Imperial College London
Department of Electrical and Electronic Engineering

Towards A Truly Wearable, Non-Invasive Respiration Monitoring System

By

Guangwei Chen

2016

A Thesis Submitted for the Degree of
Doctor of Philosophy in
Electrical and Electronic Engineering of
Imperial College London

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

I hereby confirm that all materials in this thesis which is not my own work has been properly acknowledged.

Guangwei Chen

February 2016

Abstract

Non-intrusive medical-grade accuracy respiratory monitoring is one of the two unresolved challenges in design of wearable vital signs monitoring system, besides blood pressure monitoring. The challenge of respiratory monitoring includes automatic breath detection, respiratory phase classification, and extraction of other respiratory related parameters. The main problem is that the sensed respiratory physiological signals could be heavily corrupted by artefacts and interference from non-relevant sources such as movement and other unwanted physiological signals collected at the same time. This thesis proposes a novel system which can detect respiratory activity from tracheal sounds recorded by a miniature wearable sensor with limited power budget, and can be tailored for variety of clinical scenarios by combining novel state-of-the-art techniques at both hardware and software level.

The respiratory monitoring system presented in this thesis consists of two parts: one is a hardware platform for sensing breath sounds and the other is an algorithm for respiratory activity analysis. The low power sensor platform has been optimised for breath sound data acquisition by a combination of: a custom engineered acoustic chamber and a MEMS microphone; and optional power efficient data compression techniques which include the use of novel fixed length adaptive sample size (FLASS) frames optimised for Bluetooth® Smart wireless transmission protocol. The respiratory activity algorithm is based on 3-stage backward adaptive approach, using previous detected respiration's parameters to fine tune the system and then balancing the system's sensitivity and specificity in various ways to suit the need of different applications. These include sleep apnoea diagnosis, sudden unexpected death prevention in both epilepsy and babies, early warning scoring in hospitals and management of chronic conditions such as chronic obstructive pulmonary disease and asthma.

A clinical trial focused on sleep apnoea detection was conducted in the National Hospital for Neurology and Neurosurgery, London, UK, in order to verify the respiratory monitoring system. Questionnaire results show that all participants scored the novel wearable respiratory sensor in the range 4 to 5, where 5 represents the best possible scoring in a number of attributes related to wearability. Experimental results show the proposed automatic apnoea and hypopnoea

detection algorithm has 88.6% sensitivity and 99.6% specificity when comparing its performance with a specialist clinician considered as a gold standard. In comparison a state-of-the-art automatic ambulatory sleep diagnosis system has 14.3% sensitivity and 99.3% specificity.

List of Contents

Abstract **2**

List of Contents **4**

List of Figures **11**

List of Tables **16**

List of Publications **17**

List of Awards **19**

Acknowledgements **20**

Introduction **21**

Chapter 1 Non-Invasive Respiratory Monitoring: A Review **25**

1.1 Motivation 25

1.1.1 Early Warning Scoring 25

1.1.2 Sudden Death Prevention 26

1.1.3 Respiratory Conditions Management 27

1.1.3.1 Sleep Apnoea Diagnosis 27

1.1.3.2 Management of Other Respiratory Diseases 28

1.1.5 Human Behaviour Monitoring 28

1.2 Review of State-of-the-art 29

1.2.1 Airflow Sensing 29

1.2.1.1 Bioacoustic Measurement 29

1.2.1.2 Respiratory Gas Sensing *30*
1.2.2 Respiratory Effort Sensing 32
1.2.2.1 Respiratory Transthoracic Plethysmography *32*
1.2.2.2 Electrical Impedance Pneumography (EIP) *34*
1.2.2.3 Mattress Based Systems *34*
1.2.3 Hybrid Parameters Sensing 36
1.2.3.1 Non-contacting Vision/Imaging *36*
1.2.3.2 Electrocardiogram Derived Respiration (EDR) *37*
1.2.3.3 Photoplethysmogram Derived Respiration (Blood Gas Sensing) *39*
1.3 Design Considerations for a Wearable Respiratory Monitoring System 42
1.3.1 Respiratory Features 42
1.3.1.1 Respiratory Rate *42*
1.3.1.2 Lung Volumes *43*
1.3.1.3 Apnoea and Hypopnoea *45*
1.3.2 Performance and Usability 48
1.3.2.1 Performance *48*
1.3.2.2 Usability *51*
1.3.2.3 Further Discussion *52*
1.4 Conclusions 52

Chapter 2 Sensing Breath Sounds 54

2.1 Origin of Breath Sounds 54
2.1.1 Normal Breath Sound 55
2.1.2 Adventitious Breath Sounds 56
2.1.2.1 Snore and Stertor *56*
2.1.2.2 Wheeze and Stridor *56*
2.1.2.3 Crackles *56*
2.1.3 Sensing Location 57
2.2 Characteristics of Tracheal Breath Sound 58
2.2.1 Signal Bandwidth 58

2.2.2 Breath Spectral Chart 60
2.2.3 Temporal Characteristics 63
2.2.3.1 Attack 64
2.2.3.2 Sustain 64
2.2.3.3 Release 66
2.2.3.4 Breath Length 68
2.2.3.5 Breath Rhythm 68
2.3 Acquisition of Breath Sounds 70
2.3.1 Review of Acoustic Sensors 70
2.3.1.1 Piezoelectric Microphone 70
2.3.1.2 Electret Condenser Microphone 71
2.3.1.3 Microelectromechanical System (MEMS) Microphone 74
2.3.1.4 Summary 76
2.3.2 Air-coupling Acoustic Chamber 76
2.3.3 Design of Analogue Frontend 77

Chapter 3 Truly Wearable Wireless Biomedical Monitoring System Design 80

3.1 Introduction 80
3.2 Power Supply 81
3.2.1 Battery Selection 81
3.2.2 Supply Voltage 82
3.2.3 Maximum Continuous Current 83
3.2.4 Maximum Pulse Current Capability 83
3.2.5 Effective capacity and lifetime 83
3.3 Hardware Platform 85
3.3.1 Data Acquisition 85
3.3.2 Wireless Transmission Controller 85
3.4 Wireless Transmitter 86
3.4.1 Quality of the Packet Transmission 86

3.4.2 Hardware Overhead 87
3.4.3 Air Data Rate vs Effective Data Rate 87
3.5 Design of Wearable Respiratory Sensor 88
3.5.1 Node design 88
3.5.2 Optional data compression 90
3.5.3 Power performance results 90
3.6 Summary 93

Chapter 4 Data Reduction in Respiratory Sounds 94

4.1 Overview of respiratory sound compression – How much can we benefit from that? 95
4.1.1 Lossy Respiratory Sound Compression 95
4.1.2 Lossless Breath Sound Compression 96
4.2 A Lossless Low-Complexity Respiratory Sound Compression Method for Wireless Wearable Sensors 100
4.2.1 Compression Requirements 100
4.2.2 Principle of Breath Sound Prediction 100
4.2.3 Implementation of Breath Sound Synthesis Predictor 101
4.2.4 Fixed Length Adaptive Sample Size (FLASS) Frame 104
4.2.5 Backward Adaptive Residuals Encoding 106
4.2.5.1 Why Backward Adaptive? 106
4.2.5.2 Backward Adaptive Rice Parameter 107
4.2.6 Selective Encoding for Initial Non-Fully Predicted Residuals 108
4.2.7 Results 109

Chapter 5 Automatic Adaptive Respiratory Event Detection and Analysis Algorithms 112

5.1 Background 112
5.1.1 Sound Amplitude Based Methods 112
5.1.2 Machine Learning Based Methods 113
5.1.3 Sound Energy Based Methods 115
5.2 Overview of Proposed Event Detection Framework 118
5.3 Real Time Breath Detection Algorithm 119
5.3.1 Generation of Multi-Channels Signals 119
5.3.2 Real-time Segmentation Algorithm 120
5.3.2.1 Analogue-Digital Conversion 120
5.3.2.2 Basic Routine 120
5.3.2.3 Preliminary Start Point of Breath-like (BL) Segment 123
5.3.2.4 Preliminary End Point of Breath-like (BL) Segment 123
5.3.2.5 Slope Monitoring 124
5.3.2.6 Turning Point Detection 124
5.3.2.7 Verification of Breath-like (BL) Segment Start Point 125
5.3.2.8 Verification of Breath-like (BL) Segment End Point 127
5.3.2.9 Channel Noise Analysis 127
5.3.3 Breath Discriminator – A Scoring System 129
5.3.3.1 Breath Spectral Chart Mapping 129
5.3.3.2 Breath-like (BL) Segment Length 131
5.3.3.3 Channels Trend 131
5.3.3.4 Transient Condition 132
5.3.3.5 Time to Apex 132
5.3.3.6 Release slope 132
5.3.3.7 Signal Onset 132
5.3.3.8 Dynamic Signal Mean Threshold 133
5.3.4 Breath Discriminator – Interpretation 134
5.3.5 Real Time Implementation 136
5.4 Apnoea Detection Algorithm 137
5.4.1 Overview of Apnoea Detection 137

5.4.2 Specificity Improvement based on Respiratory Rhythm Recognition and Prediction 138
5.4.2.1 Respiratory Phase Classification 141
5.4.2.2 Lung Volume Estimation 144
5.4.3 Classification of the Respiration Relative Sleep Disorder Event 145

Chapter 6 Clinical Test on Sleep Apnoea and Hypopnoea Studies 148

6.1 Introduction 148
6.2 Methods 149
6.2.1 Device 149
6.2.2 Participants 149
6.2.3 Procedure 150
6.3 Data Analysis 151
6.4 Results 153
6.4.1 Breathing Exercises Data 153
6.4.2 Sleep Data 156
6.4.2.1 Sleep Apnoea Event Classification 157
6.4.2.2 Respiratory Rate 157
6.4.3 Device Comfort 157
6.5 Discussion 158
6.5.1 Main Findings 158
6.5.2 SOMNO Performance 159
6.5.3 Results Validation 160

Conclusion 161

Contributions 161

Further work 163

Bibliography **165**

Glossary of Terms **185**

List of Figures

1-1: Test subject wearing the Masimo's Cloth rainbow Acoustic Monitoring™ system and close-up photo of its acoustic sensor [90] 31

1-2: Oro-nasal airflow sensing for respiratory monitoring 32

1-3: Respiratory inductance plethysmography (RIP) elastic bands for tracking rib cage and abdomen movement (left) and two different electrical outputs from the elastic bands during contraction and expansion (right). 33

1-4: A test subject wearing the state-of-the-art electrical impedance pneumography based respiratory monitoring system (Sensium Healthcare SensiumVitals®) and close-up photo of their patch sensor [102]. 35

1-5: Clinical infrared thermography setup for infant respiratory monitoring [117] 37

1-6: An illustration of an electrocardiogram (ECG) waveform with PQRST characteristics in one complete heartbeat cycle. 38

1-7: Absorption spectra for three main types of haemoglobin: Methaemoglobin (metHb), oxyhaemoglobin (HbO2), deoxyhaemoglobin (Hb), carbaminohaemoglobin (CO_2Hb) and carboxyhaemoglobin (COHb)[131]. 40

1-8: Typical light absorption waveform (black trace) from finger, where the AC component is generated by pulsatile arterial blood flow; the DC component is generated by constant arterial blood flow, venous blood flow, capillary blood flow and tissue. 40

1-9: A test subject wearing an existing state of the art polysomnography system (SOMNOscreen™ RC, SOMNOmedics GmbH) and an additional pulse oximeter (Pulsox-300i, Konica Minolta) [155], comprising two finger oximeters; oro-nasal flow sensors; thoracic and abdominal expansion bands; and ECG. 47

1-10: Example of polysomnography data indicating obstructive sleep apnoea. 47

2-1: Overview of Respiratory Mechanical System 55

2-2: Men and Women's Chokers (Left is men's necklace from Earkandie's Celebrity Couture Collection[169] and right is women's choker from Sonoor Jewelry Concepts[170]) 58
2-3: Typical tracheal breath sound and airflow signal recording (Source: [175]) 59
2-4: Envelope signal generator for breath sound root mean squared (RMS) power estimation 61
2-5: Breath frequency chart from previous work [182] obtained from 5 volunteers, where CH1 = 500 Hz, 600 Hz;CH2 = 650 Hz, 750 Hz; CH3 = 800 Hz, 900 Hz. 62
2-6: Redefined breath frequency chart obtained from 18 volunteers, including the ones in Figure 2-5, where CH1 = 300 Hz, 400 Hz ; CH2 = 450 Hz, 550 Hz ; CH3 = 600 Hz, 700 Hz. 62
2-7: Extended breath frequency chart obtained from 18 volunteers, including the ones in Figure 2-5, where CH1 = 300 Hz, 375 Hz ; CH2 = 425 Hz, 500 Hz ; CH3 = 550 Hz, 625 Hz; CH4 = 675 Hz, 750 Hz. 63
2-8: Temporal characteristics of a breath sound 64
2-9: Histogram of breath attack duration 65
2-10: Histogram of breath attack slope 65
2-11: Histogram of breath shivering period during the sustain region 66
2-12: Histogram of breath release duration 67
2-13: Histogram of breath release slope 67
2-14: Histogram of breath length 68
2-15: Histograms of breath pause for two cases: solid line for the breath pause within one respiratory cycle; dashed line for the breath pause between two adjacent respiratory cycles. 69
2-16: The relationship between the adjacent breath pauses 69
2-17: Typical frequency response of a piezoelectric microphone [188] 70
2-18: (a) Photo of electret condenser microphone viewed from two different angles; (b) Its cross-sectional structure 71
2-19: Example of 50 Hz noise received by an electret condenser microphone due to poor shielding [182] 72
2-20: Equivalent circuit of a condenser microphone: a JFET transconductance buffer connected to an electret condenser source 73

2-21: Typical frequency response of an electret condenser microphone [192] 73
2-22: Photo of a MEMS microphone (left) and cross-sectional structure of its condenser component (right) [194] 75
2-23: Typical frequency response of a capacitive MEMS microphone [197] 75
2-24: Cross-sectional view of the conical air-coupling acoustic chamber and its surrounding components placed on top of skin 76
2-25: Circuits of analogue frontend for breath sound recording proposed in [201] 78
2-26: Improved circuits of analogue frontend for breath sound recording with better performance on power supply noise rejection. 79
3-1: Wearable device nodes connecting the biological world to the electronic world [202] 80
3-2: Measured discharge characteristics of the small size batteries from Table 3-1. Zinc-Air batteries are tested in a combination of two in series with no air management. 2.1 V is considered as the discharge voltage. Left: Battery maximum supported pulse current vs pulse hold time. Right: Battery lifetime vs average current consumption. 84
3-3: MSP430 based wireless acoustic respiratory sensor node presented here. The platform can be contrasted with the general node architecture presented in Figure 3-1 89
3-4: Measured current consumption profiles of the two transceivers with Vdd = 2.4 V. Left: CC2500, with no FEC, packet size with overhead is 568 bits, 34 samples. Right: nRF24L01+, with Auto-Ack, packet size with overhead is 321 bits, 18 samples. Section **A** is the overall transmission overhead, **B** is the actual wireless transmission time and **C** the optional Auto-Ack reception 91
4-1: Simplified overview of ADPCM with multiple predictions. 97
4-2: An example of how simple polynomial prediction affects breath sound signals in both time (a) and frequency (b) domain. Grey plots are the original continuous-time breath sound signals; circular markers represent the sampled signal with the last two samples replaced by their first order prediction; and the black plots correspond to the interpolated signal. 101
4-3: Probability distribution of the residuals resulting from breath only sound signals obtained from the trachea, against their first three orders polynomial predictions (a) (b) (c) and the proposed prediction (d). 103

4-4: FLASS Frame for verbatim and compressed data fitting in a Bluetooth® Smart (BLE) notification packet. The efficiency of the protocol for actual data transmission is 48.78%.....104
4-5: Compression encoder competition scheme.....105
4-6: Probability distribution of the residuals producing minimum Rice code length with respect to the chosen Rice parameters. The data is randomly selected (from subject 1 in Figure 4-7) 2 hours continuous tracheal sound recording, including breaths, breath pauses and ambient sound noise.....108
4-7: Compression performance for 30 different subjects. All compression methods are implemented with backward adaptive Rice code parameter, except the MPEG-4 ALS which uses its own proprietary encoding approach.....110
5-1: Overview of the respiratory event detection framework 118
5-2: Decision tree of breath detection system. The label of each decision making flow refers to its corresponding block shown in Figure 5-1.....119
5-3: Basic routine for a segmentation algorithm. Note that the actual duration of each step is not to be scaled from the length in the diagram. The difference on the length indicates the order of steps finishing.....121
5-4: An example of a new breath-like (BL) segment being identified. This example assumes the selected main channel has the highest signal-to-noise ratio (SNR) within the shown time frame. In reality, the main channel may be switched at the transit from peak seeking state to valley seeking state.....121
5-5: Flowchart of turning points detection.....124
5-6: An example of breath-like segment with ambiguous starting. When ①, ② and ③ conditions are met, point 'C' replaces the point 'A' as the verified start point of the breath-like (BL) segment. 127
5-7: Noise floor estimation from envelope signals of an arbitrary channel over two respiratory cycles. Note that the marked segments are separated by preliminary start and end point instead of the verified ones.....128
5-8: Limiting area of spectral energy distribution for a three-channel breath detection system.....131
5-9: Graphic user interface of the real time breath and apnoea detection algorithm.....137
5-10: Overview of apnoea detection mechanism.....137

5-11: Example of how to upgrade a pair of lower level confidence class breath candidate (BC) segments by recognising respiratory rhythm. Each box represents detection of a BC segment along the time axis. 139
5-12: Recursive routine for respiratory phase classification. 142
5-13: Examples of output of BC segments obtained from 4 different kinds of sleep apnoea/disorder events: hypopnoea, central apnoea, obstructive/mixed apnoea and hyperventilation. 147
6-1: Wearable respiratory activity monitor worn by myself [155] and a volunteer. 149
6-2: Illustration of the SOMNO and Wearable Apnoea Detection Device (WADD) output signals showing an apnoea event: (top) raw signals from the different SOMNO sensors, (middle) processed envelope signals from channel 1(350Hz – 450 Hz), (bottom) WADD raw signals. 154
6-3: Difference between visually determined respiratory rate and the corresponding automatic value calculated by WADD against the mean of two methods. Total number of samples is 38071. 158

List of Tables

1-1: Comparison of various respiratory monitoring methods 49

2-1: Comparison of initial works on tracheal breath sound measurements and bandwidth results. 59

3-1: Specifications of four non-rechargeable disposable batteries potentially suitable for wearable sensor nodes with varying physical sizes and battery technologies [202] 82

3-2: Measured current consumption of the hardware platform and its constituent parts, excluding the wireless transmitter. Note that the CC2500 system has a 10 MHz clock while the nRF24L01+ system uses an 8 MHz clock. 91

3-3: Current consumption of the sensor nodes with compression enabled and disabled. Only some of the miniature batteries listed in Table 3-1 are capable of powering the system for 24 hours. 92

4-1: Rice code format and the length of its each field. 107

4-2: Average compression performance 110

5-1: List of machine learning algorithms tested on breath detection 114

5-2: Description of the confidence levels given by the breath discriminator. All the segments with non-zero confidence levels are considered breath candidates (BC). 135

5-3: Sensitivity and specificity of single respiratory phase detection with confidence level 5 with respect to various ranges of tidal volumes [201] 136

6-1: Summary of performance for the WADD and SOMNO across the seven breath exercises (as detailed in procedure) 155

6-2: Summary of performance for the WADD, SOMNO and clinician scorer systems for detection of apnoea and hypopnoea in 10 s epochs of overnight recordings 156

List of Publications

Patent Application

- E. Rodriguez-Villegas, G. Chen, E. Aguilar-Pelaez, "Feature Characterization for Breathing Monitor," WO/2012/007719, 19 Jan, 2012.

Journals

- G. Chen, S. A. Imtiaz, E. Aguilar–Pelaez, and E. Rodriguez–Villegas, "Algorithm for heart rate extraction in a novel wearable acoustic sensor," *Healthcare Technology Letters,* vol. 2, pp. 28-33, 2015.
- E. Rodriguez-Villegas, G. Chen, J. Radcliffe, and J. Duncan, "A pilot study of a wearable apnoea detection device," *BMJ open*, vol. 4, p. e005299, 2014.
- E. Aguilar-Pelaez, G. Chen, and E. Rodriguez-Villegas, "Technique for interference reduction in battery powered physiological monitoring devices," *Biomedical Engineering, IEEE Transactions on*, pp. 3197-3203, 2012.

Peer Reviewed Conference Papers

- G. Chen, I. de la Cruz, and E. Rodriguez-Villegas, "Automatic lung tidal volumes estimation from tracheal sounds," in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, 2014, pp. 1497-1500.
- G. Chen and E. Rodriguez-Villegas, "System-level design trade-offs for truly wearable wireless medical devices," in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, 2010, pp. 1441-1444.

Book Chapters

- G. Chen, E. Rodriguez-Villegas, and A. J. Casson, "Wearable Algorithms: An Overview of a Truly Multi-Disciplinary Problem," in *Wearable Sensors*, E. Sazonov and M. R. Neuman, Eds., 1st ed Oxford: Academic Press, 2014, pp. 353-382.

Medical Conference Communications

- E. Rodriguez-Villegas, E. Aguilar-Pelaez, G. Chen, and J. S. Duncan, "Apnea Detector to Prevent SUDEP," in *28th International Epilepsy Congress*, Budapest, 2009, pp. 37-38.
- J. S. Duncan, E. Rodriguez-Villegas, E. Aguilar-Pelaez, and G. Chen, "Wearable Apnea Detection Device to Prevent Sudden Death," in *62nd Annual Meeting of the American-Epilepsy-Society*, Seattle, 2008, p. 77.

List of Awards

- XPRIZE Nokia Sensing XCHALLENGE 2014, Distinguished Award. Team Members: Guangwei Chen, Syed Anas Imtiaz, Esther Rodriguez-Villegas
- Institution of Engineering and Technology (IET) Innovation Awards 2009, winners of Information Technology category under the entry of: "A truly wearable device for real time monitoring of breathing and heart rate". Also selected finalist in other three categories: Emerging Technologies, Measurement in Action and Product Design. Team Members: Guangwei Chen, Eduardo Aguilar-Pelaez, Esther Rodriguez-Villegas, John Duncan.

Acknowledgements

Firstly, and most importantly, I would like to thank my parents for all the support and motivation throughout the years, your patience and trust have been invaluable.

On a more professional note I would like to thank my supervisor Esther Rodriguez-Villegas for all her assistance and guidance. I would also like to thank the other people who have contributed to this wearable respiratory monitor project: Edurado, John, Jeremy, Helen, Brett and Diane as well as all the volunteers and patients that took part in the clinical trials. Special thanks to Anas, Stuart and Romanos. Without their help, Esther and I would not see this technology won the XPRIZE competition.

I would like to acknowledge and pass my appreciation to all those in the labs at Imperial: Alex, Richard, Lojini, Amir, Yan, Hugh, Elena, Justin, Saam, George, Siavash, Mohan, Zhou, Majd, Romanos, Zhenyu, Kan, Liucheng and of course Wiesia. Rest who I've forgotten but you are appreciated nonetheless.

Finally, all my thanks to my girlfriend Bidan Huang.

Introduction

Respiration is a vital physiological process which allows humans to obtain oxygen during the inhalation phase and discard carbon dioxide through exhalation. Cessation of respiration, also known as apnoea, is harmful to human health in general and can be fatal in prolonged[1] extreme cases as it may cause brain damage due to both low level of blood oxygen saturation (hypoxemia) and high level of blood carbon dioxide concentration (hypercapnia) [1, 2]. There are two main reasons why apnoea might occur: medication side effects [3-5]; disease caused by infections, respiratory, cardiac and neurological disorders [6-10]. For example, apnoeic events are very common on patients after receiving anaesthesia in recovery wards. This is the reason why they have to be continuously monitored by nurses in order to prevent respiratory depression which could lead to death [11-13]. Apnoea can also be caused by upper airway obstruction. This results on a common sleep disorder which increases the risk of hypertension, cardiovascular disease, stroke, daytime sleepiness, motor vehicle accidents [14, 15] and diminished quality of life[16]. In the context of neurological disorders, apnoea is generally not obstructive but central in nature and is potentially life threatening, being known to be a cause of Sudden Unexpected Death in Epilepsy (SUDEP) and Sudden Infant Death Syndrome (SIDS). These account for over 1000 deaths per year in the UK [17-20].

Unfortunately, breathing is arduous to observe continuously since there is no reliable non-invasive portable device available to detect it without clinical settings. The current gold standard of unattended respiratory monitoring generally invokes nasal airflow sensing to detect changes on airflow in front of the nostrils and mouth; and a pulse oximeter to estimate the arterial blood oxygen saturation. However, it is hard to guaranty their signal quality which can be affected by many external factors, including subject's movement amongst others. In the worst but very common cases, such as when the subject is sleeping or unattended, the sensors would not get back to the right position if their correct placements change due to the movement.

[1] Generally over 5 minutes. The exact time depends on the subject's respiratory vital capacity and pre-existing medical conditions.

Because of this standard clinical guidelines require more physiological channels to be recorded in clinical settings for sleep disorder studies[21]. The combination of all the recommended physiological recording channels is called polysomnography, and is the ultimate gold standard for sleep studies. Besides nasal airflow and blood oxygen saturation, the polysomnography system normally involves monitoring of brainwaves (EEG), heart waves (ECG), and respiratory effort (chest/abdomen movement). Breathing activity is evaluated based on the combination of signals obtained from at least three different sensors, where pulse oximeter, nasal/oral thermistors and chest/abdomen belts are chosen in most of the cases. However, even when comprehensive physiological signals are recorded to improve the diagnostic yield, the polysomnography system is still highly vulnerable to body movements, because these generate interference and in many cases signal loss from the wired sensors. Hence, there are situations in which doctors are not able to even assess whether something is clinically significant- for example, an apnoeic episode is happening, or the patient is moving, or both. Furthermore, from the point of view of the patients, in the best case scenario, they have to wear a palm size 'wearable' box wired up to different sensors which restrict them from free movement and significantly compromises their sleep quality. Added to this there is no portable apnoea monitoring system that can detect apnoea automatically with high enough sensitivity and specificity. Hence, current polysomnography systems are only suitable for off-line, time consuming, clinical studies.

Therefore, there is a real need for a wearable device that can be used to real time monitor and process breathing patterns automatically for extended periods of time, while the user is still being able to carry out a normal life. The design targets for the respiratory monitoring system are improved reliability when compared to the existing polysomnography systems in terms of sensitivity and specificity; and suitability for long term use, not only for low-intensity clinical, but also non-clinical day to day situations. The device attached to human subject has to be a lightweight, miniature and non-invasive wireless solution which not only can measure the correct respiratory rate, but also detect apnoea events that last longer than a pre-established time threshold, reliably and subsequently perform task in response, such as raising an alarm, even in the case of epilepsy patients who may be simultaneously having a seizure. For people with risk of sudden death like epileptic or intensive care unit patients, the apnoea alarm may be sent to their nearest nurse or family member so that somebody can assist the person if necessary. The apnoea sensitivity of the device is particularly crucial for sudden death prevention, since

one missed apnoea detection could be fatal. However, false alarms should also be avoided since, although not leading to fatal situations, they would make the device unattractive to end users who want to be carrying out a relatively normal life. For sleep disorders monitoring, real time signal processing or raising an alarm would not be necessary. Instead, the device should incorporate functionality to record the specific time of apnoeic events and the breathing signal around that time. This information would be of more value for the doctor since it would enable them to analyse the event afterwards.

The aim of this PhD was to create a novel wearable respiratory monitoring system that can potentially replace or even outperform the current gold standard, considering all the challenges described above. The thesis is structured as follows:

Chapter 1 covers the background on respiratory monitoring, presenting the current state of the art and a comprehensive discussion on the challenges behind existing methods. The chapter also explains why bioacoustics breath detection is chosen as the sensing modality to create the novel wearable breathing monitoring system.

Chapter 2 focuses on the theory behind breath sounds; review the different systems used by researchers to collect these sounds together with the different approaches devised for breath/apnoea detection based on bioacoustic sensing; and concludes presenting a new set of acoustic breathing signal characteristics.

Chapter 3 analyses the different key factors that come into play when aiming to create a truly wearable biomedical monitoring system. A generic non-invasive wireless solution by using off-the-shelf components is presented, which provides the data acquisition platform to develop and validate the wearable respiratory monitoring technology presented in this thesis.

Chapter 4 presents a novel breath sound compression algorithm which is easy to implement on an ultra-low power Bluetooth® Smart platform and allows a considerable amount of power saving in the overall breath sound acquisition system. The data compression algorithm comprises new ways of predicting breath sound data, dividing data into frames as well as encoding them.

Chapter 5 reviews the state-of-the-art on automatic breath detection and introduces the main breath detection algorithm created during the course of this PhD State machines and flowcharts

have been employed to illustrate the operation of the algorithm. An early pilot study was conducted to assess the performance of the core (first two) stages of the breath detection algorithm. These initial results, presented in this chapter, show how the core stage of algorithm is able to detect potentially critical apnoea events reliably, but at the expense of a still unacceptable rate of false alarms. The second part the chapter focuses on the proposed solution for reducing false apnoea alarms in the last stage of the detection algorithm.

Chapter 6 analyses the results obtained from a clinical trial carried out to evaluate the performance of the new system. The trial took place in the sleep study room of the Harris Neuromedical Intensive Care Unit at the National Hospital for Neurology and Neurosurgery, Queen Square, London. The results from the trial were evaluated using different metrics and the performance was compared against the gold standard for automatic diagnosis of sleep apnoea: visual marking of events by the human expert. Additionally, the trial was also used to evaluate the performance for automatic evaluation of an FDA approved ambulatory polysomnography system. Overall, the results show how the novel technology was significantly superior to existing methods having around 90% agreement with the clinical expert.

Chapter 1

Non-Invasive Respiratory Monitoring: A Review

1.1 Motivation

The human respiratory system anatomically consists of the nasal/oral cavity, trachea, bronchus and lungs. Breathing is caused by effort of the diaphragm and intercostal muscles drawing air through the nasal/oral cavity down the trachea, bronchus and into the two lungs where gas is exchanged with the circulatory system. Although the breathing rhythm can be altered or even stopped by controlling the respiratory muscles voluntarily, the function of the respiratory system is maintained and regulated by the autonomic nervous system which operates without the need of consciousness and is powerful enough to override any harmful breath holding action so as to prevent untoward hypoxic effect. Any failure in either the physical or neurological parts of the respiratory system degrades the overall performance of the human body and can lead to unexpected consequences, including hypertension, obesity, coronary artery disease, type-2 diabetes [22], and in some cases, even death [23-25]. In the following section, four different applications are discussed in order to support the case for novel wearable respiratory monitoring techniques.

1.1.1 Early Warning Scoring

The Early Warning Score (EWS) System is commonly used as a benchmark for patients in intensive care conditions [26]. The EWS system performs measurements on temperature, blood oxygen levels, consciousness, heart rate and the respiratory rate at least once per hour. An index derived from the above parameters by an EWS algorithm can be used to detect the patient deterioration. Previous studies [27-30] have shown that the conventional automated EWS systems have a false alarm rate of 90%, due to incorrect measurements. Recently, the

performance of EWS systems has been improved by using experimental machine learning techniques, however, the best achievable combination of sensitivity and specificity is still below 75% [31, 32]. One of the main issues in current EWS systems is the inaccuracy of the respiratory rate measurements in such systems. This explains why the respiratory rate is still quantified directly by nurses in most clinical ward setups rather than depending on automatic measurement. Normally, the measurements are performed by a nurse standing next to a patient with the bed sheet lifted and by counting the number of movements of the abdomen/chest for 30 seconds [33]. The respiratory rate is then extrapolated to one minute by doubling the measurement result. This method is inevitably prone to human errors and distracts nurses from their normal routines, but is still the only trusted procedure of such nature in most hospitals. It has thus been demonstrated that a reliable and easy-to-use automatic respiratory monitoring system can be used to aid nurses in performing respiratory rate measurements on patients.

1.1.2 Sudden Death Prevention

Sudden death resulting from apnoea (i.e. cessation of breathing) can be caused by a number of physiological conditions including epilepsy where the condition is known as SUDEP [25], for several unexplained reasons including sudden infant death syndrome (SIDS, as known as cot death) [34], respiratory diseases especially when there is retention of carbon dioxide (CO_2) [35], and neuromuscular diseases causing respiratory failure [36].

Focusing for example on epilepsy, apnoea is common in epileptic seizures mostly in complex partial and tonic-clonic seizures. SUDEP kills at least 500 individuals per year in the UK and over 5,000 in the USA [37, 38]. Apnoea is the leading contender as the cause of SUDEP, the other being cardiac asystole, which results in secondary apnoea [9, 39, 40]. Prompt detection of apnoea is necessary for life-saving intervention to be initiated with cardiopulmonary resuscitation. At present there is no reliable non-invasive method to detect apnoea that is suitable for long term use in low clinical intensity or domestic environments, with the consequence that there are thousands of deaths that could have been prevented.

Epilepsy affects approximately 1% of the population (500,000 people in the UK with 32,000 new cases each year, 50 million people worldwide) [41]. SUDEP occurs most commonly when the person is unsupervised and at night [9, 42]. Alarm systems are considered a potential preventive measure in unsupervised patients [9]. Existing alarm systems are primarily based on detecting

vigorous rhythmic movements that might be caused by a seizure, as opposed to the cessation of external respiration that leads to death. Hence, they do not work with seizures such as complex partial ones that do not cause vigorous movement but can cause apnoea [37]. Additionally they suffer from other limitations: cost, long time lag to the trigger of the alarm, or most importantly very poor performance due to signal artefacts causing false alarms [43]. Breathing monitors used in other clinical contexts such as pulse oximeters are unsuitable because of the large number of false alarms they create and poor "time sensitivity", as in the majority of cases the delay from the start of an apnoea to oxygen desaturation is too long, which results in the loss of valuable time in summoning help.

In the context of sudden death in infants, infants who are born prematurely or who have low birth weight have up to four times the risk of sudden death syndrome than those infants born at term [44]. Moreover, for those with identified problems such as reflux and infection, apnoeas can also be frequent [45, 46]. In the UK alone, the number of infants registered as sudden infant death has kept constant, at an approximate level of 600 cases every year, for last decade [18-20]. Existing baby monitoring systems are again highly sensitive to artefacts, and in most cases, lack adequate sensitivity too. A system with the automatic sudden death prevention alarm functionality would have an important reassuring value for parents, especially for those with premature or low birth weight babies.

1.1.3 Respiratory Conditions Management

1.1.3.1 Sleep Apnoea Diagnosis

Sleep apnoea is the third most common respiratory disorder, after asthma and chronic obstructive pulmonary disease (COPD). It affects between 2% to 10% of the adult population and 1% to 5% of the paediatric population [47, 48]. In patients with chronic heart failure, the prevalence of breathing abnormalities during sleep is at least ten times greater, and they are associated with a 400% increased risk of mortality [49]. The indirect medical costs of under-diagnosed adult sleep apnoea patients, in the years preceding the diagnosis, is estimated to increase by up to twofold, even after correcting for chronic disease status [50]. This, added to the potential social consequences in the form of accidents, increased morbidity and impact on work efficiency, makes the condition a major public health issue [47].

The main problem with diagnosis of sleep apnoea is the scarcity and high cost of diagnostic sleep clinics where patients can be admitted for single night polysomnography studies; in some cases this can lead to waiting times of over a year in fully developed areas [51], because there is still no simple and reliable sleep study system that can be trusted to be used outside the clinics without the aid of fully trained specialists. With the limited diagnosis time window, it is estimated that overall wrong diagnosis and missed diagnosis rate of the sleep apnoea could be between 20% and 90% [52-54]. Moreover, monitoring children for the disease is even more problematic than adults because of the implications it has for children to have to spend the night in a clinic with a set of uncomfortable sensing devices attached to them. Therefore, there is a great need for an easy and very cost effective method to reliably pre-diagnose sleep apnoea within the home environment, and to identify those patients who should attend longer (i.e. more than one night) polysomnography study.

1.1.3.2 Management of Other Respiratory Diseases

Apart from apnoea detection, the ability to determine, with high accuracy, breathing related parameters is also useful for diagnosis and management of other respiratory diseases such as asthma and chronic obstructive pulmonary disease (COPD). According to the World Health Organization (WHO), there are approximately 235 million asthma and 64 million COPD patients in the world [55, 56]. Most of the respiratory disease diagnosis approaches are performed in a one-time measurement way because of their dependency on spirometry and auscultation. Hence, there is a need for a reproducible and accessible method for analysis of breathing related parameters which provides more comprehensive information for longer duration. Such system could be used by doctors to remotely diagnose or provide a first assessment of patients, in order to: prioritise face to face appointments, reduce waiting times, reduce healthcare costs, allow evaluation of patients that cannot easily attend clinics (such as those living in rural or underdeveloped areas) and improve diseases management [51]. On line with the trend towards the decentralisation of healthcare this should be offered in larger numbers to either the patient or a health enthusiast at lower cost.

1.1.4 Human Behaviour Monitoring

Apart from the healthcare applications briefly described above, a wearable respiratory system could be an invaluable tool for human behavioural studies. Monitoring of human behaviour in

free individuals is arguably one of most popular applications of wearable technology. Most wearable behaviour recognition systems are accelerometer based, and hence can only track down various movements of the body [57]. These provide a limited amount of information and consequently have a narrow applicability. With the help of an accurate respiration monitoring system, it might be possible to examine the pattern of breathing and characteristics of individual breaths and use this to detect a range of human activities and behaviours, such as talking [58], singing [59], deglutition [60], smoking [61], drowsing [62], snoring [63], teeth grinding [64], and emotional changing [65]. Knowing these behaviours can be beneficial for some of applications, like mental disorder monitoring, smoking cessation program monitoring and drowsy driving detection. This has been however not possible so far because there was no unobtrusive reliable wearable respiratory monitoring system that the patient could be sent home with, for continuous monitoring for extended periods of time, and was moreover, able to provide continuous access to high quality respiratory signals, as well as very accurate automatic evaluation of breath pattern and individual breath's characteristic; whilst operating continuously for a long time without disturbing the user.

1.2 Review of State-of-the-art

1.2.1 Airflow Sensing

1.2.1.1 Bioacoustic Measurement

The history of acoustic based respiratory monitoring only begins in the 1810s when physicians started using the traditional stethoscope to hear lung sounds caused by air turbulence in the respiratory airway, with the aim of diagnosing lung diseases. However, these early stethoscopes had strong attenuation above 300 Hz [66] which made them suboptimal for this purpose, since, as later on was proven, the acoustic breathing signal extends far beyond this frequency. This is made worse by the fact that the human ear is not very sensitive to the sound with a low pitch frequency [67].

In the middle of the 20th century, when the new era of digital signal recording began, some researchers [68-73] attempted to study the frequency content of breath turbulence sound measured from the respiratory airway. Gavriely et al. [74] were the first ones to comprehensively unscramble the entire power spectral distribution of breath sounds from both the chest wall and

the trachea. Later, further studies [66, 75, 76] suggest that trachea measurement is better for breath sound detection, because it provides more spectral information and lung sounds from the chest share the same bandwidth as heart sounds. Moreover, the specific position of the suprasternal notch, this is, where the clavicles join the sternum, is the best region suggested to measure tracheal breath sounds due to its diminished noise from head movements and carotids. It was not however until 1993, when Nissan and Gavriely [77] used a microcomputer to collect significant amount of breath samples, which they used to develop the earliest automatic breath detection system. This is the first system that is able to automatically detect breath and capture its corresponding audio spectrum in real time for further analysis by examining convergence of averaged spectra of breath sound signals that are collected from different locations such as trachea and chest.

Since then, automatic interpretation of the acoustic respiratory signal gains more and more attention. However, the progress has been slow and it is still a very immature field of research. Two main reasons that have limited its progress are: the size of sensing systems (which in its turn affects other aspects, such as the type of experiments that can be done with them), and the quality of the data that can be extracted from them, as the breathing signal, unless it is under controlled conditions, is heavily corrupted by artefacts. Others in the past had tried to detect breathing by monitoring its associated sound but with cumbersome apparatus, such as wired microphones or accelerometers; often keeping subjects in still positions and/or in a silent environment with generally very limited (or non-existent) success in automatic detection of breathing segments especially in the presence of artefacts [78-88]. Figure 1-1 shows a state-of-the-art commercial system (Masimo's Cloth rainbow Acoustic Monitoring™ system) based on acoustic sensing for breathing rate monitoring in post-operative patients. Their recent clinical reports [85, 89] claims that breath sound contains enough information for evaluation of respiratory rate with the 3% higher accuracy compared to capnography (i.e. nasal carbon dioxide concentration sensing, see next section 1.2.1.2). However, this is in very restricted clinical conditions for monitoring of anaesthetized patients.

1.2.1.2 Respiratory Gas Sensing

The amount of air displaced from the lungs is also of significant diagnostic value. The quantitative estimation of respiratory flow and lung volumes started in the 1850s by using a

Figure 1-1: Test subject wearing the Masimo's Cloth rainbow Acoustic Monitoring™ system and close-up photo of its acoustic sensor [90]

mechanical based method known as spirometry [91]. The function of a spirometer is similar to that of an anemometer: the instrument registers the mechanical movements caused by air pressure. It was not until the 1860s however that Tyndall discovered the infrared absorption property of CO_2 [92]. This was used to quantitatively estimate the continuous expired CO_2 (i.e. end-tidal CO_2), and the technique known as capnography, became available for respiratory monitoring. Capnography is currently widely used in Intensive Care Unit (ICU) and operating theatre settings. It utilises a facemask over the nose and mouth connected to an air tube with a CO_2 sensor attached. This allows recording of exhaled CO_2 concentration with extremely high accuracy [93]. This sophisticated capnography does however receive very low feedback on user experience due to the uncomfortability of the design [94].

A simplified system to quantify gas exchange associated with respiration, which avoids the use of facemasks relies on using oro-nasal airflow sensors in a configuration such as the one in Figure 1-2. The sensors measure changes on the properties of the air surrounding the nose and mouth and from there the airflow can be estimated. Besides CO_2, these properties may include pressure, temperature, and humidity. One of the issues with this kind of configuration is that, as can be seen in Figure 1-2, unlike the face mask, the oro-nasal airflow sensors do not have a sealed air collector. Because of this the gas dilution of the unsealed air collector inevitably leads to signal fading in the case of, for example, subject's movement. Added to this, in the absence of respiratory effort information, in some shallow breathing cases, the airflow sensors may also pick up tiny signals originated by the beating of the heart during a period of respiratory pause, and wrongly detect it as respiration. All of this results on the mask-free approach being less reliable for accurate evaluation of tidal volume [95]. In spite of the disadvantages, this approach is able to detect the presence of breath in most circumstances. Because of this, and since it is also

much more comfortable to wear than anything involving a facemask, it has become an essential component of polysomnography systems for sleep studies. Still, due to its shape and the location of the sensors it is not used for in-domestic monitoring.

1.2.2 Respiratory Effort Sensing

Along with airflow sensing techniques, a different approach that is being used for respiratory monitoring is based on sensing the expansion and contraction of the respiratory cavity. This approach however suffers from similar bottlenecks in the presence of movement artefacts. Additionally, in the absence of breathing any automatic software relying on it is prone to falsely detecting cardiac activity as respiration in the presence of apnoea. This can somehow be compensated for by adjusting the tightness of the bands. However, this can only be done by a clinical specialist because if a band is too loose to avoid picking up cardiac activity the system loses sensitivity to shallow breath cycles.

1.2.2.1 Respiratory Transthoracic Plethysmography

Amongst all technologies for monitoring of respiration, respiratory transthoracic plethysmography (RTP) is possibly the best developed system, being widely used in a clinical settings. RTP utilises either or both chest and abdomen elastic bands to track the change of lungs volume during the process of respiration. By converting the rib cage and abdominal movements into electrical signals, respiratory plethysmography records the activity of intercostal muscles and diaphragm. From there it is possible to determine whether there is any respiratory effort.

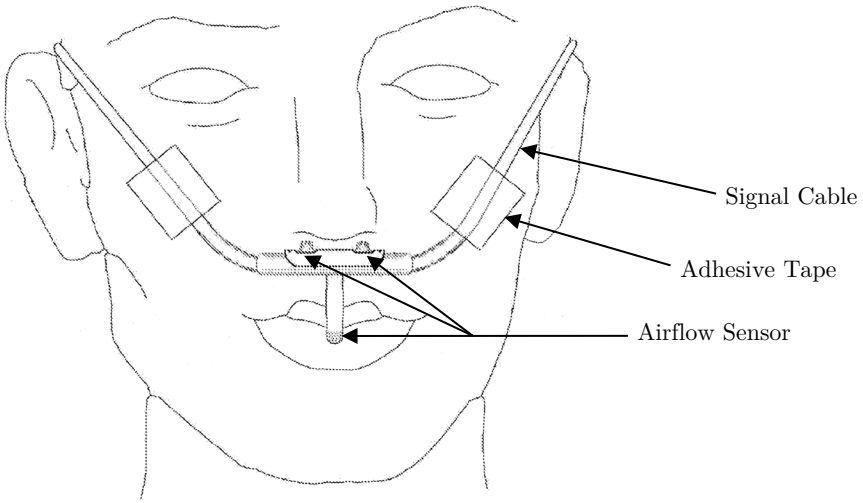

Figure 1-2: Oro-nasal airflow sensing for respiratory monitoring

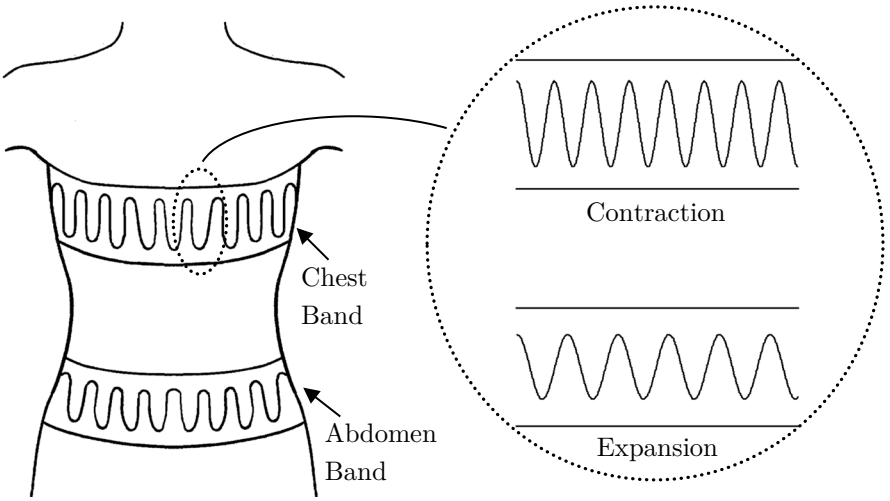

Figure 1-3: Respiratory inductance plethysmography (RIP) elastic bands for tracking rib cage and abdomen movement (left) and two different electrical outputs from the elastic bands during contraction and expansion (right).

There are three types of RTP. Each of them bases on a different sensing principle: strain gauge based, magnetometry based and inductance based. The earliest version of it was resistance strain gauge plethysmography. In resistive strain gauge plethysmography a durable conductive material is attached to an elastic band, which is able to convert expansion force into changes of the material's resistance as this gets longer and thinner. This was relative easy to make and the corresponding sensing circuit was simple. Later, Mead et al [96] used magnetometry to track the change of thoracic diameter by measuring the mutual induction between two wire coils secured to the front and back of the thorax: one to generate a magnetic field and the other one to sense the field oscillation related to their separation. In 1978, Cohn et al. [97] converted the two-coil magnetometry approach into just one-coil, introducing respiratory inductance plethysmography (RIP). The implementation is shown in Figure 1-3. A zigzag shaped inductive wire coil is attached to an elastic band. In the process of breathing, the flexible coil expands during inhalation and goes back to its shape during exhalation. The distance between each turn of the wire coil varies and consequently the inductance. By integrating the coil into a power-efficient self-oscillating circuit, the thoracic volume change can be converted to change in its resonant frequency. In general, inductance plethysmography is more sensitive to the respiratory efforts than the other two approaches, but the difference is unlikely to be clinically important [98].

Although RTP is useful in some contexts, it does have lots of limitations. Systems are cumbersome, uncomfortable, unreliable in the presence of artefacts and prone to misplacements.

Furthermore, reliability in the presence of movement is inversely proportional to the system comfortability because it is related to the tightness of the bands. Because of this, RTP cannot be used for automatic identification of apnoeas when accuracy is a concern [99].

1.2.2.2 Electrical Impedance Pneumography (EIP)

Another approach used to measure respiratory effort is based on active impedance measurements. This is generally referred to as electrical impedance tomography or electrical impedance pneumography (EIP). During inspiration, the thoracic impedance increases due to the increase in the gas volume of the chest and the length of the conductance paths. The opposite happens during expiration. An EIP system injects a low amplitude and high frequency current (maximum 100 µA of peak current at 10 kHz to meet the safety limits imposed by standards [100]) which passes through the subject's body via an electrode array (in most cases, at least 2 and up to 17 standard electrocardiogram electrodes). The variation on electrical impedance due to the changes of the respiratory cavity can hence be measured. Typical variations of thoracic impedance caused by changes in lungs tidal volume are 1 ohm/litre in average[101]. This results in approximately 100 µV/litre resolution for a maximum 100 µA injected current.

EIP requires the subject to be at rest since any muscle movement generates unwanted impedance changes that can be large enough to corrupt the respiratory signal. The most mature commercial system is shown in Figure 1-4. It has the shape of a plaster with a disposable electrode patch and a reusable sensor module. The artefact rejection strategy that they use is very simplistic and widely known to be of very limited performance, which is likely to be the reason why the validation is being done not on breathing rate comparison, but on ranges instead (i.e. the two systems agree when the parameters are above a certain range). Generally, EIP is less reliable compared to respiratory plethysmography because of its lower level of accuracy [103, 104].

1.2.2.3 Mattress Based Systems

Respiratory movements can be picked up by the ripple from a mattress top or bottom surface due to the body extrusion. Although often people refer to it as a non-contact approach, this actually requires subjects to lie on a mattress with a network of pressure sensors on its surface, either on the top or the bottom. The earliest implementation [105, 106] of this approach was based on using an air or fluid filled mattress to detect the respiratory effort via pressure changes.

Figure 1-4: A test subject wearing the state-of-the-art electrical impedance pneumography based respiratory monitoring system (Sensium Healthcare SensiumVitals®) and close-up photo of their patch sensor [102].

Later, Chow et al. [107] improved the resolution of movement detection by introducing more pneumatic segments and pressure transducers to the mattress. This requires customized mattress, which most people would not be willing to acquire for reasons of cost, which makes the whole idea less user friendly. To overcome this, another popular method to measure respiratory effort consists on placing a less sensitive thin membrane-type pressure sensor array mat on top of the mattress [108] or a highly sensitive pressure sensor array underneath [109]. Placing the sensors underneath is more unobtrusive since membrane-type pressure sensors still have a certain thickness and hardness which affects comfortability. However, the characteristics of the mattress, in terms of mattress material and thickness, have a huge impact on the performance of the under-mattress approach [109]. Although this kind of monitoring technique seems to outperform any wearable solutions as misplacement and wearability are not issues any more, it loses in flexibility since it cannot be used without the monitored subject lying on the bed and in practical terms only allows one person to use the mattress at a time. In terms of accuracy of breath detection, mattress based systems are comparable to elastic bands based methods. In both cases performance drops badly if there is any non-respiration related movement artefact.

As reported in [107, 110], compared to elastic band based systems, any movement of the body other than thorax and abdomen would be more likely to affect this type of systems if there is no sophisticated movement artefact removal strategy in place. Although this approach is commercially available as part of baby monitoring systems since 1992 [111], none of these mattress based systems has good enough accuracy to detect apnoea reliably, hence their commercial and regulatory claim is that they can only raise alarms for no movement, but not apnoea.

