

**Snoring Sounds Analysis: Automatic Detection,
Higher Order Statistics, and its Application for
Sleep Apnea Diagnosis**

by

Ali Azarbarzin

A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Department of Electrical and Computer Engineering
University of Manitoba
Winnipeg

Copyright ©2012 by Ali Azarbarzin

Abstract

Snoring is a highly prevalent disorder affecting 20-40% of adult population. Snoring is also a major indicative of obstructive sleep apnea (OSA). Despite the magnitude of effort, the acoustical properties of snoring in relation to physiological states are not yet known.

This thesis explores statistical properties of snoring sounds and their association with OSA. First, an unsupervised technique was developed to automatically extract the snoring sound segments from the lengthy recordings of respiratory sounds. This technique was tested over 5665 snoring sound segments of 30 participants and the detection accuracy of 98.6% was obtained.

Second, the relationship between anthropometric parameters of snorers with different degrees of obstruction and their snoring sounds' statistical characteristics was investigated. Snoring sounds are non-Gaussian in nature; thus second order statistical methods such as power spectral analysis would be inadequate to extract information from snoring sounds. Therefore, higher order statistical features, in addition to the second order ones, were extracted.

Third, the variability of snoring sound segments within and between 57 snorers with and without OSA was investigated. It was found that the sound characteristics of non-apneic (when there is no apneic event), hypopneic (when there is

hypopnea), and post-apneic (after apnea) snoring events were significantly different. Then, this variability of snoring sounds was used as a signature to discriminate the non-OSA snorers from OSA snorers. The accuracy was found to be 96.4%.

Finally, it was observed that some snorers formed distinct clusters of snoring sounds in a multidimensional feature space. Hence, using Polysomnography (PSG) information, the dependency of snoring sounds on body position, sleep stage, and blood oxygen level was investigated. It was found that all the three variables affected snoring sounds. However, body position was found to have the highest effect on the characteristics of snoring sounds.

In conclusion, snoring sounds analysis offers valuable information on the upper airway physiological state and pathology. Thus, snoring sound analysis may further find its use in determining the exact state and location of obstruction.

"Research is what I'm doing when I don't know what I'm doing."

Wernher von Braun

*To My Wife Zohreh,
for her constant loyalty and support.*

Acknowledgments

This dissertation would not have been possible without the support of many people who in one way or another contributed in the preparation and completion of this study.

I would like to express my gratitude to my supervisor, Prof. Zahra Moussavi, whose expertise, understanding, and patience, added considerably to my graduate experience. Prof. Zahra was abundantly helpful and offered invaluable assistance, support and guidance.

Deepest gratitude are also due to the members of the supervisory committee, Prof. Mirosław Pawlak, Prof. Attahiru Sule Alfa, Dr. Zoheir Bshouty without whose knowledge and assistance this study would not have been successful. I would like to thank Dr. Atul Malhotra, my external examiner, for his invaluable review and comments.

This study was financially supported by Tele-Communication Research Labs (TRLabs), Canada and NSERC, Canada. I thank the administration of TRILabs, especially Dr. Sergio Camorlinga.

I would like to acknowledge and thank Dr. Ellini Gianouli and 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 all my friends, especially Azadeh Yadollahi, Ardalan Aarabi, Ali Akbar Samadani, Aman Montazeri, Samaneh Sarraf, Saiful

Huq, and Davood Karimi. Their support and friendship enabled me to finish my research.

I would like to thank my parents, Mohammadali and Pouran, my brother, Touraj, and my sisters, Atefeh and Samaneh, for the support they provided me through my entire life.

Last, but not least, I must acknowledge my wife and best friend, Zohreh, without whose love and encouragement, I would not have finished this thesis.

Contents

Abstract	i
Acknowledgments	v
List of Figures	xiii
List of Tables	xx
Abbreviations	xxiii
1 Introduction	1
1.1 Sleep Monitoring Techniques and Snoring Sounds' Characteristics (Chapter 2)	4
1.2 Snoring Sound Extraction from Respiratory Sound Recording (Chapter 3)	5
1.3 Higher Order Statistics of Snoring Sound (Chapter 4)	7
1.4 Variability of Snoring Sounds as a Signature of OSA (Chapter 5)	8

1.5	Effect of Body Position, Sleep Stage, and Blood Oxygen Level on Snoring Sounds (Chapter 6)	9
2	Sleep Monitoring Techniques and Snoring Sounds' Characteristics	11
2.1	Polysomnography (PSG) and Sleep-Disordered Breathing (SDB) Home Diagnosis	12
2.2	Recording Setup	15
2.3	Acoustical Characteristics of Snoring Sounds	17
2.3.1	Definition of Snoring Sound Segment	17
2.3.2	Time and Frequency domain characteristics of Snoring Sounds	17
3	Snoring Sound Extraction From Respiratory Sound Recordings	21
3.1	Existing Algorithms	21
3.2	Snoring Sound Detection Algorithm	23
3.2.1	Data Recording	23
3.2.2	Signal Analysis	24
3.2.3	Modified Vertical box (V-Box) control chart	25
3.2.4	Feature extraction	28
3.2.5	Unsupervised Classification using Fuzzy C-means Clustering(FCM)	30
3.3	Results and Evaluation of the snoring sound detection algorithm . .	32
3.4	The effect of segmentation parameters	34

3.5	Computational Complexity	35
3.6	Discussion and Concluding Remarks on Snoring Sound Detection Algorithm	37
4	Higher Order Statistical Properties of Snoring Sounds	40
4.1	Why Higher Order Statistics of Snoring Sounds?	40
4.2	Method	43
4.2.1	Data Recording	43
4.2.2	Higher Order Statistics (HOS)	44
4.2.2.1	Definition of Bispectrum and Bicoherence	45
4.2.2.2	Bispectrum and bicoherence estimation	47
4.2.3	Feature Extraction	49
4.2.3.1	Median Bifrequency (MBF) computation	49
4.2.3.2	PMBF computation	50
4.2.3.3	Skewness and Kurtosis	50
4.2.3.4	1 st formant frequency and energy	51
4.2.3.5	Calculation of features	51
4.2.4	Statistical Analysis	53
4.2.5	Classification	53
4.3	Results and Discussion	54
4.4	Conclusion	60
5	Inter- and Intra-subject Variability of Snoring Sounds	62

5.1	Method	65
5.1.1	Data	65
5.1.2	Snoring sound extraction from respiratory sounds	65
5.1.3	Feature Extraction	66
5.1.4	Statistical Analysis	67
5.1.5	Regression and Classification Analysis	68
5.2	Results	69
5.2.1	The difference between snoring sounds of three classes within an individual	69
5.2.2	The total variation norm and classification	71
5.3	Discussion	74
6	Why Do Snoring Sounds Form Distinct Clusters?	83
6.1	Method	85
6.1.1	Data Recording	85
6.1.2	Feature Extraction	85
	Power (P)	85
	Zero Crossing Rate (<i>ZCR</i>)	86
	500Hz Sub-band normalized energy ($E_1 - E_2$)	86
	Skewness and kurtosis	86
	Formant Frequencies ($F_1 - F_3$)	87
	Crest factor (CF)	87

Spectral entropy (SE)	88
Central Tendency Measure (CTM)	88
6.1.3 Principal Component Analysis	89
6.1.4 Grouping intra-subject snoring sound segments based on PSG information	89
6.1.5 One-dimensional and two-dimensional probability density function	91
6.1.6 Statistical Analysis	94
6.2 Results	95
6.2.1 Principal Component Analysis	95
6.2.2 One-Dimensional and two-dimensional pdfs and ROC curves	97
6.2.3 Statistical Analysis	99
6.2.3.1 Body Position	99
6.2.3.2 Sleep Stage	104
6.2.3.3 Blood Oxygen Saturation	106
6.2.3.4 Finding the most effective factor changing the snoring sounds characteristic	109
6.3 Discussion	111
7 Summary and Concluding Remarks	114
7.1 Summary of Contributions	114
7.2 Future Work Recommendations	116

List of Figures

2.1	Snoring sound: Recording to clinical applications.	16
2.2	Snoring sound recorded by tracheal microphone from a heavy snorer (AHI=4.9)	18
2.3	Snoring sound recorded by tracheal microphone from a heavy snorer (AHI=5.9)	19
3.1	A typical 20s of a tracheal sound recording including breath and snoring segments and the segmentation result using V-Box algo- rithm. At this stage, all potential snoring sound segments are iden- tified. E_i denotes segment i , where i is 1, 2, 3, . . . For example, in this 20s interval, there are 6 segments.	27

-
- 3.2 Clustered feature vector using FCM method. This 2-D feature vector was obtained using a projection from 10-D feature space onto a 2-D feature space by PCA. The original 10-D feature vector was obtained by calculating 500Hz energy band distribution from a segmented sound signal. The big rectangle is a zoomed in version of the snoring cluster located around horizontal axes. 30
- 3.3 Successful removal of noise and breathing segments after clustering. Note that snoring segment E_6 was not detected and therefore there is a false alarm in this case. 33
- 4.1 An extracted snoring sound segment and its spectrogram. 44
- 4.2 Kernel density estimate of f^p (projected median bifrequency), f_1^{mp} (bifrequency along horizontal axes), and f_2^{mp} (bifrequency along vertical axes) for a typical subject. Note the asymmetry of the distributions. 52
- 5.1 Detailed flow chart of proposed algorithm. 67
- 5.2 A 1-minute snoring epoch when there is no hypoapnea event. This type of snoring is called non-apneic snoring in which the characteristics of snoring sounds do not change significantly from snore to snore. A typical snoring sound segment is shown with dashed rectangle and a breathing phase is illustrated with solid parallel lines. 71

-
- 5.3 A 1-minute snoring epoch when there are hypopneic events. This type of snoring is called hypopneic snoring in which the characteristics of snoring sounds change drastically from snore to snore. It can be clearly observed that even the duration of snoring sounds are highly variable within one minute of recording. The snoring sound segments are marked with dashed rectangles. 72
- 5.4 A 1-minute snoring epoch when there are apneic events. This type of snoring is called post-apneic snoring in which the first sound usually appears as a very loud snoring. Depending on the severity of OSA, an apneic event may be accompanied by more snoring segments with different characteristics. In this Figure, there are two apneic events each 13s long. 73
- 5.5 The intra-subject variability of snoring sound segments' peak frequency (F_p) for an individual. As seen, the variation of F_p is the lowest in non-apneic snoring class, and highest in hypopneic snoring class. However, the total variation for this snorer would be high. The horizontal axis is the number of snoring sound segment. 76
- 6.1 A scatter plot of snoring sound segments within a typical snorer in a 2-D feature space. Feature 1 and feature 2 were standardized such that they have zero mean and unit variance. 84

-
- 6.2 Top: pdf of snoring sounds' first principal component in two different classes. Sweeping threshold line was used to calculate the areas α and β . Bottom: ROC curve constructed by plotting $1 - \beta$ with respect to α 94
- 6.3 Bi-plot of transformed snoring sounds' features (the first two *PCs* with highest eigenvalues) along with original feature vector. Each vector shows the magnitude and sign of each feature's contribution to the first two *PCs*. 96
- 6.4 Scatter matrix of snoring sound segments' *PCs* categorized based on body position. The diagonal graphs show the histogram of each *PC*. 97
- 6.5 Scatter matrix of snoring sound segments' *PCs* categorized based on sleep stage. The diagonal graphs show the histogram of each *PC*. 98
- 6.6 Scatter matrix of snoring sound segments' *PCs* categorized based on blood oxygen level. The diagonal graphs show the histogram of each *PC*. 99
- 6.7 (a)1-D pdf of snoring sound segments' first *PC* categorized based on body position. (b) Contour plot of snoring sound segment of the first two *PCs* categorized based on body position. As shown in this figure, the 2-D pdf of prone position is significantly different than that of side and supine positions. 100

- 6.8 (a) 1-D pdf of snoring sound segments' first PC categorized based on sleep stage. (b) Contour plot of snoring sound segment of the first two PC s categorized based on sleep stage. As shown in this figure, the 2-D pdf of NREM, REM, and Arousal are different. . . . 101
- 6.9 (a) 1-D pdf of snoring sound segments' first PC categorized based on blood oxygen level. (b) Contour plot of snoring sound segment of the first two PC s categorized based on blood oxygen level. As shown in this figure, the 2-D pdf of NREM, REM, and Arousal are different. 102
- 6.10 The mean and standard deviation of distance measures between body position classes' pdfs among all snorers (a). The mean and standard deviation of L_1^{2D} , the highest separation is between prone and supine positions, then, between prone and side, and the lowest is between supine and side positions. (b) The mean and standard deviation of L_1^{1D} . (c) The mean and standard deviation of $AUROC$. 105
- 6.11 The mean and standard deviation of distance measures between sleep stage classes' pdfs among all snorers (a). The mean and standard deviation of L_1^{2D} , the highest separation is between NREM and REM sleep stages ($L_1^{2D} = 0.64 \pm 0.2$), then, between REM and Arousal ($L_1^{2D} = 0.63 \pm 0.16$), and the lowest is between NREM and Arousal. (b) The mean and standard deviation of L_1^{1D} . (c) The mean and standard deviation of $AUROC$ 107

-
- 6.12 The mean and standard deviation of distance measures between principal components' class densities of normal and desaturation among all snorers. (a). The mean and standard deviation of L_1^{2D} for all bivariate *PCs*. (b) The mean and standard deviation of L_1^{1D} . (c) The mean and standard deviation of *AUROC*. 108
- 6.13 The mean and standard deviation of distance measures between class densities among all snorers grouped based on all categorical variables. (a). The mean and standard deviation of L_1^{2D} , the highest separation is for body position regardless of which bivariate has been used to estimate the 2-D pdf (highest $L_1^{2D} = 0.75 \pm 0.18$), then, for blood oxygen level ($L_1^{2D} = 0.49 \pm 0.2$), and finally, for sleep stage ($L_1^{2D} = 0.47 \pm 0.18$). (b) The mean and standard deviation of L_1^{1D} . (c) The mean and standard deviation of *AUROC*. 110
- A.1 The ratio of normalized energy between ambient and tracheal recordings 147
- B.1 The mean and standard deviation of distance measures between body position classes' pdfs among all snorers (a). The mean and standard deviation of L_1^{2D} , the highest separation is between prone and supine positions, then, between prone and right, and the lowest is between left and right positions. (b) The mean and standard deviation of L_1^{1D} . (c) The mean and standard deviation of *AUROC*. 152

-
- B.2 The mean and standard deviation of distance measures between sleep stage classes' pdfs among all snorers (a). The mean and standard deviation of L_1^{2D} , the highest separation is between REM and stage 3, then, between REM and stage 1, and the lowest is between stage 2 and Arousal. (b) The mean and standard deviation of L_1^{1D} . (c) The mean and standard deviation of *AUROC*. 153
- B.3 The mean and standard deviation of distance measures between class densities among all snorers grouped based on all categorical variables. (a). The mean and standard deviation of L_1^{2D} , the highest separation is for body position regardless of which bivariate has been used to estimate the 2-D pdf, then, for sleep stage, and finally, for blood oxygen level. (b) The mean and standard deviation of L_1^{1D} . (c) The mean and standard deviation of *AUROC*. 154

List of Tables

3.1	Anthropometric information of participating individuals (BMI : Body Mass Index).	24
3.2	Classification Results and total number of breaths for a sample case with 171 snoring segments.	33
3.3	Classification results using tracheal recordings.	34
3.4	Classification Results for different values of overlap ($L = 50ms$, $\theta = 0.95$).	34
3.5	Classification Results for different values of L (Overlap=80%, $\theta = 0.95$).	34
3.6	Classification Results for different values of θ (Overlap=80%, $L = 50ms$).	35
4.1	Anthropometric Information of Participating Individuals.	44

4.2	Anthropometric information of two subsets selected for classification. Experiment A: The OSA and apneic groups were matched for gender, BMI, and height parameters. Experiment B: The OSA and apneic groups were not matched for gender, BMI, and height parameters.	54
4.3	The results of Kendall's Tau-b and Kruskal-Wallis tests for five anthropometric parameters. ns:non-significant, *: significant at level 5%, **: significant at level 1%.	55
4.4	Naïve Bayes classification results for different combination of conventional and HOS features.	60
5.1	Anthropometric Information of the study participants.	65
5.2	Number of extracted snoring sound segments from tracheal recordings.	70
5.3	Result of Kruskal-Wallis test on non-OSA group within each individual. Only 6 out of 15 non-OSA participants had more than one class of snoring sounds. NA#, HA#, PA# are the number of non-Apneic, hypopneic, and post-apneic snoring sounds, respectively. ns: not significant, and *: significant ($0.01 < p < 0.05$), **: very significant ($0.0001 < p < 0.01$) and ***: highly significant ($p < 0.0001$).	74
5.4	Result of Kruskal-Wallis test on OSA group within each individual. 9 patients had only one class of snoring and therefore they were excluded.	75
5.5	Result of correlation between AHI and TV for all features.	76

5.6	Regression analysis between TV_{F_p} and AHI.	77
5.7	Result of Linear Discriminant analysis	77
5.8	Comparison between OSA diagnosis method using snoring sounds. .	78
6.1	Anthropometric Parameters of the study participants.	85
6.2	Results of ANOVA on comparison groups of body position. ns: not significant ($p>0.05$), and *: significant ($0.01<p<0.05$), **: very significant ($0.001 <p<0.01$) and ***: highly significant ($p<0.001$). .	104
6.3	Results of ANOVA on comparison groups of sleep stage.	106
6.4	Results of ANOVA on two by two categorical variables and on all three groups of categorical variables. BP: Body Position, OL: blood Oxygen Level, SS: Sleep Stage	109
A.1	Classification results for tracheal recordings.	148
A.2	Classification results for ambient recordings.	148

Abbreviations

AHI	A pnea H ypopnea I ndex
ANOVA	A nalysis of V ariance
AR	A utoregressive
BMI	B ody M ass I ndex
CF	C rest F actor
CTM	C entral T endency M easure
DBI	D avies- B ouldin- I ndex
DFT	D iscrete F ourier T ransform
FCM	F uzzy C -means C lustering
FN	F alse N egative
FP	F alse P ositive
HOS	H igher O rders S tatistics or S tatistical
LOOCV	L eave O ne O ut C ross V alidation
LPC	L inear P redictive C oding
KWAV	K ruskal- W allis A nalysis of V ariance
MBF	M edian B i- F requency

OSA	O bstuctive S leep A pnea
PCA	P rincipal C omponent A nalysis
PFR	P eak F actor R atio
PMBF	P rojected M edian B i- F requency
PPV	P ositive P redictive V alue
PSD	P ower S pectrum D ensity
PSG	P olysomnography
REM	R apid E ye M ovement
SDB	S leep- D isordered B reathing
SE	S pectral E ntropy
TN	T rue N egative
TP	T rue P ositive
TV	T otal V ariation
UPPP	U vulopalatopharyngoplasty
V-Box	V ertical B ox
ZCR	Z ero C rossing R ate

Chapter 1

Introduction

Snoring is a common disorder. There have been several studies on the prevalence of snoring among men and women [1–3]. The commonly accepted prevalence rate of snoring is approximately 40% in men and 20% in women, which also increases by age [2–5]. Snoring can occur during natural or drug-induced sleep. It is observed more often during inspiratory and less often during expiratory phase of the respiratory cycle [6].

Snoring sound is caused either by oscillation of the structures such as soft palate, epiglottis, pharyngeal walls, the collapsible parts of the upper airway, or by turbulence of air near a partially obstructed upper airway [7–9]. The increased upper airway collapsibility is considered to aggravate snoring and other sleep-related disorders [10, 11]. Nasal obstruction [12], upper airway morphology, obesity, and excess alcohol and cigarette consumption are thought to affect snoring [13, 14]. Men are considered more susceptible to snoring than women [15]. Theoretically,

this difference can be due to bony configuration, fat deposition, or soft tissue structure [15].

Parameters such as respiratory airflow, upper airway cross-sectional area and/or diameter, and upper airway collapsibility and resistance contribute to the generation of snoring sounds; though the most important factor is airflow [16]. The upper airway morphology has an impact on the airflow by creating areas of transition between turbulent and laminar airflow. These areas could limit the airflow resulting in pressure changes and increased turbulence in the airway [17].

The snoring sounds' characteristics may vary from night to night. There are several factors accounting for variation in snore appearance from night to night or even within the same night for the same person. In addition to anthropometric factors (i.e. weight, height, smoking history, etc), as mentioned before, the breathing pathway [18], sleep stage and body position [19], natural or induced sleep [20], and the sites of narrowing in the upper airway [21] have been shown to affect the snoring sounds' characteristics.

From clinical view point, snoring is considered as a major indicative of obstructive sleep apnea (OSA) [22]. A complete cessation of breathing for at least 10s is defined as an apneic event, while a significant reduction ($>50\%$) of airflow for more than 10s is considered as hypopnea [23]. Resumption of breathing, after apnea, is usually accompanied by a sequence of snoring segments [24]. OSA is a major contributor to cardiovascular diseases [25, 26]; it causes daytime sleepiness [27–29], leads to impaired job performance [30], and increases the risk of accidents

[31–33]. There are several differences between upper airway structures of people with and without OSA. Imaging studies have shown that in general non-OSA snorers have significant narrowing at the tongue base and the hyoid bone levels, while people with OSA have narrower air space at the velopharyngeal level [34]; also, they have been shown to have a reduced muscle tone during sleep [16] and more collapsible airway during wakefulness [35] and sleep [36]. In addition to sleep apnea and other sleep disorders, snoring is considered as a risk factor for hypertension and heart disease [37–39].

The treatments for snoring and/or sleep apnea range from simple medical devices to surgical options. However, the success rate of these treatments has been reported to be low [40–42]. A treatment is usually assessed using subjective and objective assessment methods. Subjective assessments include a questionnaire filled by the snorer and his/her bed partner about the snoring frequency and loudness usually before and after surgical treatment such as uvulopalatopharyngoplasty (UPPP) [43–45]. Objective assessments include recoding snoring sounds of the person before and after the treatment (if there is any), and then comparing the characteristics of the sounds.

In an early study [46], ten features of the recorded snoring sound were examined to evaluate the level of snoring of 32 subjects before and after surgical treatment in comparison with the subjective assessments of the patients and their bed partners. The correlation between subjective and objective assessments was reported to be weak. Though, the strongest correlation was found only in supine posture,

where the bed partner was most disturbed by the loudest snoring sound [46]. More recently in another study, the subjective (Spouse Dissatisfaction Score: a numerical grade between 0 and 3) and objective (snoring sound recordings at home) evaluations of the tongue base snoring after the use of an oral appliance were compared; no correlation was found between the two different assessments [47]. This may be due to highly subjective nature of snoring perception and annoyance level. Therefore, objective assessment of snoring sounds would be an alternative method to accurately assess the snoring before and after a treatment [48, 49].

Acoustical analysis of snoring sounds can be deployed for objective assessment of snoring cause and treatment strategy. It also can help diagnose snorers with OSA, help with identifying the site of obstruction, and model the generation mechanism of snoring sounds. In this thesis, different aspects of the acoustical analysis of snoring sounds are discussed. The organization of the thesis is detailed as the followings.

1.1 Sleep Monitoring Techniques and Snoring Sounds’ Characteristics (Chapter 2)

Chapter 2 introduces sleep monitoring techniques including Polysomnography (PSG), the current gold standard to diagnose sleep-disordered breathing (SDB). This chapter also reviews some of the proposed algorithms that record snoring

sounds for OSA screening purposes. To date, there is no standard on the data recording setup in relation to the type of microphone, place of microphone, sampling frequency, bit resolution, amplification, and preprocessing. In Chapter 2, data acquisition system and its considerations are discussed. The time and frequency domain properties of snoring sounds such as existing patterns in snoring wave and their power spectrum are also discussed in Chapter 2.

The work presented in Chapter 2 has been published in:

- A.Azarbarzin and Z.Moussavi, "A Comparison between Recording Sites of Snoring Sounds in Relation to Upper Airway Obstruction," IEEE EMBS, San Diego, 2012.
- A.Azarbarzin and Z.Moussavi, "A comparison between recording sites of snoring sounds," Proc. IJAS, Toronto, May 2012.

1.2 Snoring Sound Extraction from Respiratory Sound Recording (Chapter 3)

Once the respiratory sounds are recorded and preprocessed, the snoring sound segments need to be identified (manually or automatically) for further analysis. The accuracy of this stage is important because the error at this stage can substantially change the results of consequent sound analysis. For instance, one is interested to investigate the location of obstruction using only acoustical analysis of snoring

episodes. If the snoring segments are not correctly identified, the results of localizing the apneic events will be unreliable. Therefore the snoring segments must be accurately extracted and separated from other noises such as speech, swallowing, breathing or environmental noises, i.e. opening/closing door, blanket movement, and fan noise. On the other hand, manual detection of snoring segments is a very time-consuming task as the length of an overnight respiratory sound recording is approximately 8 hours. To overcome this issue, Chapter 3 introduces an automatic and unsupervised algorithm to extract the snoring sound segments from the respiratory sound recordings.

The work presented in Chapter 3 has been published in:

- A.Azarbarzin and Z.Moussavi, "Automatic and Unsupervised Snore Sound Extraction from Respiratory Sound Signals," *IEEE Trans Biomed Eng.* Vol 58, pp. 1156-1162, 2011, DOI:10.1109/TBME.2010.2061846.
- A.Azarbarzin, Z.Moussavi, "Unsupervised Classification of Respiratory Sound Signal into Snore/No-Snore classes," *IEEE EMBS*, Buenos Aires, pp. 3666-69, 2010.

1.3 Higher Order Statistics of Snoring Sound (Chapter 4)

Second order statistical techniques such as power spectrum, and correlation-based methods are based on the assumption that the signal generating process is Gaussian and Linear. However, if this assumption does not hold, second order statistical methods will not be able to fully extract the vital information of a signal. In most of the previous studies [50–58], it was assumed that the snoring sounds were generated by a Gaussian and linear process; hence, the conventional second order measures were used for acoustical analysis of snoring sound. In this thesis, we studied whether the snoring sound signals are generated by a Gaussian and/or Linear process. In Chapter 4, the results of a detailed analysis show that most of the snoring sound signals are generated by non-Gaussian and non-linear processes. Therefore, extraction of higher order statistical (HOS) features helps obtain complementary information of the signals. Another issue discussed in Chapter 4 is the effect of anthropometric parameters (such as age, height, Body Mass Index (BMI), gender, and Apnea Hypopnea Index (AHI)) on the snoring sounds' features. The result of this part of thesis shows that anthropometric parameters affect the characteristics of the sounds to some extent.

The work presented in Chapter 4 has been published in:

- A.Azarbarzin and Z.Moussavi, "Snoring Sounds' Statistical Characteristics Depend on Anthropometric Parameters," *Journal of Biomedical Science and*

Engineering, Vol 5, pp. 245-254, 2012, DOI: 10.4236/jbise.2012.55031.

- A.Azarbarzin and Z.Moussavi, "Relationship between the Higher Order Statistical Features of Snoring Sounds and Anthropometric Factors of Snorers," SLEEP, Volume 34, Abstract Supplement, pp. A326, 2011.
- A.Azarbarzin and Z.Moussavi, "Nonlinear properties of snoring sounds," Proc. ICASSP, Prague, pp. 4316-19, 2011.
- A.Azarbarzin and Z.Moussavi, "Do Anthropometric Parameters Change the Characteristics of Snoring Sound?," IEEE EMBS, Boston, pp. 1749-52, 2011.

1.4 Variability of Snoring Sounds as a Signature of OSA (Chapter 5)

One of the key factors affecting the characteristics of snoring sounds is the severity of flow reduction (or obstruction) in the upper airway. Chapter 5 reports on the classification results of the snoring sound segments into different classes based on the severity of flow reduction in the upper airway. The sound recording was performed simultaneously with Polysomnography (PSG). Given the PSG information, the snoring sound segments were labeled as non-apneic (occurring in the absence of any apneic event), hypopneic (occurring when there was 50%-90% flow reduction), and post-apneic (occurring after completet obstruction). Next, a non-parametric statistical analysis was run within each subject among the three

aformentioned classes. The result showed that there were significant differences among three classes in terms of sounds' characteristics. Moreover, it was observed that in non-apneic class, there was a subtle variation in the characteristics of snoring sounds from breath to breath, while it was opposite in other two classes. This observation provided the motivation to use the snoring sounds' variation and develop a technique for OSA screening.

The work presented in Chapter 5 has been published in:

- A.Azarbarzin and Z.Moussavi, "Snoring Sounds Variability as a Signature of Obstructive Sleep Apnea," *Journal of Medical Engineering & Physics*, In press, 2012, DOI:10.1016/j.medengphy.2012.06.013.
- Z.Moussavi and A.Azarbarzin, "Relationship Between Obstructive Sleep Apnea And Snoring Type," *American Thoracic Society International Conference*, San Francisco, pp. A6433, 2012.

1.5 Effect of Body Position, Sleep Stage, and Blood Oxygen Level on Snoring Sounds (Chapter 6)

Snoring sounds are observed to form distinct clusters within a snorer in a multidimensional feature space. Chapter 6 investigates why this occurs. After extracting several features (12 in total), categorical variables including body position, sleep stage, and blood oxygen level were obtained from PSG score sheet. Then, the

snoring sounds were divided into different classes based on each categorical variable. For example, they were categorized based on body position to 3 classes including prone, supine, and side. The probability density function of each class of snoring sounds was estimated and finally, the separation between classes was measured. Finally, the most prominent categorical variable as well as the most affective class within each categorical variable were determined using Analysis of Variance (ANOVA). The result showed that change from any posture to prone affects the snoring sounds more than any other change. Moreover, it was found that change from any sleep stage to Rapid Eye Movement (REM) had the highest effect on the snoring sounds (within the same subject). Finally, it was observed that change in body position resulted in significantly larger effect on the snoring sounds than a change in sleep stage or blood oxygen level.

The work presented in Chapter 6 has been published in:

- A.Azarbarzin and Z.Moussavi, "Snoring Sounds Intra-Subject Variability," Submitted to Medical and Biological Engineering and Computing, 2012.
- A.Azarbarzin and Z.Moussavi, "A Comparison between Recording Sites of Snoring Sounds in Relation to Upper Airway Obstruction," IEEE EMBS, San Diego, 2012.

Chapter 2

Sleep Monitoring Techniques and Snoring Sounds' Characteristics

This chapter introduces the Polysomnography (PSG) -the current gold standard for sleep-disordered breathing (SDB)- and some proposed algorithms that deploy snoring sounds to diagnose OSA. It also reviews recording setup and preprocessing stage of this project. Finally, it discusses the basic time-frequency domain characteristics of snoring sounds.

2.1 Polysomnography (PSG) and Sleep-Disordered Breathing (SDB) Home Diagnosis

The current gold standard for SDB (sleep apnea, in particular) diagnosis is Polysomnography (PSG) [59–61]. PSG study monitors brain waves, heart rhythm, eye movements, muscle activity, respiratory airflow, and blood oxygen level. It requires the patient to spend the entire night in the sleep lab. Moreover, there are other difficulties associated with PSG such as long waiting list, connecting many wires and electrodes to the patient, and additional cost. To overcome the drawbacks of PSG, several alternative diagnosis methods have been proposed using airflow, oxygen saturation, and/or snoring [59].

Anatomic abnormalities of the upper airway play an important role in the pathogenesis of OSA [62]. These abnormalities are expected to change the acoustical properties of the respiratory sounds, e.g. snoring sounds. Several researchers investigated the possibility of using snoring sounds as a biomarker of OSA. In an early study [63], the snoring sounds of 27 PSG study patients (18 non-OSA and 9 OSA) were analyzed. They compared two measures of sound level and third-octave frequency band between the two groups of non-OSA and OSA participants. The results showed larger high frequency components for OSA group (both groups had large low frequency peaks in linear sound level around 80 Hz). Hawke Index (HI), the ratio between the overall A-weighted and linear sound levels, was positively correlated with AHI [63].

Pitch analysis was used in [64] and [65] to diagnose OSA. In [64], 10 segments of snoring were extracted from each participant. They were able to achieve a sensitivity of 92.3% and specificity of 90.7% for detecting snoring segments associated with OSA [64]. In [65], the snoring sounds of 8 OSA (236 segments) and 8 non-OSA (447 segments) participants were extracted. They quantified the pitch of snoring with mean value, standard deviation, and density for both groups. They used all three measures to separate non-OSA snoring sounds from OSA snoring sounds. They found a sensitivity of 64.4% and specificity of 58.5% for detecting OSA snoring segments [65].

The diagnostic ability of tracheal sounds was investigated in a large population of 383 patients who were referred for PSG study [66]. They used an automatic computer program to calculate tracheal sound respiratory disturbance index (TS-RDI), the number of transient falls in the time series of moving average of the logarithmic power of tracheal sound. The correlation between TS-RDI and AHI was reported to be 0.93. They used TS-RDI to diagnose OSA patients whose $AHI > 5$ or $AHI > 15$. They achieved a sensitivity and specificity of 93% and 67% for AHI cutoff value of 5 and 79% and 95% for the AHI cutoff value of 15 [66].

