crucial to obtain the necessary time resolution in measuring hearth sound component latencies.

The detection of the two main heart sounds is performed on the SE signal throughout a double thresholding process. In particular, heart sounds are segmented by applying an adaptive threshold to the SE of 1-second long signal windows. The SE peak is labelled within each segment as an indicator of the heart sound position. Differently from our previously developed algorithm, which was optimized to detect both components of each heart sound in the same recording, here only the highest SE peak is considered.

In a further stage, heart sounds are classified as S_1 (first heart sound) or S_2 (second heart sound), depending on their temporal relationships. Heart sounds occurring between the R-peak and the 18% of the corresponding RR-interval are classified as S_1, whereas heart sounds occurring after the 18% of the corresponding RR-interval are classified as S_2. RR-interval is the time interval between two consecutive R-wave peaks. This rule is based on cardiac physiology and applies to normal as well as to pathological subjects. Fig. 3 shows an example of the outcome of this phase on two cardiac cycles.

At this point, each of the two source signals is labelled as S_1 or S_2. To determine which source signal represents the acoustic expression of which side of the heart, we compute the mean latency of each heart sound with respect to the corresponding R-wave peak. Physiologically, mitral valve closes slightly before tricuspid valve. The same happens between aortic and pulmonary valves [3]. Therefore, the source signal presenting the lowest latency values is associated to the left side of the heart, the other one to the right side of the heart.

In the end, the following timing parameters are defined:

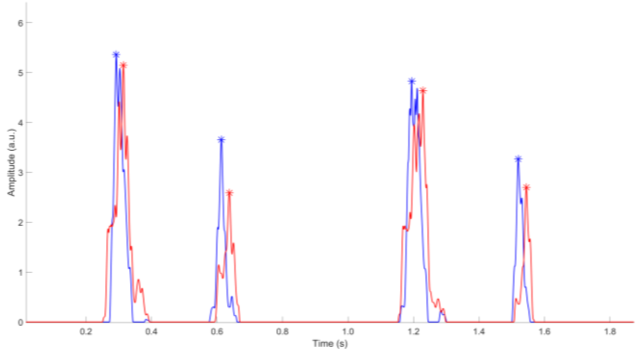

Fig. 3. Example of the outcome of the heart sounds detection phase on the source signals of two cardiac cycles of a subject. The plot shows the Shannon Energy of the signals, the stars represent the labelled heart sounds.

1) R-$S_{1,M}$: the delay between the first heart sound in the source representing the left heart and the corresponding R-peak (mitral component of S_1).

2) R-$S_{1,T}$: the delay between the first heart sound in the source representing the right heart and the corresponding R-peak (tricuspid component of S_1).

3) R-$S_{2,A}$: the delay between the second heart sound in the source representing the left heart and the corresponding R-peak (aortic component of S_2).

4) R-$S_{2,P}$: the delay between the second heart sound in the source representing the right heart and the corresponding R-peak (pulmonary component of S_2).

G. Sample Population

The algorithm was tested on a sample population counting 12 volunteers with no history of cardiopathy. All recordings lasted 10 minutes.

Methods and extent of the study were thoroughly explained to participants, who gave their assent and signed an informed consent form. The study was observational, it could neither modify subjects' health status nor expose subjects to any danger, and it conformed to the Helsinki declaration.

Because of the high variability of the quality of PCG signals as a consequence of the position of the microphones over the subject chest [15], obtaining PCG signals of good quality in all positions is not guaranteed. In this work, we considered only recordings presenting an SNR higher than 10 dB in at least 2 out of 3 PCG signals.

III. Results and Discussion

The algorithm was tested in both versions, i.e. using either FastICA algorithm or JADE algorithm for performing the Blind Source Separation phase, on the sample population. For each subject, we computed the time of occurrence of each component of the two main heart sounds with respect to the R-wave peak, thus obtaining four latency parameters.

