Multichannel analysis of normal and continuous adventitious respiratory sounds for the assessment of pulmonary function in respiratory diseases

Doctoral thesis

Manuel Lozano García

Doctoral programme in Biomedical Engineering
Department of Automatic Control
Universitat Politècnica de Catalunya

Biomedical Signal Processing and Interpretation group
Institute for Bioengineering of Catalonia

Director: Dr. Raimon Jané Campos

Barcelona, Spain
May 2016

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Escola de Doctorat

Acta de qualificació de tesi doctoral

Curs acadèmic: 2015/2016

Nom i cognoms
Manuel Lozano García

Programa de doctorat
Enginyeria Biomèdica

Unitat estructural responsable del programa
Departament ESAII, Universitat Politècnica de Catalunya

Resolució del Tribunal

Reunit el Tribunal designat a l'efecte, el doctorand / la doctoranda exposa el tema de la seva tesi doctoral titulada Multichannel analysis of normal and continuous adventitious respiratory sounds for the assessment of pulmonary function in respiratory diseases.

Acabada la lectura i després de donar resposta a les qüestions formulades pels membres titulars del tribunal, aquest atorga la qualificació:

☐ NO APTE ☐ APROVAT ☐ NOTABLE ☐ EXCEL·LENT

(Nom, cognoms i signatura)	(Nom, cognoms i signatura)
President/a	Secretari/ària

(Nom, cognoms i signatura)	(Nom, cognoms i signatura)	(Nom, cognoms i signatura)
Vocal	Vocal	Vocal

_____, _____ d'/de _____ de _____

El resultat de l'escrutini dels vots emesos pels membres titulars del tribunal, efectuat per l'Escola de Doctorat, a instància de la Comissió de Doctorat de la UPC, atorga la MENCIÓ CUM LAUDE:

☐ SÍ ☐ NO

(Nom, cognoms i signatura)	(Nom, cognoms i signatura)
President de la Comissió Permanent de l'Escola de Doctorat	Secretari de la Comissió Permanent de l'Escola de Doctorat

Barcelona, _____ d'/de _____ de _____

Diligència "Internacional del títol de doctor o doctora"

- Com a secretari/ària del tribunal faig constar que la tesi s'ha defensat en part, i com a mínim pel que fa al resum i les conclusions, en una de les llengües habituals per a la comunicació científica en el seu camp de coneixement i diferent de les que són oficials a Espanya. Aquesta norma no s'aplica si l'estada, els informes i els experts provenen d'un país de parla hispana.

(Nom, cognoms i signatura)
Secretari/ària del Tribunal

Abstract

Respiratory sounds (RS) are produced by turbulent airflows through the airways and are inhomogeneously transmitted through different media to the chest surface, where they can be recorded in a non-invasive way. Due to their mechanical nature and airflow dependence, RS are affected by respiratory diseases that alter the mechanical properties of the respiratory system. Therefore, RS provide useful clinical information about the respiratory system structure and functioning.

Recent advances in sensors and signal processing techniques have made RS analysis a more objective and sensitive tool for measuring pulmonary function. However, RS analysis is still rarely used in clinical practice. Lack of a standard methodology for recording and processing RS has led to several different approaches to RS analysis, with some methodological issues that could limit the potential of RS analysis in clinical practice (i.e., measurements with a low number of sensors, no controlled airflows, constant airflows, or forced expiratory manoeuvres, the lack of a co-analysis of different types of RS, or the use of inaccurate techniques for processing RS signals).

In this thesis, we propose a novel integrated approach to RS analysis that includes a multichannel recording of RS using a maximum of five microphones placed over the trachea and the chest surface, which allows RS to be analysed at the most commonly reported lung regions, without requiring a large number of sensors. Our approach also includes a progressive respiratory manoeuvres with variable airflow, which allows RS to be analysed depending on airflow. Dual RS analyses of both normal RS and continuous adventitious sounds (CAS) are also proposed. Normal RS are analysed through the RS intensity–airflow curves, whereas CAS are analysed through a customised Hilbert spectrum (HS), adapted to RS signal characteristics.

The proposed HS represents a step forward in the analysis of CAS. Using HS allows CAS to be fully characterised with regard to duration, mean frequency, and intensity. Further, the high temporal and frequency resolutions, and the high concentrations of energy of this improved version of HS, allow CAS to be more accurately characterised with our HS than by using spectrogram, which has been the most widely used technique for CAS analysis.

Our approach to RS analysis was put into clinical practice by launching two studies in the Pulmonary Function Testing Laboratory of the Germans Trias i Pujol University Hospital for assessing pulmonary function in patients with unilateral phrenic paralysis (UPP), and bronchodilator response (BDR) in patients with asthma. RS and airflow signals were recorded in 10 patients with UPP, 50 patients with asthma, and 20 healthy participants.

The analysis of RS intensity–airflow curves proved to be a successful method to detect UPP, since we found significant differences between these curves at the posterior base of the lungs in all patients,

whereas no differences were found in the healthy participants. To the best of our knowledge, this is the first study that uses a quantitative analysis of RS for assessing UPP.

Regarding asthma, we found appreciable changes in the RS intensity–airflow curves and CAS features after bronchodilation in patients with negative BDR in spirometry. Therefore, we suggest that the combined analysis of RS intensity–airflow curves and CAS features—including number, duration, mean frequency, and intensity—seems to be a promising technique for assessing BDR and improving the stratification of BDR levels, particularly among patients with negative BDR in spirometry.

The novel approach to RS analysis developed in this thesis provides a sensitive tool to obtain objective and complementary information about pulmonary function in a simple and non-invasive way. Together with spirometry, this approach to RS analysis could have a direct clinical application for improving the assessment of pulmonary function in patients with respiratory diseases.

Keywords

Asthma, continuous adventitious sounds, ensemble empirical mode decomposition, Hilbert-Huang transform, instantaneous frequency, multichannel respiratory sound recording, respiratory sounds, unilateral phrenic paralysis, wheezes

“Patience is to science what insanity is to art”

However, one may think that “Insanity is to science what patience is to art” is also true.

So, keep insanely on it and personal success will come to you.

Acknowledgments

No puc començar els meus agraïments d'una altra manera més que donant les gràcies al meu director de tesi, Raimon Jané, per brindar-me l'oportunitat, ja fa més de cinc anys, de formar part del grup de Processament i Interpretació de Senyals Biomèdics (BIOSPIN) de l'Institut de Bioenginyeria de Catalunya (IBEC). Gràcies per haver confiat en mi per iniciar i formar part de la Unitat Mixta amb l'Institut de Recerca Germans Trias i Pujol (IGTP) durant aquests anys. Gràcies per l'esforç que has fet i el suport que m'has donat en els moments crítics d'aquests anys, sobretot durant els dos darrers anys. Em quedo amb la teva experiència, els teus coneixements i amb les llargues i profitoses discussions científiques que hem tingut i que han sigut una peça clau en la realització d'aquesta tesi. Però, sobretot, em quedo amb la teva persona, els teus consells i el teu seny, els quals han contribuït a que prengués la millor decisió en cada moment.