Intra-Snore-Pitch-Jump (ISPJ) feature of snoring sounds was employed in another study [67] to diagnose OSA. They used the snoring sounds of 45 participants who were referred for PSG study. They divided their data set to training (16 subjects) and testing (29 subjects) to evaluate the performance of ISPJ-based OSA detection algorithm. Due to a lack of standard for separating OSA and non-OSA participants

based on their PSG AHI, they evaluated the performance of their algorithm using several thresholds i.e. 5, 10, 15, and 30. A sensitivity of 86-100% and specificity of 50-80% were achieved for ISPJ-based OSA detection algorithm [67]. This study resulted in a highly variable sensitivity and specificity among different scenarios of AHI threshold. Moreover the number of subjects in training and testing data set was small.

LPC analysis and formant frequencies of snoring sounds were used in [68] to diagnose OSA in a population of 30 OSA and 10 non-OSA participants. They extracted first three formant frequencies from snoring sounds and attempted to separate OSA from non-OSA group using these features. They achieved a sensitivity of 88% and specificity of 82% using first formant frequency with a threshold of 470 Hz [68]. One of the limitations of this study is the relatively small data set with an unmatched anthropometric parameters.

More recently, in another study, tracheal sounds and pulse oximetry were used to develop a technique for monitoring and detection of sleep apnea [69]. They recorded the respiratory sound and pulse oximetry from 66 participants. After detection of snoring and breathing episodes, they utilized five features extracted from breathing, snoring, and pulse oximetry signals. The features included energy of breathing segments, duration of breathing segments, duration of snoring segments, the amount of drop in pulse oximetry signal, and amplitude of pulse oximetry signal. The diagnostic sensitivity and specificity of the method were evaluated for different AHI cutoff values, i.e. 5, 10, 15, and 20. The sensitivity

and specificity mean values were in the range of [74.3, 91.6] and [82.4-97.8] respectively [69]. In contrast to other studies, in this study two channels of data (tracheal sound and oximetry signal) were used for OSA diagnosis.

As seen, there is a large variation in accuracy, recording setup, number of participants, AHI cutoff values, etc. This is mainly due to a lack of standard framework for different stages of algorithm. The recording setup and snoring sound characteristics are discussed in the following sections.

2.2 Recording Setup

To extract the clinical information from the snoring sounds, it is very crucial to have a reliable sound recording system. The type of microphone, place of microphone, sampling frequency, bit resolution, amplification, and preprocessing should be considered properly. Figure 2.1 shows a block diagram of a snoring sound system from recording to some applications of the signals.

In this study, two miniature omni-directional lavalier microphones (Sony: ECM-77B) were used: one was placed over the suprasternal notch of the patient's trachea (tracheal microphone), and the other was hung in the air about 20-30 cm away from the patient's head (ambient microphone). The microphones had a frequency response of 40 Hz - 20 KHz. The tracheal microphone was embedded in a chamber (diameter of 6mm) and attached to the skin using double-sided adhesive tape. A soft neck-band was also used to support the chamber throughout the night.

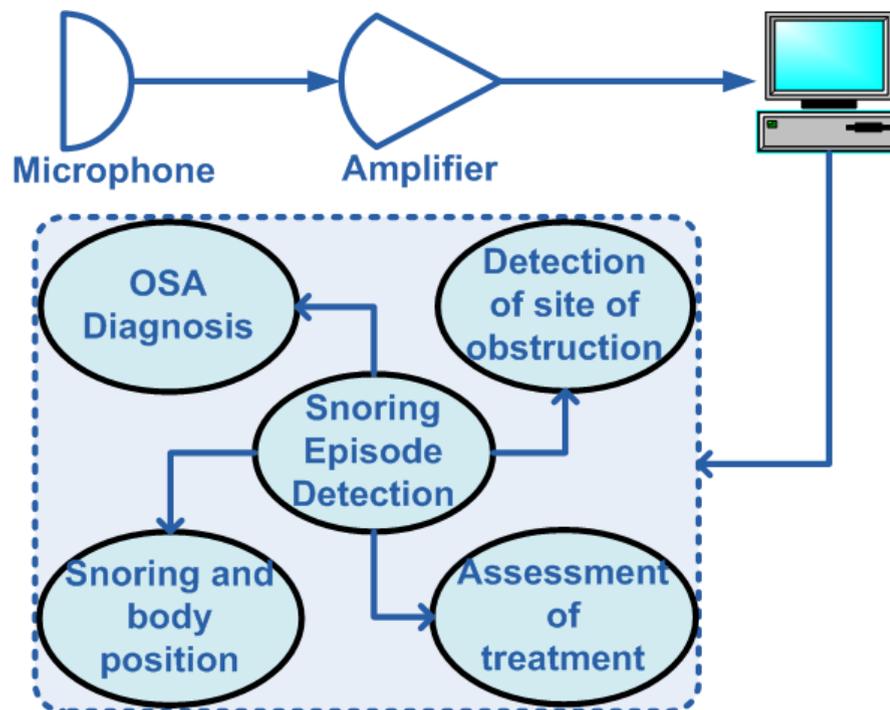


FIGURE 2.1: Snoring sound: Recording to clinical applications.

The respiratory sound signals were amplified with a gain of 200 and band-pass filtered with the cutoff frequencies of [0.5-5000 Hz] using Biopac (DA100C) amplifiers. The amplified signals were digitized at a sampling rate of 10240 Hz using NI9217 data acquisition module and a custom written LabView program. The bit resolution was 24. All recordings were performed simultaneously with the Polysomnography (PSG) at the Health Sciences Center Sleep Disorders Clinic (Winnipeg, Canada). Each recording usually took 8 hours. The study was approved by the Biomedical Research Ethics Board of the University of Manitoba and all participants gave written consent prior to data collection. Overall 68 people participated in this study.

2.3 Acoustical Characteristics of Snoring Sounds

2.3.1 Definition of Snoring Sound Segment

As known, the respiratory sounds of a snorer consist of breath, loud vibratory sounds, and/or small segments of silence [70]. The part of inspiratory and/or expiratory sound containing loud vibratory sounds (perceived as snore by human) is called snoring sound segment. The length of each segment may vary within and between subject.

2.3.2 Time and Frequency domain characteristics of Snoring Sounds

Figures 2.2 and 2.3 show two different snoring sounds with their power spectrum density (PSD) estimated using Welch [71] method. To estimate the PSD, windows of 100 ms and 50% overlap were used. As can be seen the two snoring sounds are quite different in both time and frequency domains. The snoring sound shown in Figure 2.2 has a repetitive structure in time and a flat spectrum (comb-like) below 500 Hz, while the snoring sound shown in Figure 2.3 has a noisy structure in time with 3-4 distinct peaks in its power spectrum. The dominant peaks are below 600 Hz.

There are several studies investigating the time and frequency characteristics of snoring sounds. In an early study [72], snoring sounds recorded from 10 non-apneic heavy snorers and 9 OSA patients were analyzed. They found that most of the

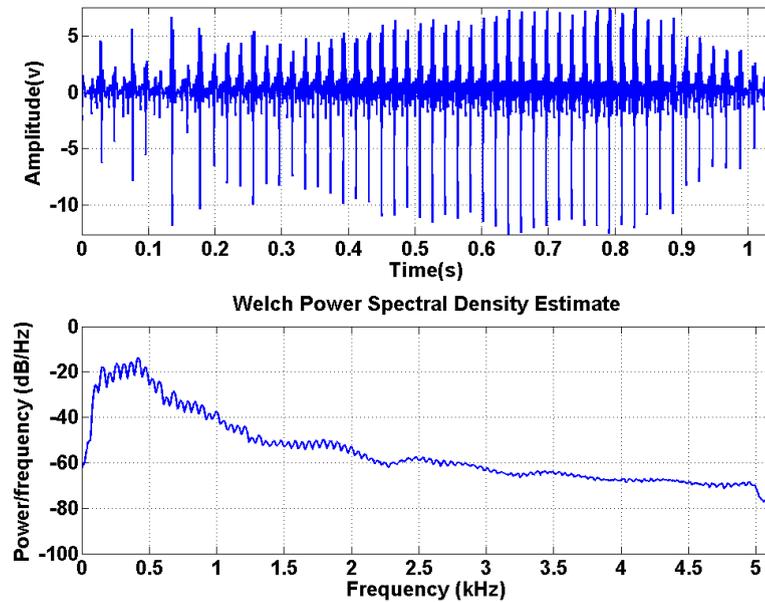


FIGURE 2.2: Snoring sound recorded by tracheal microphone from a heavy snorer (AHI=4.9)

power of snoring sounds was concentrated below 2000 Hz with a peak power below 500 Hz. The spectrum of nasal snoring sounds (when breathing through nose-only) was characterized by a series of discrete and sharp peaks with a fundamental note (similar to spectrum of a voiced sound). On the other hand the spectrum of the oronasal snoring sounds (when breathing through nose and mouth) had a mixture of sharp peaks and broad-band white noise [72].

The acoustical properties of snoring sounds in time-frequency domain were also investigated in another study [73]. They collected three sets of snoring sounds: 100 snoring sounds from 6 dogs (partial obstruction was created by implanting a balloon in the upper airway), 48 snoring sounds from 4 healthy participants while

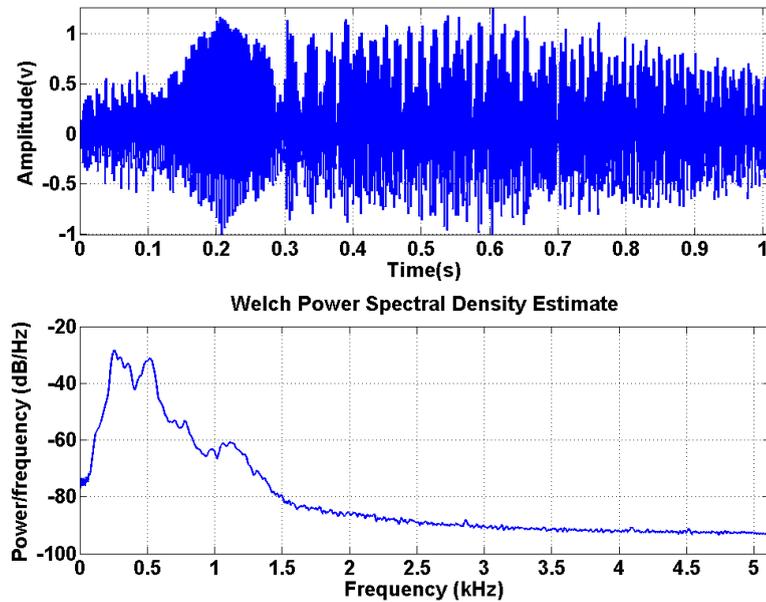


FIGURE 2.3: Snoring sound recorded by tracheal microphone from a heavy snorer (AHI=5.9)

simulating snoring sounds (24 snoring sounds by ambient microphone and 24 snoring sounds by tracheal microphones), and 400 snoring sounds from 9 non-apneic heavy snorers (all snoring sounds were recorded by an ambient microphone). The findings showed two dominant patterns of snoring sounds: simple and complex waveforms. The simple-waveform snoring sounds were characterized by a quasi-sinusoidal shape with a range of variants and only 1-3 peaks in their power spectrum (the first peak was the most prominent). On the other hand, the complex waveform snoring sounds were identified by a train of repetitive and equally-spaced sound structures with a comb-like power spectrum [73].

In another study, the snoring sounds of 17 male snorers were recorded and analyzed using power spectrum. Two different patterns in the sounds' spectra were

observed: first pattern the presence of a fundamental frequency with several harmonics and the second pattern consisted of a low frequency peak with the sound energy scattered on a narrower band of frequencies [74].

All the studies agreed on the presence of two type of snoring: one was characterized by a fundamental frequency and its harmonics (mainly happening in non-OSA snorers) and the other one was characterized by more complex frequency spectrum whose fundamental frequency and harmonics are difficult to be observed by a simple frequency analysis such as Fourier Transform [71].

Chapter 3

Snoring Sound Extraction From Respiratory Sound Recordings

As mentioned in Chapter 2, the average length of respiratory recording is approximately 8 hours with sampling frequency of 10 KHz. Hence, manual extraction of snoring sound segments is very time-consuming, and there is a need for an automatic snoring sound detection algorithm. In this chapter, an unsupervised snoring sound detection algorithm is introduced.

3.1 Existing Algorithms

There has been a limited number of studies on automatic detection and classification of snoring sounds. Energy and zero-crossing rate were used in [75] to identify the snoring segments from the recorded sounds by a microphone hung in the air. Energy and zero-crossing rate features are commonly used for silence and sound

segments identification with a reasonable accuracy but they are not discriminative enough for classification of snoring and breathing sounds segments [76, 77]. An automatic segmentation method was proposed in [78], in which the recorded sound by an ambient microphone was segmented into snoring segments, breathing, duvet noise, and silence periods using Hidden Markov Models (HMMs) and spectral-based features. The accuracy of detecting snoring segments was reported to be 82%-89% at a high computational cost [78].

In another study, the 500Hz sub-band energy distribution of the recorded sound signal was used for classification of the sound segments that were recorded by an ambient microphone [76]. The proposed algorithm was tested on 30 individuals (18 simple snorers and 12 patients with OSA). The accuracy of the method was reported to be 97.3% for simple snorers (when using only simple snorer data for training the classifier), 90.2% for simple snorers (when using both simple snorer and OSA data for training), and 86.8% for OSA patients. However, the accuracy of the method dropped significantly (from 97.3% to 86.8%) when the data set used for training was changed from simple snorer data to a combination of simple snorers and OSA patients data [76].

More recently, a classification method was proposed in [77] to classify snoring and breathing sounds. The number of zero crossings, energy of the signal, normalized autocorrelation coefficient at 1 ms delay and the first predictor coefficient of LPC analysis were deployed to label the snoring-related sounds into three classes: snoring (voiced non-silence), breathing (unvoiced non-silence), and silence segments.

The accuracy was reported to be 90.7% (the microphone was hung in the air). The accuracy increased when the conventional noise reduction methods were applied before the classification. In another study [79], zero-crossing rate, logarithm of the signal's energy, and first formant frequency were used to classify the breathing and snoring sound segments (recorded by tracheal and ambient microphones). The accuracy of that method was reported to be 95.7% and 93.2% for tracheal and ambient recordings, respectively.

3.2 Snoring Sound Detection Algorithm

We propose an unsupervised snoring classification algorithm, in which the learning procedure is based on Fuzzy C-means clustering. The algorithm is divided into two major parts: I) sound segmentation using Vertical-Box algorithm, and II) feature extraction using Principal Component Analysis (PCA) and classification of snoring and no snoring segments using Fuzzy C-means (FCM) clustering. The novelty of this method is its online automatic snoring sound extraction, high accuracy and low computational cost.

3.2.1 Data Recording

Data of 30 snorers (7 females) with an average age of 50.6 years ($STD = 9.96$) were used in this study. The snorers' anthropometric information for this part of study is shown in Table 3.1; seven of the patients were simple snorers ($AHI = 2.3 \pm 1.5$)

TABLE 3.1: Anthropometric information of participating individuals (BMI : Body Mass Index).

Group	Number of Patients	Age	BMI	AHI
OSA Patients	23(7 females)	49.9 ± 10.2	34.1 ± 7.2	26.1 ± 22.9
Simple Snorers	7 (no females)	53.1 ± 9.3	30.0 ± 3.8	2.3 ± 1.5

and the rest were diagnosed with OSA ($AHI = 26.1 \pm 22.9$) based on the PSG assessment.

As mentioned in Chapter 2, simultaneously with the PSG data recording, participants' respiratory sounds (including snoring and breathing sounds) were recorded by tracheal and ambient microphones. However in this chapter, only the results of tracheal recording are discussed, and the comparison between ambient and tracheal recording is discussed in Appendix A. A short period of the entire night recording (around 15 minutes, mostly when the patient was snoring) was randomly selected for the validation of the method; there were 5665 snoring sound segments in total. The sound segments were manually annotated as snoring or no-snoring segment by visual and auditory inspection of the spectrogram of each segment.

3.2.2 Signal Analysis

Signal analysis was performed in three steps: segmentation using a modified version of the vertical box (V-Box) control chart [80], feature extraction using PCA [81], and clustering using FCM [82, 83].

3.2.3 Modified Vertical box (V-Box) control chart

The sound signals were first band-pass filtered in the frequency range of 150-5000 Hz as the fundamental frequencies of the snoring sound are mostly below 5000 Hz [57, 84]. The 150 Hz low cutoff frequency was used to remove heart sound effect. The V-Box chart is a detection algorithm based on the concept of a moving vertically trimmed box along the time axis of the data; in each box (time window) it counts the number of past observations falling into a vertical box with the width L and height $2H$. In the original V-box algorithm [80], as long as the total count is above a certain threshold, no indicator is given. Once it goes below the threshold, a signal will be given indicating a change in the mean of a sequential process. The V-Box control chart was adopted for the sound signal segmentation purpose. The original algorithm was used for detecting change in the mean of a sequential process in a nonparametric framework [80]. The modified algorithm for breathing and snoring segmentation is summarized as the following steps:

1. At each time point $n \geq L$, a vertical box of the form $B(L, H, n, Y_n) = [n - L, n] \times [-H, H]$ is defined where Y_n is the last observation, n is the time point, $L > 1$ denotes the width of the V-Box(number of past observations to be taken into account), and $2H$ is the height of V-Box.
2. At any current time point, $n \geq L + 1$, the number of past observation falling into $B(L, H, n, Y_n)$, $b_{LH}(n)$, is counted.

3. If $b_{LH}(n) \leq L \times \theta$, mark a change in the process. The parameter $0 < \theta < 1$ should be chosen such that the fraction of observations located in the box $B(L, H, n, Y_n)$ equals $L \times \theta$.

The above algorithm used in this study is different from the original version from two aspects: a) the current observation, Y_n , determines the vertical position of V-Box (i.e. $B(L, H, n, Y_n) = [n - L, n][Y_n - H, Y_n + H]$), and b) in the original V-Box algorithm box moves forward when a new observation is obtained as it was designed for sequential real-time process. However, in this study, the data was not sequential; hence, the overlap between the successive boxes was set to be 80% (moving forward by 102 samples). Different events in the recorded sounds, such as breathing, snoring or biological noise (i.e. cough, talking, swallowing), were identified by the above algorithm. The sound segments separated by a time-lag less than 200 ms were merged. Figure 3.1 shows a typical 20s of a recording by the tracheal microphone and the resulted segments by the modified V-Box control chart algorithm. In order to have an accurate segmentation, the V-Box parameters needed to be tuned properly based on the sound signal amplitude and noise level. The parameters L, θ were not sensitive to the level of sound signal segments; hence, they were selected experimentally as: $L = 50ms, \theta = 0.95$ for all data. However, H was very sensitive to the level of breath and snoring sounds. Therefore, for every subject, H was adaptively tuned based on the sound signal's energy in each epoch.

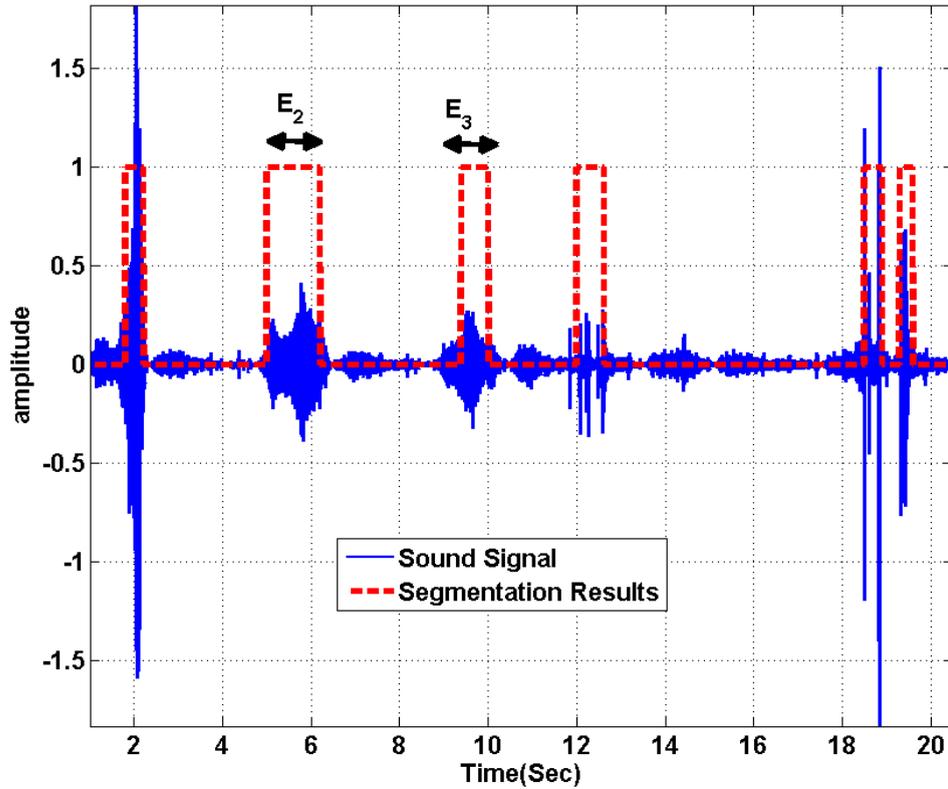


FIGURE 3.1: A typical 20s of a tracheal sound recording including breath and snoring segments and the segmentation result using V-Box algorithm. At this stage, all potential snoring sound segments are identified. E_i denotes segment i , where i is 1, 2, 3, \dots . For example, in this 20s interval, there are 6 segments.

Because the snoring and breathing sounds' levels are not constant over the entire night, in order to identify all potential snoring sound segments, H needs to be adjusted adaptively. Small H results in large detection error while large H affects the overall accuracy as some of the shallow snoring segment may not be detected. The initial value of H is set to zero for all data in order to identify all sound activities including snoring and breathing at the beginning. After clustering, there exist three clusters representing high, medium, and low amplitude sound segments.

The new value of H , H_{new} , can be updated as:

$$H_{new} = \frac{M_1 + M_2}{a} \quad (3.1)$$

where M_1 is the mean of absolute value of the medium amplitude sound segments and M_2 is the mean of absolute value of the high amplitude sound segments. a was selected experimentally as $a = 8$. It should also be mentioned that snoring clusters in the 2-D feature space are linearly concentrated around the horizontal axis.

3.2.4 Feature extraction

In this part of study, the same feature vectors as in [76] were used. However the novelty of our method compared to the method proposed in [76] is in its unsupervised classification and high accuracy without the need for training set and learning process. The 0-5000Hz frequency range was divided into 500Hz sub-bands (The first sub-band is actually between 150-500 Hz as the signals were previously band-pass filtered in the frequency range of 150-5000 Hz), and the average normalized energy in each sub-band for each detected sound segment by V-Box algorithm was calculated for further processing and feature extraction. To explain the feature extraction procedure, a typical signal is shown in Figure 3.1, in which 6 segments ($E_1 - E_6$) including both breathing and snoring segments have been identified. Let's assume we wanted to find the corresponding feature vector for the k^{th} segment (i.e. ε^k). First, the short-time Fourier transform (STFT)

of each segment was calculated using 50ms windows with 50% overlap between successive windows. Next step was to calculate the elements of feature vector ε^k , in which element i^{th} of the feature vector was computed as

$$\varepsilon_i^k = \frac{\sum_{j=1}^{N_k} \sum_{f=500(i-1)}^{500i} |Y^k(j, f)|^2}{\sum_{j=1}^{N_k} \sum_{f=0}^{5000} |Y^k(j, f)|^2}, \quad i = 1, \dots, 10 \quad (3.2)$$

where N_k is the total number of overlapping windows in the k^{th} segment and $Y^k(j, f)$ is the STFT of the j^{th} frame of the k^{th} segment. Then, PCA [81] was deployed to reduce the dimension of the feature space. The two largest eigenvalues were found to be much higher than the rest, resulting in a two-dimensional feature vector, $\hat{\varepsilon}^k$, for snoring and no-snoring classification. Hence, new feature vectors were computed in the following two steps:

1. The covariance matrix, COV, of feature vectors was calculated on the 60s intervals of the original recorded signal.

$$COV = \frac{1}{N-1} \sum_{k=1}^N (\varepsilon^k - \bar{\varepsilon})(\varepsilon^k - \bar{\varepsilon})^T \quad (3.3)$$

where, $\bar{\varepsilon}$ is the mean of snoring feature vector and N refers to the number of sound segments in each interval.

2. The two largest eigenvalues and the corresponding eigenvectors (the columns of matrix W) of the covariance matrix (COV) were calculated to form the new feature vector as

$$\hat{\varepsilon}^k = W^T \varepsilon^k \quad (3.4)$$

Figure 3.2 shows a projection of the 10-dimensional feature vector of a 200s segmented sound signal into 2-dimensional feature vectors.

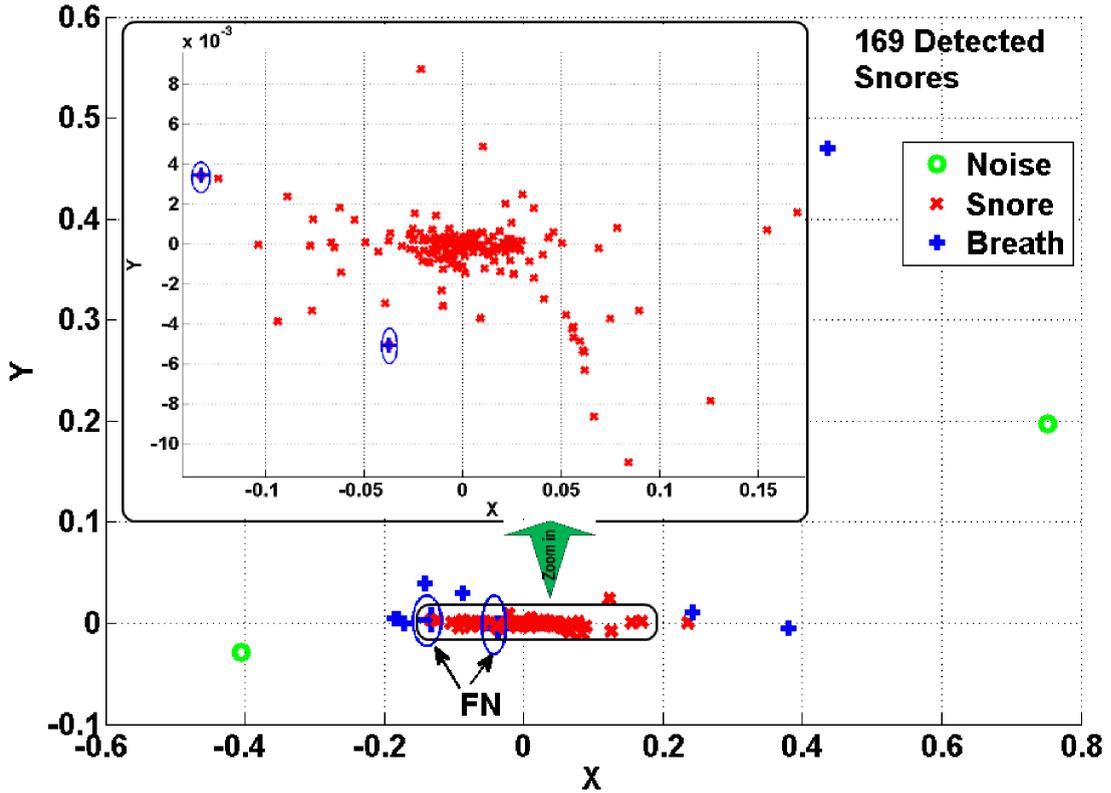


FIGURE 3.2: Clustered feature vector using FCM method. This 2-D feature vector was obtained using a projection from 10-D feature space onto a 2-D feature space by PCA. The original 10-D feature vector was obtained by calculating 500Hz energy band distribution from a segmented sound signal. The big rectangle is a zoomed in version of the snoring cluster located around horizontal axes.

3.2.5 Unsupervised Classification using Fuzzy C-means Clustering(FCM)

Once the feature vectors were formed, the FCM algorithm [85] was deployed to label each segment as snoring, breathing, or noise, in which the breath and

noise clusters were considered as no-snoring class. FCM algorithm was developed by Dunn [85], improved by Bezdek in [82, 83], and used extensively for different applications such as pattern recognition, data analysis, and image processing [86]. To obtain the clusters, the following objective function needs to be minimized:

$$\sum_{k=1}^N \sum_{j=1}^C u_{kj}^m \|\hat{\varepsilon}^k - c_j\|^2, \quad 1 \leq m < \infty \quad (3.5)$$

where u_{kj} is the degree of membership of $\hat{\varepsilon}^k$ in the cluster j , c_j is the two-dimensional center of the cluster j , m is a real number greater than 1, and C is the number of clusters ($C = 3$ in this study). Fuzzy clustering is an iterative optimization problem, in which the aforementioned objective function is being minimized and in each iteration the degree of membership, u_{kj} , and cluster centers, c_j , will be updated. The update values are calculated as

$$u_{kj} = \frac{1}{\sum_{i=1}^C \left[\frac{\|\hat{\varepsilon}^k - c_j\|}{\|\hat{\varepsilon}^k - c_i\|} \right]^{\frac{2}{m-1}}} \quad (3.6)$$

$$c_j = \frac{\sum_{k=1}^N u_{kj}^m \hat{\varepsilon}^k}{\sum_{k=1}^N u_{kj}^m} \quad (3.7)$$

The iteration will stop if

$$\max_{kj} |u_{kj}^{(i+1)} - u_{kj}^{(i)}| < \epsilon \quad (3.8)$$

where $0 < \epsilon < 1$ is a termination criterion and i is the iteration step.

3.3 Results and Evaluation of the snoring sound detection algorithm

The proposed method was used to extract the snoring sound segments. Figure 3.2 shows the clustered sound segments (feature vector) using FCM method.

As shown in Figure 3.2, the snoring cluster is concentrated around the horizontal axis in a linear shape. This property was seen in almost all subjects, implying that the 500Hz sub-band energy of the snoring sounds is consistent for all snoring events during a 60-second interval. Figure 3.3 shows the segmentation result of the same sample sound, shown in Figure 3.1, after clustering and snoring sound extraction; note that all undesired sound activities are removed from the originally segmented signal. The performance of the detection algorithm was evaluated in terms of accuracy and positive predictive value (PPV). The results are shown in terms of True Positive (TP), False Negative (FN), False Positive (FP), and overall accuracy and PPV. The accuracy and PPV were calculated as

$$Accuracy = \frac{TP}{TP + FN} \times 100 \quad (3.9)$$

$$PPV = \frac{TP}{TP + FP} \times 100 \quad (3.10)$$

Table 3.2 shows the classification result for a sample subject. The feature space of this subject is shown in Figure 3.2. Table 3.3 shows the results of classification using tracheal recording. The results of classification for the ambient recording

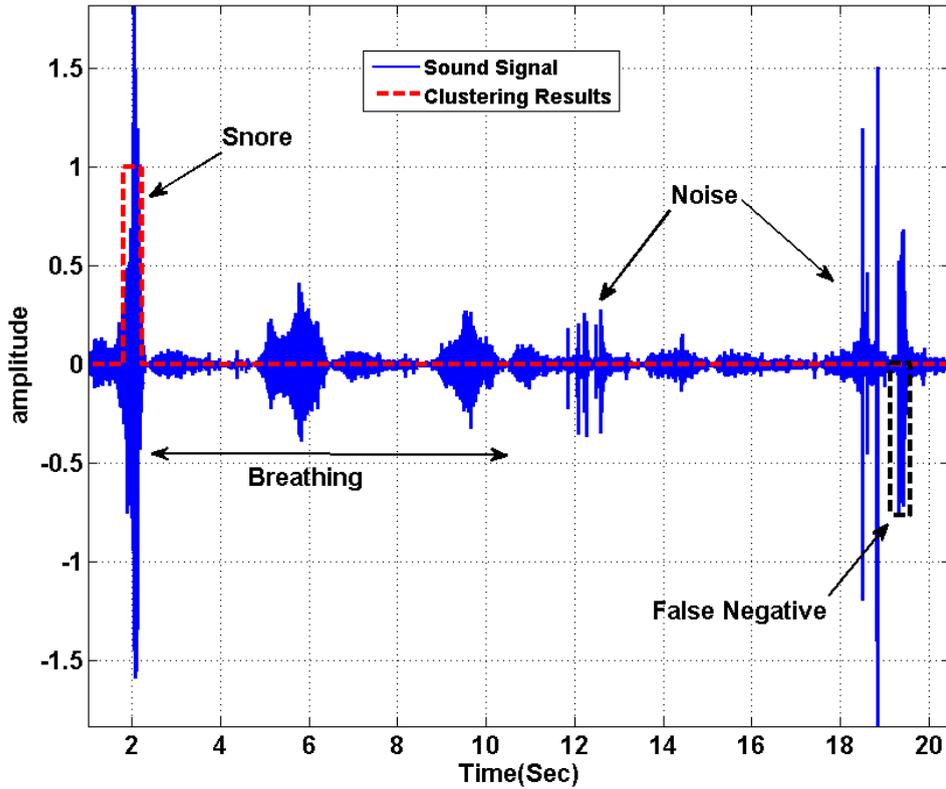


FIGURE 3.3: Successful removal of noise and breathing segments after clustering. Note that snoring segment E_6 was not detected and therefore there is a false alarm in this case.

