Respiratory Sound Analysis for Flow Estimation During Wakefulness and Sleep, and its Applications for Sleep Apnea Detection and Monitoring

by

Azadeh Yadollahi

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Electrical and Computer Engineering
University of Manitoba

Winnipeg

To My Family,

for their support, patience and kindness.

This research was financially supported by TRLabs, Winnipeg.

Contents

Abbreviations

AASM: American Academy of Sleep Medicine

AHI: Apneahypopnea index

ANOVA: Analysis of variance

ASAD: Acoustical sleep apnea diagnosis

AUC: Area under the curve

BMI: Body mass index

COPD: Chronic obstructive pulmonary disease

CSA: Central sleep apnea

Exp: Expiration

Insp: Inspiration

L: Logarithm of the difference between maximum and minimum values of the signal

L_r: logarithm of the difference between the average values of the upper and lower $r\%$ of the signal

L_{Avg}: Logarithm of the average power of the signal

L_{std}: Logarithm of the standard deviation of the signal

LogVar: Logarithm of the tracheal sound variance

OSA: Obstructive sleep apnea

pdf: Probability density function

PSG: Polysomnography

RIP: Respiratory inductance plethysmography

ROC: Receiver operating curve

List of Figures

List of Tables

Abstract

Tracheal respiratory sounds analysis has been investigated as a non–invasive method to estimate respiratory flow and upper airway obstruction. However, the flow–sound relationship is highly variable among subjects which makes it challenging to estimate flow in general applications. Therefore, a robust model for acoustical flow estimation in a large group of individuals did not exist before. On the other hand, a major application of acoustical flow estimation is to detect flow limitations in patients with obstructive sleep apnea (OSA) during sleep. However, previously the flow–sound relationship was only investigated during wakefulness among healthy individuals. Therefore, it was necessary to examine the flow–sound relationship during sleep in OSA patients.

This thesis takes the above challenges and offers innovative solutions. First, a modified linear flow–sound model was proposed to estimate respiratory flow from tracheal sounds. To remove the individual based calibration process, the statistical correlation between the model parameters and anthropometric features of 93 healthy volunteers was investigated. The results show that gender, height and smoking are the most significant factors that

affect the model parameters. Hence, a general acoustical flow estimation model was proposed for people with similar height and gender.

Second, flow–sound relationship during sleep and wakefulness was studied among 13 OSA patients. The results show that during sleep and wakefulness, flow–sound relationship follows a power law, but with different parameters. Therefore, for acoustical flow estimation during sleep, the model parameters should be extracted from sleep data to have small errors. The results confirm reliability of the acoustical flow estimation for investigating flow variations during both sleep and wakefulness.

Finally, a new method for sleep apnea detection and monitoring was developed, which only requires recording the tracheal sounds and the blood's oxygen saturation level (S_aO_2) data. It automatically classifies the sound segments into breath, snore and noise. A weighted average of features extracted from sound segments and S_aO_2 signal was used to detect apnea and hypopnea events. The performance of the proposed approach was evaluated on the data of 66 patients. The results show high correlation $(0.96, p < 0.0001)$ between the outcomes of our system and those of the polysomnography. Also, sensitivity and specificity of the proposed method in differentiating simple snorers from OSA patients were found to be more than 91%. These results are superior or comparable with the existing commercialized sleep apnea portable monitors.

Acknowledgments

This dissertation would not have been possible without the guidance and the help of several individuals who in one way or another contributed in the preparation and completion of this study.

First and foremost, my utmost gratitude to my supervisor, Prof. Zahra Moussavi, whose encouragement, supervision and support enabled me to complete my research. She was not only a great academic mentor to guide me through my studies and research, but also her help and support in my personal life was invaluable.

This study was financially supported by TeleCommunication Research Labs (TRLabs), Canada. I thank the administration of TRLabs, especially Dr. Sergio Camorlinga.

It is a pleasure to thank Prof. Attahiru Sule Alfa, Prof. Newman Stephens, Dr. Magdy Younes and Dr. Ellini Gianouli for their valuable inputs and comments. I also wish to thank Prof. Neil Popplewell for his helpful discussions.

I would like to acknowledge and thank Wayne Thompson, Health Sciences Center Sleep Disorders Clinic staff and Misericordia Health Center Sleep Center staff for their invaluable help and assistance in clinical data recording.

I would like to show my gratitude to my colleagues Saiful Huq, Aman Montazeri and Ali Azarbarzin for their priceless help in recording data from volunteers and patients, especially in the brutal winter nights of Winnipeg.

I am indebted to the anonymous volunteers who patiently accepted to participate in the study and encouraged me to continue my research.

I owe my deep gratitude to my precious friends, Nazila Hafezi, Samaneh Sarraf, Mahmoud Azimaee, Fereshteh Amini, Hossein Pourreza, Shekoufeh Saboktakin, Behtash Babadi, Arash Abadpour and Maliheh Mokhtari. Their support and friendship enabled me to hurdle the difficult situations through my life and study.

It is an honor for me to thank my family members. No matter how far they are from me, without their continues support and kind wishes, I could not reach this point.

Last, but not least, I offer my regards and blessings to my friends in Winnipeg, and around the world. Their friendship is my most precious treasure.

Chapter 1

Introduction

This thesis investigates the flow-sound relationship in healthy individuals and patients with obstructive sleep apnea (OSA) during sleep and wakefulness. The premise of this research stems from our previous study [1] in acoustic flow estimation from tracheal breath sounds. In this thesis, the acoustic flow estimation method has been investigated in detail during sleep and wakefulness, its application on sleep apnea detection has been investigated and a technology using an acoustical method for monitoring and diagnosis of sleep apnea has been developed, patented and implemented.

This thesis has three major sections focused on flow–sound relationship in sleep and wakefulness and its application on sleep apnea detection. In chapter 2 the flow–sound relationship is investigated in healthy individuals during wakefulness. The model developed for acoustical flow estimation is discussed, and the the effects of different anthropometric features on the model parameters is examined. Furthermore, the possibility of replacing

the calibration process by using the average values of the model parameters in a matched population with similar anthropometric data with the test subject has been investigated. Flow–sound relationship during sleep and in patients with OSA is an important issue which has not been investigated before. This was investigated in this thesis and the results are presented in chapter 3.

One of the most important applications of acoustical flow estimation is sleep apnea detection. In chapter 4, we report on the development of an acoustical method for sleep apnea diagnosis and monitoring. The method uses the results of the proposed flow estimation algorithm to estimate variations in flow amplitude during sleep. The method was verified on data of 66 patients, who were referred to the sleep lab for full night Polysomnography (PSG) study. The results of acoustical OSA detection are reported and compared with those of PSG accepted as the gold standard. These results are discussed in Section 4.2.

Finally chapter 5 outlines the conclusions and future objectives of this research. Details of our previous research on acoustical flow estimation are presented in Appendix A. In this appendix, investigation of different features of tracheal sound and selection of the best feature are discussed. In Appendix B, we report on the development of a new method for automatic classification of breathing and snoring sound segments based on the energy, zero crossing rate and formants of the sound signals. These results can be used in future to improve the performance of our developed system for acoustical sleep apnea diagnosis.

1.1 Background

Respiratory flow measurement is usually essential for respiratory sound assessment and recording, and its relationship with respiratory sounds is of great importance for researchers and physicians. Airflow is usually recorded by spirometry devices such as pneumotachograph, nasal cannulae connected to a pressure transducer, heated thermistor or anemometry. Airflow can be also measured indirectly by means of detecting chest or abdominal movements using respiratory inductance plethysmography (RIP), strain gauges or magnetometers [2]. Application of these techniques has some disadvantages, such as changing the breathing pattern [3-6], difficulties in implementing them in children with neurological impairments or inaccuracies due to misplacement of the sensor [7].

Tracheal respiratory sound analysis is a simple and non–invasive way to study the pathophysiology of the upper airways. It has been used for acoustical flow estimation [1, 8-10] and investigation of the upper airways abnormalities such as wheezes [11-13], tracheal stenosis [14], and airway obstructions [15-18]. However, since multiple sources contribute to tracheal sound generation, the sound characteristics and flow–sound relationship are highly variable between individuals. This makes the use of a general flow–sound model for accurate flow estimation in a large population of subjects very challenging.

Tracheal breath sounds are generated by the passage of turbulent air in the upper airways. Different mechanisms including turbulence of air flow, jet formation and pressure fluctuations in the upper airway walls contribute to the production of sounds and

fluid-induced vibrations [19, 20]. The vibrations are transmitted to the skin through the tracheal wall and tissue beneath the skin, and can be picked up by a microphone placed over trachea. Normal tracheal breath sound has a broad–band spectrum with several peaks; it has been shown that shape and peaks of the spectral curve change with the geometry and pathology of the upper airway [19-24], while its amplitude and energy change with the amount of breathing flow [1, 9, 10, 19, 25].

The flow–sound relationship has been studied extensively in the literature, and the effects of flow rate on various temporal and spectral features of tracheal sounds were examined. The main features include average power [9, 10, 25-28], sound's envelope [8], entropy [29] and mean frequencies [30, 31]. In one of our previous studies, we investigated different tracheal sounds features, and found the logarithm of sound variance ($LogVar$) a robust feature for flow estimation [1].

Inconsistency of flow–sound relationship between individuals imposes the need for calibration of flow–sound model parameters for every individual as the primary step for acoustic flow estimation. This has been the major drawback of acoustic flow estimation because calibration may not always be applicable easily, specifically when assessing young children, patients with neurological impairments and/or patients in emergency conditions. In most of the previous studies, it was assumed that a few cycles of breath sounds with known flow at each flow rate were available to derive the model coefficients [8, 10, 26-28]. In our previous study [1], a new method of flow estimation was proposed that simplified

the calibration routine significantly; it used only one cycle of breath sound at medium flow rate. The method proposed in [1] used a modified linear model for describing the relationship between flow and $LogVar$ of the tracheal sound at different flow rates. The model's parameters were derived from only one breath sound cycle with known flow at medium flow rate. The results showed that the model was able to follow the flow variation from very low to high flow rates with an error of less than 10%. Details of that research are presented in Apendix A.

The method developed in [1] removed the need for calibration at different flow rates, but did not eliminate it completely. Flow–sound relationship changed with individuals' anthropometric features. Although the underlying mechanisms are not clear, height, age and gender have been found to change the spectral features of respiratory sounds [22, 32]. In the first part of this thesis we pursued two goals to improve the performance of flow estimation algorithms. First, we investigated the significance of different anthropometric features on the flow–sound model parameters. Then, we examined whether it was possible to replace the calibration process by using the average values of the model parameters in a matched population (with similar anthropometric data) with a test subject. Hence, we recorded tracheal sound and flow signals from a large group of healthy individuals including both smokers and non–smokers. Then, we investigated the statistical effect of important anthropometric features of the participants such as height, weight, age, gender, etc. on the flow–sound model parameters. The statistical results were then used

to generalize the flow–sound model, and examine the flow estimation accuracy. The results are presented in Section 2.1.2.

One of the main applications of acoustical flow estimation is to detect the abnormalities of the upper airways and examine the obstructions in the upper airways during sleep. As mentioned, in previous studies the relationship between tracheal sound and flow were studied extensively [19-24, 1, 9, 10, 25]. However, these studies were all focused on data of healthy individuals during wakefulness. The relationship between tracheal sound and flow may change when one sleeps. The relationship may also be different in healthy individuals and OSA patients. Therefore, a crucial step in application of acoustical flow estimation for sleep studies is to have a detailed investigation of flow–sound relationship among OSA patients during sleep.

In the second part of this thesis, we investigated the flow–sound relationship during sleep. First, we examined the flow–sound relationship in patients with OSA and examined how the relationship in every individual changed from wakefulness to sleep. Then, the accuracy of acoustical flow estimation during sleep was studied. For this application, the flow–sound model parameters were estimated during sleep and wakefulness. Different calibration algorithms based on the sleep and wakefulness data were applied and the method's performance in flow estimation was examined. The results are presented in Section 3.2.

Sleep apnea is a common and serious respiratory disorder. It is defined as periods

of airflow cessation (apnea) or reduced airflow by more than 30% (hypopnea) associated with a minimum of 4% drop in blood's oxygen saturation level (S_aO_2) [33, 34]. There are three types of sleep apnea: Obstructive, central and mixed sleep apnea [34]. The most common one is obstructive sleep apnea (OSA), in which respiratory effort exists but there is no resulting respiratory airflow. Central sleep apnea (CSA) is less common, in which respiratory effort does not exist due to the dysfunction of central drive mechanisms and mixed apnea is a combination of both obstructive and central sleep apnea [34, 35]. The severity of sleep apnea is usually measured by apnea–hypopnea index (AHI) which shows the number of apnea and hypopnea events per hour, although the extent of oxygen desaturation and frequency of arousals or any cardiac arrhythmias that may occur as a result of the sleep apnea/hypopnea events are also indicators of sleep apnea severity [34]. OSA is highly prevalent in the general population, approaching about 24% of men and 9% of women aged $30 - 60$ years old with $AHI \geq 5$, while the prevalence of OSA syndrom, defined as $AHI \geq 5$ and excessive daytime sleepiness is present in at least 4% of men and 2% of women in the general adult population [36]. The main consequences of sleep apnea are daytime sleepiness [37-39], increased risk of cardiovascular and cerebrovascular disease [40-43], traffic accidents [44-47] and impaired quality of life [48, 49].

Full night PSG is considered as the gold standard method for sleep apnea diagnosis [34, 50]. However, the high cost of PSG, its time consuming and labor intensive nature and the high prevalence of the sleep apnea disorder have resulted in worldwide long waiting lists

of patients resulting in delaying their timely access to treatment, while there is increasing evidence in the literature to indicate that untreated OSA is associated with significantly increased morbidity and likely mortality [34]. The above mentioned complications have persuaded researchers to look for portable monitoring devices that can detect sleep apnea with comparable accuracy with the PSG but with smaller number of sensors, and eliminate the need for lengthy in lab monitoring for some patients. There are a variety of portable devices for monitoring sleep apnea. Some use only one signal such as nasal airflow [51-55], S_aO_2 [56, 57], respiratory sounds [18, 16, 58] or a combination of 2 to 4 signals [59-65].

The main signals used in most of the current portable monitoring devices are either the nasal airflow or S_aO_2 signals. However, nasal airflow may fail to give an accurate estimate of breathing flow rate due to the misplacement of the sensor during the night or in the cases of mouth–breathing. Use of S_aO_2 as the only signal for sleep apnea diagnosis [56, 57] is not currently recommended by American Academy of Sleep Medicine (AASM) due to its limited specificity and sensitivity [66]. On the other hand, tracheal respiratory sounds analysis during sleep can reveal useful information about the changes in the behavior of the upper airways and the breathing flow rate.

The diagnostic performance of tracheal respiratory sound and S_aO_2 signals for apnea/hypopnea detection has been compared in [16]. It was shown that tracheal sound analysis has higher sensitivity than pulse oximetery (S_aO_2 signal), while S_aO_2 signal showed higher specificity. Therefore, it can be concluded that the combination of both

signals may result in higher sensitivity and specificity for sleep apnea detection. In Section 4.2, we present the results of a new ambulatory device (acoustical sleep apnea detection, ASAD) for detection of sleep apnea using tracheal respiratory sounds and blood S_aO_2 level. The method is simple, fast, and can analyze 8–hours of data (during the entire night) in less than 15 minutes.

Chapter 2

Flow–Sound Relationship during Wakefulness

In this chapter we investigate the flow–sound relationship in healthy individuals during wakefulness. The goals of this part of the research are to investigate the significance of different anthropometric features on the flow–sound model parameters, and to examine whether it is possible to replace the calibration process by using the average values of the model parameters in a matched population for an individual with similar anthropometric data. Hence, we looked into the impacts of important anthropometric features such as height, weight, age, gender, ethnicity on flow-sound model parameters. We recorded tracheal sound and flow signals from a large group of healthy individuals including both smokers and non–smokers, and investigated the statistical effect of anthropometric data of the participants on the flow–sound model parameters. The statistical results were then used to generalize the flow–sound model, and examine the flow estimation accuracy.

Table 2.1: *Patients' anthropometric information (BMI is the body mass index).*

Ethnicity	Gender	Weight(kg)	Height(cm)	Age	Neck(cm)	BMI	Chest(cm)	Waist(cm)
Caucasian	M(23)	71.8 ± 10.1	173.3 ± 7.5	43.0 ± 17.6	38.5 ± 3.0	23.9 ± 3.4	99.7 ± 7.4	85.3 ± 9.9
	F(33)	64.2 ± 13.0	165.0 ± 6.1	36.8 ± 11.8	34.3 ± 2.5	23.6 ± 4.8	92.6 ± 6.7	79.4 ± 13.6
	All(56)	67.3 ± 12.4	168.4 ± 7.8	39.3 ± 14.7	36.0 ± 3.4	23.7 ± 4.3	95.5 ± 7.8	81.8 ± 12.5
Middle Eastern	M(7)	76.0 ± 9.3	172.6 ± 6.3	26.9 ± 2.8	38.6 ± 2.1	25.4 ± 2.0	104.0 ± 7.9	88.1 ± 6.8
	F(4)	53.0 ± 9.2	159.9 ± 5.9	27.3 ± 3.0	32.6 ± 2.7	20.6 ± 2.5	86.3 ± 6.6	72.0 ± 6.1
	All(11)	67.6 ± 14.6	168.0 ± 8.7	27.0 ± 2.7	36.4 ± 3.7	23.7 ± 3.2	97.5 ± 11.4	82.2 ± 10.2
Chinese	M(14)	73.6 ± 12.2	171.1 ± 6.2	31.0 ± 6.3	38.2 ± 3.6	25.1 ± 3.2	106.4 ± 11.3	88.8 ± 8.7
	F(3)	59.7 ± 2.5	155.3 ± 18.9	30.0 ± 6.2	31.3 ± 3.1	25.6 ± 7.0	87.8 ± 4.2	76.0 ± 2.6
	All(17)	71.1 ± 12.3	168.3 ± 10.7	30.8 ± 6.1	37.0 ± 4.3	25.1 ± 3.8	103.1 ± 12.6	86.5 ± 9.3
South Asian	M(8)	68.6 ± 9.5	169.0 ± 4.6	33.9 ± 8.4	39.1 ± 2.1	24.1 ± 3.8	95.5 ± 5.5	83.4 ± 11.0
	F(1)	49.0 ± 0.0	160.0 ± 0.0	19.0 ± 0.0	31.0 ± 0.0	19.0 ± 0.0	83.0 ± 0.0	64.0 ± 0.0
	All(9)	66.4 ± 11.0	168.0 ± 5.2	32.2 ± 9.3	38.2 ± 3.3	23.5 ± 3.9	94.1 ± 6.6	81.2 ± 12.1
All Ethnicities	M(52)	72.3 ± 10.4	171.9 ± 6.6	36.2 ± 14.0	38.5 ± 2.9	24.5 ± 3.2	101.4 ± 9.0	86.3 ± 9.4
	F(41)	62.4 ± 12.6	163.7 ± 7.7	34.9 ± 11.5	33.8 ± 2.6	23.3 ± 4.8	91.4 ± 6.8	78.1 ± 12.7
	All(93)	68.0 ± 12.4	168.3 ± 8.2	35.6 ± 12.9	36.5 ± 3.6	24.0 ± 4.0	97.0 ± 9.5	82.7 ± 11.6

With minor modifications, this chapter was published in IEEE Transaction of Biomedical Engineering [67].

2.1 Method

2.1.1 Data

Data of this part were recorded from 93 (52 males) healthy individuals (35.6 ± 12.9 years old) with no record of respiratory problems. This study was approved by the Ethics board of University of Manitoba prior to the study, and all participants signed the consent form before starting the experiment. Participants' anthropometric data including ethnicity, age, gender, height, weight, neck size, chest size, waist size, smoking habit and physical activity rate were recorded prior to the experiments, and are summarized in Table 2.1.

The respiratory tracheal sounds were recorded by a Sony (ECM-77B) microphone embedded in a chamber (diameter of $6mm$) and placed over the suprasternal notch of the

subjects neck. Sound signals were amplified with a gain of 200 and band-pass filtered with the cutoff frequencies of $[0.5Hz - 5kHz]$ using Biopac (DA100C) amplifiers. Respiratory flow was also measured simultaneously by a pneumotachograph (Fleisch No.3) connected to a differential pressure transducer (Biopac, TSD127). Subjects were instructed to breathe at 3 different flow rates, $L1, L2$ and $L3$, corresponding to flow rates of less than $0.5L/s$, $0.5 - 0.9L/s$ and $0.9 - 1.3L/s$, respectively, followed by $10s$ of breath–hold at the end of experiment. To have a consistent setup during the recordings, all subjects were asked to sit upright, not to move their neck and to place their mouth securely on the mouthpiece of the flow-meter and breathe orally. A nose clip was used to prevent any air escaping through the nasal passage.

2.1.2 Flow–sound Relationship

Tracheal sounds, recorded by a microphone over trachea, represent the vibrations due to the sound generated in trachea and motions of the tracheal wall that are transmitted to the skin. The main sources producing sounds and tracheal wall vibrations include turbulence of airflow in trachea and interaction between tracheal wall and airflow. The sounds power (proportional to the sound's amplitude square) generated by turbulent flow in a solid pipe changes with the speed of flow to the power of 6 or 7.7 [68]. On the other hand, studying the sound generated in a model of trachea and glottis, it was shown that the sound amplitude changes with the flow to the power of 3.0 ± 0.2, which is compatible

with the jet noise production in a pipe [69]. In another study [25], a model of trachea and 4 bronchial branches was investigated; the generated sound showed to change with flow to power of 2.

Another source of tracheal wall's vibration is the fluctuation of turbulent pressure in the tube, which is related to the pressure drop in the gas along the tube. The pressure drop is related to the tube's length, diameter, density of fluid, viscosity of fluid and flow rate [19, 70]. To reach skin's surface, the vibrations pass through the soft tissues of neck that has a low–pass filtering effect on the signal [71]. Considering these studies, it can be assumed that flow–sound relationship follows a power law, while the model parameters change with respect to the contributions of different components involving in the sound generation.

2.1.3 Flow–Sound Model

The relationship between flow and sound follows a power law; hence, we may express it as $E_s = kF^{\alpha}$ where, E_s and F represent the tracheal sound's average power and flow, respectively. Therefore, a linear relationship exists between the flow and sound's average power in logarithm scale:

$$\log F = a \log E_s + b, \tag{2.1}$$

where, $a = 1/\alpha$ and $b = -\log k/\alpha$. The sound's average power (E_s) can be estimated by calculating the signal's variance. We have shown previously that the logarithm of

the sound's variance follows the changes in the breathing flow [1] (details of the method and results can be found in Apendix A). The preliminary studies of the sound's average power and flow for different subjects showed high correlation between the two signals among all subjects. The R–values were 0.93 and 0.90 for inspiration and expirations phases, respectively, which confirms the validity of the proposed linear model 2.1.

Tracheal sounds were bandpass filtered in the frequency range of $[70 - 1000]Hz$ to remove the low- and high–frequency noises including ambient noises, muscle movements and motion artifacts, while keeping the main frequency components of tracheal sounds. The bandpass filtered sound signals were divided into windows of $20ms$ duration with 75% overlap between adjacent windows. The values of window size and overlap for segmenting the tracheal sound signal were selected based on the results of our previous studies on acoustical flow estimation [1, 29].

Heart sounds reduction

In studying respiratory sounds heart beat is an unavoidable source of interference that changes both frequency and time characteristics of the respiratory sounds. Most of the heart sounds energy is concentrated in the frequency range of $20 - 200Hz$ [72], which overlaps with the low frequency components of respiratory sounds. In most of the acoustical flow estimation methods [27, 26, 28], the tracheal sound was analyzed for frequency range above $300Hz$, where it is almost free of the heart sounds effect. However, we found

that for very shallow breathing where the tracheal sound has very low intensity, it is important to consider the frequency range below $300Hz$ for flow estimation. Because heart sounds interferes with the respiratory sounds in the frequency range below 300 Hz and also that our pilot studies showed a high inaccuracy in very low flow estimation due to heart sounds, therefore we investigated several methods to cancel the effect of heart sounds prior to flow estimation.

As the effect of heart sounds on lung sounds is much more pronounced than that on tracheal sound, all of the heart sounds cancellation techniques have been applied to lung sounds. Several methods based on adaptive filtering [73-76], wavelet denoising [77, 78], adaptive thresholding and 2-D interpolation of lung sounds in the time-frequency domain [79] and removing heart sounds-included segments from the wavelet coefficients of lung sounds and then reconstructing the signal by auto regressive or moving average models [80] have been proposed for heart sounds reduction from lung sounds.

We employed some of the above mentioned techniques but the results were not satisfactory for flow estimation purpose. This is probably due to the fact that by applying most of these techniques for heart sounds cancellation, some of the heart sounds still remain in the respiratory sound signal. On the other hand, some other techniques alter the respiratory sound signal. Although these changes are not much noticeable by auditory means, they increase the error of flow estimation at shallow breathing. Therefore, instead of heart sounds cancellation from the tracheal sound, we used a technique for canceling

the effect of heart sounds on the calculated features of the tracheal sound.

The first step in all of the methods for canceling heart sounds from respiratory sounds except the wavelet de-noising method is to localize the segments including heart sounds. In order to remove the effects of heart sounds on the calculated tracheal sound energy, the segments including heart sounds were localized and tracheal sound enery (variance of tracheal sound in time domain) was calculated for the segments void of heart sounds. Then, spline interpolation was applied to estimate the values of tracheal sound energy in the segments including heart sounds. This technique effectively cancels the effect of heart sounds on the estimated energy of the tracheal sound [1].

Estimating the model parameters

The model parameters a and b (Eq. 2.1) must be derived through a calibration process for every subject. In previous studies it was shown that the increase in sounds average is not similar at different flow rates [25]; hence, using the same model at all flow rates will cause over/under estimation of flow at the lower/higher flow rates than the flow rate used for calibrating the model [27, 28]. Therefore, we modified the model as:

$$\begin{aligned} \log F &= \bar{E}_s/\bar{E}_{base} \times a \times \log(E_s) + b \\ &= \bar{E}_s \times \left[a/\bar{E}_{base}\right] \times \log(E_s) + b, \end{aligned} \tag{2.2}$$

where $[\bar{E}]$ is the average function, E_s is the sound's variance in the overlapping windows of current breath cycle, and E_{base} is the sound's variance in the breath cycle used for calibrating the model. Since, the sound generation mechanisms are different during inspiration and expiration, different parameters should be extracted for each phase. Hence, $a_{ins}, b_{ins}, a_{exp}$ and b_{exp} represent the model parameters during inspiration and expiration phases, respectively where $a_{ins} = a/\bar{E}_{base-ins}$ and $a_{exp} = a/\bar{E}_{base-exp}$ are normalized values of a parameter during inspiration and expiration, respectively. This normalization also cancels the effects of sound variations between individuals. In the first stage of this study, these parameters were derived from one breath sound cycle with known flow similar to that in previous study [1]; these parameters are called "*self–calibrated*" model parameters, hereafter.

2.1.4 Statistical Analysis of model parameters and anthropometric features

To investigate the impacts of anthropometric features on model parameters ($a_{ins}, b_{ins}, a_{exp}$ and b_{exp}), for each feature the subjects were grouped differently. Table 2.2 shows the grouping criteria for each feature and the number of subjects in every group is presented in parenthesis. Statistical analysis of variance (ANOVA) was performed to examine whether the average of self–calibrated model parameters (Eq. 2.2) were significantly different ($p < 0.05$) in the groups within each anthropometric feature. The results of

Table 2.2: *Patients' anthropometric information (BMI is the body mass index).*

Parameter	G_0	G_1	G_2	G_3	G_4
Ethnicity	Chinese(17)	Caucasian(56)	Middle Eastern(11)	South Asian(9)	
Gender	Female(41)	Male(52)			
Age	$< 30(38)$	$[30-40)(28)$	$[40-50)(16)$	$\geq 50(11)$	
BMI	$(18.5-24.9](55)$	$\leq 18.5(5)$	$[25-29.9](25)$	$> 30(8)$	
Body-Fat	(33)	(3)	(24)	(16)	(17)
Height(cm)	$< 160(6)$	$[160-170)(48)$	$\geq 170(39)$		
Weight(kg)	$< 50(6)$	$[50-70)(48)$	$[70-90)(39)$		
Neck(cm)	$< 32(6)$	$[32-36)(48)$	$[36-40)(39)$		
Chest(cm)	$< 90(6)$	$[90-100)(48)$	$[100-110)(39)$		

Table 2.3: *Results of p-values obtained by ANOVA for investigating significance of different anthropometric features on the self–calibrated model parameters.*

Parameter	a_{ins}	b_{ins}	a_{exp}	b_{exp}
Smoking	0.695	$< 0.001^*$	0.072	0.050^*
Height	0.885	0.005^*	0.384	0.004^*
Gender	0.694	$< 0.001^*$	0.151	0.002^*
Age	0.525	0.255	0.096	0.721
BMI	0.450	0.057	0.087	0.425
Ethnicity	0.630	0.060	0.515	0.047
Body-Fat	0.457	0.113	0.139	0.092
Weight	0.369	0.168	0.219	0.795
Neck size	0.853	0.402	0.035	0.152
Chest size	0.503	0.396	0.465	0.742

this test, shown in Table 2.3, indicate that during both respiratory phases, b–parameters (b_{ins} and b_{exp}) were significantly different for smoking, height and gender features. However, other anthropometric features did not show any significant impact on the model parameters within different groups.

2.1.5 Alternative Model Calibration Approaches

Once, the anthropometric features with significant effect on model parameters were found, we investigated an alternative calibration approach, and compared its effect on flow estimation error with respect to self–calibration approach (when the self–calibrated model parameters were used). In the alternative calibration approach, we used the average values of the self–calibrated model parameters in a group of subjects with similar anthropometric features as the *reference calibration* parameters. These parameters were used as the reference model to estimate flow in the other groups of subjects. This approach will be called *group–calibrated* model, hereafter. Height, gender and smoking were found to have significant effects on the self–calibrated model parameters (Table 2.3). However, for reasons which will be discussed later, we only included non–smoker individuals in this approach. Height and gender were used to divide the non–smoker participants into four groups of $FH_{G1}, FH_{G2}, MH_{G1}$ and MH_{G2} representing females with height of $\leq 170cm$, females with height of $> 170cm$, males with height of $\leq 170cm$ and males with height of $> 170cm$, respectively.