También quiero dar las gracias a José Antonio Fiz, neumólogo del Hospital Germans Trias i Pujol (HUGTiP). Visionario, entusiasta, rebelde e impetuoso. De ti me quedo con tu carácter polifacético, tu pasión por la ciencia, tus ganas de aprender e innovar constantemente, tus buenas ideas y consejos que, junto a los de Raimon, han contribuido a que los estudios presentados en esta tesis llegasen a buen puerto. De ti también me quedo con tu amistad, tus reflexiones personales, tu filosofía y tu gusto por la buena música. ¡Los viejos rockeros nunca mueren! Sobre todo si toman agua de Lanjarón…

También quisiera dar las gracias a Miquel Àngel Gassull y Manel Puig, antiguo y actual director del IGTP, por aceptar mi contratación en el IGTP, en el marco de la Unidad Mixta con el IBEC, durante cuatro años y medio. Igualmente, agradezco a Josep Morera y Juan Ruiz, antiguo y actual jefe del servicio de neumología del HUGTiP, el apoyo que me han ofrecido durante estos años.

Gracias Rosa y Aurora, enfermeras de la Unidad de Pruebas Funcionales Respiratorias del HUGTiP, por vuestra ayuda en el reclutamiento de pacientes y en la realización de los test espirométricos. Sin vuestra paciencia y dedicación habría sido imposible llevar a cabo las sesiones de registro de sonidos respiratorios en el hospital. Aprovecho también la ocasión para dar las gracias a todos los pacientes que, de forma voluntaria, aceptaron participar en los estudios de acústica pulmonar.

No puedo olvidarme de todos mis compañeros del IGTP y del IBEC. Gracias Txell y Edurne por vuestra compañía en el despacho del IGTP y por vuestros buenos deseos. Gracias a todos los compañeros del grupo BIOSPIN del IBEC, tanto a los que me han acompañado desde el principio (Leo, Puy, Oiane, Luís, Bea, Abel y Jordi) como a los que han coincidido conmigo en algún momento (Ainara, Christian, Andrés, Sara, Javier, Joan, Mireia,…) por las comidas, cafés, cenas, congresos y tantos buenos momentos compartidos durante estos años. Gracias por haber contribuido a crear buen rollo y buen ambiente de trabajo en el laboratorio y gracias, también, por vuestros consejos y por los debates científicos tan interesantes que hemos tenido. No sé si habremos conseguido arreglar el mundo durante

los cafés de media mañana, pero ¡buenas ideas de negocio hemos tenido! Me quedo, ante todo, con lo buenas personas que sois. Espero poder compartir todavía buenos momentos con muchos de vosotros.

Quiero agradecer también los buenos momentos compartidos con compañeros de otros grupos del IBEC, los que conocí en el Máster de Ingeniería Biomédica y los que he conocido en las clases de inglés, los seminarios y otras actividades que se han organizado.

I would like to warmly thank all of the members of the Respiratory Muscle Laboratory at King's College London, for their enthusiastic welcome at my arrival in London, and their support and patience during the past six months. I hope we will share great moments during the next months, now that I am free!

También quiero dar las gracias a todos mis buenos amigos, por estar siempre ahí y por poder disfrutar de vuestra compañía, aunque no pueda ser tan a menudo como nos gustaría. En especial, gracias José por pasar toda una vida juntos, siempre incondicional. Gracias Javi, el hombre de la acústica musical, no pulmonar, y Jona, el hombre de los sueños e ilusiones, por pasar casi toda una vida juntos. Gracias Eddie por recordarme lo que es el auténtico espíritu científico (eres un crack). Gracias Carlos porque, aunque sé que te tienen muy ocupado (por cierto, felizmente ocupado), también sé que siempre puedo contar contigo. Gracias Jessi, admiro tu capacidad de esfuerzo y sacrificio. Gracias Núria, nunca pierdas tu espíritu aventurero. Gracias Inma, admiro tu valentía y tu optimismo ante las adversidades.

Quiero dedicar un agradecimiento especial a mis padres y a mi hermano. Gracias a mis padres por su apoyo incondicional y por creer en mí, por enseñarme que, con esfuerzo y sacrificio, puedo conseguir todo lo que quiera, por dejarme decidir con total libertad aquello a lo que me he querido dedicar y por apoyarme siempre, en cada uno de los pasos que he dado a lo largo de mi vida, sabiendo que era lo mejor para mí, aunque a veces haya supuesto alejarme un poco de ellos. Gracias también a mi hermano, por ayudarme y apoyarme en las decisiones importantes. Gracias por hacerme saber que, haga lo que haga y esté donde esté, estáis siempre conmigo.

Por último, mil gracias a ti, Noe, como me gusta decir, mi compañera de vida. Tú mejor que nadie has vivido esta etapa conmigo. Gracias por tu paciencia, por aguantar mis momentos de locura y subsanarlos con tus dosis de cordura, por tu constante ánimo, por creer plenamente en mí, por tu inmenso cariño y, sobre todo, por tu comprensión, tu apoyo y por la libertad que me has dado a la hora de tomar decisiones que nos incumbían a los dos. Gracias por estar en Londres conmigo. Gracias por hacerme el día a día mucho más fácil. Gracias por compartir tu vida conmigo. En definitiva, gracias por darme tantos momentos de felicidad.

Gracias a todos los mencionados y a los que, de una manera u otra, han contribuido a hacer de esta etapa, no sólo una gran experiencia profesional, sino también personal.

Preface

This doctoral thesis was begun five years ago in the framework of a joint research unit between the Institute for Bioengineering of Catalonia (IBEC, Barcelona, Spain) and the Health Sciences Research Institute of the Germans Trias i Pujol Foundation (IGTP, Badalona, Spain), and with the collaboration of the Pulmonology service of the Germans Trias i Pujol University Hospital (HUGTiP, Badalona, Spain). We aimed to develop an alternative and more sensitive tool for assessing pulmonary function in respiratory diseases based on the analysis of respiratory sounds (RS). To do this, a number of advanced signal processing techniques have been developed in this thesis, allowing us to propose a new approach to RS analysis that combines the analysis of both normal RS and continuous adventitious sounds. However, the applicability of the proposed approach in clinical practice has been a major objective from the beginning of this thesis. For that reason, many practical aspects have also been taken into consideration—such as the number of sensors and the respiratory manoeuvres—to define an RS recording protocol that can be launched in pulmonary function testing laboratories.