TABLE 3.2: Classification Results and total number of breaths for a sample case with 171 snoring segments.

Data Set	TP	FP	FN	Number of Breaths	Number of Noises
Sample case	169	16	2	324	2

are presented in Appendix A. As can be seen, the overall accuracy and PPV of the proposed snoring sound detection algorithm is 98.6% and 94.8%. The performance of the algorithm remained more or less the same when it was applied only to data of either OSA patients (98.8%) or simple snorers (98.4%).

TABLE 3.3: Classification results using tracheal recordings.

Data Set	TP	FP	FN	Accuracy	PPV
Simple Snorers and OSA	5588	304	77	98.6	94.8
OSA	3816	203	48	98.8	94.9
Simple Snorers	1772	101	29	98.4	94.6

TABLE 3.4: Classification Results for different values of overlap ($L = 50ms$, $\theta = 0.95$).

Overlap(%)	Accuracy (%)	PPV (%)
50	97.4	94.4
70	97.5	93.8
80	98.6	94.8

TABLE 3.5: Classification Results for different values of L (Overlap=80%, $\theta = 0.95$).

L(ms)	Accuracy (%)	PPV (%)
25	98	93.3
50	98.6	94.8
100	97.5	94.5

3.4 The effect of segmentation parameters

To investigate the effect of the three L , θ , and overlap parameters on the accuracy of the algorithm, data of both simple snorers and OSA patients recorded by the tracheal microphone were used. As shown in Tables 3.4 and 3.5, the variation of the accuracy and PPV was low with respect to changes of L and overlap parameters. In addition, as shown in Table 3.6, the accuracy increased for higher values of θ up to an optimal point at $\theta = 0.95$. On the other hand, PPV decreased slightly as θ increased (Table 3.6).

TABLE 3.6: Classification Results for different values of θ (Overlap=80%, $L = 50ms$).

θ	Accuracy (%)	PPV (%)
0.7	88.3	96.9
0.75	92.3	96.6
0.8	94.7	95.8
0.85	95.8	95.7
0.9	96.6	95
0.98	97.6	92.5

3.5 Computational Complexity

The computational cost of our proposed algorithm is relatively low. This measure was derived for a typical respiratory sound (S_r) with length N . The algorithm consists of three parts: Segmentation, Feature extraction, and Clustering. For the segmentation part, there exists a window with length L which moves along the signal with a certain overlap. Therefore, the total number of overlapping windows would be approximately $\frac{N}{L-Overlap}$. In each window the number of observation falling into the VBox will be calculated. This takes $2L$ basic operations. The total number of operations to calculate $b_{LH}(n)$ is:

$$2L \times \frac{N}{L - 0.8L} = 10N \quad (3.11)$$

$b_{LH}(n)$ will be compared with a threshold for all time points (n). Therefore the number of operations for segmentation would be:

$$10N + 2N = 12N \quad (3.12)$$

Assume that the algorithm detects K sound segments each with length N_s . Note that the lengths of segments are not necessarily equal but for simplicity, we assume all segments have the same length (N_s). The first step in feature extraction is to calculate the STFT (window length= L_s , overlap= $0.5L_s$, $L_s = 50ms$) of each segment. The computational cost for STFT is as below:

$$2 \frac{N_s}{L_s} \times L_s \log_2 L_s = 18N_s \quad (3.13)$$

To calculate ε_i^j , we need the following number of basic operations:

$$2 \frac{N_s}{L_s} \times 500 \times 18N_s + 2 \frac{N_s}{L_s} \times 5000 \times 18N_s = \frac{1.98 \times 10^5}{L_s} N_s^2 \quad (3.14)$$

As mentioned before, ε_i^j should be calculated for all segments and 10 frequency sub bands. Therefore, the total cost for feature extraction would be:

$$10K \times \frac{1.98 \times 10^5}{L_s} N_s^2 \quad (3.15)$$

We can assume that the length of each segment is between $0.2s$ and $3s$ ($0.2F_s < N_s < 3F_s$). Therefore the computational cost of feature extraction has an order of $O(K)$. For the PCA calculation, we need to compute the covariance matrix and eigenvalues of a 10D feature vector with the following number of basic operations:

$$(10 + 1 + 100)K + 100 = 111K + 100 \quad (3.16)$$

Finally, we have a 2D feature vector for clustering in which the computational cost is negligible but has an order of $O(K)$. Therefore the total cost would be:

$$12N + 10K \times \frac{1.98 \times 10^5}{L_s} N_s^2 + 111K + 100 \quad (3.17)$$

Consequently, the total computational cost has an order of $O(N)$ assuming $K \propto N$.

3.6 Discussion and Concluding Remarks on Snoring Sound Detection Algorithm

A new unsupervised algorithm for snoring sound detection from a record of breath and snoring sounds was developed and tested on respiratory sound data of 30 patients. Given that, 5665 snoring segments were extracted from the data of all subjects, and that the number of predictors was two (snoring vs. no-snoring), the statistical tests have a sufficiently high power to hold the reliability of the results. It is worth noting that total number of breathing segments is a bit higher than that of snoring segments.

In most of the previous studies on automatic snoring sound detection [75–78], the recordings were done by a microphone hung in the air in the vicinity of the patient (the ambient microphone). When using an ambient microphone, breath sounds are usually not recorded, and mostly loud snoring sounds are recorded. Hence, identifying the snoring segments in data recorded by an ambient microphone would be much easier than that in data recorded by a tracheal microphone. This is

because classification of data recorded by an ambient microphone would reduce to classify snoring vs. silence, while it is snoring versus breathing in case of data recorded by a tracheal microphone.

In this part of study, snoring sound detection and extraction were investigated when the recorded data was collected by a tracheal microphone and included snoring sound, breathing sound and other noises. The proposed algorithm was also applied to the data recorded by an ambient microphone for the sake of comparison with previous studies (Appendix A).

In the method proposed in [76], the accuracy of the classification was reported to be 97.3% when only simple snorer data was used for training and this value dropped to 86.6% when only OSA data was used for training. In contrast, the results of the proposed method in this chapter show a very small variation in the accuracy of the classification ($\leq 0.4\%$ in tracheal microphone and $\leq 4.8\%$ in ambient microphone) in case of using data of either group of patients, indicating robustness and insensitivity of the proposed method to the severity associated with AHI.

The algorithms used in [75–78] were based on supervised classification with specific parameters, which needed to be tuned for each subject. In contrast, the proposed method in this study is based on an unsupervised clustering algorithm, and the parameters are adjusted adaptively. The effect of segmentation parameters including V-Box width (L) and overlap between successive windows for a fixed θ was negligible (Table 3.5), which implies the algorithm is robust and not

sensitive to the width and overlap of the windows. On the other hand, for a fixed L and overlap parameter, the accuracy and PPV showed an opposite trend with respect to variation of θ (Table 3.6). However the variation in the PPV for $\theta > 0.8$ was very minute; hence, θ was selected such that accuracy was maximized.

The main advantages of the proposed algorithm are its high accuracy, robustness and insensitivity to AHI, unsupervised operation, and low computational cost. It should be emphasized that an accurate snoring sound extraction method is essential for diagnosis and treatment of different snoring-related disorders such as OSA. Such method can also be used as pre-processing tool in a variety of studies such as those that investigate the relationship between sleep stages and snoring sound characteristics, the relationship between body position and snoring sound intensity, and identify the person's best sleeping position, in which the least number of snoring segments occur.

Chapter 4

Higher Order Statistical Properties of Snoring Sounds

In Chapter 3, an automatic snoring sound detection algorithm was introduced. The extracted snoring sound segments can be used to obtain important clinical information about the upper airway state over night.

4.1 Why Higher Order Statistics of Snoring Sounds?

As mentioned in Chapter 1, different tasks such as investigation of obstruction in the upper airway [87, 88], assessment of the outcome of surgical treatment [44, 45, 89], and classification of snorers as simple snorer or OSA patients [67, 90, 91] utilize acoustical analysis of snoring sounds. Most of the signal processing techniques used for the acoustical analysis, such as autocorrelation/autocovariance function [70, 90], power spectrum density (PSD) [73, 92], and autoregressive

(AR) modeling [90, 92] are based on the assumption that a linear model can represent snoring sound and also that the signal-generating process is Gaussian. Furthermore, these 2nd order statistical techniques do not analyze the information contained in the signal's phase. (Note that, henceforth, the terms linear, nonlinear, Gaussian, and non-Gaussian signal/segment are being used interchangeably with linear, non-linear, Gaussian, and non-Gaussian signal-generating process).

If the signal of interest, i.e. snoring sounds, violates the above assumptions, one should take into account a complementary technique. Higher order statistical (HOS) techniques reveal information on not only amplitude of a signal, but also its phase. Furthermore, if a non-Gaussian signal is received along with additive Gaussian noise, a transformation to higher order cumulants domain would be blind to the noise; hence, achieving a cleaner estimate in noisy environments. Thus, the bispectrum and bicoherence can efficiently reveal and quantify any nonlinear relationship among the harmonic peaks (such as phase coupling) [93]. In fact, the basic property used in higher order spectral analysis is that for a stationary Gaussian process, all cumulants of order larger than two are zero. It is worth noting that HOS analysis has been used as a tool for screening OSA among snorers [91, 94].

Anthropometric parameters such as gender, height, BMI, and AHI are hypothesized to change the properties of snoring sounds. If this hypothesis is true then in order to use snoring sounds to develop a technique for classification anthropometric parameters need to be matched. In this chapter, first, the Gaussianity and

linearity assumptions of snoring sounds are investigated. Second, the relationship between snorers' anthropometric parameters and characteristics of snoring sounds are reviewed. Third, the result of an acoustical task (OSA detection) with anthropometrically matched and unmatched groups is given. Since many researchers have used the snoring sound features for OSA screening, it is important to investigate whether the same features are also sensitive to anthropometric parameters, i.e. weight, height, gender, etc., among different people. It is important because if they are sensitive, then it implies the classification accuracy might be partly due to unmatched groups in terms of anthropometric parameters.

Two features called Median Bifrequency (MBF) and Projected Median Bifrequency (PMBF), and several conventional features such as skewness, kurtosis, 1st formant frequency, and energy of the snoring sound segments were calculated. Then, the statistical relationship between these features and anthropometric parameters was investigated using Kendall's Tau-b test [95] and Kruskal-Wallis Analysis of Variance (KWAV) [96, 97]. Lastly, a Naïve Bayes classifier [98] was run on the features to examine the feasibility of OSA screening using both HOS and common spectral features. For this part of the study, two subsets of the participants were selected: one with matched anthropometric parameters between the OSA and simple snorer groups and the other one with unmatched anthropometric parameter between the two groups.

4.2 Method

4.2.1 Data Recording

Data of 57 individuals (15 females, 51.5 ± 11.8 y), who were snorers, were selected for this study. The participants' anthropometric information of this study is shown in Table 4.1. The AHI value of each participant was determined by the PSG study scored by the sleep lab technicians. As known, the respiratory sounds of a snorer consist of breath, loud vibratory sounds (perceived as snore by humans), and/or small segments of silence [70]. We call the part of respiratory sound containing snore (or loud vibratory sounds) as snoring sound segment. The length of each snoring sound segment varies within and between the subjects.

The algorithm proposed in Chapter 3 was used to extract the snoring sound segments in a semi-automated manner [99]. An example of the selection method is as the following: the PSG data provided information about the time (e.g. 3:00-3:45 am) when the patient X was snoring (snoring interval). Given this information the snore detection algorithm proposed in [99] was run on the interval. Although the method's accuracy was over 98%, to ensure 100% accuracy of snoring sound segments, all of the detected snoring sound segments were validated by visual (spectrogram) and auditory means, and the misclassified cases were removed.

Most (99%) of the extracted snoring sounds occurred during inspiration. The extracted snoring sound segments for each patient were used to estimate the bispectrum and derive the desired features. Figure 4.1 shows an extracted snoring

TABLE 4.1: Anthropometric Information of Participating Individuals.

Group	Number of subjects	Age	BMI	AHI	Height
OSA	42 (9 females)	52.2 ± 12.4	33.9 ± 6.7	35.5 ± 33.1	173.9 ± 10.5
Simple Snorers	15 (6 females)	49.5 ± 10.4	30.1 ± 4.0	2.4 ± 1.3	168.4 ± 12.1

sound segment along with its spectrogram. It should be noted that the bispectral analysis was only performed on the snoring sound segments.

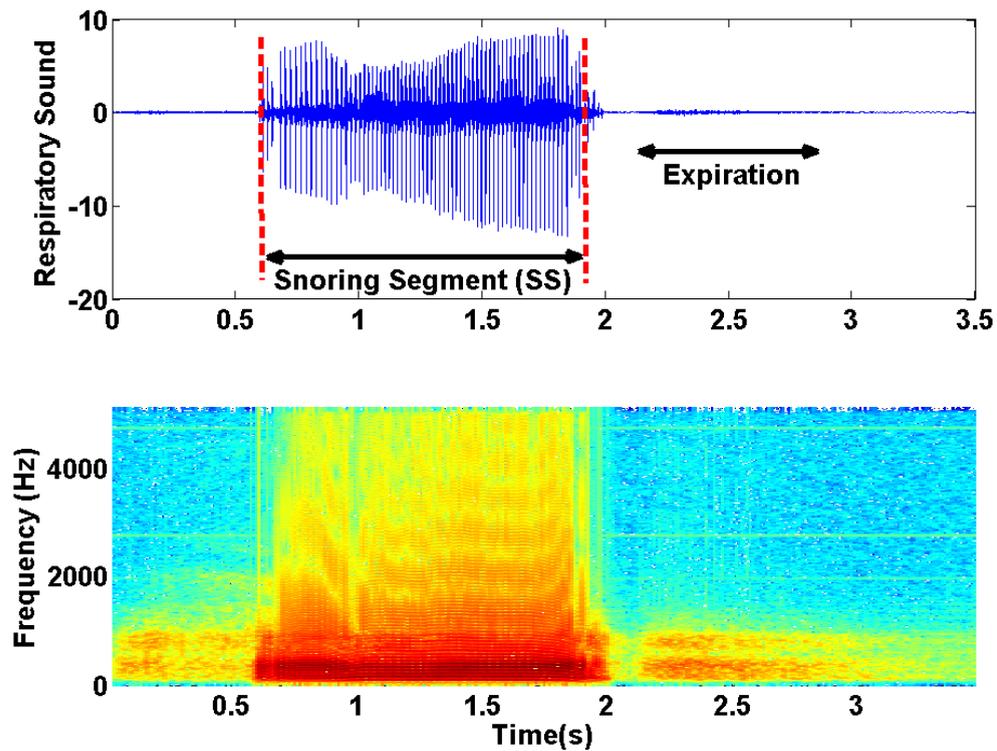


FIGURE 4.1: An extracted snoring sound segment and its spectrogram.

4.2.2 Higher Order Statistics (HOS)

Assume that $s(n)$ is an extracted snoring sound segment (in general a random process). The key assumption underlying the HOS analysis is that the process $s(n)$ is stationary in some sense [93]. Snoring sounds are non-stationary in nature

[92]. Hence, all the HOS measures such as bispectrum and bicoherence should be calculated on a short time-windowed version of the signal to ensure stationarity (wide-sense) of the snoring sound segments.

4.2.2.1 Definition of Bispectrum and Bicoherence

The 2nd and 3rd order cumulants of a zero-mean stationary process are defined as:

$$c_2(k) = E \{s^*(n) s(n+k)\} \quad (4.1)$$

$$c_3(k, l) = E \{s^*(n) s(n+k) s(n+l)\} \quad (4.2)$$

where $s(n)$ is a zero mean stationary process, k , l , and m are different time increments, $*$ refers to complex conjugate operator, and c_2 , c_3 , c_4 denote 2nd, 3rd, and 4th order cumulants, respectively [100]. The 2nd and 3rd order polyspectrum are defined as the Fourier Transform of c_2 and c_3 , respectively [100]:

$$P(f) = \sum_{k=-\infty}^{+\infty} c_2(k) e^{-j2\pi f k} \quad (4.3)$$

$$B(f_1, f_2) = \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} c_3(k, l) e^{-j2\pi f_1 k} e^{-j2\pi f_2 l} \quad (4.4)$$

where $P(f)$, $B(f_1, f_2)$ represent the PSD and bispectrum, respectively. Note that the PSD is real valued, nonnegative, and a function of one variable (discrete frequency, f). On the other hand, the bispectrum is a function of two variables (discrete bifrequencies, f_1 , f_2) and has complex values. Bicoherence is another

useful statistical measure, which is defined as [93]:

$$bic(f_1, f_2) = \frac{B(f_1, f_2)}{\sqrt{P(f_1 + f_2)P(f_1)P(f_2)}} \quad (4.5)$$

A linear and stationary random process, $s(n)$, can be represented as the output of a linear system (impulse response: $h(n)$) excited by an independent and identically distributed (*iid*) noise, $e(n)$. The power spectrum and bispectrum of the output ($s(n)$) can be simplified as:

$$P(f) = \sigma_e^2 |H(f)|^2 \quad (4.6)$$

$$B(f_1, f_2) = \mu_3 H(f_1) H(f_2) H^*(f_1 + f_2) \quad (4.7)$$

and therefore bicoherence will be constant as:

$$bic^2(f_1, f_2) = \frac{\mu_3^2}{\sigma_e^6} \quad (4.8)$$

where $e(n)$ is an *iid* noise with zero mean, variance σ_e^2 , and 3rd moment μ_3 ($\mu_3 = E\{e^3(n)\}$) and $H(f)$ is the Fourier Transform of $h(n)$ [101].

Equation (4.8) shows that for any linear signal, squared bicoherence is constant and independent of the bifrequencies (f_1, f_2) . If the squared bicoherence is zero, signal $s(n)$ is Gaussian or non-skewed with a symmetric distribution because μ_3 or equivalently skewness is also zero [93, 101].

We used the Hinich's method to test for non-skewness (loosely called Gaussianity)

and linearity of the snoring sound segments [101]. First, we tested the Gaussianity hypothesis (H: the bispectrum is zero). If H holds the process is Gaussian; hence, signal generating process is linear. Otherwise, the process is non-Gaussian and needs to be tested for its linearity. If the bicoherence is constant, the process is linear, otherwise, it is nonlinear. This procedure was repeated for all snoring sound segments and all body positions.

4.2.2.2 Bispectrum and bicoherence estimation

In practice, the number of sound samples is finite; hence, the HOS measures need to be estimated from available data. The bispectrum of the signal $s(n)$, can be estimated using *direct* or *indirect* approaches. In this part of study, the direct approach [93], which is an extension of the Welch technique for power spectrum density estimation, was used to estimate the bispectrum in the following steps:

1. The signal $s(n)$, $n = 0, \dots, N - 1$, is divided into K overlapping segments, each of length M . Let the k^{th} segment of $s(n)$ be $s_k(n)$, $n = 0, \dots, M - 1$. It is worth noting that for snoring sound analysis, we used 100 ms time windows with 50% overlap to ensure the stationarity assumption.
2. Calculate the zero-mean signal of each segment as:

$$s_{zk}(n) = s_k(n) - \frac{1}{M} \sum_{n=0}^{M-1} s_k(n) \quad (4.9)$$

3. Multiply the zero-mean signal by the Hanning window, $w(n)$, to control the effect of spectral leakage.

$$s_{wk}(n) = s_{zk}(n) w(n) \quad (4.10)$$

4. Compute the discrete Fourier transform (DFT) of each segment:

$$X_k(l) = \frac{1}{M} \sum_{n=0}^{M-1} s_{wk}(n) e^{-j\frac{2\pi nl}{M}} \quad (4.11)$$

The raw bispectral estimate ($\hat{B}_k(l, m)$) can be calculated as:

$$\hat{B}_k(l, m) = X_k(l) X_k(m) X_k^*(l+m) \quad (4.12)$$

where l, m are the discrete frequencies.

5. The consistent estimate of bispectrum ($\hat{B}(l, m)$) can be obtained by averaging raw estimates over all segments.

$$\hat{B}(l, m) = \frac{1}{K} \sum_{k=0}^{K-1} \hat{B}_k(l, m) \quad (4.13)$$

Consequently, the squared bicoherence can be derived from bispectrum as below [102]:

$$\widehat{bic}^2(l, m) = \frac{|\hat{B}(l, m)|^2}{\hat{P}(l) \hat{P}(m) \hat{P}(l+m)} \quad (4.14)$$

The discrete bispectrum has many symmetries in (l, m) plane. It is only needed to calculate $\hat{B}(l, m)$ in the non-redundant region or principal domain (D) which is defined as: $D = \{0 < f_1 \leq \frac{f_s}{2}, 0 < f_2 \leq f_1, 2f_1 + f_2 \leq f_s\}$ [103]. Where (f_1, f_2) are the bifrequencies (in Hertz) correspondent to the normalized bifrequencies (l, m) .

4.2.3 Feature Extraction

Suppose that we estimated the bispectrum ($\hat{B}(f_1, f_2)$) in D . This section details on deriving two new features defined in Section 4.1: 1) The Median Bifrequency (MBF) feature, which is a 2-D feature denoted as (f_1^{mp}, f_2^{mp}) , and 2) PMBF, which is a 1-D feature denoted as f^p .

4.2.3.1 Median Bifrequency (MBF) computation

MBF is the bifrequency where the L_1 norm [104] of $\hat{B}(f_1, f_2)$ becomes half of the L_1 norm of $\hat{B}(f_1, f_2)$ over all bifrequencies in D . The procedure is detailed in the following steps:

1. Calculate the summation of $|\hat{B}(f_1, f_2)|$ at all bifrequencies in D .

$$B_T = \sum_{f_1} \sum_{f_2} |\hat{B}(f_1, f_2)|, \quad f_1, f_2 \in D \quad (4.15)$$

2. Set $f_1 = 0$.

3. For all bifrequencies (f_1, f_2) satisfying the condition $\{0 < f_2 \leq f_1, 2f_1 + f_2 \leq f_s\}$ calculate:

$$SB(f_1, f_2) = \sum_{f_1} \sum_{f_2} |\hat{B}(f_1, f_2)| \quad (4.16)$$

4. Check if $SB(f_1, f_2) \geq \frac{1}{2} B_T$

If **YES**, end the algorithm and $(f_1^{mp}, f_2^{mp}) = (f_1, f_2)$ If **NO**, increase f_1 and go to step 3. (Note that: $f_1^{max} = \frac{f_s}{2}$.)

4.2.3.2 PMBF computation

Once the MBF is computed, the PMBF, f^p , can be determined by the projection of (f_1^{mp}, f_2^{mp}) onto the line $\{f_2 = f_1, f_1, f_2 \in D\}$ corresponding to the diagonal slice of the bispectrum. Equivalently we have:

$$f^p = \frac{f_1^{mp} + f_2^{mp}}{2} \quad (4.17)$$

4.2.3.3 Skewness and Kurtosis

Let $s(n)$ be a zero-mean random process. Skewness (γ_1) and kurtosis (γ_2) are defined as:

$$\gamma_1 = \frac{c_3(0, 0)}{\sigma_s^3} \quad (4.18)$$

$$\gamma_2 = \frac{c_4(0, 0, 0)}{\sigma_s^4} \quad (4.19)$$

where σ_s is the standard deviation of $s(n)$ and $c_3(0, 0)$ and $c_4(0, 0, 0)$ are its zero-lag 3rd and 4th order cumulants respectively [105].

4.2.3.4 1st formant frequency and energy

Energy (E) and first formant frequency (F_1) were obtained from each snoring sound segment. Linear predictive coding (LPC) [71] was used to estimate F_1 . To meet stationarity assumption, $s(n)$ was divided into 100 ms overlapping frames (50% overlap and Hanning window). In each frame, the autoregressive (AR) model of the signal was estimated and the roots of AR model were calculated. To select the AR model order, we used the optimum order model (optimum order = $f_s(KHz) + \gamma$, $\gamma = 4, 5$ & $f_s \in [6 - 18]KHz$) suggested in [106]. Therefore, we selected an AR model of 14 to estimate first formant frequency of each frame. F_1 was estimated by taking median over all frames.

4.2.3.5 Calculation of features

As mentioned in Section 4.2.1, the number of snoring sound segments is different for each patient. Let us denote i^{th} snoring sound segment of patient X by s_i^X , $i = 1, \dots, I^X$. First, $(f_1^{mp}, f_2^{mp})_i^X, (f^p)_i^X, (\gamma_1)_i^X, (\gamma_2)_i^X, (E)_i^X, (F_1)_i^X$ for all segment were calculated resulting in a finite number of observations for each feature. Then, the sample median of each feature set was estimated. This procedure was repeated for all 57 individuals yielding a 57×2 matrix for MBF feature and a 57×1 vector for PMBF, energy, 1st formant frequency, skewness, and kurtosis

features. The reason we used median instead of mean is the insensitivity of median to outliers; it is known that when the data is not symmetrically distributed, the median outperforms the mean in measuring the middle range of data [107]. Figure 4.2 shows the sample density of $(f_1^{mp})^X$, $(f_2^{mp})^X$, and $(f^p)^X$ estimated by kernel method [108]. As shown in Figure 4.2, the data is skewed; therefore, the median is a better estimate of the middle range of the data than the mean value in this case.

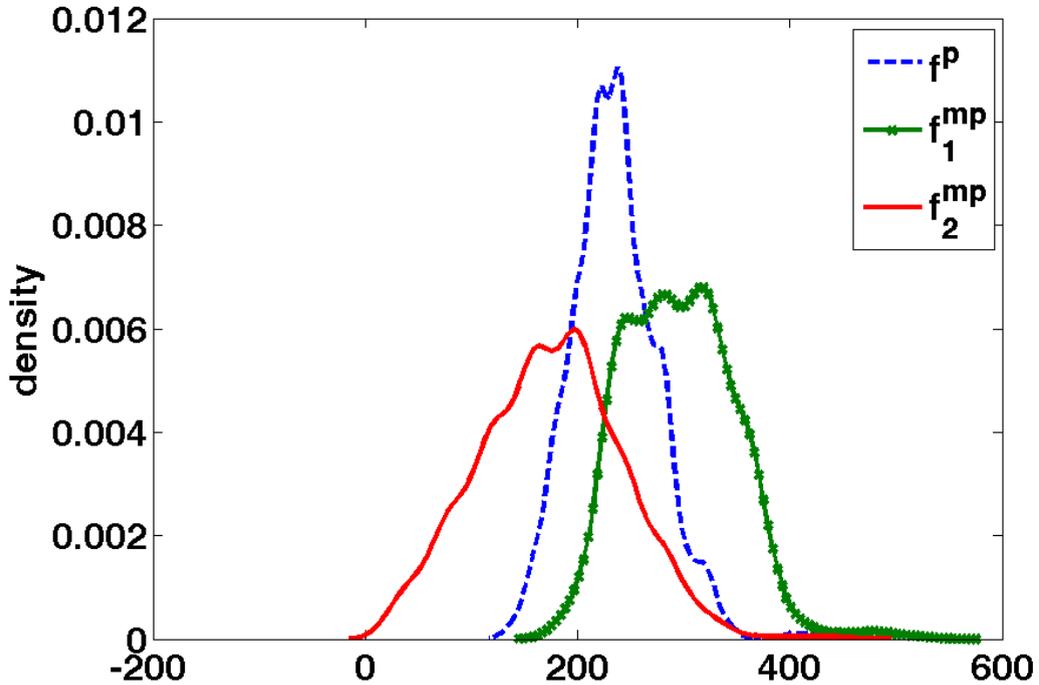


FIGURE 4.2: Kernel density estimate of f^p (projected median bifrequency), f_1^{mp} (bifrequency along horizontal axes), and f_2^{mp} (bifrequency along vertical axes) for a typical subject. Note the asymmetry of the distributions.

4.2.4 Statistical Analysis

To investigate the effect of anthropometric parameters such as age, gender, height, BMI, and AHI on the features, we ran statistical tests assuming the significance level as $p = 0.05$. Since the distribution of the features deviated from normal distribution, the Kendall's Tau-b test (nonparametric counterpart of Pearson correlation) [95] was used to measure the correlation among continuous anthropometric parameters and HOS features. The one-way KWAV [96, 97] was also used to compare the median of features between men and women.

4.2.5 Classification

Naïve Bayes classifier [98] was used to evaluate the ability of our feature set to discriminate the subjects to snorers with OSA and snorers without OSA or the so called "simple snorer" groups. Particularly, we were interested to compare the ability of snoring sound features to be used as a signature of OSA when the groups of OSA and non-OSA were matched (Experiment A) and unmatched (Experiment B) in terms of anthropometric parameters.

Therefore, we performed two experiments: **Experiment A:** We selected a subset of our database including 22 apneic and 6 simple snorers that were matched in terms of gender, BMI, height and AHI. **Experiment B:** Another subset with the same number of participants (28 including 21 apneic and 7 simple snorers) with unmatched anthropometric parameters was used for classification. Table 4.2 shows the anthropometric profile of both experiments.

TABLE 4.2: Anthropometric information of two subsets selected for classification. Experiment A: The OSA and apneic groups were matched for gender, BMI, and height parameters. Experiment B: The OSA and apneic groups were not matched for gender, BMI, and height parameters.

Group	# of subjects	Age	BMI	Height	AHI
Experiment A					
OSA	22(no female)	47.2± 11.4	33.5± 6	176.4±2.4	36.5±35
Simple Snorers	6(no female)	50.6± 5.8	33.8± 5	176.6±3.2	1.8±1.3
Experiment B					
OSA	21(7 females)	49.4± 10.6	34.7±7.3	173.9±12.2	27.8±23.3
Simple Snorers	7(no female)	53.1± 9.3	30±3.8	178.3±7.6	2.3±1.5

We used the energy, 1st formant frequency, MBF, PMBF, skewness, and kurtosis as our features for linear discriminant analysis. Several combinations of the features were examined and the performance was evaluated using the Leave-One-Out Cross-Validation (LOOCV) technique [109, 110]. The LOOCV is a common technique when the number of observations (subjects in this case) is relatively small; it helps to prevent over-fitting. In the LOOCV, one observation is used as testing set and the rest (27) is used as training set. This procedure is repeated for all observations (28) and the average performance is measured. It is worth noting that the Euclidean metric was used to compute the distance.

4.3 Results and Discussion

All snoring sound segments were found to be non-Gaussian, while their linearity varied during the night. In fact, for each snorer, there existed some linear snoring sound segments as well as some non-linear ones. We also noticed that the linearity of the snoring sound segments varied among different body positions within each

TABLE 4.3: The results of Kendall's Tau-b and Kruskal-Wallis tests for five anthropometric parameters. ns:non-significant, *: significant at level 5%, **: significant at level 1%.

Features	H	BMI	Age	AHI	Gender
γ_1	ns	ns	ns	ns	ns
γ_2	ns	ns	ns	ns	ns
f^p	**	*	ns	*	*
f_1^{mp}	*	*	ns	*	ns
f_2^{mp}	*	ns	ns	ns	*
E	ns	ns	ns	**	*
F_1	ns	*	ns	ns	ns

subject. However, this result was not consistent among all subjects. Furthermore, not everybody slept and/or snored in all positions. It was shown that the body position during sleep changes both duration and intensity of snoring sounds [19]. However, we did not find a consistent relationship between the sleeping position and the change in the linearity of snoring sound generating process.

It is known that if a signal is non-Gaussian, the 2^{nd} order statistical techniques are only able to extract partial information from the signal [93, 94]. Therefore, we used HOS measures to develop new features such as MBF and PMBF from existing data. We also extracted common HOS features such as skewness and kurtosis from the snoring sound segments. It was found that there was a significant relationship between frequency based features such as f_1^{mp} , f_2^{mp} , and f^p and all anthropometric parameters except age. As shown in Table 4.3, four out of five anthropometric parameters (height, BMI, AHI, and gender) significantly affected the HOS features of the snoring sound segments, while gender and BMI were significant parameters affecting energy and first formant features.

The height of individuals was observed to be a significant factor influencing the value of f^p ($p < 0.01$), f_1^{mp} , and f_2^{mp} ($p < 0.05$). There was a negative relationship between height and these frequency related features. The taller the individuals, the lower frequency components were in their snoring bispectrum. The height has been shown to affect the tracheal sound spectral features [111]. It was reported that the tracheal sounds in children had higher frequency components than in healthy adults. In another study [112], it is shown that the anatomy of the trachea determines the characteristic features of tracheal sounds. However, there was no study confirming the change in the features of snoring sound segments due to the height. Based on our findings, the MBF and PMBF features of the extracted snoring sound segments are negatively related to the height of individuals. Assuming that taller individuals have taller neck, this result implies that the characteristics of snoring sound segments reflect resonances (existing in snoring sound) that depend on the upper airway's length.