The subjects in one group were assumed as the training dataset and other subjects were considered as the test dataset. First, we evaluated sensitivity of the group–calibrated model for estimating flow in a group of subjects with similar anthropometric data. For this purpose, the self-calibrated model parameters of all the subjects in train dataset except one were averaged and used to estimate flow of the left-out subject. This procedure

was repeated for all subjects in the train dataset. The second validation was performed to evaluate the results of group–calibrated model for estimating flow of the subjects with different anthropometric data. For this task, the average of self–calibrated model parameters of all subjects in the train dataset were used to find the *group–calibrated* model parameters. These parameters were applied to estimate flow from data of the subjects in the other groups, or test datasets. The procedure was repeated until all groups were considered as the train dataset once.

In all tests, flow estimation error was calculated as:

$$Error = \frac{mean(F - F_{est})^2}{mean(F^2)} \times 100, \tag{2.3}$$

where, F and F_{est} represent real and estimated values of flow, respectively. The error was averaged at different flow rates of less than $0.5L/s$, $0.5 - 0.9L/s$, $0.9 - 1.3L/s$ and for all flow rates.

2.2 Results

Figures 2.1–a and 2.1–b show the tracheal sound and flow signals recorded from one of the participants (male, 21 years old, height $170cm$). The data shows that by increasing flow, the tracheal sound's amplitude and correspondingly its energy increase. This is more clearly evident in the zoomed out portions of the signals which are presented in the Fig. 2.1–d and 2.1–e. Figure 2.1–c shows the results of acoustic flow estimation using

tracheal sound signals. It is clear that the estimated flow (Fig. 2.1–c and 2.1–f) follows the variations of the real flow (Fig. 2.1–b and 2.1–e). Also, when calibrating the model based on one breath cycle of the subject, a quantitative estimation of the recorded flow was achieved.

In the first flow estimation approach, the flow–sound model was calibrated individually and the data from one breath cycle of every subject was used to calculate the model parameters. Figure 2.2 shows the average and standard deviation values of flow estimation error among all individuals based on the self–calibration approach. The error values were averaged within and between subjects at three different flow rates, $L1, L2$ and $L3$. Also, the average of error values at all flow rates were calculated (Avg) to present the overall error for each subject. The results show that the average errors for all flow rates and their average during inspiration and expiration phases (except level $L3$ during inspiration) were less than 10% which confirms the results of our previous study with a smaller database [1] (details can be found in Apendix A).

In this study we recorded data from smoker and non–smoker participants. The average and standard deviation values of flow estimation errors in two groups of smoker and non–smoker participants are presented in Fig. 2.3. The results showed the average error values at different flow rates were higher for smokers compared to non–smokers. However, these differences were not significant ($p > 0.05$), except for error of $L1$ during inspiration ($p = 0.001$). The higher errors among smokers indicate that the flow–sound relationship

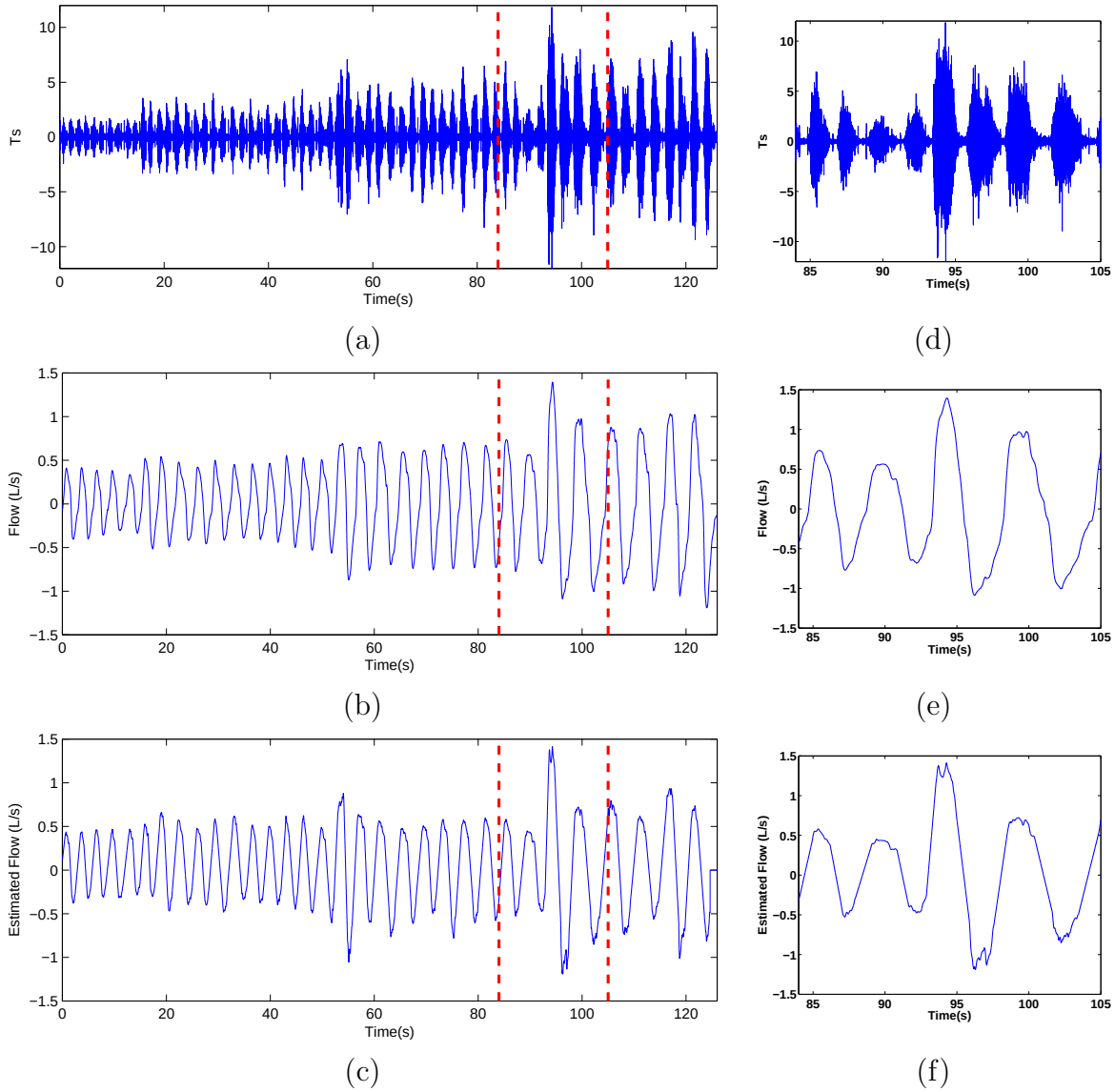


Figure 2.1: *Data recording of one of the subjects presenting a) recorded tracheal sound, b) recorded flow signal, and c) the estimated flow signal. The zoomed out plots show d) recorded tracheal sound, e) recorded flow and f) the estimated flow in the time period marked by dashed lines.*

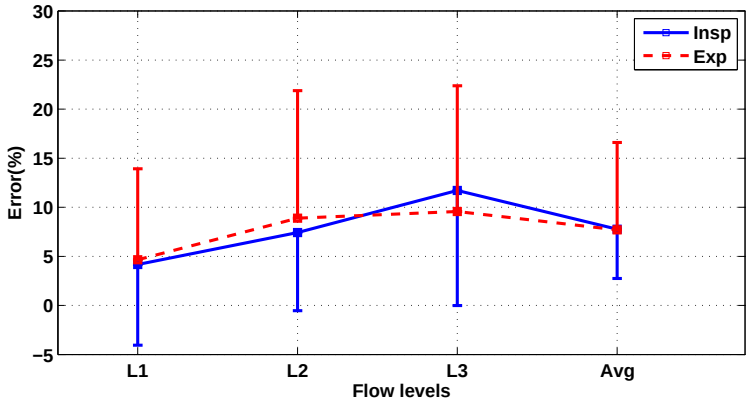


Figure 2.2: *Mean and standard deviation values of flow estimation error when the model was calibrated individually. The error values were averaged between subjects at three different flow rates, $L1, L2$ and $L3$, corresponding to flow rates of less than $0.5L/s$, $0.5-0.9L/s$, $0.9-1.3L/s$ and for all flow rates (Avg). The results are shown for inspiration (Insp) and expiration (Exp) phases, separately.*

at different flow rates is more variable within a smoker subject compared to that within a non–smoker subject. Therefore, when calibrating the flow-sound model parameters of a smoker individual at a specific flow rate, the same parameters may not be valid at other flow rates.

The model parameters (b_{ins} and b_{exp}) were shown to be significantly different between smokers and non–smokers (Table 2.3). Moreover, the flow estimation errors with self calibrations were higher for smokers (Fig 2.3–a and 2.3–b). Therefore, for the rest of this study, we only considered non–smoker participants and examined the effects of height and gender on the self–calibrated model parameters of non–smoker individuals.

ANOVA analysis was repeated for the model parameters of non–smoker participants and the results were similar (Table 2.4). Height and gender had significant effect on b_{ins}

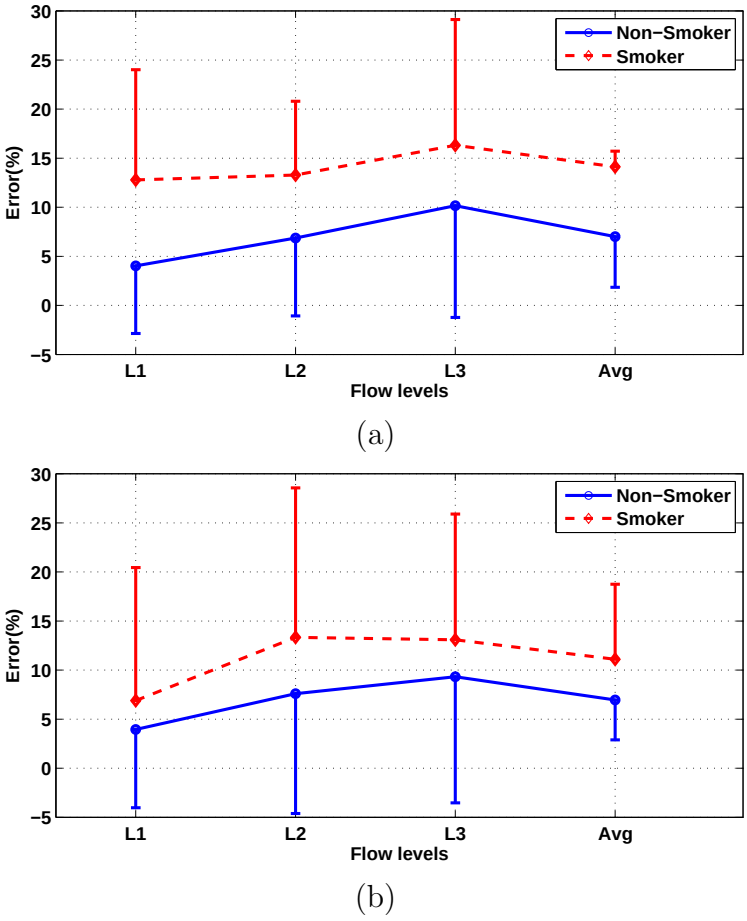


Figure 2.3: *Mean and standard deviation of flow estimation error for smokers and non-smokers during a) inspiration and b) expiration phases. The model parameters were calibrated individually and the error values were averaged within and between subjects at three different flow rates, $L1, L2$ and $L3$, corresponding to flow rates of less than $0.5L/s$, $0.5 - 0.9L/s$, $0.9 - 1.3L/s$ and for all flow rates (Avg).*

Table 2.4: *Results of p-values obtained by ANOVA for investigating significance of different anthropometric features on the self–calibrated model parameters of non–smoker individuals.*

Parameter	a_{ins}	b_{ins}	a_{exp}	b_{exp}
Height	0.870	$< 0.001^*$	0.713	0.002^*
Gender	0.335	$< 0.001^*$	0.365	0.002^*
Age	0.693	0.721	0.093	0.998
BMI	0.615	0.055	0.157	0.286
Ethnicity	0.641	0.241	0.677	0.159
Body-Fat	0.419	0.116	0.137	0.036
Weight	0.274	0.142	0.568	0.853
Neck size	0.901	0.297	0.157	0.160
Chest size	0.701	0.220	0.876	0.524

and b_{exp} with $p \leq 0.001$ and $p = 0.002$, respectively. However, they were not significantly different for a_{ins} nor for a_{exp}, as expected.

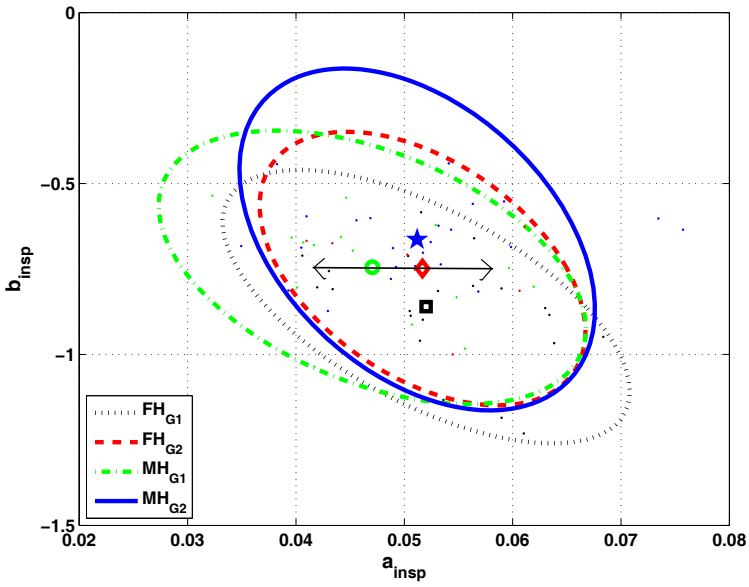
Height and gender were used to divide the non–smoker participants into four groups of $FH_{G1}, FH_{G2}, MH_{G1}$ and MH_{G2} representing females with height of $\leq 170cm$ (FH_{G1}), females with height of $> 170cm$ (FH_{G2}), males with height of $\leq 170cm$ (MH_{G1}) and males with height of $> 170cm$ (MH_{G2}), respectively. Figures 2.4–a and 2.4–b show scatter plot of the model parameters for the above groups. To have a better visualization of the parameters' variations in different groups, an ellipse was fitted to the data points in each group. Average of the model parameters in $FH_{G1}, FH_{G2}, MH_{G1}$ and MH_{G2} groups were marked with square, diamond, circle and star markers. From the results it can be concluded that for both respiratory phases, the average of b–parameters changes with the height variations. Moreover, the average values of b–parameters of taller females (FH_{G2}, diamond marker) and shorter males (MH_{G1}, circle marker) are similar (highlighted with

arrows in the Fig. 2.4). On the other hand, the averages of a–parameters do not change as much as the averages of b–parameters change in different groups (except for MH_{G1}, circle marker, during inspiration). This complies with the statistical results presented in Table 2.3.

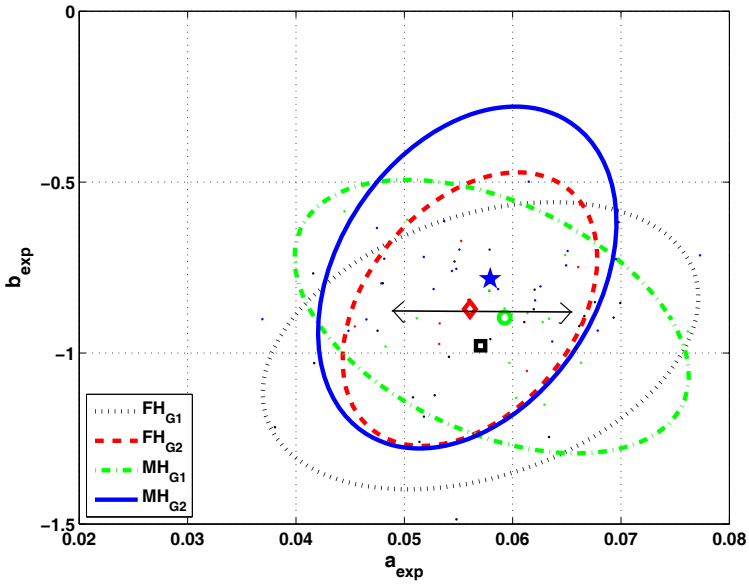
Table 2.5 shows the flow estimation errors, when the model parameters were set as the average of the parameters of the subjects in one group (train dataset) and tested on the data of the subjects in the other groups (test datasets). The results show that the lowest flow estimation errors were achieved when data of tall males (MH_{G2}) was used as the train dataset. In this setting, the calibration parameters were found to be $a_{ins} = 0.051$, $b_{ins} = -0.664$, $a_{exp} = 0.058$ and $b_{exp} = -0.783$. Average and standard deviation of the flow estimation errors at different flow rates based on selecting the model parameters from the data of MH_{G2} participants are shown in Fig. 2.5.

2.3 Discussion

Tracheal sounds analysis is a practical and non–invasive tool for investigating the pathophysiology of the upper airways and acoustical flow estimation [1, 8-10, 31]. Turbulence of air and pressure fluctuations in trachea generate sounds and vibrations, which are transmitted to the skin, and can be picked up by a microphone. The trachea can be modeled as a collapsible tube with a glottis at one end and the bronchi at the other end [25]. Therefore, sound characteristics change with airflow, air density, physiological

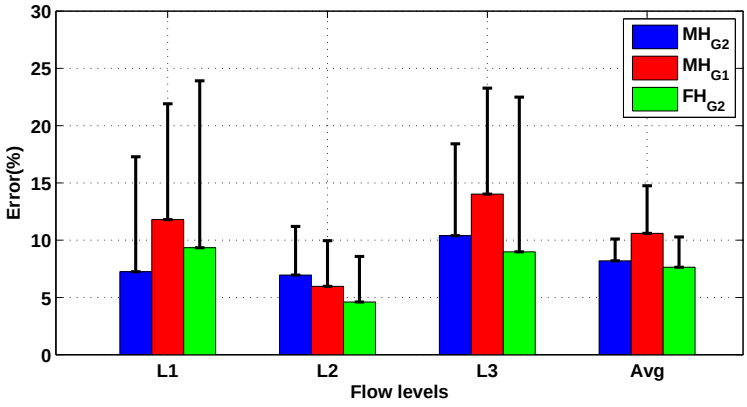


(a)

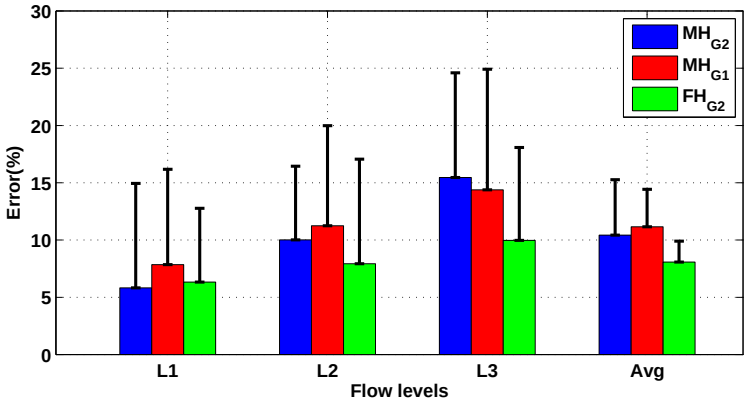


(b)

Figure 2.4: *Scatter of model parameters in different groups during a) inspiration and b) expiration phases. The average of the model parameters in $FH_{G1}, FH_{G2}, MH_{G1}$ and MH_{G2} groups were marked with square, diamond, circle and star markers.*



(a)



(b)

Figure 2.5: *Errors of flow estimation based on model parameters calibrated from data of males with height of* $> 170cm$ *(*MH_{G2}*), a) inspiration and b) expiration.*

Table 2.5: *Errors of flow estimation for different train and test datasets. $FH_{G1}, FH_{G2}, MH_{G1}$ and MH_{G2} representing females with height of $\leq$ 170cm (FH_{G1}), females with height of $>$ 170cm (FH_{G2}), males with height of $\leq$ 170cm (MH_{G1}) and males with height of $>$ 170cm (MH_{G2}), respectively. NA stands for not available.*

Inspiration				
Train \ Test	FH_{G1}	FH_{G2}	MH_{G1}	MH_{G2}
FH_{G1}	10.99%	14.06%	16.10%	NA
FH_{G2}	5.08%	5.89%	NA	10.54%
MH_{G1}	6.51%	NA	11.45%	12.74%
MH_{G2}	NA	3.52%	8.48%	6.09%
Expiration				
Train \ Test	FH_{G1}	FH_{G2}	MH_{G1}	MH_{G2}
FH_{G1}	15.35%	16.57%	20.95%	NA
FH_{G2}	10.97%	6.99%	NA	16.22%
MH_{G1}	8.81%	NA	13.15%	16.13%
MH_{G2}	NA	6.17%	9.44%	7.97%

and anatomical features of trachea such as length, diameter, wall thickness, stiffness and glottis size [19, 25, 68-71]. Hence, we investigated the effects of different anthropometric features of a large group of subjects on the flow–sound model parameters. The main motivation for this investigation was to remove the calibration need of model parameters for every individual in flow estimation applications.

Most of the models proposed for flow estimation are valid at only one flow rate and should be calibrated at different flow rates [8-10, 26-28, 31]. When the flow rate increases, the epiglottis diameter enlarges [25]; this changes the flow–sound model parameters. We improved the flow–sound model by adjusting the slope coefficient (a) of the linear model (Eq. 2.1) to compensate for the sound variations at different flow rates. The normalization

factor was defined as the average of the sound variance in every breath cycle divided by the average in the base cycle that was used to calibrate the model parameters. Using the modified linear model (Eq. 2.2), we removed the necessity to calibrate the model at different flow rates. However, it still required at least one breath cycle of every subject with known flow to derive the model parameters.

The main goal of this study was to examine the possibility of removing the calibration process completely by finding a set of general model parameters for individuals with similar anthropometric data. In previous studies [19, 70] it was claimed that fluctuation of turbulent pressure in the trachea is another source of tracheal wall's vibration. These fluctuations are related to the pressure drop, Δp, along the tube:

$$\Delta p \approx 0.241\ L\ \rho^{0.75}\ \mu^{0.25}\ F^{1.75}\ d^{-4.75}, \tag{2.4}$$

where L, ρ, μ, F and d are the tube's length, density of fluid, viscosity of fluid, flow and tube's diameter, respectively [19, 70]. Hence, tracheal sounds and turbulent pressure fluctuations in trachea change with the physical characteristics of trachea such as length and diameter that depend on individual's anthropometric data. We recorded data from 93 healthy individuals with different ethnicity, height, gender, weight, neck size, chest size, waist size, smoking and exercising habits to determine the most significant anthropometric features that affect the flow–sound model parameters.

When the model was calibrated for every subject individually, the average flow esti-

mation errors were less than 10% (Fig 2.2), which is similar to our previous results on a smaller database [1]. In the previous studies, acoustic flow estimation was applied on data of healthy non–smoker individuals and the method's performance on data of smokers were not examined. Data of this study included smoker individuals, and the results show that the average errors were higher for smokers compared to those of non–smokers (Fig 2.3). This indicates that for smoker participants, variability of the model parameters within a subject at different flow rates is higher. It should be noted that the models describing flow–sound relationship (Section 2.1.2) are based on the assumption of having turbulent flow in a tube with smooth surface. However, smoking has been shown to alter the pathology of airways, deposits semi–micro particles in the airways, induces airway remodeling and airway wall fibrosis [81, 82]. These consequences change the tracheal walls' smoothness, stiffness and thickness, which complicates using flow–sound data at a known flow rate to predict the flow-sound fluctuations at other flow rates, even for the same subject. Therefore, the presented flow–sound relationships do not apply to smokers, and a separate study is required to develop a model for acoustic flow estimation in smokers.

As expected, the results of ANOVA revealed that smoking, gender and height significantly change the b–parameter of the model (but not the a–parameter) consistently in both respiratory phases. Considering Eqs. 2.1 and 2.4, it is clear that the parameter b is associated with the factors multiplying with the flow component such as air density, tracheal length and diameter, while the parameter a depends on the power term (α) of

the flow component. The power term (α) and consequently parameter $a = 1/\alpha$ can be considered independent of the physical characteristics of trachea. This justifies why the parameter a did not change significantly with the variations in anthropometric data of the subjects (Table 2.3). On the other hand, the parameter b is related to trachea's length and diameter, which are known to be correlated with body height [83]. This explains the significant effects of individuals' height on the parameter b (Table 2.3).

Gender is the other factor that significantly affected the parameter b. The authors are not aware of any study which investigates the length of trachea in males and females with similar height. However, in a study on the pathophysiology of the pharyngeal airways, it was shown that the length of pharyngeal airway was significantly longer in men compared with that in women [84]. This may explain the significant differences in the b parameters among male and females. Smoking is the other factor that changed the b parameter by remodeling the physical and mechanical characteristics of tracheal wall as mentioned before. Consistency of the statistical results in both respiratory phases supports the importance of the body height and gender in defining the flow-sound relationship in healthy individuals.

We used gender and height parameters to divide the non–smoker subjects into 4 groups of FH_{G1}, FH_{G2}, MH_{G1} and MH_{G2} representing females with height of $\leq 170cm$, females with height of $> 170cm$, males with height of $\leq 170cm$ and males with height of $> 170cm$, respectively. The average of the individual model parameters (a, b) in each group was

calculated as representative parameters of the group. Again it was shown that the average a–parameters did not change significantly in different groups (except for a_{ins} in MH_{G1}), while the variations in the average b–parameters for different groups were affected by the body height (Fig. 2.4). Furthermore, the average b–parameters of the taller females (FH_{G2}) were similar to those of the shorter males (MH_{G1}), which complies with the hypothesis that airway length is smaller in females compared to that in males.

In order to remove the self–calibration need, we investigated the possibility of defining a general model for flow estimation in subjects with similar height and gender. We derived the group–calibrated model parameters by averaging the self–calibrated model parameters in every group (train dataset). The group–calibrated model parameters were then tested by estimating flow from data of the subjects in the same group (train dataset) and other groups (test datasets). The flow estimation errors were calculated for different train and test datasets. It was found that the errors were smaller in the train datasets compared to those of the test datasets as expected (Table 2.5). Moreover, the results were similar when either FH_{G2} or MH_{G1} were selected as train datasets (Table 2.5); this confirms the results presented in Fig. 2.4. The smallest flow estimation errors were achieved when data of tall males (height of more than $170cm$, MH_{G2}) were used as the train dataset. Studying the pathology of trachea in tall males to justify these results was beyond the scope of this study, but its examination may be helpful to enhance our knowledge of tracheal sound generation and improve the flow estimation model in future.

In summary, we investigated flow–sound model in detail and in a large group of individuals with different anthropometric data. The results have shown that gender, height and smoking are the most significant factors in modifying flow–sound model parameters. However, it is possible to use the self-calibrated model parameters extracted from data of subjects with similar height and gender to derive a general model for flow estimation. This technique will remove the calibration process for every individual, which will have a significant impact on acoustic flow estimation methods.

Chapter 3

Flow–Sound Relationship during Sleep

One of the main applications of acoustical flow estimation is sleep apnea diagnosis. In those applications, tracheal sounds are used to estimate the respiratory flow rate, and determine the reduction or cessation of air to the lungs [15, 16, 58]. In previous studies the relationship between tracheal sound and flow were studied extensively [19-24, 1, 9, 10, 25]. Tracheal sound amplitude and energy change with the amount of respiratory flow [1, 9, 10, 19, 25] and different models were proposed to estimate flow from tracheal sound [1, 9, 10, 8]. However, those studies were all focused on data of healthy individuals during wake. The relationship between tracheal sound and flow may change from wake to sleep. The relationship may also be different in healthy individuals and OSA patients. Therefore, a crucial step in application of acoustical flow estimation for sleep apnea diagnosis and monitoring is to investigate the flow–sound relationship among OSA patients during sleep.

The goal of this part of the thesis is two fold: first, to investigate the flow–sound relationship in patients with sleep apnea disorder, and examine how the relationship in every individual changes from wake to sleep. Second, to investigate the accuracy of acoustical flow estimation during sleep. Hence, the flow–sound model parameters were estimated during both sleep and wake. Different calibration algorithms based on the sleep and wake data were proposed and the method's performance in flow estimation was examined.

3.1 Method

3.1.1 Data

Data of this study were recorded from 16 (3 females) individuals suspected of OSA, who were referred to the Miserecordia Hospital Sleep Disorders Clinic for sleep studies. Patients were recruited randomly with no strict limitations in terms of age, gender or BMI. The study was approved by the Ethics board of University of Manitoba prior to the clinical trial. The patients' anthropometric detailed information is shown in Table 3.1.

The respiratory tracheal sounds were recorded by a Sony (ECM-77B) microphone embedded in a chamber (diameter of $6mm$) and placed over the suprasternal notch of the subject's neck. The chamber was attached to the skin with a double sided adhesive tape. The microphone and chamber were held in place with a soft neckband, which was fastened gently around patient's neck to assure the microphone would not be misplaced

Table 3.1: *Patients' anthropometric information (BMI is the body mass index). NA stands for not available.*

Patient (Gender)	BMI	Age	Weight(kg)	Height(cm)	AHI			
					Total	Supine	Left	Right
$S1(M)$	30.3	33	89.5	172	47.4	52	43.5	27.7
$S2(M)$	23.5	27	82.9	187.7	0.6	0	1.2	0
$S3(M)$	32.9	43	104.2	178	17.5	93.6	0	2
$S4(M)$	30.7	46	108.6	188	28.1	41.4	0	3.6
$S5(F)$	29.9	47	87.3	171	20.5	55.7	22.2	0
$S6(F)$	32.9	59	88.5	163.9	43.2	66.2	NA	13.9
$S7(F)$	37.1	56	101.1	165	26.5	40	NA	1.1
$S8(M)$	27.8	44	87	177	36.6	96.8	NA	11.9
$S9(M)$	29.5	49	91.4	176	7.9	17.1	1.9	19.2
$S10(M)$	37	48	128	186	12.6	28.3	2.1	0.8
$S11(M)$	26.4	38	88.6	183.2	0.3	0.3	0.8	0
$S12(M)$	37.1	59	111.9	173.6	1.1	20	0	0
$S13(M)$	43.5	41	125.7	170	116.9	132.9	NA	111.7
$S14(M)$	25.8	54	88.4	185	15.7	42.9	9.7	0
$S15(M)$	30.5	38	91.3	173	6.8	19.8	0	0
$S16(M)$	28.9	50	93.5	180	3.8	12.3	0.6	0.6
Average	31.5	45.8	98.0	176.8	24.1	45.0	6.8	12.0
Std	5.1	9.0	14.0	7.7	29.0	37.3	13.2	27.9

during the night. None of the patients reported any complaints due to the neckband or the microphone setting. Sound signals were amplified with a gain of 200 and band-pass filtered with the cutoff frequencies of $[0.5Hz - 5kHz]$ using Biopac (DA100C) amplifiers. Respiratory flow was measured simultaneously by a face mask pneumotachograph (Fleisch No.3) connected to a differential pressure transducer (Biopac, TSD127). The Respimed full face mask was sealed around patient's nose and mouth. During the recording, the CO_2 level in the mask was monitored to make sure the amount of recirculated CO_2 that patient would inhale did not increase beyond the normal range of $50mmHg$. An accelerometer was taped to the top of face mask, on the forehead, to monitor the head position during sleep. The blood's oxygen saturation level (S_aO_2 signal) was recorded with a Masimo finger probe ($5N040$) connected to a Masimo pulse oximeter (Radical signal extraction pulse oximeter). PSG study was running simultaneously with our recording; we used the EEG recordings to monitor the patients' sleep stage. Figure 3.1 shows arrangement of the sensors for recording data during sleep.