1.2.3 Hybrid Parameters Sensing

The approaches described above can only sense one respiratory parameter. The following section introduces standalone approaches that are able to measure more than one parameter.

1.2.3.1 Non-contacting Vision/Imaging

Research on non-contacting type of respiratory measurement goes back to the 1970s. It started by studying the interaction between different kinds of signals and the human body in order to extract respiratory parameters. Those signals included radio frequency signals [112], ultrasonic signals [113, 114] and visible/non-visible light [115, 116]. Most of the approaches emit a certain kind of field directed towards the observed body. By detecting variations in phase shifts and round trip time of the reflected signal, the distance change between the source and the test subject's chest wall can be measured, hence the movement. A few like [116] try to passively measure the infrared wave released from the human body and detect the movement of the heat source.

Non-contacting approaches are not limited to respiratory effort sensing. Abbas et al. [117] recently developed a different method to monitor respiration based on real-time infrared thermography, as shown in Figure 1-5. Like the thermistor based nasal oral airflow sensing technique, this technique uses a high resolution infrared camera to remotely detect the thermal signature of respiration, which is a tiny temperature variation from 0.3 ℃ to 0.5 ℃ around the nostrils opening area. Provided there is no other heat source or wind acting as artefacts, Abbas suggested that this small temperature variation can be treated as the thermal signature of respiration. However, the prerequisites could be hardly guaranteed in a non-controlled environment. Moreover the infrared thermal sensing technique requires a reasonable humidity level to work.

Another non-contact based approach which tries to compensate for the movement associated drawbacks of these ones, is based on a complicated computer vision technique called optoelectronic plethysmography [118]. Both human body and face tracking technologies are on-going popular research topics and are divided into two categories: marker and marker-less. Marker techniques require distinct makers attached to the tracked subject. In general marker-less vision based approaches are less reliable and practical than the marker based ones. Furthermore this

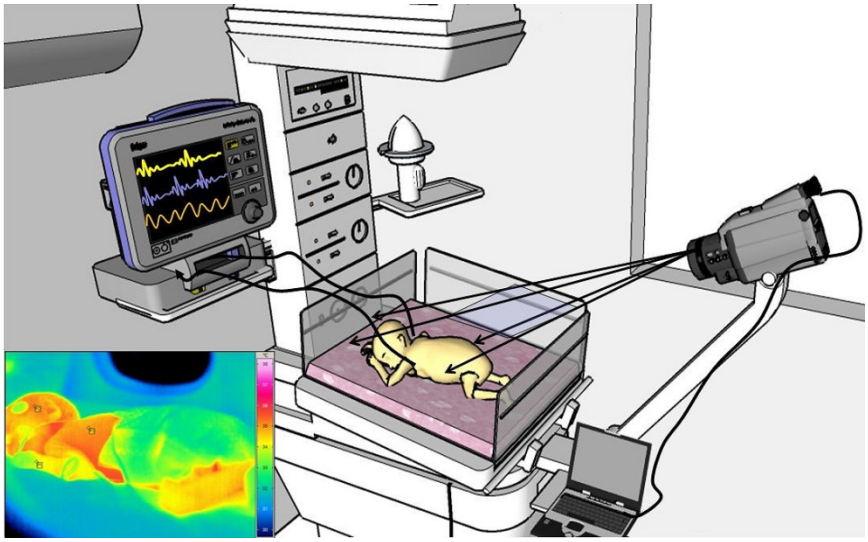

Figure 1-5: Clinical infrared thermography setup for infant respiratory monitoring [117]

kind of non-contacting vision/imaging based sensing techniques requires more than one set of specialised sensing equipment to cover the whole monitoring environment. This results on very limited clinical use, and not being still suitable to use in non-controlled free-living environments.

1.2.3.2 Electrocardiogram Derived Respiration (EDR)

Another way of monitoring the presence of breath is to observe the short-term irregularity of the heart electrocardiogram (ECG) caused by respiratory activity. This method has been used since mid-1900s [119]. Similar to EIP, the electrocardiogram derived respiration (EDR) approach observes the movement caused by chest expansion and contraction from ECG electrodes, however without the need of any external current injection. The actual small amplitude ECG signal generated from the heart acts as the stimulation from inside the body.

By observing the variation of the peak-to-peak amplitude change of the ECG signals, the relative chest/abdomen excursion can be estimated, though the resolution is lower than its active mode EIP approach. In addition to sensing impedance variation, the EDR approach measures the tiny heart rate variability which is in synchrony with respiration: the instantaneous heart rate is faster during inspiration and slower during expiration. This phenomenon is called respiratory sinus arrhythmia (RSA) and was first discovered by Carl Ludwig in 1847 [120]. In [121] it was concluded that there are two possible mechanisms to explain RSA: 1) blood pressure changes influenced by respiratory movements; 2) autonomic nervous system.

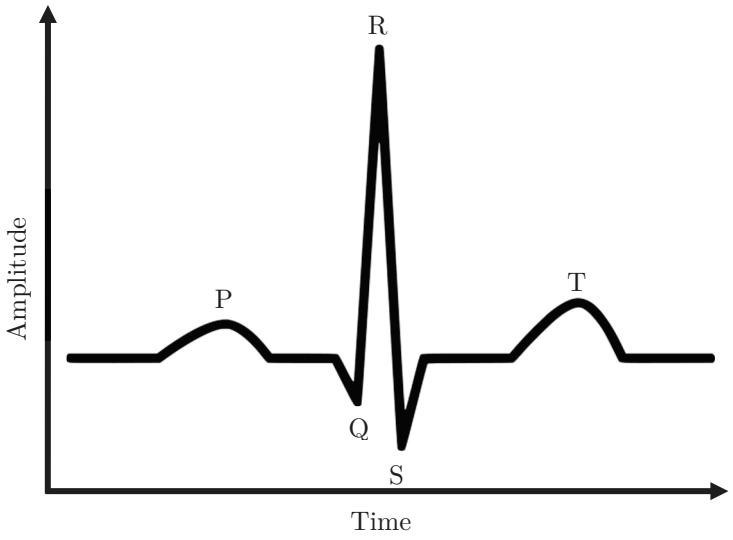

Figure 1-6: An illustration of an electrocardiogram (ECG) waveform with PQRST characteristics in one complete heartbeat cycle.

A typical ECG waveform is shown in Figure 1-6 with its five characteristic features denoted as P, Q, R, S and T. Each peak and valley are generated by contraction and expansion of various heart valves in different blood pumping stages. The amplitude of the ECG signal, specifically the QRS points (Figure 1-6), are affected by the chest expansion and contraction due to respiratory events. Furthermore the RSA effect translates on a reduction of the interval between two adjacent R points, during inspiration, and an increase during expiration. Hence, by analysing the amplitude and frequency modulation of the ECG signal, it is possible to detect the presence of breath. This has been extensively researched, but the majority of systems have used multiple ECG channels which are too cumbersome as well as difficult to attach, to be considered a wearable technology. Although some research [122-124] has shown that it is possible to extract respiratory rate from a single EEG channel, most of the reported methods eliminate the data corrupted by movement artefacts in their validation methodologies and have proved this in very short data samples.

Additionally, the presence and intensity of RSA is dependent on two factors: age and health status. And some researchers [125-127] have already shown sinus arrhythmia could still be observed at the beginning and at the end of a voluntary breath hold, which could give false breath detection around the apnoea events. An unbiased review [128] tested 6 different EDR methods on a single EEG channel data obtained from a free-moving subject and showed that the lowest achievable error rate between observed and derived measurements was between 18%

to 30%. A recent report [129] showed that EDR systems are generally highly susceptible to the subject's position since median errors on breath rate estimation could go up to 87% in certain positions. Also, this approach was still sub-optimal from the point of view of wearability because it still required two electrodes (measuring the channel and a reference) positioned far from each other. Generally, ECG based respiratory monitoring is difficult to use, requiring a fully trained specialist to put on the sensors, and heavily suffers from interference caused by body movements, therefore it cannot be adapted for long term monitoring in daily life situations.

1.2.3.3 Photoplethysmogram Derived Respiration (Blood Gas Sensing)

Respiration can be monitored indirectly by observing the change in the amount of respiratory gas that diffuses through the alveoli into the arterial blood. In the blood, there is a type of protein called haemoglobin which can be of two main types: one type carries gas like oxygen, carbon dioxide and carbon monoxide; the other type cannot carry any gas. The forms that carry oxygen, carbon dioxide, carbon monoxide are called oxyhaemoglobin (HbO_2), carbaminohaemoglobin (CO_2Hb) and carboxyhaemoglobin (COHb) respectively. A form that can carry oxygen but without bonding to oxygen is called deoxyhaemoglobin (Hb). Only HbO_2 and CO_2Hb can offload their gas quickly and reverse back to Hb which can be bonded with oxygen once again. The release of carbon monoxide CO from COHb is not easy and can take place very slowly (that is why CO is toxic to the human body). The last form is called Methaemoglobin (metHb) which cannot carry any gas. In most cases, metHb and COHb account for less than 3% of the total amount of haemoglobin, and the remaining functional haemoglobin that can carry oxygen dominates in the blood [130].

Following the development of non-invasive photoplethysmography during the 1970s, monitoring blood gas level became a reality by mean of the pulse oximeter which nowadays is already a lightweight portable solution. The functionality of pulse oximeters is based on the different light absorption properties of different types of haemoglobin, as shown in Figure 1-7. Figure 1-8 shows a typical trace of light absorption from a finger. This is obtained subtracting the initial pass-through light intensity and the received attenuated light intensity. The AC component of the trace represents the absorption from pulsatile arterial blood due to heartbeat; while the DC component represents the absorption from the constant part of arterial blood, venous blood, capillary blood and tissue.

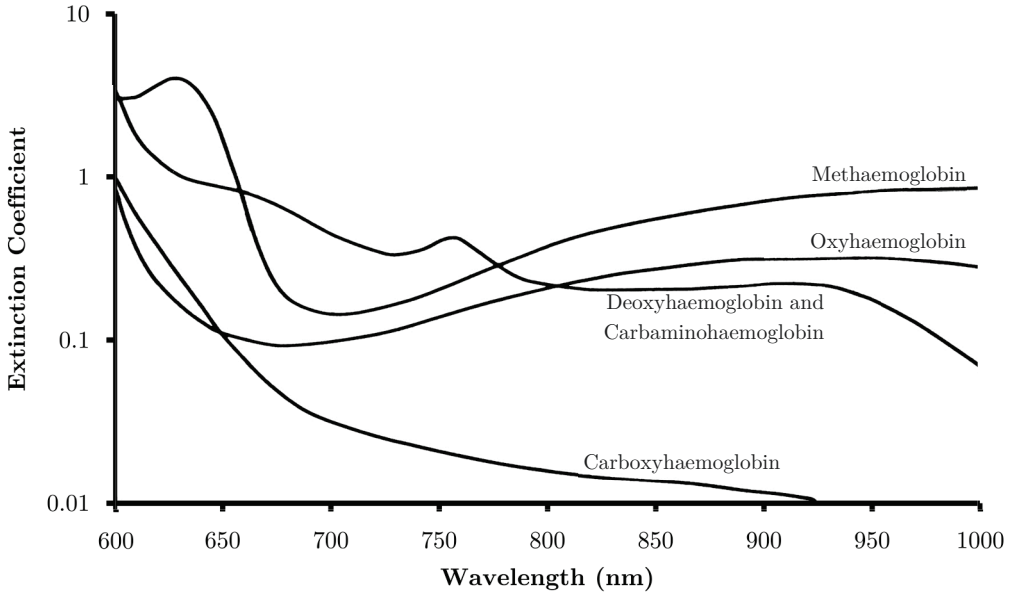

Figure 1-7: Absorption spectra for three main types of haemoglobin: Methaemoglobin (metHb), oxyhaemoglobin (HbO2), deoxyhaemoglobin (Hb), carbaminohaemoglobin (CO_2Hb) and carboxyhaemoglobin (COHb)[131].

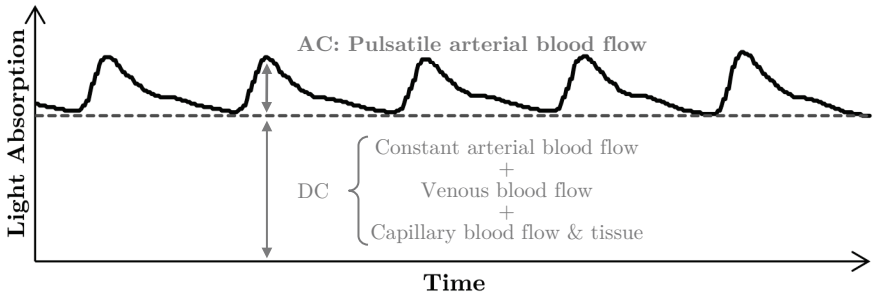

Figure 1-8: Typical light absorption waveform (black trace) from finger, where the AC component is generated by pulsatile arterial blood flow; the DC component is generated by constant arterial blood flow, venous blood flow, capillary blood flow and tissue.

Recent research efforts [132-134] have proven that it is possible to extract breathing rate information from pulse oximetry by observing the RSA from its pulsatile light signal - the same technique used in the ECG based approach. However, using the RSA obtained from the pulse oximeter to detect breathing has very limited sensitivity/specificity, even after undergoing subject specific training/calibration [135]. As illustrative numbers, 20% of recordings have been reported to be invalid due to detachment of the oximeter [136]; and 75% of the false alarms in

intensive care units are caused by pulse oximeter artefacts (with only around 8% of the total overall being real alarms [137, 138]).

In addition to the RSA information, photoplethysmography provides the level of arterial oxygen saturation (SaO_2) by examining the ratio of the oxygenated and deoxygenated oxygen carrier molecules in blood (i.e. HbO_2 and Hb). Thanks to the fact (see Figure 1-7) that HbO_2 has higher absorption at infrared light (around 940 nm) and Hb/CO_2Hb has higher absorption at red light (around 660 nm), by sensing the ratio between these two light absorptions at the fingertip or ear, where there is easy access and denser arterial blood volume, it is possible to estimate SaO_2 provided the blood has much less metHb and COHb.

SaO_2 can be approximated by the following formula:

$$SaO_2 \approx SpO_2 = \frac{HbO_2}{\text{Total Haemoglobin}} = f(I_{AC_RED}, I_{AC_IR}) \tag{1.1}$$

where

SpO_2	represents the saturation of peripheral oxygen, indicating an approximation of SaO_2.
HbO_2	is the concentration of oxyhaemoglobin
Hb	is the concentration of deoxyhaemoglobin
I_{AC_RED}	is the AC amplitude of received pulsatile intensity at red light
I_{AC_IR}	is the AC amplitude of received pulsatile intensity at infrared light

Because the light travelling through the testing sample is scattered inside the tissue and blood, the Lambert-Beer law which relates the attenuation of the light to the characteristics of the material through which this is traveling cannot be applied to obtain the function, *f*, shown in equation (1.1) directly. Thus, the actual function is interpolated from a look-up table which was obtained from experimental invasive oxygen saturation measurements on a certain number of healthy people.

SpO_2 measurements are commonly used for evaluating heart rate and detecting blood oxygen desaturation events. Anything more than a 3% change from the baseline of a pre-event blood oxygen saturation level is considered as a blood oxygen desaturation event. This may indicate the absence of breathing provided there is no misplacement of the oximeter.

Clinical validations [139, 140] have shown that the SpO_2 results from oximetry have a delayed response to hypoxemia events, which are defined as SaO_2 below 90%. The latency can be as long

as 20 seconds, depending on the distance between the sensor and arteries, as well as the subject's medical and metabolic condition. The rate of fall in SaO2 during apnoea varies from individual to individual, changes between awake and sleep [141, 142], and also depends on the duration and frequency of apnoeic events. These factors weaken the correlation between apnoea and peripheral blood saturation level. This means that not every apnoeic event triggers SpO_2 desaturation. Therefore, although a SpO_2 desaturation event can be treated as a critical respiratory distress signal, accurate detection of apnoea cannot be based solely on it.

1.3 Design Considerations for a Wearable Respiratory Monitoring System

Previous sections have shown how most of state-of-the-art respiratory monitoring techniques require a well-controlled clinical environment to achieve their claimed performance. To create a robust clinical grade wearable respiratory monitoring system, a minimum set of features have to be targeted. The following section discusses the requirements for the different features putting them within the context of the gold standard in the clinical setting. The ultimate goal of this work was to maintain a performance comparable to this gold standard and to eliminate the need of a clinical personal, in order to maximise the new technology usability.

1.3.1 Respiratory Features

1.3.1.1 Respiratory Rate

Respiratory rate is defined as the number of respirations (inhalation-exhalation cycles) occurring during a specific time frame. Practically the time frame should be 1-minute in order to avoid misleading readings for slow and irregular respiratory conditions. Acquiring respiratory rate is essential in respiratory monitoring because most respiratory failure events can be detected from this parameter.

Abnormally fast respiratory rate at rest indicates hyperventilation (fast and deep breathing) or tachypnoea (fast and shallow breathing) both of which are respiratory dysfunctions. Hyperventilation makes the removal of CO_2 too fast, leading to blood vessel constriction due to blood alkalization, resulting in insufficient oxygen delivery to the body, causing anoxic effect, eventually self-promoting the autonomic nervous system for faster respiration [143]. Tachypnoea,

on the other hand, can be the first sign of patient deterioration associated with acidosis, infections or cardiopulmonary disease.

Slow respiratory rate, on the opposite side, could cause low blood oxygen saturation level (hypoxemia), and in some cases even increase the level of CO_2 in blood (hypercapnia). Both hypoxemia and hypercapnia can be life-threatening respiratory depression conditions. With a 1-minute evaluation window in non-exercise conditions, high respiratory rate could reveal respiratory dysfunction in most cases, but low respiratory rate does not always lead to adverse situations because it depends on the efficiency of the individual's respiratory system. However, by using different time windows, other than 1 minute, to analyse breath patterns (like for detecting repetitive apnoea/hypopnoea as will be discussed in the next section), respiratory failure could be foreseen. The average respiratory rate of a healthy adult is 18±4 breaths per minute (bpm) with a upper limit of 24 bpm to indicate hyperventilation/tachypnoea syndrome [144] and a lower limit of 8-10 bpm to indicate hypoventilation [145, 146].

As mentioned in the motivation section, in critical clinical environments, respiratory rate is still done by the nurses because there is no automatic way to reliably monitor it. The clinically acceptable accuracy of respiratory rate for an automatic monitoring system is ±20%, which is considered to be as good as the manual measurement performed by a nurse [147]. In addition, because the fastest normal respiratory rate is 22 bpm and the threshold for hyperventilation is 24 bpm, the reasonable accuracy between 22 bpm to 26 bpm should be within ±2 bpm in order to avoid raising any false alarms in those cases.

1.3.1.2 Lung Volumes

The term lung volumes is used to refer to volumes of air associated to respiration. There are different definitions for lung volumes. For example, lung tidal volume (V_t) accounts for the normal volume of air displaced between normal inhalation and exhalation; forced vital capacity (FVC) is the maximum volume resulting from a maximally forced expiratory effort after full inspiration; forced expiratory volume (FEV) is similar to FVC but with the volume quantified in a specific time interval.

Adequate V_t is required to ensure fresh air is not just staying in the respiratory airway, but sufficient oxygen is dispersing into the lungs. The critical volume, called physiological dead space, is around 2 ml/kg of body weight or 150 ml in average [148]. If the V_t is constantly lower

than the dead space, fresh air will not be able to reach the lungs and perfuse the alveoli, causing respiratory depression. On the other hand, FVC and FEV are the most important measurements of lung function. Normal healthy adults have 3 to 5 litres of FVC, where around 70 - 80% of the FVC should be obtained from FEV in the first second [149]. Decreased FVC indicates restrictive lung diseases which is often associated with obesity [150]; while low FEV implicates obstructive lung diseases such as asthma and chronic obstructive pulmonary disease (COPD).

Unfortunately, accurate lung volumes measurements can be only obtained continuously using a respirometer connected with a facemask which is not a comfortable method because the tube connected to facemask artificially increases subject's dead space because of additional volume of air need to pass the tube but not participate in respiratory gas exchange, hence the subject has to breath deeper than their normal way. Sophisticated and expensive respirometers in the hospital are able to achieve the gold standard accuracy $\pm 3\%$ or ± 50 ml, whichever is greater [151]. Besides respirometry, lung volumes could in principle be extracted using other sensing modalities for airflow sensing, such as the one [152] reported, by integrating the airflow over the duration of a breath. The quality of the model for quantitative airflow measurements does however affect the accuracy. When lung volumes, especially tidal volume, are to be calculated by automatic means, accurate airflow determination is not the only factor affecting accuracy. The temporal characteristics of the signal, and most specifically, precise determination of integration times are also equally important. These integration times are not difficult to determine visually, mostly when dealing with artefact free segments of the signal. The problem becomes significantly more challenging when lung volumes have to be calculated automatically.

Other non-airflow based respiratory monitoring techniques like respiratory effort sensing and oximetry are also generally inaccurate or unsuitable for lung volumes determination. The respiratory effort method relies on calibrating each movement related reading to a corresponding lung volume. Hence, variations of the movement signal can be correlated to change of lung volumes in the case of FVC and FEV measurements. However, the lower the lung volumes are the lower resolution of this method is since it becomes more affected by heartbeat, body movements or obstructed airways. On the other hand, pulse oximetry only provides information on blood oxygen saturation level and heart rate derived respiratory sinus arrhythmia which cannot be correlated to lung volumes.

1.3.1.3 Apnoea and Hypopnoea

Apnoea, i.e. cessation of breathing, is defined as no air exchange happening between the respiratory system and the external environment. Any event with tidal volume smaller than the dead space is considered an apnoea. On the other hand, hypopnoea is defined as an event with exceedingly shallow breathing or a low respiratory rate which prolongs longer enough to cause a drop of arterial oxygen saturation. Fatal conditions like hypoxemia and hypercapnia can be caused by not only long prolonged apnoea events, but also frequent repetitive short apnoea and hypopnoea events, depending on the person's medical condition. However, both apnoea and hypopnoea, have no strict consensual quantitative non-invasive detection criteria such as minimum tidal volume or event duration [153].

Taking the field of sleep studies as an example, the latest American Academy of Sleep Medicine (AASM) guidelines [153] for sleep apnoeas scoring, conservatively recommends some criteria with arbitrary thresholds to detect nonfatal apnoea:

a) An event with a reduction of the amplitude of the signal associated to tidal volume by more or equal to 90% from pre-event baseline.
b) The duration of the 90% drop is at least 10 seconds for adults and the length of 2 respiratory cycles from pre-event breathing pattern for children.

There are three types of apnoea during sleep: obstructive apnoea, central apnoea and mixed of the two. The AASM suggests that classification of the apnoea could be done based on observation of respiratory effort. For measurement purposes central apnoea is defined as an event that shows absence of respiratory effort during the entire period, indicating failure of the neurological CO_2 sensing feedback mechanism. In contrast, obstructive apnoea is defined in the opposite way: the respiratory effort has to be continuous throughout the entire apnoea period. In this case, the respiratory muscles keep making respiratory efforts to change the lungs volumes, but they do not provide enough pressure to pass sufficient air across the obstructed upper airway. This implies that respiratory effort based methods are intrinsically incompetent to detect obstructive apnoea events when airflow data is missing, though in some cases paradoxical obstructive apnoea can be detected because it generates 90° out of phase chest and abdomen movements. On the other hand, airflow sensing based methods could detect apnoea but may not

be able to classify the event. A mixed apnoea is considered to be an event that is led by central apnoea, and ends by obstructive apnoea.

In the case of hypopnoea, the AASM gives the following quantitative criteria as the recommended guideline for classification [153]:

a) An event with a reduction of the amplitude of the signal associated to tidal volume by more or equal to 30% for adults and 50% for children from pre-event baseline.
b) The duration of the drop meeting the first criterion lasts for at least 10 seconds for adults and the length of 2 respiratory cycles from pre-event breathing pattern for children.
c) Oxygen desaturation is more than 3% from pre-event baseline or subject arouses from sleep within the events.
d) If a portion of the event can be classified as apnoea, that portion should not be classified as hypopnoea. The rest of the event is classified as hypopnoea.

Although hypopnoea can be also classified into the same three categories: obstructive, central and mixed, it is difficult to distinguish between them because all of them still have some respiratory effort. AASM guidelines suggest that the only reliable way to distinguish them would be to measure oesophageal pressure excursions. This cannot be done by any non-invasive approach since they do not always reflect the magnitude of respiratory effort [154]. Hence hypopnoea is difficult to classify non-invasively unless there are obvious signs of upper airway obstruction, such as desynchrony in thoracic and abdominal exclusions.

The sleep apnoea/hypopnoea detection guidelines require a polysomnography sleep study system as shown in Figure 1-9 which comes with at least three different sensors, in order to sense airflow, respiratory effort and oxygen saturation. As an example, Figure 1-10 depicts a snapshot of sleep data obtained from the gold standard system. In this case a physician is able to identify an apnoea event because the flow signal drops below 10% of the pre-event values. With the additional movement information from the abdomen and thorax band, the apnoea event can be marked as obstructive type. However, this is not an option due to the constraint in the system size imposed by wearability and usability in domestic/non-clinical environment for mobile health application. Therefore a new set of detection rules had to be created that match as much as possible the current gold standard.

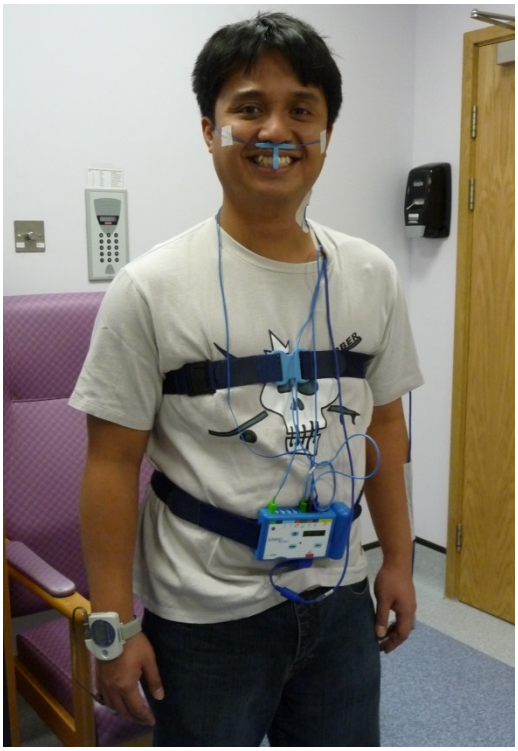

Figure 1-9: A test subject wearing an existing state of the art polysomnography system (SOMNOscreen™ RC, SOMNOmedics GmbH) and an additional pulse oximeter (Pulsox-300i, Konica Minolta) [155], comprising two finger oximeters; oro-nasal flow sensors; thoracic and abdominal expansion bands; and ECG.

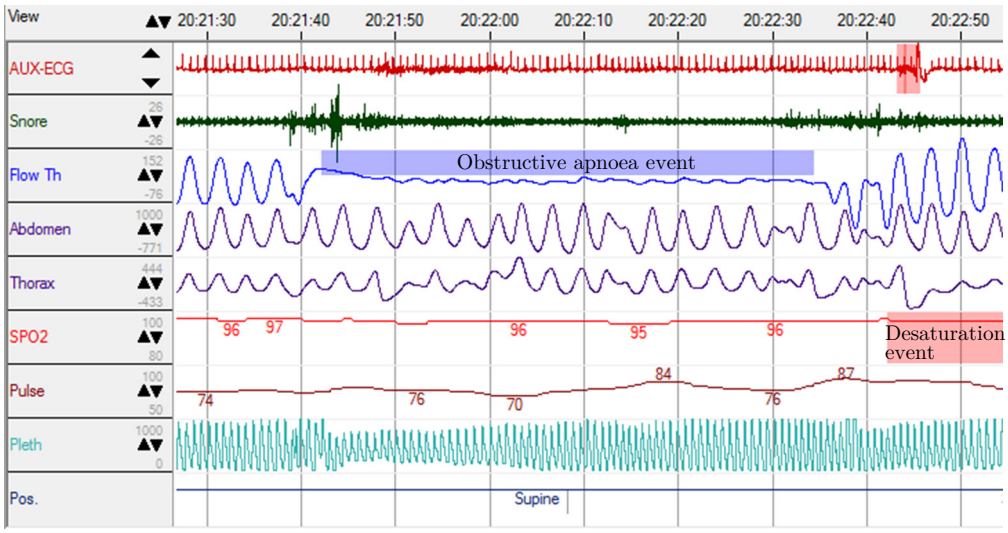

Figure 1-10: Example of polysomnography data indicating obstructive sleep apnoea.

1.3.2 Performance and Usability

In order for respiratory monitoring to work in a non-controlled environment, two factors have to be traded off: performance and usability. As an illustration of the reasons for this, the most accurate respiratory system is the polysomnography which benefits from comprehensive data provided by a number of sensors (see Figure 1-9). However, polysomnography is never used outside in-clinic sleep studies due to its vulnerability to sensors misplacement/detachment problems and movement artefacts, amongst others. And, even in the sleep clinics, physicians have to manually sift through all the raw data to search for anomalies rather than using any of the automatic sleep diagnosis software provided by the polysomnograph manufacturers [156]. One of the reasons for this is that polysomnography systems require a certain degree of agreement between the signals obtained from the different sensors in order to guarantee their accuracy, and hence can be easily affected by any failure in their sub-systems [157, 158]. As a summary, high cost of the system, discomfort caused by the sensors connected to the body, and the high amount of information that needs to be analysed, are main disadvantages of the polysomnography system. Therefore, polysomnography did not meet the wearability criteria for this work, and hence only less intrusive approaches with lower number of sensor nodes were considered. Under the wearability constraints, the ideal respiratory system should only adopt one approach amongst the ones shown in section 1.2 whilst still achieving reasonable performance in terms of the requirements listed in the last section. The approach has to be acceptable and accessible to the general public. Table 1-1 shows the ideal performance and usability matrix for individual non-invasive respiratory monitoring approaches.

1.3.2.1 Performance

In relation to respiratory rate evaluation in an ideal environment, airflow based approaches are the most accurate and RSA based approaches are the least reliable when compared to others. This is because the definition of respiration based on gas exchange and the existence of RSA could be present at the beginning or at the end of apnoea (see section 1.2.3.2). Respiratory effort sensing could generate errors when the movement caused by shallow breathing is similar to the movement caused by the heartbeat. Amongst the rest of respiratory effort approaches, movement tracking on a mattress is less reliable because it is extremely sensitive to movement artefacts.

	Methodology	Respiratory Rate	Tidal Volume	Apnoea: Detection	Apnoea: Classification	Hypopnoea Detection	Usability Issues
Airflow	Bioacoustic	★★	★☆	★★	★	★	sensitive to ambient sound; sensor misplacement
Airflow	Respiratory Gas Sensing	★★	★★	★★	☆	★	unsealed approach is sensitive to movement and other sources of airflow; flow sensor is highly intrusive, uncomfortable and easily misplaced
Respiratory Effort	Transthoracic Plethysmography	★☆	★	★☆[1]	★☆	☆	sensitive to movement; loose elastic band causes signal fading and misplacement; tight elastic band causes comfortability issue
Respiratory Effort	Electrical Impedance Pneumography	★☆	★	★☆[1]	★★	☆	sensitive to movement; electrodes misplacement; require reasonable distance between electrodes resulting on larger device size
Respiratory Effort	Mattress	★	☆	★	★☆	☆	extremely sensitive to movement; mattress thickness affects performance
Hybrid	Non-Contacting (Respiratory Effort + Airflow)	★★	★☆	★★	★★	★	sophisticated duplicated setup to achieve reasonable coverage; sensitive to changes on the media between subject and receiver; lack of robust non-stationary human/face tracking support to make it work
Hybrid	ECG Derived Respiration (RSA + Respiratory Effort	★	☆	★☆	★★	☆	sensitive to movement; electrodes misplacement; require reasonable distance between electrodes resulting larger device size
Hybrid	PPG Derived Respiration (RSA+ SpO_2)	★	×	★	☆	★	extremely sensitive to movement and ambient light; sensor misplacement

Legend:

★★ As accurate as the gold standard theoretically.

★☆ Mostly as accurate as gold standard, but not in some physiological cases.

★ Have indirect information to perform the related measurement. In some cases this may lead to less accurate results.

☆ Have unsuitable information to perform the related measurement. Mostly inaccurate.

× Have too little information to perform the related measurement accurately.

Table 1-1: Comparison of various respiratory monitoring methods

[1] Require to track both chest and abdomen movements in order to detect paradoxical breathing.

Tidal volume is calculated from the integral of airflow over time. Only approaches providing direct airflow information can guarantee the accuracy. Methods that rely on indirect measurements are less accurate. For example, when there is an obstruction in the airway, without knowing the oesophageal pressure, it is impossible to obtain accurate airflow measurements based on chest/abdomen excursions alone. EDR and EIR based approaches have limited resolution compared to inductance based transthoracic plethysmography. Especially EDR, being passive, generates smaller and noisier signals. Others like pulse oximetry have no airflow or respiratory effort related information so they are not able to give a reading on tidal volume. Of all indirect approaches, bioacoustic sensing has the potential to outperform the others because the measured signals created by airflow have a closer relationship to airflow thus providing more information about respiration [159].

Apnoea detection is based on detection of absence of respiratory gas exchange. According to the AASM's guidelines, any tidal volume smaller than 90% of pre-event average with certain duration can be considered as apnoea. Given the relaxed definition of apnoea, most respiratory monitoring techniques can achieve the accuracy in controlled, artefact free environments. For respiratory effort sensing, both chest and abdomen movements have to be monitored in order to distinguish paradoxical breathing from normal breathing. Since mattress based approaches provide less accurate measurements on tidal volume compared to respiratory effort sensing methods, they have less accuracy on apnoea detection. Both ECG and PPG based approaches rely on RSA and have difficulty on estimating the duration of apnoea. Hence they do not work as well as the others.

In terms of classification of an apnoeic event as central apnoea, obstructive apnoea or mix of the two, this could be done by respiratory pattern recognition and respiratory effort sensing. By examining the respiratory pattern before and after an apnoeic event, it is possible to tell whether the apnoea is central or obstructive in type. For patients who have obstructed airways, breathing tends to be deeper and harder before or after the apnoea, because their autonomic nervous systems force them to do that in order to avoid further drop of blood oxygen saturation. But for most central apnoeas, the breaths close to the event tend to be normal or reduce in strength. Mixed apnoea can be classified as if there is no sign of abnormal deep breathing before the apnoea event and this resumes with deep breaths at the end of the apnoea.

Using respiratory effort to determine the apnoea type is not straightforward. Most respiratory effort sensing techniques have difficulty on differentiating the movement caused by heartbeat or respiratory muscles. Although bioacoustic sensing does not provide information on respiratory effort, the sounds of pre-event breathing can be used to determine whether there is an obstruction or not. This is further explained in Chapter 2.

The last feature to be considered is hypopnoea detection, for which there is no consensual guideline. Considering the AASM guidelines as the gold standard, both airflow and blood oxygen saturation data are required. The only method to obtain blood oxygen saturation level non-invasively is PPG. However PPG does not provide any airflow related data. Amongst the rest of the approaches, airflow sensing is likely to be better in detection of hypopnoea by considering just reduction on airflow, than respiratory effort sensing.

1.3.2.2 Usability

As Table 1-1 shows, different respiratory monitoring approaches pose different kinds of issues related to usability varying from comfortability, to ease of attachment or aesthetics associated to their use.

Bioacoustic sensing is the less sensitive of all to movement artefacts; the worse being mattress and PPG based systems. Furthermore, different sensing modalities are sensitive in more or less extended to other kinds of artefacts. For example, in some cases, tiny movements caused by heartbeat may be mistaken as respiration if ECG referencing signals are not present. In addition, unsealed airflow sensing and PPG based systems could suffer from other sources of airflow and ambient light artefacts respectively. On the other hand, well designed bioacoustic airflow sensing should be less affected by movement artefacts, but it is vulnerable to noisy environments.

In terms of wearablility, non-contacting approaches seem to be perfect. However, they require sophisticated duplicated field emitting and receiving equipment to enable remote monitoring on a free living subject in a predefined living space. Moreover, they have to work with a reliable human body or face tracking algorithm running on a dedicated PC. Any changes on the media between the remote sensing unit and monitoring subject could lead to errors. Examples of these are dust, humidity and moving objects. Mattress based approaches can only be used during sleep which limits their applicability.

Other approaches have at least one sensor attached to the subject, which leads to the problem of potential sensor misplacement. In particular, transthoracic plethysmography requires not too loose and not too tight elastic bands to work properly. Loose elastic bands cause signal fading and misplacement, while tight elastic bands impair comfortability. Multi-sensor approaches, such as those involving ECG electrodes are difficult to use correctly without trained specialist help. Some single sensor approaches, such as PPG and unsealed airflow sensing, are well-known for particularly suffering from sensor misplacement problems. Compared to the others, placing the sensor for bioacoustic sensing is easier: on top of chest or front/side of the neck, more precisely, on top of the lungs or the trachea.

In the interest of aesthetics hiding the sensors inside clothes or embedding them into accessories would be desirable. From this point of view both sealed and unsealed airflow direct sensing approaches are the worst of the all.

1.3.2.3 Further Discussion

Designing a robust respiratory monitoring system for daily life use is not easy. In terms of performance for the respiratory monitoring features listed in section 1.3.1, airflow based sensing achieves the highest scores. However, non-contacting approaches are not practical because there is no reliable human body and face tracking technology to ensure that the targeted measurement is valid all the time. They also constrain the subject to a specific location. Out of the rest of airflow based monitoring methods, bioacoustic sensing provides the highest resilience to body movement artefacts is the less intrusive and has the potential for the lowest complexity. Therefore, this work focused on a bioacoustic sensing approach to be used in a non-controlled environment.

1.4 Conclusions

The aim of the work in this PhD was to create a novel clinical grade domestic respiratory monitoring system which could be used to monitor and process breathing patterns in real time for extended periods of time with excellent performance, while the user was still being able to carry out a normal life. The reason why bioacoustic measurement was chosen was because all the other sensing approaches resulted on methods that were either cumbersome or had a poor performance, mostly in the presence of movement artefacts outside laboratory environments. Therefore, any potential issues that could be caused by daily life indoor activities had to be

addressed in this work. Due to the intrinsic nature of airflow sensing techniques, some events like hypercapnia and hypoxemia, are however not detectable using this approach. However, this does not invalidated the approach because instead of detecting the existence of those events directly, warnings could be issued in the cases of abnormal respiratory rates and over a certain number of apnoea or hypopnoea events which could give users sufficient time to prevent life threatening situations from happening.

Chapter 2

Sensing Breath Sounds

2.1 Origin of Breath Sounds

The origin of human breath sounds has been studied in the field of fluid dynamics for several decades. The respiratory system can be divided into two sections: upper airway and lower airway. These are depicted in Figure 2-1. The upper airway is formed by the nasal cavity, oral cavity, pharynx and upper part of the larynx entering into the thyroid cartilage which holds the glottis. The lower airway refers to the parts below the glottis which include lower part of the larynx, trachea, primary bronchi and lungs. The whole respiratory cycle starts with air entering the upper airway through the nose or mouth, reaching the lower airway, and then the lungs where oxygen and carbon dioxide exchanges occur. Finally the exchanged air is exhaled in opposite direction back to the atmosphere.

During respiration, intercostal muscles from the chest, abdomen and thoracic diaphragm provide expansive and contractive forces which are applied to the lungs, creating negative and positive pressure in the respiratory tract which maintains the airflow for gas exchange. Inside the airway, the outer layer of air close to the airway wall flows slower than the one in the inner layer due to surface friction. Because the speed of the outer layer of air follows an exponential decay related to the overall airflow rate, larger airflows create much larger speed differences between the inner and outer layers of air. Once the difference in speed reaches above a certain threshold, a turbulence flow is generated [160]. Because of the chaotic way that air travels along the respiratory airways, it also generates vibration, and consequently an audible sound. Depending on the mechanical status of the respiratory system, two kinds of breath sounds can be produced: normal breath sounds and adventitious breath sounds. The origin of the two is described in the following section.

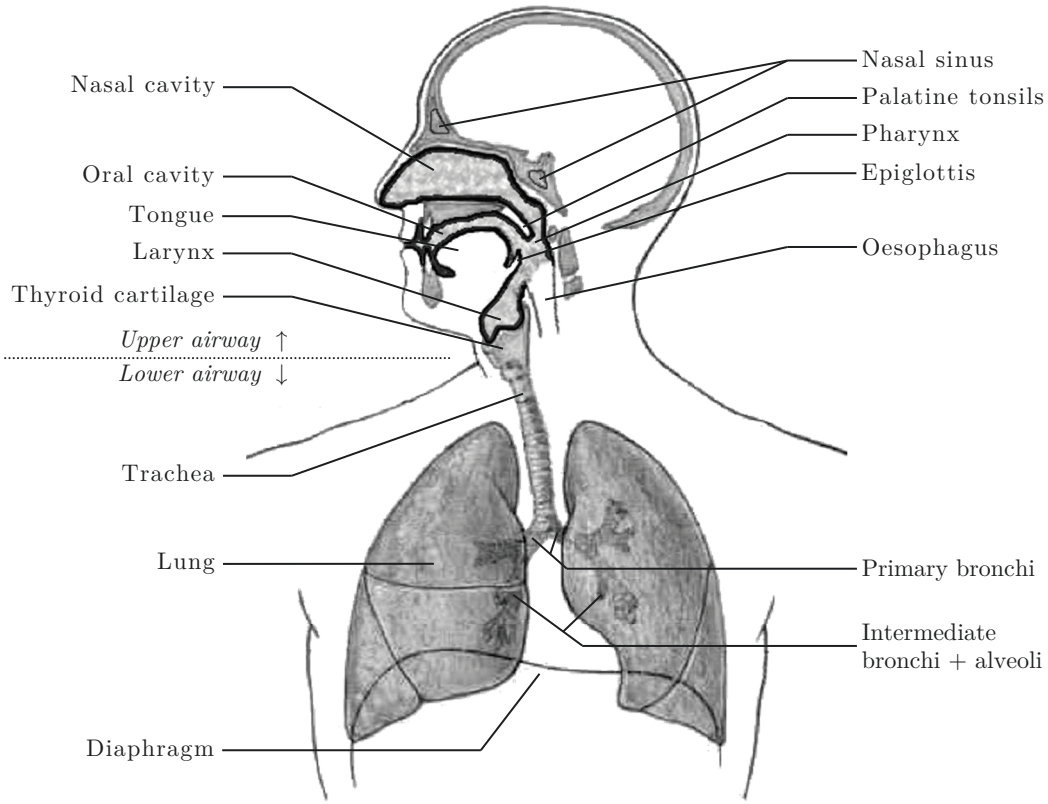

Figure 2-1: Overview of Respiratory Mechanical System

2.1.1 Normal Breath Sound

The origin of breath sounds can be explained by approximating the respiratory system as multiple straight cylindrical tubes and applying one of the principles of fluid dynamics, the Reynolds number [161]. For sound to be generated the Reynolds number has to reach a certain critical threshold shown in equation (2.1):

$$R = \frac{ud}{v} \tag{2.1}$$

where u is the average airflow rate, d is the diameter of the respiratory airway, and v is the kinematic viscosity of the air. Assuming the kinematic viscosity is constant, the critical Reynolds number is proportional to the airflow rate and the cross sectional size of the respiratory airway. Because the respiratory airways have a tree-like structure, different parts have different cross sectional sizes, gradually decreasing from the upper airway to the deep lower airway as the branches split. In particular, the trachea has lowest threshold for airflow rate to create air turbulence. In practice, literature [161] indicates that the onset of breath sound can be observed when airflow rate is as low as 200 ml/s.

2.1.2 Adventitious Breath Sounds

2.1.2.1 Snore and Stertor

Snore and stertor are low pitched obstructive breath sounds caused by irregular airflow which originates resonance in an obstructed respiratory airway [162]. Snore happens during sleep while stertor occurs during wake. Causes for upper airway obstruction include inflammation, relaxing throat muscles and collapsing tongue. These on their turn may be caused by nasal sinus, throat muscles/fat or palatine tonsils. In order for the respiratory system to maintain the blood oxygen saturation level, sufficient volume of air has to go through the reduced or blocked upper airway. During the respiratory process, the increased lumen air pressure pushes the blocking tissues around the respiratory airway, in order to enlarge the size of the latter. This causes persistent periodic drop and rise of the air pressure, leading to resonance of the upper airway structure and acoustic changes throughout the respiratory cycle. In some cases this can result on louder than normal breath sounds, with some audible harmonics in a low frequency range, which relate to the natural resonant frequency of the obstructing tissue [163].

2.1.2.2 Wheeze and Stridor

Wheeze and stridor breath sounds are like a coarse whistling tone embedded in the normal breath sound [162]. This type of breath sounds has higher pitch because of the flowing air bouncing against the wall much more quickly in obstructed airways without causing the resonance like typical of snore and stertor. Wheezing is sometimes originated in the lower respiratory airways. It is caused by obstruction of part of respiratory tree inside the lung. If this is the case wheezing sounds are not audible outside the body directly without the help of a stethoscope. Depending on how close or far away the sensing location to the affected respiratory tree is, the length of the wheezing tone varies and also the relative position within the respiratory cycle [162]. Stridor is normally generated in the larynx. The stridor sound is loud and high pitched. It can also be heard throughout the whole breath cycle [162]. Wheezing is often associated with asthma, allergy and pulmonary diseases [164]. Stridor can be caused by swallowed foreign bodies, inflammation, or tumours [165].

2.1.2.3 Crackles

Crackles can be described as normal breath sounds with added popcorn sounding noise. This noise is caused by the random volatile opening of small airways inside inflamed respiratory

bronchioles and alveoli collapsing due to sticky secretion [166]. In some cases, a cold can cause infection in the lower respiratory tract hence generating crackles during breathing [167]. In critical cases, excess of fluid in the lung which is the underlying reason for the inflammation could be an indication of a serious pulmonary disease or heart failure [168]. Similar to wheezes, crackles are produced in the lower airway, far away from the entrance of the respiratory system. Thus they can only be heard with a stethoscope placed on the chest wall.

2.1.3 Sensing Location

Breath sounds are the combination of multiple source signals generated within the entire respiratory tract including upper and lower airways. Breath sounds from one location can be transmitted through the respiratory tract to any other location of the respiratory system. For any single non-invasive sensing location such as the ear canal, neck, chest or back, the following bioacoustic signal can be obtained:

$$BS_{\text{Sensed}} \approx \frac{\left(BS_{\text{Local}} + \frac{BS_1}{\eta_1} + \frac{BS_2}{\eta_2} + \cdots + \frac{BS_n}{\eta_n}\right)}{\eta_{\text{Local}}} \tag{2.2}$$

where BS_{Sensed} is the power of the breath sound signal that can be sensed at any single position outside the body; η_{Local} is the attenuation factor of the body tissue between the sensing site and closest part of respiratory tract; BS_{Local} is the power of the breath sound signal generated at the closest to the sensing point part of the respiratory tract; $\eta_x \cdot BS_x$ is the power of propagated breath sound generated in other parts of the respiratory tract: attenuation factor of η_x. There are four general considerations that can be made for the selection of an optimal sensing location. Firstly, from equation (2.2) it is obvious that the farther away the breath sound source is from the monitoring site the larger the attenuation factors η_n will be. Also, in order to maximise the power of the sensed breath sound, the optimal sensing location should be close to the midway of the respiratory tract provided there is no significant difference on all the local sound powers generated by similar amount of airflow. Secondly, the tissue thickness between the sensing location and the underneath respiratory tract plays an important role as this determines the overall attenuation factor η_{Local}. Finally, as discussed in the previous section, thicker airways, like the trachea, have lower threshold for the airflow to generate breath sounds.

Figure 2-2: Men and Women's Chokers (Left is men's necklace from Earkandie's Celebrity Couture Collection [169] and right is women's choker from Sonoor Jewelry Concepts [170])

As a result, the suprasternal notch region, where the trachea starts seems to be the perfect sensing location, because it is midway between the upper and lower airway and the tissue above the trachea is thin. This sensing location happens to be possibly sharing a site where users are used to wear accessories. There are literatures [171, 172] confirm it is a relatively unobtrusive position for a wearable sensor since people are used to wear necklaces or chokers around that area. Figure 2-2 shows two examples for men and women wearing accessories over the neck. Both accessories have one or more pendants close to the trachea in the front or by the side of the neck. Ideally, the breath sound sensor could be embedded in this kind of jewellery.

2.2 Characteristics of Tracheal Breath Sound

2.2.1 Signal Bandwidth

Initial works on analysis of breathing from tracheal breath sounds used to be done by recording the signal obtained by a microphone with flat frequency response up to 10 kHz and carrying out off-line spectral analysis. Table 2-1 summarises the work of different groups in this area, which led to the determination of what was the optimum signal bandwidth.

Breath sound picked up from the suprasternal notch region is characterized by a broad power spectra that starts at around 100 Hz, dominating between 450 Hz to 850 Hz and becoming negligible above 1300 Hz. Figure 2-3 shows the typical spectra of expired and inspired breath signals with corresponding airflow signals in a three dimensional (time, frequency and amplitude) plot. In most of cases, the predominant spectral content of tracheal breath sound resides below 1

kHz. These include some disease symptomatic sounds such as crackles and wheezes [173, 174].

Year	Researcher	Sensor Type	Tracheal Breath Sound Spectrum		Sampling Frequency
			Measured Bandwidth	*Roll off Frequency*	
1981	Gavriely [74]	Accelerometer	75 – 2000 Hz	800 Hz	4000 Hz
1982	Charbonneau [75]	Condenser Microphone	60 – 1800 Hz	400 Hz	5120 Hz
1990	Mussell [175]	Electret Condenser Microphone	40 – 2500 Hz	500 Hz	5000 Hz
1996	Gavriely [176]	Accelerometer	75 – 2000 Hz	450 Hz	4800 Hz
1997	Pasterkamp [177]	Accelerometer	15 – 2500 Hz	800 Hz	10240 Hz
2000	Hult [178]	Electret Condenser Microphone	100 – 1300 Hz	600 Hz	5000 Hz
2000	Moussavi [179]	Accelerometer	50 – 2500 Hz	800 Hz	10240 Hz
2003	Harper [180]	Accelerometer	200 – 2500 Hz	800 Hz	10240 Hz
2007	Corbishley [181]	Electret Condenser Microphone	100 – 1300 Hz	650 Hz	11050 Hz

Table 2-1: Comparison of initial works on tracheal breath sound measurements and bandwidth results.

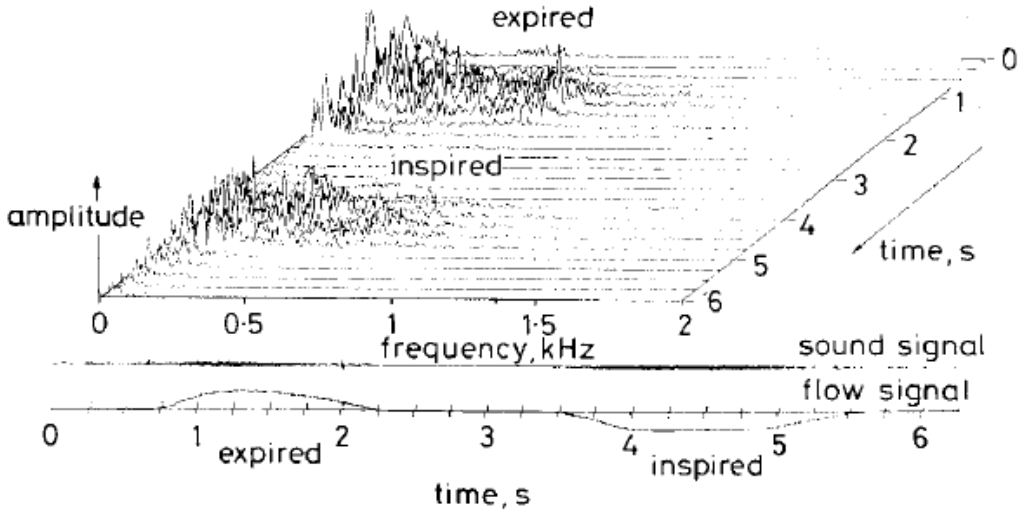

Figure 2-3: Typical tracheal breath sound and airflow signal recording (Source: [175])

Preliminary research [182] showed that frequency content higher than 1 kHz can initially be ignored for breath detection purposes because the signal-to-noise ratio is not good. Hence, this work has focused on a frequency range from 100 Hz to 1 kHz. Around 5-10 minutes of breath sounds and their corresponding airflow data was collected from 4 volunteers for better understanding of the acoustic frequency response related to airflow. The conclusions from the

initial analysis were [152, 183]: the frequency range from 350 Hz to 500 Hz has the highest sensitivity to airflow; the frequency range from 500 Hz to 750 Hz has the most linear response to airflow; and frequencies beyond 750 Hz have the lowest sensitivity. Based on this, the first low frequency band should be used for breath segmentation and the middle frequency band should be used for airflow estimation.