Two different protocols were defined and launched in the Pulmonary Function Testing Laboratory of the HUGTiP, at the beginning of this thesis, for assessing unilateral phrenic paralysis and asthma. During the past five years, the RS recording sessions with patients at the HUGTiP have run in parallel with the development of new techniques for RS signal processing and analysis at the IBEC. All the work done during these years has resulted in the following published articles, which have been collected for the submission of this thesis:

Article 1 (Chapter 2)

Title: Detecting unilateral phrenic paralysis by acoustic respiratory analysis
Authors: J. A. Fiz, R. Jané, M. Lozano, R. Gómez, J. Ruiz
Journal: *PLOS ONE*, Apr. 2014.

Article 2 (Chapter 3)

Title: Automatic differentiation of normal and continuous adventitious respiratory sounds using ensemble empirical mode decomposition and instantaneous frequency
Authors: M. Lozano, J. A. Fiz, R. Jané
Journal: *IEEE J. Biomed. Health Inform.*, vol. 20, no. 2, pp. 486–497, Mar. 2016.

Article 3 (Chapter 4)

Title: Performance evaluation of the Hilbert–Huang transform for respiratory sound analysis and its application to continuous adventitious sound characterization
Authors: M. Lozano, J. A. Fiz, R. Jané
Journal: *Signal Process.*, vol. 120, pp. 99–116, Mar. 2016.

Book chapter (Chapter 5)

Title: Analysis of normal and continuous adventitious sounds for the assessment of asthma
Authors: M. Lozano, J. A. Fiz, R. Jané
Book: *XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013*, vol. 41. Springer International Publishing, 2014, pp. 981–984.

This work was supported in part by:

- The Secretary of Universities and Research of the Department of Economy and Knowledge of the Government of Catalonia (Consolidated Research Group GRC 2014 SGR 1569)
- The Spanish Ministry of Economy and Competitiveness under grant TEC2010-21703-C03-01
- The Spanish Ministry of Economy and Competitiveness and the European Regional Development Fund under grant DPI2015-68820-R
- The Biomedical Signal Processing and Interpretation (BIOSPIN) group of the IBEC
- The Biomedical Systems and Signals group of the Biomedical Research Networking Centre in Bioengineering, Biomaterials, and Nanomedicine (CIBER-BBN)
- The Medical Technology Innovation Platform of the Institute of Health Carlos III

I am currently using the proposed approach to RS analysis as part of the project of my Long-Term Research Fellowship (LTRF 2015-5185) of the European Respiratory Society. I am performing this joint project in the Respiratory Muscle Physiology group at King's College London (KCL), under the supervision of Prof. John Moxham (KCL, Host institution) and Prof. Raimon Jané (Biomedical Signal Processing and Interpretation group, IBEC, Home institution). Together with RS, non-invasive measurements of respiratory muscle function are being used in this project to assess neural respiratory drive and breathlessness in patients with chronic obstructive pulmonary disease.

Table of contents

Abstract V
Acknowledgments IX
Preface XI
Table of contents XIII
List of figures and tables XV
Abbreviations XVII
Chapter 1: Introduction 1
1.1. Ventilation 1
1.1.1. Airways and airflow 1
1.1.2. Regulation of ventilation 2
1.1.2.1. Spontaneous periodic ventilation 2
1.1.2.2. Chemical regulation of ventilation 3
1.1.2.3. Other regulations of ventilation 3
1.1.3. Measurement of ventilation 4
1.2. Respiratory sounds 7
1.2.1. RS classification 7
1.2.1.1. Normal RS 8
1.2.1.2. Adventitious RS 9
1.2.1.3. Other abnormal RS 9
1.2.2. Production of normal RS 10
1.2.3. Transmission and distribution of RS 11
1.3. Relationship between RS and respiratory diseases 13
1.3.1. Decrease in RS intensity 13
1.3.2. Production of adventitious RS 14
1.4. RS analysis 17
1.4.1. Relationship between normal RS intensity and airflow 17
1.4.2. Detection and characterisation of CAS 18
1.5. Proposed approach to RS analysis 21
1.5.1. Multichannel recording of RS 21
1.5.2. Airflow-dependent analysis of RS 21
1.5.3. Analysis of normal RS based on the RS intensity–airflow curves 22
1.5.4. Analysis of CAS based on HHT 23
1.5.4.1. The need for time-frequency analysis 23
1.5.4.2. The concept of instantaneous frequency and HHT 24
1.6. Applications of the proposed approach to RS analysis 31
1.6.1. Assessment of UPP by RS analysis 31
1.6.1.1. Previous work on the assessment of diaphragmatic paralysis 31
1.6.1.2. Proposed approach to assessing UPP by RS analysis 33
1.6.2. Assessment of OPDs by RS analysis 33
1.6.2.1. Previous work on normal RS and OPDs 33
1.6.2.2. Previous work on CAS and OPDs 35
1.6.2.3. Proposed approach to assessing OPDs by RS analysis: application to BDR assessment in asthma 38
1.7. Hypotheses 41
1.8. Objectives 43
1.8.1. Main objective 43
1.8.2. Specific objectives 43
Chapter 2: Detecting unilateral phrenic paralysis by acoustic respiratory analysis 45

Chapter 3: Automatic differentiation of normal and continuous adventitious respiratory sounds using ensemble empirical mode decomposition and instantaneous frequency 57

Chapter 4: Performance evaluation of the Hilbert–Huang transform for respiratory sound analysis and its application to continuous adventitious sound characterization 71

Chapter 5: Analysis of normal and continuous adventitious sounds for the assessment of asthma 91

Chapter 6: Discussion 97

Chapter 7: Conclusions 103

7.1. Future work 105

Bibliography 107

Publications derived from this thesis 115

List of figures and tables

Figure 1. Conducting airways and respiratory zone **1**
Figure 2. Elements of ventilation regulation **2**
Figure 3. Typical volume-time and flow-volume curves obtained during forced spirometry **4**
Figure 4. Classification of RS **7**
Figure 5. Frequency band of tracheal and lung sounds **8**
Figure 6. Production site of tracheal and lung sounds **10**
Figure 7. Relationship between RS and respiratory diseases **13**
Figure 8. Power spectrum of an RS signal containing CAS **19**
Figure 9. Spectrogram of an RS signal containing CAS **19**
Figure 10. Recording points for lung and tracheal sounds **22**
Figure 11. Progressive respiratory manoeuvre **22**
Figure 12. Proposed approach to CAS analysis based on HHT as compared to previous approaches based on TFDs **23**
Figure 13. RS decomposition by ensemble empirical mode decomposition (EEMD) **27**
Figure 14. IF and IE sequences of an RS signal containing CAS **27**
Figure 15. TFDs for RS analysis **28**
Figure 16. Motor innervation of the diaphragm by the phrenic nerves **31**
Figure 17. RS recording points to assess UPP **33**
Figure 18. Proposed approach to assessing BDR in OPDs by RS analysis **38**