The results of the Kendall's Tau-b test on BMI groups shows that BMI is significantly associated with the value of f^p ($p < 0.05$), f_1^{mp} ($p < 0.01$) as well as $F_1(p < 0.05)$. As known, obesity is a factor strongly associated with the presence of OSA [113]. Obese individuals with sleep apnea have been shown to have more (about 42%) fat in their cervical region than normal subjects as well as non-obese individuals with OSA [114]; thus, resulting in pharyngeal area narrowing. It is also known that higher BMI is associated with increased level of leptin (a hormone

produced by the adipose tissue and has also actions on the respiratory centre control) [115]. Therefore, our observed changes in the acoustical properties of the snoring sound segments due to BMI can be explained by both anatomical and hormonal changes of the upper airway.

It was also found that AHI and gender were significantly correlated with energy and frequency-based HOS features of the snoring sound segments. As shown in Table 4.3, the individuals with higher AHI had lower frequency-based features (f^p and f_1^{mp}) ($p < 0.05$) and higher energy ($p < 0.01$). The female snorers of this study were observed to have higher frequency-based features (f^p and f_2^{mp}) ($p < 0.05$) and lower energy feature ($p < 0.05$) than the male snorers. Although there was no study investigating the gender effect on the snoring sounds, this observation is congruent with findings reported in two studies focused on breath and lung sounds [116, 117]. According to those studies, breath and lung sounds in healthy women contain higher frequency components than in men. It has also been shown that men have higher pharyngeal and supraglottic resistances than women [118]. Therefore, given that the size and mechanical properties of pharynx are significantly different between men and women [119], the snoring sounds of women and men can be expected to be significantly different as the results of our study indicate. Moreover, these might be also a reason for greater incidence of OSA in men [118, 119].

Two of the frequency-based HOS features (f^p and f_1^{mp}) were found to be significantly different in snorers with different AHI. This result is congruent with

previous studies. In people with OSA, the lateral pharyngeal muscular wall is usually narrower [120]. Therefore, minimum area of the airway has been shown to be significantly smaller in apneic individuals than non-OSA people. The size of airway plays a major role in the frequency components of the sound produced by the flow turbulence in the airway. This explains the change in the frequency based HOS feature of the snoring sound segments between snorers with OSA and simple snorers.

It was observed that for some of the anthropometric parameters (e.g. BMI and gender) two of the frequency-based features (f_1^{mp} , f_2^{mp} , and f^p) were significantly correlated, while the third one was not significantly correlated. f^p is linearly related to the summation of f_1^{mp} and f_2^{mp} (projection of two). If both have a significant correlation with a parameter, then we expect that f^p would be also significant (as in the case with Height parameter) but having one of them significantly correlated with a parameter, does not necessarily lead to a significant correlation of f^p and that parameter. The reason that only one of the coordinates of MBF is significant depends on the bispectrum of the snoring sound segment. As an example, let us compare MBF for 4 snoring sound segments of participant 4 (P4) and participant 6 (P6). P4(BMI=24.4): $f_1^{mp}=[320, 320, 360, 320]$ and $f_2^{mp}=[240,40,160,160]$. P6(BMI=47.1): $f_1^{mp}=[240, 240, 200, 250]$ and $f_2^{mp}=[180,160,180,120]$. It is clear that BMI significantly changed f_1^{mp} but not f_2^{mp} . In fact, the difference between the bispectral information of the two sets of snoring sound segments is well extracted using the 2-D MBF feature which is an advantage of bispectral analysis.

One important point is that these frequency changes due to small changes in the airway size may not always be detectable by spectral analysis of the sounds. However, as known, HOS techniques complement the information obtained from 2nd order statistical techniques, i.e. power spectral analysis. Hence, we propose using a combination of HOS techniques and conventional acoustical techniques increases the diagnosis accuracy of OSA. In this work, we tried to verify this point by applying a simple classifier to our feature set. We partitioned our database into two sets to compare two scenarios, one when the height, gender, and BMI are matched between the two groups of snorers with OSA and simple snorers, and the other one when those parameters are not matched. We observed an increase in the accuracy of classification when the parameters were matched.

Table 4.4 illustrates the results of classification. Several combinations of features were used as input to the naïve Bayes classifier. Results demonstrate that the highest sensitivity and specificity occurred when a combination of both conventional feature (Energy) and HOS feature set (f_1^{mp} and skewness) was used. This combination resulted in sensitivity of 93.2% (87.5%) and the specificity of 88.4% (86.3%) for experiment A (B). As shown in Table 4.4, for experiment A, the sensitivity and specificity values for only HOS features were 75.9-94.1% and 74.6-81.9%, respectively. On the other hand, using only energy and formant frequency resulted in a sensitivity and specificity of 78.2% and 72.1%, respectively.

As expected, overall, the sensitivity and specificity decreased when an unmatched subset was used for classification.

TABLE 4.4: Naïve Bayes classification results for different combination of conventional and HOS features.

Feature set	Sensitivity (%)		Specificity (%)	
	Experiment A	Experiment B	Experiment A	Experiment B
f^p, f_1^{mp}, γ_2	77.2	75.2	80.3	79.1
f^p, f_1^{mp}, γ_1	84.2	81.3	74.6	78.6
f^p, γ_1, γ_2	75.9	77	81.9	87.3
$f_1^{mp}, \gamma_1, \gamma_2$	94.1	91.2	74.6	71.2
f^p, γ_1, E	85.9	82.1	88.1	84.4
f_1^{mp}, γ_1, E	93.2	87.5	88.4	86.3
E, F_1	78.2	80.5	72.1	65.3

To compare our work with a recently published work [91], we matched the anthropometric parameters of snorers with OSA and simple snorer groups. Moreover, our recordings were performed using a microphone placed over trachea. Therefore, our recorded sounds have a higher signal to noise ratio than those recorded by an ambient microphone. We also improved the sensitivity and specificity of OSA diagnosis among snorers by simultaneous usage of HOS and conventional features. However a major difficulty in our study was to find a larger population with matched anthropometric parameters to validate the results of our analysis.

4.4 Conclusion

In this chapter, the relationship between anthropometric parameters and the 3rd and 4th order statistical features derived from the snoring sound segments were investigated. In summary, we investigated statistical correlation of these features along with the zero-lag HOS features with different anthropometric parameters. An important contribution of the statistical investigation is on the application of

snoring sound for OSA identification among snorers. Since the common features of snoring sounds used in classification are sensitive to anthropometric parameters, the results of classification may change when the two groups of apneic and controls are matched for those parameters.

Chapter 5

Inter- and Intra-subject Variability of Snoring Sounds

In Chapter 4, the result of classification between non-OSA and OSA (AHI>5) groups was shown. We also investigated to what extent anthropometric parameters may change the results if they were not matched between the groups. It was shown that the classification accuracy slightly increased when the aforementioned anthropometric parameters were matched. In a follow up study we became interested to investigate how variable the snoring sounds are not only between the groups but also within a snorer over a single night.

This chapter reports on sequential variability of snoring sounds during sleep, and also on classification of OSA patients into 4 groups with a nearly-matched anthropometric parameters. Moreover, instead of using the sound features directly, the variability of features (both inter- and intra-subject variabilities) was used to

estimate the severity of OSA. The results were compared to those presented in Chapter 4.

As known, mild reduction in airflow is usually associated with snoring, while more marked reduction in airflow is associated with partially obstructed airways resulting in hypopneas and heavy snoring. Complete obstruction (zero flow) is called apnea [6]; obviously, there is no snoring during apnea. We hypothesize that the flow limitation (or airway obstruction) is reflected on the snoring sounds characteristics differently during non-apneic, hypopneic, and post-apneic snoring.

Snoring sounds have been used in several studies for the diagnosis of OSA [65, 67, 70, 91, 94, 121, 122]. These studies used both linear and nonlinear acoustical techniques to analyze the snoring sounds. Snoring sound intensity [49], power spectral measures and formant frequency [121, 122], and pitch detection [65, 67] are based on classical linear techniques. On the other hand, phase coupling extraction [70, 94], and nonlinear mode interaction [91] techniques are based on a nonlinear model of snoring sounds. None of the above studies have considered intra-subject variability of snoring sounds. The number of participants in these studies was at most 40 snorers but more importantly, the number of snoring sound segments was less than 40 segments per subject (as in [91, 121]) which does not capture the intra-subject variability.

In this part of the study, we extracted the snoring sound segments of non-OSA and OSA snorers, who went through full-night sleep assessment by PSG. Using the PSG score sheet, we marked the snoring sound segments as non-apneic, hypopneic, and

post-apneic classes. Then, we calculated the second order and non-linear features of the segments, and ran a non-parametric statistical test (Kruskal-Wallis analysis of variance) to characterize the difference among the three classes to investigate the above hypothesis.

We hypothesized that the variation of snoring sounds over time is associated with Apnea-Hypopnea Index (AHI), and can be used as a classification tool. Therefore, we measured the variation of snoring sounds' features using their total variation norm [123]; then used the regression analysis to fit a linear function on the total variation norm and the AHI values of participants. Finally, the linear discriminant analysis and leave-one-out cross validation were used to classify the participants into four classes: non-OSA ($AHI < 5$), mild OSA ($5 < AHI < 15$), moderate OSA ($15 < AHI < 30$), and severe OSA ($AHI > 30$).

The contributions of this part of the thesis compared to previous studies are: comparing and characterizing the snoring sound segments among non-apneic, hypopneic, and post-apneic classes (using several features), investigating the variability of snoring sound segments within each class (using total variation analysis), and investigating how the variability of snoring sound segments within each individual reflects on his/her OSA severity (using regression analysis).

TABLE 5.1: Anthropometric Information of the study participants.

Group	Number of subjects	Age	BMI	AHI
Non-OSA	15 (6 females)	49.5±10.4	30.1± 4.0	2.3±1.3
Mild OSA	13 (4 females)	49.8±13.9	30.2± 3.1	8.6±2.2
Moderate OSA	15 (3 females)	52.8±14.2	34.8± 6.4	22.7±4.7
Severe OSA	14 (2 females)	53.8± 8.7	36.3± 8.2	74.2±30.1

5.1 Method

5.1.1 Data

The snoring sounds of 57 snorers (15 females, 51.5 ± 11.8 y) were used for this part of study. Out of the 57 participants, 13 were diagnosed with mild OSA, 15 with moderate OSA and 14 with severe OSA, and 15 were non-OSA snorers. The PSG study was scored by the trained sleep lab technicians and approved by the referring physician. AHI, Age and BMI variations of the participants are shown in Table 5.1. It should be noted that only tracheal recordings were deployed in this part of study.

5.1.2 Snoring sound extraction from respiratory sounds

Figure 5.2 shows a 60s record of respiratory sound with an example of snoring sound segment. The length of each snoring sound segment varies within and between the snorers. This variability can be observed between the snoring sound segments shown in Figures 5.2 to 5.4, (the snoring sound segments were extracted using the automatic snore detection algorithm proposed in [99]).

To ensure that all snoring sound segments were identified correctly in terms of the definition of snore and their start and end times, all the detected segments were also validated by visual and auditory means in the time-frequency domain. The occasional misclassified cases were removed from the database. It is worth noting that the number of misclassified cases was 572 which form about 2.4% of the total snoring sound segments. Using the PSG score sheet of each participant, all snoring sound segments were marked as post-apneic, hypopneic and non-apneic. Note that each of these classes may have included snoring sound segments of a few consecutive breathing phases; however, the post-apneic class had snoring sound segments of at most 3 consecutive breathing phases. It should also be noted that labeling the snoring sound segments was solely done for the first part of study (investigating variability within a snorer using non-parametric test). For the second part, the regression and classification analysis were blind to the label of the classes. Details of proposed algorithm are illustrated in Figure 5.1.

5.1.3 Feature Extraction

All the snoring sound segments were passed through a band-pass filter with the frequency range of 150-2000 Hz (to remove the effect of heart sounds and high frequency noises) as most of the power of snoring sounds is concentrated below 2000 Hz [72]. The power spectrum density (PSD) of the snoring sounds segments were calculated using Welch method [71] in windows (Hanning) of 100 ms with 50% overlap. The following features were extracted from each snoring sound segment:

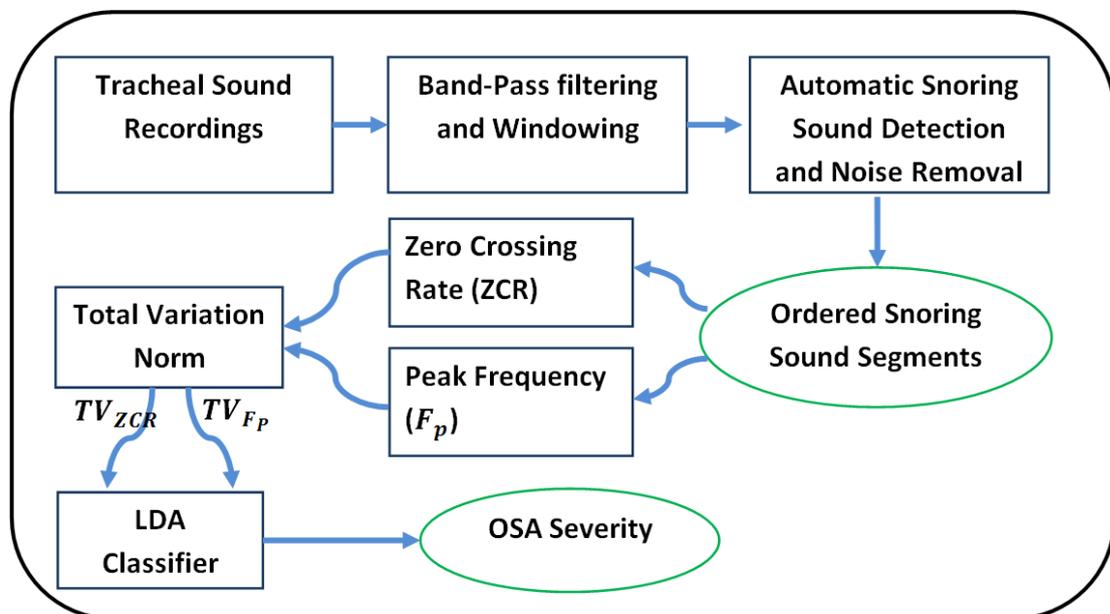


FIGURE 5.1: Detailed flow chart of proposed algorithm.

1) average power (P_{avg}), 2) Zero Crossing Rate (ZCR) which is usually used to categorize the speech [124] and also snoring sound [67] into voiced and unvoiced categories, 3) the frequency of the spectral peak with the lowest frequency (F_0), 4) the frequency of the peak with maximum power (F_p), and 5) the spectral entropy (SE) [125] which is a measure of flatness of PSD and can be calculated as:

$$SE = - \sum_f P_n(f) \cdot \ln(P_n(f)) \quad (5.1)$$

where $P_n(f)$ refers to normalized PSD at discrete frequency f .

5.1.4 Statistical Analysis

Our first hypothesis was that the snoring sounds characteristics of an OSA snorer are different among three classes of non-apneic, hypopneic, and post-apneic. To

investigate this hypothesis, the Kruskal-Wallis nonparametric test [96, 97] (a nonparametric counterpart of one-way analysis of variance) was used to compare the median of snoring sound features of each class of non-apneic, hypopneic and post-apneic episodes for each participant's snoring sound data. The variation of snoring sound segments was also characterized segment by segment.

5.1.5 Regression and Classification Analysis

We also hypothesized that the variability of snoring sounds over time that occur during hypopnea and/or after apnea is high, while for non-apneic class is low. Assume two snorers: snorer 1 with only non-apneic class of snoring (low AHI), and snorer 2 with non-apneic, hypopneic, and post-apneic classes of snoring sound segments (high AHI). We expect the total variability of snoring sound features for snorer 1 would be less than that of snorer 2. As mentioned in the previous section, the analysis for this part of study is without the knowledge of snoring sounds' class label. To investigate this hypothesis we calculated the total variation norm (TV_i) [123] of features for each individual as the following:

$$TV_i = \frac{1}{N} \sum_{k=1}^{N-1} |f_i(k+1) - f_i(k)|, \quad i = 1, \dots, F \quad (5.2)$$

where f_i is one of the extracted features and TV_i is the total variation norm corresponding to that feature. F denotes total number of features and N represents the total number of snoring sound segments for each snorer. We treated the variation of each feature as a new feature and used all TV_i s for our classification.

We hypothesize that the total variation norm of some of the sound features to increase as AHI increases and vice versa. In fact, the more variability leads to higher amounts of total variation norm of some snoring sound characteristics. To investigate this hypothesis, we calculated the strength of association between total variation norm and AHI using nonparametric correlation analysis. We also fitted a linear regression function to AHI-TV pairs (for the feature giving highest correlation). Then, TV norms were used to classify the snorers into 4 classes of non-OSA snorer, mild OSA, moderate OSA and severe OSA. Linear discriminant analysis [126], and leave-one-out cross validation technique [109] were used for classification.

5.2 Results

5.2.1 The difference between snoring sounds of three classes within an individual

As shown in Table 5.2, only 1.8% of the snoring sounds were in the post-apneic class, whereas 30.8% of the snoring sounds occurred during hypopnea and 67.4% of snoring sound segments were during the time with no indication of apnea or hypopnea. This is partially due to the fact that the majority of people referred to PSG are either non-OSA or with mild OSA, thus few apneic events

Figures 5.2 to 5.4 show three 1-minute epochs of snoring sounds for different classes: a record of non-apneic snoring sounds (Figure 5.2), a record of hypopneic

TABLE 5.2: Number of extracted snoring sound segments from tracheal recordings.

Type of Snoring		Number of Snoring Segments
	Non-Apneic	15504 (67.4%)
	Hypopneic	7080 (30.8%)
	First Post-Apneic	208 (0.9%)
Post Apneic	Second Post-Apneic	154 (0.7%)
	Third Post-Apneic	47 (0.2%)

snoring sounds (Figure 5.3), and an epoch containing apneic events and post-apneic snoring sound (Figure 5.4). We suggest the sequence of snoring sound segments occurrence over time can predict the state of upper airway in terms of obstruction (or severity of flow limitation in the upper airway).

Tables 5.3 and 5.4 demonstrate the result of Kruskal-Wallis nonparametric test among the possible three classes of post-apneic, hypopneic, and non-apneic snoring within each individual. Some snorers had only two classes of snoring sounds while there were few of them with only one class of snoring sound (we exclude them for this part of study). Table 5.3 shows the comparison among non-OSA individuals whereas Table 5.4 demonstrates the results among patients with OSA (including mild, moderate, and severe OSA).

As shown in Tables 5.3 and 5.4, there is a significant difference between some of the characteristic features of snoring sound segments among the three classes. For instance, consider subject 49 (AHI=81.9) in OSA group with 130 non-apneic, 85 hypopneic, and 22 post-apneic snoring sound segments. For this patient, the median of P_{avg} , ZCR , F_p were significantly different among the three classes ($p < 0.00001$). This can be observed in both groups of OSA and non-OSA individuals

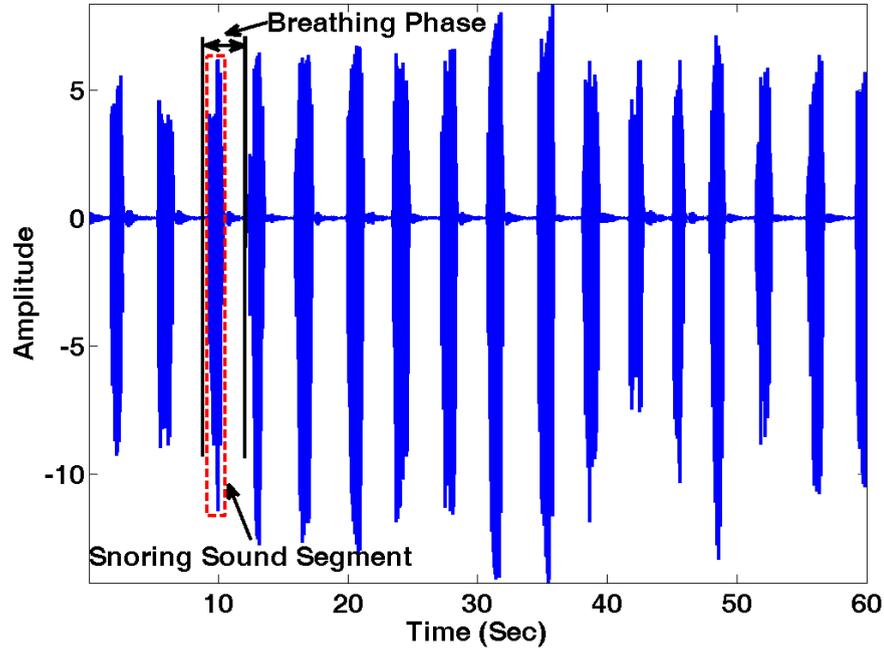


FIGURE 5.2: A 1-minute snoring epoch when there is no hypoapnea event. This type of snoring is called non-apneic snoring in which the characteristics of snoring sounds do not change significantly from snore to snore. A typical snoring sound segment is shown with dashed rectangle and a breathing phase is illustrated with solid parallel lines.

(at least for some sound features).

5.2.2 The total variation norm and classification

In addition to the difference between median of snoring sounds' features among the snoring sound classes of a snorer, the results show that the variability of the snoring sound features within each class is also different between the classes. For example, as shown in Figure 5.5, for an OSA snorer, the variation of F_p in hypopneic and post-apneic classes was high (even between two neighboring snoring segments), while it was negligible for non-apneic class of snoring (of the same subject). As a result, regardless of the label of each class (i.e. non-apneic, hypopneic, and

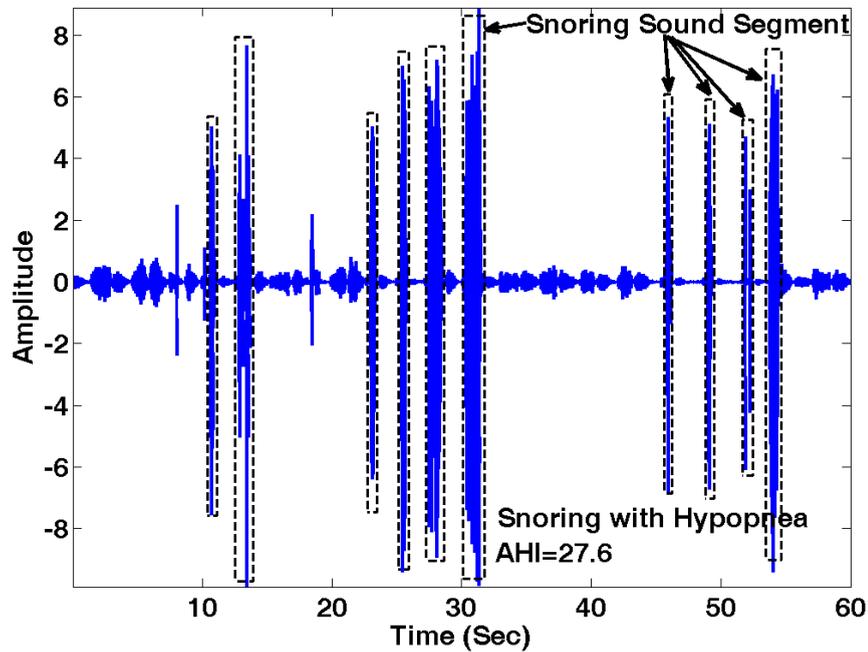


FIGURE 5.3: A 1-minute snoring epoch when there are hypopneic events. This type of snoring is called hypopneic snoring in which the characteristics of snoring sounds change drastically from snore to snore. It can be clearly observed that even the duration of snoring sounds are highly variable within one minute of recording. The snoring sound segments are marked with dashed rectangles.

post-apneic), the total variation norm would be large for this participant. Now let us consider a non-OSA snorer with one class of snoring sounds (i.e. non-apneic): because the variation among consecutive snoring segments is low, the total variation norm would be small. Hence, without knowing the label of each class, this feature is able to estimate the severity of OSA.

Table 5.5 shows the correlation values between AHI and total variation for all features. As shown in Table 5.5, TV_{Fp} and TV_{ZCR} had the strongest correlation with AHI. We ran a regression analysis on (AHI, TV_{Fp}) and (AHI, TV_{ZCR}) pairs for 57 participants. Table 5.6 shows the result of linear regression analysis.

The accuracy of 4-class (non-OSA, mild OSA, moderate OSA, and severe OSA)

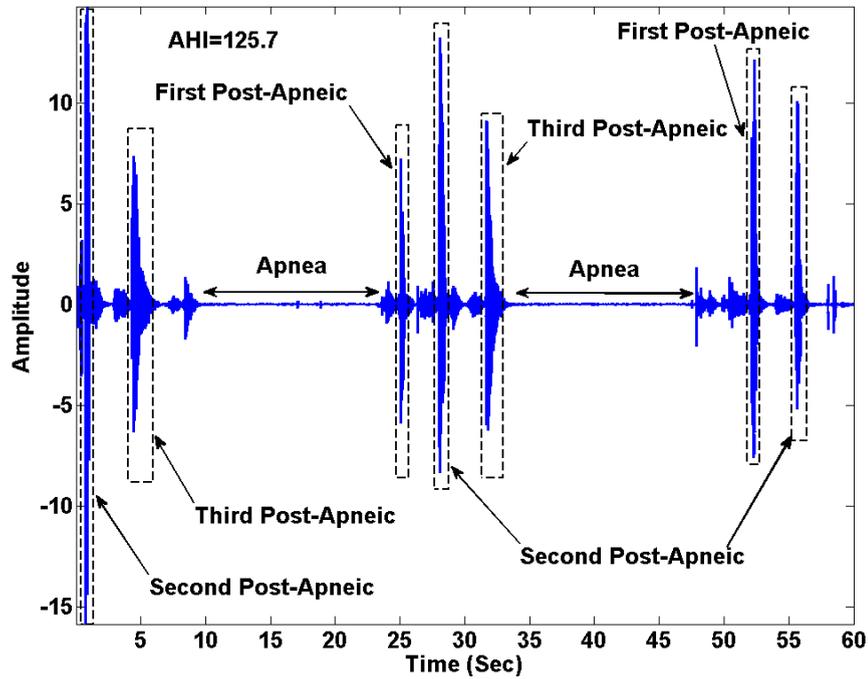


FIGURE 5.4: A 1-minute snoring epoch when there are apneic events. This type of snoring is called post-apneic snoring in which the first sound usually appears as a very loud snoring. Depending on the severity of OSA, an apneic event may be accompanied by more snoring segments with different characteristics. In this Figure, there are two apneic events each 13s long.

classification problem was 68.4% (64.9%) using only TV_{F_p} (only TV_{ZCR}). We combined TV_{F_p} and TV_{ZCR} to classify the individuals to the aforementioned four classes of snorers. Table 5.7 shows the result of leave-one-out cross validation with linear discriminant classification. As can be calculated from Table 5.7, the accuracy for 4-class classification is about 77.2%. However, if the goal is set to classify only non-OSA ($AHI < 5$) and OSA ($AHI > 5$) snorers, the results are 92.9% sensitivity, 100% specificity and 96.4% accuracy. Or if the goal is to detect people with moderate and severe OSA and refer them for PSG or further sleep study, the sensitivity and specificity would be 87% and 96%, respectively.

TABLE 5.3: Result of Kruskal-Wallis test on non-OSA group within each individual. Only 6 out of 15 non-OSA participants had more than one class of snoring sounds. NA#, HA#, PA# are the number of non-Apneic, hypopneic, and post-apneic snoring sounds, respectively. ns: not significant, and *: significant ($0.01 < p < 0.05$), **: very significant ($0.0001 < p < 0.01$) and ***: highly significant ($p < 0.0001$).

No.	AHI	NA#	HA#	PA#	P_{avg}	ZCR	SE	F_0	F_p
19	0.8	153	39		ns	**	ns	**	**
4	1.1	295	45		**	***	***	ns	***
45	2.8	220	102		ns	***	**	ns	ns
47	3	338	95		ns	ns	ns	***	**
54	3.4	98	56		**	*	**	ns	*
25	3.5	193	80	13	ns	***	***	***	***

Several studies attempted to use snoring sounds to diagnose OSA but there has always been lack of a standard framework to compare different techniques. Table 5.8, however, summarizes some of the techniques in terms of number of participants, location of microphone, features, AHI threshold, and accuracy.

5.3 Discussion

The results show that the non-apneic, hypopneic, and post-apneic snoring sound segments are significantly different; since the number of snoring segments in this study was large, the statistical results are reliable as they have enough statistical power. This shows that snoring sounds are associated with the level of upper airway obstruction (or airflow). In fact, during non-apneic snoring, there is minor variation in airflow between consecutive respiratory cycles. However, during hypopnea the airflow changes significantly and influences the snoring sounds' characteristics.

TABLE 5.4: Result of Kruskal-Wallis test on OSA group within each individual.
9 patients had only one class of snoring and therefore they were excluded.

No.	AHI	OSA severity	NA#	HA#	PA#	P_{avg}	ZCR	SE	F ₀	F _p
12	29	Moderate	202	123		ns	ns	*	ns	ns
53	20.8	Moderate	160	86		ns	**	ns	***	***
3	6.3	Mild	385	89	7	ns	**	***	***	**
22	27.6	Moderate	362	195		*	***	***	***	**
34	121.4	Severe		152	36	**	**	ns	ns	*
50	22.6	Moderate	125	188	27	*	***	ns	ns	***
48	15.2	Moderate	321	138		**	ns	***	***	ns
56	49.1	Severe	129	140	5	**	ns	***	**	***
8	77.5	Severe	139	241	46	**	*	***	ns	ns
52	55.7	Severe	111	112	11	***	**	ns	***	***
16	21.8	Moderate	223	250		***	ns	***	***	ns
20	63.5	Severe		98	55	***	ns	ns	***	ns
2	29.5	Moderate	324	113	2	***	ns	ns	ns	ns
39	125.7	Severe	98	150	29	***	ns	ns	ns	ns
33	34	Severe	349	100		***	***	***	***	ns
13	29.2	Moderate	358	232		***	ns	***	**	***
17	22.1	Moderate	301	120		***	***	***	***	***
40	89	Severe	153	98	70	***	***	***	ns	***
32	8.9	Mild	379	161		***	*	***	***	***
42	83.4	Severe	275	132	81	***	***	***	*	***
43	23.3	Moderate	334	101		***	***	***	*	***
11	33.1	Severe	278	127	5	***	ns	***	***	ns
29	5.9	Mild	164	51		***	***	***	***	ns
10	9.4	Mild	128	75		***	ns	***	**	**
9	7.1	Mild	274	150		***	ns	ns	ns	ns
31	13.9	Mild	278	111		***	*	***	ns	ns
28	37.2	Severe	388	149		***	***	***	**	***
6	21.8	Moderate	346	173		***	***	***	**	***
30	9.7	Mild	336	57		***	***	***	*	***
51	7.8	Mild	139	74		***	***	***	***	***
49	81.9	Severe	130	85	22	***	***	ns	ns	***
26	18.4	Moderate	140	87		***	***	***	ns	*
23	16.8	Moderate	326	106		***	***	***	*	ns

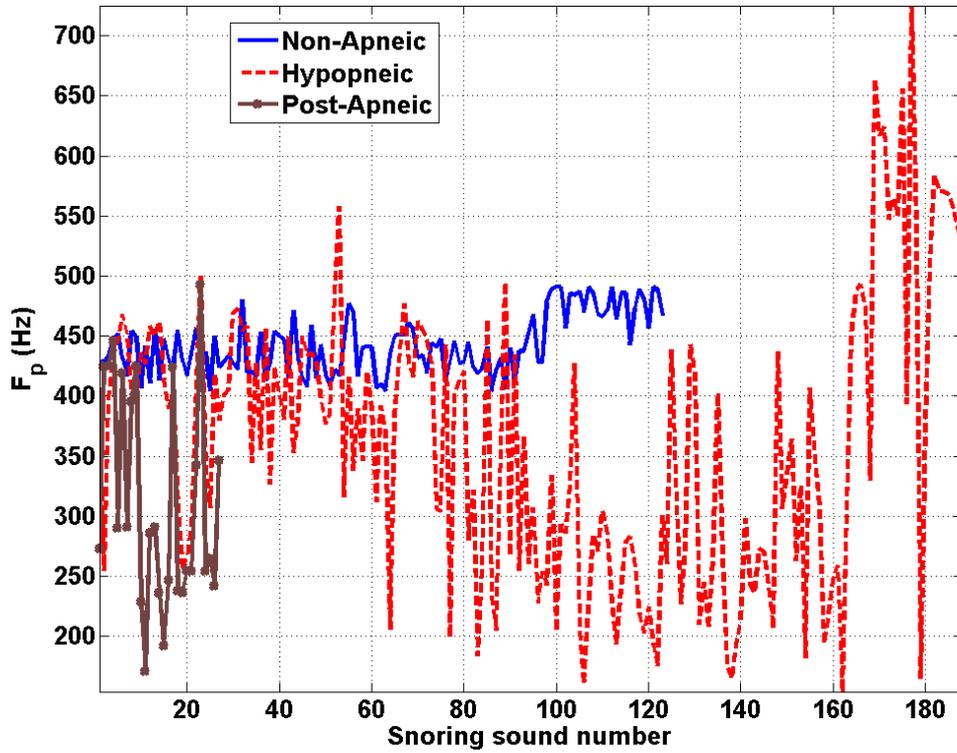


FIGURE 5.5: The intra-subject variability of snoring sound segments' peak frequency (F_p) for an individual. As seen, the variation of F_p is the lowest in non-apneic snoring class, and highest in hypopneic snoring class. However, the total variation for this snorer would be high. The horizontal axis is the number of snoring sound segment.