2.2.2 Breath Spectral Chart

Due to the variations in the structure of the respiratory system between different people, normal breath sounds from different people result on different power spectra. Hence, looking for a common feature in the frequency spectrum can be a challenging problem. In acoustics, one of such features is called timbre. An expert physician is able to determine some lung diseases from certain types of breath timbre during auscultation. Some researchers [184] have also found that some animals can even recognise individual humans by the unique timbre of their breath sounds. The breath timbre not only varies from people to people. It can also change with airflow speed and direction. For example, the pitch of expiration is generally higher than the pitch of inspiration, but this is not always the case. Moreover a person's timbre could change from time to time following changes of the respiratory airway geometry due to different physical activities.

Previous work[1] [182] analysed the common characteristics of breath sounds by mean of a breath frequency chart (Figure 2-5). In the original breath frequency chart, three arbitrary frequency ranges were selected for breath sound characterisation. These were evenly divided within the 500 Hz to 900 Hz frequency band with 50 Hz frequency separation. The chart showed the relationship amongst the energies of these three frequency bands by plotting the energy of the third band against the second band with both normalised by the one from the first band. The energy of each frequency band was estimated from the integral of its root mean squared (RMS) power generated with an envelope signal generator as depicted in . The channels were separated by a 6^{th} order Butterworth bandpass filter. The filtered signals were full-wave rectified and then smoothed by a lowpass filter with a cut-off frequency not lower than the fastest repeat rate of the shortest breaths. In case of adults, the shortest breath can be as short as 0.5 seconds. Hence, a 2 Hz lowpass filter was used to smooth the envelope signals in order to estimate their corresponding RMS power.

[1] Note that I am the author of that MSc project.

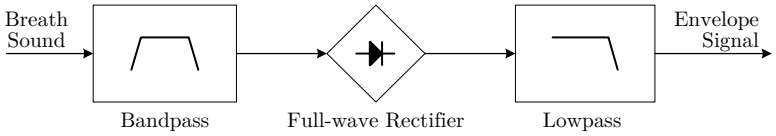

Figure 2-4: Envelope signal generator for breath sound root mean squared (RMS) power estimation

In that study, 5 healthy adults' tracheal recordings are used. The subjects were 4 males and 1 female with an age range of 22 to 24. They were asked to sit still and breathe normally. Breath samples were picked up at least 10 minutes apart from each other. The data generated a breath frequency chart as the one shown in Figure 2-5. This chart showed that all breath samples were clustered together within a triangle shape region, hence sharing similar power spectral distribution in those three frequency bands. However, previous work did not explain why all breath sounds shared similar power spectral distribution. A simplified explanation behind this is that the breath sound is a harmony produced by a bunch of airways acting like wind instruments inside the respiratory tree, as previously stated in equation (2.2). Depending on the diameter of the respiratory branches, air flowing through them generates sounds with different pitch, which combined together become breath sounds. Therefore, as long as the constitution of the 'wind orchestra' does not vary from individual to individual a lot, their breath sounds share similar distribution in frequency spectrum.

This PhD extended the breath characterisation work stated above, with 13 more adults were tested, including 10 males and 3 females with an age range of 22 to 28. This made the total number of subjects to 18, tested in the same still pose condition. By repeating Huq's tracheal breath sound analysis [185], characterisation of the breathing signal led to the confirmation that the optimal signal to noise ratio (SNR) occurred in a frequency range between 300 Hz to 700 Hz, i.e. 200 Hz variation from the original estimated bands. As Figure 2-6 shows, including the previous 5 volunteers' breath data together with the additional 13 healthy subjects, shifting down the bands by 200 Hz makes the breath sound more discriminable since the plots are more concentrated in smaller area. This is due to the improvement on the SNR which leads to better results after normalisation.

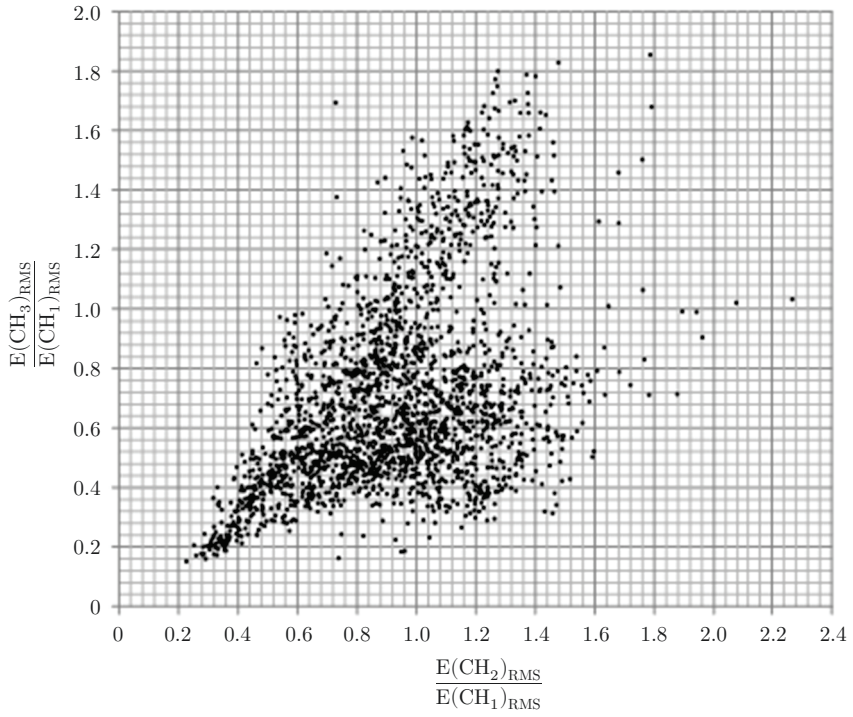

Figure 2-5: Breath frequency chart from previous work [182] obtained from 5 volunteers, where $CH_1 = [500\text{ Hz}, 600\text{ Hz}]$;$CH_2 = [650\text{ Hz}, 750\text{ Hz}]$; $CH_3 = [800\text{ Hz}, 900\text{ Hz}]$.

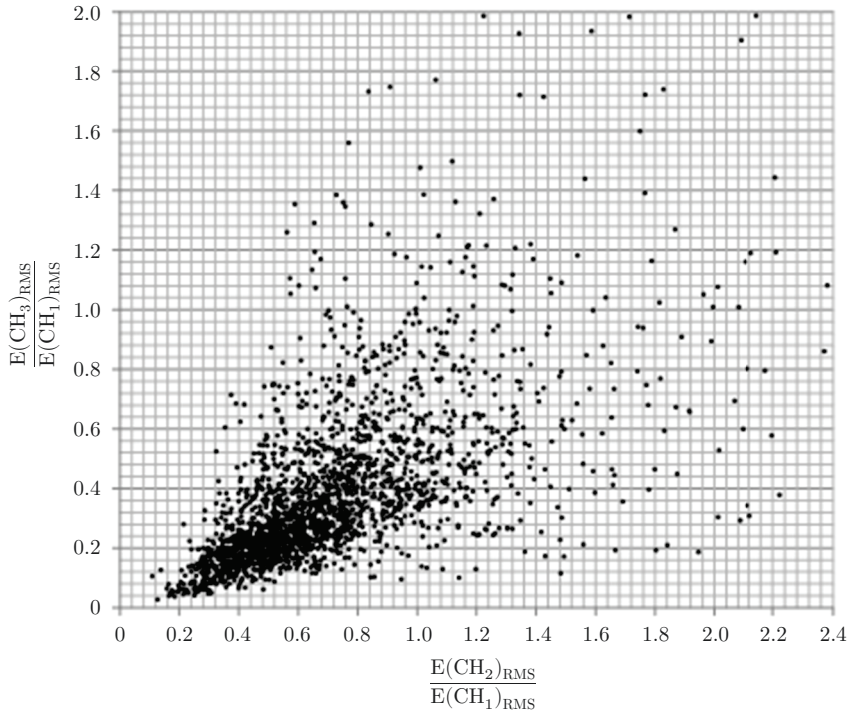

Figure 2-6: Redefined breath frequency chart obtained from 18 volunteers, including the ones in Figure 2-5, where $CH_1 = [300\text{ Hz}, 400\text{ Hz}]$; $CH_2 = [450\text{ Hz}, 550\text{ Hz}]$; $CH_3 = [600\text{ Hz}, 700\text{ Hz}]$.

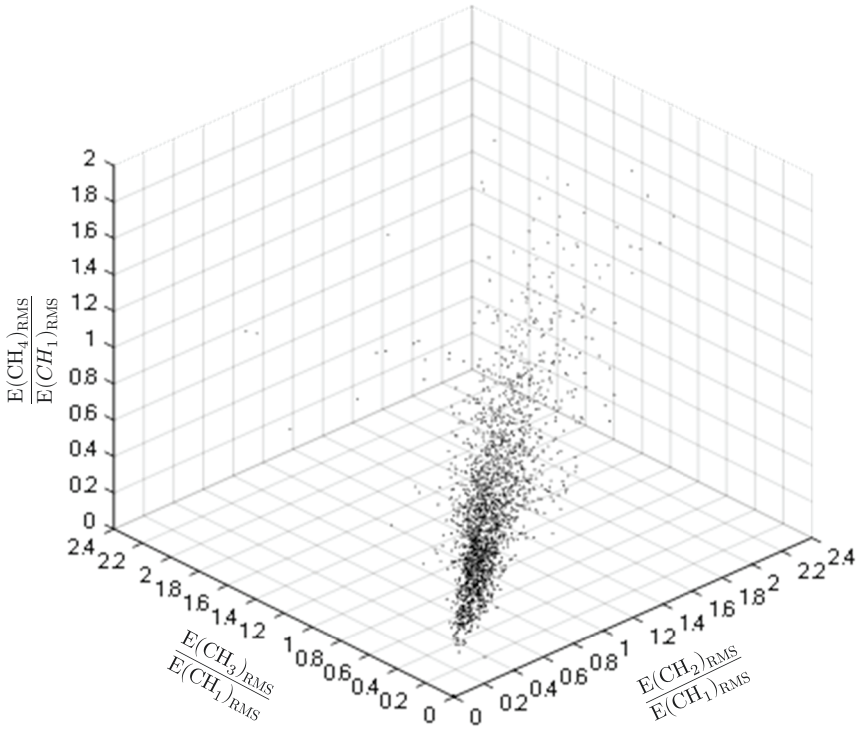

Figure 2-7: Extended breath frequency chart obtained from 18 volunteers, including the ones in Figure 2-5, where $CH_1 = [300\text{ Hz}, 375\text{ Hz}]$; $CH_2 = [425\text{ Hz}, 500\text{ Hz}]$; $CH_3 = [550\text{ Hz}, 625\text{ Hz}]$; $CH_4 = [675\text{ Hz}, 750\text{ Hz}]$.

Furthermore, the concept of the breath spectral chart can be extended multi dimensionally by introducing one or more frequency bands within the bandwidth of the breath sound, hence effectively increasing the resolution of each band. As an example, Figure 2-7 shows the breath spectral chart for 4 channels. The 3-dimensional chart reveals the breath sound samples are concentrated within a cone shaped region.

2.2.3 Temporal Characteristics

Although preliminary research mentioned temporal characteristics of breath sounds [182], these were not comprehensively analysed. Using the data from the 18 volunteers mentioned in the previous section, temporal characteristics were extracted from the 300 Hz to 400 Hz frequency channel. This one was chosen because it empirically demonstrated to have better sensitivity to airflow. However, similar temporal characteristics could be observed from the rest of the channels, with delayed or slower response.

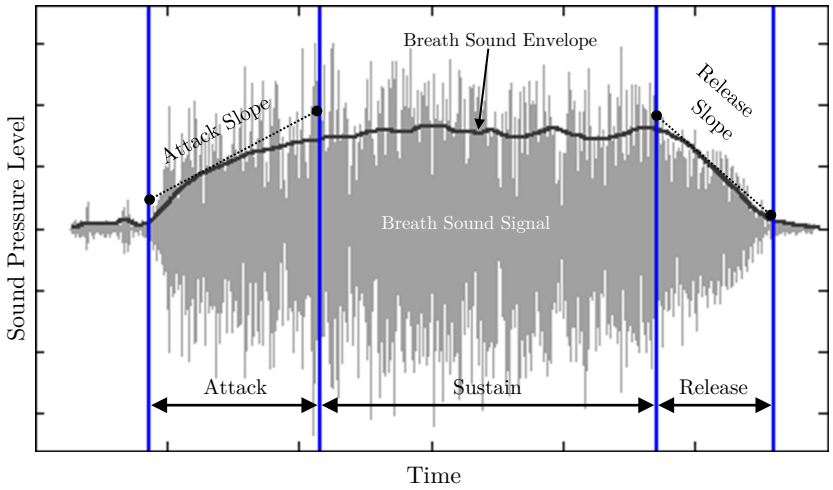

Figure 2-8: Temporal characteristics of a breath sound

As shown in Figure 2-8, a typical breath sound can be segmented into three parts according to their energy level: attack, when respiratory effort is initialising; sustain when the effect reaches peak and steady state; and release when the effort is stopping. In most cases, spontaneous breaths have distinct characteristics on those temporal features. These are further explained in the following.

2.2.3.1 Attack

The attack region for the purposes of breath detection was defined between the start of the breath and the first point when the slope of the breath envelope declined to 30% of its maximum slope. Detailed description of how to locate the breath start position can be found in Chapter 5. The length of the moving time window for slope calculation was set to 66.7 ms.

The breath samples obtained from the 18 volunteers showed that the majority of spontaneous breaths have attack time between 50 to 500 ms, with typical value of 210 ms, as shown in Figure 2-9. The acceleration during breath initialisation can be determined from the slope of the attack. A histogram of breath attack slope is shown below in Figure 2-10. It can be seen how it mostly varies from approximately 25 to 95 dB Sound Pressure Level (SPL), peaking at 55 dB SPL. The minimum human audible SPL is 20 uPa.

2.2.3.2 Sustain

Sustain region was defined as the time period when the respiratory effort reaches peak and maintains a steady level of airflow [186]. The sustain region sits between the end of the attack and the start of the release sections. The definition of when the sustain section ends is described

in the next section (i.e. the start location of release). For most breaths, the sustain region of their envelopes is smooth without any sudden descent. However, in some particular cases, the respiratory effort can quiver to the level below 30% of their peak value during the sustain period.

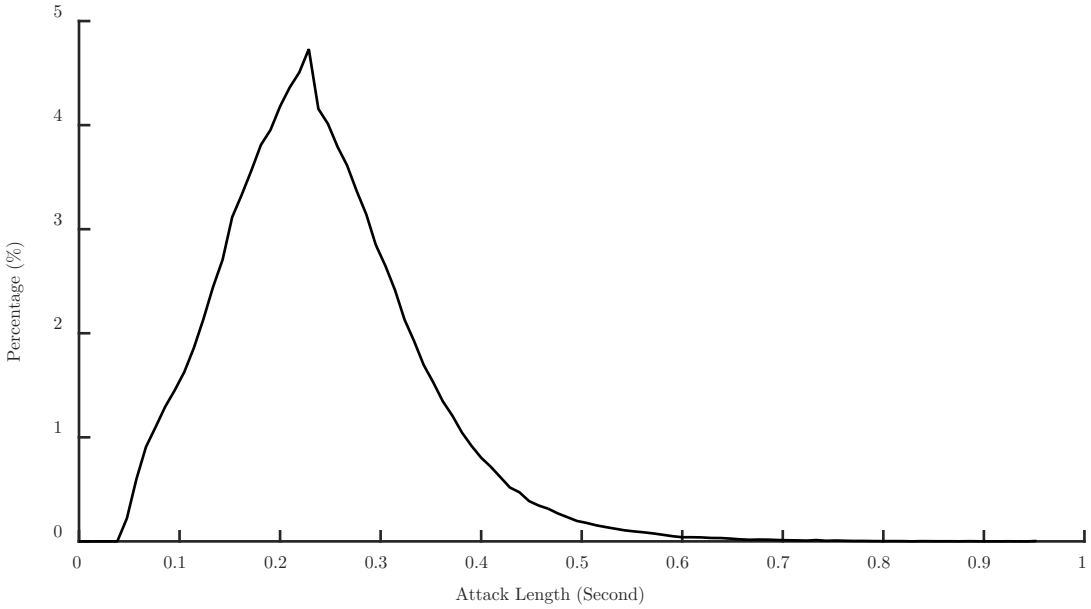

Figure 2-9: Histogram of breath attack duration

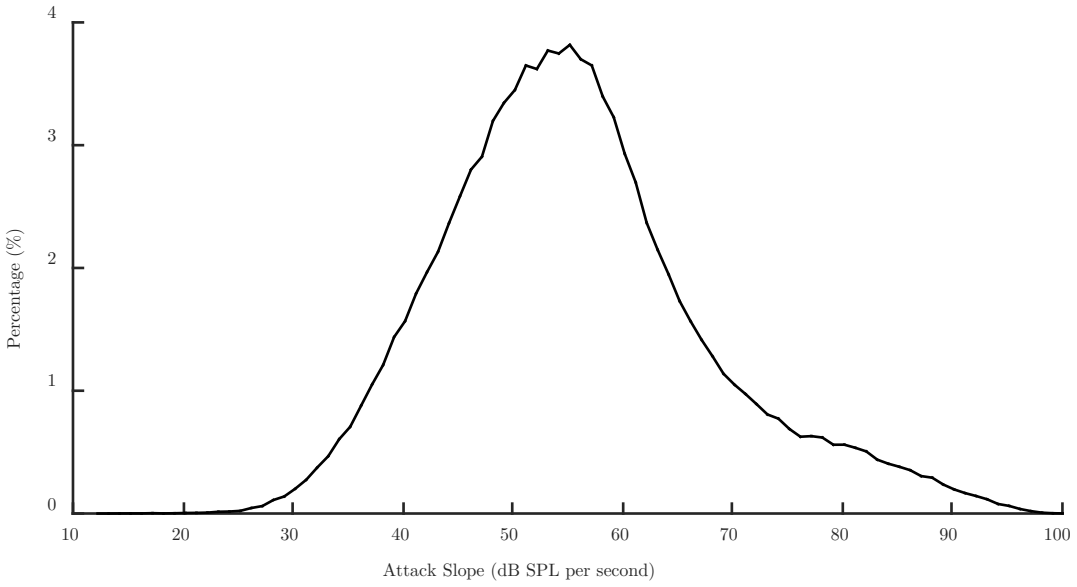

Figure 2-10: Histogram of breath attack slope

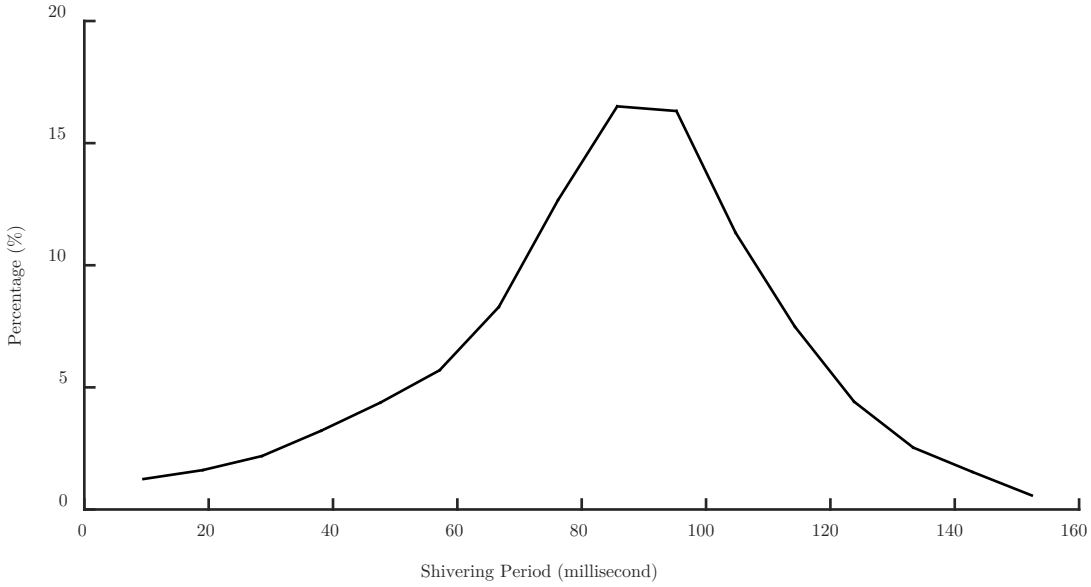

Figure 2-11: Histogram of breath shivering period during the sustain region

This may impose a challenge for separation of the breath sound signal from ambient noise by using a simple threshold strategy. Within the breath samples obtained from 18 subjects, only 1.6% of breaths had shivering sustain which indicates airway pressure fluctuation due to lack of adequate respiratory effort compared to airflow resistance [186]. Most of the quiver happens in the second half part of the breath. The observed shivering period between the envelope signals dropped below the 30% point and returned to the same level in less than 160 ms. The histogram of occurrence of the shivering period is given in Figure 2-11. In the rest of the text, this kind of breath segments is referred as shivering breath segments.

2.2.3.3 Release

The release region of a breath segment starts when the respiratory effort retracts till it eases. The search of the release region starts near the last position when the envelope signal dropped below 45% of its peak and does not recover within 160 ms (i.e. not to be considered as the sustain region of a shivering breath). The start of release is defined as the last time the envelope of the signal dropped below 70% (-3 dB) of the mean of the section between the end of its attack and the search starting point. Once the start of the release region is defined, the next step is to search for the absolute maximum of the release slope, with the same time window length used in the attack slope evaluation which is 66.7 ms. The end of the release region is found when the absolute value of the negative slope decreases below 30% of its maximum value.

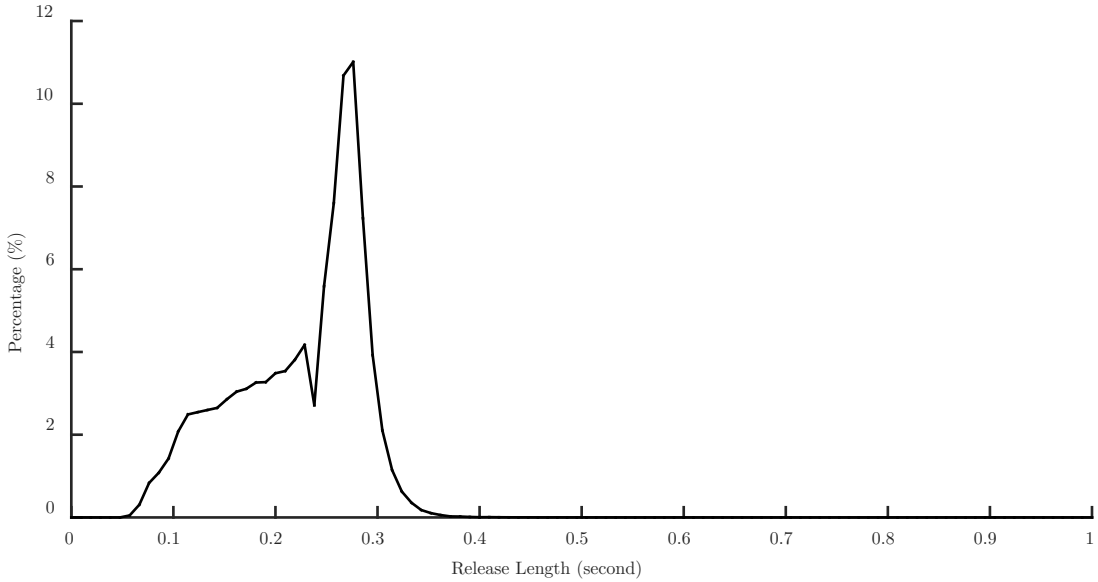

Figure 2-12: Histogram of breath release duration

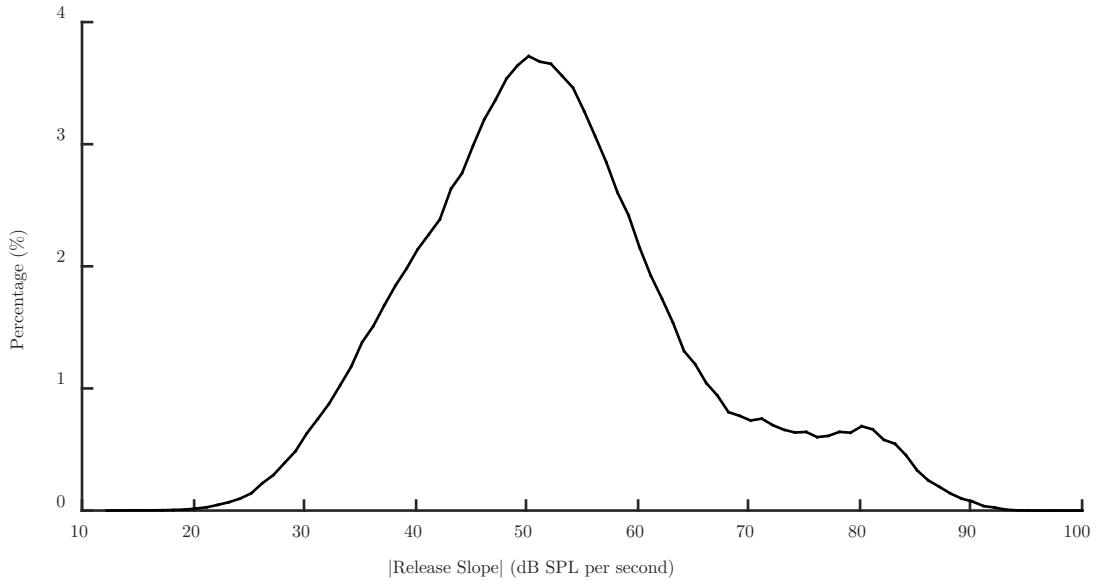

Figure 2-13: Histogram of breath release slope

The statistical characteristics of the breath release duration and slope are given in and respectively. Compared to the temporal features of the breath attack, the results show how the length of release is more concentrated at a longer duration, 276 ms, and also has a lower value of slope at around 50 dB SPL. The physiological explanation for this is that in most cases, the speed of respiratory muscles retracting their efforts is slower than the speed of initiating.

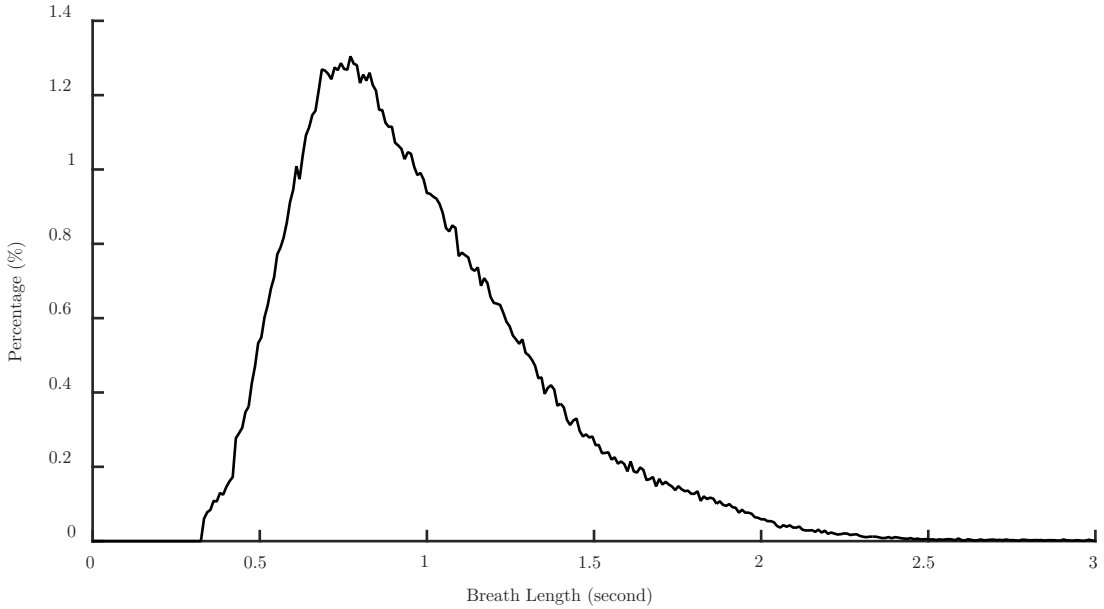

Figure 2-14: Histogram of breath length

2.2.3.4 Breath Length

The total length of a breath sound is the sum of the duration of the attack, sustain and release time periods. Statistical results reveal that the majority of spontaneous breaths durations for adults are between 0.4 to 2.5 seconds, as shown in . The minimum breath length is proportional to the minimum amount of air exchange required for supporting life and inversely proportional to respiratory rate. From observation, in the case of voluntarily breathing adults, the duration can be higher than 2.5 seconds, but it is unlikely to be shorter than 0.4 seconds unless the subject is hyperventilating.

2.2.3.5 Breath Rhythm

The term breath rhythm is used to quantify how inspiration and expiration occur within a respiratory cycle as well as the time distance between two adjacent respiratory cycles. To help explain this, two types of breath pauses are considered: the time gaps between inspiration and expiration within a respiratory cycle and the time gaps for a respiratory cycle to repeat. From observation using the available breath data set, the interval between an inspiration and the following expiration has smaller variations and concentrates below one second. In comparison, the interval between an expiration and the following inspiration has a broader distribution, being the dominating factor in respiratory rate changes. Figure 2-15 shows the comparison of the histograms for these two intervals. To analyse the relationship between the breath intervals, Figure 2-16 plots the distribution of the ratio of expiration-inspiration interval over the

following inspiration-expiration interval. Although the expiration-inspiration interval has broader distribution, the results show how it is generally several times longer than the inspiration-expiration interval.

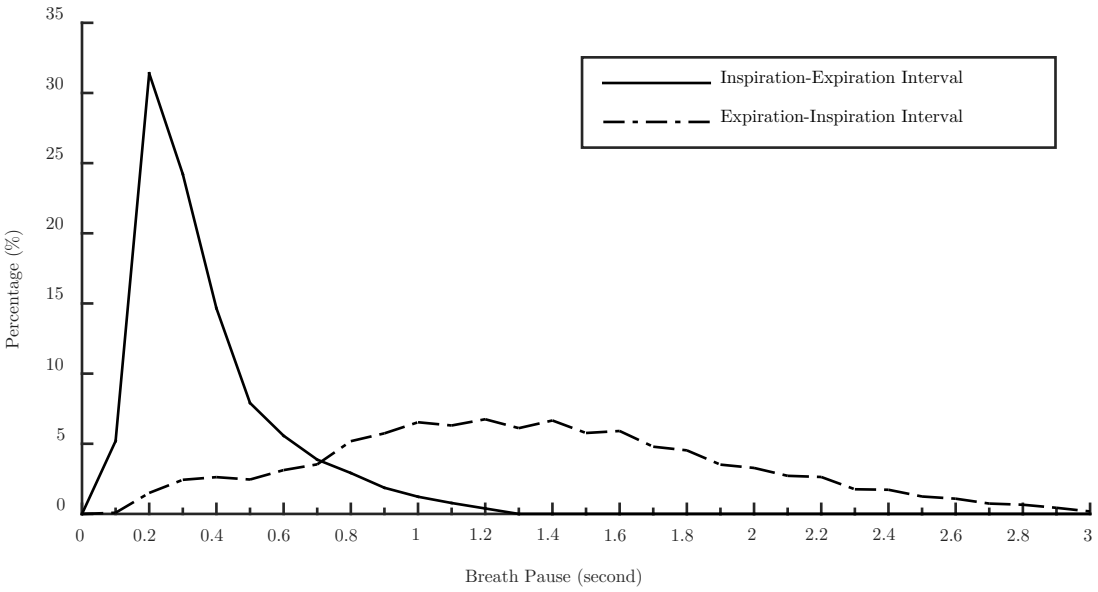

Figure 2-15: Histograms of breath pause for two cases: solid line for the breath pause within one respiratory cycle; dashed line for the breath pause between two adjacent respiratory cycles.

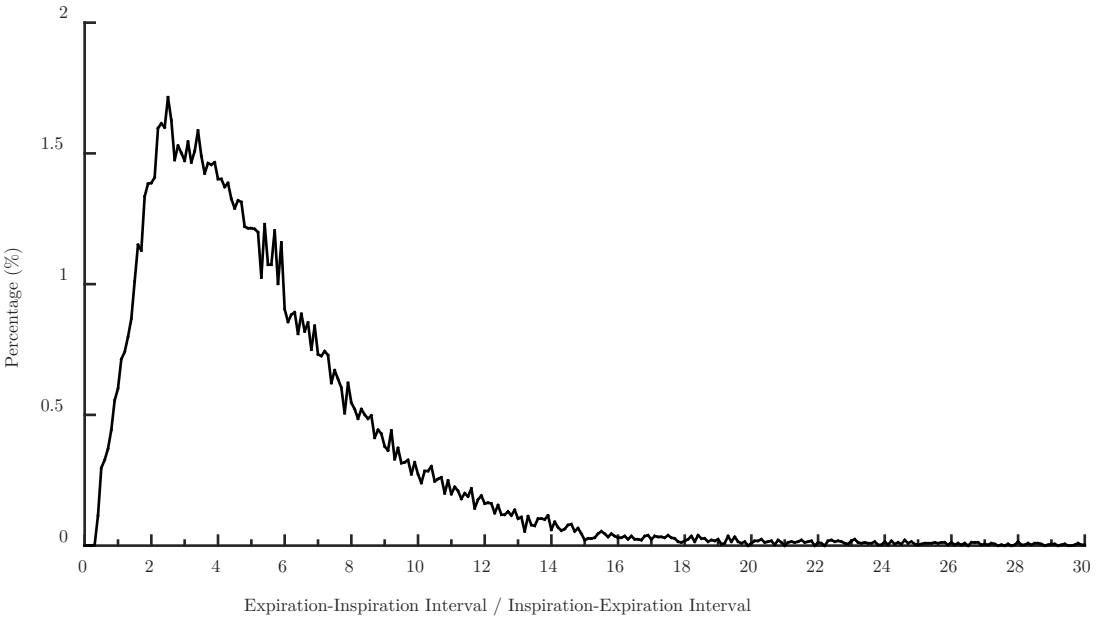

Figure 2-16: The relationship between the adjacent breath pauses

2.3 Acquisition of Breath Sounds

2.3.1 Review of Acoustic Sensors

2.3.1.1 Piezoelectric Microphone

A piezoelectric microphone is a contact type of microphone, which makes use of the piezoelectric effect in certain materials to measure dynamic changes in mechanical variables. When a piezoelectric material is stressed or pulled, it can convert this mechanical movement into electrical energy. Usually the sensors have effective flat frequency response up to a few kHz [187] which is enough for breath sounds measurement. This can be done by fixing them to the skin, and directly sensing the skin vibration caused by airflow inside the respiratory airway. Because the tissue is the only transmission media between the sensor and the airway, this kind of sensor has direct access to the energy generated by the breath air turbulence. For this reason, piezoelectric accelerometers were once the preferred transducer to analyse the characteristics of respiratory sounds. High sensitivity to any mechanical vibration, nevertheless, makes the sensors much more vulnerable to mechanical noise. Similar to most of EEG and EKG systems, the amplitude of mechanical noise could be much stronger than the breath sound, which could saturate the input circuitry and lead to failure of the whole system. Moreover, as shown in , a typical frequency response of a piezoelectric microphone has higher gain after its usable frequency range. This makes the design specification of the signal conditioning circuit tougher since the filter order has to be higher.

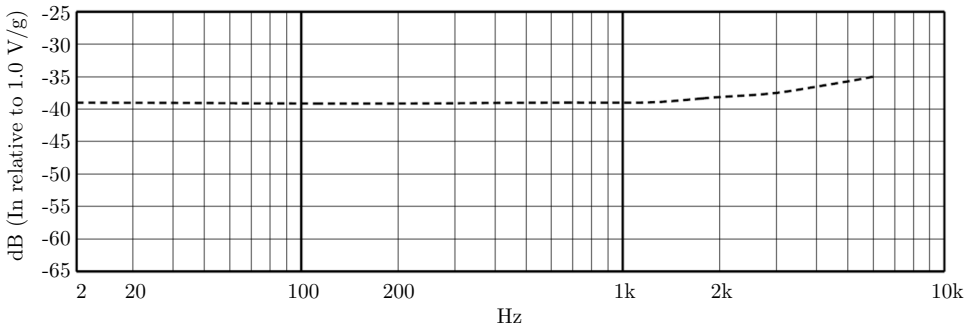

Figure 2-17: Typical frequency response of a piezoelectric microphone [188]

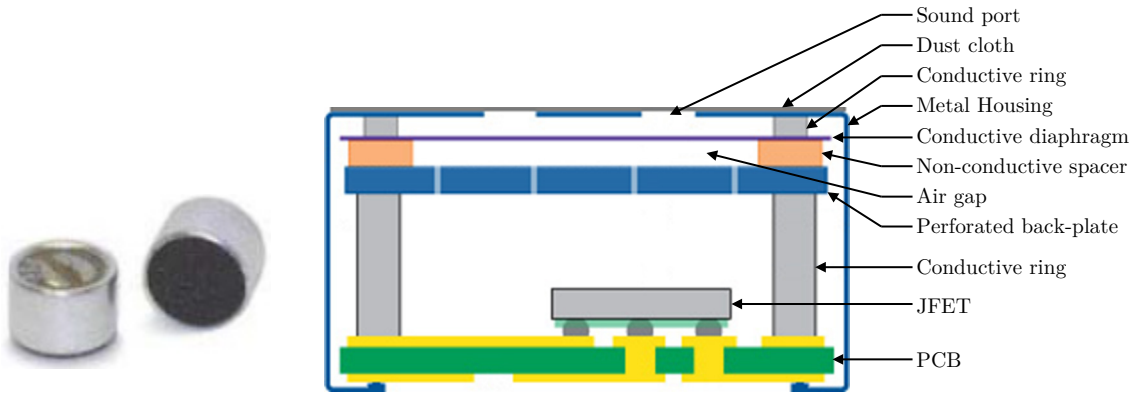

Figure 2-18: (a) Photo of electret condenser microphone viewed from two different angles; (b) Its cross-sectional structure

2.3.1.2 Electret Condenser Microphone

Instead of detecting mechanical vibration, a condenser microphone uses one plate of the capacitor as a diaphragm to sense air vibrations. These produce changes in the distance between the microphone plates. By mainly sensing the air vibration electret condenser microphones are able to better reject mechanical vibrations at the cost of losing some breath sound energy due to the extra transmission media getting involved. In earlier days, most condenser microphones required high DC voltages, of even hundreds of volts, to polarise the diaphragm in order to generate the corresponding output voltage due to changes in the capacitance. But later, with the invention of the thin electret (i.e. permanent electric polarization) film, the transducer itself no longer needs power supply [189]. For most small size electret condenser microphone, this thin electret layer of material can be adhered to one of the main building blocks of the condenser microphone - either to a flexible non-conductive diaphragm or a perforated metallic back-plate, while keeping a very small separation between these two in order to form a variable capacitor, as depicted in . Sound pressure wave causes deformation of the diaphragm, changing the distance between the electret layer and its counterpart. As the electret layer keeps a constant amount of charge, displacement due to movement will generate variations in the electret capacitance given by:

$$\Delta C = \varepsilon \frac{\text{Plate Area}}{\Delta\text{Distance}} \tag{2.3}$$

where ε is the permittivity of the material between two plates, i.e. the air. As long as the overall charges Q on the permanent electric plate keep constant, this change in capacitance induces a voltage difference between the diaphragm and the conductive plate given by:

$$\Delta V = \frac{Q}{\Delta C} \quad \Rightarrow \quad \Delta V \propto \Delta\text{Distance} \tag{2.4}$$

An electret condenser, based on this principle, is very low cost and hence is widely used. It can be seen from the previous equations how a clear advantage of this sensing method is that it does not require power. A typical usable frequency range of an electret condenser is up to 10 kHz [190]. However, using an electret condenser as a microphone also poses four disadvantages:

a) Some external mechanical vibration could still pass into the diaphragm through its circular support material due to inertial effect- though it is far less sensitive than the piezoelectric approach since there is less contact area between the vibration energy source and the diaphragm.
b) Because of the small size of the microphone diaphragm, the condenser source has a very low output capacitance, typically in the scale of 10 pF [190]. Most condenser microphones come with a transconductance buffer to reduce its output impedance. In most of cases, it requires an additional external resistor to form the common-source amplifier.
c) The interconnection between the sensitive diaphragm and the transconductance buffer relies on the metal housing of the microphone, which could affect the resilience to interference from external electrical noises if it is not properly shielded [182]. This is shown in Figure 2-19 which illustrates interference from the 50Hz mains.
d) The electret material can slowly discharge over time causing a loss of sensitivity to sound signal [190].

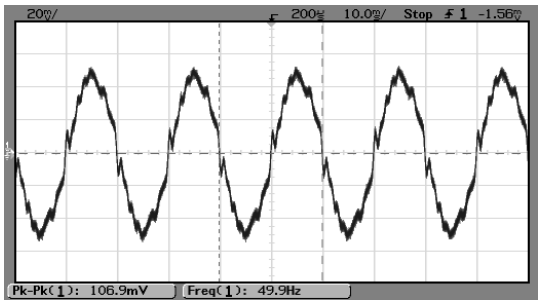

Figure 2-19: Example of 50 Hz noise received by an electret condenser microphone due to poor shielding [182]

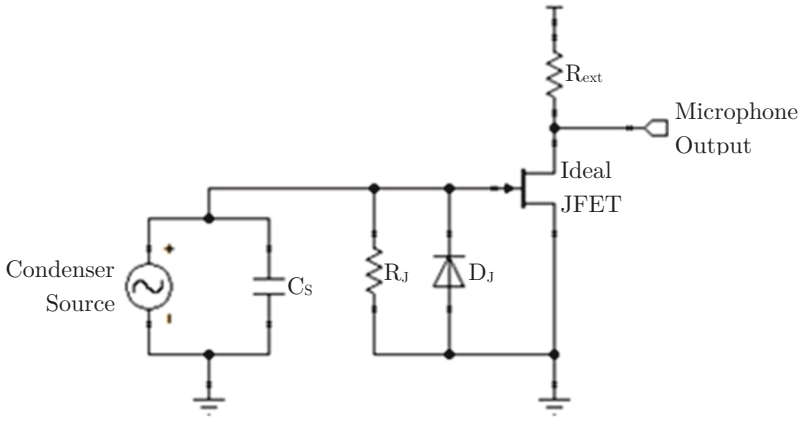

Figure 2-20: Equivalent circuit of a condenser microphone: a JFET transconductance buffer connected to an electret condenser source

Furthermore, because the output capacitance of the condenser is very small due to the small amount of electret charge; added to the input resistance of the transconductance buffer, this generates an intrinsic high pass filter which can attenuate the wanted sound signal if its cut-off frequency becomes higher than the lower frequency of interest. In order to prevent for this to happen and to decrease the intrinsic high pass frequency of the microphone condenser, a junction gate field-effect transistor (JFET) is commonly used because of its high input impedance. To buffer the signal from this kind of microphone, the most conventional method is to use a simple off-the-shelf n-channel JFET integrated circuit as a common source structure with a gain between 1 and 2. A typical JFET can be modelled as an ideal JFET with an input resistor R_J and a reverse biased diode D_J as shown in Figure 2-20. According to [191], the input impedance R_J of the JFET gate can reach several 10 MΩ and is up to few 100 MΩ. Considering the condenser source capacitance C_S is few tens of pF, this results on a high pass frequency between 100 Hz to 200 Hz, which is low enough for the sound signals of interest. Therefore, the

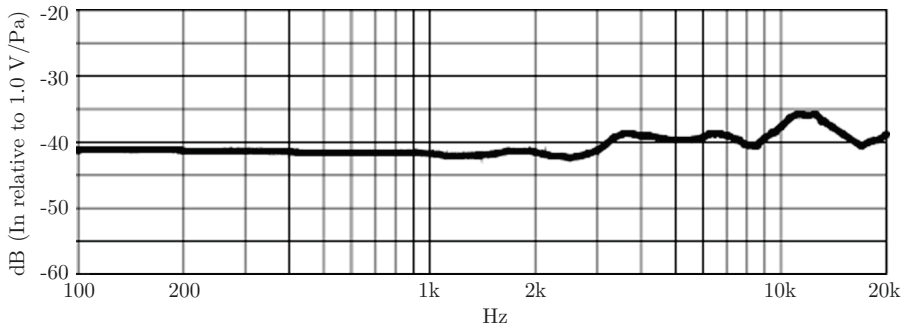

Figure 2-21: Typical frequency response of an electret condenser microphone [192]

usable frequency range is between 100 Hz to 10 kHz, as shown in Figure 2-21. For breath sounds recording purposes, electret condenser microphones have an advantage in their effective sensitivity, size, durability, and cost.

2.3.1.3 Microelectromechanical System (MEMS) Microphone

With MEMS being used ever more frequently in a variety of different applications, solving the problem of reliability over a wide range of temperatures, vibration and shock on a silicon-based substrate enables more sophisticated microphone structures. Since 1950's, researchers have been developing MEMS microphones using piezoelectric, piezoresistive, photo-optic and capacitive approaches [193]. The capacitive approach is the main stream of MEMS microphone development because of its advantages in terms of cost, noise performance and sensitivity. Thanks to silicon lithography and micromachining techniques, the capacitive microphone structure mainly consists of a moveable diaphragm and a fixed back-plate which can be small and weightless (generally less than 1 mm in diameter and 10 µm in thickness) [194]. In the microscopic environment when the mass of the pressure sensing diaphragm is comparable to the mass of the surrounding air, the air resistance become dominant opposing to the undesired mechanical vibration energy [194]. Not only that, the problematic electret layer is no longer necessary since polarising the tiny silicon-based diaphragm that can be as thin as 1 µm does not require hundreds volts of electrical potential. Depending on the sound pressure sensitivity and noise performance requirements, in some cases, this bias voltage could be provided by a small coin-cell battery [195]. In most cases, especially for off-the-shelf MEMS microphone, a bias voltage of up to 10 volts is generated by an on-chip DC-to-DC booster in order to guarantee its performance.

Besides the benefits resulting from miniaturisation, different manufacturers are able to add their proprietary features to the diaphragm holding structure such as suspension spring and free-floating in order to reduce the tensile stress at the edge of the diaphragm due to the mechanical vibration applied to the microphone [196]. A typical MEMS microphone has 10 times less mechanical sensitivity compared to an electret condenser microphone [196]. Figure 2-22 shows an MEMS microphone with two condenser units both of which have a free-floating diaphragm inside. The free-floating diaphragm is lifted by electrostatic repulsion when a bias voltage is applied through the only contact connection between it and the rest of the condenser structure. The sound port of the main condenser is opened through the bottom of the microphone package, while the other condenser is sealed inside the package which can be used to sense any residual

mechanical noise and cancel it out with the output of the main opened condenser. The application-specific integrated circuit (ASIC) next to the condensers hosts sophisticated circuits for bias voltage generation, noise filtering and signal amplification which provide a more uniform frequency response (see Figure 2-23), higher sound pressure sensitivity, higher signal-to-noise ratio and lower output impedance, compared to the electret condenser microphone. The grounded metal housing provides electrical shielding and mechanical protection for the delicate microchip and the MEMS structure.

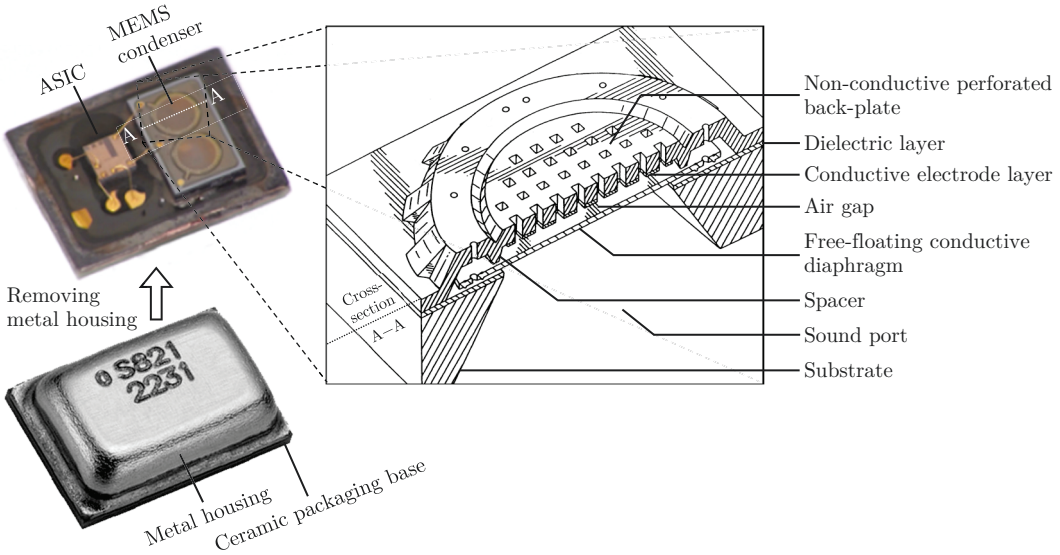

Figure 2-22: Photo of a MEMS microphone (left) and cross-sectional structure of its condenser component (right) [194]

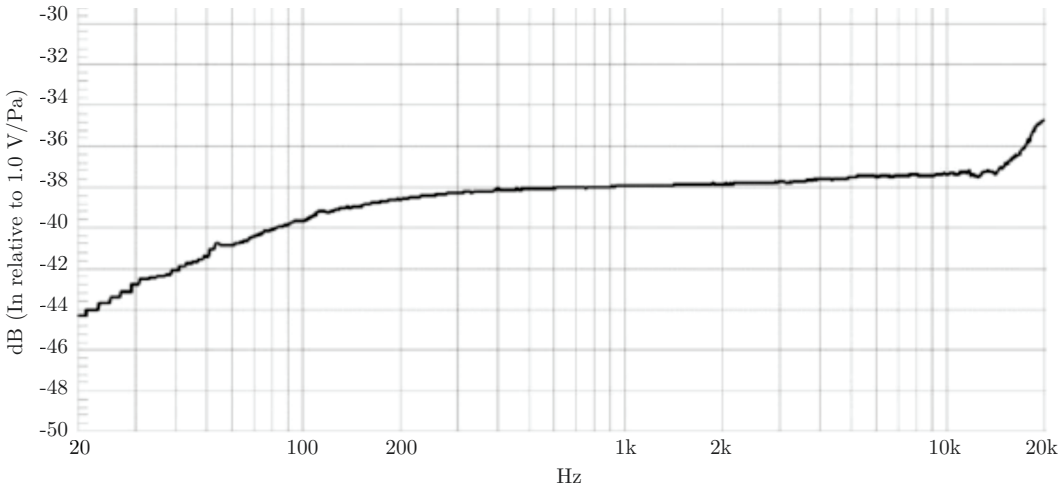

Figure 2-23: Typical frequency response of a capacitive MEMS microphone [197]

2.3.1.4 Summary

To sum up, the MEMS microphone has many advantages over the other microphones. The sophisticated micro-structure of the MEMS sensing component makes it less sensitive to mechanical noises such as body movements. The MEMS manufacturing technique similar to the microchip fabrication process ensures that its frequency response has less variation from sample to sample. Finally, the MEMS packaging provides better shielding for electromagnetic interference. Therefore, the MEMS microphone has been adopted in this work.