Table 1. Analysis of CAS in OPDs **36**

Abbreviations

AO	Airway obstruction
BDR	Bronchodilator response
CAS	Continuous adventitious sounds
CO_2	Carbon dioxide
COPD	Chronic obstructive pulmonary disease
CORSA	Computerised respiratory sound analysis
EMD	Empirical mode decomposition
EEMD	Ensemble empirical mode decomposition
FEV_1	Forced expiratory volume in 1 second
FFT	Fast Fourier transform
FVC	Forced vital capacity
HHT	Hilbert-Huang transform
HS	Hilbert spectrum
HUGTiP	Germans Trias i Pujol University Hospital
IE	Instantaneous envelope
IF	Instantaneous frequency
OPD	Obstructive pulmonary disease
PSD	Power spectral density
RID	Reduced interference distribution
RMS	Root mean square
RS	Respiratory sounds
TFD	Time-frequency distribution
TLC	Total lung capacity
UPP	Unilateral phrenic paralysis
VC	Vital capacity
VRI	Vibration response imaging
WVD	Wigner-Ville distribution

Chapter 1: Introduction

1.1. Ventilation

Ventilation, one of the functional processes of the respiratory system, consists in moving air from outside the body into the lungs, and vice versa. The aims of ventilation are to transport oxygen to the alveoli within the lungs and to eliminate the carbon dioxide (CO_2) produced by cell respiration [1].

1.1.1. Airways and airflow

Ventilation comprises two phases:

- *Inspiration*: air flows from the atmosphere into the lungs
- *Expiration*: air flows from the lungs to the atmosphere

Inspiration is an active process, even during quiet breathing, which requires the activation of the inspiratory muscles, mainly of the diaphragm and, to a lesser extent, the external intercostal muscles. Contraction of these muscles expands the rib cage, and this expansion is transmitted to the lungs through the pleural cavity. Lung expansion causes the intrapulmonary pressure to fall below the atmospheric pressure, thus making air enter the lungs through a series of tubular structures, called conducting airways (Figure 1) [2], [3].

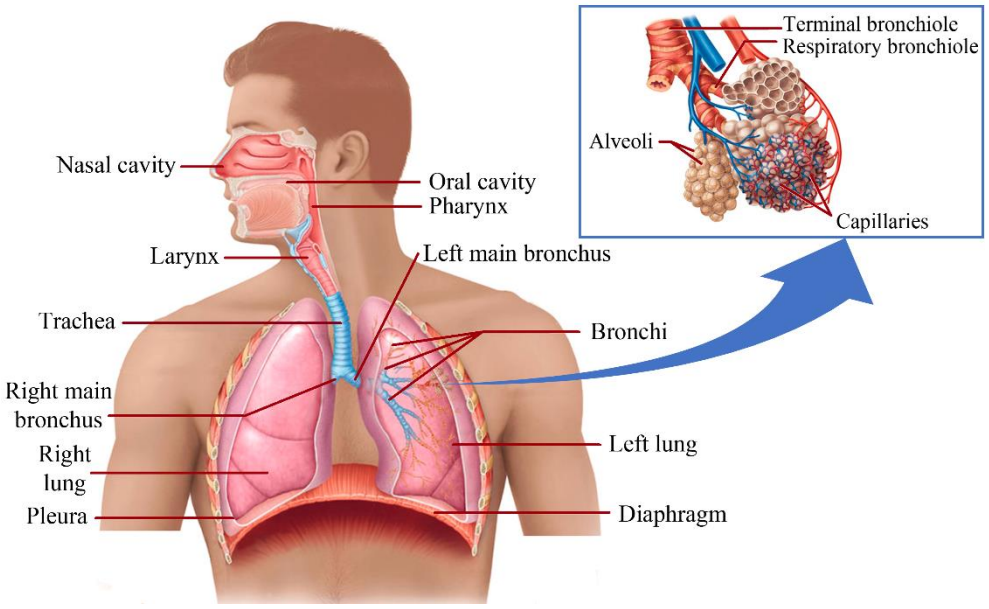


Figure 1. Conducting airways and respiratory zone (see box)

Inspired air enters the nasal cavities and the pharynx and moves downwards through the tracheobronchial tree, which progressively branch out, starting from the trachea down to the terminal bronchioles. Conducting airways do not participate in gas exchange but rather clear and improve the quality of inspired air as it is carried to the alveoli. The respiratory zone, where gas exchange occurs,

starts with the division of terminal bronchioles into respiratory bronchioles and includes alveolar ducts and alveoli as well.

After each inspiration, a fraction of the inhaled air (about 150 mL) remains in the conducting airways, constituting the anatomical dead space. However, most of the inhaled air (about 2.5 to 3 L) goes into the respiratory zone and takes part in gas exchange [2].

Different from inspiration, normal expiration is a passive process. Relaxation of the diaphragm and the external intercostal muscles causes the lungs to compress due to the elasticity of their tissue. Lung compression causes the intrapulmonary pressure to rise above the atmospheric pressure, causing air to flow from the lungs to the atmosphere.

1.1.2. Regulation of ventilation

The three basic elements of the regulation of ventilation are the central controller (nervous system), effectors (respiratory muscles), and sensors (chemical and mechanical receptors) (Figure 2).