The sound, flow, position data and S_aO_2 signals were simultaneously digitized at a sampling rate of 16 kHz using National Instruments data acquisition module ($NI9217$). While the patients were awake, they were asked to breathe at different positions with their normal breathing rate to provide a reference at each position. There was no control over the patient's position during sleep. The flow–sound relationship during sleep and wake were compared for the available body positions during sleep.

Figure 3.1: *Arrangement of the sensors for recording data during sleep.*

3.1.2 Flow–sound Relationship

Tracheal sound's average power and flow were shown to have a power relationship during wake. In this study, we investigated flow–sound relationship during sleep and wake at different positions.

wake

At the beginning of the study, the subjects were asked to breathe at sitting, supine, lateral left and right positions. For every position, tracheal sound signals were investigated manually to select the portions void of background noises including speech, movements, etc. Tracheal sounds were bandpass filtered in the frequency range of $[100 - 1000]Hz$ to remove the low- and high–frequency noises, while keeping the main frequency components of tracheal sounds. Tracheal sound's average power was calculated in windows of $50ms$ with 75% overlap between the adjacent windows. In every breath cycle, only the stationary sound segments corresponding to the upper 40% of the peak flow were considered for investigation.

Tracheal sound average power and flow are assumed to follow a power relationship:

$$\log E_s = a \log F + b, \tag{3.1}$$

where E_s and F represent tracheal sound average power and flow, respectively. A line was fitted to the logarithm of the sound's average power and the logarithm of flow by mini-

mizing the mean square error. Parameter a was calculated for different respiratory phases and different body positions. To make it clear that parameter a shows the relationship between sound's average power and flow, it will be called Avg_a hereafter.

Sleep

The EEG signals of the PSG recording were used to detect the patient's sleep status and stage. During sleep, the sound signals were monitored manually, and the periods void of noise and snore sounds were selected. If the patient slept at different positions, the data of each position was selected separately. Table 3.2 shows the duration of the extracted data for every patient at different positions during sleep and wake. Similar to the wake, average power of the tracheal sound was calculated and the slope of the line fitted to the flow and sound's average power were estimated (Avg_a). For every subject and at every available position, the variation of parameter Avg_a from wake to sleep was investigated.

The variation of parameter Avg_a from wake to sleep might be related to variations in physiology of the upper airways during sleep. Consequently, it may indicate severity of sleep apnea. Hence, for every available position, the difference in Avg_a from wake to sleep was calculated:

$$da = Avg_a(Sleep) - Avg_a(Wake), \tag{3.2}$$

where, $Avg_a(Sleep)$ and $Avg_a(Wake)$ represent values of Avg_a during sleep and wake, respectively. For every position, the correlation between da and two different values

of AHI were calculated: AHI_P and AHI_T which represent patient's AHI score in the corresponding position and the total AHI value regardless of sleep positions. Since the number of observations were small, the non–parametric Spearman algorithm was applied to calculate the correlation coefficient (r) and p–value.

During sleep, the peak flow changes during time. It might also be different from those during wake. A change in peak flow may change the parameter Avg_a; therefore, the peak flow of wake and sleep recordings were also compared. For every position, logarithm of the peak flow was subtracted from the parameter Avg_a (normalization) to compensate for the variations in the peak flow. Then, for every subject, variations in the the normalized values of Avg_a were compared between wake and sleep.

3.1.3 Flow Estimation

The modified linear model was applied to estimate flow from tracheal sound variance:

$$\begin{aligned} \log F &= \bar{E}_s/\bar{E}_{base} \times a \times \log(E_s) + b \\ &= \bar{E}_s \times \left[a/\bar{E}_{base}\right] \times \log(E_s) + b, \end{aligned} \tag{3.3}$$

where $[\bar{E}]$ is the average function, E_s is the sound's variance in the overlapping windows of the current breath cycle, and E_{base} is the sound's variance in the breath cycle used for calibrating the model. Two approaches were applied to calibrate the model and estimate the model parameters. In the first approach, we investigated the possibility of using

Table 3.2: *Duration of available data for every subject at different positions during sleep and wake.*

Subject	State	Position					
		Sitting	Supine	Head Left	Body Left	Head Right	Body Right
S1	Wake	11.8s	19.7s	25.4s	27.3s	22.8s	20.7s
	Sleep	-	221.2s	-	-	29.6s	-
S2	Wake	21.0s	43.0s	27.8s	28.5s	33.0s	34.4s
	Sleep	-	-	-	69.7s	133.3s	-
S3	Wake	16.9s	368.2s	21.0s	13.0s	21.0s	15.8s
	Sleep	-	-	-	-	-	149.6s
S4	Wake	7.9s	39.2s	30.7s	13.7s	22.6s	37.6s
	Sleep	-	555.2s	-	-	-	-
S5	Wake	31.0s	30.6s	22.5s	40.4s	25.2s	22.2s
	Sleep	-	-	-	222.5s	-	-
S6	Wake	-	-	15.0s	-	-	-
	Sleep	-	-	-	-	-	-
S7	Wake	27.0s	237.5s	21.9s	25.0s	34.0s	33.0s
	Sleep	-	23.9s	-	-	-	-
S8	Wake	17.7s	31.5s	5.3s	15.0s	-	13.0s
	Sleep	-	53.7s	-	-	-	-
S9	Wake	10.0s	65.9s	13.0s	43.5s	21.5s	30.5s
	Sleep	-	-	-	-	-	205.0s
S10	Wake	-	87.3s	7.9s	8.0s	12.0s	-
	Sleep	-	180.0s	-	-	-	-
S11	Wake	21.0s	20.0s	21.0s	29.0s	19.0s	22.0s
	Sleep	-	394.1s	182.6s	-	-	-
S12	Wake	57.0s	43.6s	31.7s	18.0s	30.0s	21.1s
	Sleep	-	69.6s	55.3s	-	-	90.8s
S13	Wake	-	30.0s	13.9s	10.8s	12.3s	13.0s
	Sleep	-	169.5s	-	-	-	-
S14	Wake	10.3s	13.4s	9.5s	12.0s	13.0s	14.4s
	Sleep	-	-	-	179.5s	-	-
S15	Wake	20.0s	16.8s	17.0s	8.7s	13.0s	8.4s
	Sleep	-	57.9s	-	-	-	-
S16	Wake	-	45.0s	-	-	-	-
	Sleep	-	-	-	-	147.2s	-

the patient's data during wake to calibrate the model and estimate flow during sleep. Therefore, the model was calibrated using one breath cycle with known flow recorded during wake at similar body position. In the second approach, one breath cycle with known flow during sleep was used to calibrate the model and estimate parameters a and b. To compare these approaches, model parameters a and b during wake and sleep were compared and their variations were investigated. In addition, the flow estimation error using different calibration approaches were estimated and compared:

$$Error = \frac{mean(F - F_{est})^2}{mean(F^2)} \times 100, \tag{3.4}$$

where, F and F_{est} represent real and estimated values of flow, respectively.

3.2 Results

The relationship between tracheal sound average power and flow were investigated during wake and sleep. Parameter Avg_a represents slope of the line fitted to the tracheal sound average power and flow in the log–log scale. For every subject and for every available position, the variations in parameter Avg_a and the peak flow during sleep and wake are shown in Fig. 3.2. Every line in Fig. 3.2 relates to the parameter Avg_a or peak flow of a subject during wake and sleep. Lines with different colors and styles represent different body positions: red solid line for supine (Sp), blue dashed line for lateral left (Lt) and black dotted line for lateral right (Rt) positions. Data of different individuals are marked

with various markers. The vertical solid lines with downward triangle markers represent average and standard deviation of each variable among all subjects.

Figures 3.2–a and 3.2–b show the variations in parameter Avg_a from wake to sleep during inspiration and expiration phases, respectively. During inspiration, for majority of individuals and body postures, the parameter Avg_a increased from wake to sleep ($p < 0.001$). This is clear from positive slope of the lines in Fig. 3.2–a and the increase in average of Avg_a parameter from wake to sleep. This trend was not seen during expiration ($p = 0.36$, Fig. 3.2–b).

The relationship between flow and tracheal sound average power may change with the peak flow. The average of peak flows for every individual and position are presented in Fig. 3.2–c and 3.2–d for inspiration and expiration phases, respectively. The results indicate that for both respiratory phases, the peak flow may increase, decrease or remain similar from wake to sleep; no consistent pattern was seen. For every sleep or wake situation, the normalized Avg_a are shown in Fig. 3.2–e and 3.2–f. Comparing the results presented in Fig. 3.2–a and 3.2–b with those of Fig. 3.2–e and 3.2–f, it is clear that normalization with respect to peak flow did not change the variation pattern of parameter Avg_a from wake to sleep. The results of correlation between between da ($Avg_a(Sleep) - Avg_a(Wake)$) and AHI values for different positions and respiratory phases are presented in Table 3.3.

Figure 3.3 shows the variations in flow–sound model parameters a and b (Eq. 3.3) during sleep and wake. For both respiratory phases, the parameters a and b showed either

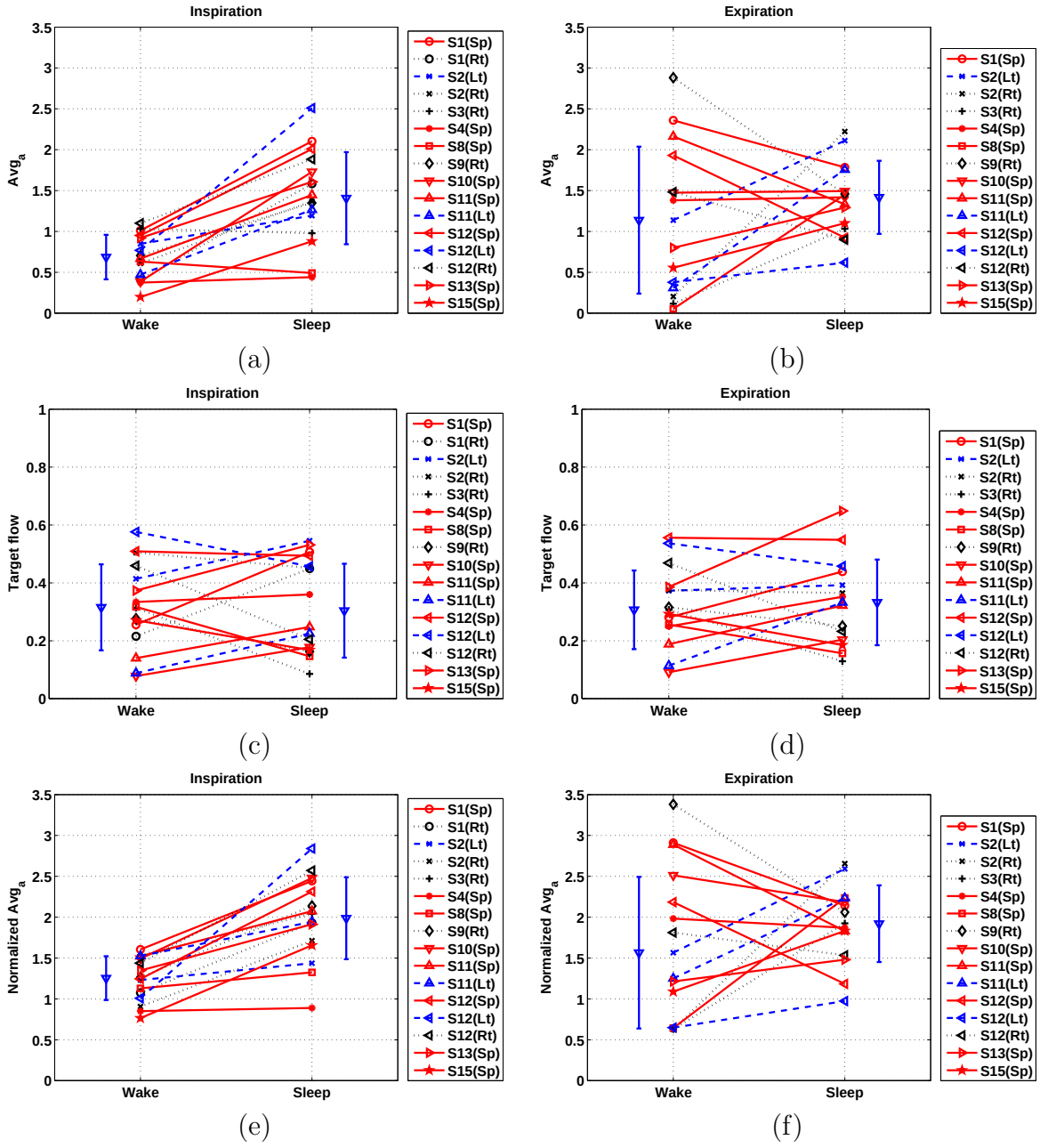


Figure 3.2: *Variations in parameter Avg_a (slope of the line fitted to the tracheal sound average power and flow in the log–log scale) and peak flow during sleep and wake. Lines with different colors and styles represent different body positions: red solid line for supine (Sp), blue dashed line for left (Lt) and black dotted line for right (Rt) positions. Data of different individuals are marked with various markers. The vertical solid lines with downward triangle markers represent average and standard deviation of each variable among all subjects. Parameter Avg_a during a) inspiration, and b) expiration, peak flow during c) inspiration and d) expiration, normalized values of parameter Avg_a during e) inspiration, and f) expiration.*

Table 3.3: *Spearman correlation coefficients between da* $(Avg_a(Sleep) - Avg_a(Wake))$ *and AHI values for different positions and respiratory phases. r and p show the correlation coefficients and significance of the results.* AHI_P *and* AHI_T *represent AHI values at the corresponding position and the total AHI independent of position, respectively.*

Position	Inspiration $[r(p)]$		Expiration $[r(p)]$	
	da, AHI_P	da, AHI_T	da, AHI_P	da, AHI_T
Supine	$-0.24(0.58)$	$-0.17(0.70)$	$0.48(0.24)$	$0.50(0.22)$
Right	$1.00(0.33)$	$0.10(0.95)$	$-1.00(1.00)$	$-0.40(0.75)$
Left	$-1.00(1.00)$	$0.50(1.00)$	$-1.00(1.00)$	$-1.00(0.33)$
All	–	$-0.11(0.67)$	–	$-0.10(0.72)$

increase, decrease or negligible variations from wake to sleep (Fig. 3.3). The variations in parameter a were not significantly different from wake to sleep with the p–values of 0.89 and 0.98 for inspiration and expiration, respectively. On the other hand, the variations in parameter b were slighty different from wake to sleep with p–values of 0.11 and 0.02 for inspiration and expiration, respectively. The results of correlation between between da and db $[da = a(Sleep) - a(Wake), db = b(Sleep) - b(Wake)]$ and AHI values for different positions and respiratory phases are presented in Table 3.4.

In this study two approaches were used to calibrate the flow–sound model and estimate the model parameters. Samples of the recorded and estimated flow of subject $S11$ during sleep at lateral left position are shown in Fig. 3.4. Figure 3.4–a presents the flow estimation results when the model parameters were calibrated using data of the same subject at the lateral left position during wake (first calibration approach), while Fig. 3.4–b shows the results when the model was calibrated using the recordings during sleep.

Figure 3.5 shows the results of flow estimation error based on the second calibration

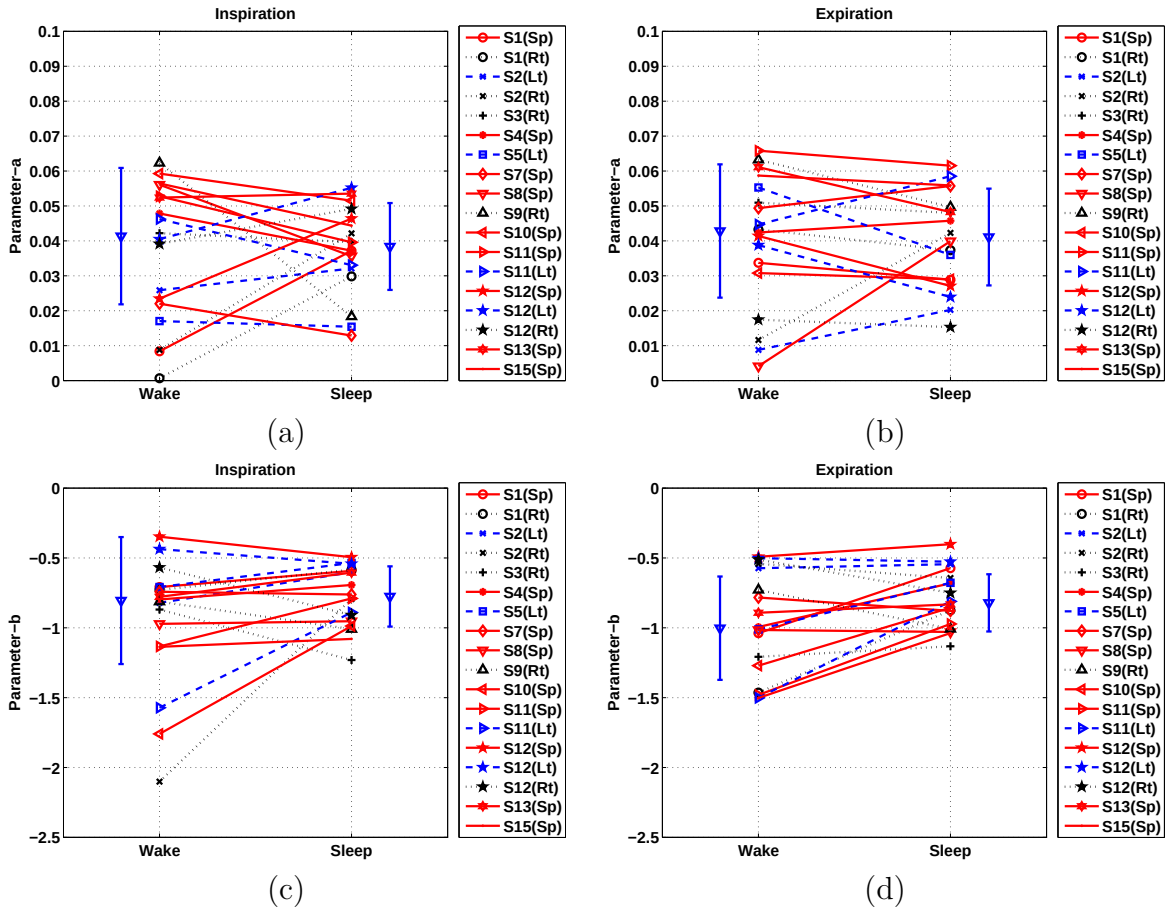


Figure 3.3: *Variations in flow–sound model parameters a and b (Eq. 3.3) during sleep and wake. Lines with different colors and styles represent different body positions: red solid line for supine (Sp), blue dashed line for left (Lt) and black dotted line for right (Rt) positions. Data of different individuals are marked with various markers. The vertical solid lines with downward triangle markers represent average and standard deviation of each variable among all subjects. Parameter a during a) inspiration, and b) expiration, parameter b during c) inspiration and d) expiration.*

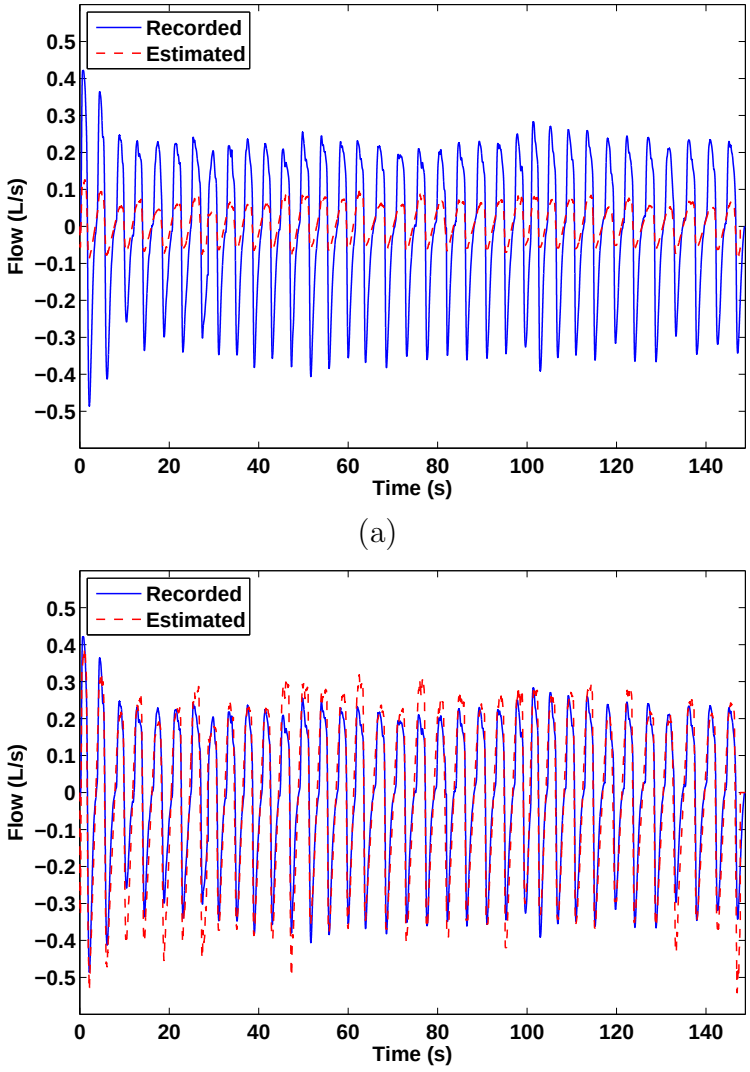


Figure 3.4: *Example of flow estimation results during sleep from data of subject S*11 *at left body position. Flow–sound model was calibrated based on a) awake and b) sleep data.*

Table 3.4: *Spearman correlation coefficients between da and db and AHI values for different positions and respiratory phases. $da = [a(Sleep) - a(Wake)]$, $db = [b(Sleep) - b(Wake)]$ and AHI_P and AHI_T represent AHI values at the corresponding position and the total AHI independent of position, respectively.*

Position	Inspiration $r(p)$			
	da, AHI_P	da, AHI_T	db, AHI_P	db, AHI_T
Supine	0.23(0.55)	0.28(0.46)	−0.02(0.98)	0.08(0.84)
Right	0.50(1.00)	−0.30(0.68)	1.00(0.33)	−0.30(0.68)
Left	0.50(1.00)	0.40(0.75)	−1.00(0.33)	−0.80(0.33)
All	–	0.08(0.75)	–	−0.20(0.44)
Position	Expiration $r(p)$			
	da, AHI_P	da, AHI_T	db, AHI_P	db, AHI_T
Supine	0.15(0.71)	0.12(0.78)	−0.53(0.15)	−0.38(0.31)
Right	−0.50(1.00)	−0.70(0.23)	0.50(1.00)	0.60(0.35)
Left	−1.00(0.33)	−1.00(0.08)	−0.50(1.00)	−0.40(0.75)
All	–	−0.20(0.43)	–	0.05(0.84)

approach (calibrating the model using sleep data) for different individuals. For subjects with data in more than one body position, the errors at different positions were averaged. The results show for majority of the subjects (except for $S9$ during inspiration and $S15$ during expiration) the flow estimation errors were less than 10% during both respiratory phases. The total flow estimation errors were averaged among different individuals; they were found to be $3.32 \pm 2.80\%$ and $4.65 \pm 4.18\%$ during inspiration and expiration phases, respectively.

3.3 Discussion

One of the main application of the acoustical flow estimation is to determine the change (reduction) of breathing flow during sleep, and detect apnea and hypopnea events during

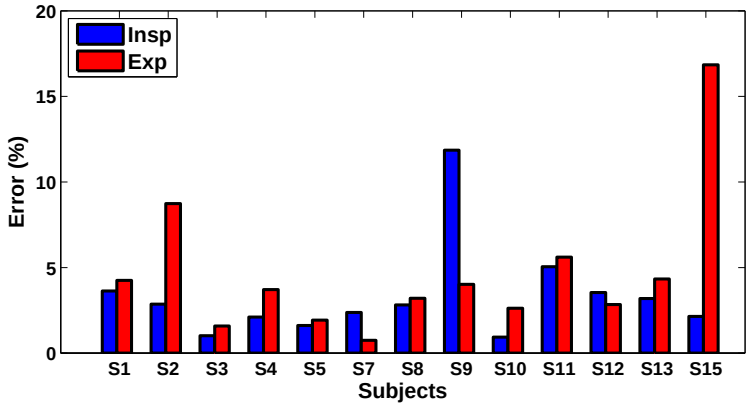


Figure 3.5: *Results of the flow estimation error based on second calibration approach for different individuals during inspiration and expiration phases.*

sleep. However, all the models developed to describe flow–sound relationship and estimate flow rate have been based on data recorded from healthy individuals during wake. The breathing mechanisms during sleep and wake are different. Hence, it is probable that the model developed based on wake data may not be valid during sleep, or may require some modifications for accurate performance.

In this study, we investigated the variations in flow–sound relationship during wake and sleep. We recorded tracheal sound, flow, and body positions from 16 patients referred for full night sleep study. At the beginning of the recording, the subjects were asked to breathe at different body positions to have samples of wake data at every position. Anatomy of the upper airways changes with posture; hence, the flow–sound relationship during sleep and wake were compared in similar body postures. It should be noted that during sleep, there was no control over the patient's position.

We recorded flow with a sealed full face mask that covered patient's nose and mouth.

In this setting the patients were re–breathing their own CO_2. We monitored the CO_2 level in the face mask to make sure the CO_2 level was in the normal range (less than $50mmHg$). Due to the delays in ventilating the air in the face mask, the water vapor content of the breathing air is more than that of the normal breathing without mask. This issue changes the air density and alters the flow–sound relationship. However, we used the same setting for sleep and wake; hence, we may ignore the variations due to gas density from wake to sleep. The use of face mask also increases upper airway resistance, which limits our findings.

Because of the limitations imposed by the use of face mask, we considered patients with BMI of less than 40 and younger than 60 years old, who would have less difficulties to breathe with the mask. We had no control on the gender or AHI of the recruited patients. We recorded data from 26 patients, 10 of whom had difficulties in tolerating the use of face mask and could not fall asleep; hence, they were excluded from the analysis. We recorded data in the hospital with no control on the environmental noises and conditions. In another study we compared the signal to noise ratio of two microphones placed in the air and on the patient's neck; the results showed SNR of the ambient recording is about $10dB$ less than the neck recording which can be regarded as an indicative of the ambient noise amplitude (Section B.2). Another type of noise can vary with the amount of flow and interact in the flow–souund model as an additive term. Furthermore, the use of face mask made it difficult for patients to fall asleep, and consequently increased patients'

movements. Theses factors increase the noise level of the recorded signal, and deteriorate the signal to noise ratio. Among the 16 patients who did fall asleep, data of 3 patients were noisy and it was impossible to extract reliable information from their data. These patients (subjects $S6, S14$ and $S16$) were ruled out from the investigations. Considering these factors, applying noise removal techniques rather than band–pass filtering the signal can enhance the sound quality and improve the method's performance.

Tracheal sound average power changes with flow rate following a power law model. Tracheal sounds average power and flow were investigated in the log–log scale and a line was fitted to the data points. The results of Fig. 3.2–a indicate that for most of the subjects during inspiration, slope of the line (Avg_a) increased from wake to sleep. The same patterns were observed after normalizing for peak flow variations during sleep and wake (Fig. 3.2–e). Increase in Avg_a, and specifically its variations from below 1 to greater than 1, shows that the shape of flow-sound curve changes from convex to concave curve during wake and sleep, respectively; this indicates variations in the mechanisms controling flow-sound relationship during wake and sleep. This is a very important finding that once verified can be used to investigate the variaitons in the underlying physiology of upper airway, such as its collapsibility, during sleep and wake.

Tracheal sounds are generated from turbulence of airflow in the upper airways. The power of Jet noise generated in an infinite solid pipe was shown to change with flow to power of 6 or 7.7 [68]. On the other hand, tracheal sounds intensity was shown to change

with flow to power of between $1.5 - 2$ [25, 9], which is less than what was shown for solid pipes (around $6 - 8$). Collapsibility of upper airways, finite length of the airways and bifurcations at the end of the airways could be the main factors which deviates trachea from a solid pipe model and changes flow-sound power model from the expected model.

Furthermore, the lower power of flow (Avg_a) in our results and its variations from wake to sleep can be due to the contribution of other sources in generating tracheal sounds. Trachea is a collapsible tube; in patients with sleep apnea, its stiffness is smaller, which causes narrowing of upper airway and flow limitation [85]. With narrowing in the upper airways, the neuromuscular activity becomes crucial to maintain the airway stability. However, the magnitude of the genioglossus muscles' reflex to negative upper airway pressure decreases during sleep and the delay in the response increases from wake to sleep [86]. The delayed restoring forces can increase the instability of the airways, and produce flutters in the airways which may cause snoring as well [87]. The same flutter phenomenon is used to describe the generation of wheezes durng airway obstructions [21]. The incorporation of flutters in tracheal sound generation may explain the smaller vlaues of flow power (Avg_a) in our results and variations in the flow power during wake and sleep.

AHI is a rough estimation of sleep apnea severity, and we hypothesized that the variations in the Avg_a may be correlated to the AHI. To investigate this, the correlation between da (difference between values of Avg_a during sleep and wake) and AHI values

were calculated. Statistical analysis showed that the correlation between da and AHI were not significant (Table 3.3). This can be due to the lack of enough observations or due to the fact that AHI is not a good representative of upper airway characteristics such as its collapsibility [85]. During expiration, values of Avg_a either increased, decreased or remained similar from wake to sleep. Similar to inspiration, the correlation between da and AHI values were not statistically significant (Table 3.3).