2.3.2 Air-coupling Acoustic Chamber

Without worrying too much about mechanical vibrations caused by body movements, this section focuses on how to secure the breath sound transmission path between the skin surface and the MEMS microphone, and ambient sound isolation. Targeting these two factors, an air-coupling acoustic chamber (as shown in Figure 2-24) is designed that also optimized transmission of the wanted bioacoustic signal from the skin surface to the MEMS microphone. The acoustic chamber has two openings. The bottom one is sealed to the skin above the source of breath sounds using a medical adhesive tape. The other end is sealed to a PCB that holds a MEMS microphone on the other side. There is a through-hole at the centre of the PCB linking the sound port of the MEMS microphone to the acoustic chamber. The size of the through-hole is designed as recommended by the microphone manufacturer, as big as the sound port plus the component placement tolerance. Hence the hole is assumed to cause little change on the frequency response of the microphone. However, when it comes to the design of the air-coupling chamber, its dimensions have to be carefully decided to avoid any change to the frequency response within the breath sound band. In [198, 199], Wodicka et al. modelled the microphone

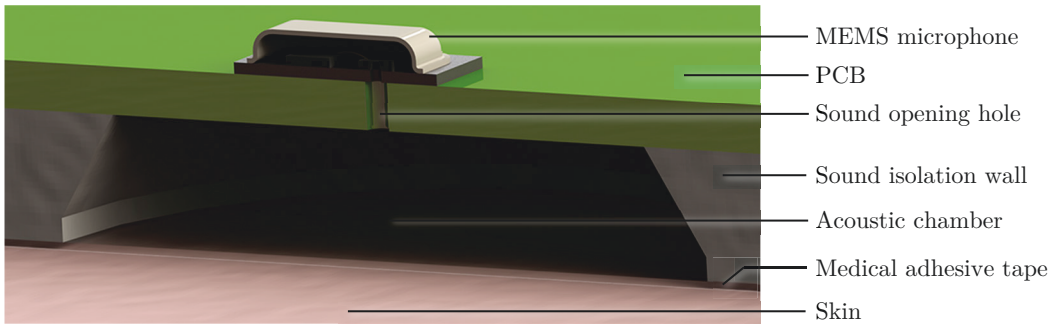

Figure 2-24: Cross-sectional view of the conical air-coupling acoustic chamber and its surrounding components placed on top of skin

air-coupler as a low pass filter. The geometry of the air-coupling chamber such as the cavity depth and bottom diameter can cause different effects on the frequency response above 500 Hz. It is suggested that a conical chamber has less attenuation than a cylindrical shape on the frequency band between 500 Hz to 1500 Hz. The depth of the air coupler should also be as small as possible, since bringing the microphone closer to its source can improve the sensitivity on the frequency band above 500 Hz by minimising the energy lost during travelling of the sound wave through the air. However, a minimum of 1 – 2 mm clearance between the PCB and the skin is required to prevent them touching each other. The overall sound energy collected by the air-coupler is proportional to the area of the bottom opening. Larger bottom opening diameters result on more sound energy concentration at the microphone end, but it also makes the outline of the sensor bigger. Because one of the main objectives of this work was to minimise the size of the sensor, the base outline of the sensor was designed to have a size matching the minimum dimension of a PCB (around 22 mm) which could hold all the necessary components for signal processing and wireless transmission plus 2 mm thickness of the sensor housing on all sizes. In this case, the maximum bottom opening diameter was 15 mm, leaving at least 5.5 mm clearance to the outline of the sensor housing for skin attachment purposes. This clearance also provided the space for building a sound isolation wall between the acoustic chamber and the outside world. This eliminated another problem of the acoustic chamber, as presented in [199, 200]: the fact that the breath sound signals recorded can easily become corrupted by air-borne sounds generated from the ambient environment. To maximise the sound isolation, the acoustic chamber had to be completely sealed, both in the PCB and skin sides. In order to attach the sensor to the skin and seal the acoustic chamber, a 3M medical double-sided polyethylene adhesive tape was used. This adhesive was chosen because it proved it could attach firmly both to the plastic sensor housing and the human skin for over 24 hours without causing skin anaphylaxis.

2.3.3 Design of Analogue Frontend

Since there is an ASIC inside each MEMS microphone, the generated breath sound signal can be analogue or digital (including pulse-density modulated analogue) depending on what kind of feature is added into the chip. To date, off-the-shelf digital MEMS microphones are not very power-efficient and consume 500 µA to 1.5 mA more current than analogue ones because manufacturers target mobile phone and tablet markets which have different power budget. Since

one of the goals of this work was to minimise the power consumption, it was reasonable to choose an analogue MEMS microphone, working together with an external analogue frontend that gives better performance to power consumption ratio. This also provided the additional freedom to partly process the sound signal in the analogue domain to ease the requirements of the conversion and/or transmission later on.

The output of MEMS microphones may have different dynamic range than the input range of the chosen analogue-to-digital converter and also a wider range of frequency content compared to the breath sound band of interest. An analogue frontend is required to amplify and bandpass filter the microphone output. As an example, implemented in [201], Figure 2-25 shows an analogue frontend starting with a 2nd order cascade high pass filter which eliminates out-of-band low frequency components generally caused by mechanical movements and heart sounds; followed by an amplifier which conditions the normally very weak signals for the subsequent blocks; and a low pass anti-aliasing filter to reject any image caused by frequency components higher than the Nyquist frequency (i.e. half of the sampling frequency) at the analogue-to-digital conversion (ADC) stage. This can be also implemented using a multiple-feedback filter topology as shown in Figure 2-26. The benefit of the second approach is that all the passive components are not connected to the power supply rail, resulting on a higher power supply rejection ratio.

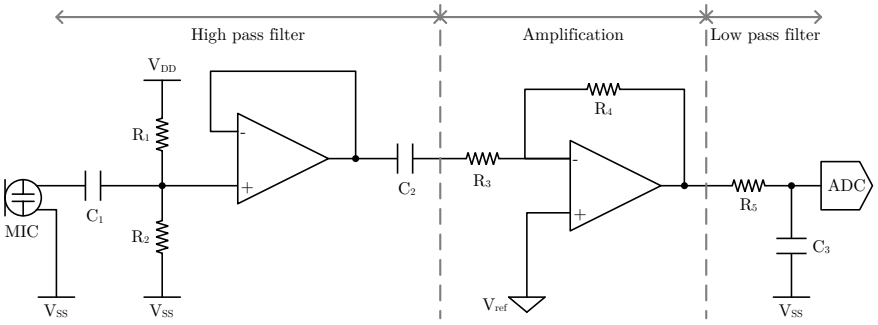

Figure 2-25: Circuits of analogue frontend for breath sound recording proposed in [201]

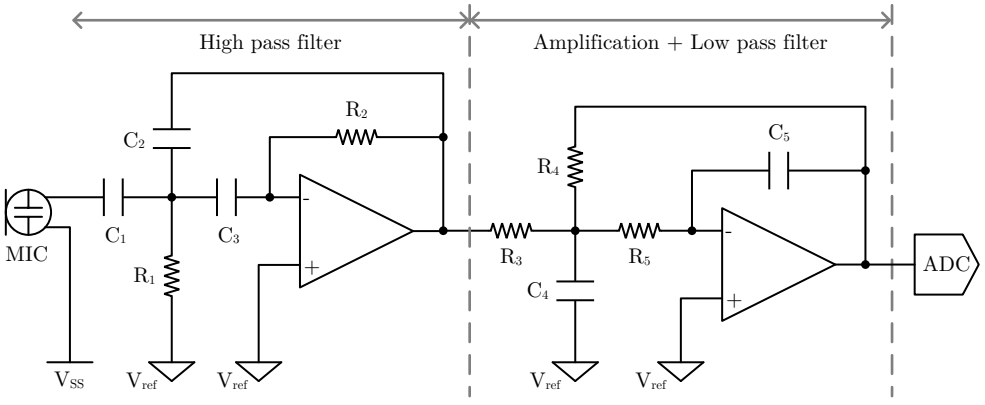

Figure 2-26: Improved circuits of analogue frontend for breath sound recording with better performance on power supply noise rejection.

Chapter 3

Truly Wearable Wireless Biomedical Monitoring System Design

3.1 Introduction

Wearable biomedical monitoring devices are quickly emerging as a new form of platform for ubiquitous philological monitoring of the human body. As illustrated in Figure 3-1, a wearable biomedical monitoring system includes at least one highly miniaturized sensor node which attach to the body, records one (or potentially more) physiological parameters and may perform online signal processing before wirelessly transmitting the recorded signals to a base station which can be either a dedicated receiving device or an existing off-the-shelf consumer product such as a smartphone or PC.

For end users there are a number of features that successful wearable sensors must include: it is essential that they are easy-to-use, socially acceptable, and long lasting. The power consumption of the sensor node is a critical factor in realising all of these features as the current drawn by the sensor sets the physical size of the battery required, which determines the device size and operating lifetime, which in turn affects the ease of use. Unfortunately, for creating simple pick

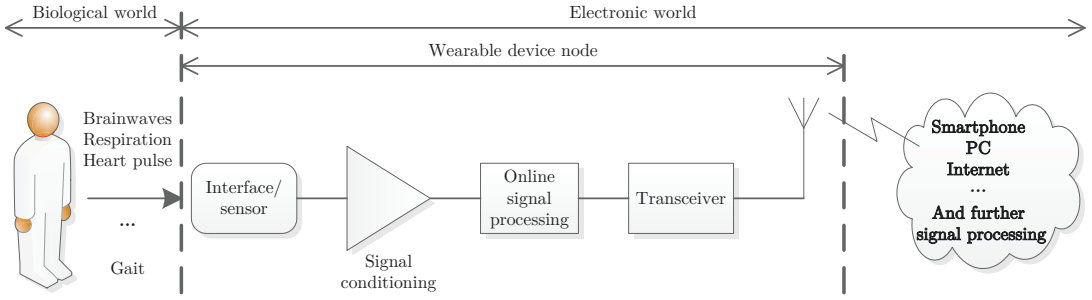

Figure 3-1: Wearable device nodes connecting the biological world to the electronic world [202]

up and use biomedical monitoring devices, the power level still falls far short according to the state-of-the-art presented in Chapter 1. Substantial improvements in systems power consumption are required to realise units that can be reliably re-used session after session without having to worry about the device stopping to work due to the battery. This is a great source of frustration for users, and limited battery life is the major obstacle to the widespread deployment of wearable sensor systems today. Strategies for maximising the operational lifetime in next generation devices are therefore essential. This relies on the close fusion of the application requirements, the sensor node design, the signal processing design, and the electronic design in order to realise the lowest possible levels of power consumption and to maximise the battery lifetime.

This chapter starts with an introduction and trade-offs analysis of the critical factors that affect the design of miniature wearable devices, specifically their volume and operating lifetime regarding to the selection of batteries- This has been published by the author in [202, 203]. Although this analysis is generic and hence applicable to a variety of wearable systems, it is subsequently explained how it was used in the design of the respiratory monitoring device created in this thesis. The detailed power performance of a current sensor node based upon Texas Instrument's popular MSP430 microcontroller is analysed. This study is useful to demonstrate the design trade-offs present, and the benefits of on-sensor node signal processing in terms of reducing data rate, maximising operational lifetime, and in increasing the range of battery technologies that are suitable for use with a particular senor.

3.2 Power Supply

3.2.1 Battery Selection

The most important key to the success of a wearable medical device is the minimization of its size, which directly affects its discreteness and comfortability. As surface mount components are nowadays small compared to batteries, leaving aside the application specific sensor, the size of batteries will dominate the overall volume of the device. But besides the size, there are four additional battery specification parameters that need to be considered in low power wireless design: supply voltage, maximum continuous current, maximum pulse current capability and effective capacity.

Group	Name	Chemical Composition & Energy Density	Nominal Voltage	Nominal & Max. Cont. Current	Nominal Capacity	Weight	Size (diameter x height)
Cylindrical Cell	Duracell DA675 (Size 675) [204]	$Zn(OH)_4$ 1673 Wh/L	1.4 V	2mA 16 mA	600 mAh	1.8 g	11.6 mm x 5.4 mm
Button Cell	Duracell DA13 (Size 13) [205]	$Zn(OH)_4$ 1673 Wh/L	1.4 V	0.9mA 6 mA	290 mAh	0.8 g	7.9 mm x 5.4 mm
Cylindrical Cell	Xeno XL-050F (½ AA) [206]	$LiSOCl_2$ 1200 Wh/L	3.5 V	1mA 50 mA	1200 mAh	9.2 g	14.5 mm x 25.2 mm
Coin Cell	Renata CR2430 [207]	$LiMnO_2$ 580 Wh/L	3 V	0.5mA 4 mA	285 mAh	4.1 g	24.5 mm x 3.0 mm

Table 3-1: Specifications of four non-rechargeable disposable batteries potentially suitable for wearable sensor nodes with varying physical sizes and battery technologies [202]

To guide the investigation of battery selection, Table 3-1 summarises the specifications of four off-the-shelf primary batteries from three different size groups, cylindrical cell, button cell and coin cell, which are potentially suitable for powering wearable sensor nodes. Rechargeable Lithium batteries are not listed because commercial available forms with similar size have lower energy densities (typical 200 Wh/L) than the chemistries listed here. For higher energy density (typical 400 Wh/L) with the rechargeable options, the minimum physical sizes available are generally bigger, similar to mobile phone batteries. The table shows five battery specifications that are critical to consider for low power wireless design. The physical size which, as discussed above dominates the device volume, and the energy storage capacity, typically expressed in mA-hours, are the well-known parameters. However, all batteries also have an internal resistance and this means that the energy stored cannot be optimally discharged into all possible loads. This leads to further important battery parameters which are discussed below.

3.2.2 Supply Voltage

The supply battery voltage must meet the operation requirements of the electronic circuitry used in the wearable device. Importantly, the supply voltage is not the same as the battery nominal voltage as there will be an internal voltage drop in the battery due to the internal resistance, and this drop will vary depending on the current draw. Most off-the-shelf low power microcontrollers and transceivers today require somewhere between 1.7 V and 3.6 V and batteries which supply a voltage outside this range must be used with a DC-to-DC voltage converter or stacked in series to increase the delivered voltage. However, these techniques

reduce the battery lifetime due to either the extra power consumption from the additional circuitry or due to the increased internal impedance.

3.2.3 Maximum Continuous Current

The maximum continuous current (also referred to as the maximum average current supply), $I_{avg(max)}$, limits the average current draw from the battery. In theory a 200 mAh battery can provide 1 mA for 200 hours, or 200 mA for 1 hour. However each battery actually has a maximum supported current draw and if more current than this is drawn the effective capacity will not be the full 200 mAh reported value. The average current required by the system must be smaller than the corresponding $I_{avg(max)}$ for the specific battery. Otherwise more than one battery will have to be used in parallel, possibly together with diodes to prevent any non-rechargeable batteries from inadvertently charging. Again, the battery must be able to provide this current without a significant drop in its voltage supply because of the internal impedance.

3.2.4 Maximum Pulse Current Capability

F short periods a battery can provide more than $I_{avg(max)}$, up to a maximum pulse value $I_{pulse(max)}$. $I_{pulse(max)}$ must be large enough to guarantee that variations in the supply voltage, which will drop if the current consumption increases, will not exceed the operating range of any circuit in the wearable sensor. The value of $I_{pulse(max)}$ is a function of how long the pulse must be provided (i.e. the hold time) and Figure 3-2 shows typical maximum values of pulse current with various pulse widths for the batteries listed in Table 3-1. There is a clear decrease in the maximum pulse current that can be provided as the duration of the required pulse increases. This is of particular importance when selecting the radio transceiver block and protocol as this generally determines both the peak current draw and how long it is required for.

3.2.5 Effective capacity and lifetime

Combining the above effects allows the effective battery capacity (C_{eff}) to be found. This is a critical parameter for long term monitoring applications and it will generally be smaller than the nominal battery capacity, unless the system average current consumption (I_{sys}) matches the manufacturer's recommended value (also referred to as the nominal current or the standard

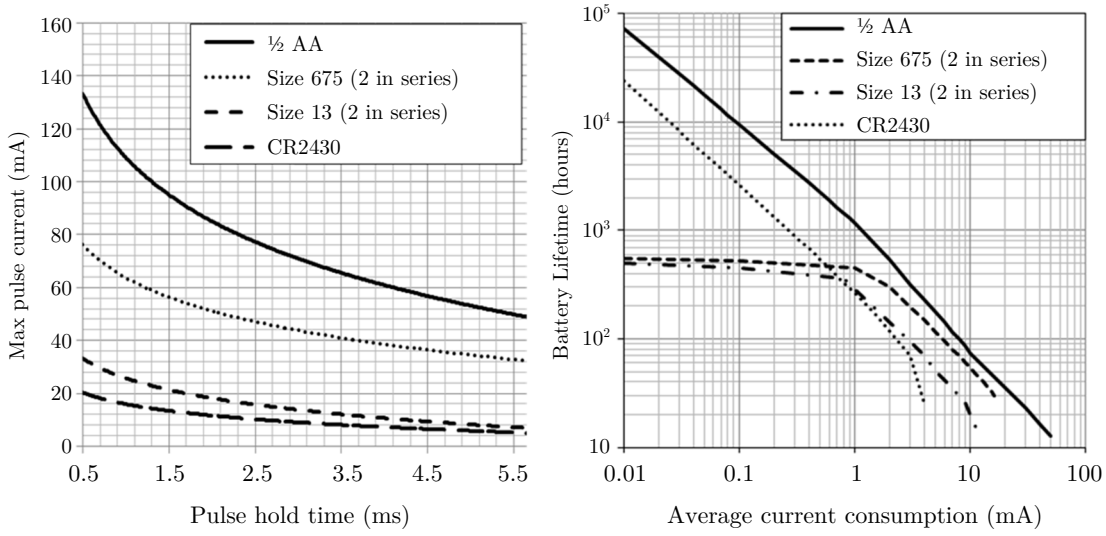

Figure 3-2: Measured discharge characteristics of the small size batteries from Table 3-1. Zinc-Air batteries are tested in a combination of two in series with no air management. 2.1 V is considered as the discharge voltage. Left: Battery maximum supported pulse current vs pulse hold time. Right: Battery lifetime vs average current consumption.

discharge current). The value of C_{eff} is hence a function of I_{sys} and it will be lower than its nominal value when I_{sys} is larger than the nominal current.

Based on this, in order for a certain battery to be a viable option for a specific system it must meet the system requirements not only in terms of size and supply voltage but also in terms of the average and maximum dynamic currents that the system needs to operate. Once the necessary conditions are met the lifetime of the battery, $LT(I_{\text{sys}})$, can be determined as:

$$LT(I_{\text{sys}}) = \frac{f_{C_{\text{eff}}}(I_{\text{sys}})}{I_{\text{sys}}} \tag{3.1}$$

Figure 3-2 also shows a plot of the measured battery lifetime $LT(I_{\text{sys}})$ versus the average current consumption I_{sys} for the examples of batteries presented. From this it is clear how, although the cylindrical cell battery used requires more volume it also provides a much better performance in all other aspects. Comparatively the Zinc-Air button cell batteries have similar performance only if the effective average current drain is higher than 1 mA. This is due to their high self-discharge characteristic when exposed to air, and for applications with ultra low effective power consumption (i.e. $I_{\text{sys}} < 1\text{mA}$). Zinc-Air batteries should be used with air management so that only a limited (but necessary) amount of air goes into the batteries. Alternatively, in cases when

the continuous current is not a limiting factor, the coin cell battery may be preferable since it has a superior combined performance in terms of self-discharge rate and pulse current capability.

3.3 Hardware Platform

The hardware platform in a wearable sensor node is responsible for collecting the physiological data (data acquisition) and for packaging and passing this data to the wireless transmitter. It must also control the transmitter operation (wireless transmission control) and there are a number of design factors that need to be studied.

3.3.1 Data Acquisition

As shown in Figure 3-1, a general wearable sensor may contain signal conditioning circuitry (typically an amplifier and an Analog-to-Digital Converter (ADC)), and some optional signal processing blocks to reduce the amount of data for wireless transmission. In the signal conditioning block there is at least one anti-aliasing filter which restricts the bandwidth of the input signal before passing it to the ADC. For wearable applications seeking to minimize size, the anti-aliasing filter may be realized as a 1^{st} order passive RC lowpass filter. This is a physically very small circuit, but its use will come at the cost of requiring a higher ADC sampling frequency to ensure that the analogue input signal is correctly represented in the digital domain. In turn this implies that the amount of data passed through the system is increased, leading to an increase in power consumption due to the data multiplication. If wanted, and as considered in section 3.5.2, downsampling of the data can be performed in the digital domain to obtain data compression. It is to avoid similar data multiplication that Successive Approximation Register (SAR) ADCs tend to be preferred in wearable applications, rather than Σ-Δ ones which are intrinsically based upon oversampling. In addition, SAR converters can operate with a lower peak dynamic current.

3.3.2 Wireless Transmission Controller

After acquisition, the physiological data is temporarily stored in a buffer and packetized to be sent to the wireless transmitter. During this process the controller decides on the structure of the transmission protocol, including the total size of a packet frame (L_{frame}) and the size of the packet header. Both limit the amount of physiological data (L_{data}) that can be carried in each

over-the-air data packet and mean that the full over-the-air rate available to the wireless transmitter cannot be used to transmit useful physiological data. The utilisation of the protocol (η_{protocol}) for the data can be expressed as:

$$\eta_{\text{protocol}} = \frac{L_{\text{data}}}{L_{\text{frame}} + T_{\text{latency}} \cdot R_{\text{air}}} \tag{3.2}$$

where R_{air} is the over-the-air data rate and T_{latency} is the total latency introduced by the data acquisition, packet packaging and transceiver interfacing.

Note that although longer packet sizes increase the data throughput and relieve the communication overhead they also increase the transmission time, exponentially increasing the risk of interruption from undesired Radio Frequency (RF) interferences. This error rate will be explored in more detail below. Further, to avoid extra transceiver control complexity and overhead, the length of each data packet could be designed to match the size of the buffer present in the transceiver being used. Using this optimal packet size also prevents the buffer from overflowing.

3.4 Wireless Transmitter

Finally the packetized physiological data is passed to the wireless transmitter which sends it out to the sensor node base station. There are three factors that dominate the power consumption at this stage: the quality of the packet transmission, the hardware overhead and the over-the-air data rate.

3.4.1 Quality of the Packet Transmission

Unstable transmission quality can cause unnecessary retransmission overhead or even packet loss. The successful transmission rate (η_{tx}) for one packet can be calculated as:

$$\eta_{\text{tx}} = (1 - \text{PER}) \cdot \varphi_{\text{QOL}} = (1 - \text{BER})^N \cdot \varphi_{\text{QOL}} \tag{3.3}$$

where PER is the Packet Error Rate; BER is the Bit Error Rate; N is the bit length of a packet; and φ_{QOL}, the Quality of Link (QOL) factor, which estimates the probability of the RF channel being clear throughout the transmission process. It decays with longer transmission times.

From equation (3.3) it can be seen how a long packet size is not advisable. Also, the QOL of radio transceivers is affected by the transmission range and the transmission power. Long transmission ranges decrease the strength of the RF signal and hence lower the QOL. Generally, the transmission RF power can be tuned to trade-off QOL and power consumption.

In addition to this, to compensate for the BER and improve the link quality chip manufacturers have introduced different features into transceivers so as to strengthen their error tolerance capability. Forward Error Correction (FEC) and automatic acknowledgement (Auto-Ack) are two examples which will be used in the test cases considered below. However the use these can come with extra cost, the largest one being FEC which sacrifices half of the available bandwidth to create redundancy.

3.4.2 Hardware Overhead

In order to allow the wireless transmitter and receiver to recognize each other, and to synchronize, the transceiver appends a preamble signal (with size L_{preample}) and its own transmission identification data (with size L_{txid}) to the front of every packet automatically. Once the packet reaches the receiver these extra bits are examined and removed by the packet handling hardware. Again this results in not all of the over-the-air data rate being available for useful data transmission. The hardware efficiency η_{hw} is a key factor in this and can be calculated as:

$$\eta_{\text{hw}} = \frac{L_{\text{data}}}{L_{\text{frame}} + L_{\text{preamble}} + L_{\text{txid}} + (T_{\text{cal}} + T_{\text{switch}}) \cdot R_{\text{air}}} \tag{3.4}$$

where the last term in the denominator accounts for a certain overhead generated by the time required to calibrate the phase lock loop (PLL) module of the RF synthesizer before transmission: T_{cal} is necessary in battery powered systems in order to avoid the frequency drift caused by variations in the supply voltage; and T_{switch} is the time taken by the transitions between different low power states.

3.4.3 Air Data Rate vs Effective Data Rate

Given all of these factors the effective data transmission rate can be calculated, and this must be sufficient for the wanted application. Moreover, for the same amount of data, the faster the

data rate the lower the effective power consumption as the transmitter can be turned off for more of the time. However, the overheads discussed in previous sections decrease the effective data rate and must not be ignored in power management. The effective bandwidth (R_{eff}) of the transmission process can be estimated as:

$$R_{\text{eff}} = \eta_{\text{protocol}} \cdot \eta_{\text{tx}} \cdot \eta_{\text{hw}} \cdot R_{\text{air}} \tag{3.5}$$

3.5 Design of Wearable Respiratory Sensor

Taking the design constraints from the previous Sections 3.2, 3.3 and 3.4 and applying them to the design of a real respiratory monitoring sensor platform, the following sections demonstrate how they impact the node design and performance. Furthermore, an example of on-sensor node signal processing shows the practical impact for improving the operating lifetime. The architecture of the wearable sensor is shown in Figure 3-3, and is based upon the popular MSP430 microcontroller to be representative of many current sensor nodes.

3.5.1 Node design

The wearable node has a single channel which starts at the high pass filter, normally used to eliminate out-of-band low frequency components caused by wearer's mechanical movements. This is followed by an amplification block which conditions the normally very weak signals for the subsequent blocks, a low pass anti-aliasing filter and a 10 bit SAR ADC which is part of the MSP430 chip. Before the ADC block, the components in the analogue frontend (detailed design can be found in Section 2.3.3) consume much less power compared to the rest in the system. Unless all the analogue circuits were implemented on a single application specific microchip, which is usually not a choice due to the high cost of fabrication process, a combination of single general purpose microchip containing multiple operational amplifiers and some passive components such as surface mounted resistors and capacitors is used in most of system design cases. Hence, for the purpose of minimising PCB size of the sensor node, it is not advisable to adopt high order filters at the cost of using more surface mounted components.

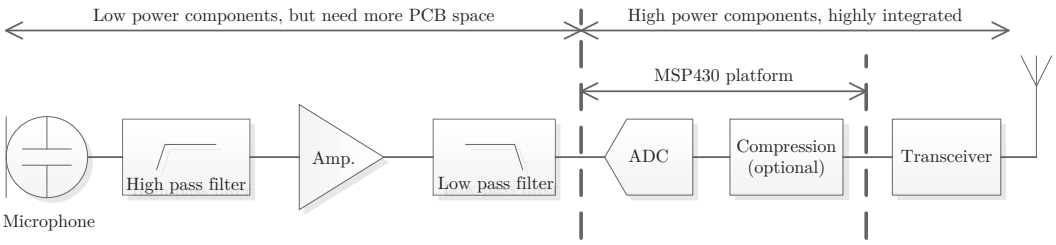

Figure 3-3: MSP430 based wireless acoustic respiratory sensor node presented here. The platform can be contrasted with the general node architecture presented in Figure 3-1

After the ADC, the analogue sound signals are converted to the digital form, which are processed by the MSP430 chip into packets and ready for wireless transmission. Here two variants on this basic system are investigated: one using a Texas Instruments CC2500 transceiver as the transmitter stage; and one using a Nordic nRF24L01+ transceiver. Both of these operate in the 2.4 GHz band and are connected to the MSP430 by the Serial Peripheral Interface (SPI) with their supported maximum 10 MHz and 8 MHz clocks respectively. The CC2500 has an option to enable FEC, while the nRF24L01+ provides hardware support for automatic acknowledgement (Auto-Ack).

Texas Instrument's SimpliciTI protocol stack is used with both transceivers, although it was originally designed for the CC2500 only. SimpliciTI is a low power light-weight wireless network protocol dedicated for battery operated devices that require long battery life. From [208] the header of a non-encrypted SimpliciTI packet frame contains 96 bits. Because the nRF24L01+ has half the buffer size of the CC2500 (256 bits), one field in the header is used to store the byte length of the user data and this has been modified for the nRF24L01+. In the user data field, a time stamp is added for the recording, 20 bits in length, to each packet as an additional header. Given the memory constraints of the used MSP430 the rest of space in a packet allows the CC2500 to transmit 34 samples (340 bits), whereas the nRF24L01+ can transmit 18 samples (180 bits). In total, the packet size is 456 bits for the CC2500 and 294 bits for the nRF24L01+.

To maximize the data throughput, minimize latency, and allow the transmitters to be duty cycled to reduce the average power, the air data rates of the transceivers are set to their maximum values: 500 kbits/s for the CC2500 and 2 Mbits/s for the nRF24L01+. As recommended in [209], the preamble signal of the CC2500 is set to 96 bits. The nRF24L01+ has 8 bits fixed preamble length [210] and both the CC2500 and nRF24L01+ have extra identifica-

tion data appended by hardware with lengths of 16 and 19 bits. Overall, assuming 100% of QOL, the effective data rate can be calculated from equations (3.2) – (3.5) as 106 kbits/s and 338 kbits/s respectively.

3.5.2 Optional data compression

As discussed in Section 2.2, to cover the physiological range a bandwidth of at least 1 kHz is required. However, due to the use of a 1st order passive RC anti-aliasing filter, the sampling frequency is pushed to 4 times of the filter's cut-off frequency, which is at 4 kHz. The oversampling technique can reduce the effect of aliasing and improve the signal-to-noise ratio (SNR). Given this oversampling we again investigate two variants on the system: one where we transmit all of the collected data, and one were we first downsample the already digitised data by a factor of 2 with a higher order digital anti-aliasing. Effectively, this sacrifices a little bit more power for digital signal processing, in return of not only saving PCB area by keeping the simplicity of the analogue section, but also having less data to transmit. This also explores the impact of even modest data compression on our node lifetime.

The downsampling is implemented with a 10th order Kaiser window digital FIR low pass filter running on the MSP430. In this experiment, hardware acceleration unit such as hardware multiplier which can reduce the computational burden for this type of processing is not used in order to make the case more generic, as most of the low power microcontroller has very limited peripheral resources.

Further choices of data compression are discussed in Chapter 4.

3.5.3 Power performance results

For the core system, excluding the wireless transmitter, Table 3-2 shows the current consumption of each block present. The typical current profiles of the two transceivers are illustrated in Figure 3-4, where label '**A**' marks the section corresponding to the transmission overhead, '**B**' marks the actual wireless transmission times and '**C**' indicates the optional Auto-Ack reception.

Blocks Name		Peak Current (mA)	Duty Cycle	Effective Current (mA)
Sensor		0.27	100%	0.27
Amplifiers		0.36	100%	0.36
10-bit SAR ADC		1.10	19.9% (45 s per sample)	0.22
Compression by downsampling (optional)	*10 MHz (CC2500)*	2.50	57.5% (130 s per sample)	1.44
	8 MHz (nRF24L01+)	2.10	84.7% (163 s per sample)	1.51
Total	*10 MHz (CC2500)*	4.23	–	2.29
	8 MHz (nRF24L01+)	3.83		2.36

Table 3-2: Measured current consumption of the hardware platform and its constituent parts, excluding the wireless transmitter. Note that the CC2500 system has a 10 MHz clock while the nRF24L01+ system uses an 8 MHz clock.

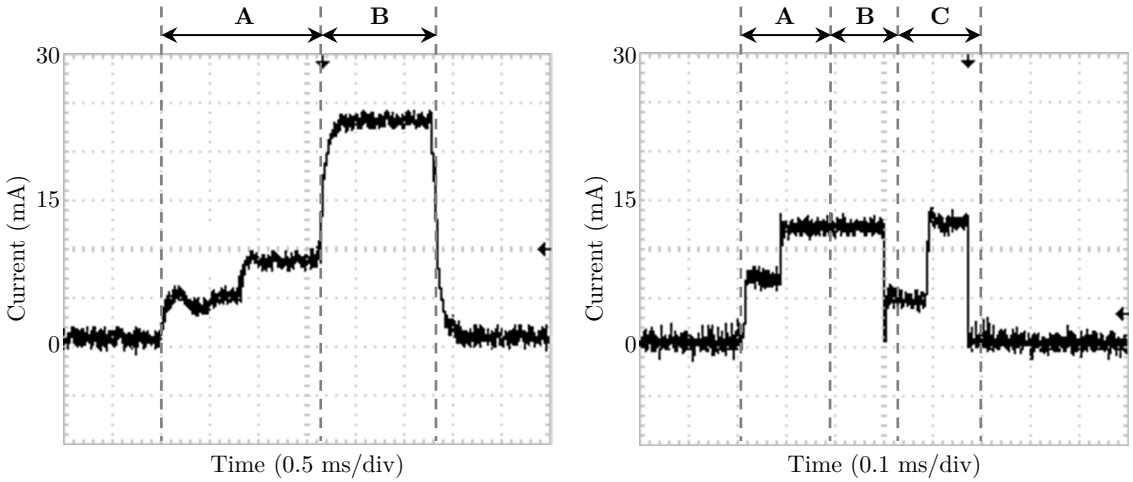

Figure 3-4: Measured current consumption profiles of the two transceivers with Vdd = 2.4 V. Left: CC2500, with no FEC, packet size with overhead is 568 bits, 34 samples. Right: nRF24L01+, with Auto-Ack, packet size with overhead is 321 bits, 18 samples. Section **A** is the overall transmission overhead, **B** is the actual wireless transmission time and **C** the optional Auto-Ack reception.

Combined with the wireless transmitter, the two systems are compared in Table 3-3 which shows the current consumption of the systems when tested over 24 hours using the different batteries given in Table 2. It can be seen how some configurations only pass the 24 hour test with certain batteries. Not all miniature batteries are suitable for long term monitoring! This has direct implications on the kind (and number) of batteries that may be required in a specific design, and consequently on the device size. For 24 hour monitoring applications with 22 kbits/s data rate, the smallest wireless configuration can be achieved by using one lithium manganese dioxide CR2430 3V battery.

Turning on the data compression leads to substantial reductions in the effective current drawn by the system. Reductions by 20–30% are achieved, and in the CC2500 case without FEC the reduction is 53%. In all of the cases this reduction allows more of the battery technologies from Table 3-1 to be used, giving greater freedom in the system design and in the optimization of the device size. In the best case the effective current consumption of the entire system goes down to 3.5 mA, equivalent to a net power consumption of 334 nW/bit. It is important to highlight however that only reductions in the effective current are achieved by the data compression. There are no substantial differences in the peak currents and in some cases this may now become the limiting factor in the system design.

Platform	**Configuration**	**Effective Current (mA)**	**Peak Current (mA)**	**Passed Batteries**
MSP430 & CC2500	Compression off, FEC off	15.8	24.3	XL-050F
	Compression off, FEC on	18.5	24.3	XL-050F
	Compression on, FEC off	7.4	25.6	XL-050F, DA675
	Compression on, FEC on	13.5	25.6	XL-050F, DA675
MSP430 & nRF24L01+	Compression off, Auto-Ack off	4.5	14.3	XL-050F, DA675, DA13
	Compression off, Auto-Ack on	5.6	15.2	XL-050F, DA675, DA13
	Compression on, Auto-Ack off	3.5	15.1	XL-050F, DA675, DA13, CR2430
	Compression on, Auto-Ack on	3.8	16.1	XL-050F, DA675, DA13

Table 3-3: Current consumption of the sensor nodes with compression enabled and disabled. Only some of the miniature batteries listed in Table 3-1 are capable of powering the system for 24 hours.

3.6 Summary

Designing a wearable sensor node involves a careful set of trade-offs between the electronic components used, the battery technology selected and the implementation of any real-time signal processing. This Section has demonstrated that in battery selection, which dominates the end physical size of the device, both the average and peak current draws have to be taken in to consideration. The presented numbers can be used as a realistic guide for system designers when distributing their power budget and when estimating the size and kind of battery required to operate their device for a certain length of time. Further, quantitative measured results have shown for two different transmitters the potential benefits of simply onboard downsampling data by a factor of two reduce up to 53% in the total system power consumption. Moreover, this was done with the MSP430 active for up to 72% of the time. Clearly if more data reduction could be provided, or if the signal processing platform could be turned off more of the time to reduce its effective power, even greater increases in operational lifetimes could be provided.

The challenge, of course, is in realising accurate data reduction algorithms that can operate within the limited power budgets available. Further discussion on potential suitable data reduction algorithm on breath sound is given in Chapter 4.

Chapter 4

Data Reduction in Respiratory Sounds

Although the breathing monitoring prototype created as part of this PhD does not require any data compression of the sensed signals because the power provided by Zin-Air batteries is enough to operate the chosen transmitter, this comes at the cost of being restricted in the choice of receivers.. A more flexible system can be created if the existing implementation could eventually be ported to the now widely accepted ultra-low power protocol Bluetooth® Smart (formally as known as Bluetooth® Low Energy or BLE). In order to make this possible a significant challenge needs to be overcome. This relates to the actual data rate, which in practice is much lower than the maximum advertised 1 Mbit/s and is very much dependent on the implementation of the Bluetooth host controller because this will limit the number of communication events in a period of time. The reported maximum effective data rate varies from 15.8 kbit/s for the case of Apple iOS device to 57.2 kbit/s for Android device [211]. Unfortunately, the worst case is smaller than the required data rate for breath sound data which is 22 kbit/s. This does not even include the overhead for the packet identification (for detail see Section 3.4.2). Therefore, without data compression, it is impossible to acquire the breath sound data using Bluetooth® Smart in some cases, for example when any Apple product is used as a receiver.

Data compression has been widely used in portable devices to reduce the amount of digital data that needs to be stored or transmitted in order to extend the battery lifetime. Audio compression is possible because adjacent audio signal samples tend to present redundancies [212]. So far, data compression has been briefly mentioned in Section 3.5, where downsampling was used in the example respiratory sensor prototype. This chapter discusses other possible compression approaches that have been investigated throughout the course of this thesis and can be used in breath sound acquisition systems.

4.1 Overview of respiratory sound compression – How much can we benefit from that?

Sound compression techniques can be split into two groups: lossy and lossless approaches. Both approaches have been well developed for all kinds of sound data, especially music. However, research on respiratory sound specific data compression has received little attention.

4.1.1 Lossy Respiratory Sound Compression

Yadollahi, et al. [213] have developed a lossy compression method on respiratory sound with an objective to improve the level of SNR for the use of telemedicine applications. The method used an adaptive transform coding-based technique to quantise each respiratory sound sample into a shorter length representation. They showed that the adaptive quantisation technique provided up to 11dB better SNR than downsampling at the same of compression rate. The optimal analysis window size is 12.5 ms for this approach [213], which is equivalent to 28 samples with a sampling frequency of 2205 Hz. However, for every 1 second of data it required 19.8% of an Intel® dual cores CPU time, on average. Therefore, the approach is still too complicated resulting in considerable memory usage, as well as computational power requirements for complex matrix operations, which it is difficult to implement on a low power wireless sensor platform for the purpose of reducing transmission power. Moreover the quantisation noises introduced into the original respiratory sound signal affect the following data analysis, creating uncertainty.

To avoid the uncertainty created by the lossy compression during the data analysis stage, there is another route to reduce the size of the respiratory sound data. Most of the data analysis approaches can be divided into three main steps: signal pre-processing, feature extraction and feature classification. The remaining information derived from the first two steps is generally less than the original input signals. If that is the case and the power consumption required by those steps is smaller than the power saving benefited from the reduction on data transmission, the compression can be achieved by just keeping the derived information. As an example, a part of obstructive apnoeas and hypopnoeas screening algorithm[1] [214] has been used to compress respiratory sounds. Similar to downsampling, the compression technique keeps the maxima of

[1] For further detail of the algorithm, please see Section 5.1.1.

sound signals within a fixed length moving time window at the rate of 1Hz [214]. This results in a significant compression rate on respiratory sound data without risking degradation of the detection performance. However, it is impossible to recover the sounds from these compressed 1 Hz data for the purpose of auscultation. During this PhD, initial work on compression also focused on exploring lossy compression techniques which can preserve the performance of the automatic respiratory events detection algorithm described in Chapter 5, as well as providing a solution to reconstruct the sound signals for the purpose of auscultation. The overall idea was to only transmit the outputs from the envelope generators. This would still make it possible to run the algorithm in the receiver but at the cost of losing important information from the raw sound signal. Using this technique, for a 3-channel detection system, a compression ratio of 1:0.143 could be achieved by converting the original sound data with 2205 Hz sampling frequency down to 3 bandpass filtered channel envelope output signals at 105 Hz. Although compressing the respiratory sound data in this way does not affect the data analysis, it is not easy to reconstruct signals with the same quality level, in terms of SNR, as presented by Yadollahi [213]. In collaboration with Aguilar Pelaez, a method [201] to reconstruct the sound signal was proposed. The method combines a set of bandpass filtered white noise carriers which were amplitude modulated by the multi-channel envelope signals. The audio perception of the reconstructed sound signals was similar to the original one, but the reconstruction error was still high.

4.1.2 Lossless Breath Sound Compression

Adaptive differential pulse code modulation (ADPCM) [215] is the most common technique used in lossless audio compression. By observing historical values of input samples, ADPCM codes the differences between a predicted value and the present input sample, which tend to be less redundant. Hence, the performance of ADPCM highly depends on the predictor that it uses. Unfortunately, it has been proved that no lossless data compression can always guarantee making any data smaller if the data cannot be fully predicted. Therefore, some audio compression applications, such as Shorten [216], always employ more than one predictor to increase coverage of prediction and maximise its compression rate. In practice, the average lossless audio compression rate that can be achieved for vocal sound is 60 - 70% of the original size [216].

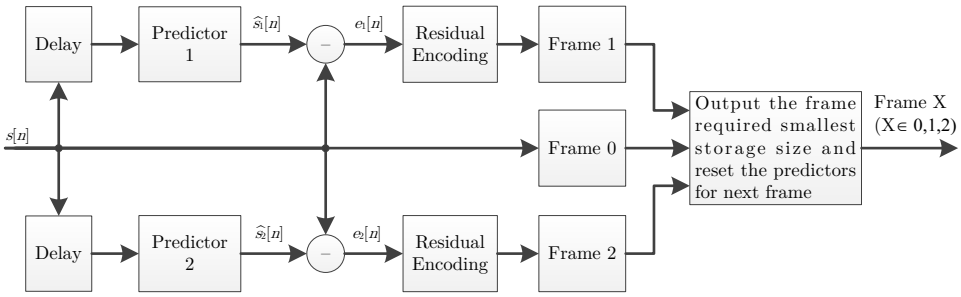

Figure 4-1: Simplified overview of ADPCM with multiple predictions.

Figure 4-1 shows the high level model of common lossless ADPCM with multiple predictions. In the figure, there are two predictors which use delayed audio samples to generate the estimated value of the present sample $s[n]$, where n is the sample number. After subtracting them, the resulting residuals are encoded by adaptive variable length codes (such as Huffman code) according to its Gaussian distribution characteristic (as described in [217-219]) and collected by frames, i.e. Frame 1 and Frame 2 as shown in Figure 4-1. Frame 0 stores the verbatim values of input samples which have not been compressed. In order to maximise the efficiency of the residual encoding, this process has to be optimised according to the statistical characteristic obtained from a predefined number of residuals, which is called compression frame conventionally. However, the actual storage sizes of the frames resulted from different predictors are different. In order to achieve the best compression rate, only the smallest frame obtained from the best performed predictor remains as output of the compressor and the other frames are discarded. The verbatim predictor is used as a reference for the other predictors, which prevents the compressor from compressing data without a smaller output.

The choice of predictors is very limited in embedded systems due to their weak computational power. A common technique to predict the audio input signals is to search a set of filter coefficients allowing the filter to produce the least mean squares of the error signal (difference between the desired and the actual signal). High performance predictors using high order filters with optimal filter coefficients obtained by adaptive algorithms (such as least mean squares or Levinson-Durbin) which aim to generate minimum prediction errors are not power-efficient on a size limited low power wearable wireless sensor platforms. As an example, MPEG-4 Audio Lossless Coding (ALS) which is the world leading industrial standard (ISO/IEC 14496-3:2009) for lossless audio encoding uses Levinson-Durbin adaptive algorithms for generating the adaptive filter coefficients, which is designed to be run on a personal computer [220]. Although

the CPU load is a few percent of the full capacity, it is still too power-hunger to be used on a low power system, taking account of the power consumption used by the personal computer CPU can be tens of watts or more. Unfortunately, there is no low power implementation of MPEG-4 ALS encoder which is suitable for miniature wearable sensors. Most of the implementations require powerful processors similar to the one in smartphones. Recently, an FPGA implementation of an improved variant has been published in [221], which has a form factor of computer graphic card.

The memory of the miniaturised wearable sensor is insufficient for buffering packets for retransmission in cases of busy wireless channels, since packet loss problems are inevitable when a crowded free license wireless frequency band is used. In order to minimise the data loss, caused by any unrecoverable wireless transmission problem, it is advisable to make each packet self-unpackable and disallow its compressor from having access to previous packets' contents so that the data reconstruction is guaranteed not to be affected. Unfortunately, reliable low power wireless transmission systems tend to also have limited payload size in each packet to avoid increasing packet error rate which is proportional to the packet's length. In particular, Bluetooth® Smart only has 20 bytes when maximum throughput is required. This small payload size limits the compression frame size which can no longer be defined by a fixed amount of samples or duration, and also makes high order adaptive predictors impractical because they have more overhead from two sources: storage for coefficients of the adaptive filter and more non-fully predicted samples at the beginning of the frame when there are not enough historical data available. Also, when the goal is to minimise power consumption, the use of simple fixed polynomial extrapolative prediction seems reasonable, because it is well known for its fast processing speed and compression effectiveness. This has been demonstrated in [216, 222].

Let's consider a zero-order polynomial predictor, which predicts each input sample input$[n]$ to be zero.

$$\hat{s}_0[n] = 0 \tag{4.1}$$

$$e_0[n] = s[n] - \hat{s}_0[n] = s[n] \tag{4.2}$$

where $\hat{s}_0$ is the estimated value of the input s, given by the zero-order predictor and e_0 is the error between them, and is so called residue.

A first order polynomial predictor[1] gives prediction $\hat{s}_1$ as the predecessor $s[n\text{-}1]$ of the current input $s[n]$, which gives a first order residue e_1 as shown in the following:

$$\hat{s}_1[n] = s[n-1] \tag{4.3}$$

$$e_1[n] = s[n] - \hat{s}_1[n] = s[n] - s[n-1] \tag{4.4}$$

By substituting Equation (4.2) into (4.4), we have:

$$e_1[n] = e_0[n] - e_0[n-1] \tag{4.5}$$

Following the same approach, it is not difficult to discover that the residues obtained from higher order polynomial predictors can be deduced recursively by their lower order predictors [216].

$$e_i[n] = e_{i\text{-}1}[n] - e_{i\text{-}1}[n-1] \tag{4.6}$$

where i is the order of the polynomial prediction, and $i > 0$.

Hence, the computation of polynomial extrapolative prediction is arithmetically simple and does not involve any multiplications, hence being suitable for implementation in any embedded system with a requirement of extremely fast responsive speed. In terms of compression rate this technique can achieve up to 50% when using low order prediction if the input signal trend is stationary [216]. Hence, this simple prediction works effectively on sound data when the respiration is not present, but it does not perform as good during breathing, especially when the sampling rate of the wearable sensor is not high compared to the signal bandwidth. Although it is possible to improve the compression performance, by running higher order polynomial extrapolative predictors in parallel, the benefit of the technique then becomes extremely limited.

The dynamics of breathing signals create more uncertainty, which may result more residual after the prediction. To solve the problem, a better predictor is required, which has to be able to mimic the breath sound without the need of complicated adaptive filtering and access to comprehensive historic data. With those conditions, compression performance does not have to be affected by the small frame size constraint.

[1] It should not to be confused with the degree of polynomials which defines linear polynomials as first order.

The following sections describe novel compression method that addresses the challenges of all previously reported approaches, preserving the breath sound quality whilst achieving a reasonable compression rate which allows the original breath sound data to be transmitted from the sensor using Bluetooth® Smart.

4.2 A Lossless Low-Complexity Respiratory Sound Compression Method for Wireless Wearable Sensors

4.2.1 Compression Requirements

The targeted breath sounds data for compression are the signals feeding to the respiratory event detection system shown in Figure 5-1. Consequently, the focus of this work was to provide a simple approach to losslessly compress the 10-bit resolution tracheal sound signal after passing through the 200 Hz – 1.1 kHz bandpass filter as described in Chapter 3. Because the compressed breath sounds data are intended to be wirelessly sent in a packet form three constraints were set for the compression method:

1) Each compression frame should not be larger than the maximum buffer size of the chosen transmitter. Therefore each compression frame could be uploaded to the transmitter's buffer and sent out directly as one packet without invoking extra complexity of the control for packet handling. In addition, smaller packet size decreases the chance of wireless transmission being affected by the traffic intensity of the used frequency channels.
2) Each compression frame had to be randomly accessible, which meant that the decompression process should not use any information from any previous compression frame. This prevents any lost packet during the wireless transmission affecting the decompression.
3) The power saving from the data rate reduction should be larger than the power required by the compression implementation.

4.2.2 Principle of Breath Sound Prediction

In order to be able to improve the breath sound prediction, the limitation of the fixed polynomial predictor has to be known. The simplest and most effective predictor is a first order polynomial interpolation as presented in Equation (4.3).

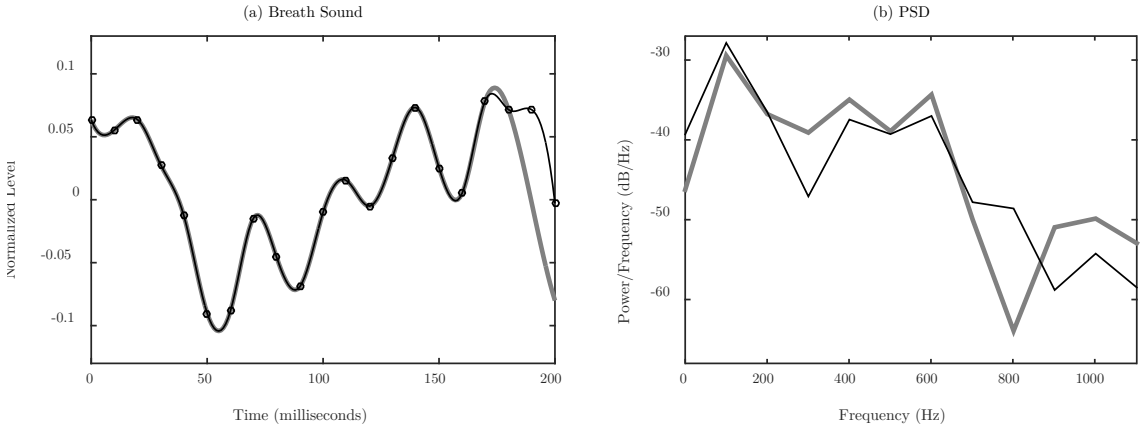

Figure 4-2: An example of how simple polynomial prediction affects breath sound signals in both time (a) and frequency (b) domain. Grey plots are the original continuous-time breath sound signals; circular markers represent the sampled signal with the last two samples replaced by their first order prediction; and the black plots correspond to the interpolated signal.