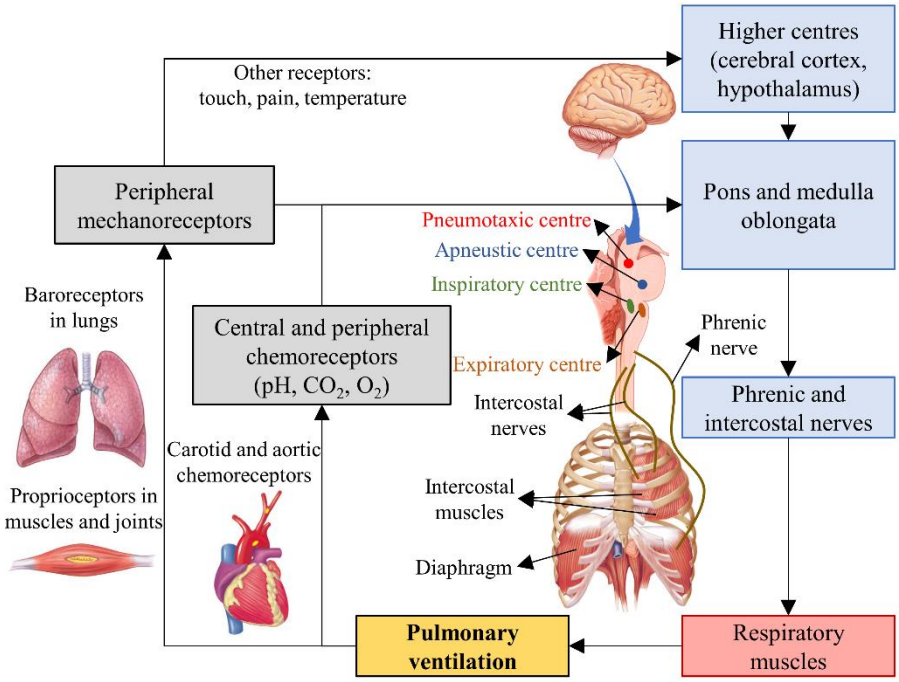


Figure 2. Elements of ventilation regulation (adapted from [3])

1.1.2.1. Spontaneous periodic ventilation

Central controller automatically regulates both the depth and rate of ventilation through the respiratory centres located in the medulla oblongata and pons. Medullary respiratory centres include the dorsal and ventral respiratory neurons, which are the inspiratory and expiratory centres, respectively. The rhythm of ventilation is mainly determined by spontaneous periodic impulses generated by the inspiratory centre. These impulses travel through the phrenic and intercostal nerves and cause the inspiratory muscles to contract. Besides its spontaneous activation, the inspiratory centre is stimulated from the apneustic centre, which is located in the lower pons and contributes to produce regular ventilation [3].

Impulses from the inspiratory centre also stimulate the pneumotaxic centre within the upper pons, which in turn inhibits both apneustic and inspiratory centres, resulting in termination of inspiration and initiation of passive expiration. A new cycle of ventilation is then started by spontaneous activation of the inspiratory centre. Normal ventilation cycles last around 3 to 5 seconds, producing a normal ventilation rate of 12 to 20 cycles per minute [2].

The medullary expiratory centre is only activated during deep ventilation. In this case, the inspiratory centre activates the expiratory centre, thereby stimulating expiratory muscles, including the internal intercostal and abdominal muscles [2].

1.1.2.2. Chemical regulation of ventilation

The main objective of ventilation is to maintain blood oxygen and CO_2 and blood pH within their normal range of values. Specific groups of neurons, called chemoreceptors, are located in the medulla (central chemoreceptors) or the carotid and aortic bodies (peripheral chemoreceptors), where they detect changes in gas content and pH of blood and regulate ventilation according to these changes.

Chemoreceptors are more sensitive to changes in blood CO_2 than to changes in blood oxygen because CO_2 directly affects blood pH, the values of which must be maintained within normal parameters for the correct functioning of cells. Therefore, changes in blood CO_2 are responsible for most of the changes in ventilation. Blood CO_2 is mostly regulated by central chemoreceptors. When blood CO_2 levels increase, the CO_2 crosses the blood-brain barrier and lowers the pH of the cerebrospinal fluid. The decreased pH activates the central chemoreceptors, which stimulate the respiratory centre. As ventilation increases, CO_2 is eliminated, and blood pH reaches the normal levels.

Blood oxygen usually does not affect ventilation regulation. Oxygen begins to affect ventilation when its blood concentration falls below half its normal value [1]. In this case, peripheral chemoreceptors are responsible to stimulate the respiratory centre in order to increase ventilation, and thus blood oxygen.

1.1.2.3. Other regulations of ventilation

Ventilation can also be affected by responses from higher brain centres and peripheral mechanical receptors [3]:

- The *limbic system* of the brain can accelerate ventilation in response to emotions
- The *cerebral cortex* allows the depth and rate of ventilation to be voluntarily controlled
- *Baroreceptors* in lung tissue detect stretching during inflation of the lungs. The activation of these receptors inhibits the inspiratory centre, thus preventing overinflation of the lungs. During expiration, these receptors are no longer activated, allowing the inspiratory centre to activate again
- *Proprioceptors* in muscles and joints detect muscle contractions and body movements, and their action potentials stimulate the respiratory centre
- Other receptors of *touch*, *pain*, and *temperature* can also stimulate the respiratory centre through higher brain centres

1.1.3. Measurement of ventilation

Lung function tests are a set of respiratory manoeuvres that are performed using standardised equipment to determine how well the lungs work. In particular, spirometry and body plethysmography are those most often used to assess ventilation [4].

Spirometry measures how much air can be inhaled and exhaled (lung volumes) and how quickly this can happen (airflow) (Figure 3). The spirometry manoeuvre consists of performing a maximal inspiration from tidal breathing followed by a maximal expiration. It can be performed in a relaxed or a forced manner.

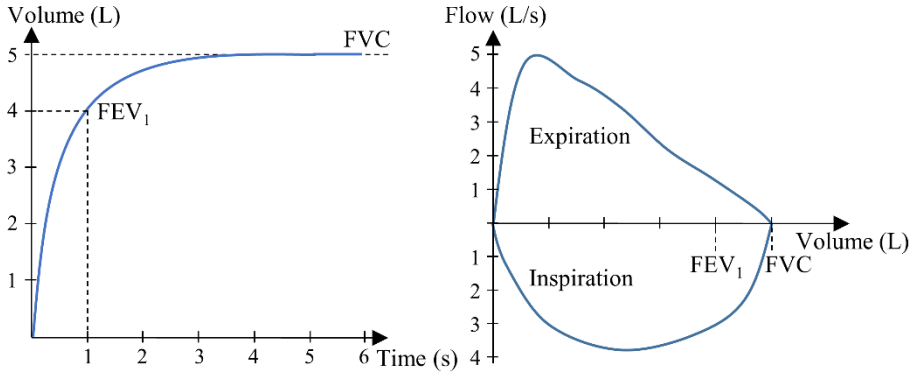


Figure 3. Typical volume-time and flow-volume curves obtained during forced spirometry. FEV_1, forced expiratory volume in 1 second; FVC, forced vital capacity.