TABLE 5.5: Result of correlation between AHI and TV for all features.

Feature	r	p
P_{avg}	0.53	<0.000001
ZCR	0.7	<<0.000001
SE	0.6	<0.000001
F_0	0.66	<<0.000001
F_p	0.72	<<0.000001

One can assume the upper airway as a collapsible tube, through which the air flows with different velocity. Once the airway collapses, it stays obstructed as in OSA or dynamically reopens and closes as in non-OSA snorers [16]. The resumption

TABLE 5.6: Regression analysis between TV_{F_p} and AHI.

Model	Standard Error	t	p
$AHI = 5.06 + 0.41 TV_{F_p}$	[1.86, 0.027]	[2.7, 15]	[0.009, 4.37×10^{-21}]
$AHI = 5.1 + 4.36 \times 10^3 TV_{ZCR}$	[2.48, 341.6]	[2.05, 12.8]	[0.045, 4.43×10^{-18}]

TABLE 5.7: Result of Linear Discriminant analysis

Cross-validated ¹	Non-OSA	Mild OSA	Moderate OSA	Severe OSA
Non-OSA	15	0	0	0
Mild OSA	2	10	1	0
Moderate OSA	1	2	10	2
Severe OSA	0	1	4	9

¹ Each case is classified by the function derived from all cases other than that case.

of breathing after apnea is usually accompanied by an abrupt change in airflow [127], pressure [128], and loud inspiratory sounds (post-apneic snoring) [24, 72].

Our results are congruent with those reported in a snoring sound modeling study; using a mechanical model of upper airway, it was shown that in snore generation mechanism, the effect of airflow was dominant [16]. This was observed from the results of this study on a larger group of snorers with sufficient number of snoring segments for each snorer. Another important observation was the change in the snoring sounds characteristics especially during hypopnea between two consecutive respiratory cycles. We speculate that this change might be due to a change in the site of obstruction from the current breathing cycle to the next during hypopneic events.

One of the main reasons that we chose total variation norm to represent the variability was that this norm does not depend on the body position. Let's consider

TABLE 5.8: Comparison between OSA diagnosis method using snoring sounds.

Study	Subjects	Recording Site	Method	AHI threshold (# of Classes)	Sen ¹	Spec ²
[65]	16	Over the neck beside the crycothyroid notch	Mean, standard deviation, and density of pitch	10 (two-class)	64.4	58.5
[66]	383	Anterior neck over the trachea	disturbance index and oxygen desaturation index	5(two-class)	93	67
				15(two-class)	79	95
[67]	16	40-70 cm away from the patient	Intra-Snore-Pitch-Jump probability	5(two-class)	100	50
				10(two-class)	83-91	67
				15(two-class)	90-95	70-80
				30(two-class)	86-93	60-73
[121]	40	30 cm above the mouth	Formant Frequency	10(two-class)	88	82
[91]	40	30 cm above the mouth	Wavelet bicoherence analysis	10(two-class)	85	90.7
Our method	57	Over suprasternal notch of trachea	TV norm of F_p and ZCR	5 (two-class)	92.9	100
				4-class	77.2 (accuracy)	

¹ Sensitivity² Specificity

the snoring segments and classes for the participant shown in Figure 5.5: the non-apneic class of snoring did not vary significantly over time; this would result in the small total variation. Suppose that the body position of this subject changed at some point during the night. As a result, the change in the upper airway status might have shifted F_p for a few respiratory cycles, however, in the absence of apnea or hypopnea, after those cycles in which the body's position changed, the variation of F_p becomes minimum once again. Consequently, the plausible shift in F_p due to a change in the body position would not significantly affect the total variation norm because it affects only a couple of respiratory cycles.

According to Table 5.4, the two features F_p and ZCR were highly significant among three classes of snoring sounds only for 20 patients out of 33 patients. Therefore, using these two features, we are able to label a snoring sound segment as non-apneic, hypopneic, and post-apneic with a moderate accuracy. This did not affect the accuracy of classification of snorers as mild OSA, moderate OSA, severe OSA, and non-OSA (the second part of study). The first goal was to emphasize that the characteristics of snoring sounds vary within a subject depending on the presence of an apneic event. However the second goal was to determine the severity of OSA without any knowledge of snoring segments' labels. This was achieved by using a measure quantifying the sequential variation of snoring sound segments. The important point in the second stage is the amount of sequential variation in those features which was shown to increase as the severity of OSA increases. Summation of sequential variation of these features resulted in the total variation norm that

could discriminate an OSA patient from a simple snorer with a high accuracy.

In a previous study [72], the snoring sounds of 10 non-apneic heavy snorers and 9 snorers with different severity of OSA were investigated. According to their findings, all OSA snorers had loud inspiratory snores during first breath after apnea [72]. However, according to the results of this study, only 0.9% of the snoring segments in our database were labeled as first post-apneic snoring sounds; in other words only 13 out of 42 OSA snorers had snored after apnea. If we calculate this number among only severe snorers, around 5% of snoring sound segments are labeled as first post-apneic snoring sound, which is still a small percentage. Therefore, we suggest that this type of snoring sounds is not of clinical value for screening OSA.

Variation of snoring sound segments' duration (SED), separation (SES), and average power (SEP) were used in a previous study [129] to discriminate simple snorers from OSA patients. The number of snorers in that study was relatively low (30 snorers) and the snoring sounds was recorded by ambient microphone. Ambient recordings are known to be noisier and less discriminative in terms of snoring sound detection [99]. The authors employed a snoring episode identification algorithm (proposed in [130]) to identify the snoring episodes, however, the accuracy of detection was reported to be around 86.8%-97.3% resulted in many misclassified cases in the database. The misclassified cases can vary significantly the result of next stages (e.g. feature calculation, regularity measure, and etc.). In general, it was found that the variation of SED, SES, and SEP was higher in OSA

snorers than simple snorers. However, no information on intra-subject variation of the features and classification scheme was provided. The coefficient variation measure was used to quantify the regularity of SED, SES, and SEP. One major limitation of the coefficient variation is its high sensitivity to sudden change in the body position (as it happens repeatedly during sleep) and noise-like sound existing in the ambient recordings.

We suggest that snoring sounds occurring during hypopnea are of useful and more important clinical value. Due to high variation in airflow and plausible change in the site of obstruction from one breath cycle to the next, the characteristics of snoring sounds during hypopnea is highly variable. On the other hand, the airflow in non-OSA snorers does not vary significantly from breath to breath; thus, the characteristics of snoring sounds do not vary significantly. As a result, one can measure the change of snoring sounds over the sleep time, and estimate the AHI of the snorer based on the amount of variation from one snoring episode to the next. Obviously, this method does not work for OSA patients, who do not snore but the majority of people with OSA do snore at least for a period of time per night. In fact, out of the 68 participants in this study only 11 did not have enough snoring sounds.

In this part of study we used the AHI values as the gold standard for OSA severity. However, the AHI value might be slightly different depending on the definition of hypopnea [131] and from night to night. These limitation of AHI as the gold standard would affect the AHI cut-off value. This, in turn, would affect the

validation of any classification particularly for those people with marginal AHI. For example one of the misclassified cases had an AHI of 15.3; this subject was grouped as moderate, but obviously if AHI value was slightly less, this participant would have been grouped as mild OSA.

In Chapter 4, we averaged the snoring sounds' features within each snorer without considering the intra-subject variability of them. In fact, by averaging, we may smooth away some important information underlying the snoring sounds. We achieved an accuracy around 90% on a population of 28 snorers. On the contrary, in this chapter, we deployed total variation to capture the sequential variation of snoring sounds and we obtained an accuracy of 96.4% discriminating non-OSA snorers and people with OSA ($AHI > 5$). This suggests that intra-subject variations of snoring sounds can be used to extract complementary information about the upper airway state over night.

In summary, we used the linear regression to estimate the AHI values from the total variation features with the highest correlation. We also employed two total variation norms (TV_{F_p} , TV_{ZCR}) to classify the snorers into four classes; the AHI values were used to verify the result of the classification. The results confirm our hypotheses and suggest that the characteristics of snoring sounds change significantly in presence of flow limitation (or airway obstruction). Depending on the severity of flow limitation, the change in snoring sounds from breath to breath is different. This time-dependent intra-subject variation of snoring sounds can be used for OSA screening.

Chapter 6

Why Do Snoring Sounds Form Distinct Clusters?

In Chapter 5, the sequential intra-subject variability of snoring sounds were investigated and it was observed that sequential variation of sounds' features can be used to estimate the AHI. However, there are other parameters such as body position, sleep stage, and blood oxygen level to affect the snoring sounds during sleep. Figure 6.1 shows a scatter plot of snoring sound segments in a 2-D feature space (features 1 and 2 were normalized to have zero mean and unit variance). As clearly shown in this figure, the snoring sound segments form two distinct clusters. Therefore, we questioned: What are the plausible causes of forming clusters of snoring sounds for each individual?. This part of study aimed to seek answer for the aforementioned question using the available information from PSG study and snoring sound analysis. In particular, the relationship between the snoring

sounds characteristic features and some physiological parameters such as sleep stage, blood oxygen level, and body position are investigated in this chapter.

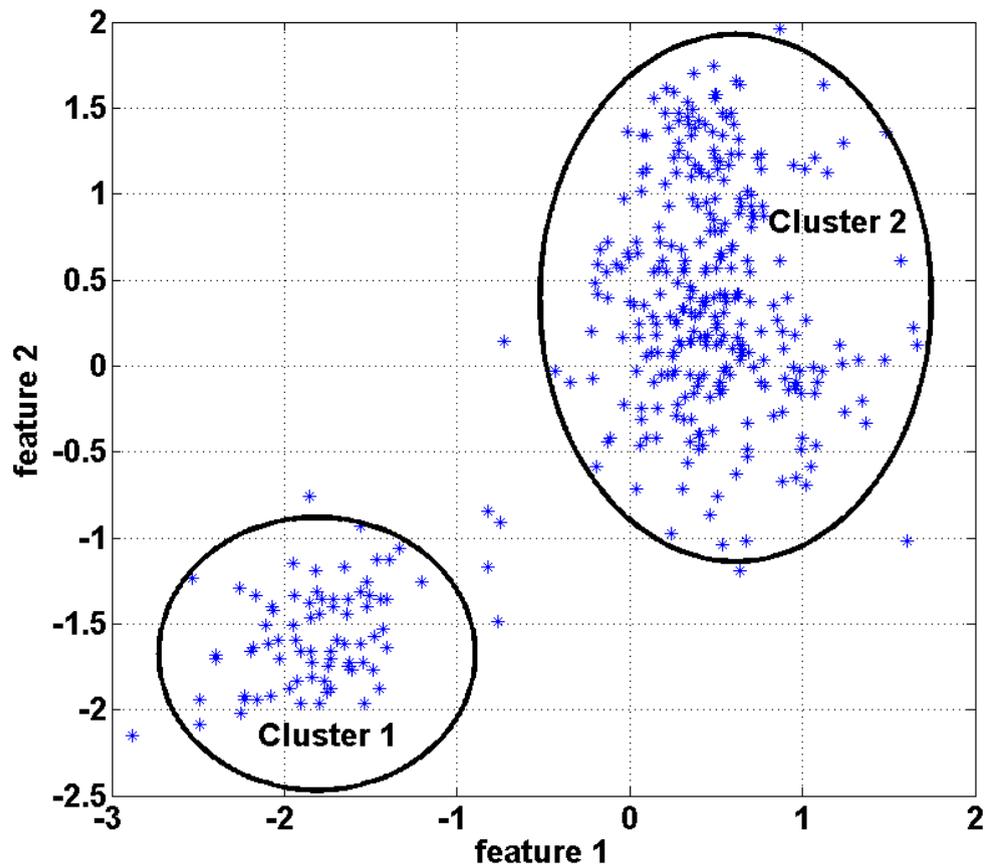


FIGURE 6.1: A scatter plot of snoring sound segments within a typical snorer in a 2-D feature space. Feature 1 and feature 2 were standardized such that they have zero mean and unit variance.

TABLE 6.1: Anthropometric Parameters of the study participants.

Group	Number of subjects	Age	BMI	AHI
OSA	42 (9 females)	52.2± 12.4	33.9± 6.7	35.5±33.1
Non-OSA	15 (6 females)	49.5±10.4	30.1± 4.0	2.4±1.3

6.1 Method

6.1.1 Data Recording

In total, snoring sound segments of 57 snorers were used for this study. Table 6.1 shows the anthropometric parameters of the snorers.

The algorithm proposed in [99] was run on each individual's respiratory sounds to extract all snoring sound segments. It is worth noting that the same snoring sound data set explained in Section 5.1.2 was used in this chapter.

6.1.2 Feature Extraction

All snoring sound segments were first band-pass filtered in the frequency range of 150-2000 Hz (to remove the effect of heart sounds and high frequency noises) resulting zero-mean signals. Twelve features (conventional linear and HOS features) were extracted from each snoring sound segment. The following features were then calculated for each segment:

Power (P) The signal's power was computed as:

$$P = \frac{1}{N} \sum_{n=1}^N |s(n)|^2 \quad (6.1)$$

where N is the length of each snoring sound segment.

Zero Crossing Rate (ZCR) It was defined as

$$ZCR = \frac{1}{2N} \sum_{n=1}^{N-1} |sgn[s(n+1)] - sgn[s(n)]| \quad (6.2)$$

where $sgn[.]$ represents the sign function. ZCR is a time domain feature representing the number of times that the amplitude of a signal changes sign. ZCR has been mainly used for detecting voiced and unvoiced parts of the speech [124]. It has also been used for categorizing snoring sound into voiced and unvoiced categories [67].

500Hz Sub-band normalized energy ($E_1 - E_2$) The 0-2000Hz frequency range was divided into 500Hz sub-bands (4 sub-bands in total) and the average normalized energy in each sub-band was calculated. The sub-band energy features were originally proposed in [130] to detect the snoring sound segments from respiratory sounds. In this study, the first 2 sub-bands energy $E_1 - E_2$ were used because the sounds' major power lies below 1000Hz as discussed in Appendix A.

Skewness and kurtosis They were defined as:

$$\gamma_1 = \frac{c_3(0,0)}{\sigma_s^3} \quad (6.3)$$

$$\gamma_2 = \frac{c_4(0,0,0)}{\sigma_s^4} \quad (6.4)$$

respectively, where σ_s is the standard deviation of $s(n)$ and $c_3(0, 0)$ and $c_4(0, 0, 0)$ are its zero-lag 3rd and 4th order cumulants respectively [105]. It should be noted that skewness is a measure of asymmetry of the probability distribution function (pdf) while kurtosis is a measure of peakedness of the pdf.

Formant Frequencies ($F_1 - F_3$) The first three formant frequencies [132] were calculated from each snoring sound segment using linear predictive coding (LPC). To meet stationarity assumption, $s(n)$ was divided into 100 ms overlapping frames (50% overlap and Hanning window). In each frame, the autoregressive (AR) model with order 14 of the signal was estimated and the roots of AR model were calculated. To select the AR model order, we used the optimum order model (optimum order = $f_s(KHz) + \gamma$, $\gamma = 4, 5$ & $f_s \in [6-18]KHz$) suggested in [106]. Then, the first three formant frequencies were estimated by taking median over all frames. Therefore, there were three formant frequencies namely F_1 , F_2 , and F_3 for each snoring sound segment.

Crest factor (CF) The snoring sound signal, $s(n)$, was divided into 100 ms windows (with 50% overlap between successive windows) and the crest factor were calculated for each frame (as below):

$$CF = \frac{V_{99}}{V_{rms}} \quad (6.5)$$

where V_{99} is the 99th centile as a measure of peak value and V_{rms} is the root mean square value [88]. The median of CF 's over all frames was used as the equivalent CF for each snoring sound segment.

Spectral entropy (SE) SE measures the flatness of the spectrum. Larger values of SE correspond to the broader spectral contents [133]. It is computed as:

$$SE = - \sum_f P_n(f) \cdot \ln(P_n(f)) \quad (6.6)$$

where $P_n(f)$ refers to normalized power spectrum density (PSD) at discrete frequency f . The PSD was estimated using Welch method [71] using Hanning windows of size 100 ms and 50% overlap between the successive windows.

Central Tendency Measure (CTM) CTM is a variability measure from second order difference plots [125]. CTM is calculated as below:

$$CTM = \frac{1}{N-2} \sum_{n=1}^{N-2} \delta(n) \quad (6.7)$$

where $\delta(n)$ is defined as:

$$\delta(n) = \begin{cases} 1, & \sqrt{(s(n+2) - s(n+1))^2 + (s(n+1) - s(n))^2} < \rho \\ 0, & \textit{otherwise} \end{cases} \quad (6.8)$$

Parameter ρ defines the radius of a circular region around the origin in second order difference plots. We used $\rho = 1$ for this part of study.

6.1.3 Principal Component Analysis

As explained in Section 6.1.2, 12 sound features were extracted from each snoring sound segment resulting in a 12-D feature vector. However, some of these features may be redundant. Therefore, to reduce the dimensionality of the feature space, the Principal Component Analysis (PCA) [81] was deployed. For each snorer, there existed a data matrix with M_i rows (number of snoring sound segment for snorer i) and 12 columns (total number of features). Transforming the data using PCA resulted in a $M_i \times q$ matrix. It should be noted that q ($q < 12$) is the number of principal components with highest eigenvalues. In fact, q is the number of first principal components that explains 80% of the total variation in the data and can be calculated as:

$$\frac{\sum_{j=1}^q e_j}{\sum_{j=1}^{12} e_j} \geq 0.8 \quad (6.9)$$

where e_j denotes the eigenvalues of the covariance matrix of dimension j of data set. The first q principal components were the eigenvectors corresponding to the highest eigenvalues explaining 80% of the total variation. The rest of eigenvectors were, in fact, ignored.

6.1.4 Grouping intra-subject snoring sound segments based on PSG information

For each snoring sound segment, physiological parameters including body position, sleep stage, and blood oxygen level were extracted from PSG score sheet. All these

parameters were categorical variables with 2-6 classes. Body position had 4 classes namely *supine*, *prone*, *left*, and *right*. Sleep stage included 6 classes: *stage 1-4*, *Arousal*, and *REM*. Lastly, blood oxygen level had two classes of *normal* and *desaturation* (drop greater than 4% in blood oxygen level). Therefore, in addition to q sound features, every snoring sound segment had three categorical features for body position, sleep stage, and blood oxygen level. Without loss of generality, we merged *left* and *right* position into a new class called *side* position. We also merged *stage 1-4* into a new class called *NREM*. Hence, this resulted in 3 classes of *supine*, *prone*, and *side* for body position, 3 classes of *NREM*, *Arousal*, and *REM* for sleep stage, and 2 classes of *normal* and *desaturation* for blood oxygen level. It is worth noting that we investigated all sleep stages, i.e. stage 1-4, REM, and arousal as well as all sleeping positions, i.e. left, right, prone, and supine. However, due to similarity of some of these classes and to give a broader picture, we merged some of the classes. The result of categorical variables without merging is presented in Appendix B. It should also be noted that some snorers had only one class of snoring sounds in each categorical variable and the number of classes shown above is the maximum possible number of classes. Some snorers had only snored in one position or one sleep stage.

6.1.5 One-dimensional and two-dimensional probability density function

Depending on the categorical variable, there existed different number of classes of snoring sound segment within every snorer. To examine if the categorical variable in question is really affecting the snoring sound segments, we can measure the overlap between one-dimensional (two-dimensional) pdf of the existing classes of snoring sounds for a single feature (a bivariate feature vector). For example, one can examine the effect of blood oxygen level on the snoring sounds by 1) estimating the one-dimensional (1-D) pdf of a single feature (e.g. $PC_1, PC_2, \dots, \text{or } PC_q$) for both classes and 2) measuring the overlap between estimated pdf of the classes. As a result, if the class densities do not overlap that means blood oxygen level does not affect the extracted feature. However, the amount of overlap quantifies how much the categorical variable is affecting that sound feature. Two-dimensional (2-D) case is similar to that of 1-D. First, we need to estimate the 2-D pdf of a bivariate feature vector (e.g. $(PC_1, PC_2), \dots, (PC_{q-1}, PC_q)$, or etc.) for all classes (e.g. normal and desaturation), then, we need to measure the overlap between 2-D pdfs of the classes and quantify how affective the categorical variable is.

Consider a general case when we have c classes (e.g. $c = 2$ for blood oxygen level and $c = 3$ for sleep stage and body position variables). The detail of this part of algorithm comes in the following:

1. For each class of snoring sound segment, the 1-D and 2-D pdf (\hat{f}) were estimated using Kernel Density Estimation method [108]. That is, there

were q transformed sound features (PCs) that were selected based on the criterion introduced in Section 6.1.3. The pdfs were estimated as:

$$\hat{f}_k(X; H) = \frac{1}{N_k^i} \sum_{n \in \text{Class}(k)} K_H(X - X_n), \quad k = 1, \dots, c \quad (6.10)$$

where N_k^i is the number of snoring sound segments in class k for snorer i , $\text{Class}(k)$ denotes the all members of class k . In 1-D case, $K_H(x) = \frac{1}{H} K(\frac{x}{H})$ where H is the bandwidth and in 2-D case $K_H(x) = |H|^{-\frac{1}{2}} K(H^{-\frac{1}{2}}x)$ where H is a 2×2 bandwidth matrix which was selected optimally using mean integrated squared error [134], $K(\cdot)$ is the kernel function (we used Normal kernel in this paper). X denotes the point where we want to estimate the pdf and X_n refers to the sample points in class k . In 1-D case, $X = \{PC_j^k, j = 1, \dots, q\}$ while in 2-D case, $X = \left\{ (PC_u^k, PC_v^k)^T \mid u < v, u = 1, \dots, q-1 \text{ \& } v = 2, \dots, q \right\}$ where PC_j^k is the j -th principal component (transformed feature) in the class k .

2. Now we need to measure the distance between class densities. If the pdfs do not overlap, it means that the c classes of snoring sound segments form c distinct clusters. On the other hand if they fully overlap, that means the categorical variable does not have any effect on the snoring sound segments. We used three measures to quantify the class discrimination capability of principal components: L_1 distance [104] between every two pdfs both in 1-D (L_1^{1D}) and 2-D cases (L_1^{2D}) as well as the area under Receiver Operating

Characteristics (ROC) curve [135] (*AUROC*) in 1-D case. The L_1 distance in general case is defined as:

$$L_1 = \int_{-\infty}^{+\infty} |f(X) - g(X)| dX \quad (6.11)$$

where X is a multivariate vector and $f(\cdot)$ and $g(\cdot)$ are multivariate pdf functions. To obtain a number in the range of $[0, 1]$, we divided the above norm by 2. To approximate the above integration, we used trapezoid rule [136].

The ROC curve was used to visualize the class discrimination capability of a principal component. After constructing the ROC curve, the area under the curve was calculated. Figure 6.2 shows how we constructed the ROC curve. The top graph in Figure 6.2 shows an example of a two overlapping pdfs describing the distribution of feature (PC_1) in two classes. Consider the threshold line shown in Figure 6.2 sweeping from leftmost of the graph to the rightmost of the graph. Each time we measure the areas α and β . Once we swept the threshold line over the entire PC axis and calculated the aforementioned areas, we can plot $1 - \beta$ with respect to α . The area under the resulted curve will be our third measure of separability between class densities. This was only performed for 1-D case.

Consequently, we obtained 3 distance measures: L_1^{1D} , L_1^{2D} , and *AUROC*. Fully overlapping densities will result in $L_1^{1D} = 0$, $L_1^{2D} = 0$, and *AUROC* = 0.5, on

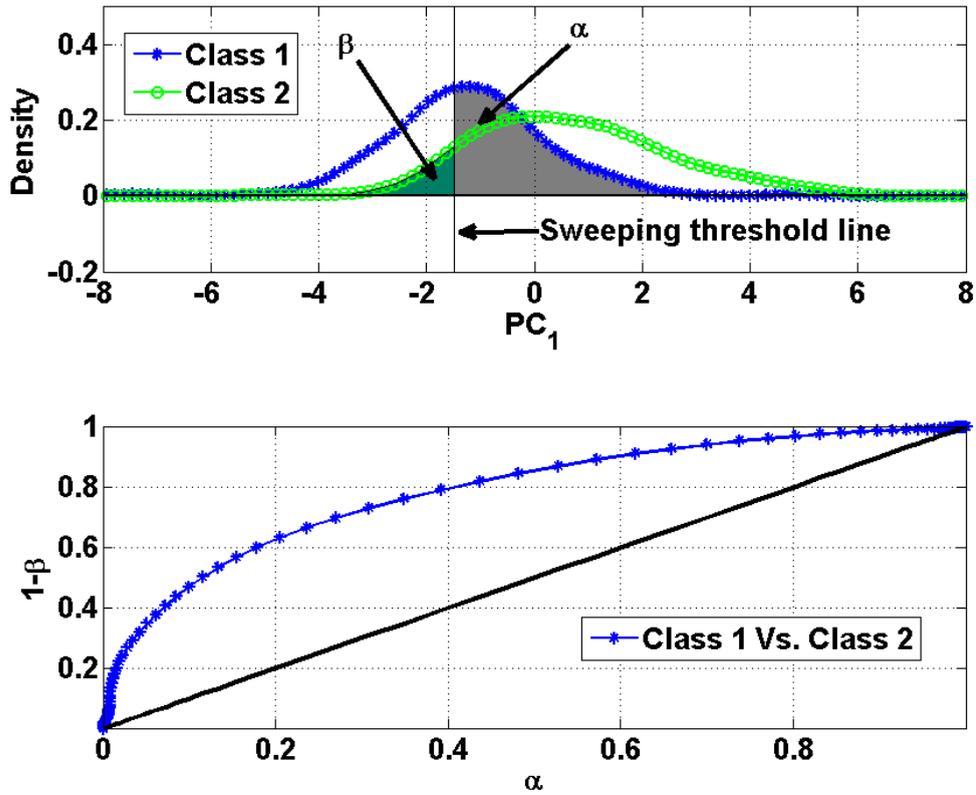


FIGURE 6.2: Top: pdf of snoring sounds' first principal component in two different classes. Sweeping threshold line was used to calculate the areas α and β . Bottom: ROC curve constructed by plotting $1 - \beta$ with respect to α .

the other hand, fully separated densities will result in $L_1^{1D} = 1$, $L_1^{2D} = 1$, and $AUROC = 1$.

6.1.6 Statistical Analysis

As for the body position variable, there are three possible comparisons between the pdfs: Supine vs. Prone, Supine vs. Side, and Prone vs. Side. For sleep stage variable, there are also three possible comparisons between the pdfs: NREM vs. REM, NREM vs. Arousal, and REM vs. Arousal. Finally, for blood oxygen level,

there is one possible comparison between the pdfs: Normal vs. Desaturation. For every categorical variable, we measured the L_1^{1D} , L_1^{2D} , and $AUROC$ for every possible comparison mentioned above. This was repeated for all study subjects' data. Once these measures were calculated for all snorers, we used Analysis of Variance (ANOVA)[97] to investigate any statistically significant differences between the groups of each categorical variable.

Lastly, to find the categorical variable with the highest effect on the snoring sounds, we rearranged all the comparisons in one group, that is, Body Position={Supine vs. Prone, Supine vs. Side, Prone vs. Side}, Sleep Stage = {NREM vs. REM, NREM vs. Arousal, REM vs. Arousal}, and Blood Oxygen={Normal vs. Desaturation } and visualized the L_1^{1D} , L_1^{2D} , and $AUROC$ for the rearranged groups of Body Position, Sleep Stage, and Blood Oxygen. In addition, we ran an ANOVA test on these groups.

6.2 Results

6.2.1 Principal Component Analysis

It was observed that the first 4 principal components explain more than 80% of the total variation in the original feature space. Therefore, we selected $q = 4$. Hence, in the reduced feature space, every snoring sound segment could be identified by a 4-D transformed sound feature (PC_1 , PC_2 , PC_3 , PC_4) and a 3-D categorical

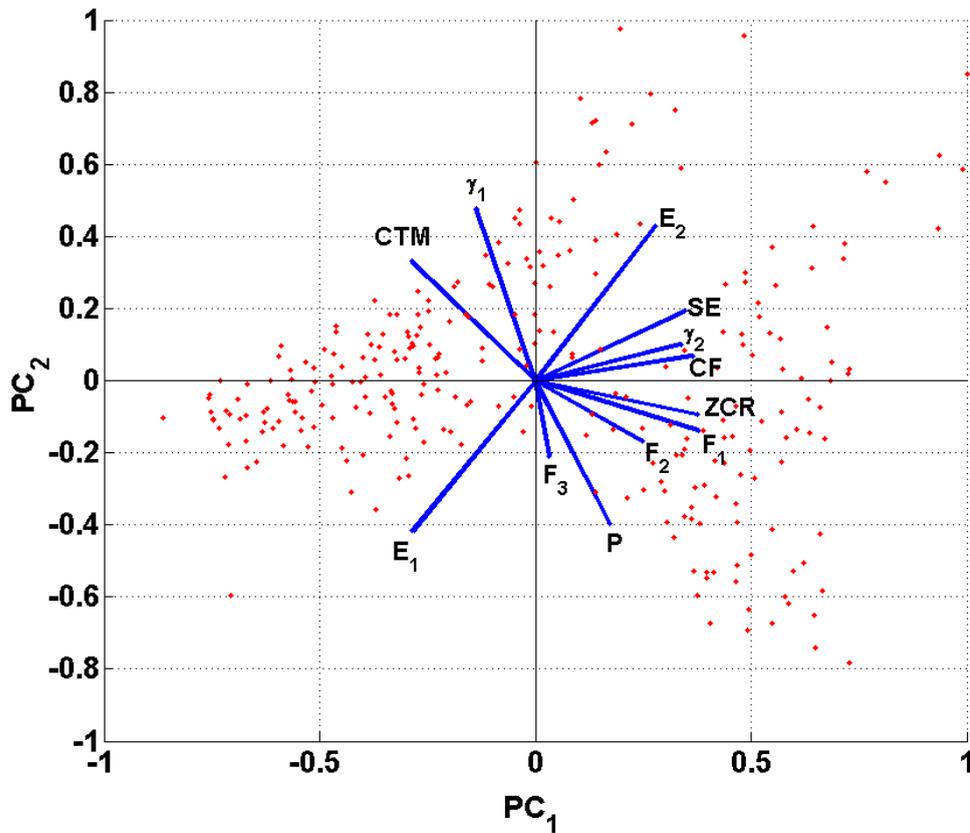


FIGURE 6.3: Bi-plot of transformed snoring sounds' features (the first two PC s with highest eigenvalues) along with original feature vector. Each vector shows the magnitude and sign of each feature's contribution to the first two PC s.

feature (body position, sleep stage, blood oxygen). Figure 6.3 shows a 2-D bi-plot of reduced feature space along with original features. The bi-plot shows the magnitude and sign of each feature's contribution to the first two principal components (PC_1 , PC_2), and how each snoring sound segment is represented in terms of those components.

Figures 6.4 to 6.6 show the scatter matrix of snoring sound segments grouped based on body position, sleep stage, and blood oxygen level, respectively. The most distinct clusters were observed in Figure 6.4 where the snoring sounds were

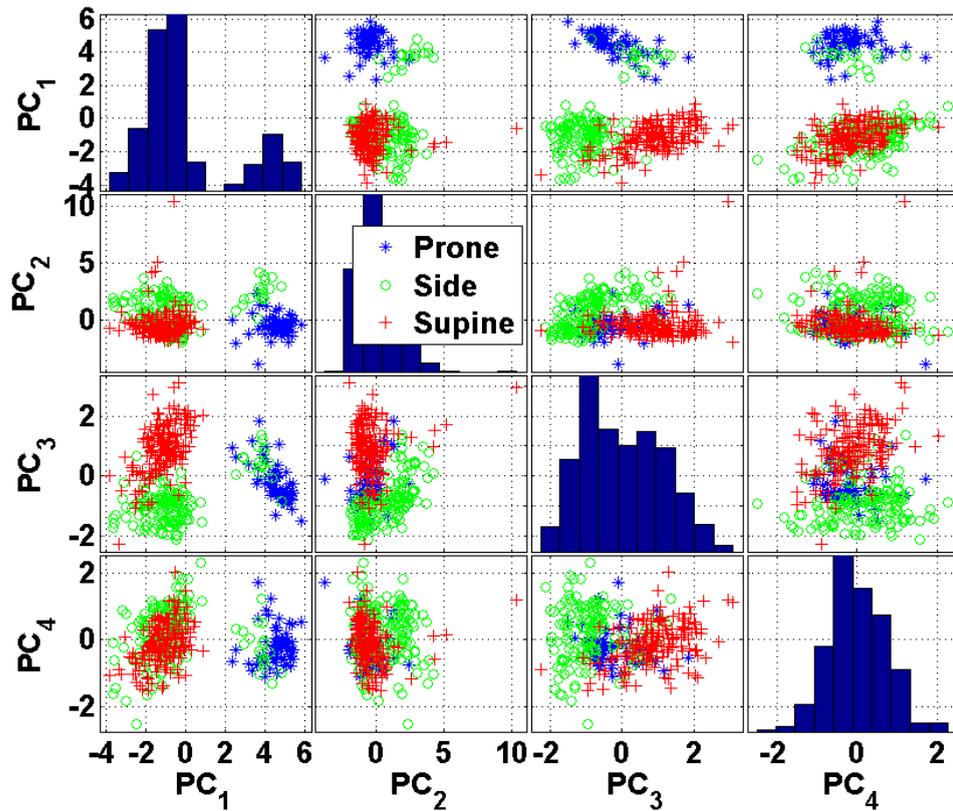


FIGURE 6.4: Scatter matrix of snoring sound segments' PC s categorized based on body position. The diagonal graphs show the histogram of each PC .

grouped based on the body position. However, this cannot be generalized by observing only one snorer. The average separation (results of Section 6.1.6) can be more conclusive in intra-subject separation of snoring sounds.