One of the main applications of tracheal sounds analysis is acoustical flow estimation. For sleep apnea monitoring, it is not necessary to have the quantitative estimation of flow rate. For these applications, the relative values of flow would be enough to detect reductions (hypopnea) and cessations (apnea) of flow rate. Having said that, it is still of interest to examine the accuracy of flow estimation algorithm during sleep, as it will be an indicator of the reliability of relative flow estimation results. More importantly, it is helpful to verify the possibility of calibrating the model at wake and apply it to the sleep data. Therefore, we used two approaches to calibrate the flow–sound model, and find the model parameters. In the first approach the goal was to study the possibility of using the wake data to calibrate the model and estimate flow during sleep; if successful, this will greatly simplify the flow estimation algorithm. In the second approach, we used sleep data to calibrate the model parameters and estimate flow during sleep. The results show that independent of the sleep position, the model parameters change from wake to sleep (Fig. 3.3). Furthermore, the correlation between da, db and AHI values were not

statistically significant (Table 3.4).

The models based on two calibration approaches were used to estimate flow during sleep. The results show that the model based on sleep calibration follows the variation of flow signal with small errors of less than 5% during both respiratory phases (Fig. 3.4–b and 3.5). On the other hand, when the model was calibrated using wake data, the estimated flow during sleep was following the relative variations in the real flow rate (Fig. 3.4–a). However, due to differences in the model parameters during sleep and wake, this model can not give an accurate estimation of the quantitave flow rate (Fig. 3.4–a) and the estimation errors (not presented here) were high.

Chapter 4

Acoustical Sleep Apnea Detection

One of the main applications of acoustical flow estimation is to detect the abnormalities of the upper airways and examine the obstructions in the upper airways. The diagnostic performance of tracheal sound and S_aO_2 signals for apnea/hypopnea detection has been compared in [16]. It was shown that tracheal sound analysis has higher sensitivity than pulse oximetery (S_aO_2 signal), while S_aO_2 signal showed higher specificity. Therefore, it can be concluded that the combination of both signals may result in higher sensitivity and specificity for sleep apnea detection. In this chapter we report on the development of a new ambulatory device (acoustical sleep apnea detection, ASAD) for detection of sleep apnea using tracheal respiratory sounds and blood's S_aO_2 level. The results are presented and compared with those of full night PSG study for comparison. With minor modifications, this chapter was published as a journal paper [15].

4.1 Method

4.1.1 Data

Tracheal respiratory sounds were recorded by a small omni–directional microphone (Sony ECM-77B) inserted in a chamber, and attached to the patient's neck over the suprasternal notch with a double sided adhesive tape. The microphone and chamber were held in place with a soft neckband, which was fastened gently around patient's neck to assure the microphone would not be misplaced during the night. None of the patients have reported any complaints due to the neckband or the microphone setting. The S_aO_2 signal was recorded with a Masimo finger probe ($5N040$) connected to a Masimo pulse oximeter (Radical signal extraction pulse oximeter). The sounds were amplified and lowpass filtered with 5 kHz cutoff frequency using Biopac (DA100C) amplifiers. The S_aO_2 signal and filtered tracheal sounds were simultaneously digitized at a sampling rate of 10240 Hz by National Instruments data acquisition module ($NI9217$). The digitized signals were saved in a file every 3 minute resulting approximately 140 files for an entire night of recording.

Sixty six patients (48 males) going through full–night PSG study at the Health Sciences Center Sleep Disorders Clinic (Winnipeg, Canada) gave written consent to participate in this study. Subjects were recruited randomly with no limitations in terms of age, gender or BMI. The study was approved by the Ethics board of University of Manitoba prior to the clinical trial. The patients' anthropometric detailed information is shown in Table 4.1.

Table 4.1: *Patients' anthropometric information (BMI is the body mass index).*

Group	Parameter	Age ($\mu \pm \sigma$)	BMI ($\mu \pm \sigma$)	AHI ($\mu \pm \sigma$)
Females (18)	Average	51.9 ± 13.4	31.5 ± 5.5	10.3 ± 11.4
	Range	$[29 - 87]$	$[22.5 - 43.4]$	$[0.2 - 37.1]$
Males (48)	Average	51.2 ± 11.4	32.1 ± 6.7	28.9 ± 33.8
	Range	$[25 - 73]$	$[20.4 - 56.8]$	$[0.8 - 125.7]$
All (66)	Average	51.4 ± 11.9	31.9 ± 6.4	23.6 ± 30.3
	Range	$[25 - 87]$	$[20.4 - 56.8]$	$[0.2 - 125.7]$

4.1.2 Signal Analysis

The energy of respiratory tracheal sounds in logarithmic scale has been shown to change with respiratory flow rate. In this part, the logarithm of the tracheal sound variance ($LogVar$) was used to estimate the relative respiratory flow and the percentage of respiratory flow reduction or complete lack of flow. The procedure of finding apnea-hypopnea events was implemented in three steps: 1) the tracheal sound signal was analyzed to find the sound and silent segments, 2) the S_aO_2 signal was investigated to find the periods including drops in the blood's oxygen level, 3) the tracheal sound segments corresponding to the periods with reduced S_aO_2 level were further examined and automatically classified into breath, snore and noise segments; their temporal information along with the features of S_aO_2 signal were used to determine the apnea and hypopnea events. Details of the proposed method (Fig. 4.1) are discussed in the following sections.

Automatic Sound Segmentation

The first step in analyzing respiratory sounds is to remove the effects of low– and high–frequency noises. When recording respiratory sounds over the trachea, heart sounds

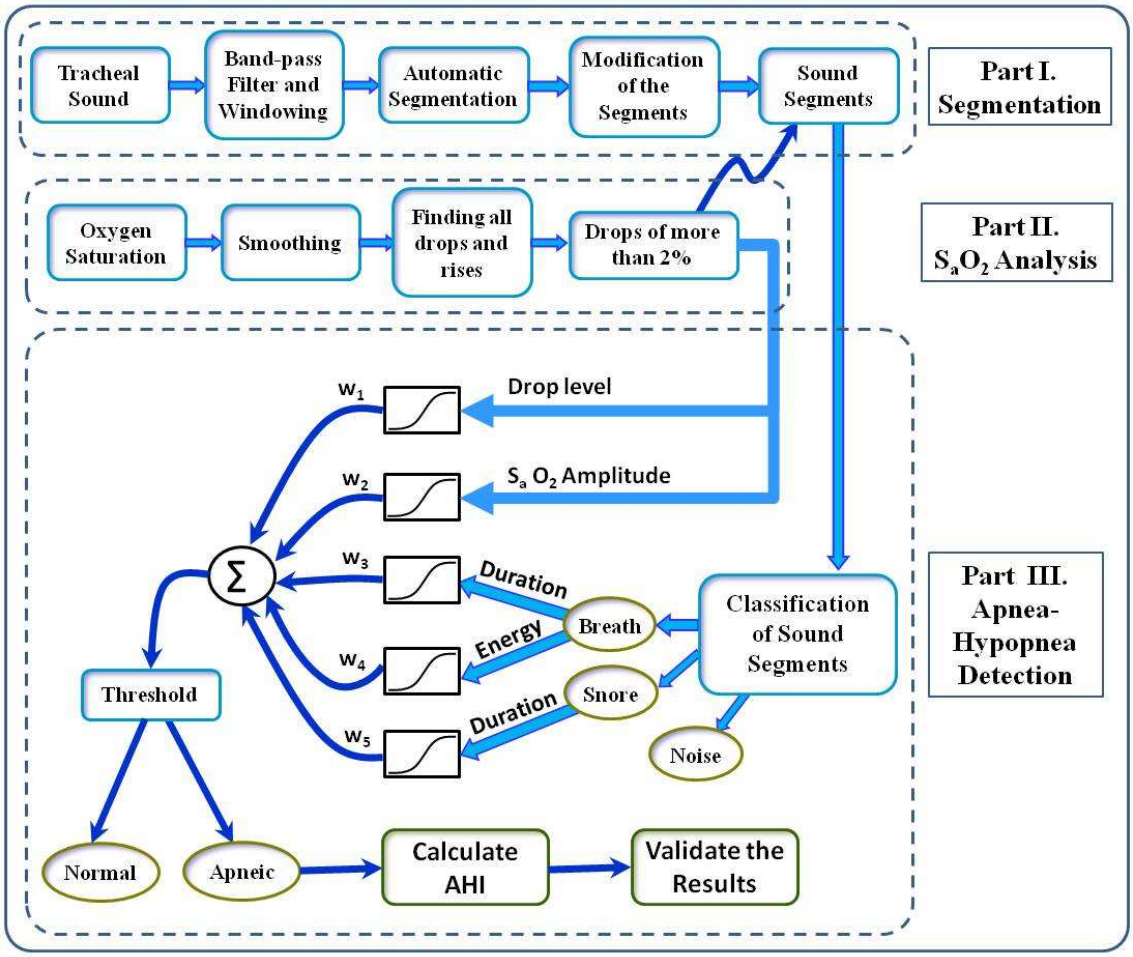


Figure 4.1: *Schematic of the proposed algorithm.*

are the main inevitable noises that are picked up by the microphone. Heart sounds are low frequency signals, and overlap with the tracheal sounds in the frequency range below $200Hz$ [21]. Different methods have been proposed for heart sounds detection and reduction [88, 89, 80]. However, all of these methods were verified on the data of healthy subjects during wakefulness when subjects were breathing normally, and other noises such as snore, movements or blanket noises were not present. Furthermore, those methods were developed for heart sounds cancellation from lung sound, which is a low frequency signal compared to tracheal sound. Since tracheal sounds have considerable energy components in the frequency range of above 200 Hz and below 1000 Hz [21, 90], in this study the recorded sounds were first bandpass filtered by a Butterworth filter of order 5 and cutoff frequency of $[200 - 1000]$ Hz to reduce the effects of heart sounds and high frequency noises, while including the main frequency components of breath and snore sounds.

The bandpass filtered sound signals were divided into windows of $20ms$ duration with 75% overlap between adjacent windows. The values of optimum window size and overlap for segmenting the tracheal sound signal were selected based on the results of our previous studies on acoustical flow estimation [1, 29]. Energy or amplitude of tracheal sounds were usually used to find the sound segments and the breathing cycles [7, 91-93]. In this study, $LogVar$ was calculated in each window which represents the signal's energy. Figure 4.2 shows a typical recorded tracheal sounds and the estimated $LogVar$ signal. The data of this particular subject were recorded with the same device during sleep, but including a

face mask pneumotachograph connected to a pressure sensor for direct measurement of flow signal for comparison with estimated flow. Comparing the recorded flow (Fig. 4.2–c) and the estimated $LogVar$ signals (Fig. 4.2–b), it can be seen that $LogVar$ follows absolute values of flow signal (Fig. 4.2–a). Figure 4.2–d shows the result of estimating relative flow from the $LogVar$, which is closely related to the corresponding recorded flow. Note that the amplitude of the estimated flow does not represent the actual amount of flow in L/s as it is the relative flow without calibration. For more details on calibration method see [29].

The median of the $LogVar$ values of all the windows of the bandpass filtered tracheal sound was calculated and used as a threshold to automatically classify each window either as a sound or silent window. Then, if two successive windows classified as sounds were not farther than the length of one window size (20 ms) apart, they were combined together, i.e., the silent portions in the middle were ignored. This process was continued by merging the small segments, i.e., equal to 20 ms, with their adjacent close segments. Since the duration of respiratory phases, i.e., inspiration/expiration, is not usually more than a second, the segments longer than $2s$ were divided into smaller segments (Fig. 4.1–Part I). Figure 4.3 shows the tracheal sound signal along with the segmentation results. The segmentation vector (dashed line) is multiplied by (-1) in successive segments for clarity purposes.

The segmentation performance was verified by manual auditory and visual inspection

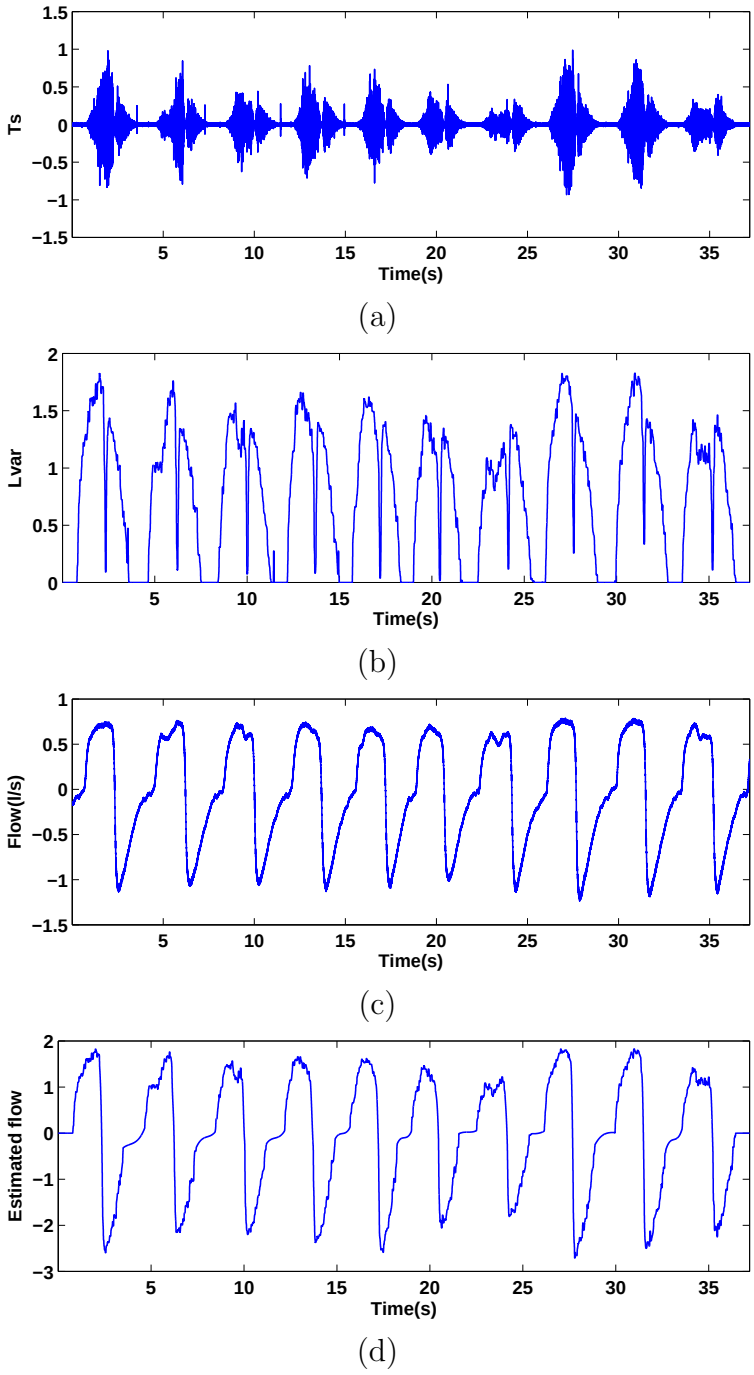


Figure 4.2: *A typical period of a) the recorded tracheal signal, b) along with its calculated $LogVar$, c) its associated respiratory flow measured by face mask pneumotachograph, and d) the estimated flow from $LogVar$.*

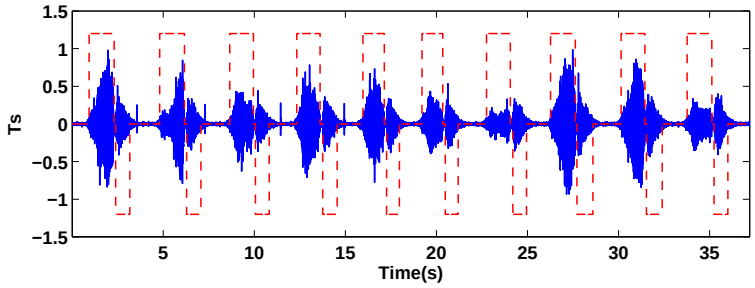


Figure 4.3: *Samples of the recorded tracheal sounds with the segmentation results. The segmentation vector (dashed line) is multiplied by* (-1) *in successive segments for clarity purposes.*

of the sound signals in the time–frequency domain. The automatic segmentation results were compared with those of the manual detection in terms of absolute delays in detecting the start and end of each sound segment, difference in the duration of each segment and the number of missed segments.

Analysis of S_aO_2 signal

Both cessation and reduction of airflow should be associated with at least 4% drop in S_aO_2 signal for being counted as an apnea/hypopnea event. S_aO_2 signal is a very low frequency signal; hence, to have a fast method it is more efficient to start the data analysis by finding the drops and rises of this signal. The S_aO_2 signal was smoothed with a median filter in windows of 150 ms. The falling and rising step changes in the S_aO_2 signal were found automatically by taking the derivative of the signal. The step changes in S_aO_2 signal which last less than 20s were merged together to get the start and end point of the fall and rise in the S_aO_2 signal (Fig. 4.1–Part II). Figure 4.4–a shows a period of S_aO_2

signal which contains a drop and a rise. The detected start and end points of the drop and rise in the S_aO_2 signal were marked by triangle and square markers, respectively.

The sound signals within the periods of S_aO_2 drop were subsequently analyzed to examine whether an apnea or hypopnea occurred in that period or not. To ensure that marginal apnea and hypopnea events would not be missed, we considered all the drops of more than 2%. On the other hand, there is usually a delay between the occurrence of apnea or hypopnea event and the drop in S_aO_2 signal; to consider this the sound segments were examined from 10 s prior to the drop of S_aO_2 signal (Fig. 4.1–Part II).

Apnea-Hypopnea detection

Sound segments classification

When recording nocturnal tracheal sounds over the neck, in addition to breath sounds other sounds such as snores and different noises including oral noises, ambient sounds, speech and blanket movements are also captured by the microphone. Oral noises are generally short in duration with large amplitude, movements are long in duration with high amplitudes and speech signals have very large amplitudes compared with the breath sounds. Snore sounds are occurred in different parts of the respiratory cycle and they usually have higher amplitude than breath sounds. When analyzing the sound signals to estimate the respiratory flow, the presence of these additional sounds is a major problem that has to be handled carefully. Therefore, a smart function was developed to first

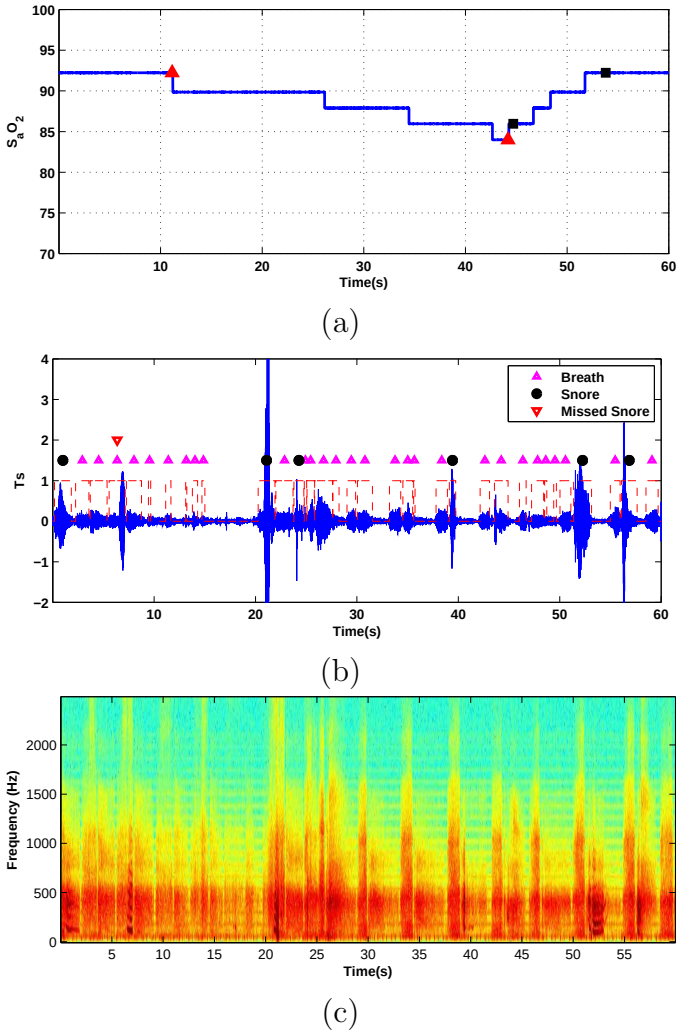


Figure 4.4: *Examples of the recorded signals during hypopnea. a) S_aO_2 signal with a drop and rise, (start and end points of the detected drop and rise were marked by triangle and square markers, respectively) b) the corresponding tracheal sound signal with segmentation vector (red dashed line) and classification results. c) spectrogram of the tracheal sound.*

classify the tracheal sound segments into breath, snore and noise segments. The program uses the sound segments' energy and duration to classify them into different groups of breath, snore and noise.

The amount of change in the tracheal sound's energy due to the flow variation is different among different people. Therefore, a period of few minutes of breath sounds segments of the subject at the beginning of each recording (when he/she was awake) was used to derive the energy level and the duration of normal breathing of the subject as a reference (self–calibration stage). To classify the sound segments, the segments which were close to each other and could be considered as pairs of successive breathing cycles, were marked. Energies of the pair segments were compared with each other. If the ratio was larger than the average plus standard deviation of the normal breath sound segments' energies (extracted in the self–calibration stage), the segment with greater energy was marked as snore and the other segment was labeled as breath. On the other hand, if the energies of the pair segments were similar and each segment's energy was in the range of normal breath sound segments' energy, they were marked as breath segments. Then, the single segments (with no segment close to them to form a pair) were marked as either snore or breath based on their energies and durations. Finally, the remaining single segments, which were shorter or longer than normal breath sounds, were labeled as noise representing oral or movement noises (Fig. 4.1–Part III).

Figure 4.4–b shows an example of the tracheal sound recorded during hypopnea (cor-

responding to the S_aO_2 signal of Fig. 4.4–a) along with the sound segments' classification results. The classification algorithm missed one snore segment and labeled it as breath segment (red downward triangle). Spectrogram of the recorded sounds is presented in Fig. 4.4–c showing the temporal and spectral changes in the energy components of the breath and snore segments. Also note the time span of $[15 - 20]s$, where there is a silent period with no detectable breathing.

Finding apnea/hypopnea events

Since neither respiratory flow nor snore sounds are present during apnea events, these episodes can be easily detected by finding the periods, in which the sounds energy is below 10% of the reference value (extracted during self–calibration stage) and its duration is more than $10s$. However, detecting hypopnea events is more complicated; they can be either very shallow breathing episodes for more than $10s$, short durations of normal breathings with periods of no–breathing in between or a combination of mildly shallow breathing and snoring that indicates partial airway obstruction (Fig. 4.4). All of these conditions may result in a deficiency in breathing and a drop in the blood's S_aO_2. However, these conditions are different for each person, and depending on position, sleep stage, etc. they can change during the night for the same person.

The following features of the sound segments and S_aO_2 signal were investigated to distinguish different situations that correspond to apnea/hypopnea event:

- The total energy of the breath sound segments (Eng_{Br}) in each period

- The duration percentage of the breath sound segments in each period (Dur_{Br})
- The duration percentage of the snore sound segments during each period (Dur_{Snr})
- The amount of drop in S_aO_2 signal (Drp_{Sat})
- The amplitude of S_aO_2 signal (Amp_{Sat})

Each feature was transformed with either the sigmoid functions $S_1(t)$ or $S_2(t)$ to represent the importance and contribution of each feature in the occurrence of an event:

$$S_1(t) = \left(1 - e^{\frac{t-a}{b-a} \times (x_2 - x_1) + x_1}\right)^{-1} \tag{4.1}$$

$$S_2(t) = 1 - S_1(t). \tag{4.2}$$

where a and b show the variation range of each parameter and $x_1 = -10$ and $x_2 = 10$ are two constants, for which the function $(1 - e^{-t})^{-1}$ converges to 0 and 1, respectively. Figure 4.5 shows the sigmoid functions S_1 and S_2. The transformed values were weighted and added together:

$$y = \sum_{i=1}^{5} w_i S(f_i), \tag{4.3}$$

where f_i represents different features (Eng_{Br}, Dur_{Br}, Dur_{Snr}, Drp_{Sat} and Amp_{Sat}), $S(.)$ is the sigmoid function (S_1 or S_2) and w_i is the weighting value of each feature. The limits of each feature (a, b), w_i and choice of sigmoid function were determined heuristically and based on the preliminary information regarding the importance of different features and

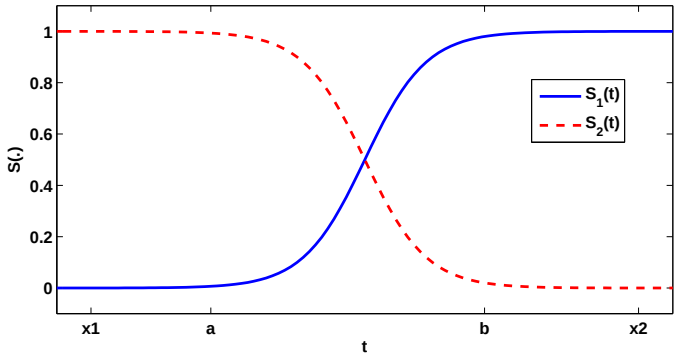


Figure 4.5: *The sigmoid functions $S_1(t)$ and $S_2(t)$ which were used to transform different features.*

their association in occurrence of apnea or hypopnea events. The value of y was compared with a threshold of Thr_{Event}; if it was less than the threshold, the period with the drop in S_aO_2 signal was considered as normal; otherwise, it was counted as an apnea/hypopnea event (Fig. 4.1–Part III). To find the Thr_{Event}, different values in the range of $[.2 - .9]$ were used as thresholds for finding apnea/hypopnea events and calculating the number of events per hour (AHI_{ASAD}). The AHI_{ASAD} values were used to classify the subjects into simple snorers and OSA patients, and the results were compared with the classification results based on the AHI values of the PSG study (AHI_{PSG}) that was manually calculated by the sleep lab technicians. The value of threshold, for which the highest accuracy was achieved, was selected as the Thr_{Event}.

The calculated Thr_{Event} was applied to re–estimate the subjects' apnea and hypopnea events. The AHI_{ASAD} and AHI_{PSG} values were compared in terms of linear correlation and Bland-Altman statistical measure [94]. Bland-Altman measure is designed to measure

the agreement between two methods that investigate the same property, and it has been widely used in sleep apnea studies to validate the performance of portable monitoring devices.

One of the main applications of sleep apnea portable monitors is to screen the patients and separate OSA patients from simple snorers for advanced diagnosis. In the last evaluation, the performance of the estimated AHI_{ASAD} values in classifying the subjects into two groups of simple snorers and OSA patients was investigated. Since PSG is considered as the gold standard, the subjects are usually grouped into simple snorers and OSA patients depending on their AHI_{PSG}. However, there is no standard threshold of AHI_{PSG} for such grouping. Researchers have used different values of AHI between 5 to 20 as the threshold between simple snorers and OSA patients [54-56, 16, 59-64]. Hence, in this study we investigated grouping of the patients with the AHI_{PSG} values of $5, 10, 15$ and 20 as the threshold, and determined what AHI_{ASAD} would correspond to those of PSG with the highest accuracy.

For each of the above mentioned four AHI_{PSG} thresholds, data were divided into train and test data sets to find the best value of AHI_{ASAD} corresponding to the selected AHI_{PSG} threshold that gives the best classification of subjects. 6–fold algorithm [95] was used to divide the subjects into train and test data sets. The patients were randomly clustered into 6 groups (11 patients in each group); data of 5 groups were selected as the train data set and the sixth group was considered as the test data set. The training data

was used to find the corresponding threshold of AHI_{ASAD} which was applied to classify the subjects in the test data set and find the classification sensitivity and specificity. The receiver operating curve (ROC) and the area under the curve (AUC) were also calculated to evaluate the classifier's performance. This process was repeated for all 6 folds as test data set and the sensitivity, specificity and AUC results were averaged. Finally, to remove the classifier's bias to the choice of train and test data sets, the whole process was repeated 200 times and the results were averaged among all trials.

4.2 Results

In the first step of the processing, the recorded signals were segmented into sound and silent segments. We have used a thresholding based technique to have a fast algorithm for detecting windows of sound (with the fixed length of 20 ms); this was followed by a smart post–processing to merge the windows and determine continuous segments of sounds with variable lengths that corresponded to different cases such as breath, noise or snore. Table 4.2 shows the mean and standard deviation values of the delays, duration errors and missed segments for 3059 breath and 1557 snore segments of 16 subjects. The errors were averaged for all the segments of every subject and among different subjects. The results indicate the method detects more than 96% of the sound segments of different lengths and intensities correctly. Considering that data were recorded in real condition with no control on the position and sleep situation of the subjects or the ambient noise, the

Table 4.2: *Mean and standard deviation values of the automatic segmentation errors.*

Segment	Delay(s)		Duration(s)	Missed(%)
	Start	End		
Breath	0.250 ± 0.216	0.216 ± 0.198	0.275 ± 0.178	3.42
Snore	0.253 ± 0.307	0.305 ± 0.214	0.305 ± 0.214	3.10

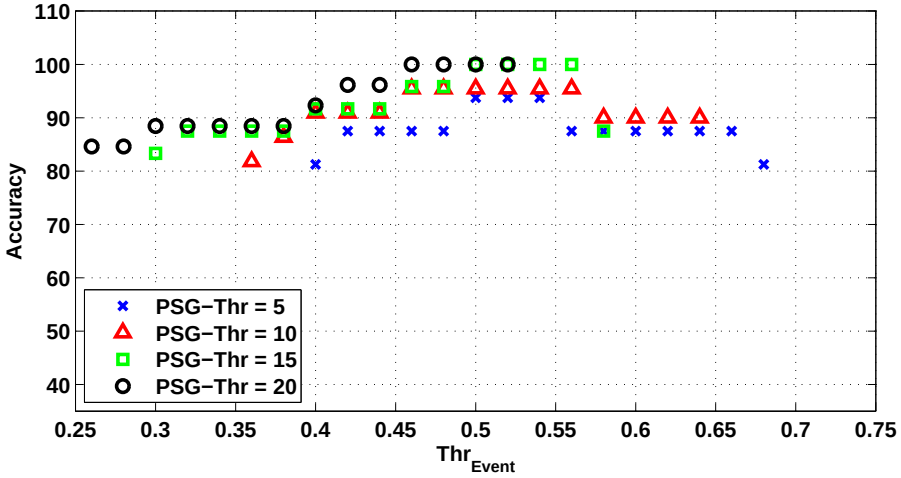


Figure 4.6: *Classification accuracy of the method for different values of* Thr_{Event}.

results are promising and reliable in detecting the sound segments with a high accuracy. Moreover, the segmentation algorithm is fully automatic and fast; the latter two are important factors in studying the overnight data of the patients.