The most important spirometric parameters from a clinical point of view are:

- *Forced vital capacity (FVC)*, the volume of air that is exhaled from a maximal inspiration to a maximal expiration during a forced spirometry
- *Forced expiratory volume in 1 second (FEV_1)*, the volume of air that is exhaled during the first second of a forced spirometry
- *FEV_1/FVC ratio*
- *Total lung capacity (TLC)*, the total volume of air that can be contained in the lungs

These parameters, which are usually expressed as a percentage of the reference values according to sex, age, and height [5], provide relevant clinical information for distinguishing different types of abnormal ventilatory patterns, including obstructive and restrictive patterns [4]. An obstructive pattern is characterised by reduced expiratory airflows, due to airway obstruction (AO). Despite airflow limitation, FVC is usually normal or reduced to a lesser extent than FEV_1, which leads to a reduction in the FEV_1/FVC ratio. A decrease in the FEV_1/FVC ratio below the normal range of values (0.7–0.8) is a clear indicator of AO [6].

On the other hand, a restrictive pattern is characterised by reduced lung volumes (<80% of predicted), including FEV_1, FVC, and TLC, whereas expiratory airflows may be normal. In this case, the FEV_1/FVC ratio is normal or even elevated, since FVC may be reduced to a larger extent than FEV_1 [7].

However, a reduced FVC is not always due to a restrictive pattern. In severe AO, as in emphysema, lung hyperinflation leads to an increased residual lung volume and a reduced FVC, which may be erroneously interpreted as a restrictive pattern. For this reason, TLC—the summation of FVC and the residual lung volume—should be measured to confirm or exclude the restrictive pattern. Since spirometry does not measure the residual volume, additional tests, such as gas dilution technique or body plethysmography, are required for this purpose [4].

Major spirometric parameters are used as a first simple tool for detecting an abnormal ventilatory pattern. Spirometry can help to detect obstructive lung diseases, such as asthma and chronic obstructive pulmonary disease (COPD), at an early stage and thus guide their management [8]. However, in order to make a more accurate diagnosis, additional studies are required. For example, a bronchodilator response (BDR) test is usually performed when airflow obstruction is suspected from spirometric parameters. This test consists in inhaling around four puffs of a bronchodilator medication, usually a short-acting β2-agonist such as salbutamol, and repeating the spirometry after 15 minutes [7]. An increase in FEV_1 of >12% and >200 mL from the baseline value is considered as a positive BDR [6], [7]. Bronchodilator therapy can often reverse airflow limitation in patients with asthma but usually is not helpful for patients with COPD.

Even though spirometry is the most widely used test to assess ventilation, and especially for AO assessment, it has also been a subject of controversy. The standard BDR criterion, based on the percentage increase in FEV_1, is highly influenced by the baseline FEV_1 [6], so that subjects with a low baseline FEV_1 are more likely to have a greater BDR than subjects with a high baseline FEV_1. In this regard, a study published in 2012 [9] concluded that BDR could be omitted in patients with normal baseline spirometry and an FEV_1 above 90% of that predicted, since only 1% of these patients had a positive BDR. Subsequently, another study performed in children [10] showed that 10–12% of children with normal baseline spirometry had a positive BDR, and therefore baseline FEV_1 should not affect the decision to perform a BDR test.

The true significance of various levels of BDR and the usefulness of the standard BDR cutoff have also been a source of debate, especially for diagnosing asthma [11], [12] and differentiating between asthma and COPD [13]. In fact, using FEV_1 to assess BDR, especially in patients with COPD, is quite paradoxical, since "*It would be the equivalent of defining essential hypertension as an increase in BP that must not respond to antihypertensive therapy, and then testing antihypertensive agents using the unmodifiable BP as the only outcome*" [8].

Spirometry is a simple and useful technique for supporting the diagnosis, but a specific diagnosis should not be made based only on spirometric parameters [7].

1.2. Respiratory sounds

Respiratory sounds (RS) are mechanical waveforms—that is, vibrations—that are produced during breathing, mainly by airflow patterns and pressure changes within the airways. Lung auscultation is a key step in physical examination, since RS provide useful clinical information about the respiratory system structure and functioning. Respiratory diseases may alter airflow patterns, airway dynamics, and regional lung volume distribution, which in turn affect the characteristics of RS [14], [15]. Over the past few years, advances in sensors and signal processing techniques have made RS recording and analysis an objective, sensitive, and powerful tool for assessing pulmonary function.

1.2.1. RS classification

Much effort has been made, both in the past and recently, to unify and standardise nomenclature in lung auscultation [15]–[17], especially by the American Thoracic Society and the European Respiratory Society [18], [19]. Nonetheless, no standard terminology exists at present to describe RS, and there is still some discrepancy in the literature [17].

According to two recent and relevant publications about RS [15], [17], the most widely accepted categorisation of RS is based on their characteristics and recording site (Figure 4).

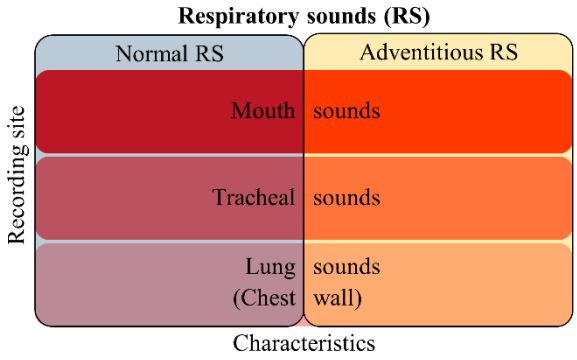


Figure 4. Classification of RS.

Besides location, the following characteristics are mainly used to distinguish between different types of RS:

- *Frequency*, measured in hertz (Hz), refers to the speed of vibration of RS. RS are formed by a range of frequencies that determine how RS are heard. The major or fundamental frequency of an RS signal is usually called pitch, with high-pitched RS mainly containing high frequencies and low-pitched RS containing low frequencies
- *Intensity*, usually measured in decibels (dB), refers to the loudness of RS, which is directly related to the amplitude of RS signals
- *Duration*, usually measured in milliseconds (ms), is the time period during which RS are heard

1.2.1.1. Normal RS

Normal RS are those heard at the chest, neck, or mouth of healthy people during breathing. Although normal RS signals are random in nature, their characteristics highly depend on the recording site.

When recorded at large airways, such as the trachea (tracheal sounds) (Figure 5A), normal RS can be heard during inspiration and expiration. These RS cover a wide frequency range, from around 70 Hz to around 1500 Hz. However, most of their power is concentrated in the frequency range up to 850–1000 Hz, with a sharp drop in power above these frequencies [15], [20] (Figure 5B).

Normal RS recorded at the chest wall (lung sounds) (Figure 5C) are much louder during inspiration than during expiration. In contrast to mouth and tracheal sounds, lung sounds are affected by the low-pass filtering effects of the chest. Accordingly, although lung sounds have a frequency band from around 70 Hz to around 1000 Hz, most of their energy extends up to 200–250 Hz, with a sharp decrease in power above 250 Hz [15], [20], [21] (Figure 5D).