Replacing the actual input signals s with their predicted values $\hat{s}_1$, is equivalent to replacing the last sample period of the signal with a stationary DC like signal and momentarily suppressing the AC frequency components. For illustration, Figure 4-2 shows an example of a sampled breath signal with the last two samples replaced by their first order polynomial prediction (i.e. one sample delay version $s[n\text{-}1]$). This leads to a series of changes on its power spectral density. It is clear that the energy close to the DC is boosted and there is some attenuation on the predominant spectral contents in the higher frequency range. To remove the effect brought by the polynomial prediction, a high pass filter with a gain slightly larger than unity can be applied on the prediction. Removing the excess in DC power introduced by the predictor and compensating the loss of AC frequency components, helps in reducing the prediction error. The cut-off frequency should be the same as the high pass setting used in Section 2.3.3 in order to avoid too much attenuation of the breath sound signal.

4.2.3 Implementation of Breath Sound Synthesis Predictor

The challenge in synthesizing breath sound signals based on previous data lies on designing a filter that can rectify the errors caused by the first order polynomial predictor whilst generating an output in time with a similar spectral content to the previous breath. The group delay of the filter is crucial for optimal results in time domain. In order for the prediction to work, the

average group delay of the filter in the passband should not be larger than one sample. Otherwise the filter would try to estimate the previous sample but not the current one. A low complexity low order infinite impulse response (IIR) filter has minimum group delay, hence being the ideal candidate for this application. It was found that the optimal group delay was 0.5 samples. This helped to reduce the errors introduced by the predictor in the passband. A second order Chebyshev Type II band pass filter was chosen. The passband of the filter was matched to the bandwidth of the breath sound signal, from 100 Hz to 1 kHz, since the predominant spectral content of tracheal breath sounds is within that range.

The discrete time function of a second order non-unity gain IIR band pass filter is given by:

$$y[n] = \alpha \cdot (B_1 \cdot s[n] + B_2 \cdot s[n-1] + B_3 \cdot s[n-2]) - A_2 \cdot y[n-1] - A_3 \cdot y[n-2] \tag{4.7}$$

where α is the gain; B_x and A_x are the coefficients of the Chebyshev Type II band pass filter.

By substituting the predicted value $\hat{s}_1[n]$ for the current sample $s[n]$ within Equation (4.3), the prediction function becomes:

$$\hat{\hat{s}}[n] = \alpha \cdot ((B_1 + B_2) \cdot s[n-1] + B_3 \cdot s[n-2]) - A_2 \cdot y[n-1] - A_3 \cdot y[n-2] \tag{4.8}$$

Assuming that the band pass filtered prediction $\hat{\hat{s}}[n]$ correctly follows the actual breath sound signal $s[n]$ and the filtered version $y[n]$ of the original signal is similar to its original, $s[n]$, $y[n]$ could be approximated by $\hat{\hat{s}}[n]$ and Equation (4.8) could be simplified further to:

$$\hat{\hat{s}}[n] \cong \alpha \cdot (B_1 + B_2) \cdot s[n-1] + \alpha \cdot B_3 \cdot s[n-2] - A_2 \cdot \hat{\hat{s}}[n-1] - A_3 \cdot \hat{\hat{s}}[n-2] \tag{4.9}$$

Hence the effective prediction errors at the sensor's output, known as prediction residuals for the breath signals can be calculated as:

$$\begin{aligned} \hat{\hat{e}}[n] &= s[n] - \hat{\hat{s}}[n] \\ &= s[n] - \left(\alpha \cdot ((B_1 + B_2) \cdot s[n-1] + B_3 \cdot s[n-2]) - A_2 \cdot \hat{\hat{s}}[n-1] - A_3 \cdot \hat{\hat{s}}[n-2]\right) \end{aligned} \tag{4.10}$$

Substituting $\hat{\hat{s}}[n-1]$ by $\hat{\hat{e}}[n-1]$, the function can be transformed into another IIR filter:

$$\hat{\hat{e}}[n] = B_1' \cdot s[n] + B_2' \cdot s[n-1] + B_3' \cdot s[n-2] - A_2 \cdot \hat{\hat{e}}[n-1] - A_3 \cdot \hat{\hat{e}}[n-2] \tag{4.11}$$

where the new filter coefficients are:

$$B_1' = 1; \qquad B_2' = A_2 - \alpha \cdot (B_1 + B_2); \qquad B_3' = A_3 - \alpha \cdot B_3 \tag{4.12}$$

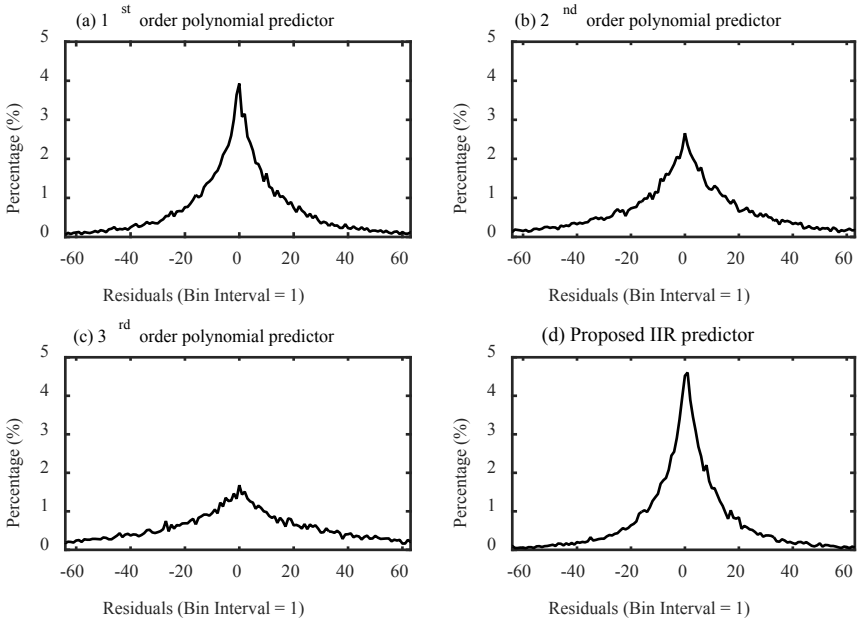

Figure 4-3: Probability distribution of the residuals resulting from breath only sound signals obtained from the trachea, against their first three orders polynomial predictions (a) (b) (c) and the proposed prediction (d).

The reconstruction of the breath signal can be implemented inversely by adding the results from Equation (4.9) and its corresponding residuals $\hat{\hat{e}}$, provided the same set of initial conditions of the synthesis filter are used for each independent frame of data. If the breath sound synthesiser tracks down the signal successfully, $\hat{\hat{e}}$ should be small and negligible. The initialisation time of the IIR filter with the transfer function shown in (4.11) only require the same time as a 4-tap finite impulse response (FIR) filter since $\hat{\hat{e}}[n-1]$ and $\hat{\hat{e}}[n-2]$ are small. Figure 4-3 shows the probability distribution of the residuals obtained from 8 pieces of randomly selected tracheal sound signals, generated when a monitored subject is breathing (i.e. breath pauses are discarded), against their first three polynomial predictions and the proposed prediction. The residuals have a Laplace distribution centred around zero. The higher Q factor of Figure 4-3(d) indicates that the proposed predictor performs better than the rest. In the first three figures, it is clear that increasing the order of the polynomial prediction does not help in reducing the residuals too much. However, by applying the breath sound synthesis filtering on the first order polynomial prediction, the residuals' redundancy has greatly reduced.

4.2.4 Fixed Length Adaptive Sample Size (FLASS) Frame

In order to compress an audio waveform efficiently, the traditional approach is to group a pre-defined number of samples in a compression frame, and then analyse their statistical characteristics to configure the adaptive predictors and residual encoder accordingly. In such case, each frame contains the same amount of samples, but has a different length in terms of actual data storage size. This results in two problems: 1) the size of the frame, defined by sample count or time duration, has to be carefully selected since the level of similarity of the signal characteristics within the frame directly affects both the prediction modelling and the residual coding performance; 2) compression frames with a fixed amount of samples may be too long or too short for optimal wireless transmission. To tackle these issues, in this work the definition of compression frame length has been revised to match the storage size of the maximum payload that can be transmitted wirelessly and efficiently. The number of samples stored in the frame depends on the redundancy of the input data. In the following this will be referred to as Fixed Length Adaptive Sample Size Frame, or FLASS frame.

The structure of the FLASS frame is shown in Figure 4-4, for both verbatim and compressed forms. Both forms start with a 3-bit field called "Type" to indicate which predictor (or no prediction in the case of verbatim) has been used for the frame. The rest of the FLASS frame is dedicated to data storage. For compressed FLASS frame, the data section has been divided in

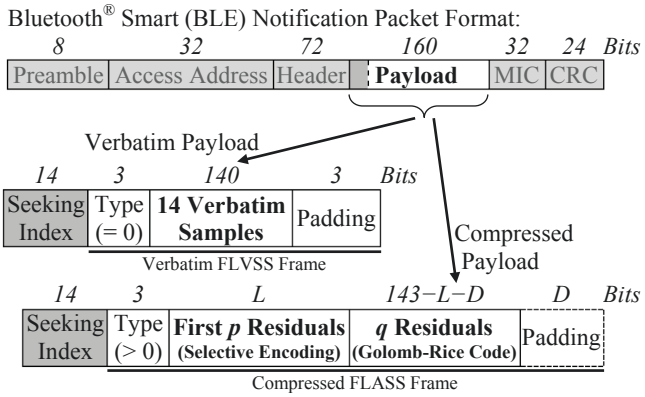

Figure 4-4: FLASS Frame for verbatim and compressed data fitting in a Bluetooth® Smart (BLE) notification packet. The efficiency of the protocol for actual data transmission is 48.78%.

two parts. The first part stores p numbers of residuals resulting from the non-fully predicted samples during the predictor's initialisation. For the proposed predictor, p is set to be 4 as, since mentioned above, the initialisation duration is equivalent to a 4-tap FIR filter. For the polynomial predictors, this p can be their orders. The remaining part contains q sets of residuals from the fully predicted ones, this being encoded in the most efficient way, which is further discussed in Section 4.2.5.

In order to fit as many samples as possible in a FLASS frame, which can on its turn be fitted in the packet's payload, a compression encoder competition scheme is introduced as shown in Figure 4-5. This consists of one verbatim encoder and multiple compression encoders that compete against each other. Each encoder is designed to cope with a specific kind of input signal. For instance in this case, there is a dedicated encoder specifically optimized for breathing sound signals and a general encoder optimized for the rest of non-breathing sound signals, which tend to be more stationary. As soon as any FLASS frame is full, its corresponding predictor and encoder stop working, which makes the system more power efficient. As a result, only the last standing FLASS frame is sent out for wireless transmission. If there is more than one FLASS frame surviving the competition, the most easily interpretable one wins. Once the winner frame is generated, all encoders and FLASS frames are reset to prevent them from accessing previous input data.

Under this scheme, even with small size compression frames, the system automatically seeks for an optimal solution for segmenting the original input data into fragments of variable length with similar amount of redundancy, and passing them into a fixed storage size FLASS frame that can be perfectly fitted in the wireless packet.

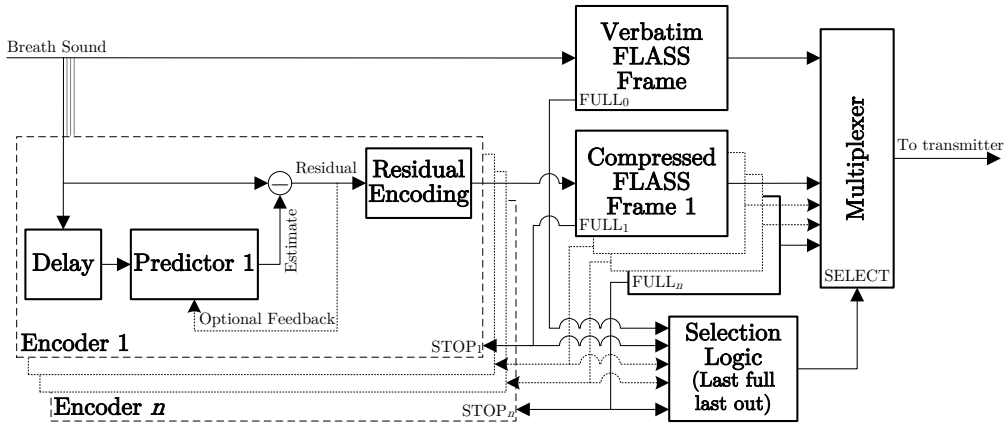

Figure 4-5: Compression encoder competition scheme

4.2.5 Backward Adaptive Residuals Encoding

4.2.5.1 Why Backward Adaptive?

For fast and low-complexity encoding of residuals with a Laplace distribution, variable-length Rice code (known as Golomb-Rice code) has proven to be the most efficient one [216]. During Rice encoding, each residual has its absolute value divided by 2^m:

$$p = \text{int}(\frac{|Residual|}{2^m}) \tag{4.13}$$

$$r = |Residual| \mod 2^m \tag{4.14}$$

The resulting quotient p and reminder r are encoded separately in the format shown in Table 4-1, including one additional bit to keep the sign status before dividing. The Rice encoder expects the quotient part to be much smaller than the reminder part, hence the quotient is coded as a binary representation of p-bit length '0's string ending with a single separator '1', and then the m-bit length reminder is attached to the end directly without any modification. Rice codes are self-delimited and no additional space is required to separate them. The total length of each Rice code is then given as:

$$\text{Length}_{\text{rice_code}} = 2_{(\text{from sign bit and separator})} + p + m \tag{4.15}$$

In order to apply Rice encoding efficiently, the Rice parameter m has to be optimal and known by both encoder and decoder. For most lossless compression methods, as they are designed to run on powerful digital signal processing platforms, it is acceptable to use a variety of computational intensive techniques to search global optimal Rice parameters within the predefined fixed length compression frame. This is so-called forward adaptive encoding. Because the decoder does not have the data until this is uncompressed, the optimal Rice parameter obtained from the forward adaptive approach has to be attached to the compressed data frame as side information to help the decoder to reconstruct the data. This is not feasible to incorporate within the FLASS frame for two reasons. Firstly it is difficult to calculate the optimal value of Rice parameter on a low power embedded system. Secondly, even with a sub-optimal Rice parameter obtained from a light-weight algorithm that could be implemented in a low power manner, the gain from applying it over a short FLASS frame could be less than the space that is required for the side information. Therefore in this case the algorithm for searching the optimal Rice parameter cannot rely on any knowledge of uncompressed data.

Field	Sign Bit	Quotient Encoding	Separator	Fixed-Length Remainder
Length (Bit)	1	p	1	m

Table 4-1: Rice code format and the length of its each field.

4.2.5.2 Backward Adaptive Rice Parameter

It has been proven [216] that the optimal Rice parameter for each individual residual can be calculated as following:

$$m_{\text{optimal}} = \log_2\big(\log_2 \text{mean}\big(\sum|\text{Residual}|\big)\big) \tag{4.16}$$

The conventional forward adaptive approach is to find a single global optimal Rice parameter over a predefined number of residual samples within the compression frame by calculating the average value of the residuals. As it has been discussed in the previous section, it is not very efficient to implement this on a low power wireless sensor platform with very limited memory space.

Instead of trying to get a global optimal value for the Rice parameter, the approach proposed in this thesis reviews the Rice parameter sample by sample using a moving average filter which gives an estimation $\text{Residual}_{\text{moving_avg}}(n)$ of the latest status of the average $\sum|\text{Residual}|$.

$$\text{Residual}_{\text{moving_avg}}(n) \approx \text{mean}\big(\sum|\text{Residual}|\big) \tag{4.17}$$

While the encoding takes place, the moving average filter keeps computing the one sample delayed version of the mean value of the residual.

$$\text{Residual}_{\text{moving_avg}}(n) = \left(1-\frac{1}{L}\right) \times \text{Residual}_{\text{moving_avg}}(n-1) + \frac{1}{L} \times |Residual(n-1)| \tag{4.18}$$

The window length L of the moving average filter is set to half the maximum amount of samples that can be carried by the FLASS frame under the initial Rice parameter setting. To approach the optimal value of Rice parameter shown in Equation (4.16), the decision on adjusting the Rice parameter is based on the following rules:

$$m(n) = \begin{cases} m(n-1), & \text{round}(\log_2 \text{Residual}_{\text{moving_avg}}(n)) = 2^{m(n-1)} \\ m(n-1)+1, & \text{round}(\log_2 \text{Residual}_{\text{moving_avg}}(n)) > 2^{m(n-1)} \\ m(n-1)-1, & \text{round}(\log_2 \text{Residual}_{\text{moving_avg}}(n)) < 2^{m(n-1)} \end{cases} \tag{4.19}$$

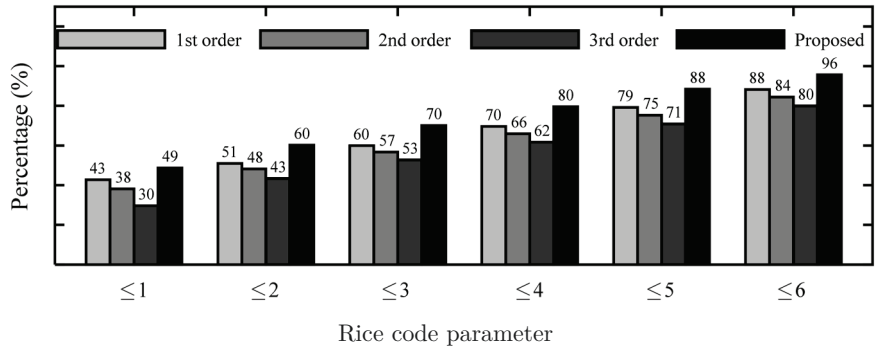

Figure 4-6: Probability distribution of the residuals producing minimum Rice code length with respect to the chosen Rice parameters. The data is randomly selected (from subject 1 in Figure 4-7) 2 hours continuous tracheal sound recording, including breaths, breath pauses and ambient sound noise.

The initial value of the Rice parameter for each predictor is defined according to the characteristics of the residuals obtained from a 2 hour recording from one subject against the corresponding predictions. In Figure 4-6, different values of Rice parameters have been applied to the residuals, and the results show the percentage of the residuals producing minimum Rice code length (i.e. the quotient part is equal to zero). Research [223, 224] has shown that the optimal Rice parameter should have around 50% chance of producing a minimum Rice code length. According to the figure, the initial values of Rice parameters are set to 2, 2, 3, 1 for first order, second order and third order predictors, and the proposed one respectively. Each encoder's moving average filter initial output is reset to 2 to the power of the corresponding Rice parameter when the encoding starts with a new empty FLASS frame.

4.2.6 Selective Encoding for Initial Non-Fully Predicted Residuals

As mentioned briefly in the last section, because the prediction methods used in this work require certain numbers of samples for initialisation for each compressed FLASS frame, a few samples at the beginning could produce large value of residuals. Therefore, the encoding method stated in Section 4.2.5 may not perform as expected. To prevent this from happening, a selective encoding technique [225] has been implemented at the beginning of each compressed FLASS frame. The non-fully predicted residuals use one extra bit during encoding to indicate whether its value is too big to be encoded by the Rice code with the predefined Rice parameter, as shown by the following logic:

IF Residual<2 THEN

Output = ‘1’ + Rice code

ELSE

Output = ‘0’ + 10-Bit uncompressed value

4.2.7 Results

The novel compression method has been tested on tracheal sound recordings from 30 adults with age range from 23 to 82. On average 12 hours of data per subject were used for analysis. These data was obtained during the course of a sleep apnoea clinical test, which is further detailed in the following chapter. The data included both wake and sleep data. The original data was stored in 16-bit WAV format, with zeros in the least significant bits (LSBs). Figure 4-7 compares the results obtained using the proposed compression framework, using polynomial and the proposed breath sound synthesis predictions, with the results obtained using the reference MPEG-4 ALS encoder [226]. The compression performance of this reference codec is only 5% below the state-of-the-art [227]. The MPEG-4 ALS codec was given the following encoding settings:

‘-l’: It indicates the actual input audio files have a bit length of 10-bit.

‘-p’: This enables long term prediction which could benefit from the repetitive respiratory sound signals.

‘-r-1’: This forces the encoder to make every compression frame random accessible.

The compression can be calculated using the following equation:

$$\text{Compression} = \frac{\text{Compressed Data Size}}{\text{Original WAV File Size} \times 10/16} \times 100\% \tag{4.20}$$

where the compressed data size should include the information for random access:

$$\text{Compressed Data Size} = \text{Size of Seeking Index} + \text{Size of Compressed Samples} \tag{4.21}$$

Table 4-2 compares the average compression performance between the low complexity predictors and the reference standard. It shows how the light-weight polynomial predictions are far away from the optimal compression rate established by the industrial MPEG-4 ALS audio compression standard regardless to their filter orders. However, the proposed approach is able to outperform the MPEG-4 ALS codec by 5.54% on average. By looking into the detail of all cases

shown in Figure 4-7, the proposed approach achieved higher compression rate for half of the cases (15 out of 30). This indicates that the way a small size FLASS frame seeks for local optimal of Rice parameters could work as efficiently as the approach of using bigger frame sizes to search the global optimal of Rice parameters. This is because the amount of samples in a FLASS frame is adapted to the entropy of the incoming breath sound data, which results on the system automatically adjusting the size of the compression frame (in terms of sample count) according to the input data, making the residuals encoding more efficient.

	MPEG-4 ALS v23	**Polynomial Predictors**		**Proposed Predictor**
		1^{st} Order	$0^{th} + 1^{st} + 2^{nd} + 3^{rd}$ Order	
Average Compression	36.46% ± 7.45%	51.52% ± 7.83%	50.03% ± 7.69%	34.44% ± 1.51%

Table 4-2: Average compression performance

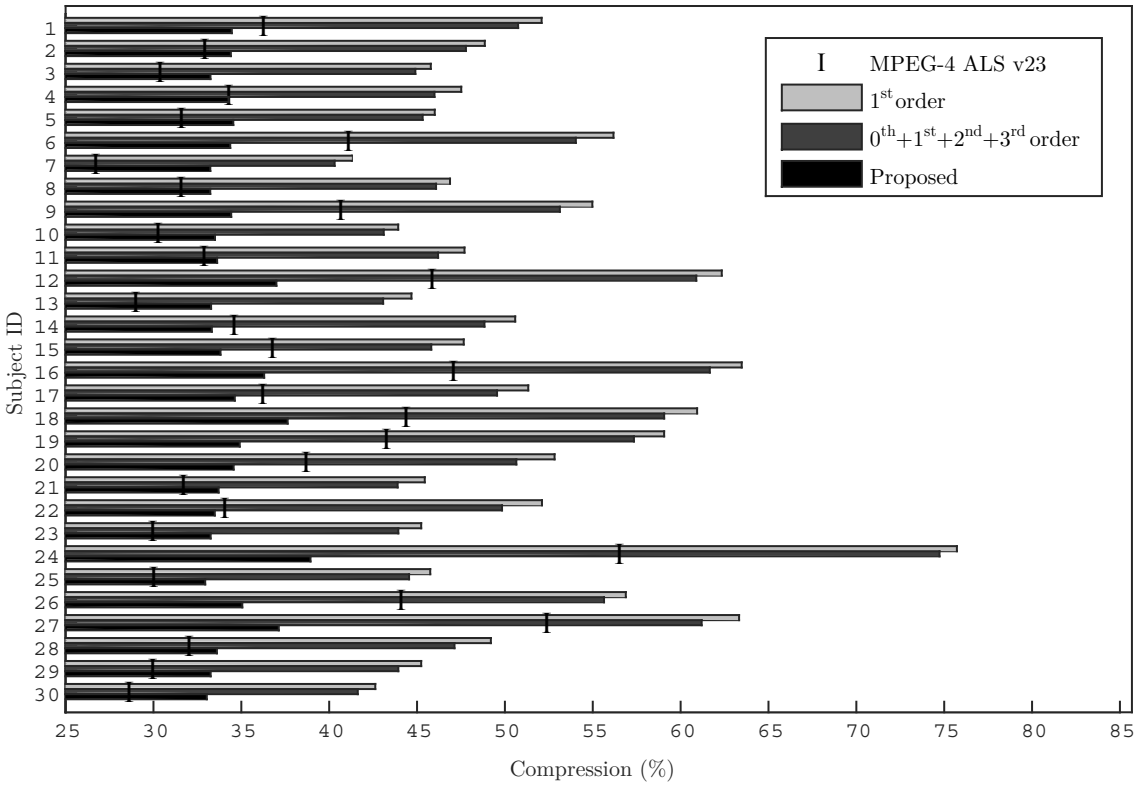

Figure 4-7: Compression performance for 30 different subjects. All compression methods are implemented with backward adaptive Rice code parameter, except the MPEG-4 ALS which uses its own proprietary encoding approach.

The compression ratio for the proposed approach is in a range between 32.94% and 38.94%, giving a data reduction of at least 60%. Taking into account that the Bluetooth Smart transmission protocol efficiency is 48.78% (see Figure 4-4), this shows that the overall air data rate can be reduced to 18.04 kbit/s from 45.1 kbit/s. Given that the maximum average power consumption of a typical Bluetooth Smart transceiver, TI CC2640, is 11 mW at its theoretical maximum air data rate 1 Mbit/s [228], the total power saving thanks to the reduction of data transmission is 298 uW. In order to implement the proposed predictor and encoding scheme, the computation can include an IIR filter, as shown in Equation (4.11), one moving-average filter for the proposed predictor, and some bit-shifting operation. This can be implemented on the 32-bit ARM® Cortex®-M3 microcontroller inside the TI CC2640 with no more than 10% of CPU usage at the speed of 10 MHz. The overall power consumption for the proposed compression method is as little as 61 uW, which results in 237 uW of net energy saving.

Chapter 5

Automatic Adaptive Respiratory Event Detection and Analysis Algorithms

5.1 Background

Regardless of the issue of how to sense respiration discussed in Chapter 2, developing a reliable non-invasive respiratory event detection algorithm for any respiratory monitoring system has been one of the most difficult unresolved challenges in physiological instrumentation research for a long time. Most related work just aim at achieving acceptable on sensitivity for breath detection in artefact-free signals and in most cases there is not even an attempt to evaluate performance in real non-controlled environments. Existing algorithms can be grouped into sound amplitude based methods, machine learning based methods and sound energy based methods.

5.1.1 Sound Amplitude Based Methods

Sound amplitude based approaches to detect breath events operate by observing the maxima of the tracheal sound signal from a 20 seconds[1] moving time window with a moving step of 1 second [214]. The relationship between the variation of the obtained maxima and the classification of respiratory events was studied. Assuming that the sound pressure levels from the apnoea/hypopnoea events were lower than the inspiration/expiration, and those problematic events were seldom over 20 seconds, there were three cases considered. A section of signals with larger fluctuation of maxima identified as 'thick' pattern could indicate sleep apnoea/hypopnoea events. A section of signals with smaller and close to no variation of maxima identified as 'thin' and 'plain' patterns could reflect normal breathing. Although they claimed the sensitivity and

[1] The authors tried a range of time window length from 20 to 60 seconds, and achieved best results by using 20 seconds.

specificity are 80% and 93% on sleep apnoea-hypopnoea event detection for ten sleep apnoea patients in the quiet sleep laboratory environment, the reference they compared to was the results from their own previous preliminary work [229] which has not been fully clinical proven[1]. This implies three things. Firstly, any apnoea or hypopnoea event longer than 20 seconds could be missed by from the detection process. Secondly, this automatic detection algorithm has been only proved that it has an ability to identify problematic sections of tracheal sound which may contain apnoeas/hypopnoea. Finally, its precision has not been proven to be good enough to evaluate location or length of the individual respiratory events accurately. Nevertheless, the technique was tested for nine control subjects and 24 patients later and a correlation between occurrence frequency of the thick pattern periods and AHI index has been found [230]. There is no other further work suggests any amplitude based algorithm is able to detect breathing or apnoea events.

5.1.2 Machine Learning Based Methods

To identify and quantify suitable sets of features for breath sound detection is not easy. Machine learning techniques can be used in order to avoid doing that manually. The techniques include Artificial Neural Networks (ANN), Naïve Bayes (NB), Random Forest (RF), Markov Models (MM or HMM), Gaussian Mixture Model (GMM), genetic algorithms (GAs), Support Vector Machine (SVM), k-Nearest Neighbour (k-NN) algorithm and Fuzzy Logic (FL) [231, 232]. Although there are a lot of machine learning techniques, the reviews [231, 232] show little researchers focusing on separating general breath sounds from unpredictable ambient noises. Instead, most of them focus on the classification of normal breath and adventitious breath sounds. The works focusing on separating adventitious breath sounds such as wheeze and crackle from normal breath sounds are not considered here.

For some of the machine learning techniques called supervised learning to work, they require certain amount of training data. The training data have to cover sufficient breath sound features in order to produce an optimal model that can be used for breath detection. Because there is no reliable automatic breath sound detection algorithm, most of the data required a specialist to manually mark the breath events in order to ensure the quality of the labelling

[1] The method invoked a human to classify the events from the pre-processed data visually. The results obtained from 10 sleep apnoea patients (none of them is control) were based on Apnoea/Hypopnoea Index (AHI) comparison, not event by event comparison. Therefore, it is impossible to quantify the accuracy of the apnoea/hypopnoea event detection.

[232]. This leads to a problem of getting sufficient high quality training data to develop a supervised machine learning algorithm. Therefore, cross validation methods are commonly used for evaluation of the algorithm performance. The accuracy of the breath detection is averaged over individual models generated by the cross validation method. Most of the algorithms are based on the changes in power spectral density of the sounds, except one recent work [233] tried serval different supervised learning techniques utilising 10 acoustic features which include almost all the discovered normal and adventitious lung sound features[1] from [234-237] and some additional features used in speech processing. Although, the trained models were able to identify seven different classes of breath sounds with up to 85.4% accuracy, the actual performance for combined breath sound detection without the classification has not been published. Some [238, 239] tried to develop unsupervised machine learning algorithms, but the dataset they analysed was very limited. These issues significantly limit the reliability of the resulting models on breath/apnoea detection in the real world situation. Table 5-1 lists the details of data being tested in the literature discussed above.

Researcher, Year	Method Type	Detail of Testing Data	Detail of Training Data
Le Cam, 2008 [238]	MM	2 respiratory cycles from one subject and 3 respiratory cycles from another subject.	N/A (Unsupervised learning)
Jin, 2009 [239]	GAs	600 seconds recording was done for each subject. The dataset consists of seven healthy subjects and 14 patients.	N/A (Unsupervised learning)
Alshaer, 2014 [233]	NB, SVM and RF	Three set of five minutes recording was extracted from the first, middle and last third of the overnight sleep study of each subject. There are 11 subjects in total.	Leave-one-subject-out cross validation was implemented. Data from one subject was used as training data, resulting 11 sets of trained models.
Palaniappan, 2015 [240]	FL	68 sections of around 10 seconds recording from RALE database were used. Different sections of recording were obtained from various subjects, but the number of subjects is unknown.	20 seconds recording was done for each subject. There were 69 subjects.

Table 5-1: List of machine learning algorithms tested on breath detection

[1] Lung sounds include the sounds obtained not only from trachea, but also other positions such as chest.

5.1.3 Sound Energy Based Methods

In earlier works [77, 241] on automatic detection of breathing features from acoustic signals, breath sounds can be recognised and captured by convergence of acoustic mean powers using an array of microphones placed on the same subject. If breathing is present, sensors from different location have similar changes on their average power. Later in 2000, a first algorithm [242] was developed by Hult et al. to detect breathing using just tracheal sound. The algorithm compares the integral of signal's power spectra from 400 Hz to 600 Hz to a certain threshold, which is able to generate two respiration states: respiratory pause and breathing.

From 2006, Z. Moussavi's research group is one of the most active groups in the field of tracheal breath sound power analysis. Yadollahi and Z. Moussavi who had previously been studying the correlation between breath sound and airflow proposed a so-called apnoea detection method [243]. Similar to Hult's [242], the method analysed the difference between average spectral powers of the tracheal breath sounds during breathing and breath hold periods from 75 Hz to 600 Hz whilst also using tracheal sound entropy related to its power spectral density. A further improvement to this method incorporates a heart sound cancellation [244]. Comparing with Hult's approach, Yadollahi's one is better because the threshold to determine respiratory pause and breathing is obtained by a nonparametric and unsupervised adequate method [243, 245]. To maximise the separability of breathing and breathing pauses, Otsu's method [245] was used to obtain the optimal threshold based on the histogram of the median value calculated from the tracheal sound entropy signals in each 200 ms frame. The statistical model assumed the sound entropy from a breath event is generally higher than the time when breathing is absent, without considering the case when noise power becomes comparable to the breath sound power. Moreover the detected respiratory pause periods demonstrated in this paper are 2, 4 and 6 seconds. This is not sufficient enough to prove that it could detect apnoea, which is defined as a respiratory pause lasting for at least 10 seconds [153].

As a continuation of Yadollahi's work [243] on apnoea detection, they turned their focus to correlating the bandpass filtered tracheal sound energy with the amount of respiratory flow with the aid of heart-sound cancellation technique taking effect on the band between 75 – 300 Hz [159, 185, 235]. The best accuracy they achieved on respiratory flow estimation is between 6.03 – 12.73% by bandpass filtering the tracheal sound in the range between 75 – 600 Hz. However,

their method assumed that there are no noise artefacts in the data. This can be confirmed when they tried to detect obstructive sleep apnoea [80] three years after the first publication of their respiratory flow related works. In this paper, they defined noise which could be caused by body movement according to the level and duration of the sound segment energy compared to adjacent segments, with additional help from the amount of drop in blood oxygen saturation. It is not difficult to see the limitation of this approach which could give false detection when a noise segment has similar energy level and duration. For this issue, they only addressed the noise caused by swallowing, thought there are many more kinds of noises that have similar energy level and duration to breath sounds. The problem has been revealed in the subsequent validation [246] that showed that the error of flow estimation from tracheal sound significantly degraded from ±8% to ±20% during awake condition rather than asleep condition. From 2010 onwards, their research and publication on breath detection has been stalled, and they have moved one step forward to work on the classification of respiratory events such as phase [83, 86] and snore[163, 247-250], which all assumed the breath detection problem has been solved and did not include the errors caused by false breath detection.

Eventually, people started to realise that breath detection cannot be solely based on either the level of amplitude or the power of the sound signal. A tracheal breath sound feature [251] based on the relationship between heart sound energy and breath sound energy has been identified for the purpose of breath detection. It utilised the fact that heart sounds are persistent during respiratory phase or respiratory pause period in most of circumstances [252]. Since the dominant frequency bands for heart sounds and tracheal breath sounds are 0 – 300 Hz [235] and 400 - 600 Hz [253] respectively, a breath sound signature can be observed from the energy ratio between these two bands. However, in the cases of heart rate variability and heartbeat pause aligning with respiratory pause, using this feature can cause detection errors. In addition, the detection method only works in clinical applications with little interference.

There is no publication on multiple signal energy features based breath detection for ambulatory use until 2007, when Rodriguez-Villegas' research group developed a novel detection algorithm [181] which outperforms the ones proposed in Hult's [242] and Yadollahi's [243]. The first generation of this detection algorithm used a frequency band from 400 Hz to 600 Hz since this has the closest correlation between sound pressure energy and respiratory flow. This frequency band also has relatively low heartbeat sound power hence resulting on a higher signal-to-noise

ratio (SNR) if the environmental noise is not taken account. Also, this detection algorithm, instead of monitoring the tracheal sound entropy, it examined the root-mean-square (RMS) power of the bandpass filtered sound signal and compared it with two threshold levels T1 and T2, where T1 > T2. If the RMS power of input signal reached the higher threshold T1, the signal was recognised as speech or breathing. If the mean power was below the lower threshold T2, it could be interpreted as absence of breathing. The middle situation between T1 and T2 could however occur when the SNR of the input signal was not high enough. In that case, the respiratory state could not be determined by a single factor and a segmentation algorithm was adopted to separate the high RMS power fragment from the input signal as a breath candidate segment. The integral and length of the breath candidate segment should then be within the range in order to confirm it was breathing. A latent problem of this algorithm was shown in Serrano's work [254] which aimed to find optimal thresholds. Unfortunately, he failed to find a single set of optimal thresholds for different recording signals due to loudness difference.

After Serrano's attempting, the algorithm was revisited throughout the course of Chen's Master project[182][1]. The main changes to the algorithms carried out as a conclusion of the results obtained during this preliminary study were:

1. Breath recognition is not just based on a single frequency band, which will be referred from no on as channel, but on a breath spectral chart (see Section 2.2.2). Every potential breath segment is evaluated by comparing ratios between the integrals of the spectral powers for different channels with subtraction of the corresponding noise floors against the pre-defined breath spectral chart. The segmentation algorithm still operated in the first channel.
2. Noise floor estimation and removal from different frequency channels was used in order to improve the noise resilience when the SNR was not sufficient. Each noise floor was calculated from its corresponding minima in mean spectral power.
3. All absolute thresholds of the breath discriminators were transformed to relative thresholds. However, most of them were still fixed and not optimal for different test subjects.

These enhancements improved the specificity of breath detection whilst still maintaining a moderate capability in identification of most breath sounds from other sounds sources. However, as the data sampled increased throughout the course of this PhD it became apparent that the

[1] Note that I am the author of that MSc project.

performance of this version of breath detection algorithm was still very limited. To solve the problems, further investigation into improving the accuracy of the algorithm was carried out which constituted the core of this PhD work. Specifically, the effort was put into making the process of feature extraction and breath sounds features recognition adaptive and more noise resilient. The aim was also to design a respiratory sound processing algorithm which did not require pre-calibration, participant specific knowledge, or any other kind of modification.

5.2 Overview of Proposed Event Detection Framework

Figure 5-1 shows a block diagram of the proposed 4 core stages (marked as a, b, c, d) respiratory event detection framework. A conditioned sound signal is taken as the input (for detail see Section 2.3). This is split into n channels ($n \geq 3$) by a group of envelop generators (stage a). Stage b segments the signals. And stages c and d extract appropriate segment features for the two-level breath sound verification system to be able to judge the likelihood of a given segment being a breath sound. Each stage of the verification process provides more information and allows the system to have improved sensitivity and specificity in breath detection. From stage b to stage d, the system gives 2 intermediate judgements which are breath-like (BL) segment and breath candidate (BC) segment; decides which segments are respiratory pauses; and classifies a subset of BC segments as inspiration, expiration and standalone. A decision tree

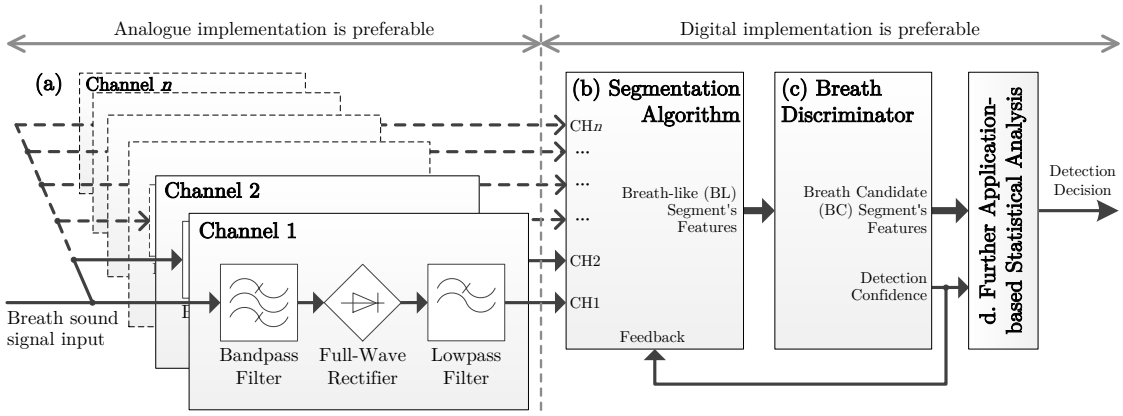

Figure 5-1: Overview of the respiratory event detection framework

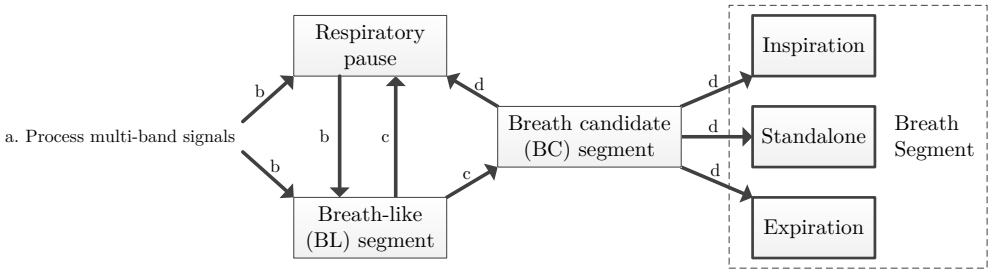

Figure 5-2: Decision tree of breath detection system. The label of each decision making flow refers to its corresponding block shown in Figure 5-1.

given in Figure 5-2 shows how intermediate judgements transform to final decisions. The labels of the decision flow diagram refer to the stage labels in Figure 5-1. Further details of the blocks inside the algorithmic framework are given in rest of this chapter. The contents described below can be found in this patent application [255].

5.3 Real Time Breath Detection Algorithm

5.3.1 Generation of Multi-Channels Signals

The main steps of multi-channel signal generation are shown at Figure 5-1. These comprise taking the conditioned sound signal using the specially designed transducer described in Section 2.3, splitting the signal into frequency bands by a set of bandpass filters, rectifying the signal, and smoothing the signals within the bands. Because this initial stage only contains filters and rectifiers which deal with the acquired breath sound signal, it would be possible and advantageous from the power point of view to use analogue implementations. According to Corbishley's works [256], the power consumption of the individual building blocks of envelope generator can be as low as 70nW for each 6th order bandpass filter [257], 600nW for each full-wave rectifier [258], and 5nW for each lowpass filter [259]. In total, 2.03 uW of power would be required by the minimum 3-channel system.

Theoretically, because the bandwidth of these multi-band signals could be as low as the cut-off frequency of the smoothing lowpass filter, the subsequence blocks can operate at much lower speed. This relieves the computational intensity of breath feature extraction and recognition stages, enabling more freedom on their design within the limited power budget and no longer being restricted in analogue domain.

5.3.2 Real-time Segmentation Algorithm

5.3.2.1 Analogue-Digital Conversion

In order to guarantee the reliability of noise rejection and due to the overall complexity of the algorithm after the generation of the multichannel signals it is preferable to perform any further computation in the digital domain. In the already rectified and smoothed bands the required sampling frequency can be significantly lower than in the original breath sound signal. However, to maintain a certain level of system responsiveness, the segmentation algorithm processes the multi-band signals at 105 Hz speed, which results on an equivalent detection time in the order of 1 ms. In a custom integrated implementation a 10-bit analogue to digital conversion, could require as little as 53 nW if a successive approximation register (SAR) analogue-digital converter (ADC) is used[260]. Together with the peripheral components of the ADC which include a reference generator shared by all ADCs and a clock generator of up to 2 kHz the digitisation stage of a customised 3-channel breath detection system could require as little as 22uW [261, 262].

5.3.2.2 Basic Routine

As Figure 5-2 shown, the first step of the so called segmentation is to initially separate breathing like (BL) segments from respiratory pauses. The basic routine of searching for the BL segment start and end points can be divided into two states, called 'valley seeking state' and 'peak seeking state', as described in a simplified timeline diagram shown in Figure 5-3. For illustration, Figure 5-4 is given as an example to show the location of all the important points helping in the search of BL segments as described below.

During the valley seeking state mean SNR estimation is performed for every channel initially by using the method provided in Section 5.3.2.9. The channel that has a highest mean SNR is taken as the main channel to drive the segmentation state machine by comparing it with a series of adaptive thresholds. The other channels assist this process by only using their trends.

Once the main channel is selected, the next step is to verify and update the last defined BL segment preliminary end point. This allows amendments when some of the beginning sections have being wrongly recognised as part of respiratory pause. Depending on the targeted application requirement, the end point verification may not be needed. But if it is required, the

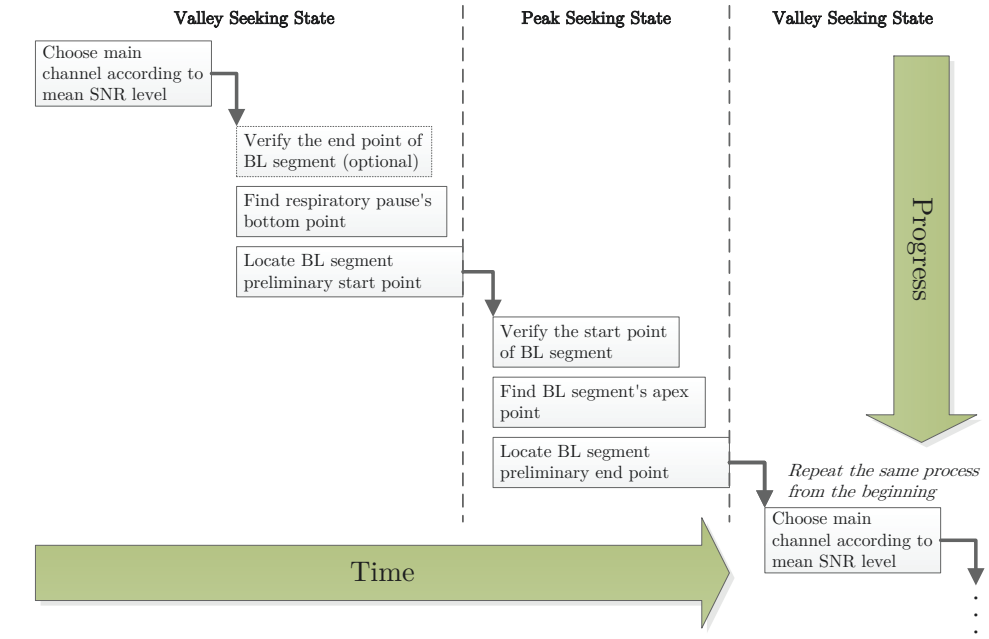

Figure 5-3: Basic routine for a segmentation algorithm. Note that the actual duration of each step is not to be scaled from the length in the diagram. The difference on the length indicates the order of steps finishing.

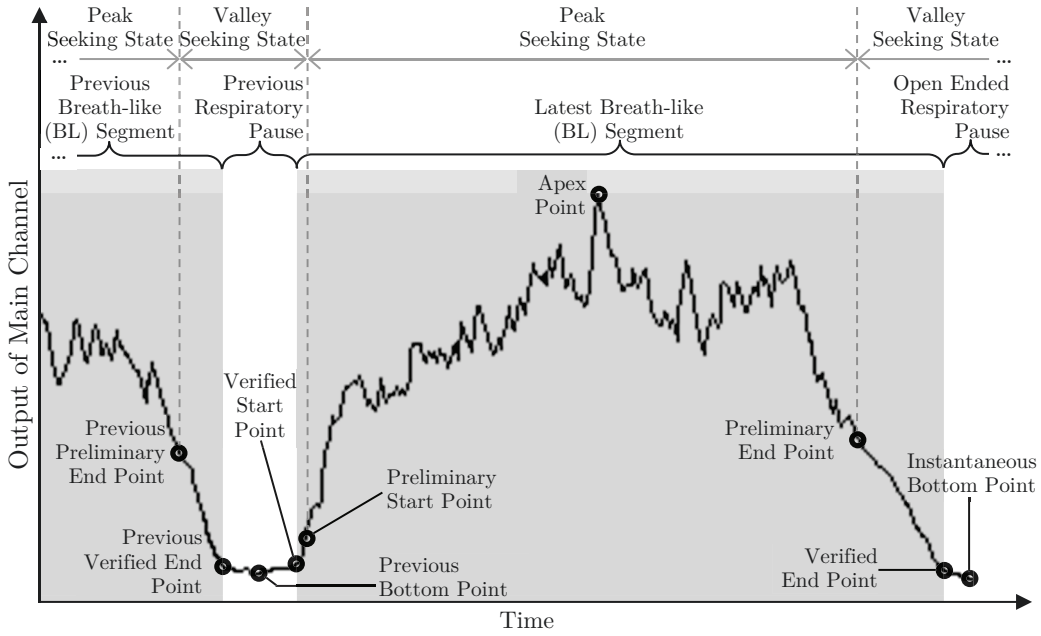

Figure 5-4: An example of a new breath-like (BL) segment being identified. This example assumes the selected main channel has the highest signal-to-noise ratio (SNR) within the shown time frame. In reality, the main channel may be switched at the transit from peak seeking state to valley seeking state.

verification process keeps postponing the BL segment end point until all the monitored channels do not look like a part of breath's release (see Section 2.2.3.3). Regardless whether the end point verification is performed or not, the algorithm seeks a point called 'bottom point' with a minimum signal level in the main channel from the moment the valley seeking state begins. The position of a new BL segment preliminary start point is being found as the point at which the signal rises above a certain value which can either be a percentage of the bottom point, or a percentage of the combination of preliminary start/end points from the previously found BL segment, or a point at which there has been a significant slope change with respect to the major channels. The bottom point is allowed to be updated when the main channel reaches a lower minimum level. This is until the BL segment preliminary start point is found.

From then, the segmentation algorithm may enter peak seeking state. This temporary start point will have to be further verified and updated to the point when the channels just start sharing some similar characteristics of breath attack (see Section 2.2.3.1). This step is compulsory because the SNR is relatively low when a breath is initiating and the detection of the preliminary start point may be false due to noise. At the same time in the verification, a maximum signal point called 'apex point' is being determined, while a preliminary end point can be chosen based on a method similar to the preliminary start point by analysing the main channel slope, or its signal against BL segment's apex point or historical preliminary start/end points. In contrast to the start point searching, the determination of the end point is much easier because the initial condition has higher SNR. Therefore, this preliminary end point may be used without verification. However, if it is required for better accuracy, the verification can be performed later when the routine returns to valley seeking state.

At this point the iteration of the segmentation routine is completed. The routine then returns to valley seeking state and begins identifying more BL segments. Any BL segments identified during this process indicates that the corresponding fraction of the original sensed signal may probably contain breath sound signal. Then the information in each of the channels within a BL segment can be assessed by the breath discriminator. The detailed implementation of each step above is provided in the following sections.

5.3.2.3 Preliminary Start Point of Breath-like (BL) Segment

Referring to the approach described above and with reference to the example depicted in Figure 5-4, which shows the outline of the signal in the channel with the highest mean SNR, the preliminary start point of the last appeared BL segment in this particular example occurs after the actual BL segment start point. The preliminary start point is the point at which the signal is ready for peak seeking state and represents also the point at which the main channel has raised to a certain proportion from a certain value which is related to the recent identified bottom point in the last respiratory pause. However these bottom points do not necessarily have to be the only reference to determine the preliminary start point, since these ones may be corrupted by noise. Therefore, there is a set of rules for the determination as listed following.

A point can be recognised as preliminary start point if:

a) The current value of the main channel is twice the value of the recent identified bottom point and the rest of the channels have positive slopes.

OR

b) The current value of the main channel is higher than the last preliminary end point and the rest of the channels have positive slopes.

5.3.2.4 Preliminary End Point of Breath-like (BL) Segment

The preliminary end point is the point at which the signal is ready for valley seeking state and represents also the point at which the main channel has fallen to a certain proportion from a certain value which is related to the apex point in the current BL segment. Similar to its counterpart above, these points cannot be identified by the apex points solely.

A point can be recognised as preliminary end point if:

a) The current value of the main channel is 30% of the recent identified apex point and the rest of the channels have negative slopes.

OR

b) The current value of the main channel is lower than the last preliminary start point and the rest of the channels have positive slopes.