Figure 4.6 shows the classification accuracy of the method for different values of Thr_{Event} for detecting apnea and hypopnea events. The classification was performed for different values of AHI_{PSG} and it can be seen that with the threshold of 0.5, the best possible performance was achieved for different cases. This threshold was used in the rest of study for finding apnea/hypopnea events and calculating AHI_{ASAD} values.

The classified sound segments and S_aO_2 signal were used to determine the occurrence of

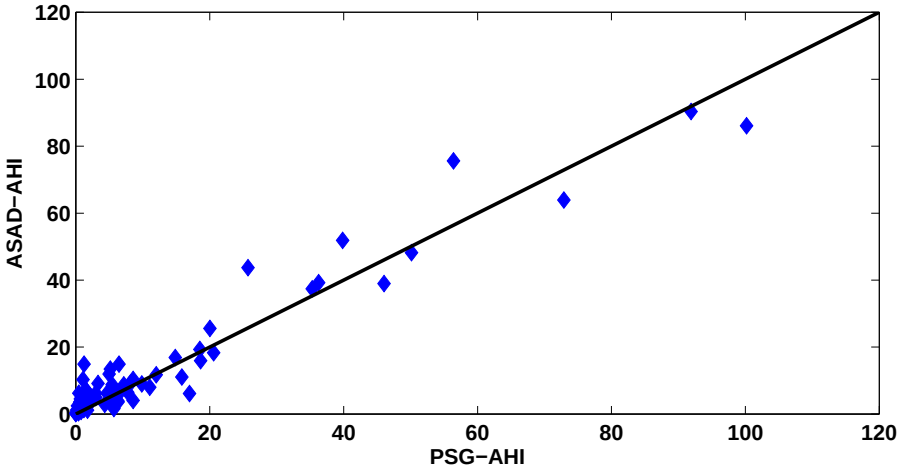


Figure 4.7: *Scatter plot of the AHI_{ASAD} and AHI_{PSG} values.*

an apnea or hypopnea event and estimate the AHI value of each subject. The AHI values of our proposed method (AHI_{ASAD}) were compared with those of the PSG study (AHI_{PSG}). Figure 4.7 shows the scatter plot of the AHI_{ASAD} and AHI_{PSG} values. The correlation ratio between the AHI_{ASAD} and AHI_{PSG} values was found to be 0.96 ($p < 0.0001$).

Bland-Altman statistical test was performed to verify the agreement between the results of ASAD and PSG systems. The average and standard deviation values were -1.56 and 5.54, respectively, and 5 out of 66 subjects were outside the 95% confidence interval (Fig. 4.8), as expected. These results confirm high correlation between the AHI_{ASAD} and AHI_{PSG} values.

Finally, AHI_{ASAD} values were used as a threshold to classify the subjects into simple snores and OSA patients. Again the AHI values of the PSG system were used as the gold standard to determine the true classes of the patients. The classification performance

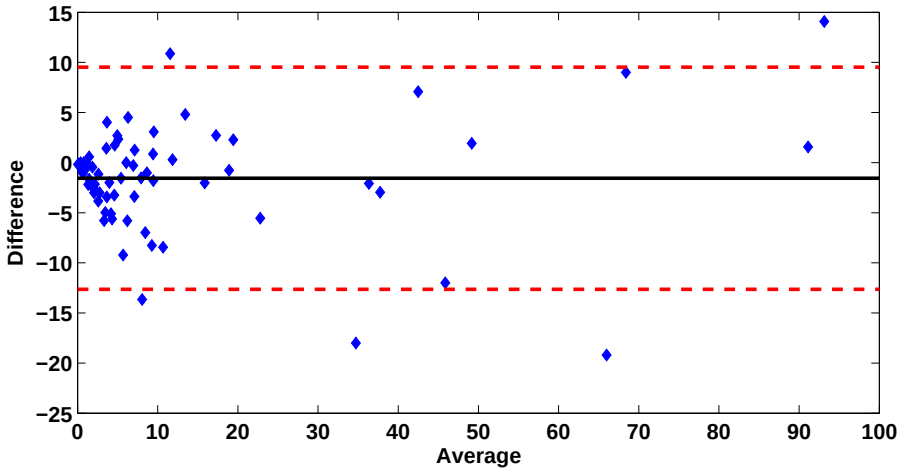


Figure 4.8: *Bland-Altman plots between the AHI_{ASAD} and AHI_{PSG}, the solid line shows the average difference and the dashed lines present the mean $\pm$ 1.96 of standard deviation (boundaries of 95% confidence interval) of the difference.*

of the method was evaluated based on specificity and sensitivity values for four different thresholds of AHI_{PSG} values $(5, 10, 15, 20)$ representing different severity levels of sleep apnea (Table 4.3). The results of Table 4.3 show that for the AHI_{PSG} thresholds of more than 10 the AUC is close to 1; this indicates the classifier has high sensitivity and specificity. For the AHI_{PSG} thresholds of more than 20 the sensitivity and specificity of the classifiers were found to be more than 91%. The high sensitivity and specificity of the proposed system's classifier was expected as the AHI values calculated by the proposed system were highly correlated with those of the PSG system.

Table 4.3: *Average ± standard deviation of specificity and sensitivity values of ASAD system for different thresholds of AHI_{PSG} and AHI_{ASAD}. The classification was repeated* 200 *times and the results were averaged.*

AHI_{PSG}	5	10	15	20
AHI_{ASAD}	8.6	13.0	18.5	23.0
Sensitivity	74.3 ± 2.7	82.8 ± 6.5	84.6 ± 7.5	91.6 ± 10.7
Specificity	82.4 ± 5.3	91.1 ± 1.9	96.0 ± 2.8	97.8 ± 0.8
AUC	0.87	0.95	0.96	0.99

4.3 Discussion

In this part of the thesis, a new automatic acoustic method was developed to detect apnea and hypopnea events with no need for respiratory flow measurement. The performance of tracheal respiratory sound and S_aO_2 signal for apnea/hypopnea detection were investigated and compared when each signal was considered alone [16]. It was shown that tracheal sound analysis had higher sensitivity than S_aO_2, while the specificity of S_aO_2 signal was higher [16]. In our proposed method, the combination of tracheal respiratory sounds and S_aO_2 signals was used to achieve higher sensitivity and specificity in sleep apnea detection and diagnosis.

In the proposed system, the sound signal recordings of the entire night (after filtering the noises such as movement noises and artefacts) are available for the user (*i.e.,* physician) to be examined by auditory and/or visual means, at the user interface of the system [96]. To increase the processing speed of analyzing the sound signals and finding apnea/hypopnea events, the system analyses only the periods of tracheal sounds that are between a drop and rise in the S_aO_2 signal, and marks the sound segments as breath,

snore and noise using the proposed method. The classifications are performed based on the information extracted from the normal breath sounds of the subject during the wake periods at the beginning of the recording. This self–calibration process is the only part of the method that requires input from the user. In each period, energy and duration of the classified breath segments are compared with the normal breathing periods extracted during the self–calibration stage to have a relative estimation of the total breathing volume. Duration of the classified snore segments, amplitude and the amount of drop in the S_aO_2 signal are the other features that are used to investigate the breathing quality. The weighted average of the features is calculated and thresholded to mark the apnea events.

The overall performance of the proposed method was evaluated by comparing its AHI values (AHI_{ASAD}) with those of the PSG (AHI_{PSG}). The correlation between the outcomes of our proposed system and PSG were found to be very high ($0.96, p < 0.0001$). According to the results of Bland-Altman test, 5 out of 66 subjects were outside of the 95% confidence interval, which is expected statistically. Among these 5 patients, 3 had high BMI values ($43.4, 47.9$ and 56.8). This explains the low performance of the method, as the sound quality degrades when there are high amounts of fats and tissue around the neck.

The AHI_{ASAD} values were used to classify the patients while the true classes were determined based on thresholding the AHI_{PSG} values as the gold standard criterion. Since, there is no standard threshold of AHI_{PSG} for classification of patients into simple

snorers and OSA patients, we used the same thresholds that are most commonly used by other researchers [16, 54-56, 59-64] as the threshold between the two groups. Hence, in this study we investigated grouping of the patients with the AHI_{PSG} values of 5, 10, 15 and 20 as the threshold, and determined what AHI_{ASAD} would correspond to those of PSG with the highest accuracy. The results are shown in Table 4.3; the closer the two thresholds are, the more correlated the results of the two systems (PSG and the proposed system) are.

For patients with mild levels of the upper airway obstruction ($AHI_{PSG} \leq 5$), the proposed system overestimates the AHI values (as presented in Fig. 4.7); this explains the low performance of the method for AHI_{PSG} threshold of 5 (Table 4.3). When increasing the AHI_{PSG} thresholds to more than 10, the AUC becomes higher than 0.95, indicating high sensitivity and specificity. For the AHI_{PSG} thresholds of more than 20, the sensitivity and specificity of the classifiers were found to be more than 91%. This high sensitivity and specificity results of the proposed system was expected as the AHI values calculated by the proposed system were highly correlated with those values calculated by PSG system. These results confirm that the AHI values calculated by our proposed system based on only two recorded signals, are good representatives of the PSG based AHI values. Thus, the proposed ASAD system may be considered as a reliable predictor of the patient's AHI and the severity level of his/her obstruction and apnea condition. The results of the proposed method are found to be better than the results of the previously proposed

portable monitoring devices [51, 54, 56, 57, 18, 60, 61, 63, 97] and similar to those reported in [55, 16, 59, 62, 64]. A detailed comparison of different methods in terms of correlation with AHI_{PSG} and classification accuracy is shown in Table 4.4.

While the accuracy of our proposed system is comparable or better than those of other current OSA monitoring systems, the main innovation of the proposed system is that it offers relative respiratory flow estimation; that can be used for several other clinical investigations such as flow limitation in patients who may also have asthma. Furthermore, since we record the respiratory breath and snore sounds, they can be used to extract clinical information regarding the physiology of upper airways and breathing pattern of the patient.

Table 4.4: *Comparison of different portable devices in estimating AHI values and classifying simple snorers from OSA patients.*

Method	Correlation	Sensitivity/Specificity (AHI_{PSG} threshold)			
		(5)	(10)	(15)	(20)
Rauscher et. al. [51]	0.90	-	-	-	-
Grover et. al. [54]	0.77	89/86	-	75/85	-
Wong et. al. [55]	-	100/43	96/71	92/86	88/100
Oslon et. al. [56]	0.74	-	-	88/70	-
Hornero et. al. [57]	-	-	82/87	-	-
Nakano et. al. [16]	0.93	93/67	-	79/95	-
Ballester et. al. [59]	-	-	95/92	-	-
Bar et. al. [60]	0.88	-	-	-	-
Zou et. al. [61]	0.92	-	-	-	-
Ng et. al. [62]	0.98	-	98/100	-	97/100
Goodrich et. al. [63]	0.9	85/67	92/88	87/82	85/94
Claman et. al. [64]	0.96	-	-	86/95	-
Ayas et. al. [97]	0.87	-	83/71	93/73	91/84
Our method	0.96	74/82	83/91	85/96	92/98

Chapter 5

Summary and Conclusion

Tracheal respiratory sounds analysis is a simple, inexpensive and non-invasive way to study the pathology of the upper airways. Recently it has attracted considerable attention for acoustical flow estimation and investigation of the upper airways abnormalities. Tracheal respiratory sounds amplitude is directly related to the breathing flow rate. However, the flow-sound relationship is highly variable among subjects which makes it challenging to estimate flow in a large population of subjects, unless the effects of different factors on the flow-sound model parameters are known.

In this thesis, a modified flow–sound model is proposed to estimate respiratory flow from tracheal sounds. The first novelty of the method is eliminating the dependency of the previous methods on calibrating the model at different flow rates. The proposed model can be calibrated at medium flow rate and is capable of estimating flow at other flow rates with an error of less than 10%. To remove the calibration process for every

individual, we investigated whether the model could be calibrated using the average values of model parameters of a data bank. Hence, we recorded tracheal sound and flow signals from 93 healthy individuals. We investigated the statistical correlation between the model parameters and anthropometric features of the subjects. The results showed that gender, height and smoking were the most significant factors that affected the model parameters. Then, we showed that a general flow estimation model could be defined for subjects with similar height and gender; hence, removing the individual calibration process. This technique simplifies the acoustical flow estimation in general applications including sleep studies and patients' screening in health care facilities.

One of the main applications of acoustical flow estimation is to examine obstruction in the upper airways. Obstructive sleep apena is characterized by periods of reduction or complete cessation of airflow during sleep. Although, acoustical flow estimation was used for sleep apnea diagnosis, in none of the previous studies, the flow–sound relationship during sleep was investigated. Furthermore, the flow–sound relationship in patients with OSA was not studied before. In the second part of the research, we recorded tracheal sound, flow data and body position from OSA patients during sleep and wakefulness. We compared the flow–sound relationship during sleep and wakefulness with respect to body position. Similar to wakefulness, the relationship between tracheal sounds average power and flow followed a power law. However, the results showed that the coefficients of this relationship changed from wakefulness to sleep. During inspiration, independent of target

flow rate, tracheal sounds were louder during sleep compared to that during wakefulness. However, such pattern did not exist during expiration.

The average power of tracheal sounds can be used to estimate the relative variations in respiratory flow and to detect apnea and hypopnea events during sleep. Another objective of this study was to investigate the variations of the flow–sound relationship and the model parameters during wakefulness and sleep. The goal was to investigate the possibility of calibrating the model parameters at wakefulness and applying them to the sleep data to achieve the quantitative estimation of flow rate. Furthermore, regardless of the dataset used to calibrate the model parameters, we investigated the accuracy and reliability of the flow estimation algorithm during sleep. Therefore, we applied two algorithms for calibrating the model parameters based on the known data recorded during sleep and wakefulness. Our results show that model parameters change from wakefulness to sleep. Therefore, if the model is calibrated using wakefulness data, the flow estimation error would be high during sleep. On the other hand, when the calibration parameters are extracted from tracheal sound and flow recordings during sleep, the flow estimation error is low. This confirms the reliability of acoustical flow estimation for investigating flow variations during sleep.

As one of the main applications of flow estimation, we applied it to sleep data to detect apnea and hypopnea events in patients with sleep apnea disorder. OSA is highly prevalent in the general population and approximately 4% of adults experience OSA

with excessive daytime sleepiness. OSA is associated with daytime sleepiness, impaired cognition, increased risk of car accidents, poor life quality and cardiovascular morbidities. The current gold standard method for sleep apnea assessment is full night PSG. Due to high cost, inconvenience and immobility of PSG, researchers have sought alternative simple and portable devices to detect sleep apnea. In this thesis, we used the results of acoustical flow estimation and developed a new method for sleep apnea detection and monitoring. The method requires only two data channels: tracheal breathing sounds and the blood's oxygen saturation level (S_aO_2). It includes an automated method that uses the energy of breathing sounds signals to segment the signals into sound and silent segments. Then, the sound segments are automatically classified into breath, snore and noise segments. The S_aO_2 signal is analyzed automatically to find its rises and drops. Finally, a weighted average of different features extracted from breath segments, snore segments and S_aO_2 signal are used to detect apnea and hypopnea events. The performance of the proposed approach was evaluated on the data of 66 patients recorded simultaneously with their full night PSG study data, and the results were compared with those of the PSG. The results show high correlation $(0.96, p < 0.0001)$ between the outcomes of our system and those of the PSG. Also, the proposed method has been found to have sensitivity and specificity values of more than 91% in differentiating simple snorers from OSA patients.

To summarize, we studied the relationship between flow and tracheal sounds in patients and healthy individuals during wake and sleep which has not been done before. Then, we

applied the results to develop a portable system for detection and monitoring of breathing disorders in patients at risk of sleep apnea. Considering the high accuracy of our developed system and its simple and portable design, it enables the physician to monitor patients at their home without the need for referring them to the sleep lab. This will significantly reduce the long waiting list of the patients which is currently a major concern in sleep studies. Furthermore, pre–screening results of our system can help the physicians to detect the patients at high risk of sleep apnea and refer them to the sleep lab for complete full–night diagnosis.

5.1 Future Work Recommendations

The main contributions of this thesis are:

1. Investigating the flow–sound relationship in healthy individuals and OSA patients during sleep and wakefulness.

2. Developing a robust flow estimation algorithm without the need for individual calibration.

3. Developing an acoustic method for sleep apnea detection and monitoring.

Considering the results of this thesis, a couple of suggestions can be recommended for future studies.

1. Investigating resistance and collapsibility of the upper airways using acoustic methods:

   The causes of OSA vary considerably between individuals; however, anatomy of the upper airway and disturbances in its neuromuscular control are some of the major factors that affect upper airway collapsibility and resistance. Age, male sex, obesity and fluid accumulation around the neck in patients with heart failure are major factors contributing to narrowing, increased resistance and collapsibility of the upper airway by altering upper airway length, wall thickness and cross–sectional area. Considering this physiological background, the results of higher sound intensities during sleep than those of wakefulness in OSA patients and the significant effects of height, gender and smoking on the flow–sound relationship in healthy individuals can lead us to hypothesize that intensity and non–linear characteristics of the tracheal sounds change according to anatomical and physiological characteristics of the upper airways; hence, sound analysis can be used to assess collapsibility and resistance of the upper airways.

   To validate this hypothesis, respiratory sounds can be recorded from healthy individuals, non–obese OSA patients, obese OSA patients and OSA patients with heart failure during both wakefulness and sleep. Sounds can be analyzed in time and frequency domains. For every subject, simultaneously with sound recordings, upper airways collapsibility and airflow resistance should be measured before and dur-

ing sleep in different body positions. Upper airways collapsibility can be measured by applying negative airway pressures and determining upper airways closing pressure. Upper airways resistance can be calculated by dividing pharyngeal pressure and airflow rate. The extracted sound features can be examined to determine the features that correlate significantly with upper airway resistance and collapsibility. The results will improve the accuracy of acoustical features for characterizing the underlying variations in the anatomy and physiology of patients' upper airways; this will be helpful to improve the phenotypic classification of the patients, and possibly tailor therapy according to the phenotype.

2. Investigating flow limitation in the upper airways:

   Flow limitation in the upper airways is a critical situation that happens during airway obstruction. The gold standard for detecting flow limitation is to monitor pharyngeal pressure and flow rate simultaneously, which is an invasive routine and challenging to apply on a regular basis. One of the outcomes of our study on investigating the flow–sound relationship during sleep and wake in patients with sleep apnea was to show that tracheal respiratory sound energy represents the pharyngeal pressure in the upper airways. Therefore, it is hypothesized that the use of sound energy and flow simultaneously can detect flow limitation in the airways.

3. Improving performance of the proposed acoustical algorithms for sleep apnea detec-

tion and monitoring:

In the proposed sleep apnea detection method, a self–calibration algorithm was applied to find the parameters of normal breathing rate of every individual. This information was extracted from the patient's data during wakefulness. Considering the results of flow–sound relationship during sleep and wakefulness, it can be concluded that a better performance would be achieved by using periods of normal breathing during sleep for self calibration stage.

On the other hand, the acoustical sleep apnea diagnosis can be improved through enhancing the algorithms for noise removal and snore detection. We developed an automatic method to classify breath and snore sound segments based on their energy, zero crossing rate and formants of the sound signals [98]. For every sound segment, the number of zero crossings, logarithm of the signal's energy and the first formant were calculated. Fischer Linear Discriminant was implemented to transform the 3–dimensional (3D) feature set to a 1–dimensional (1D) space and the Bayesian threshold was applied on the transformed features to classify the sound segments into either snore or breath classes. Three sets of experiments were implemented to investigate the method's performance for different training and test data sets extracted from different neck positions. The overall accuracy of all experiments for tracheal recordings were found to be more than 90% in classifying breath and snore sounds segments regardless of the neck position (details of the method can be found

in Appendix B). These results show that by incorporating formants in the ASAD system for snore detection, we may improve the overall performance of the method in detecting apnea and hypopnea events.

4. Modify the ASAD to give position data:

   In the proposed sleep apnea detection system, the position information was not recorded. The new system can include an accelerometer to record body position and calculate AHI values for every position. This information will help physicians to determine the severity of sleep apnea in different sleep positions, which is an important factor in assessing sleep apnea disorder.

5. Investigating sound signals to detect flow limitations during sleep:

   In this study different algorithms for automatic classification of breath sounds from snore sounds were developed. During snoring, flow limitation occurs in the upper airway and the amount of flow limitation and shape of respiratory flow during these periods convey important information for physicians. Developing algorithms to estimate respiratory flow and its limitation in the periods including snore sounds will be a practical research to follow in future.

6. Investigating flow estimation during wakefulness in COPD population:

   Studying flow–sound relationship and flow estimation accuracy in patients with chronic obstructive pulmonary disease (COPD) can be another field of research to

follow in future. COPD is a common lung disease in which airways are narrowed and limit the passage of air to and from the lungs. Sound analysis can be used to assess the amount of obstruction in the airways. Furthermore, by recording sounds from different locations over the chest, it may be possible to detect the obstruction sites.

Bibliography

[1] A. Yadollahi and Z. Moussavi, "Acoustical flow estimation: Review and validation," *IEEE Magazine in Biomedical Engineering*, vol. 26, No. 1, pp. 56–61, 2007.

[2] S. Tarrant, R. Ellis, F. Flack, and W. Selley, "Comparative review of techniques for recording respiratory events at rest and during deglutition," *Dysphagia*, vol. 12, pp. 24–38, 1997.

[3] R. Gilbert, J. Auchincloss, J. Brodsky, and W. Boden, "Changes in tidal volume frequency and ventilation induced by their measurement," *J. Appl. Physiol.*, vol. 33, p. 1972, 252-254.

[4] J. Ashkanazi, P. Silverberg, R. Foster, A. Hyman, J. Milic-Emili, and J. Kinney, "Effects of respiratory apparatus on breathing pattern," *J. Appl. Physiol.*, vol. 48, pp. 577–580, 1980.

[5] J. Hirsch and B. Bishop, "Human breathing patterns on mouthpiece or facemask during air, co2, or low o2," *J. Appl. Physiol.*, vol. 53, p. 1982, 1281-1290.

[6] C. Weissman, J. Ashkanazi, J. Milic-Emili, and J. Kinney, "Effect of respiratory apparatus on respiration," *J. Appl. Physiol.*, vol. 57, p. 1984, 457-480.

[7] Z. Moussavi, M. Leopando, H. Pasterkamp, and G. Rempel, "Computerized acoustical respiratory phase detection without airflow measurement," *Medical & Biolog Eng. & comp.*, vol. 38(2), pp. 198–203, 2000.

[8] C. Que, C. Kolmaga, L. Durand, S. Kelly, and P. Macklem, "Phonospirometry for noninvasive measurement of ventilation: methodology and preliminary results," *J. Appl. Physical.*, vol. 93, pp. 1515–1526, 2002.

[9] N. Gavriely and D. Cugell, "Airflow effects on amplitude and spectral content of normal breath sounds." *Journal of applied physiology*, vol. 80, No. 1, pp. 5–13, 1996.

[10] G. Soufflet, G. Charbonneau, M. Polit, P. Attal, A. Denjean, P. Escourrou, and C. Gaultier, "Interaction between tracheal sound and flow rate: a comparison of some different flow evaluations from lung sounds," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 4, pp. 384–391, 1990.

[11] M. Oud and E. Maarsingh, "Spirometry and forced oscillometry assisted optimal frequency band determination for the computerized analysis of tracheal lung sounds in asthma," *Physiol Meas.*, vol. 25(3), pp. 595–606, 2004.

[12] J. Fiz, R. Jan, A. Homs, J. Izquierdo, M. Garca, and J. Morera, "Detection of wheezing during maximal forced exhalation in patients with obstructed airways," *Chest*, vol. 122(1), pp. 186–91, 2002.

[13] S. Rietveld, M. Oud, L. Rijssenbeek-Nouwens, D. Vaghi, and E. Dooijes, "Characteristics and diagnostic significance of spontaneous wheezing in children with asthma: results of continuous in vivo sound recording," *J Asthma*, vol. 36(4), pp. 351–8, 1999.

[14] M. Yonemaru, K. Kikuchi, M. Mori, A. Kawai, T. Abe, T. Kawashiro, T. Ishihara, and T. Yokoyama, "Detection of tracheal stenosis by frequency analysis of tracheal sounds," *J Appl Physiol.*, vol. 75(2), pp. 605–12, 1993.

[15] A. Yadollahi, E. Giannouli, and Z. Moussavi, "Sleep apnea monitoring and diagnosis based on pulse oximetery and tracheal sound signals," *Medical and Biological Engineering and Computing*, vol. 48(11), pp. 1087–1097, 2010.

[16] H. Nakano, M. Hayashi, E. Ohshima, N. Nishikata, and T. Shinohara, "Validation of a new system of tracheal sound analysis for the diagnosis of sleep apnea-hypopnea syndrome," *Sleep*, vol. 27(5), pp. 951–7, 2004.

[17] H. Pasterkamp and I. Sanchez, "Tracheal sounds in upper airway obstruction." *Chest*, vol. 102(3), pp. 963–5, 1992.

[18] J. Cummiskey, T. Williams, P. Krumpe, and C. Guilleminault, "The detection and quantification of sleep apnea by tracheal sound recordings." *Am Rev Respir Dis.*, vol. 126(2), pp. 221–4, 1982.

[19] R. Beck, G. Rosenhouse, M. Mahagnah, R. Chow, D. Cugell, and N. Gavriely, "Measurements and theory of normal tracheal breath sounds," *Ann Biomed Eng.*, vol. 33(10), pp. 1344–51, 2005.

[20] C. Bertram, "Flow-induced oscillation of collapsed tubes and airway structures," *Respir Physiol Neurobiol.*, vol. 163(1-3), pp. 256–65, 2008.

[21] H. Pasterkamp, S. Kraman, and G. Wodicka, "Respiratory sounds, advances beyond the stethoscope," *Am J Respir Crit Care Med*, vol. 156, pp. 974–987, 1997.

[22] I. Sanchez and H. Pasterkamp, "Tracheal sound spectra depend on body height," *Am Rev Respir Dis.*, vol. 148(4 Pt 1), pp. 1083–7, 1993.

[23] S. Kraman, H. Pasterkamp, M. Kompis, M. Takase, and G. Wodicka, "Effects of breathing pathways on tracheal sound spectral features," *Respir. Physiol.*, vol. 111, pp. 295–300, 1998.

[24] V. Harper, S. Kraman, H. Pasterkamp, and G. Wodicka, "An acoustic model of the respiratory tract," *IEEE Trans. Biomed. Eng.*, vol. 48, no.5, pp. 543–550, 2001.

[25] V. Harper, H. Pasterkamp, H. Kiyokawa, and G. Wodicka, "Modeling and measurement of flow effects on tracheal sounds," *IEEE Trans. Biomed. Eng.*, vol. 50, no.1, pp. 1–10, 2003.

[26] Y. Yap and Z. Moussavi, "Acoustic airflow estimation from tracheal sound power," in *Proc. IEEE Canadian Conf. Elec. Comp. Eng. (CCECE)*, 2002, pp. 1073–1076.

[27] I. Hossain and Z. Moussavi, "Respiratory airflow estimation by acoustical means," in *Proc. Second joint EMBS/BMES Conf.*, Houston, TX, USA, 2002, pp. 1476–1477.

[28] M. Golabbakhsh, "Tracheal breath sound relationship with respiratory flow: Modeling, the effect of age and airflow estimation," Master's thesis, Electrical and Computer Engineering Department, University of Manitoba, 2004.

[29] A. Yadollahi and Z. Moussavi, "A robust method for estimating respiratory flow using tracheal sounds entropy," *IEEE Transactions on Biomedical Engineering*, vol. 53, No. 4, pp. 662–668, 2006.

[30] C. Lessard and W. Wong, "Correlation of constant flow rate with frequency spectrum of respiratory sounds when measured at the trachea," *IEEE Trans. Biomed. Eng.*, vol. 33, pp. 461–463, 1986.

[31] M. Mussell and Y. Miyamoto, "Comparison of normal respiratory sounds recorded from the chest and trachea at various respiratory air flow levels," *Frontiers Med. Biol. Engmg.*, vol. 4, pp. 73–85, 1992.

[32] V. Gross, A. Dittmar, T. Penzel, F. Schttler, and P. von Wichert, "The relationship between normal lung sounds, age, and gender," *Am J Respir Crit Care Med.*, vol. 162(3 Pt 1), pp. 905–9, 2000.

[33] American Academy of Sleep Medicine Task force, S. Quan, J. Gillin, M. Littner, and J. Shepard, "Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research; the report of an american academy of sleep medicine task force," *Sleep*, vol. 22, pp. 667–689, 1999.

[34] C. Kushida, M. Littner, T. Morgenthaler, C. Alessi, D. Bailey, J. J. Coleman, L. Friedman, M. Hirshkowitz, S. Kapen, M. Kramer, T. Lee-Chiong, D. Loube, J. Owens, J. Pancer, and M. Wise, "Practice parameters for the indications for polysomnography and related procedures: an update for 2005," *Sleep*, vol. 28(4), pp. 499–521, 2005.

[35] W. Flemons, M. Littner, J. Rowley, P. Gay, W. Anderson, D. Hudgel, R. McEvoy, and D. Loube, "Home diagnosis of sleep apnea: a systematic review of the literature," *Chest*, vol. 124, pp. 1543–79, 2003.

[36] T. Young, M. Palta, J. Dempsey, J. Skatrud, S. Weber, and S. Badr, "The occurrence of sleep-disordered breathing among middle-aged adults," *N Engl J Med*, vol. 328, pp. 1230–1235, 1993.

[37] J. Pagel, "Excessive daytime sleepiness," *Am Fam Physician*, vol. 79(5), pp. 391–6, 2009.

[38] E. Kezirian, S. Harrison, S. Ancoli-Israel, S. Redline, K. Ensrud, A. Goldberg, D. Claman, A. Spira, and K. Stone, "Study of osteoporotic in men research group, behavioral correlates of sleep-disordered breathing in older men," *Sleep*, vol. 32(2), pp. 253–61, 2009.

[39] R. Rakel, "Clinical and societal consequences of obstructive sleep apnea and excessive daytime sleepiness," *Postgrad Med.*, vol. 121(1), pp. 86–95, 2009.

[40] R. Bhattacharjee, L. Kheirandish-Gozal, G. Pillar, and D. Gozal, "Cardiovascular complications of obstructive sleep apnea syndrome: evidence from children," *Prog Cardiovasc Dis.*, vol. 51(5), pp. 416–33, 2009.