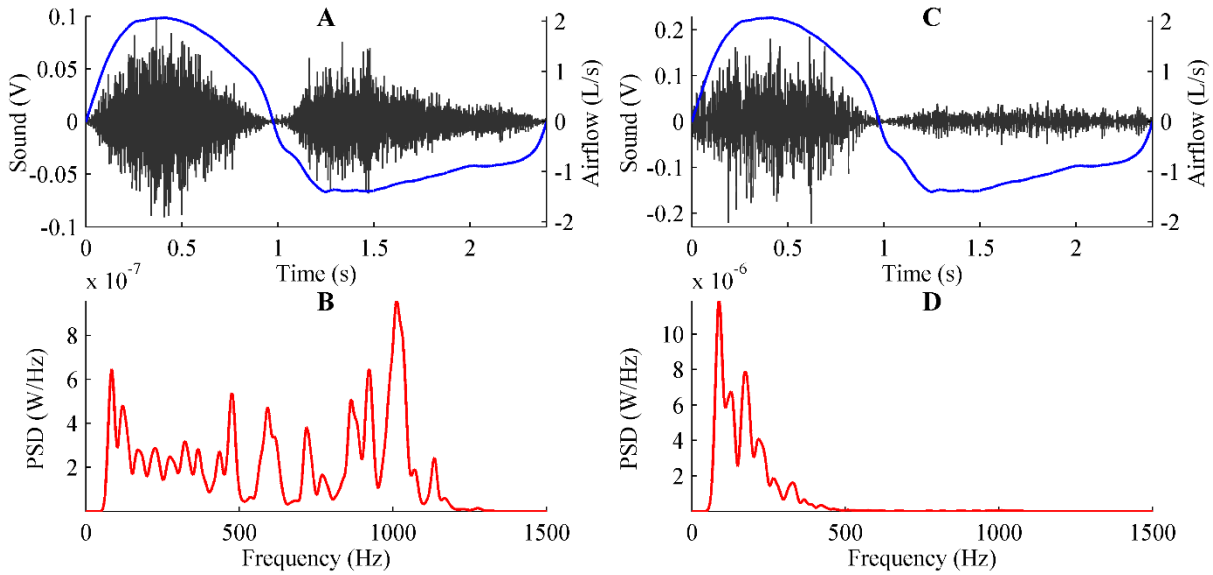


Figure 5. Frequency band of tracheal and lung sounds. (A) Tracheal sounds and airflow. (B) Power spectral density (PSD) of tracheal sounds. (C) Lung sounds and airflow. (D) PSD of lung sounds. Positive airflow values correspond to inspiration and negative airflow values, to expiration. PSDs were calculated using Welch's periodogram.

A few studies have investigated the effects of age and gender on normal RS characteristics but have not revealed major differences. Specifically, only minor changes in the spectra of RS that correlate to age have been reported [22]. However, those changes had no clinical significance, so that there is no practical need to consider the patient's age during RS analysis. Moreover, although women have been reported to have a higher power at high frequencies than men [21], [22], these differences at high frequencies are not relevant, since most of the energy of RS is concentrated at low frequencies, at which range no significant differences were found between men and women.

1.2.1.2. Adventitious RS

Adventitious RS are additional abnormal sounds that appear superimposed on normal RS and usually indicate the presence of a respiratory disorder [20]. These sounds are subdivided as follows:

- *Continuous adventitious sounds* (CAS) typically last more than 100 ms and have a sinusoidal-like or quasi-periodic waveform [15], [20]. They include:
  - *Wheezes*, musical sounds with a pitch between 100 Hz and 1000 Hz that can be inspiratory, expiratory, or biphasic. Wheezes can be monophonic or polyphonic, depending on whether they have only one frequency component or several simultaneous frequency components, respectively
  - *Stridors*, musical sounds with a pitch around 500 Hz, which are mainly inspiratory but can also be expiratory or both. Different from wheezes, stridors are better heard over the large airways, especially over the neck, and usually have shorter duration
  - *Rhonchus*, musical sounds with a pitch around 150 Hz. Their waveform can be either like pure low-pitched wheezes or similar to snores. Rhonchus are often reported as low-pitched wheezes
- *Discontinuous adventitious sounds* mainly include crackles, which are short, explosive, non-musical sounds that can be heard during both respiratory phases. Crackles are further divided into fine (around 5 ms and 650 Hz) and coarse (around 15 ms and 350 Hz). Fine crackles are mainly generated during mid-to-late inspiration, whereas coarse crackles are heard during early inspiration and throughout expiration [15]
- *Squawks* are mixed sounds that contain a short wheeze at around 200–300 Hz accompanied or preceded by crackles, mainly heard during inspiration [15]

Even though this is the most widely-accepted classification for adventitious RS, it is still controversial to date [17]. The lack of studies in large samples and of a standardised methodological approach to RS recording (e.g., number of sensors and respiratory manoeuvres) and analysis (e.g., algorithms and variables) have made it difficult to reach strong conclusions about the characteristics of different types of adventitious RS, especially with respect to CAS [23].

1.2.1.3. Other abnormal RS

In certain pathological conditions, RS characteristics may be altered, independent of whether adventitious RS are present or not. The most common abnormalities in RS are related to changes in intensity. A decrease in RS intensity can be associated to several disorders involving sound generation or transmission, as described in Section 1.3.1. Abnormal lung sounds with increased frequency and intensity and an unusual louder and longer expiratory phase are called bronchial sounds [14], [15], [24]. Although recorded at the chest, bronchial sounds are more similar to tracheal sounds than to normal lung sounds.

1.2.2. Production of normal RS

Normal RS are a consequence of turbulent airflow through the airways and changes in airflow patterns [24], [25]. Turbulent airflow occurs in large airways, in which air flows quickly and air molecules move randomly. Flow turbulence produces pressure fluctuations within the air and airway walls, which in turn produce RS [14], [26]. The intensity of both turbulence and RS increases with flow velocity. However, turbulent flow only appears above a critical flow velocity. As air flows from large airways to smaller airways, its velocity progressively decreases until it is insufficient to produce turbulence, which occurs somewhere between the lobar bronchi and the terminal bronchioles [24]. Therefore, airflow in small airways is laminar and does not produce RS.