5.3.2.5 Slope Monitoring

Referring to the approach described before the slopes of the different channels are monitored throughout the segmentation routine. This is used to determine the degree of monotonicity of the signals, i.e. whether the signal is mostly increasing or decreasing. Measuring the channel's slope can be done by taking samples at regular intervals and evaluating the difference, for instance with a time distance of 66.7 ms in this case, which is the same time window's length for the attack and release analysis mentioned in Section 0.

5.3.2.6 Turning Point Detection

In order to locate the true start and end points of BL segments, a concept of 'turning' points is introduced. A positive turning point means the transition point when slope increase dramatically, and vice versa. A positive turning point occurring before a preliminary start point could be much closer to the actual start point of the BL segment. Similarly, a positive turning point that happened right after a preliminary end point could be a better approximation of a BL segment end point. If a negative turning point is detected during the beginning of a BL segment, it may be an indication of noise corruption, making the possible breath segment looks like it has started before the preliminary start point, but it was not. These turning points can be found following a method such as the one shown in Figure 5-5, which is independent to the segmentation state machine. Changes in slopes of the envelope signals from the secondary channels are evaluated at every sample interval.

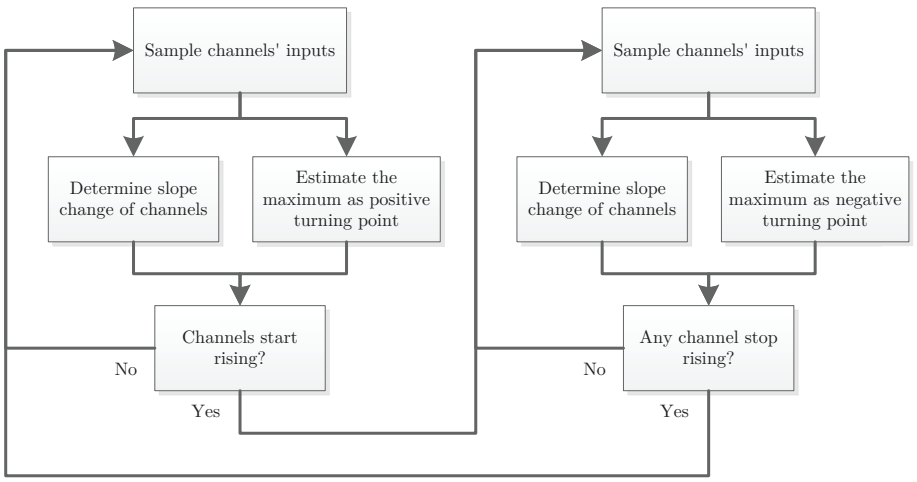

Figure 5-5: Flowchart of turning points detection

Depending on the input signals, the analysis loop starts with a search of the positive turning point (left side in Figure 5-5) if the initial slopes of all the channels are negative. Before the slopes turn to positive indicating all channels start rising, the point with maximum summed value from the channels in the slope analysis window is assumed as the positive turning point. The result of positive turning point is latched when the routine turns to search a negative turning point.

Negative turning points can be found following an equivalent procedure as depicted in Figure 5-5. In this case, the slopes of all channels are continuously being monitored until a negative trend with the same time window is found. Generally, once a negative turning point is found the search for a positive turning point will start and vice versa, reinitialising the whole process.

5.3.2.7 Verification of Breath-like (BL) Segment Start Point

Identifying when actual breath sound starts is one of the most challenging parts of the breath detection process. Any noise can trigger false detection long before the actual starting time of the true event because of weak power level of breath sound when it is initialising. Since some of the feature extraction (see Section 5.3.2.4) relies on accurate segment start point, the verification of these start point is recommended. According to the characteristics of breath attack described in Section 2.2.3.1, a certain verification time is established and applied to the signals before and after the preliminary start point.

First of all, to verify the current BL segment start point, it is necessary to ascertain that the previous respiratory pause is detected correctly. As described in Section 0, the sustain section of some breaths could be shivering. According to Figure 2-11 and Figure 2-15, the statistical results show that most shivering breath can drop below the preliminary end point and return to the same level in less than 160 ms. Therefore the following three rules can be used to determine whether both the previous BL segment and the current one belong to a single respiratory phase.

1) The previous BL segment must be classified as a BC segment by the breath discriminator, regardless of how high the confidence level was. Further details of BC segment definition can be found in Section 5.3.3 and 5.3.4.
2) The current preliminary start point is the same as the previous preliminary end point. This indicates that the main channel has reasonable power and never drops below twice of noise floor.

3) The time distance between these two preliminary points is less than 160 ms.

In this way, the method stops the verification by rewinding the start point to the previous one, effectively extending the previous identified BL segment to the current one.

If the condition above does not happen, within a time period prior to the preliminary start point, all the channels have to rise to a considerable level compared to a certain pre-established value, which is normally a function of adjacent positive turning point. This time period can be the average attack time taken by the envelope signals to rise at the beginning of an inspiration or expiration event, for example 0.2 seconds as shown in Figure 2-9. For example, referring to Figure 5-4, which represents the main channel, assuming the hypothetical scenario in which a turning point (as identified from Section 5.3.2.6) is to be evaluated as the potential start point of the BL segment, the time from the turning point to the preliminary start point should be less than the average breath attack time, which is 0.2 second. Otherwise, the point that is 0.2 second prior to the preliminary start point is taken as a potential start point. This potential start point is also shown as point 'A' in Figure 5-6 and subjected to further verification.

After the preliminary start point, the verification continues with the same time period to prevent the case shown in Figure 5-6 when noise corrupts the beginning of the segment. The method could find a wrong starting point 'A' before the preliminary start point for the segment accidentally, for example, if the noise floor had increased. Usually, the exception will happen when the noise floor is low. Because of this, verification of the potential start point found before the preliminary start point is beneficial for highly accurate applications to overcome the problem. Firstly, the change ΔAB between previous found potential start point 'A' related to next found negative turning point 'B' is recorded. After that, the method tries to find another positive turning point 'C' within a time limit of T_1 whilst at the same time searching for an apex point. Within another time limit T_2 away from the previous found potential start point, if a point 'D' is detected for which the difference ΔCD exceeds double of the previous change ΔAB, the positive turning point 'C' will replace point 'A' as the verified BL segment start point. Empirically the time thresholds for T_1 and T_2 are 0.1 second and 0.3 second respectively.

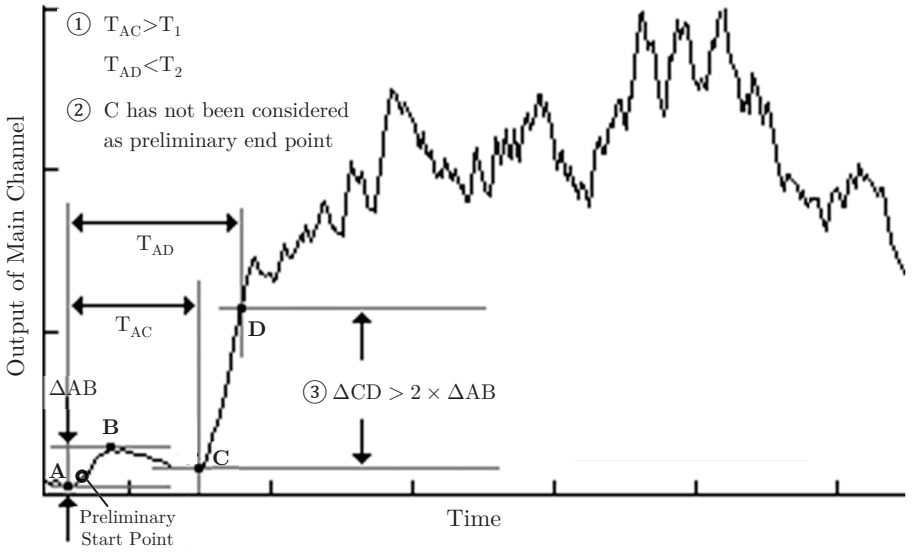

Figure 5-6: An example of breath-like segment with ambiguous starting. When ①, ② and ③ conditions are met, point 'C' replaces the point 'A' as the verified start point of the breath-like (BL) segment.

5.3.2.8 Verification of Breath-like (BL) Segment End Point

Verification of BL segment end point may not be necessary because most of the features of the BL segment have been found at the point when its preliminary end point is detected. With those features which will be detailed in Section 5.3.3, the breath discriminator is fully capable to judge the possibility of the BL segment being a breath event before it really ends. However in some cases, such as tidal volume estimation [152], a better estimation of breath ending is crucial for measurement accuracy.

When the main channel falls below the preliminary end point, the verification starts searching for a better estimation by locating the first appeared positive turning point, for a limited duration. In this case, the optimal time limit can be estimated from the findings in Section 2.2.3.3 which suggest that the average breath release time is 0.25 second. If there is no positive turning point found within 0.25 second, it is assumed that the point which is 0.25 second beyond the preliminary end point is the verified end point.

5.3.2.9 Channel Noise Analysis

Even though the signal has already been filtered there will still be a certain level of noise within the channels. From time to time the noise of the channels may vary making it difficult to segment the signals. The noise will also vary between the different channels and so is determined independently for each one of them. The easiest way to improve SNR is to subtract

the stationary part of the noises (as known as noise floor) from their corresponding channels. Empirically the noise floor can be observed at the bottom points of each channel signal within respiratory pause sections considering a certain number of previous consecutive BC segments (i.e. BL segments which passed the breath discriminator check). When identifying a pause section, it is noted that, generally, the respiratory pauses between the inspired and expired breath are shorter than the pause between each full respiratory cycle, as shown in Figure 2-16. Also, the breathing pause within a respiratory cycle may be too short for the envelope signal to reach a true minimum due to hardware delay. This can lead to wrong noise estimation if short time frames for noise evaluation are considered. The time frame for noise floor estimation and values updating should be long enough to cover at least one respiratory cycle composed of two BC segments and three respiratory pauses and hence all the relevant minima should be stored. The value of the noise floor estimates is calculated from the minimum of these buffer values and used for the entire breath cycle. The estimated noise floor can optionally be removed from the signal for each of the channels.

As an example, this is illustrated in Figure 5-7 which shows an envelope signal from one of the channels with a complete single breath including inspiration and expiration. Assuming that: the breath discriminator recognises the first part as three BC segments interleaving with one respiratory pause between each of them; in addition, during the segmentation, a segment is

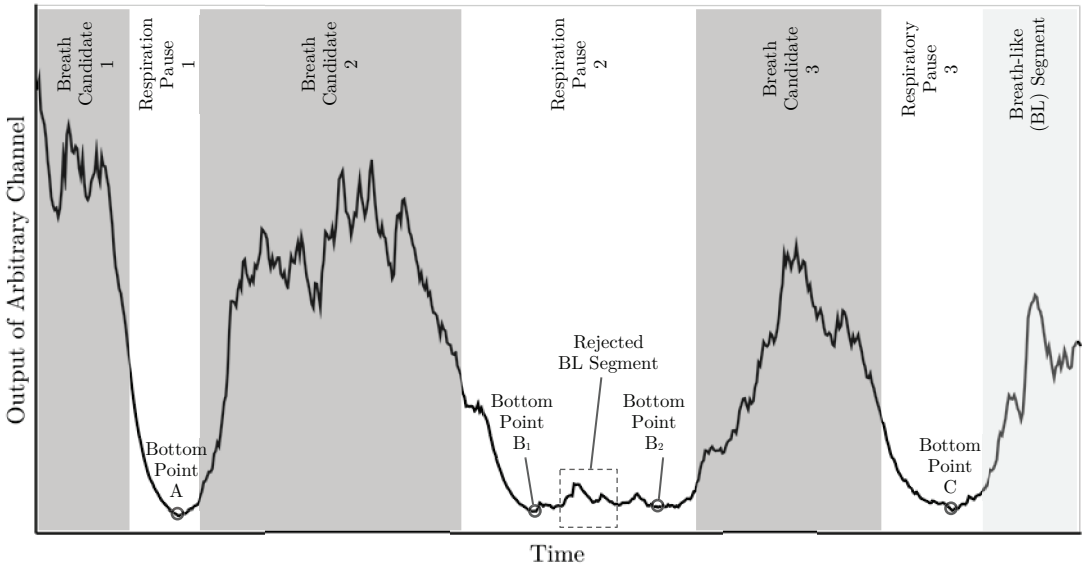

Figure 5-7: Noise floor estimation from envelope signals of an arbitrary channel over two respiratory cycles. Note that the marked segments are separated by preliminary start and end point instead of the verified ones.

accidentally identified as a BL segment but later rejected by the breath discriminator. Since the BL segment is converted to a part of respiratory pause, only the smaller bottom point is kept for noise floor evaluation. In this case, point 'B_1' is chosen over 'B_2'. When the segmentation algorithm tries to process the last BL segment, the noise floor of this particular channel is to be estimated as the minimum value of three bottom points 'A', 'B_1' and 'C'. Because the noise floor estimation only cares about the bottom points, this process does not rely on precise segmentation. As shown in Figure 5-7, preliminary start and end points defining the time for segmentation routine staying in valley seeking state and searching for bottom points are sufficient for the task.

Since the noise floors are known, the mean signal-to-noise ratio (SNR) of each channel can be evaluated by the channel average output within a breath candidate segment divided by its noise floor. The results of channels SNRs are used to determine the main channel for segmentation, as described herein.

5.3.3 Breath Discriminator – A Scoring System

At the end of the segmentation, when a BL segment is confirmed, the features of this segment will pass to the feature recognition block called breath discriminator. This increases the probability of assessing correctly whether a breath-like event is a true breath event. A set of rules to assess the BL segment against breath features is provided in the following sections. One of 4 tiers weighted scores (32, 16, 8 and 1) is given for a BL segment passing a corresponding condition. The rule for weighting conditions scores is dependent on their noise resilience. Those conditions which are easier to be affected by noise are weighted at lower score, and vice versa. The thresholds given below are dedicated for adults. For detecting children's breath, the settings need to be changed. The final gained score is converted to the confidence level for upgrading a BL segment to a BC segment. This will be discussed further in Section 5.3.4.

5.3.3.1 Breath Spectral Chart Mapping

Referring to Section 2.2.2 it is possible to compare the relationship between functions of energy of the envelope signals for identified BL segments within different channels. These are defined by spectral charts like in Figure 2-6, with predetermined boundaries. Hence if the ratios are within the boundaries the event is more likely to be a breath event than if it is outside the boundaries. These boundaries can also be different for inspiration and expiration. The method

to do this is described by the following equation. If the signal is sampled, within a segment, the power for one band after removing its noise floor is related to:

$$\mathrm{E(CH}_i)_{\mathrm{RMS}} \text{ of channel } i \propto \sum_{n=\text{BL segment start}}^{\text{BL segment end}} \mathrm{CH}_i(n) - \text{ Noise_ Floor}_i \tag{5.1}$$

where i is the channel number, n is the sample number

and $\mathrm{CH}_i(n) = \sqrt{\text{Power of channel}_i\ (n)} \approx \text{Envelope Signal}_i(n)$.

Figure 2-6 shows an example of the relationship between two channels of breath sound signal, 500 Hz to 600 Hz (Channel 2) and 650 Hz to 750 Hz (Channel 3) (both normalized against the 350 Hz to 450 Hz band (Channel 1), where the X and Y axes' values are given by:

$$X = \frac{\mathrm{E(CH_2)_{RMS}}}{\mathrm{E(CH_1)_{RMS}}} \approx \frac{\sum_{n=\text{BL segment start}}^{\text{BL segment end}} \mathrm{CH_2}(n) - \text{Noise_ Floor}_2}{\sum_{n=\text{BL segment start}}^{\text{BL segment end}} \mathrm{CH_1}(n) - \text{Noise_ Floor}_1} \tag{5.2}$$

$$Y = \frac{\mathrm{E(CH_3)_{RMS}}}{\mathrm{E(CH_1)_{RMS}}} \approx \frac{\sum_{n=\text{BL ssegment start}}^{\text{BL segment end}} \mathrm{CH_3}(n) - \text{Noise_ Floor}_3}{\sum_{n=\text{BL segment Start}}^{\text{BL segment End}} \mathrm{CH_1}(n) - \text{Noise_ Floor}_1} \tag{5.3}$$

The channels are compared using the ratios given in equations (5.2) and (5.3) above and plotted on the graph. The density of dots in Figure 2-6 is related to the probability for that signal with that particular spectral energy distribution to be caused by breathing. The figure shows that most breath signals have a pattern of distribution concentrated in a triangle area. To reduce the complexity of boundary line calculation, this can be defined as a right-angled trapezoid as shown in Figure 5-8 surrounding the concentration of dots that defines the version of most likely "breath related" distribution. The boundary region can be approximated by two linear equations with upper and lower limits on X and Y axes. By using Equations (5.2) and (5.3), any BL segments can be mapped to the spectral chart. If more than 3 channels system is used, a multi-dimension boundary region can be defined in the similar way.

The BL segments within the boundary region pass this condition. This condition can dramatically increase the sensitivity to detect absent of respiration and avoid the situation in which someone may not be breathing even though the algorithm might otherwise detect it as breathing. Therefore, the weighting for this condition is the top tier of all conditions, with a score of 32.

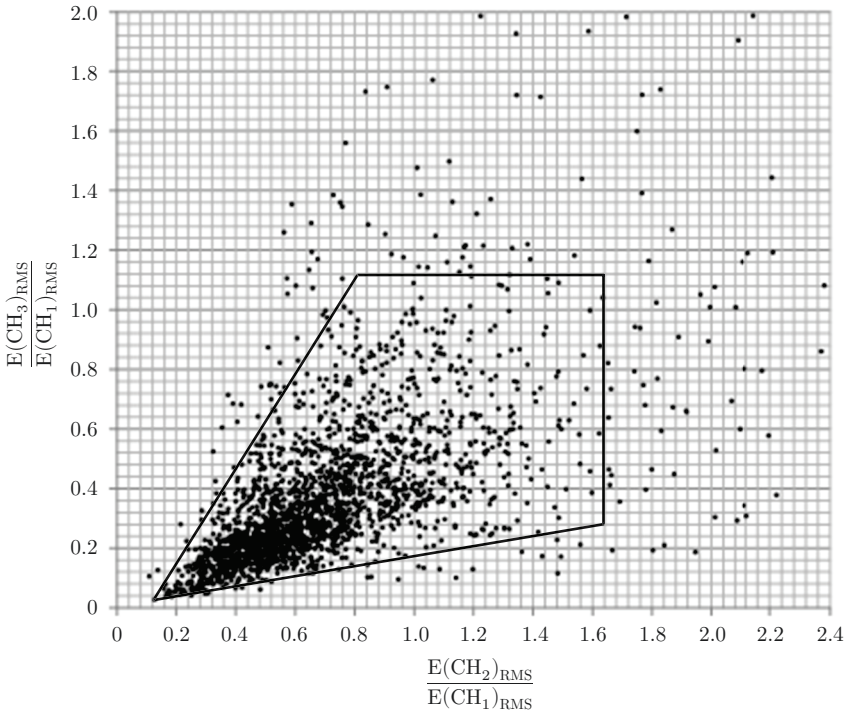

Figure 5-8: Limiting area of spectral energy distribution for a three-channel breath detection system

5.3.3.2 Breath-like (BL) Segment Length

According to the statistical distribution of breath lengths as shown in , the length of an adult's spontaneous respiratory phase (i.e. either inspiration or expiration) is in the range between 0.45 – 2.5 second. In most of cases, any segments with a length outside this range are noise. However, hyperventilation or voluntary deep breath can cause the segment outside this range. This condition is rated at 3rd tier, with score of 8.

5.3.3.3 Channels Trend

To further confirm if the BL segment is a true breath sound the shape of the breath sound envelope is assessed. In agreement with the mechanism of how breath sound is generated as discussed in Section 2.1, the system checks that some of the bands have particular and synchronous characteristics in the time domain. For instance it checks that the channels follow the same monotonicity trend between the beginning of a segment and a certain time determined by a different channel, as it is illustrated in the following expression for the secondary channels, referred to the apex position of the main channel:

$$\text{condition_trend}_i = \text{CH}_i[\text{At the apex point of CH}_{\text{main}}] - \text{CH}_i[\text{BL segment start}] > 0 \qquad (5.4)$$

where i is the channel number and $\mathrm{CH}_i \neq \mathrm{CH}_{\mathrm{main}}$.

If all the secondary channels satisfy this condition, a score of 1 is given to the BL segment.

5.3.3.4 Transient Condition

The third condition that can be evaluated to further ascertain whether an event is likely to be a breath event in the secondary channels that are close to each other in frequency, is the relationship between the envelope value of one of them at a time equal to the apex of the main channel with respect to the maximum value within the segment of the other channel in close proximity. This should exceed a predefined percentage, as shown in equation (5.5).

$$\mathrm{condition_transient}_i = \left(\frac{\mathrm{CH}_i[\text{At the apex point of } \mathrm{CH}_{\mathrm{main}}]}{\max(\mathrm{CH}_j[\text{BL segment start:BL segment end}])} \right) > 30\% \tag{5.5}$$

where i and j is the channel number, $i>j$, $\mathrm{CH}_i \neq \mathrm{CH}_{\mathrm{main}}$ and $\mathrm{CH}_j \neq \mathrm{CH}_{\mathrm{main}}$.

If all the secondary channels satisfy this condition, a score of 1 is given to the BL segment.

5.3.3.5 Time to Apex

The fourth condition is the relationship between the times at which the different channels reach their apex with respect to the segment. For example, the maximum time difference between different channels reaching their apex could be set to 5% of the segments' duration. The score for this condition is 1.

5.3.3.6 Release slope

The relationship between the negative slopes during the section of the signal corresponding to it fading out for different channels can also be used. This can be approximated by measuring the time distance between each channel's apex and the end of the BL segment. Similar to the condition above, the tolerance is ±5%. The score for this condition is 1.

5.3.3.7 Signal Onset

Another condition that can be considered is that changes of different channels instantaneous power, for example between start and apex, should correlate to the spectral energy distribution. For example the change of channel 2 could be limited to the range of 20% to 200% of $\frac{\mathrm{E}(\mathrm{CH}_2)_{\mathrm{RMS}}}{\mathrm{E}(\mathrm{CH}_1)_{\mathrm{RMS}}}$ (with noise floor compensated); and the change of channel 3 could be limited to the range of 0% to 250% of $\frac{\mathrm{E}(\mathrm{CH}_3)_{\mathrm{RMS}}}{\mathrm{E}(\mathrm{CH}_1)_{\mathrm{RMS}}}$ (with noise floor compensated), this is:

$$\text{condition_rise_CH}_i = \tag{5.6}$$

$$\frac{\max(\text{CH}_i[\text{BL segment start:BL segment end}])\text{-CH}_i[\text{BL segment start}]}{\max(\text{CH}_{\text{main}}[\text{BL segment start:BL segment end}])\text{-CH}_{\text{main}}[\text{BL segment start}]} > 20\% \times \frac{\text{E}(\text{CH}_\text{i})_{\text{RMS}}}{\text{E}(\text{CH}_{\text{main}})_{\text{RMS}}}$$

AND

$$\frac{\max(\text{CH}_i[\text{BL segment start:BL segment end}])\text{-CH}_i[\text{BL segment start}]}{\max(\text{CH}_{\text{main}}[\text{BL segment start:BL segment end}])\text{-CH}_{\text{main}}[\text{BL segment start}]} < 200\% \times \frac{\text{E}(\text{CH}_\text{i})_{\text{RMS}}}{\text{E}(\text{CH}_{\text{main}})_{\text{RMS}}}$$

where i is the channel number and $\text{CH}_i \neq \text{CH}_{\text{main}}$.

If all the secondary channels satisfy this condition, a score of 1 is given to the BL segment.

5.3.3.8 Dynamic Signal Mean Threshold

BL segments can also be processed in the time domain using a comparison with updated BC segments average power thresholds. In particular this allows artefacts which have characteristics similar to breath sound but lower average power- for example rubbing between clothes and sensor – to be discounted as breath candidates. As signal conditions change over time (such as when the subject falls asleep), a threshold, which discerns breath sounds from other breath-like noise, cannot be a fixed value. For example, high threshold values may not work when the subject is sleeping and low threshold values may reduce the discriminating capability of the system at other times. Hence, the comparison should be done with dynamically adapted thresholds.

An approach to determine and update the thresholds is to evaluate the power of the envelope signal for some previously successfully detected BC event (with or without the noise floor subtracted), which can for example be calculated as:

$$\text{P}_{i,\,\text{mean}} = \frac{\sum_{n=\text{BC segment start}}^{\text{BC segment end}} \text{CH}_i(n)}{\text{BC segment length}} \tag{5.7}$$

where i is the channel number and n is the sample number.

Then the average of the different powers can be calculated and updated by incorporating more recent events into the calculation. The average can be approximated by using an infinite impulse response (IIR) moving-average filter which provides a smoothed recent average $\text{P}_{i,\,\text{mean}}$ value.

Finally, in order to avoid the effect of strong interference skewing the filtered value (for example, speech may be detected as breath due to the airflow in trachea), A_i can also be capped at 150% of the previous calculated average value. The following transfer function (5.8) calculates the average power for every 20 BC segments, where prev_A_i is old value of Ai, and $P_{i,\,\text{mean}}(k)$ is mean power of the last detect BC segment.

$$A_i = \begin{cases} 1.5 \times \text{prev_}A_i\ , & \dfrac{P_{i,\,\text{mean}}}{\text{prev_}A_i} \geq 11 \\ \left(1 - \dfrac{1}{20}\right) \times \text{prev_}A_i + \dfrac{1}{20} \times P_{i,\,\text{mean}}\ , & \dfrac{P_{i,\,\text{mean}}}{\text{prev_}A_i} < 11 \end{cases} \tag{5.8}$$

For a breath signal segment to be selected as an actual breath the calculated power in one or a plurality of channels needs to be higher than, for example, a percentage of the mean dynamic power threshold. An empirical value would be 10% by assuming people's breath strength will not reduce below 10% of its average. The dynamic threshold is then calculated based on a fraction of the present value of A_i.

$$\text{condition_mean_power} = \frac{\sum_{n=\text{BL ssegment start}}^{\text{BL segment end}} CH_i(n)}{\text{BL segment length}} > 10\% \times A_i \tag{5.9}$$

This condition acts as the last defence of any breath-like noise being wrongly recognised as breath sound, though enforcing this condition can lead to a lower sensitive on breath detection when people starts shallow breathing suddenly, such as in the case of hypopnoea. Therefore, it is weighted at 2$^{\text{nd}}$ tier, with a score of 16.

5.3.4 Breath Discriminator – Interpretation

BL segments can be considered as breath events based on meeting a variable number of the criteria described above which will depend on the application. As described in Section 5.3.3, the maximum score that the breath discriminating system can give is 61. As listed in Table 5-2, 7 confidence levels can be proposed from the score generated by the breath discriminator. The implication of each level is also given. Segments with zero confidence level are not considered BC segment. The rule for the interpretation is based on how many scores are gained from each tier of conditions. Therefore, higher total score does not lead to a higher confidence level. For example, for a segment to be considered a BC segment, it cannot fail more than two lowest tier

BC Class	Confidence Level	Corresponding Score	Implication
First	6	61	The BL segment is definitely a breath event.
Second	5	59~60	The BL segment passes all the conditions, except one or two low tier condition(s). It could be a breath corrupted by some noise.
	4	53	The BL segment passes all the conditions, except its length is outside the range of normal breath. It could be a long deep breath or a hyperventilated breath.
	3	45	The BL segment passes all the conditions, except the average strength is below 10% of the mean power previous detect breath segment. It could be a shallow breath.
Third	2	43~44	The BL segment passes all the conditions, except the average strength is below 10% of the mean power previous detect breath segment and one or two low tier condition(s). It could be a shallow breath corrupted by some noise.
	1	29	The BL segment passes all the conditions, except the spectral energy distribution does like the one that breath sound has. It is either an abnormal breath event or a noise segment.
N/A (Not to be considered as BC segments)	0	Others not listed above	The BL segment is unlikely to be a breath event. Segments with 0 confidence level are classified as respiratory pause directly rather than BC segments.

Table 5-2: Description of the confidence levels given by the breath discriminator. All the segments with non-zero confidence levels are considered breath candidates (BC).

conditions (all with score of 1). Higher number of confidence level means higher possibility of the BL segment being a true breath event. For the purpose of explaining these further below, BC segments with similar confidence level are grouped into 3 classes. The top class BC segments are intended to be used by critical application such as apnoea detection for sudden death prevention which cannot afford any mistake. For the sake of clarity, this first class BC segment is referred to as breath segment, breath event, or respiratory phase in the rest of the text. The second and third classes are designed for less critical application

The characterisation of sensitivity and specificity on breath segment detection was carried out on four healthy volunteers in collaboration with Aguilar Pelaez [201]. In this study, the subjects were asked to breath regularly, but as shallowly and as slowly as they were able to for 15 minute periods. During these periods, tracheal sounds were recorded in synchronization with

Sensitivity of single respiratory phase detection	Specificity of single respiratory phase detection	Tidal volume range
98.1%	100%	0.20 – 0.25 L
95.8%	100%	0.15 – 0.20 L
89.1%	100%	0.10 – 0.15 L
83.7%	100%	0.05 – 0.10 L
71.9%	100%	0 – 0.05 L

Table 5-3: Sensitivity and specificity of single respiratory phase detection with confidence level 5 with respect to various ranges of tidal volumes [201]

lung volumes which were measured with a Wright Haloscale MX respirometer. In this initial study 907 BC breath segments were given the top confidence scoring, out of a total of 935 true respiratory phases. There was no false breath detections. Overall the sensitivity and specificity of the breath detection was 97%. Table 5-3 shows the detailed performance of the single respiratory phase detection for different ranges of tidal volumes.

If the detection system takes the other confidence classes of BC segments into account, the sensitivity of the breath detection can be improved, especially for the breaths with low tidal volumes. This is however at the risk of trading off specificity, i.e. having more false breath detections. Depending on the sensitivity and specificity requirements of the targeted application, second and third class BC segments can be either directly upgraded to breath events or be subject to further extra scrutiny before making a final decision. These extra evaluation steps are demonstrated in Section 5.4.

5.3.5 Real Time Implementation

The algorithm described above was implemented in C language and embedded into a Visual C++ program as shown in Figure 5-9. The program is able to save the sound data sent from the wireless sensor (described Section 3.5) as well as monitor on-line. sent. Additionally, it is also able to show the instantaneous status of the breath segment (BC segment) detection and an alarm for no breath detection over 25 seconds. A green light is displayed on the screen when a breath segment is detected and a red light if the alarm is raised. The respiratory rate based on half amount of detected breath events (assuming every two breath detections represents one respiratory cycle) is also calculated, recorded and refreshed every 30 seconds. Users are able to

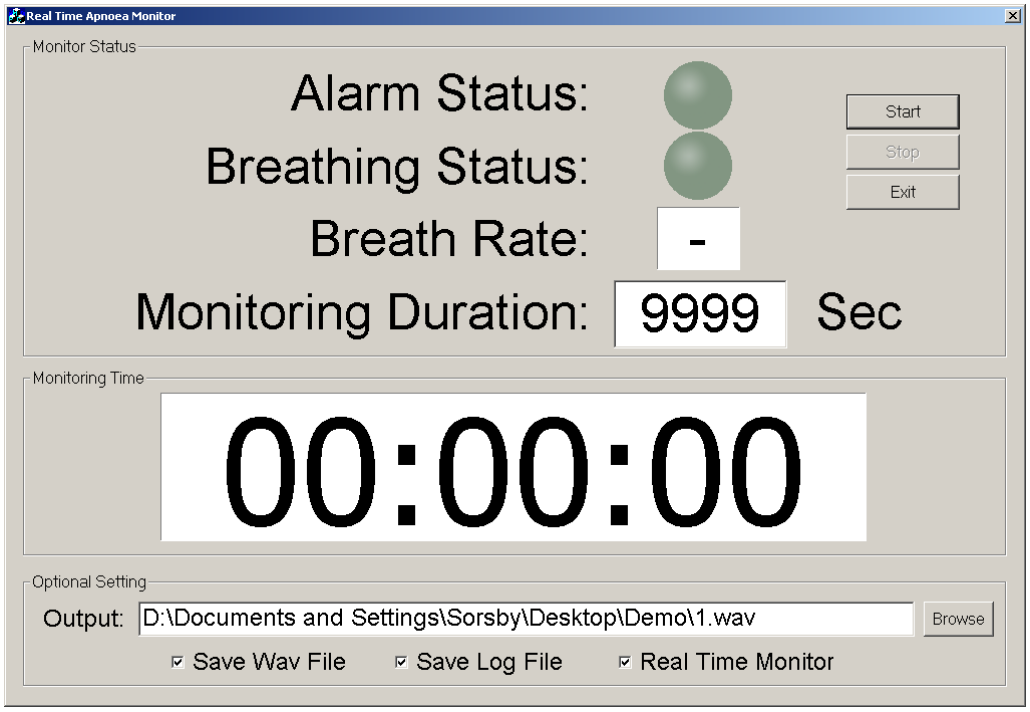

Figure 5-9: Graphic user interface of the real time breath and apnoea detection algorithm

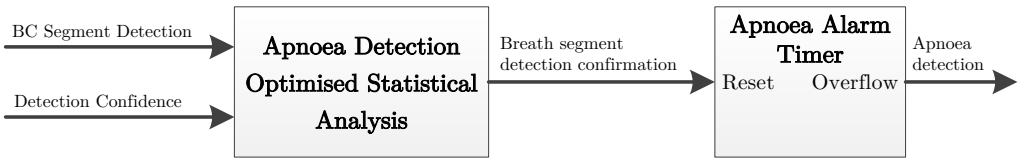

Figure 5-10: Overview of apnoea detection mechanism.

adjust the monitoring duration and choose the location of the audio file. During monitoring, every single detected apnoea event is recorded in a log file.

5.4 Apnoea Detection Algorithm

5.4.1 Overview of Apnoea Detection

One of the important applications of automatic breath detection is identification of apnoea (cessation of breathing) within a certain period of time (as shown in Figure 5-10). Depending on the final purpose of the apnoea detection, the requirements of sensitivity and specificity may vary. For example, detecting apnoea for those people under the risk of sudden death requires extremely high specificity on breath detection, or in other words, sensitivity on apnoea detection. The reason is that if there is any false breath detection during a critical true apnoea event, the

event would be 'silenced', leading to a delay in life saving interventions. However, false detections are equally important because they can make users panic for no reason and lead to loss of compliance. For the case of sleep apnoea detection, a few isolated false breath detections are not important because the disease is diagnosed base on averages over a long period of time, as well as thresholds which allow relatively wide range of variations.[263]. For the sake of clarity, since the following sections are related to apnoea detection, the sensitivity and specificity are defined as:

$$\text{Sensitivity}_{\text{apnoea}} = \frac{\text{True positive apnoeas}}{\text{True positive apnoeas} + \text{False negative apnoeas}} \tag{5.10}$$

$$\text{Specificity}_{\text{apnoea}} = \frac{\text{False positive apnoeas}}{\text{True negative apnoea detection} + \text{Number of false positive apnoea detection}} \tag{5.11}$$

A series of methods described below were created to maximise the apnoea detection specificity without compromising the sensitivity.

5.4.2 Specificity Improvement based on Respiratory Rhythm Recognition and Prediction

Based on the characteristics of respiration, if an inspiration is detected from the sound signals, an expiration must have occurred shortly after, and vice versa. Section 2.2.3.5 has already shown how regular respiration trends follow a pattern in which the time distance between inspiration and expiration is generally shorter than the other way around. This can be used to carry out time analysis on each pair of breath segments in order to increase or decrease the confidence on whether a certain class BC segment can be considered as a breath event.

For second class BC segments to be upgraded, they have to be spread over time with a similar rhythm as the adjacent detected breath segments (first class BC segments). For the sake of simplicity, from now on the term "BC segment" will refer to first or second class BC segments by default, unless otherwise stated.

For illustration, Figure 5-11 shows an example of a series of BC segments being detected over time. Each box represents detection of a BC segment in time. Its height corresponds to the

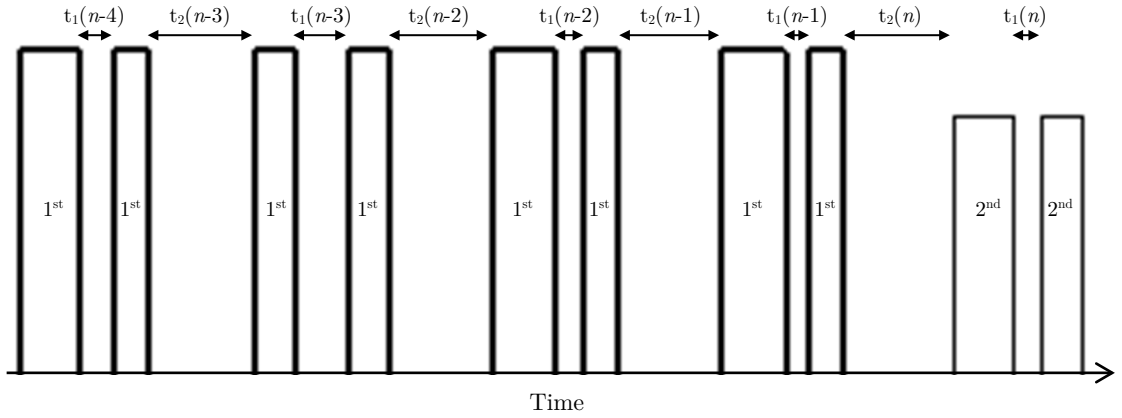

Figure 5-11: Example of how to upgrade a pair of lower level confidence class breath candidate (BC) segments by recognising respiratory rhythm. Each box represents detection of a BC segment along the time axis.

detection confidence level. The segments with shorter time distance t_1 are grouped as pairs, representing a full respiratory cycle:

$$t_1(n)=\mathrm{Seg}_{\mathrm{BC_Latter,\ Start}}(n)-\mathrm{Seg}_{\mathrm{BC_Former,\ End}}(n) \tag{5.12}$$

where n is the index of the segment pairs, $\mathrm{Seg}_{\mathrm{BC_Latter,\ Start}}(n)$ represents the start time of the last detected BC segment $\mathrm{Seg}_{\mathrm{BC_Latter}}(n)$, $\mathrm{Seg}_{\mathrm{BC_Former,\ End}}(n)$ represents the end time of the last second detected BC segment $\mathrm{Seg}_{\mathrm{BC_Former}}(n)$.

The longer time distance t_2 between each pair of BC segments is given as:

$$t_2(n)=\mathrm{Seg}_{\mathrm{BC_Former,\ Start}}(n)-\mathrm{Seg}_{\mathrm{BC_Latter,\ End}}(n-1) \tag{5.13}$$

In this example, only the last pair of BC segments is classified as second class and the rest are all first class. As shown in the figure, the last $t_1(n)$ and $t_2(n)$ match the expectation of respiratory rhythm as determined by the previous t_1 and t_2 values. In such case, it is considered safe to upgrade any second class BC segments in the last pair to first class (i.e. breath segments).

In order to implement this method, two predictors were created to estimate the range of respiratory pause times for future breath segments.. One was based on the average value of the time distance from the same group (either t_1 or t_2):

$$t_{i,\ \mathrm{mean}}(n)=\left(1-\frac{1}{m}\right)\times t_{i,\ \mathrm{mean}}(n-1)+\frac{1}{m}\times t_i(n\text{-}1) \tag{5.14}$$

where t_i is either t_1 or t_2, n is the index of the paired BC segments and m is the number of pairs to be averaged.

The other one is based on the average ratio between t_1 and t_2 as calculated below.

$$R_{mean}(n) = \left(1 - \frac{1}{m}\right) \times R_{mean}(n-1) + \frac{1}{m} \times \frac{t_2}{t_1} \tag{5.15}$$

The predicted respiratory pause times t_1' and t_2' are calculated as follows:

$$t_1'(n) = \begin{cases} t_{1,\,mean}(n) \\ \quad \text{OR} \\ \dfrac{t_2(n)}{R_{mean}(n)} \end{cases} \tag{5.16}$$

$$t_2'(n) = \begin{cases} \quad t_{2,\,mean}(n) \\ \quad \text{OR} \\ R_{mean}(n) \times t_1(n-1) \end{cases} \tag{5.17}$$

By introducing an error tolerance coefficient p, the range of pause time prediction is given by:

$$t'_{i,\,range}(n) = \pm\, p_i \times t_i' \tag{5.18}$$

where $t'_{1,\,range}$ is the range of predicted inner-pair separation time;

and $t'_{2,\,range}$ is the range of predicted pair-to-pair separation time

If $t_i(n)$ of the last detected BC segment pair is within in the range of $t'_{i,\,range}(n)$, the pair can be upgraded to first class. Empirically the chosen values for p_1 is 10%, p_2 is 50% and m is 5.

$$\text{IF} \quad t_i(n) \in t'_{i,\,range}(n) \quad \text{THEN} \quad \text{BC segments in last pair} \Rightarrow \text{Breath Segments} \tag{5.19}$$

For some apnoea detection applications such as sudden death prevention, a failsafe recovery mechanism can be added to prevent mistaken upgrade caused by failure of the respiratory rhythm prediction. If any first class breath segment does not meet the timing expectation this indicates that the prediction may not be trustworthy. In such a case the moving average predictors can be reset and the upgrade scheme temporarily deactivated until the predictors are working again and a set of correct timing conditions is found from m consecutive pairs of breath segments.

5.4.2.1 Respiratory Phase Classification

The prediction of respiratory rhythm can be used to classify breath segments (including both original and the upgraded ones) into inspiration or expiration events by means of the predicted time separations t'_1 and t'_2. The paired breath segments can be classified into three types using a recursive routine as shown in Figure 5-12: inspiration, expiration and standalone. This will be further described below.

The respiratory phase classification routine runs independently from the upgrade scheme described above and starts with a training session which identifies the respiratory rhythm from m consecutive pair of breath segments. Once the rhythm prediction model is trained, the algorithm uses the timing prediction $t'_{1,\,\mathrm{range}}$ and $t'_{2,\,\mathrm{range}}$ calculated from Equation (5.18) to classify the breath segments. Three situations can happen:

a) Full respiratory cycle

A full respiratory cycle can be confirmed if the time distance $t_1(n)$ between the last detected pair $\mathrm{BSeg}_{\mathrm{Former}}(n)$ and $\mathrm{BSeg}_{\mathrm{Latter}}(n)$ is shorter than the maximum of its predicted range$t'_{1,\,\mathrm{range}}$; and the time separation $t_2(n)$ between the pair to the last classified segment $\mathrm{BSeg}_{\mathrm{Latter}}(n\text{-}1)$ is longer than the minimum of its predicted range $t'_{2,\,\mathrm{range}}$, in addition to the condition $t_1(n) < t_2(n)$. The timing requirement for $t_1(n)$ and $t_2(n)$ is not as strict as the one presented in the upgrade scheme for the second class BC segments because the classification is not as critical as the breath detection process, since any wrong classification on the breath segments with full detection confidence will not lead to failure of the apnoea detection.

Once a pair of breath segments is classified as a full respiratory cycle, the routine restarts with a new pair of breath segments. The recorded t1 and t2 from the respiratory cycle are used to calculate the new prediction.

b) Shivering breath

As discussed in Section 0, some breath segments can present shivering characteristics. Because the segmentation algorithm analyses the breath signal sample by sample, it is difficult for it to identify and merge the shivering breath segments. Although the segmentation algorithm already makes some provisions to deal with this, it only merges segments

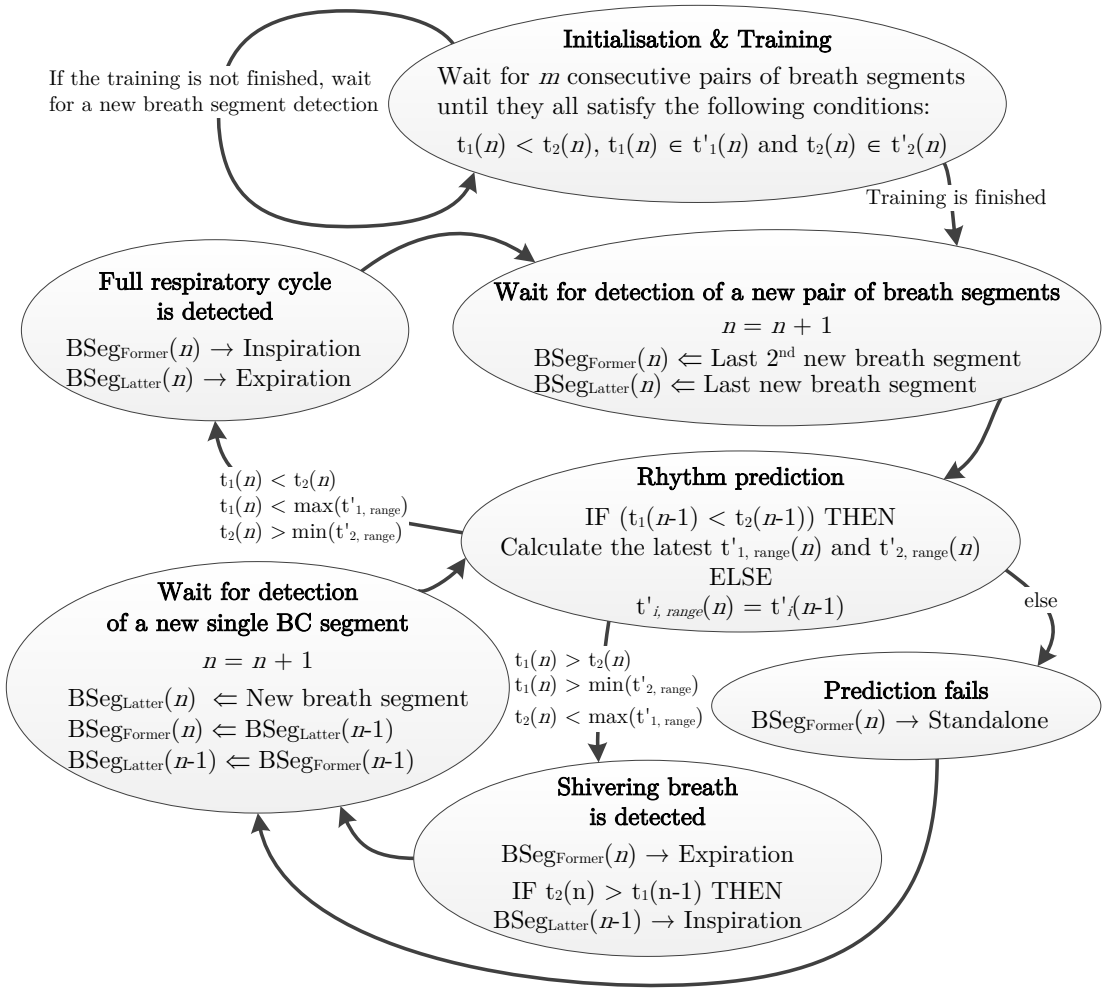

Figure 5-12: Recursive routine for respiratory phase classification.

with 160 ms separation (see Section 5.3.2.7). As shown in Figure 2-11, the actual shivering period can however be longer. Since the respiratory phase classification routine is event-based, it is easier to design a mechanism to identify those shivering breaths by examining the time separations.

When a shivering breath happens, more than two consecutive breath segments with a time separation similar to the expected t_1' and shorter than the expected t_2' can be observed. The first two breath segments are classified as a full respiratory cycle according to the first case (a). Then the shivering breath fragment $BSeg_{Former}(n)$ is paired with another new detected breath segment $BSeg_{Latter}(n)$ which belongs to another respiratory cycle. Therefore, the inner-pair separation time t_1 is expected to be longer than both the pair-to-pair separation time t_2 and the minimum of the predicted range of pair-to-pair separation time $t'_{2,\,range}$. The pair-to-pair separation time t_2 is expected to be either the pause time between inspiration

and expiration, or the shivering period during the expiratory event. In either case, t_2 is expected to be smaller than the maximum of its predicted range $t'_{1,\,\text{range}}$ for the pause time between inspiration and expiration.

Once the former breath segment $\text{BSeg}_{\text{Former}}(n)$ in the last detected pair is identified as a shivering breath fragment, it splits from the current pair and is classified as expiration, effectively joining the last identified respiratory cycle. Inside the respiratory cycle, the two with shorter time separation join together as a whole breath segment by remarking the middle breath segment $\text{BSeg}_{\text{Latter}}(n\text{-}1)$. The leftover breath segment $\text{BSeg}_{\text{Latter}}(n)$ joins the next detected breath segment as a new pair for the next round of respiratory phase classification. The timing information t_1 and t_2 calculated from this round is not taken into account for the next round.

c) Failed Prediction

For other not listed cases, the occurrence of breath segments $\text{BSeg}_{\text{Former}}(n)$ and $\text{BSeg}_{\text{Latter}}(n)$ in the last detected pair do not match the respiratory rhythm prediction. Both breath segments seem to be isolated from each other. In this case, the classifier assumes there is at least one missing breath segment detection which could link to the former breath segment $\text{BSeg}_{\text{Former}}(n)$ forming a respiratory cycle. As the phase of the missing breath segment is not known, the former breath segment $\text{BSeg}_{\text{Former}}(n)$ is classified as "standalone". And again, the leftover breath segment $\text{BSeg}_{\text{Latter}}(n)$ is grouped with the next available breath segment as a pair for further classification.

Although the assumption of missing breath segment could be wrong, as the failure of respiratory rhythm prediction could be simply caused by a change in respiratory rate being outside the prediction tolerance p_i defined in Equation (5.19), the timing information t_1 and t_2 obtained under this case are used to train the predictor providing $t_1 < t_2$.

The proposed approach to improve the apnoea detection specificity has been carefully designed for a critical vital signs monitoring application which cannot afford any degradation on apnoea detection sensitivity. For other applications such as screening of sleep apnoea, missing some apnoeas could be acceptable. In this case, the apnoea detection specificity can be increased further by considering third class BC segments for upgrade. Another possible modification that could be done is integrating the BC segments upgrade scheme into the respiratory phase

classification process. This can be carried out classifying the suitable set of BC segments first, and then upgrading the BC segments that are not classified as standalone events.

Although this approach works for most of spontaneous respiration, it may not work as expected when inspiration and expiration are evenly distributed over time. For example, athletes may voluntarily delay their expiration during exercises in order to maximise their performance. Another possibility of performance degradation could be that the detection confidence may reduce due to intensive amount of noise, leading to more breath events being identified as lower class BC segments. Under those circumstances, the respiratory rhythm, which is based on the time location of detected breath segments may not be clear or easy to estimate.

5.4.2.2 Lung Volume Estimation

A potential solution to solve the problem discussed above is to implement another approach to upgrade and classify the BC segments without using temporal features of respiration. This could be done by observing the record of tidal volume during inspiration and expiration. Since the lung volume is relatively fixed, the amount of air inhaled should match the volume exhaled over time. It has been shown [152][1] that it is potentially possible to obtain tidal volumes from the breath sound with just two calibration parameters α and β as shown in Equation (5.20).

$$\text{Volume}=\alpha \times \text{E}(\text{CH}_{\text{600Hz-700Hz}})_{\text{RMS}} + \beta \tag{5.20}$$

where $\text{E}(\text{CH}_{\text{600Hz_700Hz}})_{\text{RMS}}$ is the energy of envelope signal generated from frequency range between 600 Hz to 700 Hz over the duration of breath event.