6.2.2 One-Dimensional and two-dimensional pdfs and ROC curves

Figures 6.7 to 6.9 show the 1-D pdf of first PC with highest eigenvalue along with ROC curves and contour plots of 2-D pdfs for different classes of snoring sounds

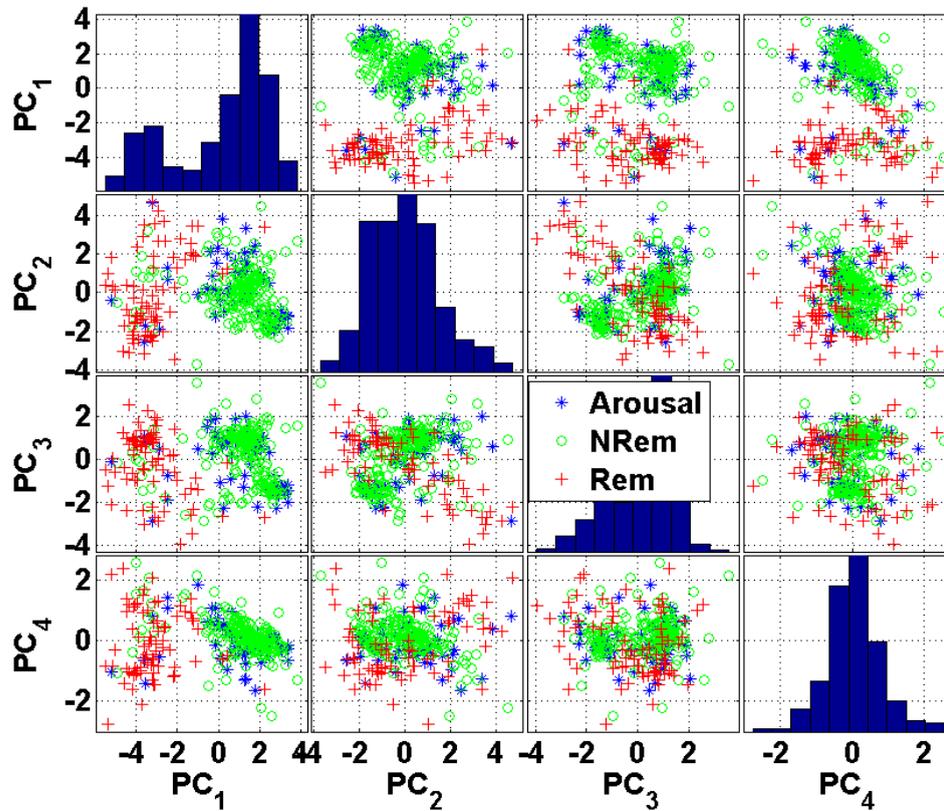


FIGURE 6.5: Scatter matrix of snoring sound segments' PC s categorized based on sleep stage. The diagonal graphs show the histogram of each PC .

within a snorer. Figure 6.7a demonstrates the 1-D pdf of PC_1 among three classes of snoring sounds. Figure 6.7b shows the contour plots of the 2-D pdf of snoring sounds grouped based on body position.

Figures 6.8 and 6.9 show the contour plots of the 2-D pdf of snoring sounds categorized based on sleep stage and blood oxygen level, respectively. As shown in Figure 6.7, both 1-D and 2-D pdfs were consistently discriminated the prone position from side and supine positions. The $AUROC$ (Figure 6.7a) was 1 for class densities supine and prone and it was very close to 1 for class densities prone and side. However, it was approximately 0.5 for class densities side and supine. This

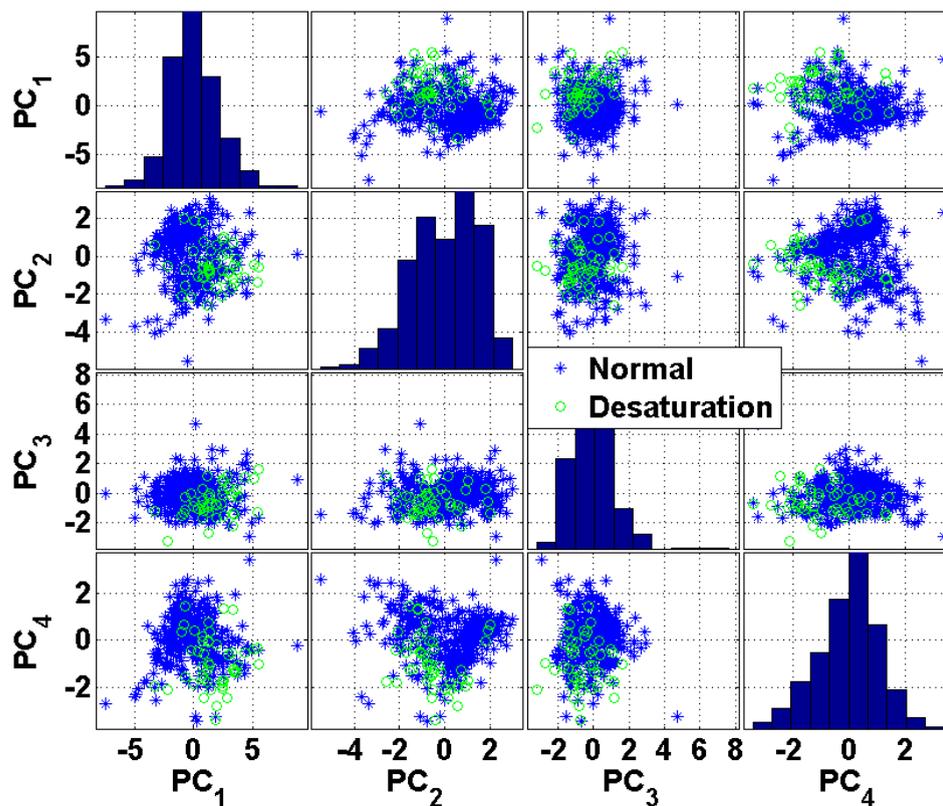


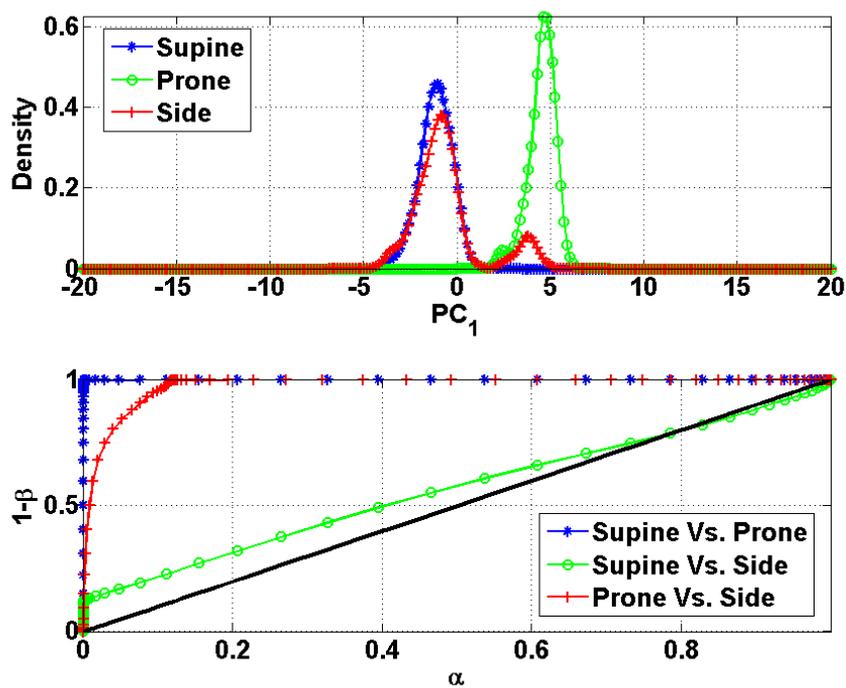
FIGURE 6.6: Scatter matrix of snoring sound segments' PC s categorized based on blood oxygen level. The diagonal graphs show the histogram of each PC .

is due to high overlap between supine and side class densities, and high separation between side and prone as well as supine and prone. The same analogy applies to Figures 6.8 and 6.9.

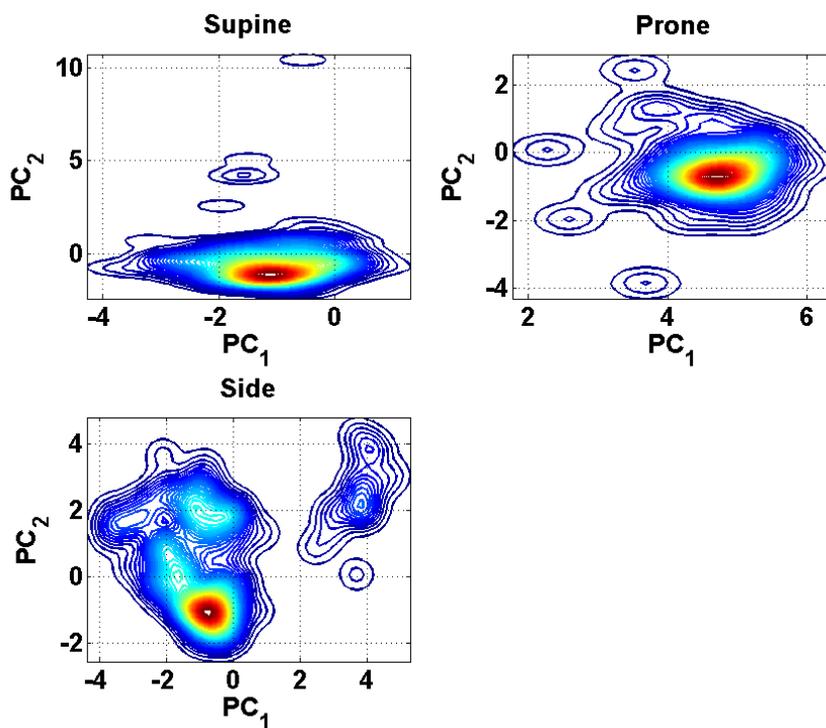
6.2.3 Statistical Analysis

6.2.3.1 Body Position

Figure 6.10 shows the mean and standard deviation of separations among density classes based on body position. As shown in this figure, prone position has the

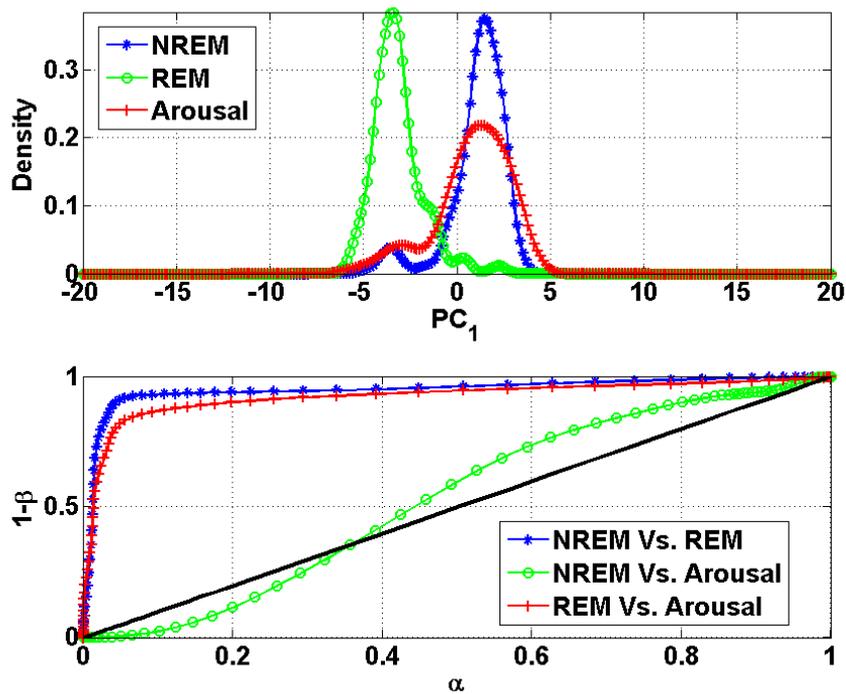


(a)

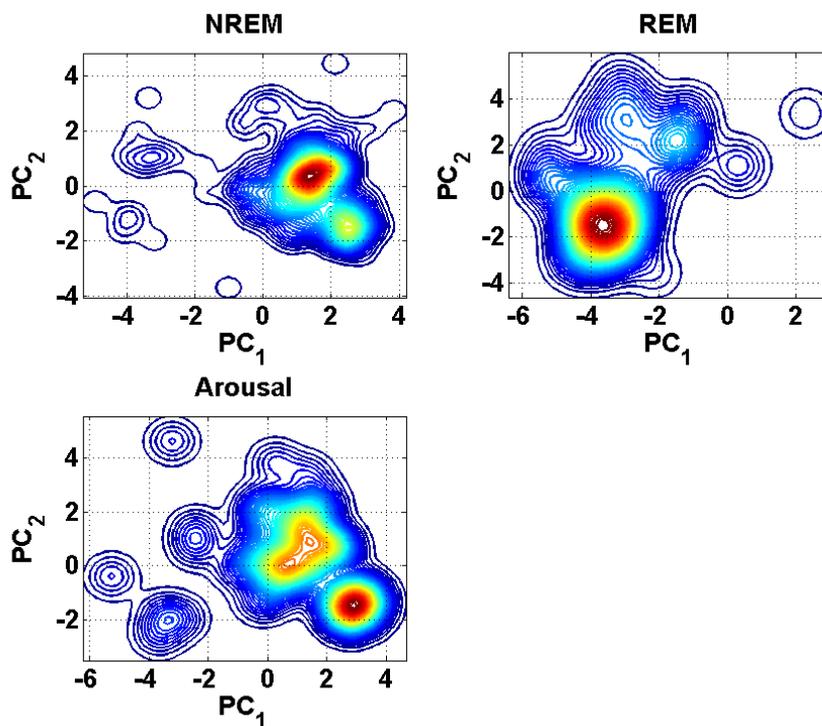


(b)

FIGURE 6.7: (a) 1-D pdf of snoring sound segments' first PC categorized based on body position. (b) Contour plot of snoring sound segment of the first two PC s categorized based on body position. As shown in this figure, the 2-D pdf of prone position is significantly different than that of side and supine positions.

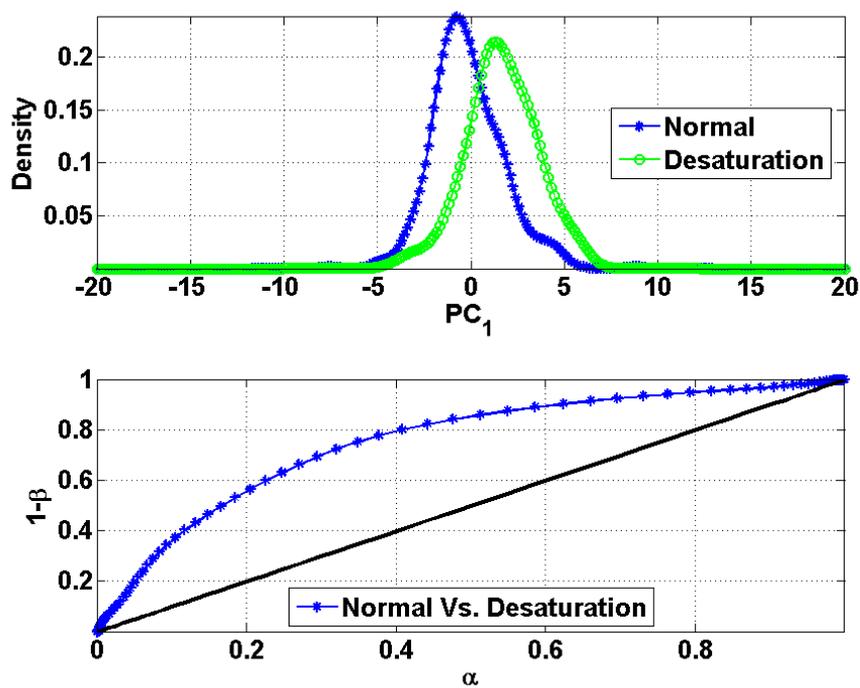


(a)

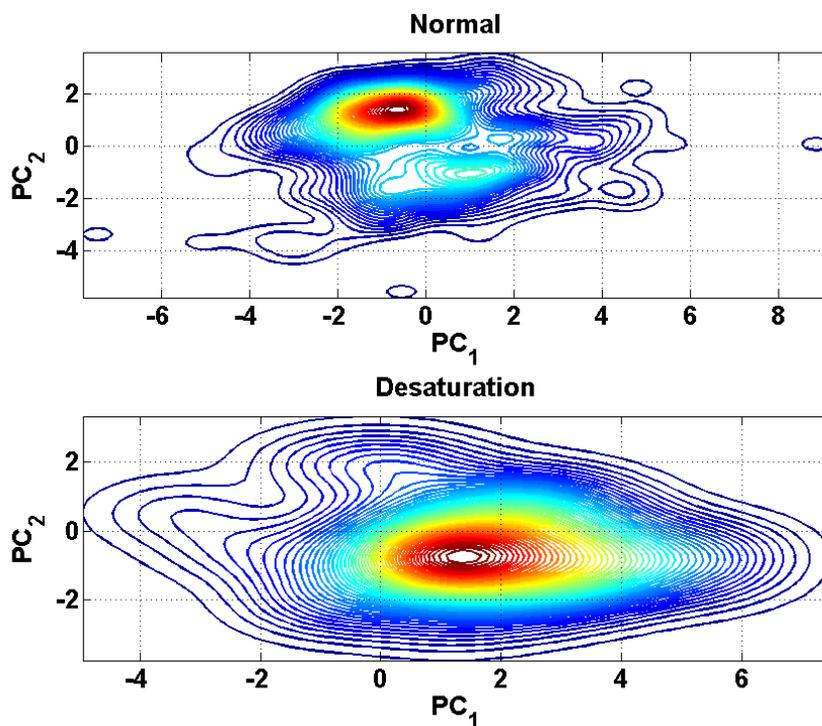


(b)

FIGURE 6.8: (a) 1-D pdf of snoring sound segments' first PC categorized based on sleep stage. (b) Contour plot of snoring sound segment of the first two PC s categorized based on sleep stage. As shown in this figure, the 2-D pdf of NREM, REM, and Arousal are different.



(a)



(b)

FIGURE 6.9: (a) 1-D pdf of snoring sound segments' first PC categorized based on blood oxygen level. (b) Contour plot of snoring sound segment of the first two PC s categorized based on blood oxygen level. As shown in this figure, the 2-D pdf of NREM, REM, and Arousal are different.

lowest overlap with the other two positions. However, the most consistent comparison is when a bivariate PC is used (Figure 6.10a). This suggests that the highest change in snoring sounds occurs when the snorer changes the body position from prone to any other position. However, there is still a high separation among all positions. Even the positions with highest overlap (supine and side) still are well separated as $L_1^{2D} = 0.72 \pm 0.19$ when we compared class densities using first two PC s. This value dropped to $L_1^{2D} = 0.57 \pm 0.13$ when we used the last two PC s (PC_3, PC_4). This supports that even the lowest significant features could moderately discriminate snoring sound segments in different body positions. The other two distance measures are not as consistent as the L_1^{2D} measure. Figure 6.10b shows the L_1^{1D} norm with the highest value for PC_1 and side and prone density comparison ($L_1^{1D} = 0.68 \pm 0.14$).

Lastly, in Figure 6.10c, the highest separation happened between prone and side ($AUROC = 0.87 \pm 0.12$). Both L_1^{1D} and $AUROC$ resulted in the lowest separation between supine and side positions using PC_4 . We also ran an ANOVA test between distance measures of each comparison (3 groups: supine vs. prone, supine vs. side, and prone vs. side) to check the equality of mean of the comparison groups. As shown in Table 6.2, there was no significant difference among means of three groups.

TABLE 6.2: Results of ANOVA on comparison groups of body position. ns: not significant ($p>0.05$), and *: significant ($0.01<p<0.05$), **: very significant ($0.001 <p<0.01$) and ***: highly significant ($p<0.001$).

Distance Measure	PC	Significance Level
L_1^{2D}	$PC(1, 2)$	ns
	$PC(1, 3)$	ns
	$PC(1, 4)$	ns
	$PC(2, 3)$	ns
	$PC(2, 4)$	ns
	$PC(3, 4)$	ns
L_1^{1D}	$PC(1)$	ns
	$PC(2)$	ns
	$PC(3)$	ns
	$PC(4)$	ns
$AUROC$	$PC(1)$	ns
	$PC(2)$	ns
	$PC(3)$	ns
	$PC(4)$	ns

6.2.3.2 Sleep Stage

Figure 6.11 shows the mean and standard deviation of class density distance among sleep stages. As shown in Figure 6.11, REM stage has the lowest overlap with NREM and Arousal classes. This is consistent among all three distance measures.

Figure 6.11a displays the L_1^{2D} mean and standard deviation using all combinations of bivariate PC s in three classes.

The highest separation occurred between NREM and REM classes of sleep stage using (PC_1, PC_4) . This resulted in: $L_1^{2D} = 0.64 \pm 0.2$. The distance between REM and Arousal classes is still high and comparable with that of NREM and REM.

The highest separation happened using (PC_1, PC_2) resulting in $L_1^{2D} = 0.63 \pm 0.16$.

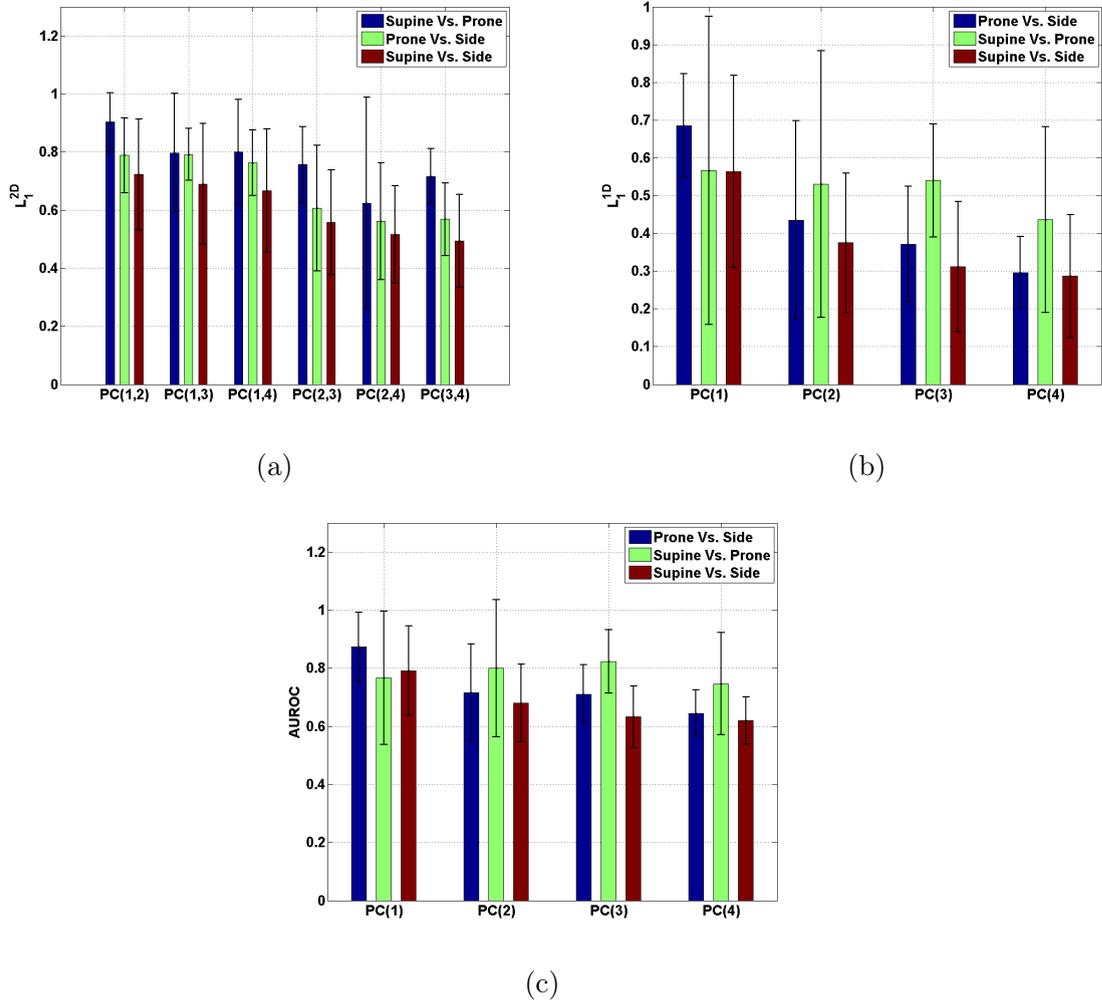


FIGURE 6.10: The mean and standard deviation of distance measures between body position classes' pdfs among all snorers (a). The mean and standard deviation of L_1^{2D} , the highest separation is between prone and supine positions, then, between prone and side, and the lowest is between supine and side positions. (b) The mean and standard deviation of L_1^{1D} . (c) The mean and standard deviation of $AUROC$.

The NREM and Arousal comparison resulted in the significantly lower values of L_1^{2D} than those of previous ones (the highest was $L_1^{2D} = 0.41 \pm 0.15$ using (PC_1, PC_2)). This was consistent among all distance measures as REM vs. Arousal and NREM vs. REM distance measures (L_1^{1D} , $AUROC$) were comparable and NREM vs. Arousal measure was significantly lower in both cases (L_1^{1D} , $AUROC$).

TABLE 6.3: Results of ANOVA on comparison groups of sleep stage.

Distance Measure	PC	Significance Level
L_1^{2D}	$PC(1, 2)$	***
	$PC(1, 3)$	**
	$PC(1, 4)$	***
	$PC(2, 3)$	**
	$PC(2, 4)$	**
	$PC(3, 4)$	***
L_1^{1D}	$PC(1)$	***
	$PC(2)$	*
	$PC(3)$	*
	$PC(4)$	***
$AUROC$	$PC(1)$	**
	$PC(2)$	ns
	$PC(3)$	*
	$PC(4)$	**

This suggests that if sleep stage changes from REM to Arousal or NREM, there would be a significant change in the characteristics of snoring sounds. The result of ANOVA between distance measures of each comparison (3 groups: NREM vs. REM, NREM vs. Arousal, and REM vs. Arousal) shows that REM sleep stage caused the significantly highest change in the characteristics of snoring sounds. This was shown in Table 6.3.

6.2.3.3 Blood Oxygen Saturation

Figure 6.12 shows the mean and standard deviation of class density distance between blood oxygen classes. As shown in Figure 6.12a, the highest separation between Normal and Desaturation classes occurred when we used (PC_1, PC_2) which resulted in $L_1^{2D} = 0.49 \pm 0.2$. The L_1^{2D} measure showed a decreasing trend

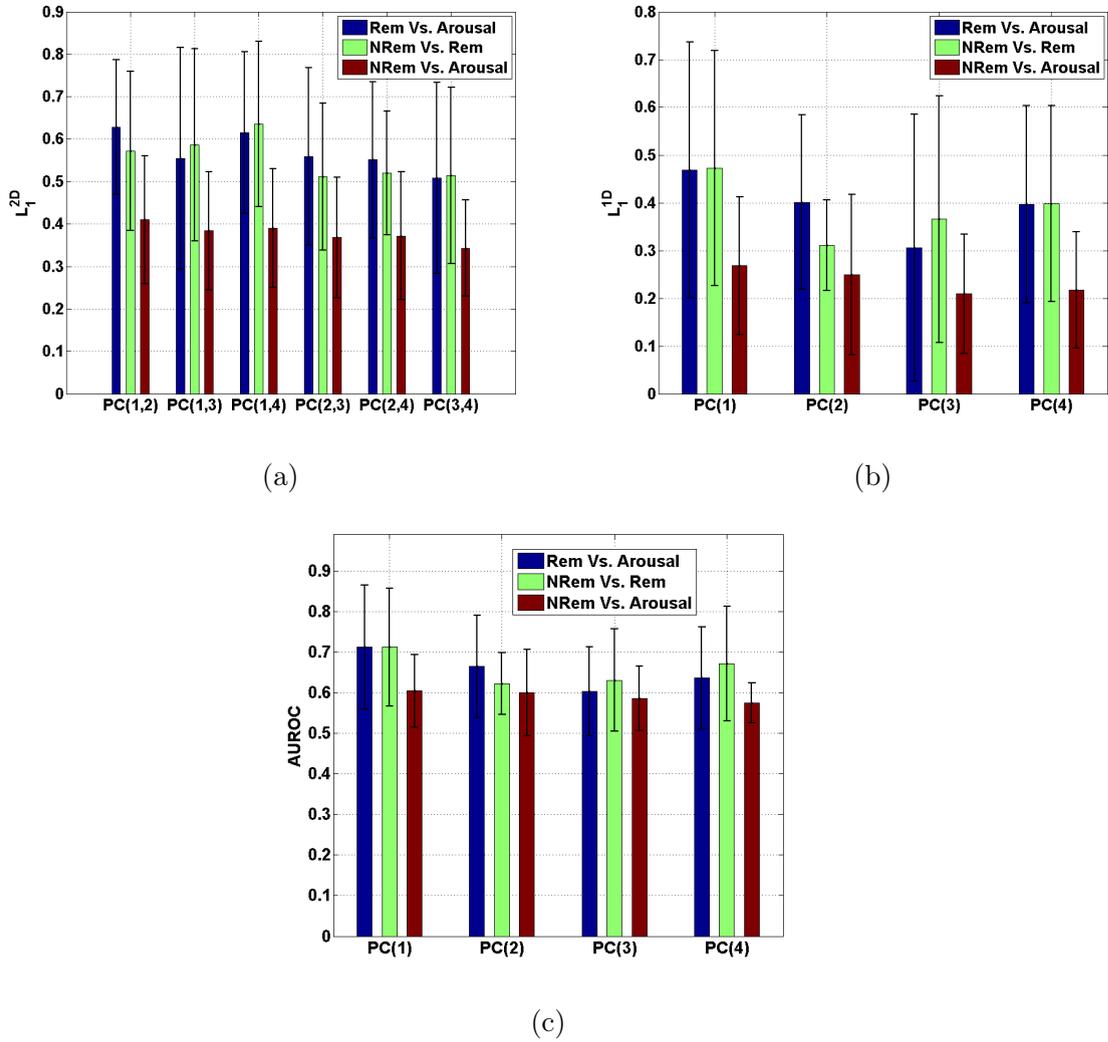


FIGURE 6.11: The mean and standard deviation of distance measures between sleep stage classes' pdfs among all snorers (a). The mean and standard deviation of L_1^{2D} , the highest separation is between NREM and REM sleep stages ($L_1^{2D} = 0.64 \pm 0.2$), then, between REM and Arousal ($L_1^{2D} = 0.63 \pm 0.16$), and the lowest is between NREM and Arousal. (b) The mean and standard deviation of L_1^{1D} . (c) The mean and standard deviation of $AUROC$.

as we used less significant PC s to estimate the density function of two classes. The lowest was using (PC_3, PC_4) which resulted in $L_1^{2D} = 0.40 \pm 0.17$. This trend was consistent among all measures of distances. The highest L_1^{1D} was for PC_1 ($L_1^{1D} = 0.33 \pm 0.19$) and the lowest was for PC_4 ($L_1^{1D} = 0.26 \pm 0.14$). The

highest $AUROC$ was for PC_1 ($AUROC = 0.66 \pm 0.12$) and the lowest was for PC_4 ($L_1^{1D} = 0.60 \pm 0.07$).