Table I reports the values of the four parameters for each subject estimated by:

1) FastICA: the algorithm described in this work, using FastICA algorithm to perform BSS.

2) JADE: the algorithm described in this work, using JADE algorithm to perform BSS.

TABLE I. VALUE OF THE ESTIMATED LATENCY FOR ALL PARAMETERS WITH THE THREE APPROACHES

Subj	$R\text{-}S_{1,M}$			$R\text{-}S_{1,T}$			$R\text{-}S_{2,A}$			$R\text{-}S_{2,P}$		
	FastICA	*JADE*	*Ref*	*FastICA*	*JADE*	*Ref*	*FastICA*	*JADE*	*Ref*	*FastICA*	*JADE*	*Ref*
01	55	55	48	74	74	61	381	377	370	399	399	384
02	40	40	40	58	59	61	385	383	380	413	400	392
03	56	59	54	84	78	77	393	394	392	413	413	413
04	49	48	50	68	63	70	398	397	399	445	424	448
05	46	49	46	65	66	62	362	362	361	371	381	391
06	65	65	52	81	80	86	421	426	409	436	432	439
07	45	44	54	64	69	71	416	399	395	432	416	423
08	49	37	18	53	57	72	375	376	372	380	402	408
09	42	43	59	67	65	94	349	346	343	359	360	386
10	31	30	37	59	56	51	362	362	355	389	390	389
11	30	30	33	55	75	78	370	373	380	388	384	398
12	50	50	42	57	56	76	356	356	355	383	382	369
ANOVA p-value	0.89			0.30			0.87			0.89		
ANOVA F crit	3.28			3.28			3.28			3.28		
ANOVA F	0.12			1.23			0.15			0.12		

3) Ref: the algorithm described in our previous work [4], applied to the single PCG signal with the highest SNR.

We statistically evaluated our results through Analysis of Variance (ANOVA). The scope of the analysis is two-sided. On one side, to assess the robustness of our methodology with respect to variations in the Blind Source Separation algorithm, which could lead to slightly different source signals. On the other side, to validate the approach by comparing it to an existing method grounding on completely different bases.

As it can be observed in Table I, with $\alpha = 0.05$, the p-value obtained for all parameters is much higher than the significance level. In particular, a p-value of 0.32 was found for the tricuspid component of S_1, whereas a p-value of almost 0.9 was found for the remaining components.

These p-values show that the application of either FastICA algorithm or JADE algorithm in the BSS phase does not lead to significantly different results from a statistical point of view. This validates the robustness of the methodology with respect to the source separation approach.

Moreover, high p-values in the ANOVA analysis suggest that the results of our novel approach are statistically consistent with the results of our previously developed one.

When comparing the actual latency values in Table I, it can be observed that the highest difference, even though not statistically significant, is found with respect to the *Ref* column, i.e. between the method based on BSS and the single-source method. This is not surprising, because the goal of the approach based on BSS is to overcome the limit of single-source phonocardiography, i.e. the inconsistency among latency values measured at different points of the chest [2]. For this reason, the small differences among the BSS-based algorithms and the single-source approach do have a physical meaning.

IV. CONCLUSIONS

In this work, we proposed an automated approach for identifying and separating the contributions of the right and left side of the heart in multi-source phonocardiography recordings. On the separated source signals, we estimated the latency of the main components of first and second heart sounds with respect to the peak of the R wave. These latencies are of potential interest in the clinical context for the analysis of the electromechanical coupling of the heart.

We tested our methodology on a sample population of 12 healthy subjects. In particular, we tried out two different algorithms for Blind Source Separation, namely FastICA and JADE. Results show that no statistically significant difference exists between the techniques, validating the robustness of the approach with respect to the source separation phase.