This preliminary work shows that the error of the tidal volume estimation could be as little as ±38 ml and up to ±123 ml. This is very close to the gold standard ±50 ml as shown in Section 0. The study above also showed that the calibration parameters α and β are not the same for inspiration and expiration. In particular, the value of β is generally very small and does not vary too much from inspiration to expiration. Therefore, for inspiration and expiration within a same respiratory cycle, the following relation can be found if the difference β between inspiration and expiration can be neglected.

$$\alpha_{\text{in}} \times \text{E}(\text{CH}_{\text{600Hz-700Hz}})_{\text{RMS, in}} + \beta_{\text{in}} = \alpha_{\text{out}} \times \text{E}(\text{CH}_{\text{600Hz-700Hz}})_{\text{RMS, out}} + \beta_{\text{out}} \tag{5.21}$$

[1] By myself in collaboration with I de la Cruz and E Rodriguez-Villegas

$$\alpha_{\text{in}} \times \text{E}(\text{CH}_{\text{600Hz-700Hz}})_{\text{RMS, in}} - \alpha_{\text{out}} \times \text{E}(\text{CH}_{\text{600Hz-700Hz}})_{\text{RMS, out}} = \beta_{\text{out}} - \beta_{\text{in}} \approx 0 \tag{5.22}$$

$$\frac{\text{E}(\text{CH}_{\text{600Hz-700Hz}})_{\text{RMS, in}}}{\text{E}(\text{CH}_{\text{600Hz-700Hz}})_{\text{RMS, out}}} = \frac{\alpha_{\text{out}}}{\alpha_{\text{in}}} = \text{constant} \tag{5.23}$$

Although De la Cruz's work [183] shows that the calibration parameters α_{in} and α_{out} are dependent on the person's position, they remained constant if the person was in one fixed position and so does the ratio between them. Based on the assumption of conservation of tidal volumes, it is possible to recognise a respiratory cycle from the frequency band limited power ratio between two adjacent BC segments.

5.4.3 Classification of the Respiration Relative Sleep Disorder Event

In addition to apnoea detection, a set of rules has been developed to differentiate the type of respiratory events from the sleep data, which are useful for diagnostic of a variety of conditions of sleep disorder conditions. In term of sleep apnoea classification, it is implemented by examining the non-top confidence level BC segments (the upgraded ones are not included) distributing over the apnoea event reported by the apnoea detection algorithm. For detecting hyperventilation during sleeping, it can be done by analysis of respiratory rate and the mean power of breath segments.

For illustration, Figure 5-13 shows 4 typical types of sleep disorder conditions which are respiration related. In those four figures, small green vertical boxes represent breath segments (with highest detection confidence) over time. Each red vertical box with a green cap represents the BC segment that does meet the mean power requirement (level 3 confidence). Other red only vertical boxes indicate the detection of BC segments with confidence level other than 1 and 3. The pink trace shows the mean power of BL segments (including the BC segments). Each abnormal condition shown in Figure 5-13 (a) (b) (c) contains a section of time duration does not have any breath segment detection. Those time durations are referred as sleep apnoea event in the following, which is generally being in 10 second epoches. The following sections declare some rules using above information to classify the sleep apnoea/disorder events.

a) Hypopnoea

A certain sequence of BC segments with reduced power for an extended period of time may correspond to a hypopnoea event. For a sleep apnoea event to be considered as a hypopnoea event, there must be more than one level 3 BC segment in every epoch.

b) Central Apnoea

A typical central apnoea event appears like a clean and silence section. This can be defined as no additional BC segment detected during a sleep apnoea event last for at least one epoch. Any BC segments detected at the beginning or the end of the sleep apnoea event is not considered as a part of central apnoea event.

c) Obstructive or Mixed Apnoea

In one implementation obstructive or mixed apnoea can be differentiated from central apnoea by evaluating increment in mean power (the pink trace in Figure 5-13) between a few last candidate events before the sleep apnoea event, as well as a few breath segments after the event. This is an indication of breath difficulty due to the blockage of respiratory airway. Another case of showing breath difficulty is that a subsequent low confidence level (≠3) BC segments happen during the sleep apnoea event. All of the events containing above conditions are classified as obstructive/mixed apnoea.

d) Hyperventilation

Although there will not be an apnoea detection as a result of hyperventilation, this kind of event is still very distinguishable. The conditions for identifying a hyperventilation event can be done by examining breath segments' mean power as well as their in instantaneous breath rate. As mentioned in Section 1.3.1.1, hyperventilation during sleeping can lead to blood vessel constriction because of blood alkalization, resulting insufficient of oxygen delivery over the body, making the respiratory system to make deeper and faster breathing. This can be revealed from the mean power of the breath segments. Together with the time separation of between every two, if it is larger than the highest normal respiratory rate at rest which is 20 per minute, the time portion when this constantly happens can be recognised as hyperventilation. In the example shown in Figure 5-13(d), the instantaneous respiratory rate calculated in this method results a rate as high as 24 per minute and the mean power of each breath is much higher than other examples shown in the same figure.

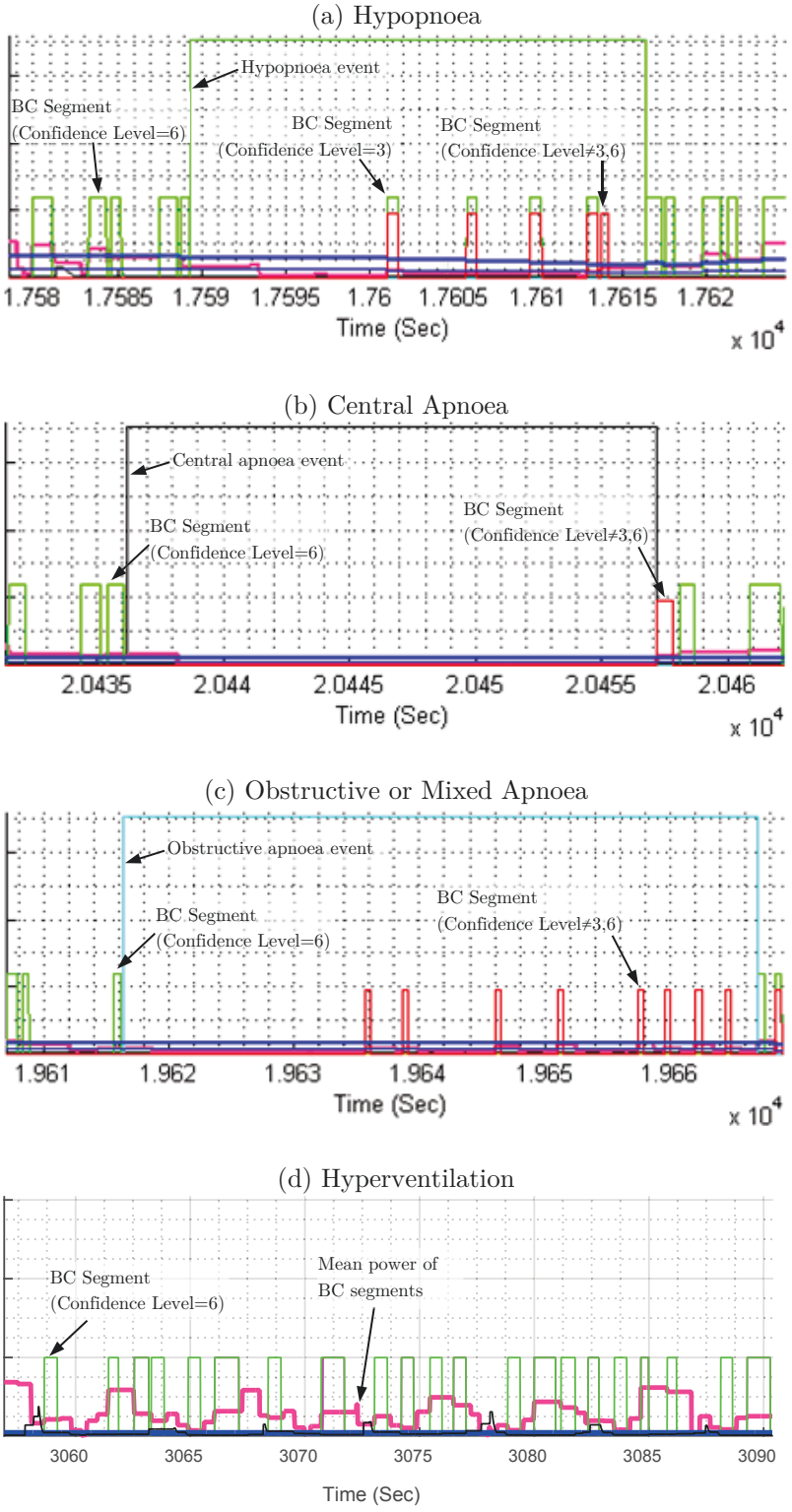

Figure 5-13: Examples of output of BC segments obtained from 4 different kinds of sleep apnoea/disorder events: hypopnoea, central apnoea, obstructive/mixed apnoea and hyperventilation.

Chapter 6
Clinical Test on Sleep Apnoea and Hypopnoea Studies

6.1 Introduction

As already covered in Chapter 1, existing techniques for monitoring respiration are not satisfactory. There is no solution that can reliably detect breathing and its absence, and which are small and discrete enough for long-term use in domestic and residential settings, as well as low-intensity clinical care areas. The original aim of this work was to develop a device that could help to prevent sudden unexpected death in epilepsy (SUDEP) in residential settings. However, it is not easy to test the proposed respiratory monitoring system in such critical setting. Because of this, a low-intensity clinical application, sleep apnoea diagnosis, became the first choice for validation of the algorithm created in this thesis.

Sleep apnoea diagnosis has always been a problem due to the lack of sleep laboratories, not enough sleep specialist time to sift through a range of patient's physiological sleep signals and the cost associated to hospital stays. The indirect costs of under-diagnosed adult sleep apnoea patients have been estimated to be US$ 80 billion per year [264]. The clinical study, carried out throughout the course of this thesis, aimed not only to evaluate performance but also to identify some specific issues that might be relevant to critical life threatening applications like SUDEP prevention. The results presented in this chapter have been published on the British Medical Journal Open [155].

Figure 6-1: Wearable respiratory activity monitor worn by myself [155] and a volunteer.

6.2 Methods

6.2.1 Device

The wireless sensor used in the study has been discussed in Chapter 3. The sensor was wireless, measured 3.74×2.4×2.1 cm, weighed 17 g, and was fixed to the skin on the neck with hydrocolloid colostomy adhesive patches of approximately 4 cm diameter (Boots, UK). The code name for this early version of prototype was WADD, which is short for Wearable Apnoea Detection Device.

As disclosed in Section 2.1.3, the preferred location was over the trachea, halfway between the lower margin of the thyroid cartilage and the suprasternal notch (Figure 6-1, left). If the skin in that location was loose, as was common in participants above 40 years of age, the device was placed anterolaterally, anterior to the sternomastoid muscle (Figure 6-1, right). The device was left in place overnight, for approximately 14 h.

6.2.2 Participants

The study was conducted in a sleep study room of the National Hospital for Neurology and Neurosurgery (UK). 20 healthy controls and 10 patients were studied, and recruited as they were sequentially admitted for diagnostic monitoring of sleep-related disorders of breathing, because these were likely to have spontaneous apnoea events. The patients and controls were not matched. Patients also had a variety of neurological conditions, including epilepsy, dementia, neuropathy and motor neuron disease. The reasons to recruit patients who had been referred for

diagnosis of possible sleep-related disorders of breathing, as opposed to those who had been already diagnosed, were twofold. First, the purpose of this study was not to evaluate the WADD for sleep apnoea diagnosis, but rather to evaluate its ability to detect individual events during controlled conditions to assess the robustness to artefact rejection and also during spontaneous sleep. Good performance on individual event identification would, however, be expected to translate in a good performance in the context of the different clinical applications. Second, non-diagnosed patients were recruited because studying diagnosed individuals would have involved either delay or interruption of their treatment. The decision on the number of patients was based on obtaining a large enough number of events that would lead to the study goals of 95% confidence interval (CI) for sensitivity and specificity values (based on clinical experience on the minimum number of apnoea events that would be expected per subject referred for sleep apnoea diagnosis, per night). A larger number of controls were included to be able to assess specificity among those who were most likely to be disease free, and also in the presence of artefacts. The patient group comprised 2 females and 8 males with: a median age of 44.5 years of age (range 25–82); a median weight of 74 kg (range 41–187); a median height of 177 cm (range 160–188); a median body mass index (BMI) of 23 kg/cm2 (range 17–61); and a median neck circumference of 40 cm (range 30–43). The control group comprised 3 females and 17 males with: a median age of 33.5 years of age (range 23–63); a median weight of 81.5 kg (range 60–120); a median height of 176 cm (range 145–185); a median BMI of 26.5 kg/cm2 (range 20–36); and a median neck circumference of 38 cm (range 34–48). Overall, 40% of the participants were overweight and 24% were obese.

6.2.3 Procedure

All participants also had simultaneous, clinically standard respiratory monitoring comprising: finger oximetry; oronasal airflow sensors; thoracic and abdominal expansion bands; and ECG; using the SOMNO polysomnography system (SOMNOscreen RC kombi. SOMNO Medics, Germany) (Figure 1-9). Additionally, to further facilitate expert interpretation of polysomnography data, a second pulse oximeter (Pulsox-300i, Konica Minolta, Japan) was attached to the free hand. After attachment of the WADD and the SOMNOmedics polysomnography system, control participants participated in a series of exercises, comprising:

T1. Normal breathing for 5 min.

T2. Shallow breathing for 5 min.

T3. Normal breathing for 45 s alternating with 15 s instructed breath holds for 5 min.

T4. Normal breathing for 30 s alternating with 30 s instructed breath holds for 10 min.

T5. As in T4 but with loud music in the background.

T6. Normal breathing while walking for 5 min.

T7. Normal breathing for 30 s alternating with 30 s instructed breath holds while lying prone for 5 min.

These exercises were designed to be representative of the worst case of artefact situations affecting the WADD following previous, very exhaustive, laboratory-based research and testing. Following the exercises participants were allowed to prepare for sleep and were left undisturbed overnight.

6.3 Data Analysis

The breathing exercises data were analysed by the automated WADD software and the automated SOMNO software. Instructed apnoeas were considered to be the 'true events'. The last 6 h of sleep were blindly analysed by the automated WADD software, the automated SOMNO software and by the experienced clinician who reviewed the raw signals from all SOMNO sensors, and had no knowledge about how WADD had been designed or worked. The reason to evaluate the last 6 h of sleep was to try to keep the same amount of sleep data in as many participants as possible in order to prevent biasing of the results. The pulse oximeter was also used by the clinician to support the diagnostic decisions and also individual event classification mostly in those cases in which the signals from the other SOMNO pulse oximeter were corrupted by artefacts. After the separate classification of WADD and SOMNO data, a further investigator compared the results.

The breathing exercises data were analysed in 15 s epochs because this was the shortest duration of an instructed apnoea. The sleep data were analysed in 10 s epochs.

Two assessments were carried out of the sleep data. In the first assessment there was no preassumption of a gold standard, and the three systems (WADD, SOMNO and expert marker) were put under test and treated indistinctively. An epoch would be classified as true positive

apnoea or true positive hypopnoea if at least two out of the three systems concurred on the classification. In the second assessment the final classification of these epochs would be that of the expert marker or, in other words, the expert marker was considered to be the gold standard deciding, and the performances of SOMNO and WADD system were evaluated. The SOMNO was evaluated as well as the WADD, as there is little or no quantitative information about the accuracy of automated polysomnography systems.

In both assessments epochs could be classified as:

A. True positive apnoea (cessation of breathing signal, with correspondent absence of respiratory airflow).
B. True positive hypopnoea (over 50% reduction in oronasal signal and in thoracoabdominal movement together with over 2% decrease in oxygen saturation).
C. False positive hypopnoea (if a system had classified a breathing epoch as a hypopnoea).
D. False positive apnoea (if a system had classified a breathing epoch as an apnoea).
E. False classification apnoea as hypopnoea (if a system had classified an apnoea epoch as hypopnoea).
F. False classification hypopnoea as apnoea (if a system had classified hypopnoea as apnoea).
G. False negative apnoea (if a system had classified an apnoea as breathing).
H. False negative hypopnoea (if a system had classified a hypopnoea as breathing).

The breathing exercises data were analysed in the same way, but the instructed apnoeas and breathing sections were considered the absolute truth and hence there was no independent expert review.

The performances of the three systems were evaluated using the following metrics:

$$\text{Sensitivity}=\frac{\text{TP}}{\text{TP+FN}} \tag{6.1}$$

$$\text{Specificity}=\frac{\text{FP}}{\text{TN+FP}} \tag{6.2}$$

where TP=true positive, TN=true negative, FP=false positive, FN=false negative.

The analysis was carried out assuming that all apnoea events were independent, since it was observed that the characteristics of the breathing signal changed as much within the same participant (depending on timing, position, external artefacts, etc), as between different participants. This was further verified by taking three random 10 min sections of the sensed breathing signals in the 30 different participants and obtaining the different correlation coefficients (2700 in total). The maximum correlation coefficient obtained from signals within the same participant was 0.05. The maximum correlation coefficient obtained from different subjects was 0.067.

For each of the two assessments (ie, not presuming a gold standard and considering the expert to be the gold standard), two different analyses were carried out. First, only apnoeas were considered to be true positives. Hence any hypopnoea would be regarded as breathing (true negative); False classification of hypopnoeas as apnoeas were reclassified as false positives; and false classification of apnoeas as hypopnoeas were re-classified as false negatives. Second, apnoeas and hypopnoeas were considered indistinctively, and hence true events of both varieties would be also considered together.

These two analyses were carried out as they would be relevant to different clinical scenarios. For example, high sensitivity for detecting apnoea would be crucial for identification of sudden apnoea if monitoring those with epilepsy; whereas for diagnosis of sleep-related breathing dysfunction, which generally relies on the Apnoea Hypopnoea Index (AHI), the differentiation between apnoea and hypopnoea might be clinically less important.

6.4 Results

6.4.1 Breathing Exercises Data

Data were available in 3956 15s epochs for the controls performing the breathing exercises (132 in total). Table 6-1 summarises the performance of WADD and SOMNO in the seven breathing exercises. Figure 6-2 illustrates examples of the signals obtained from the different sensors. Table 6-1 is divided in three parts. Part (A) and (B) quantify performance considering different scenarios for wrongly classified hypopnoeas. Although the real events were apnoeas, both systems had the ability to indicate hypopnoeas too. This resulted in some real apnoea and breathing epochs being wrongly marked as hypopnoeas. To account for these, Table 6-1(A)

shows the sensitivity and specificity when only apnoeas are considered as events (ie, hypopnoeas would be regarded as breathing). In part (B) of the table hypopnoeas and apnoeas are indistinctively considered. Part (C) illustrates the total number of epochs that fall into a specific classification for both systems. The combined sensitivity and specificity for all the exercises across all the participants for the WADD was 97.7% and 99.6% (considering hypopnoeas as breathing); or 99.2% and 99.5% (considering hypopnoeas as events). With the same criteria the sensitivity and specificity for the SOMNO was only 37.8% sensitivity, 96.5% specificity; or 62.8% sensitivity, 90.5% specificity.

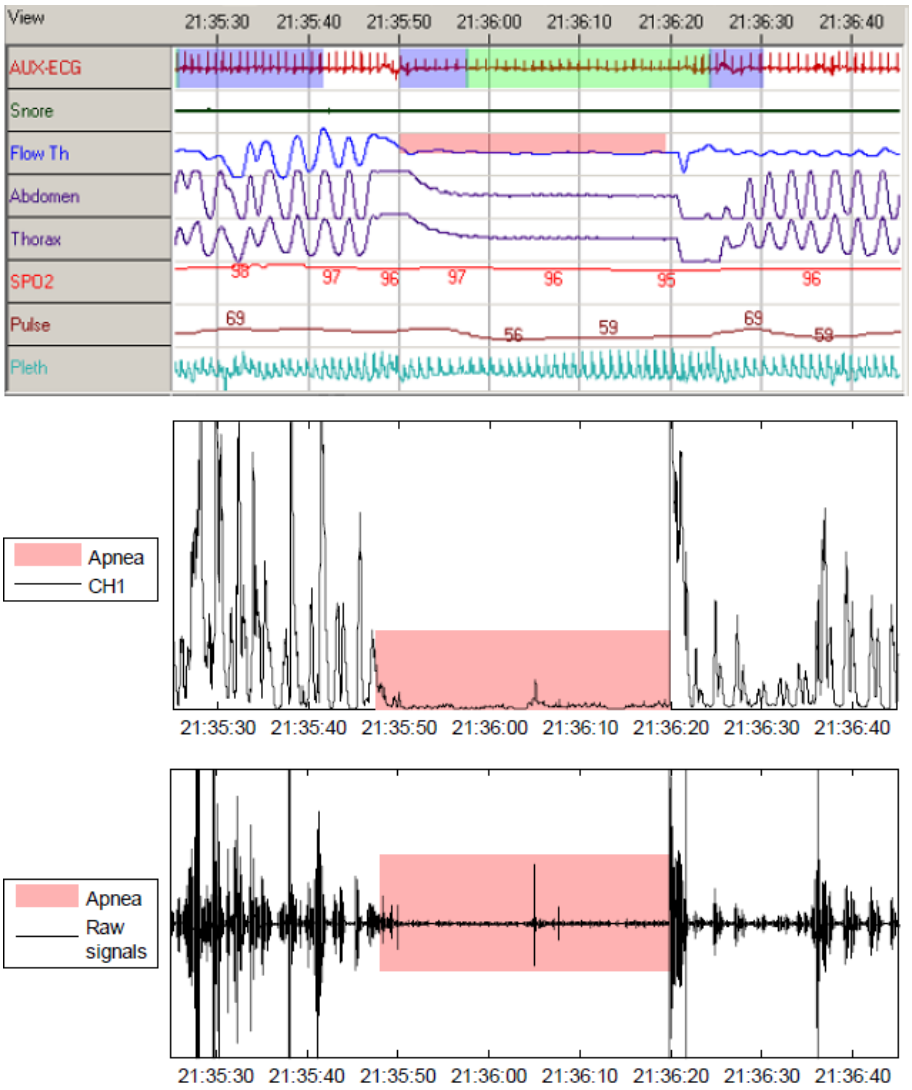

Figure 6-2: Illustration of the SOMNO and Wearable Apnoea Detection Device (WADD) output signals showing an apnoea event: (top) raw signals from the different SOMNO sensors, (middle) processed envelope signals from channel 1(350Hz – 450 Hz), (bottom) WADD raw signals.

	Sensitivity (%)		Specificity (%)	
Exercise	**WADD**	**SOMNO**	**WADD**	**SOMNO**
(A) WADD versus SOMNO performance in instructed exercises with hypopnoeas NOT considered as events				
T1	N/A	N/A	100	99.2
T2	N/A	N/A	100	90.6
T3	94.6	38.0	99.0	96.9
T4	98.9	38.8	99.7	94.5
T5	99.2	31.4	99.2	99.7
T6	N/A	N/A	100	96.5
T7	94.2	48.2	98.5	99.0
Total	**97.4**	**37.8**	**99.6**	**96.5**
(B) WADD versus SOMNO performance in instructed exercises with hypopnoeas and apnoeas indistinctively considered as events				
T1	N/A	N/A	100	99.0
T2	N/A	N/A	100	81.4
T3	96.7	66.3	99.0	87.5
T4	100	64.6	99.7	87.9
T5	99.2	59.2	99.2	95.1
T6	N/A	N/A	100	89.8
T7	99.0	64.4	97.0	93.4
Total	**99.2**	**62.8**	**99.5**	**90.5**

	TP		**TN**		**FC**		**FP**		**FN**	
Exercise	**WADD**	**SOMNO**	**WADD**	**SOMNO**	**WADD**	**SOMNO**	**WADD**	**SOMNO**	**WADD**	**SOMNO**
(C) Summary of classification of the different epochs										
T1	0	0	380	376	0	0	0	4	0	0
T2	0	0	360	293	0	0	0	67	0	0
T3	89	61	285	252	2	26	3	36	3	31
T4	356	230	363	320	4	92	1	44	0	126
T5	357	213	365	350	0	100	3	18	3	147
T6	0	0	400	359	0	0	0	41	0	0
T7	189	123	191	184	9	31	6	13	2	68
Total	**991**	**627**	**2344**	**2134**	**15**	**249**	**13**	**223**	**8**	**372**

Part (A) of the table shows the sensitivity and specificity not considering hypopnoeas as events (ie, all hypopnoeas are considered breathing). Based on this all FCs (apnoeas wrongly classified as hypopnoeas) are considered FNs; and all FPs hypopnoeas are considered TNs. Part (B) shows the sensitivity and specificity considering apnoea and hypopnoea as indistinctive events. Based on this all FC are reclassified as TPs; and all FP hypopnoeas are FPs. Part (C) details the number of epochs corresponding to a particular classification.
FC, false classification; FN, false negative; FP, false positive; TN, true negative (breathing); TP, true positive (apnoea); WADD, wearable apnoea detection device.

Table 6-1: Summary of performance for the WADD and SOMNO across the seven breath exercises (as detailed in procedure)

	SOMNO sensitivity	WADD sensitivity	Clinician sensitivity	SOMNO specificity	WADD specificity	Clinician specificity
Apnoea detection (% sensitivity and specificity)						
Controls	47.1	97.1	94.1	99.3	99.8	100
(95% CI)	(30.3 to 63.8)	(91.4 to 100)	(86.2 to 100)	(99.2 to 99.3)	(99.7/99.8)	(100 to 100)
Patients	14.7	99.4	98.1	99.5	99.5	99.9
(95% CI)	(10.8 to 18.7)	(98.5 to 100)	(96.6 to 99.6)	(99.5 to 99.6)	(99.4 to 99.6)	(99.8 to 99.9)
All	17.9	99.1	97.7	99.4	99.7	99.9
(95% CI)	(13.9 to 22.0)	(98.2 to 100)	(96.1 to 99.3)	(99.3 to 99.4)	(99.6 to 99.7)	(99.9 to 100)
Apnoea and hypopnoea combined detection (%)						
Controls	58.1	87.8	94.6	98.6	99.7	100
(95% CI)	(46.9 to 69.4)	(80.4 to 95.3)	(89.4 to 99.8)	(98.5 to 98.7)	(99.6 to 99.7	(100 to 100)
Patients	53.3	88.2	98.8	97.9	99.5	99.8
(95% CI)	(48.9 to 57.8)	(85.4 to 91.1)	(97.8 to 99.8)	(97.7 to 98.1)	(99.4 to 99.6)	(99.8 to 99.9)
All	57.8	84.1	98.2	98.4	99.5	99.9
(95% CI)	(53.8 to 61.9)	(81.1 to 87.1)	(97.2 to 99.3)	(98.3 to 98.5)	(99.5 to 99.6)	(99.9 to 100)

	SOMNO sensitivity	WADD sensitivity	SOMNO specificity	WADD specificity
Apnoea detection with the clinician scorer as gold standard reference (%)				
Controls	38.9	86.1	99.2	99.7
(95% CI)	(23.0 to 54.8)	(74.8 to 97.4)	(99.1 to 99.3)	(99.7 to 99.8)
Patients	11.7	88.9	99.5	99.4
(95% CI)	(8.3 to 15.1)	(85.6 to 92.2)	(99.4 to 99.6)	(99.3 to 99.5)
All	14.3	88.6	99.3	99.6
(95% CI)	(10.8 to 17.8)	(85.4 to 91.8))	(99.2 to 99.4))	(99.6 to 99.7)
Apnoea and hypopnoea combined detection with clinician scorer as gold standard (%)				
Controls	54.8	86.3	98.6	99.6
(95% CI)	(40/73)	(63/73)	(41 539/42 139)	(41 987/ 42 139)
Patients	49.6	80.1	98.4	100
(95% CI)	(45.4 to 53.8)	(75.2 to 82.1))	(98.2 to 98.6)	(99.9 to 100)
All	54.0	77.1	98.5	99.7
(95% CI)	(50.0 to 57.9)	(73.8 to 80.5)	(98.4 to 98.6)	(99.7 to 99.8)

Table 6-2: Summary of performance for the WADD, SOMNO and clinician scorer systems for detection of apnoea and hypopnoea in 10 s epochs of overnight recordings.

6.4.2 Sleep Data

For the sleep data 62727 10s epochs were analysed, in total, 34 true apnoea epochs and 40 true hypopnoea epochs were identified for the controls (36 and 37 if the clinician scorer was considered to be the gold standard); and 312 apnoea epochs and 181 hypopnoea epochs for the patients (342 and 200 if the clinician scorer was considered to be the gold standard). The average number of apnoea epochs for the patient group throughout the night was 32. All patients had episodes of apnoea or hypopnoea. There were only two patients who did not have any episode of apnoea. For one control, only 3.2 h of data were recorded, because of an ICT error. For one patient, only 3 h were analysed as more than one SOMNO sensor including the

nasal cannula and the pulse oximeters detached prematurely. The results in terms of sensitivity and specificity for the control group, patient group and overall are presented in Table 6-2. Table 6-2 is divided in four parts: the first and second evaluate the performance for apnoea and apnoea/hypopnoea combined detection respectively without assuming a gold standard (ie, the consensus of the majority determines a true event); and the third and fourth parts present the same evaluation but considering the expert as the gold standard.

6.4.2.1 Sleep Apnoea Event Classification

The WADD also had the added feature of being able to differentiate between central and obstructive apnoea. 90% of the central apnoeas were rightly marked as central. 96% were rightly marked as obstructive. Approximately 60% of the total apnoeas were obstructive in origin.

6.4.2.2 Respiratory Rate

The respiratory rate (per minute) was calculated for the controls' data between 1 AM to 6 AM. 38071 respiratory cycles were analysed. The results from WADD were compared with the respiratory rate evaluated from SOMNO based on visual counting of the synchronised waveforms from the oro-nasal airflow sensors as well as thoracic and abdominal expansion bands. Any sections that did not have synchronised waveforms were rejected from the analysis since it was not possible to have certainty of the accurate evaluation of the respiratory rate since those signals were likely to be corrupted by movement artefacts. Bland-Altman plots were obtained to explore whether the algorithms could have the potential to substitute visual counting (currently clinical gold standard method) for determination of breathing rate. Figure 6-3 shows the Bland-Altman plots for the 10 control subjects. This shows a bias of -0.1382, with a standard deviation of 1.5915. The limits of agreement for the difference have been defined as between bias ± two standard deviation which is -0.1382 ± 3.183. Because the results of respiratory rate were rounded into integer before the comparison, the dots in the plot have a lot of overlapping. In order to show the data density, the diameter size of the dot represents the amount of overlapping. In this data set, 94% of the dots fall within the limits of agreement.

6.4.3 Device Comfort

After the overnight study, the devices were detached and the participants scored the comfort of the devices and quality of sleep (rating 1–5, with 5 representing maximum comfort and quality).

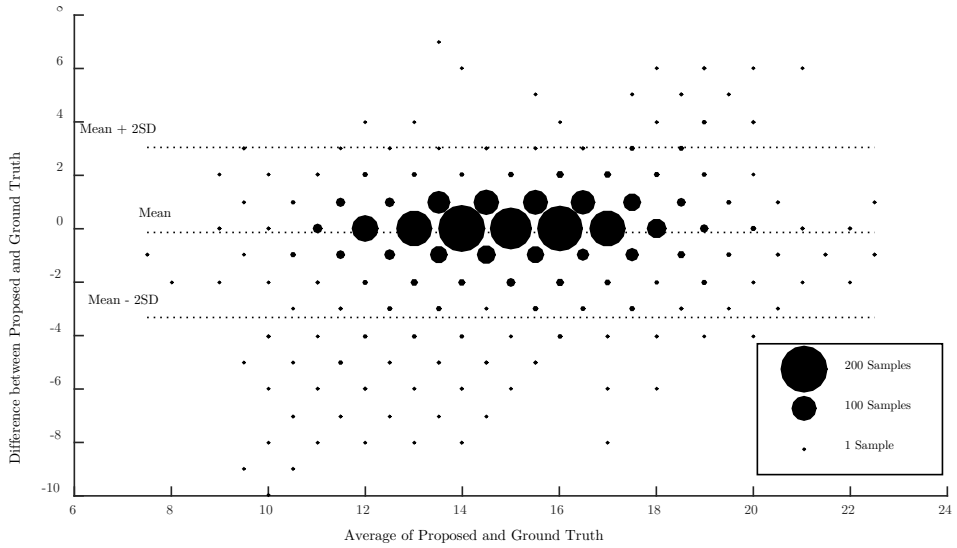

Figure 6-3: Difference between visually determined respiratory rate and the corresponding automatic value calculated by WADD against the mean of two methods. Total number of samples is 38071.

Skin irritation caused by the WADD's adhesive was also rated from 1 to 5 (5 representing no irritation, 4 mild transient, redness and 1 severe irritation). The median rating for WADD comfort was 5 (range 4–5). The median rating for SOMNO comfort was 3 (range 1–5 for controls and 2–5 for patients). The median rating for irritation caused by the WADD plaster on the neck was 5 (range 5–5 for controls and 4–5 for patients).

6.5 Discussion

6.5.1 Main Findings

WADD had very high sensitivity and specificity for detecting apnoea in 15 s epochs in a series of breathing and breath-holding exercises in a variety of conditions, including the presence of external background noise, movement and posture. The tolerability of WADD was superior to the portable polysomnography system (SOMNO) during overnight recordings.

WADD had 97.7–99.2% sensitivity to detect instructed apnoeas and 88.6–99.1% for 10 s spontaneous apnoeas during natural sleep, with similar performance in controls and patients. The WADD also detected all apnoeas over 30 s and there were only 3 over 30 s false positives. For short apnoeas, in most cases, disagreement between the clinician scorer and the WADD were caused by the WADD identifying as apnoea epochs that the expert classified as hypopnoeas.

As expected, the WADD performance was not as good when apnoeas and hypopnoeas were considered together (minimum sensitivity 77.1%).This is not surprising since the WADD was designed to identify apnoea, not hypopnoeas and the latter were detected from the transmitted signal that had already been preprocessed for apnoea detection. From the table, it can be observed that the degradation of performance was more evident in the controls because the controls had a large number of shorter hypopnoeas (under 22.5 s), which the WADD did not detect properly. In the patients, who often demonstrated apnoeas, the hypopnoea events were longer and these were detected by the WADD. Although the lower sensitivity in hypopnoea detection might in principle seem problematic if the WADD were to be used in the context of sleep apnoea diagnosis (hypopnoeas are very common events in sleep laboratories), it is worth noting that: (1) there is no other reported automatic system that gets anywhere close to this with similar specificity and apnoea detection performance; (2) the variations between different sleep laboratories due to the non-uniform definition of hypopnoeas already leads to much larger diagnostic variations than the limitation in sensitivity of the WADD [265-267]; and (3) assuming the worst case scenario for the WADD, being that a patient only had hypopnoeas throughout the night, this reduced sensitivity would be a problem that would translate to non-diagnosis of sleep apnoea for patients who with 100% sensitivity would have had a sleep apnoea hypopnoea index (AHI) between 5 and 6 (ie, very mild cases of sleep apnoea). Patients with no sleep apnoea, moderate sleep apnoea, severe sleep apnoea and those with mild sleep apnoea with AHI between 6 and 15 would have been rightly diagnosed. The median difference between the WADD calculated AHI and the one obtained by the gold standard was 0 (average=0.7).

Regarding to the respiratory rate estimation, the accuracy does not reach the clinical requirement mentioned in Section 1.3.1.1. Additional works are required to improve that.

6.5.2 SOMNO Performance

The automatic analysis of the SOMNO apparatus, a Food and Drug Administration (FDA) approved and clinically accepted system, based on assessing apnoea from a variety of different sensors, significantly differed from that of the expert marker, with an average sensitivity value of around 14%. The results obtained from the instructed apnoeas tests also showed that even in the absence of artefacts, apnoeas were not well detected by the SOMNO system, with an average sensitivity of 37.8%. This demonstrates the need for caution if relying on current

automated assessment methods for diagnosing apnoea. While performance might be improved by optimising parameters for individual patients, this is not practical for single overnight recordings or use as an alerting monitor. The WADD does not require any parameter optimisation or participant specific calibration.

The SOMNO system performance improved in the event of indistinct classification of apnoeas and hypopnoeas, but was still poor compared with the clinician scorer (54% sensitivity). This sensitivity was at the cost of reduced specificity: for every true hypopnoea detected there were approximately four false detections. Overall, the performance of the WADD in hypopnoea /apnoea combined detection was significantly better than the SOMNO, in sensitivity (77.1% vs 54% if considering the clinician scorer as a gold standard, and 84.1% vs 57.8% otherwise), but also in specificity, as the WADD only detected one false hypopnoea epoch for every four true events.

6.5.3 Results Validation

Although the number of subjects in this trial was limited, the size of the sleep study was, however, adequate to assess the potential of the technology; to determine whether the initial performance results in controlled conditions were equivalent to those obtained in real scenarios; and to inform a clinical test.

The calculation of the sensitivity and specificity has assumed that all apnoea events were independent, which for some might not be completely correct. If the data had not been pooled, and taking the expert marker as the gold-standard, in 67% of the participants the individual apnoea detection sensitivity was 100%. In 77% it was over 90%. In the remaining cases, the drop in sensitivity corresponded always to just one non-detected apnoea shorter than 15 seconds in the 6 hours night, which is clinically insignificant. The average from the individual sensitivities was 2% higher than the value obtained pooling the data. In terms of specificity 90% of the participants had values higher than 99%. Two-thirds of them were over 99.9%. The average of the individual specificities was identical to the specificity obtained pooling the data.

A different statistical analysis, possibly comparing pooled with non-pooled data, will be the subject of investigation when the technology undergoes a larger clinical trial.

Conclusion

Contributions

This thesis has presented a novel non-intrusive wearable respiratory monitoring system which can be of use in a variety of clinical and non-clinical applications such as sleep apnoea diagnosis, early warning scoring, sudden unexpected death prevention or human behavioural monitoring. The system has the form of a single miniaturised, wearable sensor unit which is easy to put on, comfortable to wear, non-intrusive, and is also capable of automatically detecting respiratory events accurately.

Chapter 1 reviewed the motivation of this PhD and the state-of-the-art of wearable non-invasive respiratory monitoring technologies. Existing technologies have at least one or more disadvantages which may include the fact that they are cumbersome, uncomfortable, difficult to keep in place, inaccurate during daily life activities, provide limited respiratory information, or have to be used under professional supervision. Based on this, key design considerations for the respiratory monitoring system targeted in this work were also discussed in this chapter.

Chapter 2 provided a concise review of the origin and characteristics of breath sounds. Suitable data acquisition approaches to collect acoustic signals for the purpose of respiratory events detection were subsequently analysed. The optimal sensing locations for a power constraint wearable sensor were then justified. Necessary considerations for hardware design as well as signal filtering techniques aiming to minimise the overall system noises were finally discussed.

Chapter 3 describes a novel system level trade-offs analysis that shows how to take into account the different design variables together with components specifications in order to optimise both the overall wearable system size, as well as to maximise the battery lifetime, and hence minimise the technology maintenance requirements.

Chapter 4 presented a novel data compression technique specifically created for breath sounds. It is shown how this breath sound compression algorithm has three advantages over convention-

al compression techniques. First, the original breath sound synthesis predictor greatly reduces the size of residual information. Second, the method modifies the classical definition of compression frame, which contains a fixed number of samples, to a fixed length adaptive sample size frame. The proposed compression frame structure overcomes the small data packet problem imposed by low power wireless transmission protocols such as Bluetooth Smart (formally known as Bluetooth Low Energy), with the amount of samples being adapted using the entropy of the incoming breath sound data. This allows the system to automatically adjust the size of the compression frame (in terms of sample count) according to the input data, making the residuals encoding more efficient. This technique was implemented together with state-of-the-art residuals encoding ones, on a TI CC2640 Bluetooth® Smart controller, demonstrating 32.94% to 38.94% compression ratio for recordings obtained from the clinical test presented in Chapter 6. To this date, Bluetooth Smart could not be used in this system without this data compression technique because the original data rate reached the limit of the practical transmission speed. In addition, the chapter showed how using this technique is also beneficial from the point of view of power.

Chapter 5 detailed the design of the novel automatic respiratory event detection and analysis algorithm. This algorithm can be divided into two main parts. The first section of the algorithm separates the acoustic inputs into the multi-channel signals. The second section focuses on feature extraction, two-level verification of breath detection and event classification. The first level of the verification tends to be conservative, making sure there are no false detection of breath events whilst at the same time providing some breath detection with insufficient confidence for further judgements. The second level was designed to be more flexible to suit the needs of different respiratory monitoring applications, by trading off sensitivity and specificity prior to classifying the respiratory events. A novel method is also proposed to group the detected breath events as full respiratory cycles or part of a respiratory cycle. During the recognition of the respiratory cycles, respiratory rate was also extracted based on events time separations and the assumption of tidal volumes conservation between inspiration and expiration. Added to this, another method is proposed to sub-classify the sections without full breath detection as hypopnoea, central apnoea, obstructive/mixed apnoea or hyperventilation.

The final chapter of this thesis, Chapter 6, presented the results of the novel breathing detection and classification algorithm validated on a clinical test. This test involved 20 controls and 10 patients and took place in the Sleep Unit of the UK National Hospital for Neurology and

Neurosurgery. A participants' survey demonstrated how all subjects felt comfortable wearing the respiratory monitoring sensor. The trial also demonstrated the very high accuracy of the algorithm in detecting all kind of apnoea events and its massive superiority when compared to state-of-the-art commercial software. Finally, a preliminary verification of automatic respiratory rate estimation was carried out for 10 control subjects. The results showed that 94% of the measurement fell within the statistical Bland-Altman limits of agreement.

Further work

There are advantages and disadvantages of the work presented in this thesis with respect to full polysomnography. The current implementation is obviously no substitute for a full night study in a sleep clinic, since it does not provide all the information that a full polysomnography system would. Currently, the main advantage of this technology is that it can be used for home assessment or monitoring, and from that point of view it is clearly superior to any other existing devices (highly resilient to artefacts, very easy to attach and durable in position, low cost, much more comfortable and accurate). Considering the restricted resources for sleep clinic referrals, this device could be a very useful tool to determine, at very low cost, for who should be referred to a specialist centre for full polysomnography. The disadvantage is that there are other parameters that could be used for extra-assessment that the device currently does not measure, such as microarousals, full cardiac activity or oxygen saturation which some believe to be needed for assessment of hypopnoea. Future research could investigate how to incorporate some of these parameters into the existing system or a modified version of it.

On the hardware implementation side, the prototype used in the current study relied on the uncompressed (but downsampled) raw data being wireless transmitted to a dedicated receiver. Changing the receiver to a dedicated mobile phone using Bluetooth® Smart protocol requires the novel breath sound compression technique to be validated in different scenarios rather than just in data obtained in a sleep study room.

Additionally, all the data analysis (including real time and off-line versions) is implemented on a PC with floating point data representation. This may not be optimal in terms of speed and power consumption if the algorithm was to be run on a portable device, hence making necessary the conversion to a fixed point implementation. Further research on the minimum bit length

required by the fixed point implementation in order to maintain the same level of performance will hence has to be carried out.

Finally, the validation is not limited to respiratory event detection. Based on the positive results obtained from the small pilot study in Chapter 6, it is expected that a fully powered clinical trial, focused on respiratory disease/condition diagnosis will follow in the future. Furthermore, research on incorporating adventitious breath sounds automatic identification algorithms [268] could be carried out to extend the functionality of the technology to other clinical applications. In the context of Sudden Unexpected Death Prevention in Epilepsy (SUDEP), further work is anticipated to be able to identify central apnoeas accurately when the signal is corrupted by the artefacts caused by seizures. With little finer tuning and comprehensive clinical evaluation this novel respiratory monitoring technology could enter commercialisation stage readily.

Bibliography

[1] J. A. Schafer and J. J. Caronna, "Duration of apnea needed to confirm brain death," *Neurology,* vol. 28, pp. 661-6, Jul 1978.

[2] M. D. Sharpe, G. B. Young, and C. Harris, "The apnea test for brain death determination," *Neurocritical Care,* vol. 1, p. 363, 01 Sep 2004.

[3] M. B. Taylor, R. M. Grounds, P. D. Mulrooney, and M. Morgan, "Ventilatory effects of propofol during induction of anaesthesia. Comparison with thiopentone," *Anaesthesia,* vol. 41, pp. 816-20, Aug 1986.

[4] G. Van Vliet, C. L. Deal, P. A. Crock, Y. Robitaille, and L. L. Oligny, "Sudden death in growth hormone-treated children with Prader-Willi syndrome," *J Pediatr,* vol. 144, pp. 129-31, Jan 2004.

[5] J. J. Murphy, T. Swanson, M. Ansermino, and R. Milner, "The frequency of apneas in premature infants after inguinal hernia repair: do they need overnight monitoring in the intensive care unit?," *J Pediatr Surg,* vol. 43, pp. 865-8, May 2008.

[6] P. Moruzzi, S. Sarzi-Braga, M. Rossi, and M. Contini, "Sleep apnoea in ischaemic heart disease: differences between acute and chronic coronary syndromes," *Heart,* vol. 82, pp. 343-347, Sep 1999.

[7] L. Cordero, R. Rau, D. Taylor, and L. W. Ayers, "Enteric gram-negative bacilli bloodstream infections: 17 years' experience in a neonatal intensive care unit," *Am J Infect Control,* vol. 32, pp. 189-95, Jun 2004.

[8] S. E. Mace, "Acute bacterial meningitis," *Emerg Med Clin North Am,* vol. 26, pp. 281-317, viii, May 2008.

[9] S. Shorvon and T. Tomson, "Sudden unexpected death in epilepsy," *Lancet,* vol. 378, pp. 2028-38, Dec 10 2011.

[10] K. Powers, "Acute Pulmonary Infections," in *Pediatric Critical Care Study Guide*, S. E. Lucking, F. A. Maffei, R. F. Tamburro, and N. J. Thomas, Eds., ed: Springer London, 2012, pp. 514-534.

[11] J. Gravenstein, "Monitoring in the post-operative period," *Current Anaesthesia & Critical Care,* vol. 2, pp. 32-36, 1991.

[12] A. Lofsky, "Sleep apnea and narcotic postoperative pain medication: a morbidity and mortality risk," *Anesthesia Patient Safety Foundation Newsletter,* vol. 17, pp. 24-5, 2002.

[13] S. M. Bhananker, K. L. Posner, F. W. Cheney, R. A. Caplan, L. A. Lee, and K. B. Domino, "Injury and liability associated with monitored anesthesia care: a closed claims analysis," *Anesthesiology,* vol. 104, pp. 228-34, Feb 2006.

[14] S. Garbarino, L. Nobili, M. Beelke, F. De Carli, and F. Ferrillo, "The contributing role of sleepiness in highway vehicle accidents," *Sleep,* vol. 24, pp. 203-6, Mar 15 2001.

[15] J. Teran-Santos, A. Jimenez-Gomez, 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. 847-51, Mar 18 1999.

[16] T. Young, P. E. Peppard, and D. J. Gottlieb, "Epidemiology of obstructive sleep apnea: a population health perspective," *Am J Respir Crit Care Med,* vol. 165, pp. 1217-39, May 1 2002.

[17] N. Hanna, M. Black, J. Sander, W. Smithson, R. Appleton, S. Brown*, et al.*, *National Sentinel clinical audit of epilepsy-related death: report 2002. Epilepsy-death in the shadows*: Stationery Office Books, 2002.

[18] O. f. N. Statistics, "Unexplained Deaths in Infancy: England and Wales, 2012," 2014.

[19] G. R. O. f. Scotland, "Infant deaths, by sex and cause, Scotland, 2003 to 2013," 2014.

[20] N. I. S. R. Agency, "Infant Deaths and Still Births: 2013 Provisional," 2014.

[21] N. A. Collop, W. M. Anderson, B. Boehlecke, D. Claman, R. Goldberg, D. J. Gottlieb*, et al.*, "Clinical Guidelines for the use of unattended portable monitors in the diagnosis of obstructive sleep apnea in adult patients," *Journal of Clinical Sleep Medicine,* vol. 3, p. 737, 2007.

[22] M. Hirshkowitz, "The clinical consequences of obstructive sleep apnea and associated excessive sleepiness," *J Fam Pract,* vol. 57, pp. S9-16, Aug 2008.

[23] C. Guilleminault, R. Peraita, M. Souquet, and W. Dement, "Apneas during sleep in infants: possible relationship with sudden infant death syndrome," *Science,* vol. 190, pp. 677-679, November 14, 1975 1975.

[24] M. R. Suchyta, T. P. Clemmer, C. G. Elliott, J. J. F. Orme, and L. K. Weaver, "The adult respiratory distress syndrome. a report of survival and modifying factors," *Chest,* vol. 101, pp. 1074-1079, 1992.

[25] E. L. So, M. C. Sam, and T. L. Lagerlund, "Postictal central apnea as a cause of SUDEP: Evidence from Near-SUDEP incident," *Epilepsia,* vol. 41, pp. 1494-1497, Nov 2000.

[26] J. McBride, D. Knight, J. Piper, and G. B. Smith, "Long-term effect of introducing an early warning score on respiratory rate charting on general wards," *Resuscitation,* vol. 65, pp. 41-4, Apr 2005.

[27] E. Institute, "Top 10 health technology hazards for 2014," *Health Devices,* vol. 42, pp. 354-80, Nov 2013.

[28] J. C. Bauer, "Automated, wireless, continuous, intelligent, and noninvasive systems for risk management and performance improvement in vital signs monitoring," *J Healthc Risk Manag,* vol. 33, pp. 12-6, 2013.

[29] K. Davidson and V. Barber, "Electronic monitoring of patients in general wards," *Nurs Stand,* vol. 18, pp. 42-6, Aug 18-24 2004.

[30] J. Welch, "An evidence-based approach to reduce nuisance alarms and alarm fatigue," *Biomed Instrum Technol,* vol. Suppl, pp. 46-52, Spring 2011.

[31] D. Clifton, D. Wong, S. Fleming, S. Wilson, R. Way, R. Pullinger*, et al.*, "Novelty Detection for Identifying Deterioration in Emergency Department Patients," in *Intelligent Data Engineering and Automated Learning - IDEAL 2011*. vol. 6936, H. Yin, W. Wang, and V. Rayward-Smith, Eds., ed: Springer Berlin Heidelberg, 2011, pp. 220-227.

[32] M. Santos, D. Clifton, and L. Tarassenko, "Performance of Early Warning Scoring Systems to Detect Patient Deterioration in the Emergency Department," in *Foundations of Health Information Engineering and Systems*. vol. 8315, J. Gibbons and W. MacCaull, Eds., ed: Springer Berlin Heidelberg, 2014, pp. 159-169.

[33] C. Subbe, H. Gao, and D. Harrison, "Reproducibility of physiological track-and-trigger warning systems for identifying at-risk patients on the ward," *Intensive Care Medicine,* vol. 33, pp. 619-624, 2007/04/01 2007.

[34] C. Guilleminault, R. L. Ariagno, L. S. Forno, L. Nagel, R. Baldwin, and M. Owen, "Obstructive sleep apnea and near miss for SIDS: I. Report of an infant with sudden death," *Pediatrics,* vol. 63, pp. 837-843, 1979.

[35] H. O. SIEKER and J. B. HICKAM, "Carbon dioxide intoxication: the clinical syndrome, its etiology and management with particular reference to the use of mechanical respirators," *Medicine,* vol. 35, pp. 389-424, 1956.

[36] B. Kelly and J. Luce, "The diagnosis and management of neuromuscular diseases causing respiratory failure," *CHEST Journal,* vol. 99, pp. 1485-1494, 1991.

[37] M. A. Nadkarni, D. Friedman, and O. Devinsky, "Central apnea at complex partial seizure onset," *Seizure,* vol. 21, pp. 555-558, 9// 2012.

[38] L. P. Sowers, C. A. Massey, B. K. Gehlbach, M. A. Granner, and G. B. Richerson, "Sudden unexpected death in epilepsy: fatal post-ictal respiratory and arousal mechanisms," *Respir Physiol Neurobiol,* vol. 189, pp. 315-23, Nov 1 2013.

[39] S. U. Schuele, M. Afshari, Z. S. Afshari, M. P. Macken, J. Asconape, L. Wolfe*, et al.*, "Ictal central apnea as a predictor for sudden unexpected death in epilepsy," *Epilepsy & Behavior,* vol. 22, pp. 401-403, Oct 2011.

[40] S. Meyer, M. G. Shamdeen, S. Gottschling, M. Strittmatter, and L. Gortner, "Sudden unexpected death in epilepsy in children," *J Paediatr Child Health,* vol. 47, pp. 326-31, Jun 2011.