[41] A. Calvin and V. Somers, "Obstructive sleep apnea and risk of stroke: time for a trial," *Nat Clin Pract Cardiovasc Med.*, vol. 6(2), pp. 90–1, 2009.

[42] N. Takama and M. Kurabayashi, "Influence of untreated sleep-disordered breathing on the long-term prognosis of patients with cardiovascular disease," *Am J Cardiol.*, vol. 103(5), pp. 730–4, 2009.

[43] G. Hamilton, I. Meredith, A. Walker, and P. Solin, "Obstructive sleep apnea leads to transient uncoupling of coronary blood flow and myocardial work in humans," *Sleep*, vol. 32(2), pp. 263–70, 2009.

[44] F. Barb, J. Perics, A. Muoz, L. Findley, J. Ant, and A. Agust, "Automobile accidents in patients with sleep apnea syndrome. an epidemiological and mechanistic study," *Am J Respir Crit Care Med*, vol. 158(1), pp. 18–22, 1998.

[45] J. Tern-Santos, A. Jimnez-Gmez, and J. Cordero-Guevara, "The association between sleep apnea and the risk of traffic accidents. cooperative group burgos-santander," *N Engl J Med.*, vol. 340(11), pp. 847–51, 1999.

[46] T. Shiomi, A. Arita, R. Sasanabe, K. Banno, H. Yamakawa, R. Hasegawa, K. Ozeki, M. Okada, and A. Ito, "Falling asleep while driving and automobile accidents among patients with obstructive sleep apnea-hypopnea syndrome," *Psychiatry Clin Neurosci*, vol. 56(3), pp. 333–4, 2002.

[47] F. Pizza, S. Contardi, S. Mondini, L. Trentin, and F. Cirignotta, "Daytime sleepiness and driving performance in patients with obstructive sleep apnea: comparison of

the mslt, the mwt, and a simulated driving task," *Sleep*, vol. 32(3), pp. 382–91, 2009.

[48] C. Fornas, E. Ballester, E. Arteta, C. Ricou, A. Diaz, A. Fernandez, J. Alonso, and J. Montserrat, "Measurement of general health status in obstructive sleep apnea hypopnea syndrome," *Sleep*, vol. 18(10), pp. 876–9, 1995.

[49] W. Flemons and W. Tsai, "Quality of life consequences of sleep-disordered breathing," *J Allergy Clin Immunol*, vol. 99(2), pp. S750–S756, 1997.

[50] A. Pack and I. Gurubhagavatula, "Economic implications of the diagnosis of obstructive sleep apnea," *Ann Intern Med*, vol. 130(6), pp. 533–534, 1999.

[51] H. Rauscher, W. Popp, and H. Zwick, "Computerized detection of respiratory events during sleep from rapid increases in oxyhemoglobin saturation," *Lung*, vol. 169(6), pp. 335–42, 1991.

[52] J. Hosselet, R. Norman, I. Ayappa, and D. Rapoport, "Detection of flow limitation with a nasal cannula/pressure transducer system," *Am J Respir Crit Care Med.*, vol. 157(5 Pt 1), pp. 1461–7, 1998.

[53] J. Salisbury and Y. Sun, "Rapid screening test for sleep apnea using a nonlinear and nonstationary signal processing technique," *Med Eng Phys.*, vol. 29(3), pp. 336–43, 2007.

[54] S. Grover and S. Pittman, "Automated detection of sleep disordered breathing using a nasal pressure monitoring device," *Sleep Breath.*, vol. 12(4), pp. 339–45, 2008.

[55] K. Wong, D. Jankelson, A. Reid, G. Unger, G. Dungan, J. Hedner, and R. Grunstein, "Diagnostic test evaluation of a nasal flow monitor for obstructive sleep apnea detection in sleep apnea research," *Behav Res Methods.*, vol. 40(1), pp. 360–6, 2008.

[56] L. Olson, A. Ambrogetti, and S. Gyulay, "Prediction of sleep-disordered breathing by unattended overnight oximetry," *J Sleep Res.*, vol. 8(1), pp. 51–5, 1999.

[57] R. Hornero, D. Alvarez, D. Absolo, F. del Campo, and C. Zamarrn, "Utility of approximate entropy from overnight pulse oximetry data in the diagnosis of the obstructive sleep apnea syndrome," *IEEE Trans Biomed Eng.*, vol. 54(1), pp. 107–13, 2007.

[58] A. Kulkas, E. Huupponen, J. Virkkala, M. Tenhunen, A. Saastamoinen, E. Rauhala, and S. Himanen, "New tracheal sound feature for apnoea analysis," *Med Biol Eng Comput.*, vol. 47(4), pp. 405–12, 2009.

[59] E. Ballester, M. Solans, X. Vila, L. Hernandez, L. Quint, I. Bolivar, S. Bardagi, and J. Montserrat, "Evaluation of a portable respiratory recording device for detecting apnoeas and hypopnoeas in subjects from a general population," *Eur Respir J.*, vol. 16(1), pp. 123–7, 2000.

[60] A. Bar, G. Pillar, I. Dvir, J. Sheffy, R. Schnall, and P. Lavie, "Evaluation of a portable device based on peripheral arterial tone for unattended home sleep studies," *Chest*, vol. 123(3), pp. 695–703, 2003.

[61] D. Zou, L. Grote, Y. Peker, U. Lindblad, and J. Hedner, "Validation a portable monitoring device for sleep apnea diagnosis in a population based cohort using synchronized home polysomnography," *Sleep*, vol. 29(3), pp. 367–74, 2006.

[62] S. Ng, T. Chan, K. To, J. Ngai, A. Tung, F. Ko, and D. Hui, "Validation of a portable recording device (apnealink) for identifying patients with suspected obstructive sleep apnea syndrome (OSAS)," *Intern Med J.*, vol. 39(11), pp. 757–62, 2009.

[63] S. Goodrich and W. Orr, "An investigation of the validity of the lifeshirt in comparison to standard polysomnography in the detection of obstructive sleep apnea," *Sleep Med.*, vol. 10(1), pp. 118–22, 2009.

[64] D. Claman, A. Murr, and K. Trotter, "Clinical validation of the bedbugg in detection of obstructive sleep apnea," *Otolaryngol Head Neck Surg.*, vol. 125(3), pp. 227–30, 2001.

[65] P. Tiihonen, A. Paakkonen, E. Mervaala, T. Hukkanen, and J. Toyras, "Design, construction and evaluation of an ambulatory device for screening of sleep apnea,"

Med Biol Eng Comput, vol. 47, pp. 59–66, 2009.

[66] A. Chesson, R. Berry, and A. Pack, "Practice parameters for the use of portable monitoring devices in the investigation of suspected obstructive sleep apnea in adults. a joint project sponsored by the american academy of sleep medicine, the american thoracic society, and the american college of chest physicians," *Sleep*, vol. 26(7), pp. 907–913, 2003.

[67] A. Yadollahi and Z. Moussavi, "A novel approach for acoustical respiratory flow estimation without the need for individual calibration," *IEEE Transaction on Biomedical Engineering*, vol. EPub, Jan 2011.

[68] H. Davies and J. Ffowcs-Williams, "Aerodynamic sound generation in a pipe," *J . Fluid Mech.*, vol. 32(4), pp. 765–778, 1968.

[69] D. Olson, M. Bogyi, D. Schwartz, and J. Hammersley., "Relationship of tracheal breath sounds to airflow," *Am. Rev. Respir. Dis.*, vol. 129, p. A256, 1984.

[70] F. White, *Fluid mechanics*. New York: McGraw-Hill, 1994.

[71] J. Fredberg, "Pseudo-sound generation at atherosclerotic constrictions in arteries." *Bull Math Biol.*, vol. 36(2), pp. 143–55, 1974.

[72] A. Sovijarvi, L. Malmberg, G. Charbonneau, J. Vanderschoot, F. Dalmasso, C. Sacco, M. Rossi, and J. Earis, "Characteristic of breath sounds and adventi-

tious respiratory sounds," *European Respiratory Review*, vol. 10 (77), pp. 591–596, 2000.

[73] V. Iyer, P. Ramamorthy, H. Fan, and Y. Ploysongsang, "Reduction of heart sounds from respiratory sounds by adaptive filtering," *IEEE Trans. Biomed. Eng.*, vol. 33, no. 12, pp. 1141–1148, 1986.

[74] M. Kompis and E. Russi, "Adaptive heart-noise reduction of respiratory sounds recorded by a single microphone," in *Proc. 14th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 1992, pp. 691–692.

[75] L. Hadjileontiadis and S. Panas, "Adaptive reduction of heart sounds from lung sounds using forth-order statistics," *IEEE Trans. Biomed. Eng.*, vol. 44(7), pp. 642–48, 1997.

[76] J. Gnitecki and Z. Moussavi, "Review: Cancellation of heart sounds from lung sounds," *IEEE, Engineering in Medicine & Biology magazine*, vol. 26(1), pp. 20–29, 2007.

[77] L. Hadjileontiadis and S. Panas, "A wavelet based reduction of heart sound noise from lung sounds," *Int. Journal of Med. Informatics*, vol. 52, pp. 183–190, 1998.

[78] I. Hossain and Z. Moussavi, "An overview of heart-noise reduction of lung sound using wavelet transform based filter," in *Proc. IEEE Eng. Med. Biol. Soc.*, 2003, pp. 458–461.

[79] M. Pourazad, Z. Moussavi, and G. Thomas, "Heart sound cancellation from lung sound recordings using time-frequency filtering," *Medical and biological engineering and computing*, vol. 44, No. 3, pp. 216–25, 2006.

[80] D. Floras, Z. Moussavi, and G. Thomas, "Heart sound cancellation based on multiscale product and linear prediction," *IEEE, Trans. Biomed. Eng.*, vol. 54(2), pp. 234–43, 2007.

[81] R. Wang, H. Tai, C. Xie, X. Wang, J. Wright, and A. Churg, "Cigarette smoke produces airway wall remodeling in rat tracheal explants," *Am J Respir Crit Care Med.*, vol. 168(10), pp. 1232–6, 2003.

[82] R. Robinson, M. Oldham, R. Clinkenbeard, and P. Rai, "Experimental and numerical smoke carcinogen deposition in a multi-generation human replica tracheobronchial model," *Ann Biomed Eng.*, vol. 34(3), pp. 373–83, 2006.

[83] I. Sanchez, A. Avital, I. Wong, A. Tal, and H. Pasterkamp, "Acoustic vs. spirometric assessment of bronchial responsiveness to methacholine in children," *Pediatr. Pulmonol.*, vol. 15, No. 1, pp. 28–35, 1993.

[84] A. Malhotra, Y. Huang, R. Fogel, G. Pillar, J. Edwards, R. Kikinis, S. Loring, and D. White, "The male predisposition to pharyngeal collapse: importance of airway length." *Am. J. Respir. Crit. Care Med.*, vol. 166(10), pp. 1388–95, 2002.

[85] M. Younes, "Contributions of upper airway mechanics and control mechanisms to severity of obstructive apnea." *Am. J. Respir. Crit. Care. Med.*, vol. 168(6), pp. 645–58, 2003.

[86] L. Huang and J. Williams, "Neuromechanical interaction in human snoring and upper airway obstruction," *J Appl Physiol.*, vol. 86(6), pp. 1759–63, 1999.

[87] J. Grotberg and S. Davis, "Fluid-dynamic flapping of a collapsible channel: sound generation and flow limitation," *J. Biomech.*, vol. 13, pp. 219–30, 1980.

[88] L. Hadjileontiadis and S. Panas, "Adaptive reduction of heart sounds from lung sounds using forth-order statistics," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 7, pp. 642–648, 1997.

[89] A. Yadollahi and Z. Moussavi, "A robust method for heart sound localization using lung sounds entropy," *IEEE Transactions on Biomedical Engineering*, vol. 53, No. 3, pp. 497–502, 2006.

[90] J. Fiz, J. Abad, R. Jane, M. Riera, M. Mananas, P. Caminal, D. Rodenstein, and J. Morera, "Acoustic analysis of snoring in patients with simple snoring and obstructive sleep apnea," *Eur. Respir. J.*, vol. 9, pp. 2365–70, 1996.

[91] P. Hult, T. Fjallbrant, B. Wranne, O. Engdahl, and P. A. P, "An improved bioacoustic method for monitoring respiration," *Technol. Health Care*, vol. 12, pp. 323–32, 2004.

[92] A. Karunajeewa, U. Abeyratne, and C. Hukins, "Silence-breathing-snore classification from snore-related sounds," *Physiol Meas.*, vol. 29(2), pp. 227–43, 2008.

[93] A. Kulkas, E. Huupponen, J. Virkkala, M. Tenhunen, A. Saastamoinen, E. Rauhala, and S. Himanen, "Intelligent methods for identifying respiratory cycle phases from tracheal sound signal during sleep," *Computers in Biology and Medicine*, vol. 39, pp. 1000–1005, 2009.

[94] J. Bland and D. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *Lancet*, vol. 1(8476), pp. 307–310, 1986.

[95] R. Duda, P. Hart, and D. Stork, *Pattern Classification.* John Wiley and Sons, Inc., 2001.

[96] Z. Moussavi, A. Yadollahi, and S. Camorlinga, "Breathing sound analysis for detection of sleep apnea/popnea events," USPTO Patent Application # : 20080243014.

[97] N. Ayas, S. Pittman, M. MacDonald, and D. White, "Assessment of a wrist-worn device in the detection of obstructive sleep apnea," *Sleep Med.*, vol. 4(5), pp. 435–42, 2003.

[98] A. Yadollahi and Z. Moussavi, "Automatic breath and snore sounds classification from tracheal and ambient sounds recordings," *Medical Engineering and Physics*, vol. 32, pp. 985–990, 2010.

[99] A. Papoulis, *Probability, random variables and stochastic processes.* McGraw-Hill Co., 1991.

[100] K. Wilson, R. Stoohs, T. Mulrooney, L. Johnson, C. Guilleminault, and Z. Huang, "The snoring spectrum: acoustic assessment of snoring sound intensity in 1,139 individuals undergoing polysomnography," *Chest*, vol. 115, pp. 762–70, 1999.

[101] F. Dalmasso and R. Prota, "Snoring: analysis, measurement, clinical implications and applications," *Eur Respir J*, vol. 9, pp. 146–159, 1996.

[102] J. Perez-Padilla, E. Slawinski, L. Difrancesco, R. Feige, J. Remmers, and W. Whitelaw, "Characteristics of the snoring noise in patients with and without occlusive sleep apnea," *Am Rev Respir Dis.*, vol. 147(3), pp. 635–44, 1993.

[103] R. Beck, M. Odeh, A. Oliven, and N. Gavriely, "The acoustic properties of snores," *Eur Respir J.*, vol. 8(12), pp. 2120–8, 1995.

[104] A. McCombe, V. Kwok, and W. Hawke, "An acoustic screening test for obstructive sleep apnoea," *Clinical otolaryngology and allied sciences*, vol. 20, No.4, pp. 348–51, 1995.

[105] J. Osborne, E. Osman, P. Hill, B. Lee, and C. Sparkes, "A new acoustic method of differentiating palatal from non-palatal snoring," *Clin. Otolaryngol.*, vol. 24(2), pp. 130–3, 1999.

[106] P. Hill, B. Lee, J. Osborne, and E. Osman, "Palatal snoring identified by acoustic crest factor analysis," *Physiol. Meas.*, vol. 20(2), pp. 167–74, 1999.

[107] S. Agrawal, P. Stone, K. McGuinness, J. Morris, and A. Camilleri, "Sound frequency analysis and the site of snoring in natural and induced sleep," *Clin. Otolaryngol.*, vol. 27(3), pp. 162–6, 2002.

[108] J. Sola-Soler, R. Jane, J. Fiz, and J. Morera, "Spectral envelope analysis in snoring signals from simple snorers and patients with obstructive sleep apnea," in *IEEE-EMBS*, Cancun, Mexico, 2003, pp. 2527–2530.

[109] N. Saunders, P. Tassone, G. Wood, A. Norris, M. Harries, and B. Kotecha, "Is acoustic analysis of snoring an alternative to sleep nasendoscopy," *Clin. Otolaryngol.*, vol. 29(3), pp. 242–6, 2004.

[110] U. Abeyratne, A. Wakwella, and C. Hukins, "Pitch jump probability measures for the analysis of snoring sounds in apnea," *Physiol. Meas.*, vol. 26, pp. 779–98, 2005.

[111] J. Sola-Soler, R. Jane, J. Fiz, and J. Morera, "Variability of snore parameters in time and frequency domains in snoring subjects with and without obstructive sleep apnea," in *IEEE-EMBS*, Shanghai, China, 2005, pp. 2583–2586.

[112] M. Cavusoglu, T. Ciloglu, Y. Serinagaoglu, M. Kamasak, O. Erogul, and T. Akcam, "Investigation of sequential properties of snoring episodes for obstructive sleep apnoea identification," *Physiol. Meas.*, vol. 29, pp. 879–898, 2008.

[113] R. Jane, J. Sola-Soler, J. Fiz, and J. Morera, "Automatic detection of snoring signals: validation with simple snorers and OSAS patients," in *IEEE-EMBS*, Chicago, IL, 2000, pp. 3129–3131.

[114] W. Duckitt, S. Tuomi, and T. Niesler, "Automatic detection, segmentation and assessment of snoring from ambient acoustic data," *Physiol. Meas.*, vol. 27, pp. 1047–56, 2006.

[115] M. Cavusoglu, M. Kamasak, O. Erogul, T. Ciloglu, Y. Serinagaoglu, and T. Akcam, "An efficient method for snore/nonsnore classification of sleep sounds," *Physiol. Meas.*, vol. 28, pp. 841–53, 2007.

[116] J. Sola-Soler, R. Jane, J. Fiz, and J. Morera, "Snoring sound intensity study with ambient and tracheal microphones," in *IEEE-EMBS*, Turkey, 2001, pp. 2032–2035.

[117] F. Dalmasso, G. Benedetto, R. Pogolotti, G. Righini, and R. Spagnolo, "Digital processing of snoring sounds," *Eur Respir J Suppl.*, vol. 11, pp. 528s–532s, 1990.

[118] A. Yadollahi and Z. Moussavi, "Acoustic obstructive sleep apnea detection," in *IEEE-EMBS*, Minneapolis, MN, 2009.

[119] X. Huang, A. Acero, and H. Hon, *Spoken Language Processing: A Guide to Theory, Algorithm and System Development.* Prentice Hall PTR, 2001.

[120] A. Yadollahi and Z. Moussavi, "Formant analysis of breath and snore sounds," in *IEEE-EMBS*, Minneapolis, MN, 2009, pp. 2563–2566.

[121] A. Jain and D. Zongker, "Feature selection: Evaluations, application, and small sample performance," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 19, No. 2, pp. 153–158, 1997.

[122] A. Jain, R. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22(1), pp. 4–37, 2000.

[123] H. Pasterkamp, J. Schfer, and G. Wodicka, "Posture–dependent change of tracheal sounds at standardized flows in patients with obstructive sleep apnea." *Chest*, vol. 110, No. 6, pp. 1493–8, 1996.

Appendix

Appendix A

Flow–sound relationship

This chapter was published as a journal paper [1].

In clinical respiratory and/or swallowing assessment, flow is usually measured by spirometry devices, such as pneumotachograph, nasal cannulas connected to a pressure transducer, and heated thermistor or anemometry. Airflow is also measured by indirect means, i.e., detection of chest and/or abdominal movements using RIP, strain gauges, or magnetometers. The most reliable measurement of airflow is achieved by a mouth piece or facemask connected to a pneumotachograph [2], but it changes the breathing pattern [3-6]. Furthermore, this device cannot be used during the swallowing assessment. Therefore, flow is usually measured by nasal cannulas connected to a pressure transducer. Potentially, this method could be an inaccurate measure of airflow because the air leaks around the nasal cannulas. In addition, if the subject breathes through the mouth, flow is not registered at all.

For these reasons, the combined use of nasal cannulas connected to a pressure transducer and the measurement of respiratory inductance plethomogoraphy to monitor volume changes have been recommended as the best approach in recording flow to assess respiratory patterns [2]. However, these techniques have some disadvantages, especially when studying young children or patients with neurological impairments. Although the application of nasal cannulas may seem a minor intrusion, it can produce agitation in children and patients with neurological impairment. In addition, applying the RIP devices is difficult in children with neurological impairment as their poor postural control and physical deformities can make it challenging to ensure stable positioning [7].

Due to these difficulties, many investigators have developed acoustical techniques to estimate flow using different features of tracheal and lung sounds [9, 10, 26-28, 30, 31]. In one of the early studies, it was found that the mean power frequency of tracheal sound increased linearly with an increase in flow but remained about the same when the flow rate was above $0.75L/s$ [30], which was also in agreement with results at higher flow rates ($1.6 - 2.1L/s$) tested later by [31]. Tracheal sounds mean amplitude, average power, mean frequency and the multiplication of tracheal sounds mean frequency and mean amplitude were investigated in [10]. It was shown that among different features proposed in [10], the average error of estimating flow from tracheal sounds average power (14 ± 6.6) was less than those of the other features [10]. On the other hand, *Gavriely* and his co-workers [9] studied the average power of tracheal and lung sounds for a power

relationship with airflow in the upper 15% of target flow region. However, they did not attempt to estimate airflow from the respiratory sounds or report the estimation errors. *Que* and his co-workers [8] have estimated flow from tracheal sounds envelope. In this method they bandpass filtered the tracheal sounds in the range of $200 - 1000Hz$ followed by applying Hilbert transform to achieve the tracheal sounds' envelope. They used the estimated flow to measure ventilation and didn't report the errors in estimating flow from tracheal sounds.

Comparing different models for estimating flow from tracheal sounds average power; it has been shown that the exponential model is superior to the polynomial and the power models [26, 28]. The average error of estimating flow from tracheal sounds average power using exponential model was found to be 9.0 ± 3.0 and 10.0 ± 4.0 during inspiration and expiration phases, respectively [28]. In these studies [27, 26, 28] the exponential model between flow and average power of tracheal sound was found to be superior to other models. Figure A.1 presents the flow estimation procedure of the methods [27, 26, 28] using the exponential model.

In another study, the tracheal sound envelope was investigated for flow estimation [8]. The tracheal sound was band-pass filtered in the range of $200-1000Hz$ and then a Hilbert transform was applied to the filtered signal. The transformed signal was used to calculate the tracheal sound envelope and to estimate the flow from the calculated envelope by a linear model. The estimated flow was then used to measure ventilation, but the flow

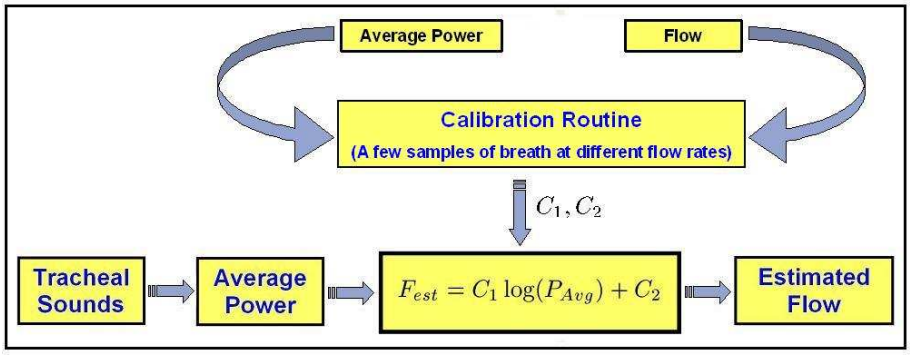


Figure A.1: *Flowchart of the previous studies for estimating flow from tracheal sound average power.*

estimation error was not reported. The flow rate in that study was constant at tidal flow and half of the recorded flow signal was used to calibrate the model [8].

All of the above mentioned methods [10, 27, 26, 28, 8] assumed that at least some samples of breath sound with known flow at each flow rate were available to derive the model coefficients for flow estimation. Capturing respiratory sounds at different flow rates for calibration may not always be possible prior to assessment especially when assessing young children, patients with neurological impairments and/or patients in emergency conditions.

In our previous study [29] a new method of flow estimation was proposed that reduced the need for calibration significantly. The method used a modified linear model describing flow and the entropy of tracheal sound relationship for flow estimation at variable rates. The coefficients of the model were derived from only one breath sound sample with known flow at medium flow rate (Figure A.2). The results showed that the entropy based model

was able to follow the flow variation with a low error of about 9% [29]. The main advantage of the entropy based model over all the previous studies is that it is robust in terms of flow variation, and that it does not need more than one breath with known flow at tidal or medium flow rate for calibration. Furthermore, with the aid of a new and simple technique [29] to cancel the effect of heart sounds on the tracheal sound, the method was also able to estimate flow at very shallow breathing, which had not been done in previous studies. These have been major improvements over the previous attempts in flow estimation.

This section investigates the robustness of the entropy based model in terms of its sensitivity to the changes of the tracheal sound signal's amplitude is presented here. In addition other parameters are investigated to see if they can reflect the flow variation as well as entropy.

A.1 Method

A.1.1 Data

The subjects for this part of study were in two age groups: 5 adults (all female) 29±8 years old and 5 children (3 female) 9.6 ± 1.7 years old [1]. Respiratory sounds were recorded using Siemens accelerometers (EMT25C) placed over suprasternal notch and the upper right lobe lung. Respiratory flow was measured by a pneumotachograph (Fleisch No.3) connected to a differential pressure transducer (Validyne, Northridge, CA). Subjects were

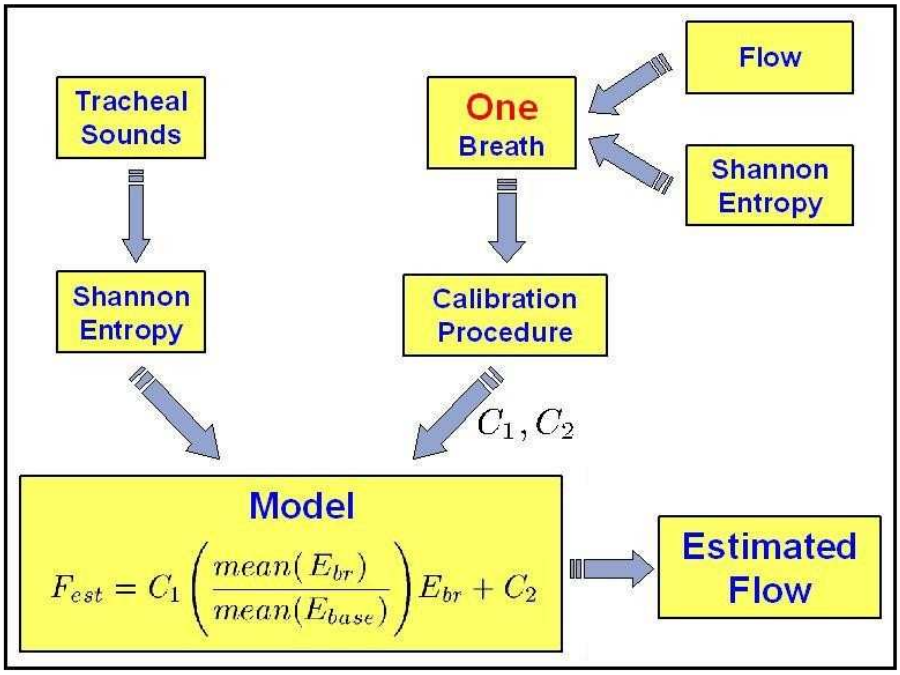


Figure A.2: *Flowchart of the flow estimation method using tracheal sound entropy.*

instructed to breathe at 5 different flow rates with 5 breaths at each target flow followed by a 10s of breath hold at the end of experiment. In this study the shallow ($< 6ml/s/kg$), low ($6 - 9ml/s/kg$), medium ($12 - 18ml/s/kg$), high ($18 - 27ml/s/kg$) and very high ($> 27ml/s/kg$) target flow rates were investigated [1]. Tracheal sound signals were used for flow estimation while the lung sound signal in correspondence with tracheal sound signal were used for respiratory phase detection using the method introduced in [7].

In [1] it was found that the best performance for estimating flow from tracheal sound entropy was achieved in the frequency range of $[75 - 600]Hz$. Tracheal sound signals were used for flow estimation while the lung sound signal in correspondence with tracheal sound signal were used for respiratory phase detection using the method introduced i [7].

In [1] it was found that the best performance for estimating flow from tracheal sound entropy was achieved in the frequency range of $[75-600]Hz$. This is in accordance with the fact that the main energy components of tracheal sound exists in the frequency range below $600-800Hz$ [72]. Thus, tracheal sound was band-pass filtered in this range followed by segmenting the band-pass filtered signal into segments of $50ms$ (512 samples) with 75% overlap between the successive segments.

A.1.2 Robustness of the Entropy Based Flow Estimation Method to Amplitude Changes

Investigation of the relationship between flow and respiratory sounds has shown that increasing flow increases the amplitude of respiratory sounds [72]. In previous studies these changes were modeled using the average power [27, 26, 28], envelope [8] and entropy [29] of the respiratory sounds. Since the method using entropy was superior to other methods and it was the most robust to flow variations, this method was chosen to further investigate the effect of flow changes on the tracheal sound's entropy.

In [29], the entire recorded signal was normalized in the range of $[-11]$ and sequestered into overlapping windows with 75% overlap between adjacent windows. Then, in each window the probability density function (pdf) of the signal was estimated. The pdf was calculated in n bins having a constant size. The bin size was chosen to be α and kept at this value for each window. Therefore, if in each window the signal $(x(t))$ lies in the range

of $[ab]$, then the number of bins is $n = (b-a)/\alpha$. Since α is constant for all windows, n depends on the range of the signal in each window. The Shannon entropy [99] in each window is calculated as:

$$E_p = -\sum_{i=1}^{n} p_i \log p_i, \tag{A.1}$$

where $\sum_{i=1}^{n} p_i = 1$.