We compared the results of our proposed BSS-based approach with the results obtained by applying a single-source algorithm for the timing of heart sounds component we previously developed. The results are consistent at a significance level of 0.05. Small differences among the BSS-based multi-source approach and the single-source approach may depend on the theoretical higher accuracy of the first. This is because the proposed BSS-based approach allows for estimating the latency directly on the acoustic signals as generated by the cardiac structures, and therefore reduces the uncertainty due to the positioning of the digital stethoscope over the chest.

In the overall, the outcomes of this work are promising for obtaining a more accurate measurement of the time of occurrence of heart sound components compared to state-of-the-art approaches.

REFERENCES

[1] Y. N. Wen, A. P. W. Lee, F. Fang, C. N. Jin, and C. M. Yu, "Beyond auscultation: Acoustic cardiography in clinical practice," *Int. J. Cardiol.*, vol. 172, no. 3, pp. 548–560, 2014.

[2] J. Fontecave-Jallon, K. Fojtik, and B. Rivet, "Is there an Optimal Localization of Cardio-microphone Sensors for Phonocardiogram Analysis?," *2019 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 3249–3252, 2019.

[3] A. C. Guyton and J. E. Hall, *Textbook of Medical Physiology*, 11th ed. Pennsylvania: Elsevier Health Sciences, 2015.

[4] N. Giordano and M. Knaflitz, "A novel method for measuring the timing of heart sound components through digital phonocardiography," *Sensors (Switzerland)*, vol. 19, no. 8, 2019.

[5] V. Nigam and R. Priemer, "A dynamic method to estimate the time split between the A2 and P2 components of the S2 heart sound," *Physiol. Meas.*, vol. 27, no. 7, pp. 553–567, 2006.

[6] R. S. Geethu, M. Krishnakumar, K. V. Pramod, and S. N. George, "Source Separation of Heartbeat Sounds for Effective E-Auscultation," *J. Inst. Eng. Ser. B*, vol. 97, no. 1, pp. 69–75, 2016.

[7] Z. Tong, I. A. Qader, and F. Abu-Amara, "Heart Sound Separation Using Fast Independent Component Analysis," *Proc. - 2015 Int. Conf. Dev. eSystems Eng. DeSE 2015*, pp. 3–6, 2016.

[8] X. Yang *et al.*, "Analyzing the phonocardiogram of ASD-patient based on independent component analysis," *Int. Conf. Signal Process. Proceedings, ICSP*, vol. 1, pp. 6–9, 2006.

[9] J. Pan and W. J. Tompkins, "A Real-Time QRS Detection Algorithm," *IEEE Trans. Biomed. Eng.*, vol. BME-32, no. 3, pp. 230–236, 1985.

[10] V. Gross, A. Dittmar, T. Penzel, F. Schüttler, and P. Von Wichert, "The relationship between normal lung sounds, age, and gender," *Am. J. Respir. Crit. Care Med.*, vol. 162, no. 3 I, pp. 905–909, 2000.

[11] Y. min Qian, C. Weng, X. kai Chang, S. Wang, and D. Yu, "Past review, current progress, and challenges ahead on the cocktail party problem," *Front. Inf. Technol. Electron. Eng.*, vol. 19, no. 1, pp. 40–63, 2018.

[12] A. Hyvärinen and E. Oja, "Indepent Component Analysis: Algorithms and Applications," *Neural networks*, vol. 13, no. 4–5, pp. 411–430, 2000.

[13] D. N. Rutledge and D. Jouan-Rimbaud Bouveresse, "Independent Components Analysis with the JADE algorithm," *TrAC - Trends Anal. Chem.*, vol. 50, pp. 22–32, 2013.

[14] J. F. Cardoso and A. Souloumiac, "Blind beamforming for non-Gaussian signals," *IEE Proceedings, Part F: Radar and Signal Processing*, vol. 140, no. 6. pp. 362–370, 1993.

[15] N. Giordano and M. Knaflitz, " Multi-source signal processing in phonocardiography: comparison among signal selection and signal enhancement techniques * ," *2019 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 6689–6692, 2019.