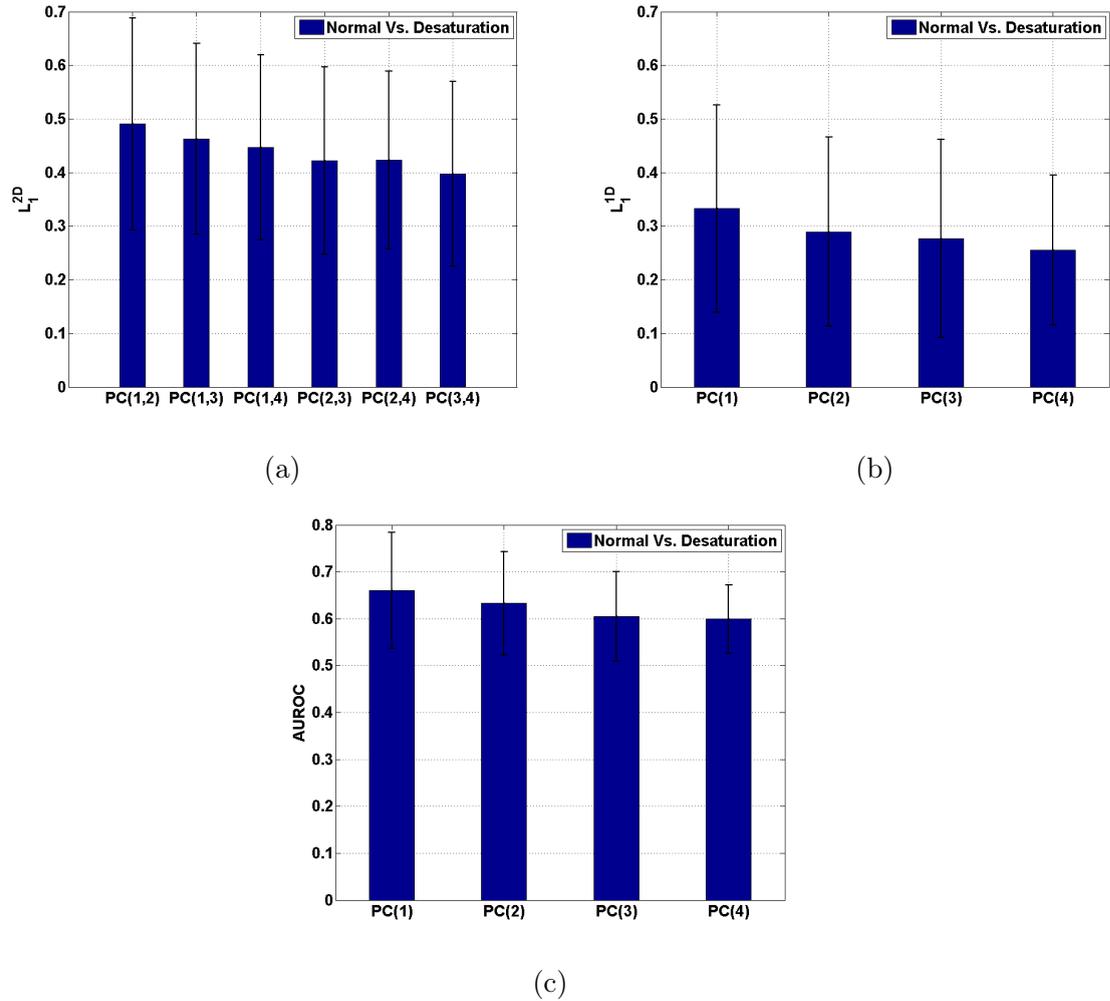


FIGURE 6.12: The mean and standard deviation of distance measures between principal components' class densities of normal and desaturation among all snorers. (a) The mean and standard deviation of L_1^{2D} for all bivariate PCs. (b) The mean and standard deviation of L_1^{1D} . (c) The mean and standard deviation of $AUROC$.

TABLE 6.4: Results of ANOVA on two by two categorical variables and on all three groups of categorical variables. BP: Body Position, OL: blood Oxygen Level, SS: Sleep Stage

Distance Measure	PC	Significance Level			
		BP vs. SS	BP vs. OL	SS vs. OL	All groups
L_1^{2D}	$PC(1, 2)$	***	***	ns	***
	$PC(1, 3)$	***	***	ns	***
	$PC(1, 4)$	***	***	ns	***
	$PC(2, 3)$	***	**	ns	***
	$PC(2, 4)$	**	*	ns	**
	$PC(3, 4)$	***	**	ns	***
L_1^{1D}	$PC(1)$	***	***	ns	***
	$PC(2)$	**	ns	ns	**
	$PC(3)$	*	ns	ns	ns
	$PC(4)$	ns	ns	ns	ns
$AUROC$	$PC(1)$	***	***	ns	***
	$PC(2)$	**	ns	ns	**
	$PC(3)$	**	ns	ns	*
	$PC(4)$	ns	ns	ns	ns

6.2.3.4 Finding the most effective factor changing the snoring sounds characteristic

One can find the most affective factor (among body position, sleep stage, and blood oxygen level) by averaging the between class density measures of each factor (e.g. averaging supine vs. prone, supine vs. side, and prone vs. side distance measures for body position). We performed this for all three categorical variables and observed the results shown in Figure 6.13. Body position has the highest effect on snoring sounds characteristics and causes the highest separation between class densities.

The other two categorical variables discriminates the snoring sounds too but the

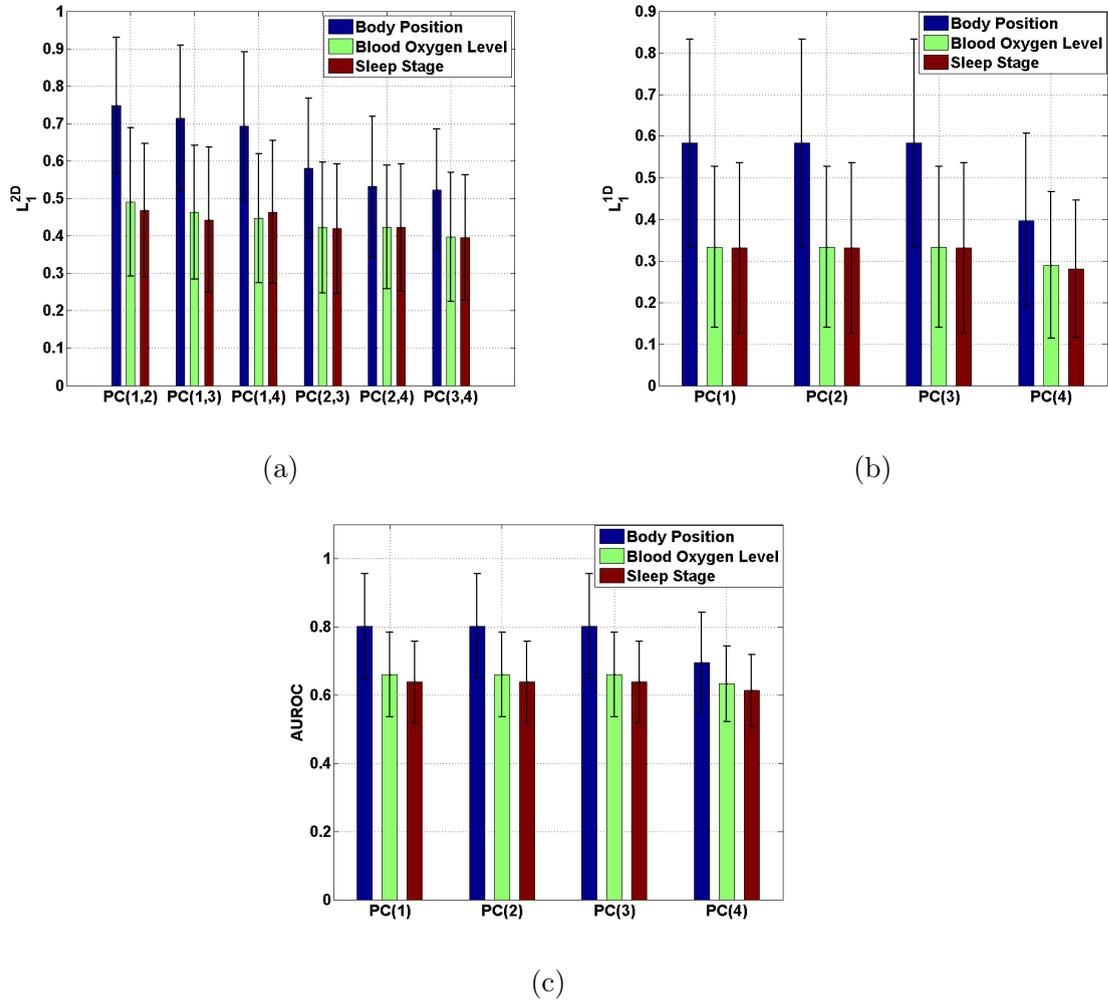


FIGURE 6.13: The mean and standard deviation of distance measures between class densities among all snorers grouped based on all categorical variables. (a). The mean and standard deviation of L_1^{2D} , the highest separation is for body position regardless of which bivariate has been used to estimate the 2-D pdf (highest $L_1^{2D} = 0.75 \pm 0.18$), then, for blood oxygen level ($L_1^{2D} = 0.49 \pm 0.2$), and finally, for sleep stage ($L_1^{2D} = 0.47 \pm 0.18$). (b) The mean and standard deviation of L_1^{1D} . (c) The mean and standard deviation of $AUROC$.

distance between class densities was not as high as that of body position. Using L_1^{2D} , the highest separation for body position, blood oxygen level, and sleep stage were 0.75 ± 0.18 , 0.49 ± 0.2 , and 0.47 ± 0.18 respectively. Using L_1^{1D} , the highest separation for body position, blood oxygen level, and sleep stage were 0.58 ± 0.25 , 0.33 ± 0.19 , and 0.33 ± 0.20 , respectively. Finally, using $AUROC$, the highest

separation for body position, blood oxygen level, and sleep stage were 0.80 ± 0.15 , 0.66 ± 0.12 , and 0.64 ± 0.12 , respectively. All three measures confirm that body position put the snoring sound segments into almost distinct clusters.

Table 6.4 shows the result of ANOVA on the distance measure of categorical variables. Columns 3-5 show the two by two comparison when we ran ANOVA to test if the mean of distance measures was significantly different between every possible pair of categorical variable. We also ran the ANOVA to test the equality of mean among all groups. The result shows that all distance measures are significantly higher if the snoring sounds were grouped based on body position than grouping snoring sounds based on sleep stage and blood oxygen level.

6.3 Discussion

In this part of the study, we investigated the effect of body position, sleep stage, and blood oxygen level on the characteristics of snoring sounds. It was observed all three parameters affected the snoring sounds to some extent, however, body position was found to be the most affective parameter that can even form distinct clusters of snoring sounds. In addition, we also looked at the effect of each specific class within the categorical variable. We observed that REM sleep stage had the higher effect on the snoring sounds compared to NREM and Arousal stages. We did not find any significant difference among different body positions' discrimination capability even though prone position seemed to be more capable of changing

snoring sounds compared to side and supine positions. Sleep stage and blood oxygen level had a moderate effect on snoring sounds on average.

We extracted a wide variety of sound features commonly used in all acoustical tasks of snoring sounds including snoring sound segmentation [99, 130, 137, 138], OSA diagnosis [64, 91, 122, 139, 140], and detection of site of obstruction [88, 141–143]. However, there existed some redundancy among the features and more importantly we faced a highly dimensional feature space. Therefore, to reduce the dimensionality and extract as much information as possible from the snoring sounds, we deployed the PCA technique to transform our feature vectors to a new informative and low dimensional feature (4-D) space.

Probability density function was used to fully describe the behavior of snoring sound segments in each category. In addition, we employed 2-D pdfs to characterize the snoring sounds more accurately. The results confirmed that using 2-D pdf, we achieved more accurate and consistent results than that of 1-D. Due to large number of snoring sound segments in each category, we did not encounter data insufficiency problem to estimate the pdfs. We used non-parametric kernal density estimation technique to obtain smooth pdfs for each category. We also used ROC curves to visualize the distance between class densities among snorers. The number of studies on the effect of body position, sleep stage, and blood oxygen level is very limited. In a previous study [19], the dependency between snoring and body position and sleep stage was investigated, however, that study was different from present study in method and purpose. They basically measured

the occurrence of snoring in different postures and sleep stages. Hence, to our best knowledge, there was no study investigating the effect of the aforementioned parameters on snoring sound.

In another study [144], it was shown that change of body position from prone to right or supine positions decreased (increased) the upper airway size in nonposition-dependent (position-dependent) OSA patients. Hence, in both cases the upper airway shape changed significantly. Our results are congruent with this previous finding as we had the highest change in acoustical characteristics of snoring sounds when turning body position from prone to other postures. To answer the question raised in the beginning of this chapter, the plausible causes of forming distinct clusters of snoring sounds within the same subject are body position, sleep stage, and oxygen saturation, however, body position is likely to be the most prominent factor.

The results of this study are important as they characterize the change of snoring sounds based on plausible states and factors. The proposed method can be used for other acoustical studies of snoring sounds as well as breathing sounds for classification and clustering purposes. In particular, a change in body position can be analogized to the change in the site of upper airway's obstruction. Hence, once the sound segments are matched in terms of body position the proposed method may be useful to detect the site of obstruction; this has been left for future studies.

Chapter 7

Summary and Concluding Remarks

7.1 Summary of Contributions

Acoustical analysis of snoring sounds is a non-invasive, inexpensive, and simple way to study the upper airway. To extract the clinical information from the snoring sounds, it is very crucial to have a reliable sound recording system and an automated algorithm to extract snoring segments from respiratory recordings. Snoring sounds are variable within and between snorers. This variability, if quantified appropriately, can reveal important information about the upper airway status. Moreover, acoustical analysis of snoring sounds can be used for OSA screening, detection of site of obstruction, snoring noise cancellation, and assessment of a treatment.

In this thesis, an unsupervised algorithm was proposed to extract the snoring sound segments from lengthy respiratory sound recording. The proposed algorithm

provides high accuracy, robustness and insensitivity to the degree of obstruction and low computational cost. It should be emphasized that an automatic, reliable, fast and accurate snoring sound extraction method is essential for diagnosis of SDB such as OSA. The proposed method can also be used as pre-processing tool in a variety of studies such as those that investigate the relationship between sleep stages and snoring sound characteristics, the relationship between body position and snoring sound intensity, and identify the person's best sleeping position, in which the least number of snoring segments occur.

We investigated statistical properties of snoring sound segments using HOS features. It was observed that snoring sounds were non-Gaussian and non-linear in general. Hence, using HOS feature complements the information extracted using conventional 2nd order methods. We also investigated the effect of anthropometric parameters on the snoring sounds. The results showed that there is an association between snoring sounds' characteristics and anthropometric parameters. Since the common features of snoring sounds used in classification are sensitive to anthropometric parameters, the results of classification may change when the two groups of apneic and controls are matched for those parameters. When HOS features were used in the classification of the apneic group, the results showed some improvements.

As mentioned, snoring sounds are variable. We measured the total sequential variability of the sounds and observed that the snoring sounds occurring during an apneic event (flow reduction) has higher variability than those occurring in the

absence of an apneic event. We compared and characterized the snoring sound segments among non-apneic, hypopneic, and post-apneic classes (using several features), and investigated the variability of snoring sound segments within each class (using total variation analysis). The results showed that the variability of snoring sound segments within each individual reflects on his/her OSA severity (using regression analysis).

Lastly, we investigated the effect of body position, sleep stage, and blood oxygen level on the characteristics of snoring sounds. It was observed all three parameters affected the snoring sounds to some extent, however, body position was found to be the most affective parameter that can even form distinct clusters of snoring sounds. The effect of sleep stage and blood oxygen level on snoring sounds characteristics were the same on average. We also observed that REM sleep stage had the higher effect on the snoring sounds compared to NREM and Arousal sleeps. We did not find any significant difference among different body positions' discrimination capability even though prone position seemed to be more capable of changing snoring sounds characteristics compared to side and supine positions.

7.2 Future Work Recommendations

Although many studies have investigated acoustical analysis of snoring sounds and its clinical application, it is still in its pioneering stage. The observation is a consequence of huge variability in the different stages of snoring sound analysis from data collection to a clinical diagnosis or assessment.

- There is not a unique definition of snoring sound in the literature. More importantly, lack of standardization for data collection setup (place of microphone), subjects under study (population size, matching their anthropometric parameters), analysis requirements, and etc. cause huge discrepancy among the published results and methods in the literature. Therefore there is a need for a guideline to standardize all stages of acoustical analysis of snoring sound.
- Apart from its clinical application, snoring sound can be used to improve the quality of sleep of snorer's bed partner. Developing a technique to adaptively reduce the social noise of snoring will be very valuable. An example would be to develop a device sensing the level of snoring sound and then based on that can appropriately change the shape of pillow or bed aiming to reduce the snoring noise.
- As known, the snoring sounds are variable within and between snorers. The snoring sound variability within snorers even during one night might be explained by the snoring generation mechanism and its variation as well as the level of obstruction and place of obstruction in the upper airway. However variability between snorers, in addition to previously-mentioned causes, may be a result of differences between upper airway structures of snorers. In fact, height, fat deposition, and other physiological factors may change the snoring sounds drastically. These factors are somehow reflected in the anthropometric parameters (such as height, BMI, and gender) of snorers.

Therefore one has to consider the effect of these parameters on snoring sounds before performing any acoustical tasks. In many published studies this effect has been ignored.

- The loud snoring sounds of habitual snorers (non-OSA or snorers without sleep disordered breathing) may convey crucial and important information on the upper airway status and snorer's respiratory system especially in children and women who snore less frequently than men.
- There have been a few studies on subjective and objective assessment of a snoring. Those studies reported a weak correlation between two types of assessment. Further investigation would be of interest to describe the difference between these two types of assessments. Moreover extracting information from snoring sounds that increases the correlation of two types of assessments would be important and need more work in this field.
- Detection of sites of obstruction using acoustical analysis of snoring sounds is still in its infancy. All the published studies have used induced sleep to investigate the ability of acoustical analysis while it has also been shown that the nocturnal snoring show a different properties than induced snoring. This field deserves further investigation. More importantly all studies recorded the snoring sounds from one location (e.g. over trachea) while multi-site recording of snoring sounds (over nose or in the ear) have not been investigated. Multi-site recordings can help extract more information from the

snoring sounds especially when there is more than one site of obstruction in the upper airway (mixed snoring).

Bibliography

- [1] E. Lugaresi, F. Cirignotta, C. Coccagna, and C. Piana, “Some epidemiological data on snoring and cardiorespiratory disturbances,” *Sleep*, vol. 3, pp. 221–224, 1980.
- [2] 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.
- [3] H. Bearpark, L. Elliott, R. Grunstein, S. Cullen, H. Schneider, W. Althaus, and C. Sullivan, “Snoring and sleep apnea. a population study in australian men,” *American Journal of Respiratory and Critical Care Medicine*, vol. 151, pp. 1459–1465, 1995.
- [4] I. Ayappa and D. Rapoport, “The upper airway in sleep: physiology of the pharynx,” *Sleep Medicine Reviews*, vol. 7, pp. 9–33, 2003.
- [5] M. M. Ohayon, C. Guilleminault, R. G. Priest, and M. Caulet, “Snoring and breathing pauses during sleep: Telephone interview survey of a united

- kingdom population sample,” *BMJ: British Medical Journal*, vol. 314, pp. 860–863, 1997.
- [6] C. Guilleminault, R. Stoohs, and S. Duncan, “Snoring (I). Daytime sleepiness in regular heavy snorers,” *Chest*, 1991.
- [7] E. Lugaresi, F. Cirignotta, and P. Montagna, “Pathogenic aspects of snoring and obstructive apnea syndrome,” *Schweiz Med Wochenschr*, 1988.
- [8] I. Gleadhill, A. Schwartz, N. Schubert, R. Wise, S. Permutt, and P. Smith, “Upper airway collapsibility in snorers and in patients with obstructive hypopnea and apnea,” *Am Rev Respir Dis*, 1991.
- [9] G. Liistro, D. Stanescu, C. Veriter, D. Rodenstein, and G. A. Tulken, “Pattern of snoring in obstructive sleep apnea patients and in heavy snorers,” *Sleep*, vol. 14, pp. 517–25, 1991.
- [10] D. W. Hudgel, “Mechanisms of obstructive sleep apnea.” *Chest*, 1992.
- [11] N. J. Douglas and O. Polo, “Pathogenesis of obstructive sleep apnoea/hypopnoea syndrome,” pp. 653–655, 1994.
- [12] T. Young, L. Finn, and M. Palta, “Chronic nasal congestion at night is a risk factor for snoring in a population-based cohort study,” *Archives of Internal Medicine*, vol. 161, pp. 1514–1519, 2001.
- [13] D. N. F. Fairbanks, “Uvulopalatopharyngoplasty complications and avoidance strategies.” *Otolaryngol Head Neck Surg.*, vol. 102, pp. 239–45, 1990.

-
- [14] J. R. Stradling and J. H. Crosby, "Predictors and prevalence of obstructive sleep apnoea and snoring in 1001 middle aged men." *Thorax*, vol. 46, pp. 85–90, 1991.
- [15] 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, pp. 1388–1395, 2002.
- [16] N. Gavriely and O. Jensen, "Theory and measurements of snores," *J Appl Physiol*, vol. 74, pp. 2828–2837, 1993.
- [17] B. Shome, L. P. Wang, M. H. Santare, A. K. Prasad, A. Z. Szeri, and D. Roberts, "Modeling of airflow in the pharynx with application to sleep apnea," *J Biomech Eng.*, vol. 120, pp. 416–22, 1998.
- [18] G. Liistro, D. Stanescu, and C. Veriter, "Pattern of simulated snoring is different through mouth and nose," *J.Appl.Phys.*, vol. 70, pp. 2736–41, 1991.
- [19] H. Nakano, T. Ikeda, M. Hayashi, E. Ohshima, and A. Onizuka, "Effects of body position on snoring in apneic and nonapneic snorers," *Sleep*, 2003.
- [20] T. Jones, M. Ho, J. Earis, A. Swift, and P. Charters, "Acoustic parameters of snoring sound to compare natural snores with snores during steady-state propofol sedation," *Clin.Otolaryngol.*, 2006.

-
- [21] S. J. Quinn, N. Daly, and P. D. Ellis, "Observation of the mechanism of snoring using sleep nasendoscopy," *Clin. Otolaryngol.*, vol. 20, pp. 360–4, 1995.
- [22] D. N. F. Fairbanks, S. A. Mickelson, and B. T. Woodson, *Snoring and obstructive sleep apnea*, 3rd ed. Lippincott Williams & Wilkins, 2003.
- [23] American Academy of Sleep Medicine, *The international classification of sleep disorders*, 2nd ed. Westchester : American Academy of Sleep Medicine, 2005.
- [24] G. Roldan and R. C. Ang, "Overview of sleep disorders," pp. 31–54, 2006.
- [25] R. Bhattacharjee, L. Kheirandish-Gozal, G. Pillar, and D. Gozal, "Cardiovascular complications of obstructive sleep apnea syndrome: evidence from children," *Prog Cardiovasc Dis.*, 2009.
- [26] R. S. T. Leung and T. D. Bradley, "Sleep apnea and cardiovascular disease," *American Journal of Respiratory and Critical Care Medicine*, vol. 164, pp. 2147–2165, 2001.
- [27] R. Rakel, "Clinical and societal consequences of obstructive sleep apnea and excessive daytime sleepiness," *Postgrad Med.*, vol. 1211, pp. 86–95, 2009.
- [28] 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, pp. 382–391, 2009.

- [29] J. Pagel, "Excessive daytime sleepiness," *Am Fam Physician*, vol. 79, pp. 391–396, 2009.
- [30] W. Flemons and W. Tsai, "Quality of life consequences of sleep-disordered breathing," *J Allergy Clin Immunol*, 1997.
- [31] F. Barbé, J. Pericas, A. Munoz, L. Findley, J. M. Anto, A. G. N. Agusti, and M. de Lluc Joan, "Automobile accidents in patients with sleep apnea syndrome. An epidemiological and mechanistic study," *Am.J.Respir.Crit.Care Med.*, vol. 158, pp. 18–22, 1998.
- [32] J. Terán-Santos, A. Jiménez-Gómez, 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, pp. 847–851, 1999.
- [33] 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, pp. 333–334, 2002.
- [34] O. J. Polo, M. Tafti, J. Fraga, K. V. K. Porkka, Y. Dejean, and M. Billiard, "Why don't all heavy snorers have obstructive sleep apnea?" *American Journal of Respiratory and Critical Care Medicine*, vol. 143, pp. 1288–1293, 1991.
- [35] P. M. Suratt, S. C. Wilhoit, and K. Cooper, "Induction of airway collapse with subatmospheric pressure in awake patients with sleep apnea," *Journal of applied physiology*, vol. 57, pp. 140–146, 1984.

- [36] D. P. White, "Occlusion pressure and ventilation during sleep in normal humans," *Journal of applied physiology*, vol. 61, pp. 1279–1287, 1986.
- [37] P. Norton and E. Dunn, "Snoring as a risk factor for disease: an epidemiological survey," *Br Med J*, vol. 291, pp. 630–632, 1985.
- [38] M. Koskenvuo, M. Partinen, S. Sarna, J. Kaprio, H. Langinvainio, and K. Heikkilä, "Snoring as a risk factor for hypertension and angina pectoris," *The Lancet*, vol. 325, pp. 893–896, 1985.
- [39] P. E. Peppard, T. Young, M. Palta, and J. Skatrud, "Prospective study of the association between sleep-disordered breathing and hypertension," *N Engl J Med*, vol. 342, pp. 1378–84, 2000.
- [40] P. Cole and H. Miljeteig, "Snoring : a review and a reassessment," 1995.
- [41] R. Taugerbeck, "Comparison of dental intraoral devices for snoring therapy: A subjective survey of their effect on mixed sleep apnea syndrome," *Sleep and Breathing*, vol. 2, pp. 20–31, 1997.
- [42] W. Engelke, W. Engelhardt, M. Mendoza-Gärtner, O. Deccó, J. Barrirero, and M. Knösel, "Functional treatment of snoring based on the tongue-repositioning manoeuvre," *The European Journal of Orthodontics*, vol. 32, pp. 490–495, 2010.
- [43] A. Labra, A. Huerta-Delgado, C. Gutierrez-Sanchez, S. A. Cordero-Chacon, and P. Basurto-Madero, "Uvulopalatopharyngoplasty and uvulopalatal flap

- for the treatment of snoring : Technique to avoid complications,” *J. Otolaryngol Head Neck Surg.*, vol. 37, pp. 256–59, 2008.
- [44] T. Jones, M. Ho, J. Earis, and A. Swift, “Acoustic parameters of snoring sound to assess the effectiveness of sleep nasendoscopy in predicting surgical outcome,” *Otolaryngol.Head Neck Surg.*, vol. 135, pp. 269–75, 2006.
- [45] T. Jones, P. Walker, M. Ho, J. Earis, A. Swift, and P. Charters, “Acoustic parameters of snoring sound to assess the effectiveness of the muller manoeuvre in predicting surgical outcome,” *Auris Nasus Larynx*, vol. 33, pp. 409–16, 2006.
- [46] A. J. Smithson, J. E. S. White, C. J. Griffiths, A. J. N. Prichardt, P. R. Close, M. J. Drinnan, H. F. Marshallt, and G. J. Gibson, “Comparison of methods for assessing snoring,” *Clin.Otolaryngol*, vol. 20, pp. 443–7, 1995.
- [47] M. M. Abo-Khatwa, E. Z. Osman, P. D. Hill, B. W. V. Lee, and J. E. Osborne, “Objective evaluation of tongue base snoring after the use of an oral appliance: a prospective case series,” *Clin.Otolaryngol*, vol. 33, pp. 592–5, 2008.
- [48] K. Wilson, T. Mulrooney, and R. R. Gawtry, “Snoring: An acoustic monitoring technique,” *The Laryngoscope*, vol. 95, pp. 1174–1177, 1985.
- [49] K. Wilson, R. A. Stoohs, T. F. Mulrooney, L. J. Johnson, C. Guilleminault, and Z. Huang, “The snoring spectrum: Acoustic assessment of snoring sound

- intensity in 1139 individuals undergoing polysomnography,” *Chest*, vol. 115, pp. 762–770, 1999.
- [50] T. Jones, A. Swift, P. Calverley, M. Ho, and J. Earis, “Acoustic analysis of snoring before and after palatal surgery,” *Eur. Respir. J.*, vol. 25(6), pp. 1044–9, 2005.
- [51] P. Hill, B. Lee, J. Osborne, and E. Osman, “Palatal snoring identified by acoustic crest factor analysis,” *Physiol. Meas.*, vol. 20, pp. 167–74, 1999.
- [52] A. Yadollahi and Z. Moussavi, “Formant analysis of breath and snore sounds,” in *IEEE-EMBS*, Minneapolis, MN, 2009.
- [53] 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, pp. 130–3, 1999.
- [54] T. Jones, M. Ho, J. Earis, and A. Swift, “Acoustic parameters of snoring sound to assess the effectiveness of sleep nasendoscopy in predicting surgical outcome,” *Otolaryngol. Head Neck Surg.*, vol. 135, pp. 269–75, 2006.
- [55] T. Jones, P. Walker, M. Ho, J. Earis, A. Swift, and P. Charters, “Acoustic parameters of snoring sound to assess the effectiveness of the muller manoeuvre in predicting surgical outcome,” *Auris Nasus Larynx*, vol. 33, pp. 409–16, 2006.

- [56] 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.
- [57] 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.
- [58] A. Yadollahi and Z. Moussavi, "Acoustic obstructive sleep apnea detection," in *IEEE-EMBS*, Minneapolis, MN, 2009.
- [59] 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.
- [60] C. Kushida, M. Littner, T. Morgenthaler, C. Alessi, D. Bailey, J. J. Coleman, L. Friedman, M. H. adn 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, pp. 499–521, 2005.
- [61] S. P. Patil, H. Schneider, A. R. Schwartz, and P. L. Smith, "Adult obstructive sleep apnea*," *Chest*, vol. 132, pp. 325–337, 2007.
- [62] S. Isono, J. Remmers, A. Tanaka, Y. Sho, J. Sato, and T. Nishino, "Anatomy of pharynx in patients with obstructive sleep apnea and in normal subjects," *J Appl Physiol*, 1997.

-
- [63] A. McCombe, V. Kwok, and W. Hawke, "An acoustic screening test for obstructive sleep apnoea," *Clinical otolaryngology and allied sciences*, vol. 20, pp. 348–351, 1995.
- [64] U. R. Abeyratne, C. K. K. Patabandi, and K. Puvanendran, "Pitch-jitter analysis of snoring sounds for the diagnosis of sleep apnea," in *IEEE-EMBS*, vol. 2, 2001, pp. 2072–2075.
- [65] J. Sola-Soler, R. Jane, J. Fiz, and J. Morera, "Pitch analysis in snoring signals from simple snorers and patients with obstructive sleep apnea," in *IEEE-EMBS*, (Houston, TX), 2002.
- [66] 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, pp. 951–957, 2004.
- [67] 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.
- [68] A. Ng, T. Koh, E. Baey, and K. Puvanendran, "Could formant frequencies of snore signals be an alternative means for the diagnosis of obstructive sleep apnea?" *Oral Platform Presentations / Sleep Medicine*, vol. 8, pp. 894–8, 2007.

- [69] A. Yadollahi, E. Giannouli, and Z. Moussavi, "Sleep apnea monitoring and diagnosis based on pulse oximetry and tracheal sound signals," *Medical and Biological Engineering and Computing*, vol. 48, pp. 1087–1097, 2010.
- [70] U. Abeyratne, A. Karunajeewa, and C. Hukins, "Mixed-phase modeling in snore sound analysis," *Med Bio Eng Comput*, vol. 45, pp. 791–806, 2007.
- [71] J. G. Proakis and D. K. Manolakis, *Digital Signal Processing*, 4th ed. Prentice Hall, 2006.
- [72] 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, pp. 635–644, 1993.
- [73] R. Beck, M. Odeh, A. Oliven, and N. Gavriely, "The acoustic properties of snores," *Eur Respir J.*, vol. 8, pp. 2120–2128, 1995.
- [74] 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.
- [75] 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.

-
- [76] 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.
- [77] A. Karunajeewa, U. Abeyratne, and C. Hukins, “Silence-breathing-snore classification from snore-related sounds,” *Physiol Meas.*, vol. 29, pp. 227–43, 2008.
- [78] 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.
- [79] A. Yadollahi and Z. Moussavi, “Automatic breath and snore sounds classification from tracheal and ambient sounds recordings,” *Medical Engineering and Physics*, vol. In Press, Corrected Proof, 2010.
- [80] E. Rafajłowicz, M. Pawlak, and A. Steland, “Nonparametric sequential change-point detection by a vertically trimmed box method,” *IEEE Transactions on Information Theory*, vol. 56, pp. 3621–3634, 2010.
- [81] L. Jolliffe, *Principal Component Analysis*, 2nd ed. New York: Springer-Verlag, Inc, 2002.
- [82] J. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum, 1981.

-
- [83] J. Bezdek, R. Ehrlich, and W. Full, "Fcm: the fuzzy c-means clustering algorithm," *Computers and Geosciences*, vol. 10, pp. 191–203, 1984.
- [84] F. Dalmaso and R. Prota, "Snoring: analysis, measurement, clinical implications and applications," *Eur Respir J*, vol. 9, pp. 146–159, 1996.
- [85] J. Dunn, "A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters," *Journal of Cybernetics*, vol. 3, pp. 32–57, 1973.
- [86] U. Kaymak and M. Setnes, "Fuzzy clustering with volume prototype and adaptive cluster merging," *IEEE Trans. on Fuzzy Systems*, vol. 10, pp. 705–712, 2002.
- [87] 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, pp. 130–3, 1999.
- [88] P. Hill, B. Lee, J. Osborne, and E. Osman, "Palatal snoring identified by acoustic crest factor analysis," *Physiol. Meas.*, vol. 20, pp. 167–74, 1999.
- [89] T. Jones, A. Swift, P. Calverley, M. Ho, and J. Earis, "Acoustic analysis of snoring before and after palatal surgery," *Eur. Respir. J.*, vol. 25, pp. 1044–9, 2005.