In order to see the effect of changes in signal range on the entropy calculation, assume that the signal $x(t)$ is multiplied by a constant factor of C_f. Entropy of a signal is based on its pdf, which for the signal $x(t)$ it is defined on n points. For simplifying the evaluation of the effects of change in the signal amplitude on its pdf assume that $C_f = m/n$, where m can be greater or less than n. Hence, the pdf (q_i) of the new signal, $\acute{x} = m/n \times x(t)$, is defined on m points and its Shannon entropy can be written as:

$$E_q = -\sum_{k=1}^{m} q_k \log q_k, \tag{A.2}$$

where $\sum_{i=1}^{m} q_i = 1$. Considering the relationship between $x(t)$ and $\acute{x}(t)$ and the constant bin-size (α) for calculating the pdf of $x(t)$ and $\acute{x}(t)$, it can be concluded that for every p_i there exist $\lceil m/n \rceil$ components in q_i satisfying:

$$p_i = \sum_{k=\lceil m/n \rceil (i-1)+1}^{\lceil m/n \rceil i} q_k, \tag{A.3}$$

where $\lceil m/n \rceil$ is the smallest integer value greater than or equal to (m/n). For simplicity

it is assumed that the probability of all components in the q_k corresponding to each p_i is the same and equals to:

$$q_k = \frac{n}{m} p_i, \tag{A.4}$$

where $i = \left\lceil \frac{nk}{m} \right\rceil$. Thus, Eq. A.2 can be written as:

$$\begin{aligned} E_q &= -\sum_{k=1}^{m} q_k \log q_k = -\sum_{k=1}^{m} \frac{n}{m} p_{\lceil \frac{nk}{m} \rceil} \log \frac{n}{m} p_{\lceil \frac{nk}{m} \rceil} \\ &= -\frac{n}{m} \log \frac{n}{m} \sum_{k=1}^{m} p_{\lceil \frac{nk}{m} \rceil} - \frac{n}{m} \sum_{k=1}^{m} p_{\lceil \frac{nk}{m} \rceil} \log p_{\lceil \frac{nk}{m} \rceil} \end{aligned} \tag{A.5}$$

Replacing $\left\lceil \frac{nk}{m} \right\rceil$ with j, it is clear that:

$$\left\lceil \frac{nk}{m} \right\rceil = j \rightarrow k = \frac{m}{n}(j-1) + 1, \cdots, \frac{m}{n} j. \tag{A.6}$$

Thus, each j corresponds to m/n components in p_i. Therefore, Eq. A.5 can be rewritten as [1]:

$$\begin{aligned} E_q &= -\frac{n}{m} \log \frac{n}{m} \sum_{i=1}^{n} \frac{m}{n} p_i - \frac{n}{m} \sum_{i=1}^{n} \frac{m}{n} p_i \log p_i \\ &= \log \frac{n}{m} + E_p \Rightarrow E_q + \log m = E_p + \log n \end{aligned} \tag{A.7}$$

From the above equation A.7 it can be concluded that multiplication of the signal, $x(t)$, by a factor (m/n) causes a term $(\log n)$ to be added to its entropy. Depending on how this factor is defined, this extra term in entropy represents an increase (or decrease) of the

original signal amplitude that is supposedly changing in correspondence to flow. Recall that n represents the signal range in each window and therefore $\log n$ is the logarithm of the signal range. Thus, instead of entropy with its high computational cost, one may use the range factor, $L = \log n$ of the signal for flow estimation and yet obtain a robust flow estimation. Three variations of the range factor to be used as parameters to reflect the changes in tracheal sound amplitude with respect to flow are described next. The parameters are used for flow estimation and their performance is compared with the entropy based method.

A.1.3 Relationship between Flow and Tracheal sound Range

Recall that n and m of Eq. A.7 are:

$$\begin{aligned} n &= \frac{\max x - \min x}{\text{binsize}}, \\ m &= \frac{\max \acute{x} - \min \acute{x}}{\text{binsize}}, \end{aligned} \tag{A.8}$$

and $L = \log n$ for the signal $x(t)$ or $L = \log m$ for the signal $\acute{x}(t)$. Since the binsize $= \alpha$ is a constant, then L represents, respectively the range of the signals x and $\acute{x}$ around their average value [1].

The previous study [29] for flow estimation, showed that the modified linear model using entropy was superior to other ones. Therefore this model was chosen to investigate

the relationship between flow and the tracheal sound range parameter:

$$F_{est} = C_1 \frac{meanL_{ph}}{meanL_{base}} L_{ph} + C_2, \tag{A.9}$$

where C_1 and C_2 are the model coefficients derived from the one breath with known flow, $L_{ph} = [L_1, \cdots, L_w]$ is a vector representing the range of the signal in each respiratory phase (inspiration or expiration), w is the number of segments in each respiratory phase and L_i is the range of the upper 40% values of the tracheal sound (x) in each segment ($L_i = \log[\max x - \min x]$). Similarly, L_{base} is the same vector that is calculated in the base respiratory phase. Base respiratory phase is the one breath that is assumed to be available with known flow to calibrate the model.

Since the maximum and minimum values might be sensitive to noise, the average of the upper and lower r% of the signal in each segment was also considered as another parameter to represent the signal range and to estimate the flow from tracheal sound. In this case, L_r was defined as:

$$L_r = \log[mean(x|x > [\max x \times (1 - r/100)]) - mean(x|x < [\max x \times (1 - r/100)])], \tag{A.10}$$

where x is the tracheal sound signal in each segment and $mean(.)$ represents the average value. Standard deviation can be another parameter to represent the range of the signal around its average value. Therefore, the logarithm of the standard deviation (L_{std}) of the

signal in each segment is also investigated as another parameter for flow estimation.

In previous studies the logarithm of tracheal sounds average power was used for flow estimation [26, 28]. In order to compare the performance of the above proposed parameters with that of previous studies, the logarithm of the average power (L_{Avg}) of the signal in each segment was also calculated. Finally the modified linear model (Eq. A.9) was used to estimate flow from L, L_r, L_{std} and L_{Avg}. Then the estimated signal was compared with the results of the flow estimation method based on tracheal sound's entropy [1].

A.1.4 Measure of Comparison between the Parameters

In order to evaluate the performance of different parameters, the error of flow estimation using any of the proposed new parameters were compared with each other and with that of the previous study in [1]. To be consistent with previous studies the error was defined as:

$$error = \frac{|mean(F) - mean(F_{ph})|}{mean(F)} \times 100, \tag{A.11}$$

where F and F_{ph} are the upper 40% values of the actual and the estimated flow in each respiratory phase, respectively. The error was then averaged within and between the subjects.

A.1.5 Heart sounds reduction

In studying respiratory sounds heart beat is an unavoidable source of interference that changes both frequency and time characteristics of the respiratory sounds. Most of the

heart sounds energy is concentrated in the frequency range of $20 - 200Hz$ [72], which overlaps with the low frequency components of respiratory sounds. In most of the acoustical flow estimation methods [27, 26, 28], the tracheal sound was analyzed for frequency range above $300Hz$, where it is almost free of the heart sounds effect. However, we found that for very shallow breathing where the tracheal sound has very low intensity, it is important to consider the frequency range below $300Hz$ for flow estimation. Because heart sounds interferes with the respiratory sounds in the frequency range below 300 Hz and also that our pilot studies showed a high inaccuracy in very low flow estimation due to heart sounds, therefore we investigated several methods to cancel the effect of heart sounds prior to flow estimation.

As the effect of heart sounds on lung sounds is much more pronounced than that on tracheal sound, all of the heart sounds cancellation techniques have been applied to lung sounds. Several methods based on adaptive filtering [73-76], wavelet denoising [77, 78], adaptive thresholding and 2-D interpolation of lung sounds in the time-frequency domain [79] and removing heart sounds-included segments from the wavelet coefficients of lung sounds and then reconstructing the signal by auto regressive or moving average models [80] have been proposed for heart sounds reduction from lung sounds.

We employed some of the above mentioned techniques but the results were not satisfactory for flow estimation purpose. This is probably due to the fact that by applying most of these techniques for heart sounds cancellation, some of the heart sounds still re-

main in the respiratory sound signal. On the other hand, some other techniques alter the respiratory sound signal. Although these changes are not much noticeable by auditory means, they increase the error of flow estimation at shallow breathing. Therefore, instead of heart sounds cancellation from the tracheal sound, we used a technique for canceling the effect of heart sounds on the calculated features of the tracheal sound.

The first step in all of the methods for canceling heart sounds from respiratory sounds except the wavelet de-noising method is to localize the segments including heart sounds. In order to remove the effects of heart sounds on the calculated features, the segments including heart sounds were localized and the features were calculated for the segments void of heart sounds. Then, spline interpolation was applied to estimate the values of the feature in the segments including heart sounds. This technique effectively cancels the effect of heart sounds on the estimated features of the tracheal sound [1].

A.2 Results

Figure A.3 shows an example of the actual flow along with the corresponding values of the parameter L (the difference between maximum and minimum values of the signal) calculated for a typical subject. Even though tracheal sound was band-pass filtered in the range of $[75 - 600]Hz$, heart sounds effects were still evident in the time-domain signal (Fig. A.3–a). Their effects were later removed from L using the heart sound cancellation method introduced in section A.1.5 (Fig. A.3–c).

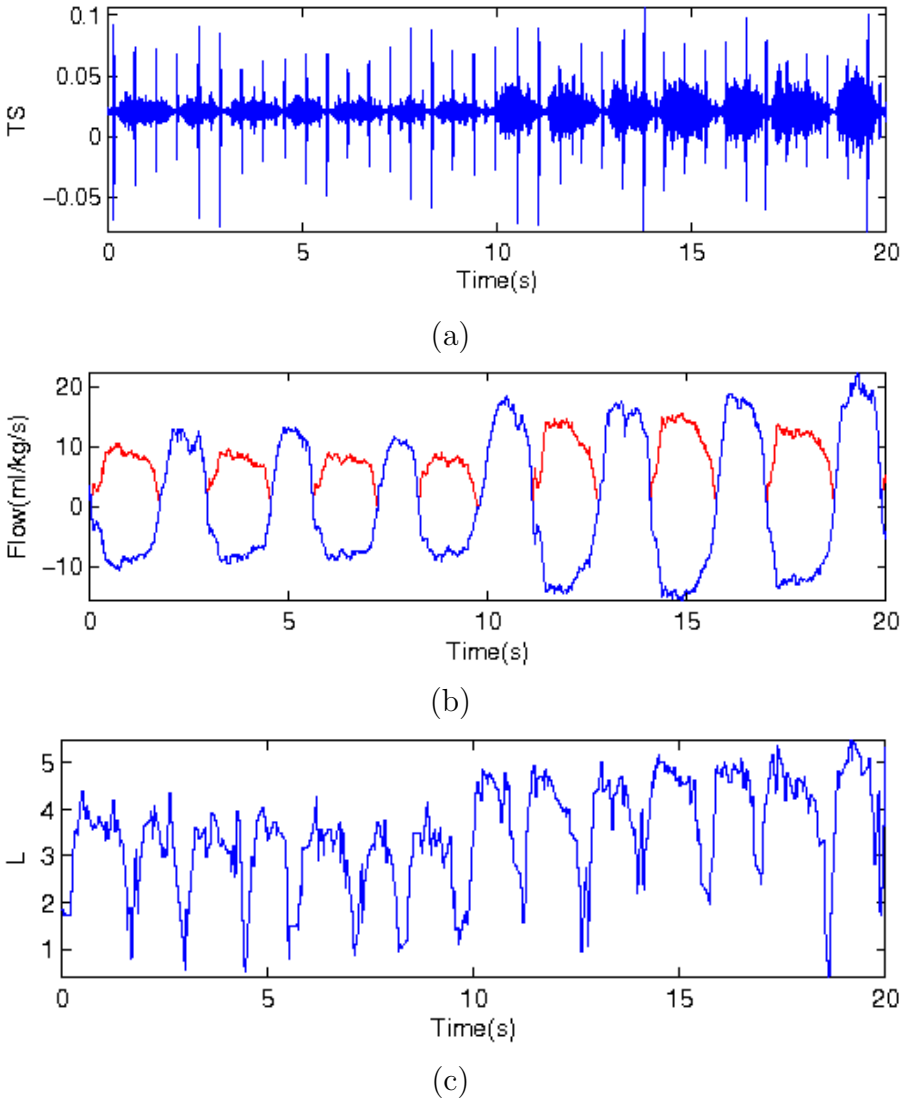


Figure A.3: *A typical example of a) the band-pass filtered tracheal sound, along with b) its correspondent actual flow (blue curve) and its absolute value (red curve) and c) the corresponding values of L.*

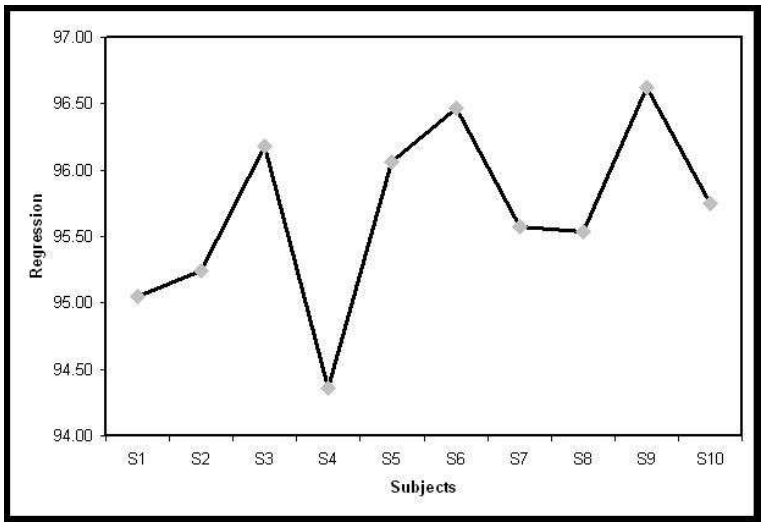


Figure A.4: *Regression coefficients between E and L for different subjects.*

In order to consider the relationship between E and L and their similarity, the regression coefficient between E and L was calculated and the results are presented in Fig. A.4. It can be observed in Fig. A.4, the regression coefficient is more than 94% for all the subjects, implying a high degree of linear similarity between E and L. Since L might be sensitive to noise, L_r was proposed. Figure A.5 presents the overall error of estimating flow from modified linear model using L_r for different values of r.

Finally the flow estimation error using any of the L_{Avg}, L, L_{std} and E parameters of the tracheal sound at different flow rates is shown in Fig. A.6. Figure A.7 shows the absolute value of the actual flow acquired from a typical subject at different flow rates along with its corresponding parameters L_{std}, L, L_{Avg} and E; it clearly depicts the incapability of L_{Avg} at shallow flow rates.

A typical example of the actual and estimated flow using the parameter L for a typical

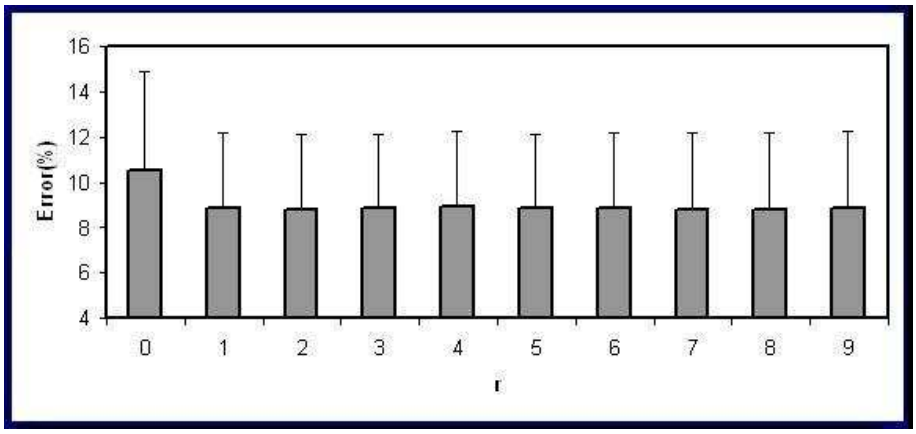
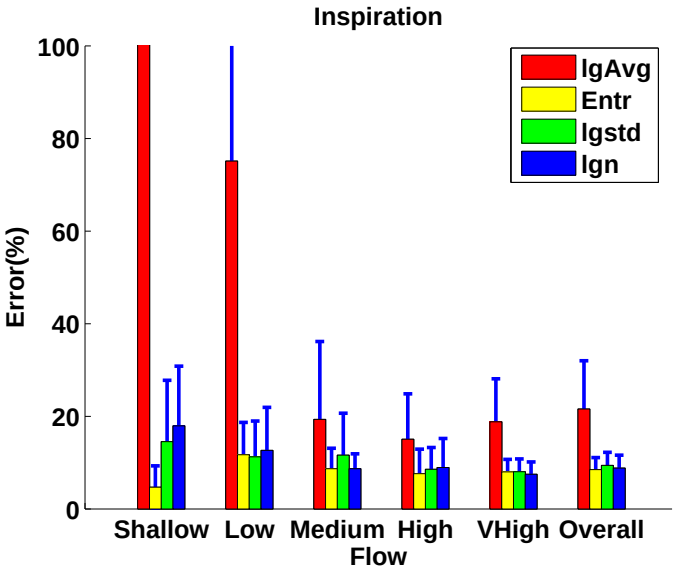


Figure A.5: *The effect of parameter r on the overall error averaged between the subjects when estimating flow from L_r.*

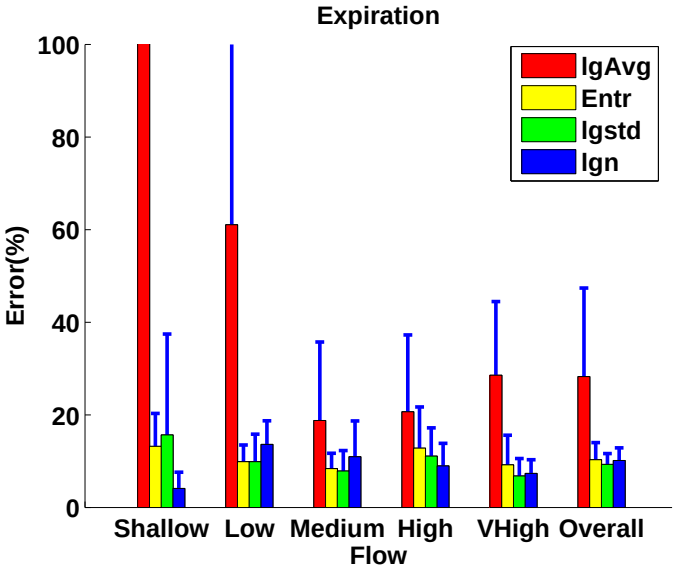
subject is shown in Fig. A.8. The zoom–in plots of the results at different flow rates are also presented for more thorough comparison of the actual and estimated flow rates.

A.3 Discussion

Our previous study in this field using entropy of tracheal sounds signal as the main characteristic feature was found to perform superior to other flow estimation methods. In search of finding a physiological representation of entropy and why it worked well for flow estimation, in this study it was shown analytically that when calculating entropy using a constant bin–size, there exists a term (L) that represents the range of the signal around its average. L was found to be capable of following the changes in the absolute values of flow (Fig. A.3); the pattern which was previously found between flow and entropy of tracheal sounds (E).



(a)



(b)

Figure A.6: *The overall error of different parameters for estimating flow from tracheal sound during a) inspiration and b) expiration.*

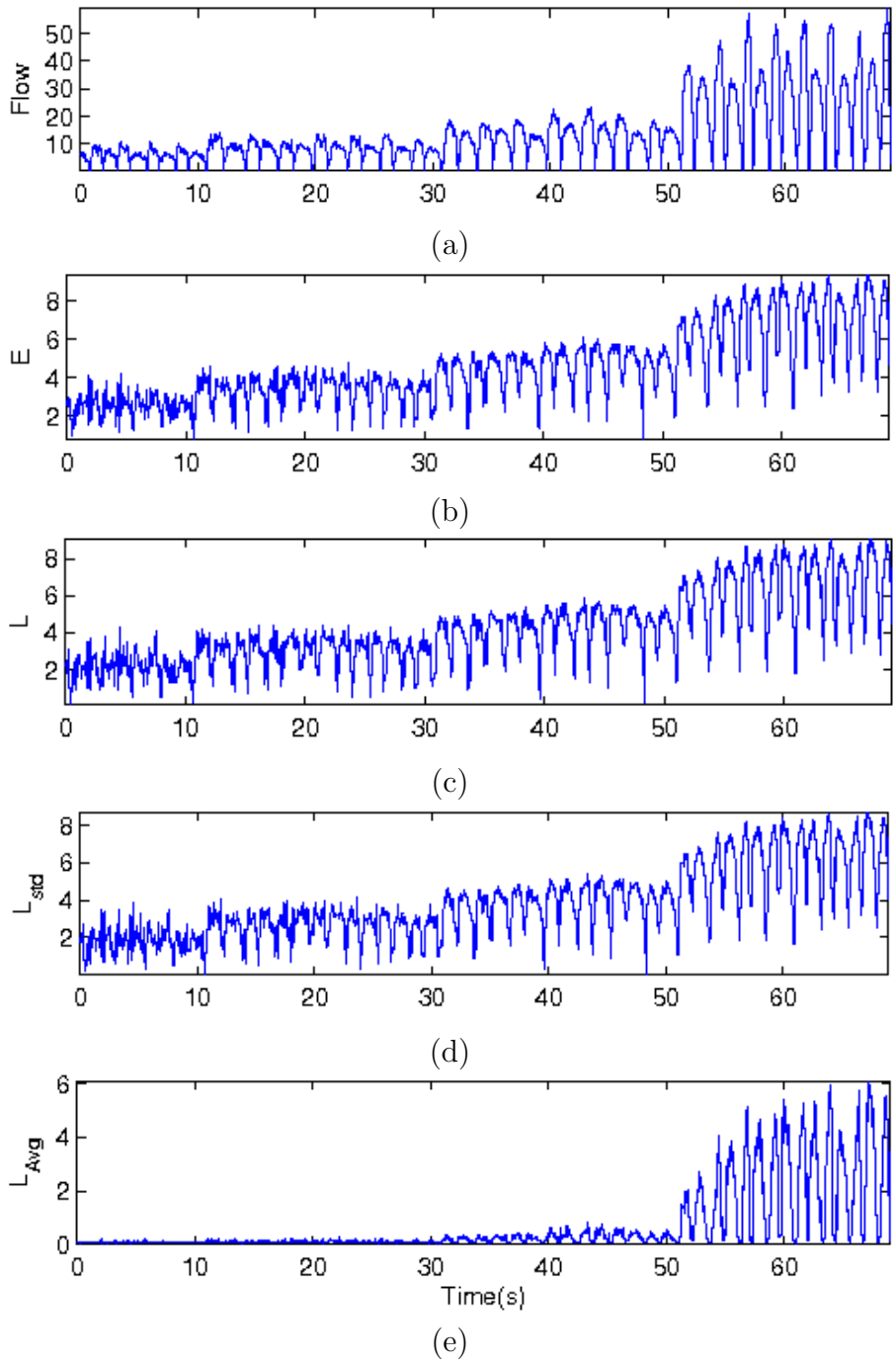


Figure A.7: *a) Absolute value of a normal subject's actual flow along with its corresponding parameters b) E, c) L, d) L_{std} and e) L_{Avg}.*

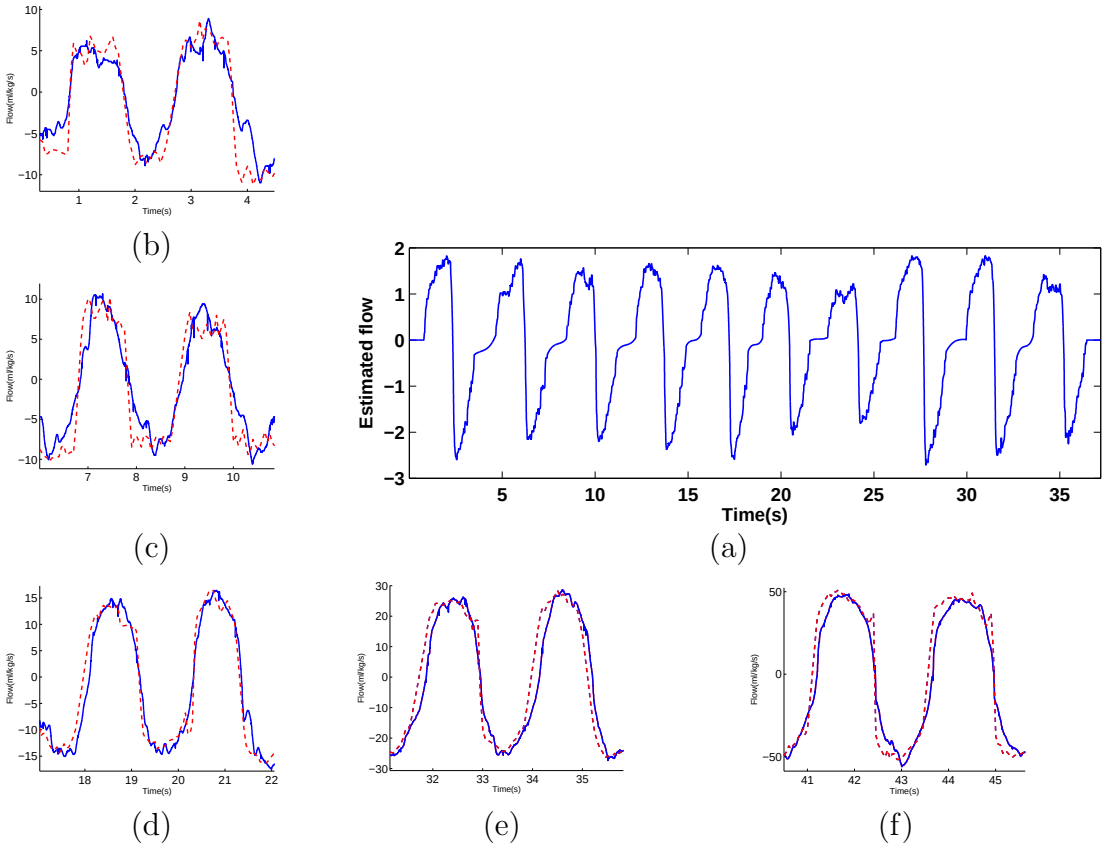


Figure A.8: *A typical example of the a) actual flow (solid blue line) and the estimated flow (dotted red line) along with the focus on the results during b) shallow, c) low, d) medium, e) high, and f) very high breathing.*

In order to consider the relationship between E and L and their similarity, the regression coefficient between E and L was calculated (Fig. A.4). As can be observed in Fig. A.4, the regression coefficient is more than 94% for all the subjects, implying a high degree of linear similarity between E and L. Thus, it can be concluded that when estimating flow from tracheal sound entropy as in [1], the main component contributing in following flow variation is in fact L which is embedded within the entropy but can be determined without entropy calculation. Hence, it provides another robust feature which involves less computation.

L represents the range of the signal around its average value in each segment and is calculated by finding the difference between maximum and minimum values of the signal. As these values might be sensitive to noise, L_r was proposed, which represents the difference between the average values of the upper and lower $r\%$ of the signal in each segment. This parameter was used for estimating flow from tracheal sound and its performance for different values of r was evaluated. For this purpose the overall error of estimating flow from modified linear model using L_r was calculated for different values of r (Fig. A.5) and the error was maximum when $r = 0$. This result was expected as for $r = 0$, L_r becomes equal to L and it is more sensitive to noise. On the other hand for values of $r > 0$ the changes in the overall error were negligible; this implies the model is not very sensitive to the changes of r. This observation might also be due to the effects of band-pass filtering the tracheal sound in the range of $[75 - 600]Hz$, which removes high

frequency noises.

Finally the flow estimation error using any of the L_{Avg}, L, L_{std} and E parameters of the tracheal sound was calculated (Fig. A.6). As can be observed, the performance of L, L_{std} and E are similar at all flow rates in both respiratory phases, while the error of flow estimation using L_{Avg} was consistently higher than those of the other parameters especially at the shallow and low flow rates. This result was expected since L_{Avg} cannot follow the flow variation as good as the other parameters. In fact, at shallow flow rates, it completely fails. This result is in agreement with the results reported in [1]. When the absolute value of the actual flow acquired from a typical subject at different flow rates was shown along with its corresponding parameters L_{std}, L, L_{Avg} and E (Fig. A.7); it clearly depicted the incapability of L_{Avg} at shallow flow rates.

Entropy of a signal represents the changes in its pdf. Therefore, accurate estimation of pdf is of great importance in calculating entropy. When calculating entropy of tracheal sound, the pdf of the signal in each segment is estimated using a kernel estimator, which is the most time consuming part of the flow estimation process [1]. Although it was shown that the method proposed for estimating flow from tracheal sound entropy can be performed in real–time, its computational cost was more than that of other methods based on average power of tracheal sound [1]. On the other hand the computational cost of estimating flow using either L or L_{std} less than that of the methods using L_{Avg}, and is much less than that of the method using entropy.

The results of using the modified linear model with the L_{std} and E parameters are very similar to that when using the parameter L. For acoustical detection of apnea and hypopnea events, this parameter (L_{std}) was chosen as the superior parameter of the tracheal sound to follow flow variation and estimate flow acoustically due to its better performance and also less computational cost compared to the other parameters.

A.4 Conclusions

The goal of this study was to investigate the relationship between flow and tracheal–sound entropy more thoroughly. The objective was to extract an embedded feature in the entropy of tracheal sound that reflects the flow variation most clearly and to investigate its robustness for flow estimation. It was shown that when calculating entropy using a constant bin–size, there exists a term (L) that represents the range of the signal around its average. As the results suggest, L follows the variation of the absolute value of flow from very shallow breathing to very high flow rate quite well. The model using this parameter resulted in flow estimation with an average error of $8.82 \pm 2.79\%$ and $10.15 \pm 2.73\%$ during inspiration and expiration respectively. Thus, this feature can be used for estimating flow with essentially the same performance as entropy but with a much lower computational cost.

Appendix B

Automatic Breath and Snore Sounds Classification

This chapter was published as a journal paper [98].

Tracheal respiratory sounds are generated by the turbulence of air in the upper airways, and their characteristics change as the flow rate, physical and mechanical properties of the upper airways change. Recording tracheal sounds during sleep may contain snore sounds. Snoring is a common condition that can affect anyone specially people above age 50 [36]. Occasional snoring that occurs usually when people are very tired or their neck being in a bad position is very common but not serious. However, serious and regular snoring is a condition that may be associated with sleep apnea [100]; hence requires a clinical study for the entire night during sleep.

When there is a narrowing or obstruction in the upper airways, the turbulence of air and the vibrations of soft palate and the upper airways walls generate snore sounds.

Upper airway narrowing, poor muscle tone in the throat and tongue, bulky throat tissue, long soft palate and/or uvula are the common causes of snoring. While it is difficult to diagnose the exact source of snoring as one may also have a combination of the causes [90, 101], the snore sound analysis has recently received more attention due to its diagnostic potentials for detecting sleep apnea [90, 101-112].

One of the most common applications of snore sound analysis is to document its association with obstructive sleep apnea (OSA) objectively. Hence, breath and snore sounds commonly recorded by a microphone placed over the patient's neck or hung in the air above the patient's head during the night. Due to the long duration of data collected in sleep studies ($6-8$ hours), there is a need for automatic and accurate identification of snore segments in a record of breath and snore sounds during the entire night.