[41] J. E. Council, "Epilepsy prevalence, incidence and other statistics," *Joint Epilepsy Council, Leeds,* 2005.

[42] R. J. Lamberts, R. D. Thijs, A. Laffan, Y. Langan, and J. W. Sander, "Sudden unexpected death in epilepsy: People with nocturnal seizures may be at highest risk," *Epilepsia,* vol. 53, pp. 253-257, 2012.

[43] S. Fulton, K. V. Poppel, A. McGregor, M. Ellis, A. Patters, and J. Wheless, "Prospective Study of 2 Bed Alarms for Detection of Nocturnal Seizures," *J Child Neurol,* vol. 28, pp. 1430-1433, Nov 2013.

[44] R. Y. Moon, R. S. C. Horne, and F. R. Hauck, "Sudden infant death syndrome," *The Lancet,* vol. 370, pp. 1578-1587, 2007.

[45] M. Page and H. Jeffery, "The role of gastro-oesophageal reflux in the aetiology of SIDS," *Early Human Development,* vol. 59, pp. 127-149, Aug 2000.

[46] J. R. Kerr, A. Al-Khattaf, A. J. Barson, and J. P. Burnie, "An association between sudden infant death syndrome (SIDS) and Helicobacter pylori infection," *Archives of Disease in Childhood,* vol. 83, pp. 429-434, November 1, 2000 2000.

[47] D. Leger, V. Bayon, J. P. Laaban, and P. Philip, "Impact of sleep apnea on economics," *Sleep Med Rev,* vol. 16, pp. 455-62, Oct 2012.

[48] B. Fauroux, "What's new in paediatric sleep?," *Paediatr Respir Rev,* vol. 8, pp. 85-9, Mar 2007.

[49] E. Shahar, C. W. Whitney, S. Redline, E. T. Lee, A. B. Newman, F. J. Nieto*, et al.*, "Sleep-disordered breathing and cardiovascular disease: cross-sectional results of the Sleep Heart Health Study," *Am J Respir Crit Care Med,* vol. 163, pp. 19-25, Jan 2001.

[50] V. Wittmann and D. O. Rodenstein, "Health care costs and the sleep apnea syndrome," *Sleep Med Rev,* vol. 8, pp. 269-79, Aug 2004.

[51] L. P. Boulet, J. Bourbeau, R. Skomro, and S. Gupta, "Major care gaps in asthma, sleep and chronic obstructive pulmonary disease: a road map for knowledge translation," *Can Respir J,* vol. 20, pp. 265-9, Jul-Aug 2013.

[52] N. J. Douglas, "Snoring and Sleep Apnoea," in *Lung Report III*, ed: British Lung Foundation, 2003, pp. 36-38.

[53] M. R. Shepertycky, K. Banno, and M. H. Kryger, "Differences between men and women in the clinical presentation of patients diagnosed with obstructive sleep apnea syndrome," *Sleep,* vol. 28, pp. 309-14, Mar 2005.

[54] N. Ahmadi, G. Shapiro, S. Chung, and C. Shapiro, "Clinical diagnosis of sleep apnea based on single night of polysomnography vs. two nights of polysomnography," *Sleep and Breathing,* vol. 13, pp. 221-226, 2009/08/01 2009.

[55] W. H. Organization. (2015, Jan 04). *Asthma Fact sheet.* Available: http://www.who.int/mediacentre/factsheets/fs307/en/

[56] W. H. Organization. (2015, Jan 04). *Chronic obstructive pulmonary disease (COPD) Fact sheet.* Available: http://www.who.int/mediacentre/factsheets/fs315/en/

[57] A. Avci, S. Bosch, M. Marin-Perianu, R. Marin-Perianu, and P. Havinga, "Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: A survey," in *Architecture of computing systems (ARCS), 2010 23rd international conference on*, 2010, pp. 1-10.

[58] T. Haderlein, S. Steidl, E. Nöth, F. Rosanowski, and M. Schuster, "Automatic recognition and evaluation of tracheoesophageal speech," in *Text, Speech and Dialogue*, 2004, pp. 331-338.

[59] J. Sundberg, "Breathing behavior during singing," *NATS Journal,* vol. 49, pp. 44-49, 1992.

[60] M. Aboofazeli and Z. Moussavi, "Swallowing sound detection using hidden markov modeling of recurrence plot features," *Chaos, Solitons & Fractals,* vol. 39, pp. 778-783, 2009.

[61] A. A. Ali, S. M. Hossain, K. Hovsepian, M. M. Rahman, K. Plarre, and S. Kumar, "mPuff: Automated detection of cigarette smoking puffs from respiration measurements," in *Information Processing in Sensor Networks (IPSN), 2012 ACM/IEEE 11th International Conference on*, 2012, pp. 269-280.

[62] S. Shea, "Behavioural and arousal-related influences on breathing in humans," *Experimental Physiology,* vol. 81, pp. 1-26, Jan 1 1996.

[63] R. Jane, J. Sola-Soler, J. A. Fiz, and J. Morera, "Automatic detection of snoring signals: validation with simple snorers and OSAS patients," in *Engineering in Medicine and Biology Society, 2000. Proceedings of the 22nd Annual International Conference of the IEEE*, 2000, pp. 3129-3131 vol.4.

[64] R. B. Crossley, "Multifunctional behavioral modification device for snoring, bruxism, and apnea," US4715367A, Dec 29, 1987.

[65] F. A. Boiten, "The effects of emotional behaviour on components of the respiratory cycle," *Biological Psychology,* vol. 49, pp. 29-51, Sep 1998.

[66] G. Charbonneau, J. Racineux, M. Sudraud, and E. Tuchais, "An accurate recording system and its use in breath sounds spectral analysis," *Journal of applied physiology,* vol. 55, p. 1120, 1983.

[67] F. W. Kranz, "The Sensitivity of the Ear as a Function of Pitch," *Physical review,* vol. 22, p. 66, 1923.

[68] V. A. McKusick, J. T. Jenkins, and G. N. Webb, "The acoustic basis of the chest examination; studies by means of sound spectrography," *The American review of tuberculosis,* vol. 72, pp. 12-34, 1955.

[69] R. Nachman, M. Nachman, and L. Vinaru, "Recording and analysis of sounds produced by human lungs during respiration," *Cellular and molecular life sciences,* vol. 25, p. 1040, 1969.

[70] R. L. Murphy and K. Sorensen, "Chest auscultation in the diagnosis of pulmonary asbestosis," *Journal of occupational medicine,* vol. 15, pp. 272-6, 1973.

[71] E. F. Banaszak, R. C. Kory, and G. L. Snider, "Phonopneumography," *The American review of respiratory disease,* vol. 107, pp. 449-55, 1973.

[72] K. Akasaka, K. Konno, Y. Ono, S. Mue, and C. Abe, "Acoustical studies on respiratory sounds in asthmatic patients," *The Tohoku journal of experimental medicine,* vol. 117, pp. 323-33, 1975.

[73] J. Werthammer, J. Krasner, J. DiBenedetto, and A. R. Stark, "Apnea Monitoring by Acoustic Detection of Airflow," *Pediatrics,* vol. 71, pp. 53-55, Jan 1983.

[74] N. Gavriely, Y. Palti, and G. Alroy, "Spectral characteristics of normal breath sounds," *Journal of applied physiology,* vol. 50, p. 307, 1981.

[75] G. Charbonneau, J. L. Racineux, E. Tuchais, and M. Sudraud, "Digital processing techniques of breath sounds for objective assistance of asthma diagnosis," *Proceedings of ICASSP 82. IEEE International Conference on Acoustics, Speech and Signal Processing,* pp. 736-8, 1982.

[76] N. Gavriely and G. Charbonneau, "Measurement of tracheal lung sounds," *Journal of applied physiology,* vol. 56, p. 817, 1984.

[77] M. Nissan and N. Gavriely, "A microcomputer based lung sounds analysis," *Computer methods and programs in biomedicine,* vol. 40, pp. 7-13, 1993.

[78] J. A. Reichert, D. A. Bloch, E. Cundiff, and B. A. Votteri, "Comparison of the NovaSom QSG, a new sleep apnea home-diagnostic system, and polysomnography," *Sleep Med,* vol. 4, pp. 213-8, May 2003.

[79] G. A. Pressler, J. P. Mansfield, H. Pasterkamp, and G. R. Wodicka, "Detection of respiratory sounds at the external ear," *Ieee Transactions on Biomedical Engineering,* vol. 51, pp. 2089-2096, Dec 2004.

[80] A. Yadollahi and Z. Moussavi, "Acoustic obstructive sleep apnea detection," *Conf Proc IEEE Eng Med Biol Soc,* vol. 2009, pp. 7110-3, 2009.

[81] H. Alshaer, G. R. Fernie, E. Sejdic, and T. D. Bradley, "Adaptive Segmentation and Normalization of Breathing Acoustic Data of Subjects with Obstructive Sleep Apnea," *Ieee*

Tic-Sth 09: 2009 Ieee Toronto International Conference: Science and Technology for Humanity, pp. 279-284, 2009.

[82] N. Ben-Israel, A. Tarasiuk, and Y. Zigel, "Nocturnal Sound Analysis for the Diagnosis of Obstructive Sleep Apnea," *2010 Annual International Conference of the Ieee Engineering in Medicine and Biology Society (Embc),* pp. 6146-6149, 2010.

[83] S. Huq and Z. Moussavi, "Automatic Breath Phase Detection Using Only Tracheal Sounds," *2010 Annual International Conference of the Ieee Engineering in Medicine and Biology Society (Embc),* pp. 272-275, 2010.

[84] M. Yamashita, S. Matsunaga, and S. Miyahara, "Discrimination between healthy subjects and patients with pulmonary emphysema by detection of abnormal respiration," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, 2011, pp. 693-696.

[85] O. Mimoz, T. Benard, A. Gaucher, D. Frasca, and B. Debaene, "Accuracy of respiratory rate monitoring using a non-invasive acoustic method after general anaesthesia," *British Journal of Anaesthesia,* vol. 108, pp. 872-875, May 2012.

[86] S. Huq and Z. Moussavi, "Acoustic breath-phase detection using tracheal breath sounds," *Medical & biological engineering & computing,* vol. 50, pp. 297-308, 2012.

[87] J. Jin and E. Sanchez-Sinencio, "A Home Sleep Apnea Screening Device With Time-Domain Signal Processing and Autonomous Scoring Capability," *IEEE Trans Biomed Circuits Syst,* Jun 3 2014.

[88] O. Yahya and M. Faezipour, "Automatic detection and classification of acoustic breathing cycles," in *American Society for Engineering Education (ASEE Zone 1), 2014 Zone 1 Conference of the*, 2014, pp. 1-5.

[89] M. A. Ramsay, M. Usman, E. Lagow, M. Mendoza, E. Untalan, and E. De Vol, "The Accuracy, Precision and Reliability of Measuring Ventilatory Rate and Detecting Ventilatory Pause by Rainbow Acoustic Monitoring and Capnometry," *Anesth Analg,* Apr 30 2013.

[90] Masimo. (2015, Jan 4). *Acoustic Respiration Rate.* Available: http://www.masimo.com/rra/

[91] L. A. Maynard, "Francis Gano Benedict—a biographical sketch (1870–1957)," *The Journal of nutrition,* vol. 98, pp. 1-8, 1969.

[92] J. Tyndall, *On Radiation: The" Rede" Lecture, Delivered in the Senate-house Before the University of Cambridge on Tuesday, May 16, 1865*: Longman, Green, Longman, Roberts & Green, 1865.

[93] S. Y. Liu, T. S. Lee, and F. Bongard, "Accuracy of capnography in nonintubated surgical patients," *Chest,* vol. 102, pp. 1512-1515, 1992.

[94] R. A. Langer, "Simple Modification of a Medium Concentration (Hudson Type) Oxygen Mask Improves Patient Comfort and Respiratory Monitoring with Capnography," *Anesthesia & Analgesia,* vol. 83, p. 202, 1996.

[95] R. Farre, J. M. Montserrat, M. Rotger, E. Ballester, and D. Navajas, "Accuracy of thermistors and thermocouples as flow-measuring devices for detecting hypopnoeas," *Eur Respir J,* vol. 11, pp. 179-82, Jan 1998.

[96] J. Mead, N. Peterson, G. Grimby, and J. Mead, "Pulmonary ventilation measured from body surface movements," *Science,* vol. 156, pp. 1383-4, Jun 9 1967.

[97] M. A. Cohn, H. Watson, R. Weisshaut, F. Stott, and M. A. Sackner, "A transducer for noninvasive monitoring of respiration," in *Proceedings of the Second International Symposium in Ambulatory Monitoring*, 1978, pp. 119-128.

[98] R. Russell and P. Helms, "Evaluation of three different techniques used to measure chest wall movements in children," *European respiratory journal,* vol. 7, pp. 2073-2076, 1994.

[99] D. E. Weese-Mayer, M. J. Corwin, M. R. Peucker, J. M. Di Fiore, D. R. Hufford, L. R. Tinsley*, et al.*, "Comparison of apnea identified by respiratory inductance plethysmography with that detected by end-tidal CO2 or thermistor," *American Journal of Respiratory and Critical Care Medicine,* vol. 162, pp. 471-480, Aug 2000.

[100] A. f. t. A. o. M. Instrumentation, *Safe current limits for electromedical apparatus*: Association for the Advancement of Medical Instrumentation, 1993.

[101] M. E. Valentinuzzi, L. A. Geddes, and L. E. Baker, "The law of impedance pneumography," *Medical and biological engineering,* vol. 9, pp. 157-163, 1971/05/01 1971.

[102] (2015, 28 Oct). *Sensium Healthcare*. Available: www.sensium-healthcare.com

[103] R. T. Brouillette, A. S. Morrow, D. E. Weese-Mayer, and C. E. Hunt, "Comparison of respiratory inductive plethysmography and thoracic impedance for apnea monitoring," *J Pediatr,* vol. 111, pp. 377-83, Sep 1987.

[104] Y. Yasuda, A. Umezu, S. Horihata, K. Yamamoto, R. Miki, and S. Koike, "Modified thoracic impedance plethysmography to monitor sleep apnea syndromes," *Sleep Medicine,* vol. 6, pp. 215-224, May 2005.

[105] J. E. Lewin, "An apnoea-alarm mattress," *Lancet,* vol. 2, pp. 667-8, Sep 27 1969.

[106] U. Sielaff, "Apnea detection device," US3727606A, 1973.

[107] P. Chow, G. Nagendra, J. Abisheganaden, and Y. T. Wang, "Respiratory monitoring using an air-mattress system," *Physiological Measurement,* vol. 21, pp. 345-354, Aug 2000.

[108] T. Agatsuma, K. Fujimoto, Y. Komatsu, K. Urushihata, T. Honda, T. Tsukahara*, et al.*, "A novel device (SD-101) with high accuracy for screening sleep apnoea-hypopnoea syndrome," *Respirology,* vol. 14, pp. 1143-50, Nov 2009.

[109] M. Holtzman, D. Townsend, R. Goubran, and F. Knoefel, "Validation of pressure sensors for physiological monitoring in home environments," in *Medical Measurements and Applications Proceedings (MeMeA), 2010 IEEE International Workshop on*, 2010, pp. 38-42.

[110] K. Watanabe, T. Watanabe, H. Watanabe, H. Ando, T. Ishikawa, and K. Kobayashi, "Noninvasive measurement of heartbeat, respiration, snoring and body movements of a subject in bed via a pneumatic method," *IEEE Trans Biomed Eng,* vol. 52, pp. 2100-7, Dec 2005.

[111] Hisense. (2014, Jan 4). *Babysense*. Available: http://www.hisense.co.il

[112] J. C. Lin and J. Salinger, "Microwave Measurement of Respiration," in *Microwave Symposium, 1975 IEEE-MTT-S International*, 1975, pp. 285-287.

[113] R. B. Northrop, "Ultrasonic respiration/convulsion monitoring apparatus and method for its use," ed: Google Patents, 1980.

[114] P. Arlotto, M. Grimaldi, R. Naeck, and J. M. Ginoux, "An ultrasonic contactless sensor for breathing monitoring," *Sensors (Basel),* vol. 14, pp. 15371-86, 2014.

[115] H. Kronberg and H. Leist, "Apparatus and method for monitoring respiration," 1993.

[116] V. Hers, D. Corbugy, I. Joslet, P. Hermant, J. Demarteau, B. Delhougne*, et al.*, "New concept using Passive Infrared (PIR) technology for a contactless detection of breathing movement: a pilot study involving a cohort of 169 adult patients," *Journal of Clinical Monitoring and Computing,* vol. 27, pp. 521-529, Oct 2013.

[117] A. K. Abbas, K. Heimann, K. Jergus, T. Orlikowsky, and S. Leonhardt, "Neonatal non-contact respiratory monitoring based on real-time infrared thermography," *Biomed Eng Online,* vol. 10, p. 93, 2011.

[118] A. Aliverti, R. Dellaca, P. Pelosi, D. Chiumello, L. Gatihnoni, and A. Pedoti, "Compartmental analysis of breathing in the supine and prone positions by optoelectronic plethysmography," *Ann Biomed Eng,* vol. 29, pp. 60-70, Jan 2001.

[119] W. Einthoven, G. Fahr, and A. De Waart, "On the direction and manifest size of the variations of potential in the human heart and on the influence of the position of the heart on the form of the electrocardiogram," *American heart journal,* vol. 40, p. 163, 1950.

[120] C. Ludwig, "Beiträge zur Kenntniss des Einflusses der Respirationsbewegungen auf den Blutlauf im Aortensysteme," *Arch. Anat. Physiol,* vol. 13, pp. 242-3027, 1847.

[121] J. A. Hirsch and B. Bishop, "Respiratory sinus arrhythmia in humans: how breathing pattern modulates heart rate," *Am J Physiol,* vol. 241, pp. H620-H629, 1981.

[122] A. E. Santo and C. Carbajal, "Respiration rate extraction from ECG signal via discrete wavelet transform," in *Circuits and Systems for Medical and Environmental Applications Workshop (CASME), 2010 2nd*, 2010, pp. 1-4.

[123] D. Widjaja, J. C. V. Perez, A. C. Dorado, and S. Van Huffel, "An improved ECG-Derived Respiration Method using Kernel Principal Component Analysis," *2011 Computing in Cardiology,* pp. 45-48, 2011.

[124] C. Avci, S. Besli, and A. Akbas, "Performance of the EDR Methods: Evaluations Using the Mean and Instantaneous Respiration Rates," in *Bioinformatics and Biomedical Engineering, (iCBBE) 2011 5th International Conference on*, 2011, pp. 1-5.

[125] M. Clynes, "Respiratory sinus arrhythmia: laws derived from computer simulation," *J Appl Physiol,* vol. 15, pp. 863-74, Sep 1960.

[126] C. T. Davies and J. M. Neilson, "Sinus arrhythmia in man at rest," *Journal of Applied Physiology,* vol. 22, pp. 947-955, 1967.

[127] M. Valentinuzzi and L. Geddes, "The central component of the respiratory heart-rate response," *Cardiovascular Research Center bulletin,* vol. 12, pp. 87-103, 1973.

[128] J. Boyle, N. Bidargaddi, A. Sarela, and M. Karunanithi, "Automatic detection of respiration rate from ambulatory single-lead ECG," *IEEE Trans Inf Technol Biomed,* vol. 13, pp. 890-6, Nov 2009.

[129] S. Dash, K. H. Shelley, D. G. Silverman, and K. H. Chon, "Estimation of respiratory rate from ECG, photoplethysmogram, and piezoelectric pulse transducer signals: a comparative study of time-frequency methods," *IEEE Trans Biomed Eng,* vol. 57, pp. 1099-107, May 2010.

[130] G. K. Pal and P. Pal, *Textbook Of Practical Physiology*, 2 Ed ed.: Orient Longman, 2006.

[131] D. Macintyre and J. Samsoondar, "Spectroscopic method for total hemoglobin measurement," ed: Google Patents, 2009.

[132] P. Leonard, T. F. Beattie, P. S. Addison, and J. N. Watson, "Standard pulse oximeters can be used to monitor respiratory rate," *Emergency Medicine Journal,* vol. 20, pp. 524-525, Nov 1 2003.

[133] D. Wertheim, C. Olden, E. Savage, and P. Seddon, "Extracting respiratory data from pulse oximeter plethysmogram traces in newborn infants," *Arch Dis Child Fetal Neonatal Ed,* vol. 94, pp. F301-3, Jul 2009.

[134] P. S. Addison, J. N. Watson, M. L. Mestek, J. P. Ochs, A. A. Uribe, and S. D. Bergese, "Pulse oximetry-derived respiratory rate in general care floor patients," *J Clin Monit Comput,* May 6 2014.

[135] B. R. Knorr-Chung, S. P. McGrath, and G. T. Blike, "Identifying airway obstructions using photoplethysmography (PPG)," *Journal of clinical monitoring and computing,* vol. 22, pp. 95-101, 2008.

[136] E. Chiner, J. Signes-Costa, J. M. Arriero, J. Marco, I. Fuentes, and A. Sergado, "Nocturnal oximetry for the diagnosis of the sleep apnoea hypopnoea syndrome: a method to reduce the number of polysomnographies?," *Thorax,* vol. 54, pp. 968-71, Nov 1999.

[137] C. L. Tsien and J. C. Fackler, "Poor prognosis for existing monitors in the intensive care unit," *Crit Care Med,* vol. 25, pp. 614-9, Apr 1997.

[138] M. Borowski, S. Siebig, C. Wrede, and M. Imhoff, "Reducing false alarms of intensive care online-monitoring systems: an evaluation of two signal extraction algorithms," *Comput Math Methods Med,* vol. 2011, p. 143480, 2011.

[139] M. J. Tobin, "Respiratory monitoring in the intensive care unit," *American Journal of Respiratory and Critical Care Medicine,* vol. 138, pp. 1625-1642, 1988.

[140] J. B. Downs, "Prevention of hypoxemia: the simple, logical, but incorrect solution," *Journal of clinical anesthesia,* vol. 6, pp. 180-181, May-Jun 1994.

[141] L. J. Findley, A. L. Ries, G. M. Tisi, and P. D. Wagner, "Hypoxemia during apnea in normal subjects: mechanisms and impact of lung volume," *Journal of Applied Physiology,* vol. 55, pp. 1777-1783, 1983.

[142] K. P. Strohl and M. D. Altose, "Oxygen saturation during breath-holding and during apneas in sleep," *CHEST Journal,* vol. 85, pp. 181-186, 1984.

[143] A. Baum, *Cambridge handbook of psychology, health and medicine*: Cambridge University Press, 1997.

[144] T. R. Gravelyn and J. G. Weg, "Respiratory rate as an indicator of acute respiratory dysfunction," *Jama,* vol. 244, pp. 1123-5, Sep 5 1980.

[145] D. Sidebotham, M. R. Dijkhuizen, and S. A. Schug, "The safety and utilization of patient-controlled analgesia," *J Pain Symptom Manage,* vol. 14, pp. 202-9, Oct 1997.

[146] A. f. H. R. a. Quality, "Acute Pain Management - Operative or Medical Procedures and Trauma," *Clinical Pharmacy,* vol. 11, pp. 309-331, Apr 1992.

[147] B. Krieger, D. Feinerman, A. Zaron, and F. Bizousky, "Continuous noninvasive monitoring of respiratory rate in critically ill patients," *Chest,* vol. 90, pp. 632-4, Nov 1986.

[148] M. Hart, M. Orzalesi, and C. Cook, "Relation between anatomic respiratory dead space and body size and lung volume," *J. appl. Physiol,* vol. 18, pp. 519-522, 1963.

[149] M. P. Swanney, G. Ruppel, P. L. Enright, O. F. Pedersen, R. O. Crapo, M. R. Miller*, et al.*, "Using the lower limit of normal for the FEV1/FVC ratio reduces the misclassification of airway obstruction," *Thorax,* September 11, 2008 2008.

[150] C. Zammit, H. Liddicoat, I. Moonsie, and H. Makker, "Obesity and respiratory diseases," *International Journal of General Medicine,* vol. 3, pp. 335-343, 10/20 2010.

[151] M. R. Miller, J. Hankinson, V. Brusasco, F. Burgos, R. Casaburi, A. Coates, *et al.*, "Standardisation of spirometry," *Eur Respir J,* vol. 26, pp. 319-38, Aug 2005.

[152] G. Chen, I. de la Cruz, and E. Rodriguez-Villegas, "Automatic lung tidal volumes estimation from tracheal sounds," in *Engineering in Medicine and Biology Society, 2014 36th Annual International Conference of the IEEE*, 2014, pp. 1497-1500.

[153] R. Berry, R. Brooks, and C. Gamaldo, "The AASM Manual for the scoring of sleep and associated events: rules, terminology and technical specifications. Version 2.0," *Darien, IL: American Academy of Sleep Medicine,* 2012.

[154] R. B. Berry, R. Budhiraja, D. J. Gottlieb, D. Gozal, C. Iber, V. K. Kapur, *et al.*, "Rules for scoring respiratory events in sleep: update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events. Deliberations of the Sleep Apnea Definitions Task Force of the American Academy of Sleep Medicine," *J Clin Sleep Med,* vol. 8, pp. 597-619, Oct 15 2012.

[155] E. Rodriguez-Villegas, G. Chen, J. Radcliffe, and J. Duncan, "A pilot study of a wearable apnoea detection device," *BMJ Open,* vol. 4, October 2014.

[156] D. L. Brown, R. D. Chervin, G. Hegeman, M. A. Smith, N. M. Garcia, L. B. Morgenstern, *et al.*, "Is Technologist Review of Raw Data Necessary after Home Studies for Sleep Apnea?," *Journal of Clinical Sleep Medicine,* vol. 10, pp. 371-375, Apr 15 2014.

[157] M. Ohayon, C. Guilleminault, J. Zulley, L. Palombini, and H. Raab, "Validation of the Sleep-EVAL system against clinical assessments of sleep disorders and polysomnographic data," *Sleep,* vol. 22, pp. 925-930, 1999.

[158] C. A. Nigro, E. Dibur, S. Aimaretti, S. González, and E. Rhodius, "Comparison of the automatic analysis versus the manual scoring from ApneaLink™ device for the diagnosis of obstructive sleep apnoea syndrome," *Sleep and Breathing,* vol. 15, pp. 679-686, 2011.

[159] A. Yadollahi and Z. M. Moussavi, "Acoustical respiratory flow. A review of reliable methods for measuring air flow," *IEEE Eng Med Biol Mag,* vol. 26, pp. 56-61, Jan-Feb 2007.

[160] J. B. West, *Respiratory physiology: the essentials*: Lippincott Williams & Wilkins, 2012.

[161] V. P. Harper, H. Pasterkamp, H. Kiyokawa, and G. R. Wodicka, "Modeling and measurement of flow effects on tracheal sounds," *IEEE Trans Biomed Eng,* vol. 50, pp. 1-10, Jan 2003.

[162] A. Boudewyns, J. Claes, and P. Van de Heyning, "Clinical practice," *European journal of pediatrics,* vol. 169, pp. 135-141, 2010.

[163] A. Yadollahi and Z. Moussavi, "Automatic breath and snore sounds classification from tracheal and ambient sounds recordings," *Medical Engineering & Physics,* vol. 32, pp. 985-990, 11 2010.

[164] C. Pullan and E. Hey, "Wheezing, asthma, and pulmonary dysfunction 10 years after infection with respiratory syncytial virus in infancy," *BMJ,* vol. 284, pp. 1665-1669, 1982.

[165] K. O. Devaney, D. J. LaFeir, A. Triantafyllou, W. M. Mendenhall, J. A. Woolgar, A. Rinaldo*, et al.*, "Inflammatory myofibroblastic tumors of the head and neck: evaluation of clinicopathologic and prognostic features," *European Archives of Oto-Rhino-Laryngology,* vol. 269, pp. 2461-2465, 2012.

[166] P. Forgacs, "The functional basis of pulmonary sounds," *Chest,* vol. 73, pp. 399-405, Mar 1978.

[167] J. P. Metlay, W. N. Kapoor, and M. J. Fine, "Does this patient have community-acquired pneumonia?: Diagnosing pneumonia by history and physical examination," *Jama,* vol. 278, pp. 1440-1445, 1997.

[168] T. Kaisla, H. Rajala, and T. Katila, "Crackles in patients with fibrosing alveolitis, bronchiectasis, COPD, and heart failure," *CHEST Journal,* vol. 99, pp. 1076-1083, 1991.

[169] (2015). *EarKandie*. Available: http://www.earkandie.com

[170] *Sonoor Jewelry Concepts*. Available: http://www.sonoor.com

[171] F. Gemperle, C. Kasabach, J. Stivoric, M. Bauer, and R. Martin, "Design for wearability," in *Wearable Computers, 1998. Digest of Papers. Second International Symposium on*, 1998, pp. 116-122.

[172] L. E. Dunne, H. Profita, C. Zeagler, J. Clawson, S. Gilliland, E. Y. Do*, et al.*, "The social comfort of wearable technology and gestural interaction," *Conf Proc IEEE Eng Med Biol Soc,* vol. 2014, pp. 4159-62, 2014.

[173] M. Mussell, Y. Nakazono, and Y. Miyamoto, "Effect of air flow and flow transducer on tracheal breath sounds," *Medical and Biological Engineering and Computing,* vol. 28, pp. 550-554, 1990.

[174] N. Gavriely and D. W. Cugell, "Airflow effects on amplitude and spectral content of normal breath sounds," *Journal of applied physiology,* vol. 80, p. 5, 1996.

[175] H. Pasterkamp, S. Patel, and G. Wodicka, "Asymmetry of respiratory sounds and thoracic transmission," *Medical and Biological Engineering and Computing,* vol. 35, pp. 103-106, 1997.

[176] P. Hult, T. Fjallbrant, S. Dahle, P. Danielsson, and P. Ask, "A method for respiration monitoring by the use of a bioacousticsignal," 2000, pp. 22-25.

[177] Z. Moussavi, M. Leopando, H. Pasterkamp, and G. Rempel, "Computerised acoustical respiratory phase detection without airflow measurement," *Medical and Biological Engineering and Computing,* vol. 38, pp. 198-203, 2000.

[178] V. Harper, H. Pasterkamp, H. Kiyokawa, and G. Wodicka, "Modeling and measurement of flow effects on tracheal sounds," *Biomedical Engineering, IEEE Transactions on,* vol. 50, pp. 1-10, 2003.

[179] P. Corbishley and E. Rodriguez-Villegas, "Breathing Detection: Towards a Miniaturized, Wearable, Battery-Operated Monitoring System," *Biomedical Engineering, IEEE Transactions on,* vol. 55, pp. 196-204, 2008.

[180] F. Dalmasso, M. M. Guarene, R. Spagnolo, G. Benedetto, and G. Righini, "A computer system for timing and acoustical analysis of crackles: a study in cryptogenic fibrosing alveolitis," *Bull Eur Physiopathol Respir,* vol. 20, pp. 139-44, Mar-Apr 1984.

[181] N. Meslier, G. Charbonneau, and J. L. Racineux, "Wheezes," *Eur Respir J,* vol. 8, pp. 1942-8, Nov 1995.

[182] G. Chen, "Wearable miniature bionic apnoea detection modelling," MSc. Thesis, Electrical & Electronic Engineering, Imperial College London, London, 2007.

[183] I. de la Cruz, "Real Time Acoustic Monitoring of Airflow Ventilation Volume," MEng Report, Electrical & Electronic Engineering, Imperial College London, London, 2010.

[184] U. Gröger and L. Wiegrebe, "Classification of human breathing sounds by the common vampire bat, Desmodus rotundus," *BMC Biology,* vol. 4, p. 18, 2006.

[185] S. Huq, A. Yadollahi, and Z. Moussavi, "Breath analysis of respiratory flow using tracheal sounds," *2007 Ieee International Symposium on Signal Processing and Information Technology, Vols 1-3,* pp. 759-763, 2007.

[186] M. J. Banner, J. B. Downs, R. R. Kirby, R. A. Smith, P. G. Boysen, and S. Lampotang, "EFfects of expiratory flow resistance on inspiratory work of breathing," *Chest,* vol. 93, pp. 795-799, 1988.

[187] H. Pasterkamp, S. S. Kraman, P. D. DeFrain, and G. R. Wodicka, "Measurement of respiratory acoustical signals. Comparison of sensors," *Chest,* vol. 104, pp. 1518-1525, 1993.

[188] M. Specialties. (2015, 28 Oct). *CM-01B contact microphone.* Available: http://www.meas-spec.com/downloads/Contact_Microphone.pdf

[189] P. Chakraborty, M. Kapoor, and K. Pavate, "Electret Condenser Microphone," *IETE Journal of Research,* vol. 21, pp. 383-385, 1975.

[190] F. Sandoval and E. Gutiérrez, "Analysis of a condenser-microphone using general-purpose circuit simulation programs," *Instrumentation and Development,* vol. 3, 1997.

[191] A. V. Rhijn, "Integrated Circuits for High Performance Electret Microphones," *The Audio Engineering Society,* p. 5719, 2003.

[192] Kingstate. (2015, 28 Oct). *Electret condenser microphone KEEG1542PBL-A.* Available: http://onecall.farnell.com/kingstate/keeg1542pbl-a/microphone-electret-condenser/dp/1502746

[193] P. J. Henning and T. Storgaard - Larsen, "MEMS condenser microphone," *The Journal of the Acoustical Society of America,* vol. 108, pp. 2495-2495, 2000.

[194] P. V. Loeppert and M. Pederson, "Miniature broadband acoustic transducer," ed: Google Patents, 2003.

[195] P. R. Scheeper, W. Olthuis, and P. Bergveld, "Fabrication of a subminiature silicon condenser microphone using the sacrificial layer technique," in *Solid-State Sensors and Actuators, 1991. Digest of Technical Papers, TRANSDUCERS '91., 1991 International Conference on*, 1991, pp. 408-411.

[196] K. Acoustics, "SiSonic Design Guide," 20 Apr 2006.

[197] Knowles. (2015, 28 Oct). *Surface mount MEMS microphone SPA2629LR5H.* Available: http://www.knowles.com/eng/Products/Microphones/Surface-mount-MEMS

[198] G. Wodicka, S. Kraman, G. Zenk, and H. Pasterkamp, "Measurement of respiratory acoustic signals. Effect of microphone air cavity depth," *Chest,* vol. 106, pp. 1140-4, 1994.

[199] S. Kraman, G. Wodicka, Y. Oh, and H. Pasterkamp, "Measurement of respiratory acoustic signals. Effect of microphone air cavity width, shape, and venting," *Chest,* vol. 108, pp. 1004-1008, 1995.

[200] M. Zanartu, J. Ho, S. Kraman, H. Pasterkamp, J. Huber, and G. Wodicka, "Air-borne and tissue-borne sensitivity of skin-radiation acoustic sensors," *The Journal of the Acoustical Society of America,* vol. 123, 2008.

[201] E. Aguilar-Pelaez, "Real time algorithms for acoustic heart rate detection and respiratory rate extraction for use in miniature wearable breathing and heart monitor," PhD. thesis, 2010.

[202] G. Chen, E. Rodriguez-Villegas, and A. J. Casson, "Wearable Algorithms: An Overview of a Truly Multi-Disciplinary Problem," in *Wearable Sensors*, E. Sazonov and M. R. Neuman, Eds., 1st ed Oxford: Academic Press, 2014, pp. 353-382.

[203] G. Chen and E. Rodriguez-Villegas, "System-level design trade-offs for truly wearable wireless medical devices," in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, 2010, pp. 1441-1444.

[204] Duracell. (2015, June). Zinc Air Battery DA675 Datasheet. Available: http://www.farnell.com/datasheets/6247.pdf

[205] Duracell. (2015, June). Zinc Air Battery DA13 Datasheet. Available: http://www.farnell.com/datasheets/6248.pdf

[206] X. Energy. (2015, June). Thionyl Chloride Lithium Battery XL-050F Specifications. Available: http://www.xenousa.com/pdf/XL-050F.pdf

[207] Renata. (2015, June). 3V Lithium Battery CR2430 Datasheet. Available: http://www.renata.com/fileadmin/downloads/productsheets/lithium/3V_lithium/CR2430_v07.pdf

[208] T. Instruments. (2010, June). SimpliciTI: Simple Modular RF Network Specification. Available: http://focus.ti.com/docs/toolsw/folders/print/simpliciti.html

[209] T. Instruments. (2010, June). CC2500 Low-Cost Low-Power 2.4 GHz RF Transceiver Datasheet. Available: http://focus.ti.com/lit/ds/symlink/cc2500.pdf

[210] N. Semiconductor. (2010, June). nRF24L01+ Single Chip 2.4GHz Transceiver Product Specification. Available: http://www.nordicsemi.com/files/Product/data_sheet/nRF24L01P_Product_Specification_1_0.pdf

[211] A. Rüst, M. Gysin, and A. Müller, "Simultaneously Connecting Devices through Bluetooth Smart," presented at the embedded world Conference, 2014.

[212] D. Salomon, "Audio Compression," in *Data Compression*, ed, 2007, pp. 719-850.

[213] A. Yadollahi and Z. Moussavi, "Respiratory Sounds Compression," *Biomedical Engineering, IEEE Transactions on,* vol. 55, pp. 1336-1343, 2008.

[214] A. Kulkas, E. Rauhala, E. Huupponen, J. Virkkala, M. Tenhunen, A. Saastamoinen*, et al.*, "Detection of compressed tracheal sound patterns with large amplitude variation during sleep," *Med Biol Eng Comput,* vol. 46, pp. 315-21, Apr 2008.

[215] T. Equipments, "40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)," *ITU-T Recommendation, G,* vol. 726, p. 59, 1990.

[216] T. Robinson, "Shorten: Simple lossless and near lossless waveform compression," Cambridge University Engineering Department, Technique Report, 1994.

[217] R. F. Rice and P. J.R., "Adaptive Variable-Length Coding for Efficient Compression of Spacecraft Television Data," *Communication Technology, IEEE Transactions on,* vol. 19, pp. 889-897, 1971.

[218] P. Yeh, R. Rice, and W. Miller, "On the optimality of a universal noiseless coder," in *AIAA Computing in Aerospace Conference, 9th*, San Diego, CA, 1993, pp. 490-498.

[219] R. F. Rice, "Some practical noiseless coding techniques, Part II," *Jet Propulsion Laboratories Publication,* vol. 91, 1991.

[220] T. Liebchen, "MPEG-4 ALS-The Standard for Lossless Audio Coding," *the Journal of the Acoustical Society of Korea,* vol. 28, pp. 618-629, 2009.

[221] D. S. Kim and J. S. Kwon, "A lossless multichannel bio-signal compression based on low-complexity joint coding scheme for portable medical devices," *Sensors (Basel),* vol. 14, pp. 17516-29, 2014.

[222] D. Salomon, *Curves And Surfaces for Computer Graphics*. New York: Springer, 2006.

[223] H. S. Malvar, "Adaptive run-length/Golomb-Rice encoding of quantized generalized Gaussian sources with unknown statistics," in *Data Compression Conference, 2006. DCC 2006. Proceedings*, 2006, pp. 23-32.

[224] P. G. Howard and J. S. Vitter, "Fast and efficient lossless image compression," in *Data Compression Conference, 1993. DCC '93.*, 1993, pp. 351-360.

[225] Y. A. Reznik, "Coding of prediction residual in MPEG-4 standard for lossless audio coding (MPEG-4 ALS)," *2004 Ieee International Conference on Acoustics, Speech, and Signal Processing, Vol Iii, Proceedings,* pp. 1024-1027, 2004.

[226] T. Liebchen. (2015, 25 Oct). *MPEG-4 ALS reference software*. Available: http://www.nue.tu-berlin.de/menue/forschung/projekte/beendete_projekte/mpeg-4_audio_lossless_coding_als/parameter/en/

[227] H. Huang, H. Shu, and R. Yu, "Lossless audio compression in the new IEEE Standard for Advanced Audio Coding," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, 2014, pp. 6934-6938.

[228] T. Instruments. (2015, CC2640 SimpleLink ultra-low power wireless MCU for Bluetooth Smart Datasheet. Available: http://www.ti.com/lit/gpn/cc2640

[229] S. L. Himanen, E. Huupponen, A. Kulkas, and E. Rauhala, "Compressed tracheal sound analysis as a screening method for obstructive apneas and hypopneas," in *Sleep Medicine*, 2007, p. S91.

[230] E. Rauhala, J. Hasan, A. Kulkas, A. Saastamoinen, E. Huupponen, F. Cameron, *et al.*, "Compressed tracheal sound analysis in screening of sleep-disordered breathing," *Clin Neurophysiol,* vol. 119, pp. 2037-43, Sep 2008.

[231] R. Palaniappan, K. Sundaraj, and N. U. Ahamed, "Machine learning in lung sound analysis: A systematic review," *Biocybernetics and Biomedical Engineering,* vol. 33, pp. 129-135, 2013.

[232] R. Palaniappan, K. Sundaraj, and S. Sundaraj, "A comparative study of the SVM and K-nn machine learning algorithms for the diagnosis of respiratory pathologies using pulmonary acoustic signals," *BMC Bioinformatics,* vol. 15, p. 223, 2014.

[233] H. Alshaer, A. Pandya, T. D. Bradley, and F. Rudzicz, "Subject independent identification of breath sounds components using multiple classifiers," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, 2014, pp. 3577-3581.

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

[235] A. Yadollahi and Z. Moussavi, "A robust method for estimating respiratory flow using tracheal sounds entropy," *Biomedical Engineering, IEEE Transactions on,* vol. 53, pp. 662-668, 2006.

[236] H. Alshaer, G. R. Fernie, and T. D. Bradley, "Phase tracking of the breathing cycle in sleeping subjects by frequency analysis of acoustic data," *International Journal of Healthcare Technology and Management,* vol. 11, pp. 163-175, 2010.

[237] A. Yadollahi and Z. Moussavi, "Automatic breath and snore sounds classification from tracheal and ambient sounds recordings," *Med Eng Phys,* vol. 32, pp. 985-90, Nov 2010.

[238] S. Le Cam, C. Collet, and F. Salzenstein, "Acoustical respiratory signal analysis and phase detection," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, 2008, pp. 3629-3632.

[239] F. Jin, F. Sattar, and D. Y. Goh, "An acoustical respiratory phase segmentation algorithm using genetic approach," *Medical & biological engineering & computing,* vol. 47, pp. 941-953, 2009.

[240] R. Palaniappan, K. Sundaraj, S. Sundaraj, N. Huliraj, and S. S. Revadi, "A novel approach to detect respiratory phases from pulmonary acoustic signals using normalised power spectral density and fuzzy inference system," *The Clinical Respiratory Journal,* pp. n/a-n/a, 2015.

[241] Z. Moussavi, M. Leopando, and G. Rempel, "Automated detection of respiratory phases by acoustical means," 1998, pp. 21-24.

[242] P. Hult, T. Fjallbrant, S. Dahle, P. Danielsson, and P. Ask, "A method for respiration monitoring by the use of a bioacoustic signal," 2000, pp. 22-25.

[243] A. Yadollahi and Z. Moussavi, "Apnea Detection by Acoustical Means," in *28th IEEE EMBS Annual International Conference*, New York City, USA, 2006.

[244] A. Yadollahi and Z. Moussavi, "A robust method for heart sounds localization using lung sounds entropy," *IEEE Transactions Biomedical Engineering,* vol. 53, pp. 497-502, 2006.

[245] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *Systems, Man and Cybernetics, IEEE Transactions on,* vol. 9, pp. 62-66, 1979.

[246] A. Yadollahi, A. Azarbarzin, A. Montazeri, and Z. Moussavi, "Acoustical flow estimation in patients with obstructive sleep apnea during Sleep," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, 2012, pp. 3640-3643.

[247] A. Azarbarzin and Z. Moussavi, "Unsupervised classification of respiratory sound signal into snore/no-snore classes," in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, 2010, pp. 3666-3669.

[248] A. Azarbarzin and Z. Moussavi, "Automatic and Unsupervised Snore Sound Extraction From Respiratory Sound Signals," *Biomedical Engineering, IEEE Transactions on,* vol. 58, pp. 1156-1162, 2011.

[249] A. Azarbarzin and Z. Moussavi, "Nonlinear properties of snoring sounds," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, 2011, pp. 4316-4319.

[250] A. Azarbarzin and Z. Moussavi, "Snoring sounds' statistical characteristics depend on anthropometric parameters," *Journal of Biomedical Science and Engineering,* vol. 5, p. 245, 2012.

[251] A. Kulkas, E. Huupponen, J. Virkkala, M. Tenhunen, A. Saastamoinen, E. Rauhala*, et al.*, "New tracheal sound feature for apnoea analysis," *Med Biol Eng Comput,* vol. 47, pp. 405-12, Apr 2009.

[252] M. De Burgh Daly, "Interactions Between Respiration and Circulation," in *Comprehensive Physiology*, ed: John Wiley & Sons, Inc., 2011.

[253] 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. 951-7, Aug 1 2004.

[254] C. M. Serrano, "Calibration, testing & optimisation of a state-of-the-art apnoea detector," MSc. Thesis, Electrical & Electronic Engineering, Imperial College London, London, 2006.

[255] E. Rodriguez-Villegas, G. Chen, and E. Aguilar-Pelaez, "Feature Characterization for Breathing Monitor," WO/2012/007719, 19 Jan, 2012.

[256] P. G. Corbishley, "Ultra low power circuits for a miniature apnoea detection device," Ph.D. Thesis, Electrical & Electronic Engineering, Imperial College London, London, 2007.

[257] P. Corbishley and E. Rodriguez-Villegas, "A Nanopower Bandpass Filter for Detection of an Acoustic Signal in a Wearable Breathing Detector," *Biomedical Circuits and Systems, IEEE Transactions on,* vol. 1, pp. 163-171, 2007.

[258] P. Corbishley and E. Rodriguez-Villegas, "A Low Power Low Voltage Rectifier Circuit," in *Circuits and Systems, 2006. MWSCAS '06. 49th IEEE International Midwest Symposium on*, 2006, pp. 512-515.

[259] E. Rodriguez-Villegas, A. J. Casson, and P. Corbishley, "A Subhertz Nanopower Low-Pass Filter," *Circuits and Systems II: Express Briefs, IEEE Transactions on,* vol. 58, pp. 351-355, 2011.

[260] D. Zhang, A. Bhide, and A. Alvandpour, "A 53-nW 9.1-ENOB 1-kS/s SAR ADC in 0.13-m CMOS for Medical Implant Devices," *Ieee Journal of Solid-State Circuits,* vol. 47, pp. 1585-1593, Jul 2012.

[261] A. Becker-Gomez, T. L. Viswanathan, and T. R. Viswanathan, "A low-supply-voltage CMOS sub-bandgap reference," *Ieee Transactions on Circuits and Systems Ii-Express Briefs,* vol. 55, pp. 609-613, Jul 2008.

[262] T. Hui Teo, L. Gin Kooi, D. Sutomo, T. Kuo Hwi, P. K. Gopalakrishnan, and R. Singh, "Ultra Low-Power Sensor Node for Wireless Health Monitoring System," in *Circuits and Systems, 2007. ISCAS 2007. IEEE International Symposium on*, 2007, pp. 2363-2366.

[263] R. B. Berry, R. Budhiraja, D. J. Gottlieb, D. Gozal, C. Iber, V. K. Kapur*, et al.*, "Rules for Scoring Respiratory Events in Sleep: Update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events: Deliberations of the Sleep Apnea Definitions Task Force of the American Academy of Sleep Medicine," *Journal of Clinical Sleep Medicine : JCSM : Official Publication of the American Academy of Sleep Medicine,* vol. 8, pp. 597-619, 10/15 2012.

[264] (2015, 28 Oct). *The price of fatigue: the surprising economic costs of unmanaged sleep apnea*. Available: https://sleep.med.harvard.edu/what-we-do/public-education

[265] S. Redline, V. K. Kapur, M. H. Sanders, S. F. Quan, D. J. Gottlieb, D. M. Rapoport*, et al.*, "Effects of varying approaches for identifying respiratory disturbances on sleep apnea assessment," *Am J Respir Crit Care Med,* vol. 161, pp. 369-74, Feb 2000.

[266] C. A. Nigro and E. E. Rhodius, "Effect of the definition of hypopnea on apnea/hypopnea index," *Medicina (B Aires),* vol. 63, pp. 119-24, 2003.

[267] W. R. Ruehland, P. D. Rochford, F. J. O'Donoghue, R. J. Pierce, P. Singh, and A. T. Thornton, "The new AASM criteria for scoring hypopneas: impact on the apnea hypopnea index," *Sleep,* vol. 32, pp. 150-7, Feb 2009.

[268] S. Reichert, R. Gass, C. Brandt, and E. Andres, "Analysis of respiratory sounds: state of the art," *Clin Med Circ Respirat Pulm Med,* vol. 2, pp. 45-58, 2008.

Glossary of Terms

AASM	American Academy of Sleep Medicine
AC	Alternating current
ADC	Analog-to-digital converter
ADPCM	Adaptive differential pulse code modulation
ALS	Audio Lossless Coding
Apnoea	Cessation of respiration for more than 10 seconds
ANN	Artificial neural networks
Auto-Ack	Automatic acknowledgment
BC	Breath candidate
BER	Bit error rate
BL	Breath-like
BLE	Bluetooth Low Energy, formal name of Bluetooth Smart
bpm	Breaths per minute, where 1 breaths include both inhalation and exhalation
Breath candidate segment	Suspected breath event detected from the first stage of breath verification process which is also referred as breath discriminator.
Breath Rate / Respiratory Rate (RR)	The number of inhalation-exhalation cycles taken within a set amount of time
Breath segment / Breath event / Respiratory phase	Herein it refers to either inspiration or expiration phase of a full respiratory cycle

Breath-like Segment	Suspected breath event detected from the early stage of breath detection system which is also referred as segmentation algorithm.
CH	Channel
COPD	Chronic obstructive pulmonary disease
dB SPL	Logarithmic measure of a sound pressure relative to the minimum sound pressure threshold of human hearing
DC	Direct current
Dead space	Dead space is the volume required to ensure fresh air not just staying in the respiratory airway, but sufficient oxygen dispersing into the lung.
ECG / EKG	Electrocardiography
EDR	Electrocardiogram derived respiration
EEG	Electroencephalography
EIP	Electrical impedance pneumography
EWS	Early warning scoring
FEC	Forward error correction
FEV	Forced expiratory volume
FIR	Finite impulse response
FLASS	Fixed length adaptive sample size
FVC	Forced vital capacity
GAs	Genetic algorithms
GMM	Gaussian Mixture Model
ICU	Intensive Care Unit
IIR	Infinite impulse response
k-NN	k-Nearest Neighbour algorithm
MEMS	Microelectromechanical systems
MPEG	Moving Picture Experts Group

PCB	Printed circuit board
PER	Packet error rate
PLL	Phase lock loop
Polysomnography	A multi-parametric sleep study/diagnostic approach which invokes comprehensive physiological signal recording.
PPG	Photoplethysmogram, which is a plethysmogram (in term of blood volume changes) detected by an optical approach. This is normally performed by a pulse oximeter.
PSRR	Power supply rejection ratio
QOL	Quality of link
Respiratory pause	Cessation of breathing for less than 10 seconds
RF	Radio frequency
RIP	Respiratory inductance plethysmography
RSA	Respiratory sinus arrhythmia, which is heart rate variability in synchrony with respiration
RTP	Respiratory transthoracic plethysmography
SaO_2	saturation of arterial oxygen
SAR	Successive approximation register
SIDS	Sudden infant death syndrome
SNR	Signal to noise ratio
SPI	Serial peripheral interface
SpO_2	Saturation of peripheral oxygen
SUDEP	Sudden unexpected death in epilepsy
SVM	Support Vector Machine
WADD	It is short for Wearable Apnoea Detection Device, which is the code name of the wireless respiratory monitoring sensor prototype used in the sleep apnoea clinical trial.

95% Confidence interval (CI)	Normally it refers to 95% confidence level which means a 0.95 probability of containing the population mean