- [90] 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.
- [91] A. Ng, T. Koh, U. Abeyratne, and K. Puvanendran, "Investigation of obstructive sleep apnea using nonlinear mode interactions in nonstationary snore signals," *Annals of Biomedical Engineering*, vol. 37, pp. 1796–1806, 2009.
- [92] F. Dalmaso and R. Prota, "Snoring: analysis, measurement, clinical implications and applications," *Eur Respir J*, vol. 9, pp. 146–159, 1996.
- [93] J. W. Fackrell, "Bispectral analysis of speech signals," Ph.D. dissertation, University of Edinburgh, Edinburgh, UK, 1996.
- [94] A. K. Ng, K. Y. Wong, C. H. Tan, and T. S. Koh, "Bispectral analysis of snore signals for obstructive sleep apnea detection," in *Proceedings of the 29th Annual International Conference of the IEEE EMBS*, 2007, pp. 6195–6198.
- [95] M. G. Kendall, "A new measure of rank correlation," *Biometrika*, vol. 30, pp. 81–93, 1938.
- [96] P. H. Kvam and B. Vidakovic, *Nonparametric Statistics with Applications to Science and Engineering*. Wiley-Interscience, 2007.
- [97] S. R. Searle, *Linear Models*. New York: John Wiley and Sons, Inc, 1971.

-
- [98] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. Wiley-Interscience, 2000.
- [99] A. Azarbarzin and Z. Moussavi, “Automatic and unsupervised snore sound extraction from respiratory sound signals,” *IEEE Trans Biomed Eng.*, vol. 58, pp. 1156 – 1162, 2011.
- [100] D. R. Brillinger, “An introduction to polyspectra,” *The Annals of Mathematical Statistics*, vol. 36, pp. 1351–1374, 1965.
- [101] M. J. Hinich, “Testing for gaussianity and linearity of a stationary time series,” *J. Time Series Analysis*, vol. 3, pp. 169–76, 1982.
- [102] D. R. Brillinger and M. Rosenblatt, “Computation and interpretation of k-th order spectra,” *Spectral Analysis of Time Series*, pp. 189–232, 1967.
- [103] M. J. Hinich and M. A. Wolinsky, “A test for aliasing using bispectral analysis,” *Journal of the American Statistical Association*, vol. 83, pp. 499–502, 1988.
- [104] L. Devroye and L. Györfi, *Nonparametric density estimation: the L1 view*. Wiley, 1985.
- [105] A. Swami, J. M. Mendel, and C. L. Nikias, “Higher-order spectral analysis toolbox user’s guide,” The MathWorks, Tech. Rep. version2, 1998.

-
- [106] J. Markel, "Digital inverse filtering—a new tool for formant trajectory estimation," *IEEE Transactions on Audio and Electroacoustics*, vol. 20, pp. 129–137, 1972.
- [107] R. Hogg, A. Craig, and J. Mckean, *Introduction to Mathematical Statistics*. Prentice Hall, 2004.
- [108] B. W. Silverman, *Density estimation for statistics and data analysis*. London ; New York: Chapman and Hall, 1986.
- [109] P. Burman, "A comparative study of ordinary cross-validation, v-fold cross-validation and the repeated learning-testing methods," *Biometrika*, vol. 76, pp. 503–514, 1989.
- [110] J. H. Venter and J. L. J. Snyman, "A note on the generalised cross-validation criterion in linear model selection," *Biometrika*, vol. 82, pp. 215–219, 1995.
- [111] I. Sanchez and H. Pasterkamp, "Tracheal sound spectra depend on body height," *Am Rev Respir Dis.*, vol. 148, pp. 1083–1087, 1993.
- [112] 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.
- [113] T. Young, M. Palta, J. Dempsey, J. Skatrud, S. Weber, and S. Badr, "The occurrence of sleep-disordered breathing among middle-aged adults," *New England Journal of Medicine*, vol. 328, pp. 1230–1235, 1993.

- [114] I. L. Mortimore, I. Marshall, P. K. Wraith, R. J. Sellar, and N. J. Douglas, "Neck and total body fat deposition in nonobese and obese patients with sleep apnea compared with that in control subjects," *American Journal of Respiratory and Critical Care Medicine*, vol. 157, pp. 280–283, 1998.
- [115] A. G. de Sousa, C. Cercato, M. C. Mancini, and A. Halpern, "Obesity and obstructive sleep apnea-hypopnea syndrome," *Obes Rev*, vol. 9, pp. 340–354, 2008.
- [116] N. Gavriely, M. Nissan, A. H. Rubin, and D. W. Cugell, "Spectral characteristics of chest wall breath sounds in normal subjects." *Thorax*, vol. 50, pp. 1292–1300, 1995.
- [117] V. Gross, A. Dittmar, T. Penzel, F. Schuttler, and P. von Wichert, "The relationship between normal lung sounds, age, and gender," *American Journal of Respiratory and Critical Care Medicine*, vol. 162, pp. 905–909, 2000.
- [118] D. P. White, R. M. Lombard, R. J. Cadieux, and C. W. Zwillich, "Pharyngeal resistance in normal humans: influence of gender, age, and obesity," *Journal of applied physiology*, vol. 58, pp. 365–371, 1985.
- [119] L. J. Brooks and K. P. Strohl, "Size and mechanical properties of the pharynx in healthy men and women," *Am Rev Respir Dis*, vol. 146, pp. 1394–1397, 1992.
- [120] R. Schwab, K. Gupta, W. Gefter, L. Metzger, E. Hoffman, and A. Pack, "Upper airway and soft tissue anatomy in normal subjects and patients with

- sleep-disordered breathing. significance of the lateral pharyngeal walls,” *Am J Respir Crit Care Med.*, vol. 152, pp. 1673–1689, 1995.
- [121] A. K. Ng, T. S. Koh, E. Baey, and K. Puvanendran, “Diagnosis of obstructive sleep apnea using formant features of snore signals,” in *IFMBE Proceedings*, ser. IFMBE Proceedings, R. Magjarevic, R. Magjarevic, and J. H. Nagel, Eds., vol. 14, Nanyang Technological University Nanyang Avenue Singapore 639798 Singapore. Springer Berlin Heidelberg, 2007, pp. 967–970.
- [122] H. Michael, S. Andreas, B. Thomas, H. Beatrice, H. Werner, and K. Holger, “Analysed snoring sounds correlate to obstructive sleep disordered breathing,” *European Archives of Oto-Rhino-Laryngology*, vol. 265, pp. 105–113, 2008.
- [123] C. R. Vogel and M. E. Oman, “Iterative methods for total variation denoising,” *SIAM Journal on Scientific Computing*, vol. 17, pp. 227–238, 1996.
- [124] L. Rabiner and R. Schafer, *Digital processing of speech signals*. Englewood Cliffs, NJ: Prentice-Hall, 1978.
- [125] M. E. Cohen, D. L. Hudson, and P. C. Deedwania, “Applying continuous chaotic modeling to cardiac signal analysis,” *Engineering in Medicine and Biology Magazine, IEEE*, vol. 15, pp. 97–102, 1996.
- [126] R. Duda, P. Hart, and D. Stork, *Pattern Classification*. John Wiley and Sons, Inc, 2001.

- [127] M. Younes, M. Ostrowski, R. Atkar, J. Laprairie, A. Siemens, and P. Hanly, “Mechanisms of breathing instability in patients with obstructive sleep apnea,” *Journal of applied physiology*, vol. 103, pp. 1929–1941, 2007.
- [128] M. Tvinnereim, P. Cole, J. S. Haight, and V. Hoffstein, “Diagnostic airway pressure recording in sleep apnea syndrome,” *Acta Otolaryngol*, vol. 115, pp. 449–454, 1995.
- [129] 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,” *Physiological Measurement*, vol. 29, p. 879, 2008.
- [130] 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.
- [131] W. H. Tsai, W. W. Flemons, W. A. Whitelaw, and J. E. Remmers, “A comparison of apnea–hypopnea indices derived from different definitions of hypopnea,” *American Journal of Respiratory and Critical Care Medicine*, vol. 159, pp. 43–48, 1999.
- [132] X. Huang, A. Acero, and H. Hon, *Spoken Language Processing: A Guide to Theory, Algorithm and System Development*. Prentice Hall PTR, 2001.
- [133] T. Inouye, K. Shinosaki, H. Sakamoto, S. Toi, S. Ukai, A. Iyama, Y. Katsuda, and M. Hirano, “Quantification of eeg irregularity by use of the entropy of

- the power spectrum,” *Electroencephalography and clinical neurophysiology*, vol. 79, pp. 204–210, 1991.
- [134] M. P. Wand and M. C. Jones, *Kernel Smoothing*. London: Chapman and Hall/CRC, 1994.
- [135] S. Theodoridis and K. Koutroumbas, *Pattern Recognition, Forth Edition*. USA: Academic Press, 2008.
- [136] K. E. Atkinson, *An introduction to numerical analysis*. Wiley, 1978.
- [137] A. Karunajeewa, U. Abeyratne, and C. Hukins, “Silence-breathing-snore classification from snore-related sounds,” *Physiol Meas.*, vol. 29, pp. 227–43, 2008.
- [138] 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.
- [139] 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.
- [140] 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.*, 2008.

- [141] S. Miyazaki, Y. Itasaka, K. Ishikawa, and K. Togawa, "Acoustic analysis of snoring and the site of airway obstruction in sleep related respiratory disorders," *Acta Otolaryngol.*, vol. 537, pp. 47–51, 1998.
- [142] 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.*, 2002.
- [143] R. J. Beeton, I. Wells, P. Ebdon, H. B. Whittet, and J. Clarke, "Snore site discrimination using statistical moments of free field snoring sounds recorded during sleep nasendoscopy," *Physiol. Meas.*, vol. 28, pp. 1225–1236, 2007.
- [144] D. Pevernagie, A. Stanson, P. Sheedy, B. Daniels, and J. Shepard, "Effects of body position on the upper airway of patients with obstructive sleep apnea," *Am J Respir Crit Care Med*, 1995.
- [145] 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.
- [146] 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.
- [147] A. Azarbarzin and Z. Moussavi, "Nonlinear properties of snoring sounds," in *The 36th International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011.

-
- [148] M. Herzog, T. Kühnel, T. Bremert, B. Herzog, W. Hosemann, and H. Kaftan, “The impact of the microphone position on the frequency analysis of snoring sounds,” *European Archives of Oto-Rhino-Laryngology*, vol. 266, pp. 1315–1322, 2009.

Appendices

Appendix A

Comparison between Site of Recording

A.1 Comparison between Sites of Recording

Currently there is no standard recommending the place of microphone for different acoustical applications of snoring sound. The place of microphone is highly variable among published literature. Locations such as larynx (attached to the skin), trachea (attached to the skin), inside nasal cannula, or in the air around the snorer's mouth have been proposed in the related literature [73, 99, 145–147]. In most studies, the snoring sounds were recorded from either over the trachea (by tracheal microphone: a microphone attached to the skin) or a point near the mouth of snorer (ambient microphone). Although the sounds picked up by two

microphones have the same sources and nature, there are some differences between their acoustical characteristics such as intensity, frequency contents, and classification accuracy.

In an early study [73], the authors compared the acoustical properties such as shape of waveform and frequency contents of simulated snoring sounds recorded from two sites: 1) 2 cm above and slightly to the right of the suprasternal notch by a piezoelectric contact sensor and 2) 20 cm away from the mouth by an electret-type condenser microphone. In fact, they asked four healthy male non-snoring participants to simulate snoring. The results of this study showed that although the sound waves had different shape, both had similar repetitive complex structures with the identical frequency and no phase shift [73]. In another study [145], snoring sounds were simultaneously recorded by two electret microphones: an omnidirectional microphone placed 10 cm away from participant's mouth, and a unidirectional microphone placed over the trachea at the level of cricoid cartilage (1 cm lateral to the median line). The data set for this study included several simulated snoring from three healthy participants and five minutes real snoring from three snorers. The intensity of ambient and tracheal snoring sounds were calculated and compared. The authors observed a high correlation between intensity values of tracheal and ambient recordings for simulated snoring sounds. They also found that the mean value of intensity of snoring sounds was higher for tracheal recordings than ambient ones. This result was not consistent for all subjects [145]. Influence of different microphone positions on the frequency features

of snoring sounds was investigated in [148]. They recorded the sounds from six positions: three points in vicinity of snorers' head and three contact microphones (body, neck, and parasternal) in five patients. The results of frequency analysis revealed a wider range of frequency in air microphones compared to contact microphones. They concluded that the contact microphones were good candidates for screening devices while ambient microphones were preferable for a natural analysis of snoring sounds [148].

The sensitivity of tracheal microphone does not depend on the body movement during sleep while ambient microphone may be dramatically influenced by the body movement and the worst case would be when the patient is in the prone posture. On the other hand, tracheal microphone may record all the movement artifacts. This increases the preprocessing time of the signal due to those artifacts. Other issues related to location of recording is the ability of keeping information underlying the recorded signal. In this appendix, the tracheal and ambient sites of recording are compared in terms of sub-band energy and ability to extract snoring sound segments from respiratory sounds.

A.2 Sub-band Energy of Ambient and Tracheal Recordings

We compared the sensitivity of two microphones in the first four 500Hz frequency sub-bands (introduced in Section 3.2.4) by taking ratio of normalized energy between two simultaneous recordings. First, for every sound segment we calculated the ratio of normalized energy for four sub-bands. Then, average of the ratios were found within each participant resulting in four values per participant. Figure A.1 shows the mean and variations of the ratios among all participants with both recordings. As shown in Figure A.1, for lower frequencies less than 1000 Hz, tracheal microphone is more sensitive than ambient microphone; however, the sensitivity of tracheal microphone significantly decreases for frequencies higher than 1000Hz.

A.3 Extraction of Snoring Sound Segment from Respiratory Sounds

This section compares the result of tracheal and ambient recordings in terms of discrimination between snoring and no-snoring sound segments as discussed in Chapter 3. Tables A.1 and A.2 show the results snoring sound detection algorithm on the tracheal and ambient recordings, respectively. As can be seen, the overall accuracy and PPV of the proposed snore detection algorithm is 98.6% and

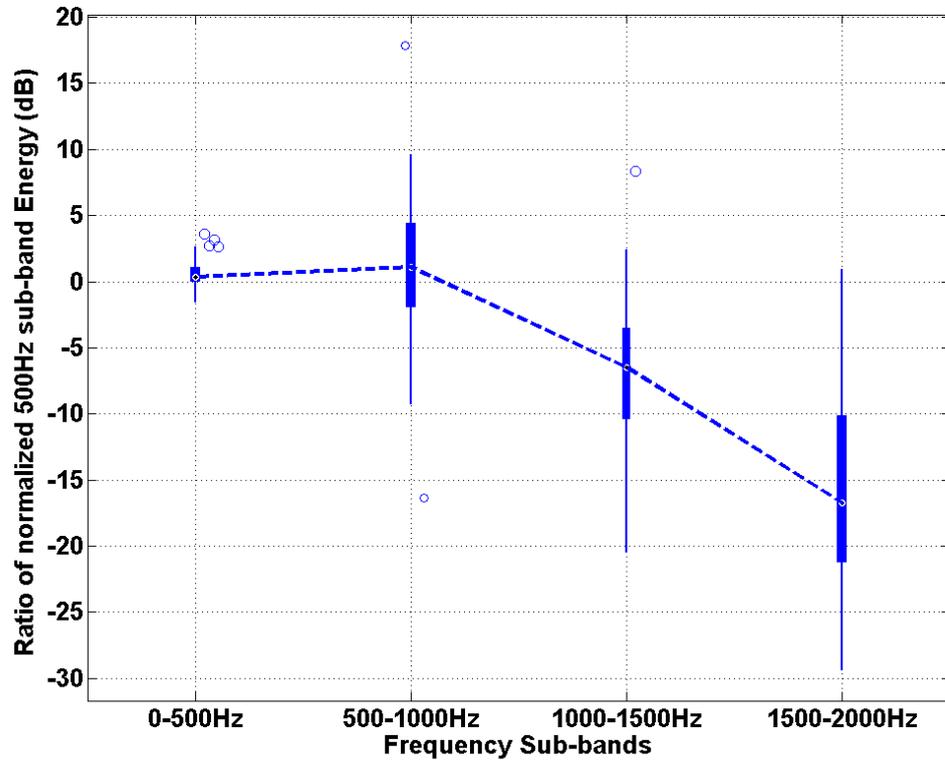


FIGURE A.1: The ratio of normalized energy between ambient and tracheal recordings

94.8% for the tracheal recordings, while they are 93.1% and 95.9% for the ambient recordings. The performance of the algorithm for tracheal recordings remained more and less the same when it was applied only to data of either OSA patients (98.8%) or simple snorers (98.4%). On the other hand, for the ambient recordings, the variation in the accuracy and PPV was larger than those of the tracheal recordings. In case of using the ambient microphone, the accuracy was 3.3% higher when the method was applied to data of simple snorers, and it dropped by 1.5% when data of the OSA patients was used.

TABLE A.1: Classification results for tracheal recordings.

Data Set	TP	FP	FN	Accuracy	PPV
Simple Snorers and OSA	5588	304	77	98.6	94.8
OSA	3816	203	48	98.8	94.9
Simple Snorers	1772	101	29	98.4	94.6

TABLE A.2: Classification results for ambient recordings.

Data Set	TP	FP	FN	Accuracy	PPV
Simple Snorers and OSA	5275	223	390	93.1	95.9
OSA	3569	104	327	91.6	97.2
Simple Snorers	1706	119	63	96.4	93.5

A.4 Conclusion

It was found that the energy of tracheal recordings is concentrated below 1000 Hz. On the other hand, although a high percentage of the sound energy of ambient recording is concentrated below 1000 Hz, it still has some components above 1000 Hz which were not detectable by tracheal recordings probably due to the chamber in which it is inserted to before being placed over the skin. On the other hand, tracheal microphone is more sensitive to low frequency components. As a result, the sound features extracted from the tracheal tracings were more discriminative between three groups than those of ambient recordings while using low frequency features. Overall, we suggest the tracheal microphone as a better choice due to four reasons: 1) insensitivity of tracheal microphone with respect to body posture, 2) concentration of a high percentage of the snoring sounds energy below 1000 Hz in which the tracheal microphone is more sensitive than ambient microphone, 3) the amount of environmental noise is higher in ambient microphone than that of tracheal microphone, and 4) ability to detect snoring sound segments more

accurately.

Appendix B

Effect of Body Position and Sleep Stage with Original Classes

This appendix discusses the results of comparison between class pdf of snoring sounds for body position and sleep stage without merging the right and left classes for body position, and stage 1-4 classes for sleep stage. The analysis is presented in Chapter 6.

B.1 The effect of body position with 4 classes

Based on PSG score sheet, every snorer had 4 possible Body positions namely *supine*, *prone*, *left*, and *right*. Similar to Chapter 6, the snoring sound segments of each snorer were categorized based on body position and then both 1-D and 2-D pdfs were estimated for each class and finally the overlap between classes were quantified using L_1^{1D} , L_1^{2D} , and $AUROC$. Figure B.1 shows the mean and standard

deviation of all distance measures among snorers. As shown in Figure B.1a, on average, turning body from prone to supine or right causes the highest change in snoring sounds' characteristics, while turning from left to right or vice-versa causes the lowest change in snoring sounds' characteristics. This is congruent with the result of Chapter 6 where the left and right positions were merged into side position.

B.2 The effect of sleep stage with 6 classes

As mentioned in Chapter 6, sleep stage included 6 classes before merging: *stage 1-4*, *Arousal*, and *REM*. The same procedure as in Chapter 6 was repeated here, and the distance measures (L_1^{1D} , L_1^{2D} , and *AUROC*) between different class densities were calculated. Figure B.2 shows the the mean and standard deviation of all distance measures among snorers. On average, snoring in stage 3 has the highest separation with snoring in REM. In fact, REM sleep seems to have highest effect on the snoring sounds' characteristics, while stage 2 and Arousal have the highest similarity among other comparisons. This is also congruent with the result of Chapter 6 where the stages 1 to 4 were merged into NREM sleep.

B.2.0.5 Finding the most effective factor changing the snoring sounds characteristic

If the body position and sleep stage classes were not merged into new classes the overlap between class densities would be different and as a result the comparison

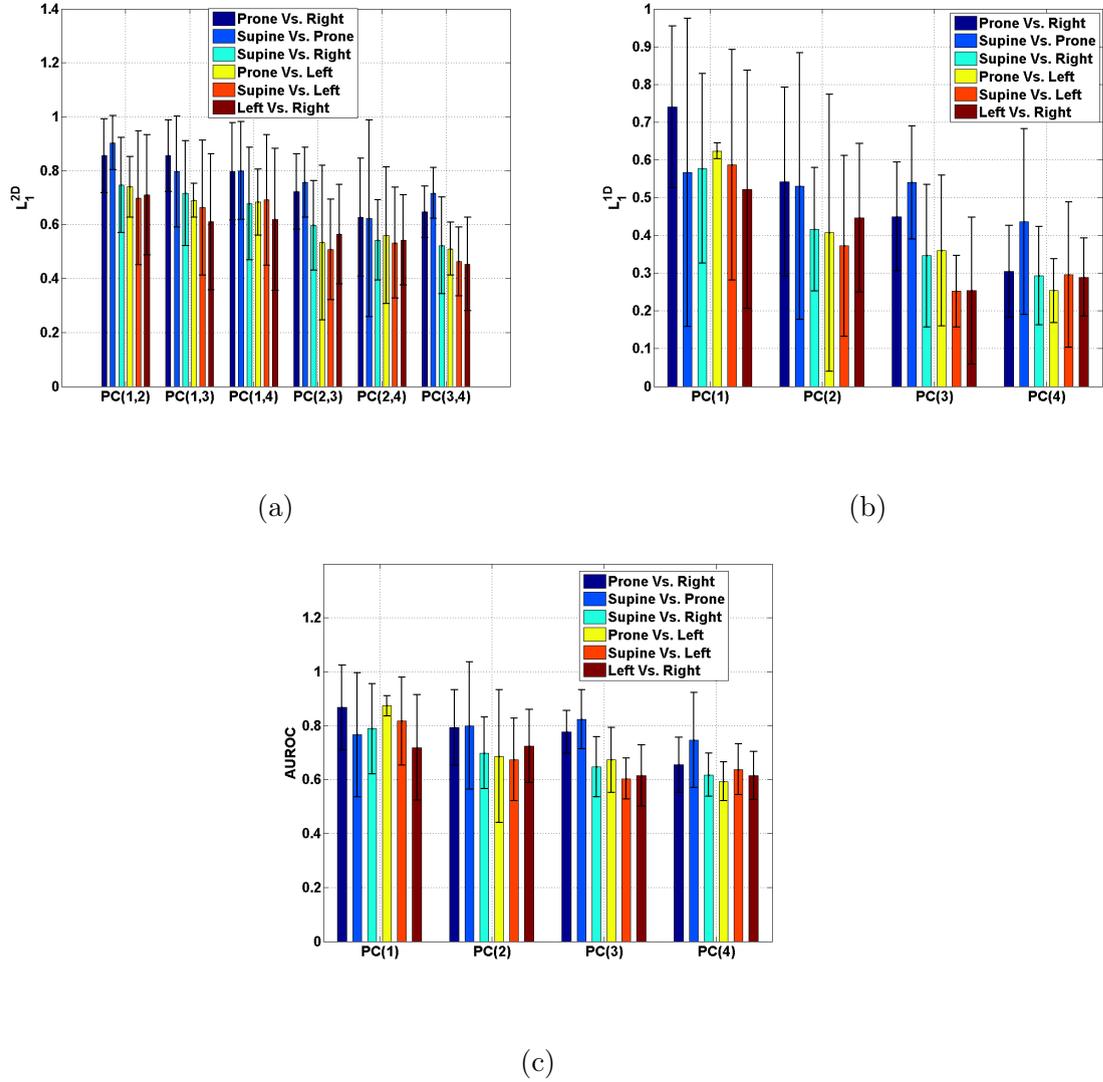


FIGURE B.1: The mean and standard deviation of distance measures between body position classes' pdfs among all snorers (a). The mean and standard deviation of L_1^{2D} , the highest separation is between prone and supine positions, then, between prone and right, and the lowest is between left and right positions. (b) The mean and standard deviation of L_1^{1D} . (c) The mean and standard deviation of *AUROC*.

among three categorical variables would be different as well. Figure B.3 shows the mean and standard deviation of distance measures for each categorical variable. Although the results slightly changed compared to Section 6.2.3.4, Figure B.3 shows that body position has the highest effect on snoring sounds compared to

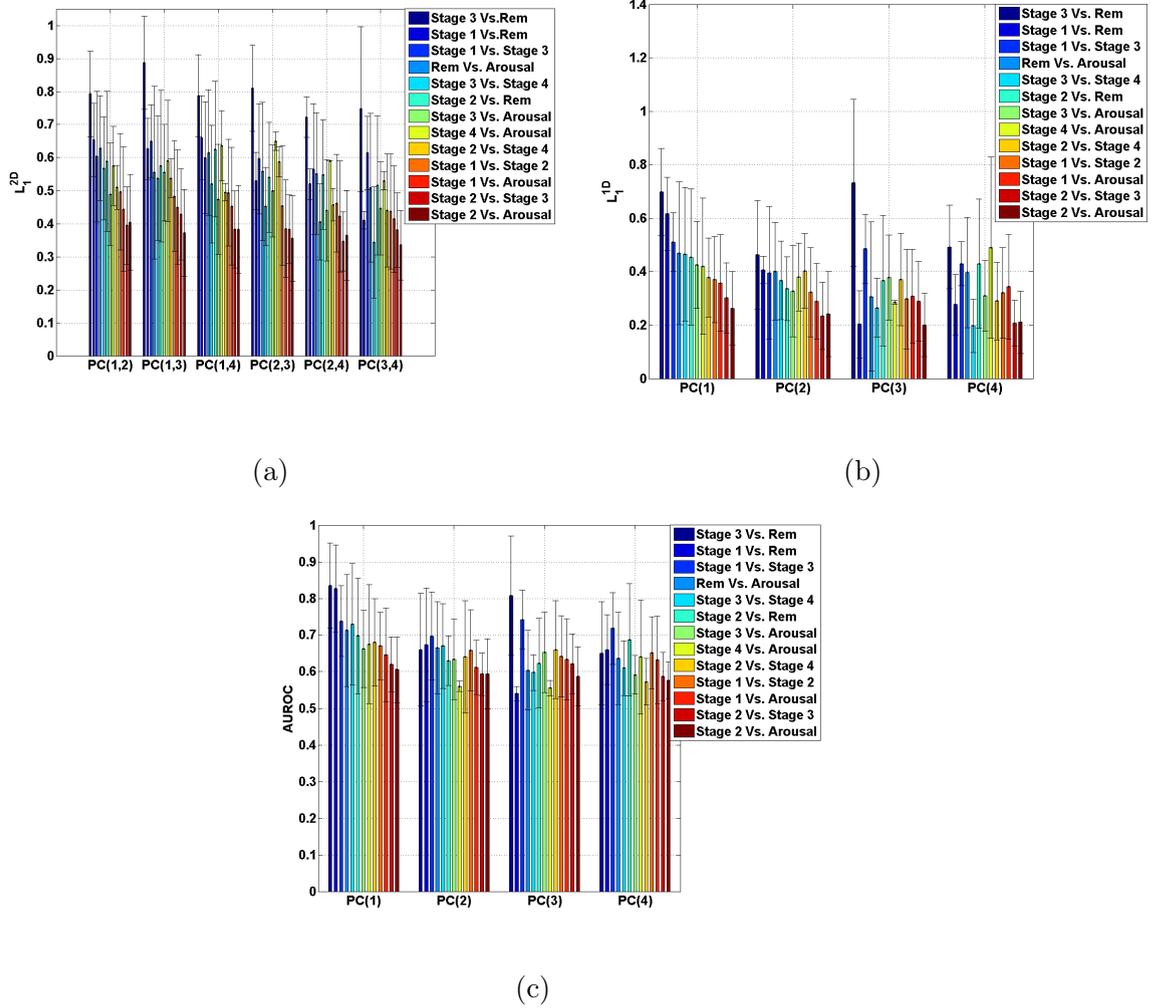


FIGURE B.2: The mean and standard deviation of distance measures between sleep stage classes' pdfs among all snorers (a). The mean and standard deviation of L_1^{2D} , the highest separation is between REM and stage 3, then, between REM and stage 1, and the lowest is between stage 2 and Arousal. (b) The mean and standard deviation of L_1^D . (c) The mean and standard deviation of $AUROC$.

sleep stage and blood oxygen level.

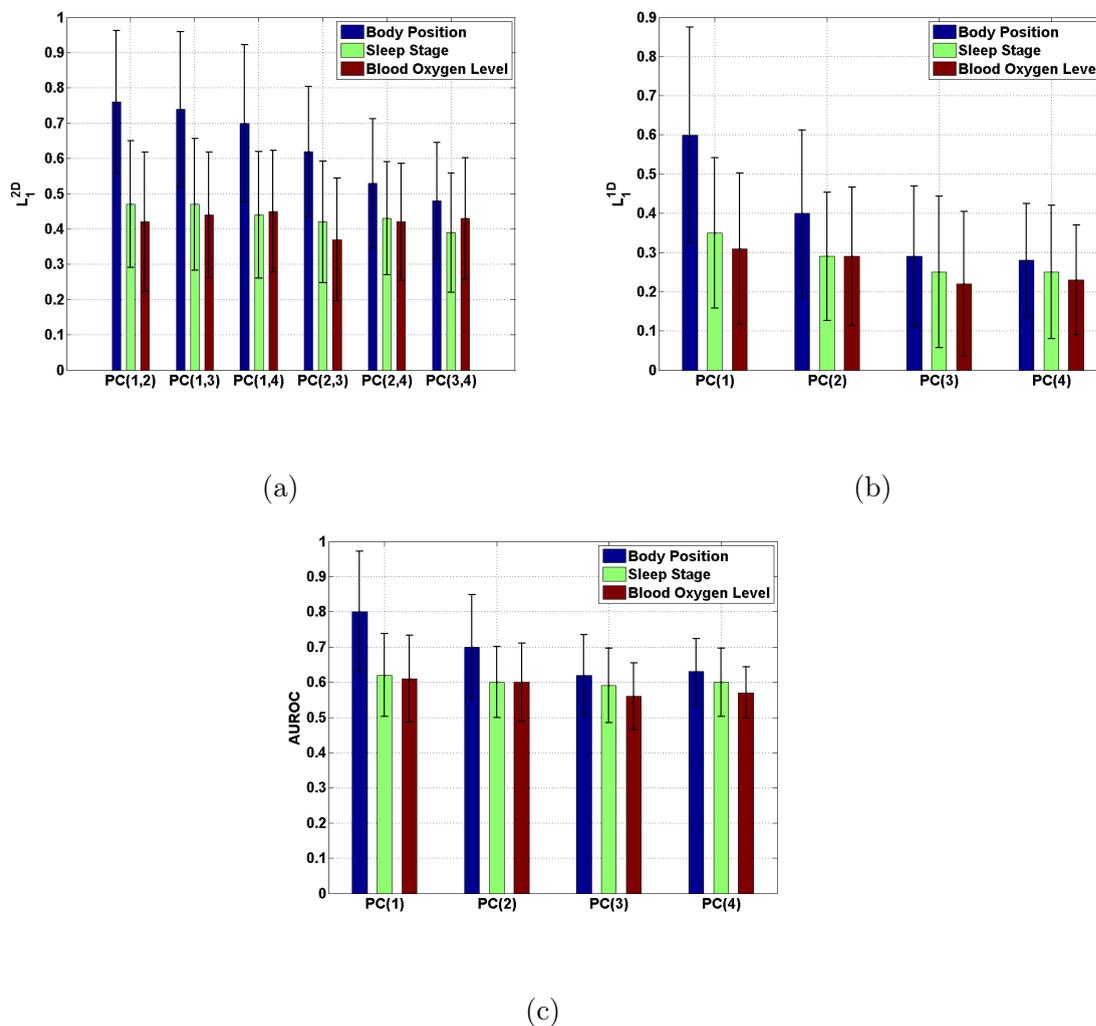


FIGURE B.3: The mean and standard deviation of distance measures between class densities among all snorers grouped based on all categorical variables. (a). The mean and standard deviation of L_1^{2D} , the highest separation is for body position regardless of which bivariate has been used to estimate the 2-D pdf, then, for sleep stage, and finally, for blood oxygen level. (b) The mean and standard deviation of L_1^{1D} . (c) The mean and standard deviation of $AUROC$.