Several methods for automatic classification of snore and breath sounds have been employed using different characteristics of the respiratory sounds, such as neural network classifier based on temporal and spectral features of sound signals [113], pitch of the sounds signals [110], Hidden Markov Model (HMM) based on Mel–frequency cepstral coefficients (MFCC) of sounds [114], spectral energy of the recorded signal in the frequency range of $[0-7500]Hz$ [115], and combination of zero crossings, signal's energy, normalized auto-correlation coefficient at $1ms$ delay and the first predictor coefficient of linear predictive coding (LPC) estimation [92]. The accuracies of these methods have been reported between 82% to 96.7% [113-115, 92]. However, in most of the above mentioned studies, the

microphone was hung in the air above the patient's head [110, 114, 115, 92]. Therefore, the recorded signal mostly included snore and ambient noises while the breathing sounds were less detectable unless the patient breathes loudly. In [113] the sound signals were recorded with a microphone placed over the trachea, and sensitivity of the detected snore sound segments was reported as 82% with the positive predictive value of 90%.

The sound signals which are recorded with microphones hung in the air or attached to the patient's neck are different due to the different distances and medias to the sound generating sources. The relationship between snore sounds from ambient and tracheal recording were investigated in [102, 116], however, the effects of the microphone position on the respiratory breath sound signals were not investigated previously. The other parameter that affects the spectral characteristics of tracheal breath sound signals is the neck position; this was not investigated in the previous studies when classifying breath and snore sound signals.

In this study, we have recorded sound signals from two microphones simultaneously; one placed over trachea (tracheal microphone) and one placed over the forehead of the patient (ambient microphone). An automatic method was developed to classify breath and snore sound segments based on their energy, zero crossing rate and formants of the sound signals. The effects of body and neck position on the classification's performance was also investigated in this study. Moreover, the classification results were compared for the sounds recorded with ambient and tracheal microphone to find the best position of

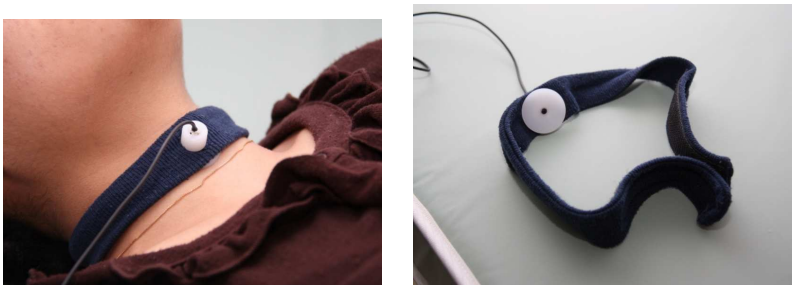
the microphone for recording sound signals during night.

B.1 Method

B.1.1 Data

Data were recorded from 23 (7 females) patients (51.3 ± 12.5 years old) suspect of OSA, who were referred to full-night sleep study at the Health Sciences Center Sleep Disorders Clinic (Winnipeg, Canada). This study was approved by the Ethics board of University of Manitoba prior to the study and all the patients signed the consent form before starting the experiment. The patients' demographic information is shown in Table B.1. Among the patients, 10 were simple snorers (AHI: 5.6 ± 5.2) and 13 were OSA patients with different levels of severity (AHI: 43.6 ± 40.7).

The respiratory tracheal sounds (including snore sounds) were recorded by two Sony (ECM-77B) microphones: one placed over the suprasternal notch of the patient's neck embedded in a chamber (diameter of $6mm$) wrapped around the neck with a soft neck band (Fig. B.1), and the second microphone was hung in the air about $20 - 30$ cm from the patient's head. The sound signals were recorded simultaneously with PSG data for the entire night. The detailed analysis of their PSG data done by sleep lab technicians was used for extracting the patients' neck positions during the night.



(a) (b)

Figure B.1: *a) The arrangement of tracheal microphone over the neck and b) the microphone in the chamber and its neck band.*

Table B.1: *Patients' demographic information. BMI stands for body mass index.*

Parameter	Age ($\mu \pm \sigma$)	BMI ($\mu \pm \sigma$)	AHI ($\mu \pm \sigma$)
Average	51.3 ± 12.5	33.6 ± 4.4	27.1 ± 35.8
Range	$[25 - 87]$	$[28.4 - 47.9]$	$[0.9 - 125.7]$

Sound signals were amplified with a gain of 200 and band-pass filtered with the cutoff frequencies of $[0.5Hz - 5kHz]$ using Biopac (DA100C) amplifiers. The amplified signals were digitized at a sampling rate of 10240 Hz using NI9217 data acquisition module.

B.1.2 Signal Analysis

It has been shown that the energy of breath sounds void of snore sounds is focused below $800Hz$ [21], while the energy of snore sounds is up to $2000Hz$ [90]. On the other hand, snore sounds have shown to have important components in low frequencies of around $100Hz$ [117, 102, 103, 90, 101]. Therefore, the recorded sounds were bandpass filtered in the frequency range of $[70 - 2000]Hz$ to remove the effects of low–and high–frequency noises, while including the main frequency components of both breath and snore sounds.

The sound and silent segments were extracted by an automated method prior to breath and snore classification [118].

To validate the snore and breath classification method, a large number of breath and snore sound segments were first manually extracted from tracheal sounds by auditory and visual inspection of the signals in the time–frequency domain. Figure B.2 shows samples of the recorded tracheal sounds in time and time–frequency domains. Breath and snore sound segments were marked in both domains for investigating the signal's characteristics. The dark colors seen during the first and second inspirations in the time–frequency domain, represent snore sounds and its harmonic formants. These harmonics represent the resonance of the upper airways during snore sound generation. However, it should be noted that the formants are not constant among different people, and they even change from snore to snore for the same person during the night.

In order to investigate the effects of the patient's neck position, the sounds were extracted and labeled from different positions using the score sheet of PSG data. For each patient the available positions including supine (lying down, face up), prone (lying down, face down), lateral left or right were determined. Assuming symmetry between the lateral left and lateral right positions with respect to the upper airways as the source sound generation, the segments extracted from the left and right positions were merged and marked as lateral position. In total, 5909 breath and 3995 snore segments in different body postures were extracted from all patients. The details of the number of breath and

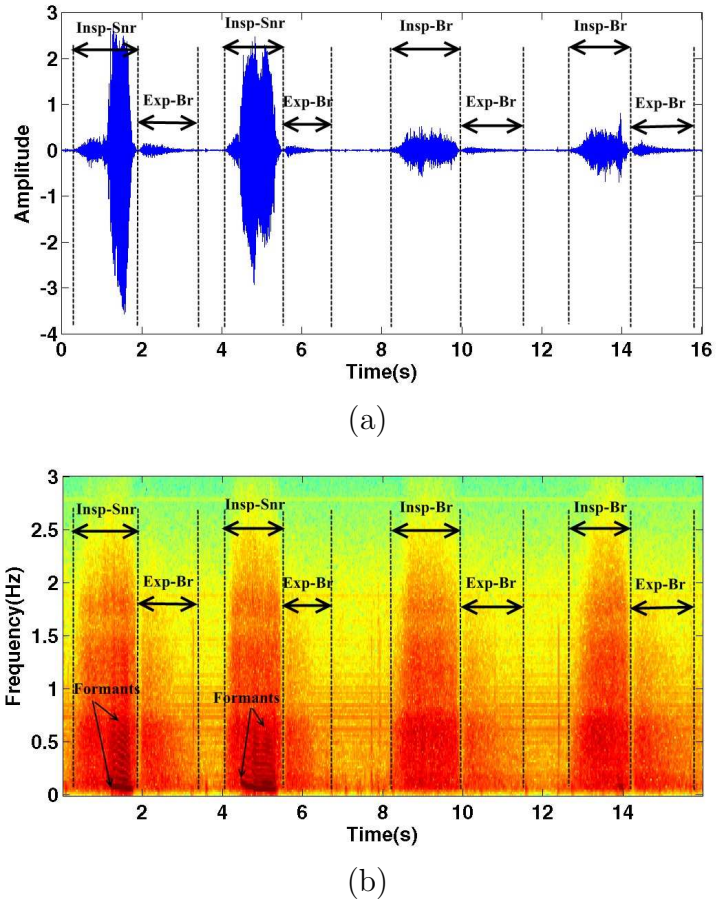


Figure B.2: *Samples of the recorded tracheal sound in a) time and b) time–frequency domains. The sound segments are extracted and marked manually. Insp-Snr, Insp-Br and Exp-Br represent inspiration segments including snore, inspiration and expiration breath segments void of snore, respectively. The dark repeating frequencies in the time–frequency representation of tracheal sounds (b) show the snore sounds' formant frequencies.*

Table B.2: *The details of the number of breath and snore segments of all subjects at different positions.*

Sound	Lateral	Supine	Prone	Total
Breath	2202	3115	592	5909
Snore	1650	2076	269	3995

snore segments at different positions are presented in Table B.2.

Feature Selection and Classification of breath and snore sounds

Three features, the sound's energy in dB, zero crossing rate (ZCR) and the first formant frequency ($F1$), were calculated from the sound segments. The number of zero crossings in each segment was calculated as:

$$ZCR = \frac{\Sigma_{k=1}^{N-1} |sign\,(x(k+1)) - sign\,(x(k))|}{2N}, \tag{B.1}$$

where N is the number of samples in each segment, $sign()$ shows the sign function and $|.|$ represents the absolute value. In each sound segment, the average of the sound signal was set to zero. Since, the number of zero crossings is proportional to the length of the signal, it was divided by N to be independent of the changes in the segment's length.

For every sound segment, linear predictive coding (LPC) was used to find the formant frequencies [119]. In every segment, sound signal was windowed with a Hamming window of $20ms$ with 50% overlap between adjacent windows and the signal in the window was estimated by an autoregressive (AR) model as described in [120]. Since, the first formant ($F1$) of the sound segments in the frequency range of below 400 Hz was found to be

significantly different between breath and snore sound segments [120], in every window of the sound segment $F1$ was estimated and their median value was calculated and considered as the $F1$ of the sound segment.

Fisher Linear Discriminant (FLD) [121], was used to transform the three features into a new 1D space. Principle component analysis (PCA) is another method, which is also commonly used for transforming features and extracting the best features [115]. Calculation of the PCA base functions is a blind process, in which the class information is not considered. On the other hand, in FLD method the transform vector is estimated by maximizing the class separability. Therefore, FLD based transformation is expected to achieve better results [122].

In FLD method, the observations $\mathbf{x}$ are transformed into a new space ($\mathbf{y} = \mathbf{w}^T\mathbf{x}$). In our case, $\mathbf{x}$ is a $3 \times n$ matrix of features extracted from the segments, and n is the number of segments. $\mathbf{w}$ and $\mathbf{y}$ are 3×1 and $1 \times n$ vectors representing the projection vector and the transformed features in the new dimension, respectively. $\mathbf{w}$ is estimated from the training data by maximizing the separability between classes after the projection ($\mathbf{y}$) [95].

To derive a classification threshold, we minimized the Bayesian error to estimate the optimum threshold. Assume that for a chosen threshold, the projected features that are smaller or larger than the threshold, are classified into classes ω_1 or ω_2. Then, the

Bayesian error [95], P_{err}, associated with the selected threshold, k, is defined as:

$$P_{err}(k) = \sum_{y_i >= k} p\left(y_i|\omega_1\right) P\left(\omega_1\right) + \sum_{y_i < k} p\left(y_i|\omega_2\right) P\left(\omega_2\right), \tag{B.2}$$

where $P\left(\omega_c\right),\ c = 1, 2$ shows the probability of each class. $p\left(y_i|\omega_c\right),\ c = 1, 2$ is the relative probability that y_i actually belongs to class $\omega_c,\ c = 1, 2$. Here, the probability functions are estimated by histogram functions. The optimum threshold value is determined by minimizing the error as:

$$Thr = \min_k \left\{P_{err}(k)\right\}. \tag{B.3}$$

Evaluation

The method's performance in classifying breath and snore sounds was investigated with and without considering patients' posture. In our dataset, 14 out of 23 patients had sleep data in both lateral and supine positions. The classifier was evaluated in three ways:

- Evaluation A:

  For each subject, data at supine position was used for training, and data at the lateral position was used as the test data set. Then, the classifier accuracy was averaged among the subjects.

- Evaluation B:

  Similar to Evaluation A, except that training and test data sets were swapped

Table B.3: *Number of breath and snore segments of* 14 *subjects at different positions.*

Sound	Lateral	Supine	Total
Breath	1795	2171	3966
Snore	1407	1633	3040

and chosen from lateral and supine positions, respectively. The number of breath and snore segments of these 14 patients at different body positions are shown in Table B.3.

- Evaluation C:

  Data of all patients (in all sleep positions) except one was considered as the training data set, and data of the left out patient was used as test data set. This procedure was repeated till all patients' data were used as the test set once. The results were then averaged among all the patients. In essence, Evaluations A and B in comparison to Evaluation C investigate the effect of body position on the classifier's performance.

In each evaluation scheme, the features of training and test data sets were normalized to be in the range of $[0, 1]$. This normalization was chosen because the three features used in this study have different characteristics; hence, their range of variations could be different without normalization, and that would affect the calculation of FLD projection vector and hamper the classification results. The above mentioned normalization avoids such errors.

For each evaluation, the classification was performed independently on the sounds recorded by either tracheal and ambient microphones, and its performance was compared in terms of sensitivity, specificity and accuracy:

$$Sensitivity = \frac{TP}{TP+FN} \times 100, \tag{B.4}$$

$$Specificity = \frac{TN}{TN+FP} \times 100, \tag{B.5}$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100, \tag{B.6}$$

where TP, TN, FP and FN are the number of true positive, true negative, false positive and false negative classified segments, respectively. Note that TP and TN refer to the number of correctly detected breath segments and the number of correctly classified snore segments, respectively.

B.2 Results

In this study, the optimum FLD threshold was estimated based on an approximation of the Bayesian error, for which the probability distributions of different classes were estimated using histogram function. To investigate the effects of the number of bins on the histogram and the corresponding estimated optimum thresholds, in Evaluation C the histogram and threshold values were estimated for different numbers of bins. Figure B.3 shows the average and standard deviation values of the threshold averaged among different

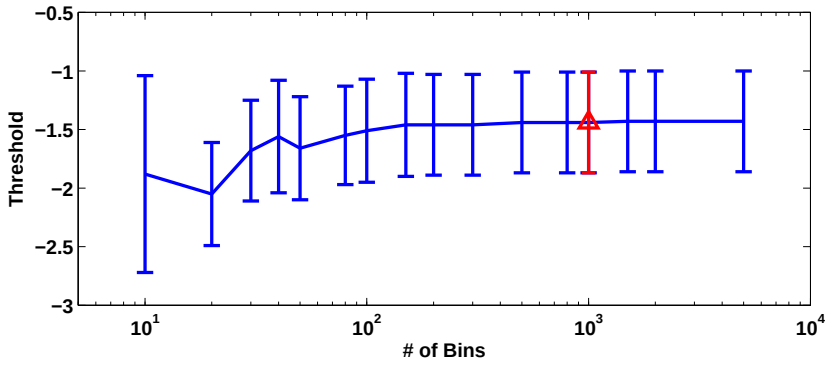


Figure B.3: *Mean and standard deviation values of the optimum threshold for different number of bins. The threshold values were calculated for Evaluation C and averaged among different subjects.*

subjects. The results show that for more than 1000 bins, the threshold reaches a plateau and further increments in the bin numbers do not change its value. Therefore, we used 1000 bins to estimate the histogram in all experiments (marked with red bar and triangle marker in Fig. B.3).

Figure B.4 shows the classification performances of the evaluations A, B and C for the sounds recorded by tracheal microphone. In Evaluation A for each patient with sleep data at two neck positions, the classifier was trained with the data of each patient in supine position and tested on their data in lateral position; while in Evaluation B the train and test data sets were swapped (Fig. B.4). As it can be seen in Fig. B.4 and Table B.4, the sensitivity and specificity results of the classifiers were similar by both ways of evaluations, and the change in the training and test data sets did not affect the performance of the classifier noticeably. The student *t–test* was performed on the classification results of the Evaluations A and B and the *p*–values for accuracy, sensitivity and specificity of

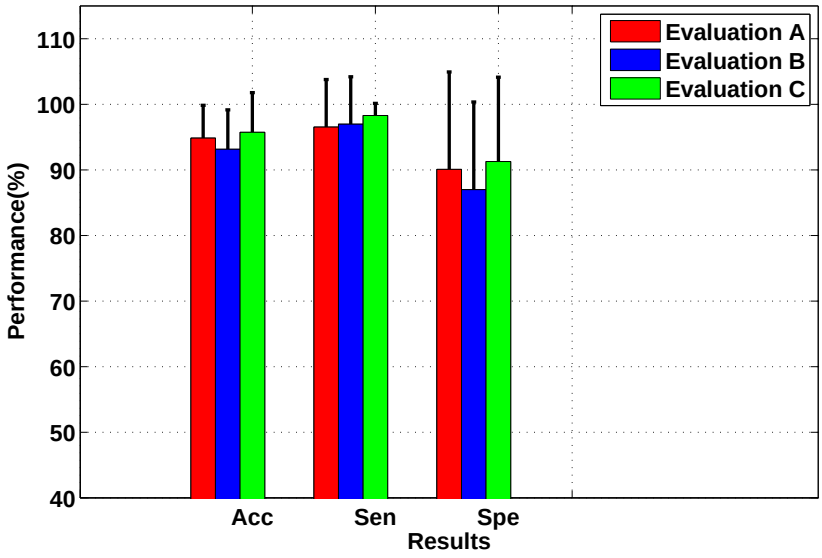


Figure B.4: *Classification results of Evaluations A, B and C for tracheal recording. Sen, Spe, and Acc represent sensitivity, specificity and accuracy values, respectively.*

the classifier were 0.37, 0.88 and 0.53, respectively; this shows they were not significantly different.

Since in most portable monitoring devices the patient's position is not known, in Evaluation C the training and test data sets were selected from different patients, blind to their neck positions. The accuracy, sensitivity and specificity results are shown in Fig. B.4 and Table B.4 for tracheal sound recording. The results show that the performance of the classifier in terms of accuracy, sensitivity and specificity values are similar to those of the Evaluations A and B which again confirms the stability of the method with respect to the changes in the neck position.

In this study sound signals were recorded simultaneously with two microphones: ambient and tracheal microphones. Figure B.5 shows samples of the ambient sounds in the time

Table B.4: *The classification results of Evaluations A, B and C for tracheal sound recordings. Sen, Spe, and Acc represent sensitivity, specificity and accuracy values, respectively.*

Evaluation	Accuracy	Sensitivity	Specificity
A	$94.8\% \pm 5.0\%$	$96.5\% \pm 7.3\%$	$90.1\% \pm 14.8\%$
B	$93.1\% \pm 6.0\%$	$97.0\% \pm 7.2\%$	$87.0\% \pm 13.4\%$
C	$95.7\% \pm 6.1\%$	$98.3\% \pm 1.8\%$	$91.3\% \pm 12.9\%$

and frequency domains which were recorded simultaneously with the same tracheal sound signal shown in Fig. B.2. Comparing the signals shown in Figures B.2 and B.5, it is clear that the tracheal sound recording is less affected by the ambient noise. In the spectrum of tracheal sounds, breath and snore sound segments are clear and the formants of the snore sounds can be detected. On the other hand, although breath sounds were audible in the ambient recording with lower intensity compared to those of tracheal recording, it is hard to detect breath and snore sounds' formants from the ambient sound's spectrum. Signal to noise ratio (SNR) values of breath and snore sound segments were estimated for ambient and tracheal recordings and the SNR values of tracheal recordings were found to be larger than those of the ambient recordings with 15 dB and 8 dB differences for snore and breath sound segments, respectively.

The results of Evaluations A, B and C for the signals recorded by the ambient microphone are shown in Fig. B.6 and Table B.5. Comparing the classifier performances for tracheal and ambient recordings (Tables B.4 and B.5), it is evident that in all experiments, the performance of the method using tracheal sound is superior to when using the sounds by ambient microphone. Furthermore, in all evaluations the standard deviation values

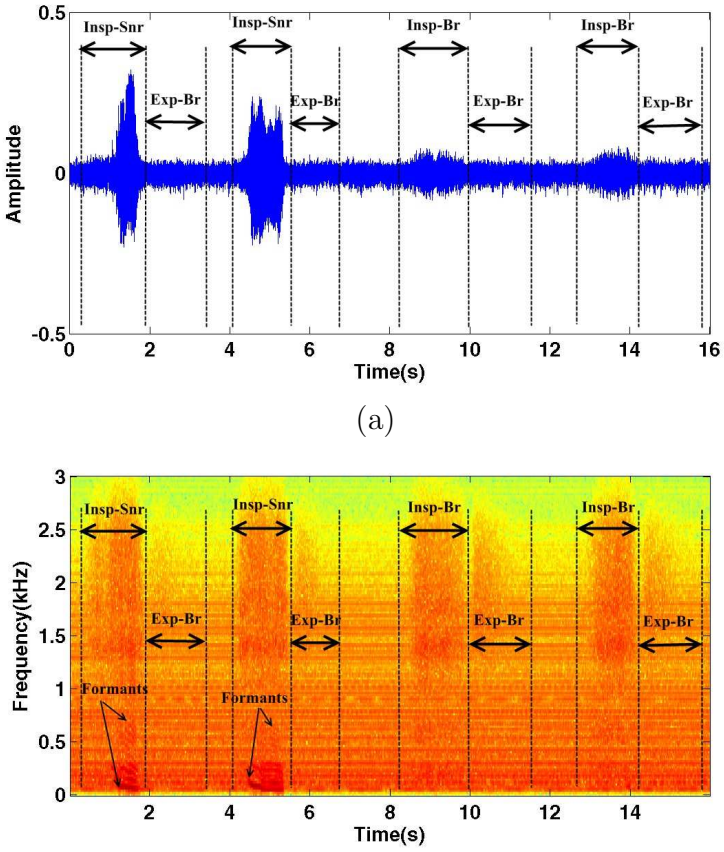


Figure B.5: *Samples of the ambient sound in a) time and b) time–frequency domains which were recorded simultaneously with the tracheal sound presnted in Fig. B.2. The sound segments are extracted and marked manually. Insp-Snr, Insp-Br and Exp-Br represent inspiration segments including snore, inspiration and expiration breath segments void of snore, respectively. The dark repeating frequencies in the time–frequency representation of ambient sounds (b) show the snore sounds' formant frequencies.*

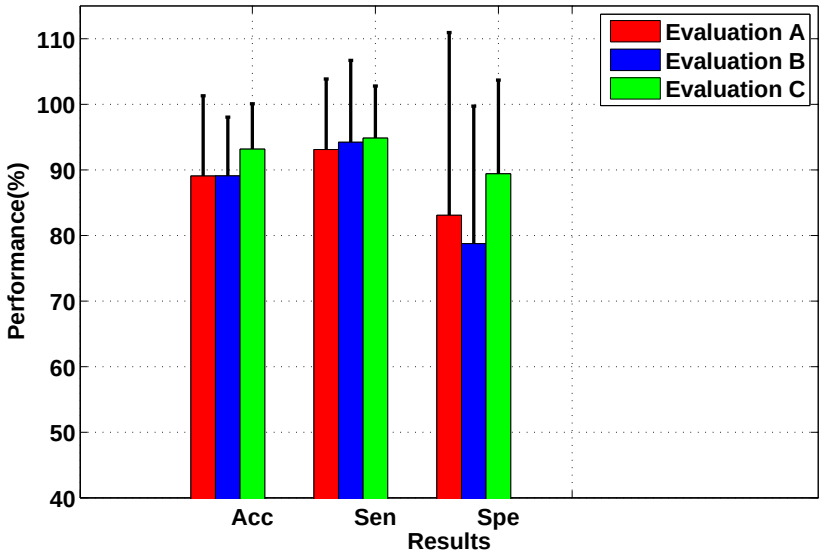


Figure B.6: *Classification results of Evaluations A, B and C for ambient recording. Sen, Spe, and Acc represent sensitivity, specificity and accuracy values, respectively.*

Table B.5: *The classification results of Evaluations A, B and C for ambient microphone. Sen, Spe, and Acc represent sensitivity, specificity and accuracy values, respectively.*

Evaluation	Accuracy	Sensitivity	Specificity
A	$89.1\% \pm 12.2\%$	$93.1\% \pm 10.8\%$	$83.1\% \pm 27.8\%$
B	$89.1\% \pm 8.9\%$	$94.2\% \pm 12.5\%$	$78.7\% \pm 21.0\%$
C	$93.2\% \pm 6.9\%$	$94.8\% \pm 7.9\%$	$89.4\% \pm 14.3\%$

are smaller for tracheal recording. Student *t–test* was performed to compare the classification results of Evaluation C for ambient and tracheal recordings and the p–values were $.01, .05$ and $.18$ for accuracy, sensitivity and specificity values, respectively. Although, the specificities were not significantly different for the two recordings, the accuracies and sensitivities of the classifier are significantly different.

B.3 Discussion

In this paper an automated method is presented to classify snore and breath segments from respiratory sounds recordings. For feature extraction and classification we used FLD to transform the features into a new 1D space providing the maximum distance between the classes. It was claimed in [21] that when respiratory sounds are recorded over the neck, the sounds characteristics change with respect to the body and neck position. In a previous study, it was shown that the intensity of respiratory tracheal sounds change with neck position and in the supine position the inspiratory and expiratory sounds are louder among both OSA patients and simple snorers [123]. We were interested to examine whether the body position affects the classification performance and if it does, to what degree. Hence, the performance of the method was investigated under different body postures.

In Evaluations A and B each patient was considered individually, and their data in supine and lateral positions were used to form training and test data sets. These positions were examined as they were the most common positions among our patients. The test and train data sets were swapped in the two experiments. From the results presented in Fig. B.4 and Table B.4, it can be observed that the sensitivity and specificity of the method in both experiments were similar, and in general there was not any significant difference in the classification performance due to the body posture. This was also confirmed by the high p–values (more than 0.3) of the *t–test* between the sensitivity, specificity and

accuracies of Evaluations A and B. The robustness of the classifier to the changes in the body posture could be due to the incorporation of other features rather than only the intensity of the sound segments, which overcomes the changes in the sounds intensities in different positions. Based on the insensitivity of the method's performance to the neck positions, the train and test data sets can be selected with no limitations on the neck position; this simplifies the classification process as it means that having *a priori* information on the body posture is not necessary for an accurate breath–snore classification. This is an important outcome of this study which to our knowledge has not been investigated in the previous studies for breath–snore classification.

The sensitivity and specificity results of our proposed method for classifying breath and snore sounds segments from tracheal recording were more than 90% (Table B.4, Evaluation C) which are superior to those of [113] (Table B.6). On the other hand, it should be noted that in all experiments, the sensitivities of the proposed method were higher than its specificities. Considering the definitions of TN and TP in equations (B.4) and (B.5), the higher sensitivities indicate better performance in detecting breath sound segments and that some of the snore sounds were misclassified and marked as breath. This could be due to the similarities in the energies of breath and snore sound segments for loud breaths sound segments.

In previous studies, different features of sound signals in temporal and spectral domains were used for breath and snore sounds classification [113-115, 92], where the majority of

those features represent the changes in the energy of sound signals in different frequency ranges and/or their linear combinations [114, 115, 92] (Table B.6). However, there are different mechanisms in the generation of breath and snore sounds; that can be reflected in their formants [120]. Hence, in this study we have included $F1$ as another feature to overcome the possible similarities between the intensities of breath and snore sound segments. Our preliminary results showed that when using only sounds energy and ZCR for classifying breath and snore sound segments, sensitivity and specificity of the classifier were $99.7\% \pm 0.7\%$ and $78.7\% \pm 16.6\%$, respectively. The noticeable difference between sensitivity and specificity results confirms the incapability of sound segments' energy in distinguishing between breath and snore sounds segments when they are similar in terms of loudness. However, the first formant frequency, $F1$, of the breath and snore sounds represents the resonance frequency of the upper airways, and that is different during sound generation of breath and snore sounds [120]. Hence, its contribution improved the specificity fo the classifier in detecting snore sound segments.

In most of the previous studies for breath/snore classification [110, 114, 115, 92] (except [113]), the sound signals were recorded with an ambient microphone (Table B.6). In this study, we have recorded sound signals simultaneously with ambient and tracheal microphones. The classification performance was compared for the same breath and snore sound segments extracted from tracheal and ambient recordings (Fig. B.4, Table B.4 and Fig. B.6, Table B.5). The results show that in all experiments, the performance of the

Table B.6: *Comparison of the classification results of different studies proposed for breath–snore classification.*

Study	Recording	Accuracy	Sensitivity	Specificity	PPR
22 temporal and spectral features [113]	Trachea	–	82.3%	–	90.8%
39–dimensional feature vector of energy and MFCC [114]	Ambient	89.0%	–	–	–
Spectral energy of the recorded signal [115]	Ambient	88.5%	–	–	96.2%
Zero crossings, signal's energy [92]	Ambient	95.5%	–	–	–
Our method: Zero crossings, signal's energy, first formant	Trachea	95.7%	98.3%	91.3%	95.2%
Our method: Zero crossings, signal's energy, first formant	Ambient	93.2%	94.9%	89.4%	94.4%

classifier was superior for the tracheal recordings. Furthermore, the standard deviation values were smaller for tracheal sound recordings indicating less variability in the results among different subjects. We have also used the features extracted from both tracheal and ambient recordings for classification of the sound segments and the accuracy; sensitivity and specificity values of the Evaluation C for the ambient recording were $95.6\% \pm 6.3\%$, $98.0\% \pm 2.6\%$ and $91.1\% \pm 14.1\%$, respectively. These results are very close to those of the tracheal recordings (Tables B.4 and B.5), which confirms that including data of ambient recordings in the classifier does not improve the classification performance.

The results of *t–test* analysis on the classifier performance for the ambient and neck recordings show that while the specificity values were not significantly different ($p =$

0.18), the accuracy and sensitivity values were significantly different ($p = 0.01$ and 0.05, respectively). The significant differences in sensitivity results and correspondingly in the accuracy values, confirms that tracheal recording is a better setting for breath and snore sound segments detection. On the other hand, according to the similar specificities of the classifier, it can be concluded that if the snore sound detection is of interest, tracheal and ambient recording will be similar in terms of classification. However, it should be noted that while ambient recording would be more convenient for the patient, the intensity of snore sounds will be much higher in tracheal recording [102, 116]; hence, it can be better quantized to different levels of severity.

In addition, when comparing the accuracies of the proposed method in detecting breath and snore sound segments from ambient recordings with the previous studies, our results (Table B.5, Evaluation C) were found to be similar to those reported in [115, 92] and superior to the results in [114].