Normal tracheal sounds are produced by turbulent inspiratory and expiratory airflows in upper airways. However, normal lung sounds are related to turbulent airflows and changes in airflow patterns in the central airways [14], [27]. The branching structure of the central airways and the different airway diameters cause abrupt changes in the direction of airflow and force the airstream to be separated into layers with different velocities. Therefore, as air flows towards peripheral airways, the airflow pattern becomes irregular, and circular airflows or vortices appear. This airflow pattern also contributes to produce normal lung sounds.

In contrast to tracheal sounds, the inspiratory and expiratory components of normal lung sounds have different origins [15], [20], [26]–[31]. Inspiratory lung sounds are produced within the lobar and segmental airways, with the contribution of even more peripheral airways [24]. However, expiratory lung sounds are produced within more proximal airways, as airflows from small airways progressively converge within larger airways (Figure 6). Since inspiratory lung sounds are produced within more peripheral airways—that is, more locally and deep underneath the chest wall—they can be clearly heard over the entire chest wall surface. Nevertheless, the intensity of expiratory lung sounds gradually decreases from the upper airways to the base of the lungs.

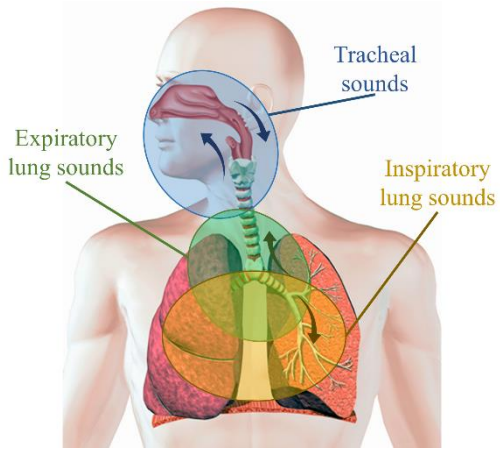


Figure 6. Production site of tracheal and lung sounds.

1.2.3. Transmission and distribution of RS

RS transmission depends on the mechanical properties and geometry of the solid structures within the thorax, including airways, lung parenchyma, and chest wall [32]. As RS are transmitted through different structures, they are attenuated and filtered. The amount of RS energy that reaches the surface of the chest wall mainly depends on the acoustic impedance of the different media, that is, their opposition to the flow of energy. If the acoustic impedance of two different media are very different, most of the RS energy is absorbed rather than transferred. In contrast, the more similar the impedances, the higher the amount of transferred energy [14].

The asymmetric transmission and distribution of RS was analysed in several early studies in the ‘80s and ‘90s [32]–[38]. RS cause large airway walls to oscillate at low frequencies, allowing most of the energy to be transferred to the lung parenchyma and towards the chest surface. However, airway walls become more rigid at high frequencies, allowing RS to remain within the airways and follow an airway pathway [33]. Accordingly, the high-frequency components of normal RS and wheezes are better transmitted through the airways than through the lung parenchyma and are more clearly heard over the trachea than at the chest wall surface. Thus, tracheal sounds could contain information relevant for assessing several respiratory diseases, such as asthma [20], [39].

The lung parenchyma, which mainly includes alveoli, small airways, and capillaries, was modelled as an uniform mixture of gas (alveoli) and water (lung tissue) [26], [33]–[35]. This model was used to analyse how RS are transmitted through the lung parenchyma, showing that the absorption of RS energy highly depends on frequency. The lung parenchyma acts as a low-pass filter, thus attenuating high-frequency RS and allowing low-frequency RS to propagate towards the chest surface. Further, transmission of low-frequency RS through the lung parenchyma is minimally affected by changes in lung volume or gas density [35].

The chest wall, which contains bones, muscles, fat, and skin, is thinner but also denser than the lung parenchyma. Consequently, most of the RS energy is lost at the interface of these two media. In addition, the heterogeneity of the chest wall and the differences in its composition between people are responsible for the high variability of RS amplitude at the chest wall surface. For example, RS are poorly transmitted to chest surface areas overlying bones, and therefore these areas should be avoided when recording RS [26].

Despite the high variability of the spatial distribution of RS amplitude over the chest surface, some particular patterns have been reported in the literature. On average, inspiratory sounds are around 10 dB louder than expiratory sounds at the chest surface [36]. Accordingly, most studies have focused only on inspiratory sounds to analyse the distribution of RS amplitude.

Significant differences in RS distribution have been reported both vertically and horizontally. Specifically, RS amplitude is greater in the upper lobe than over the base of the lungs anteriorly, but it

increases towards the base posteriorly [26], [37]. In the horizontal plane, RS recorded over the upper right hemithorax have been found to be significantly louder than those measured over the upper left hemithorax, both anteriorly [38] and posteriorly [35]. However, at the posterior lower chest, a lateralisation of RS to the left hemithorax has been reported [35], [36], [38], [40]. In healthy people, horizontal differences in RS distribution have been found to be slight (of a few dB) and less relevant than vertical differences [36], [40]–[42]. These differences have been mainly attributed to the asymmetric mediastinal anatomy, and especially to the asymmetric geometry of airways because of the position of the heart. Further, relevant postural effects on the distribution of RS amplitude have also been reported [40]. In particular, lung sound intensity was found to be greater over the dependent lung in lateral decubitus positions.

During the past decade, advances in sensor and computer technologies have made the multichannel analysis of RS more feasible and practical [43]. RS multichannel recording consists on placing multiple sensors on the chest surface so that RS can be simultaneously recorded at several points and large amounts of data can be recorded and processed in a straightforward and efficient way. Several multichannel approaches to RS recording and analysis have been proposed to provide more accurate data about RS localization and distribution [44], [45].

A major advantage of RS multichannel recording is the possibility of mapping RS to create respiratory acoustic thoracic images [46]. RS imaging provides detailed information about the distribution of the RS intensity, which has allowed further analysis of the asymmetries and patterns in the distribution of normal RS [41], [47]. Furthermore, this new technique for RS analysis has led to the development of a vibration response imaging (VRI) device [42], [48], which uses 40 sensors assembled into two arrays to record RS over the posterior chest surface of a person. The system maps the RS intensity from each sensor to create static and dynamic grey-scale images. However, these images have been interpreted qualitatively, thus depending on the subjectivity of the physicians. The VRI system only provides quantitative data for two or three regions of each lung (i.e., upper, middle, lower), in which the relative RS intensities are aggregated and expressed as a percentage of the total for both lungs.

Although some particular patterns in the spatial RS distribution have been proposed, the heterogeneous distribution of RS intensity remains still unclear, so that more studies still need to address this issue.

1.3. Relationship between RS and respiratory diseases

The fact that RS characteristics vary over the chest surface, even in healthy people, indicates that RS contain regional information about their site of generation and transmission path [26], either one or both of which may be affected by respiratory diseases, thus altering the RS characteristics (Figure 7).

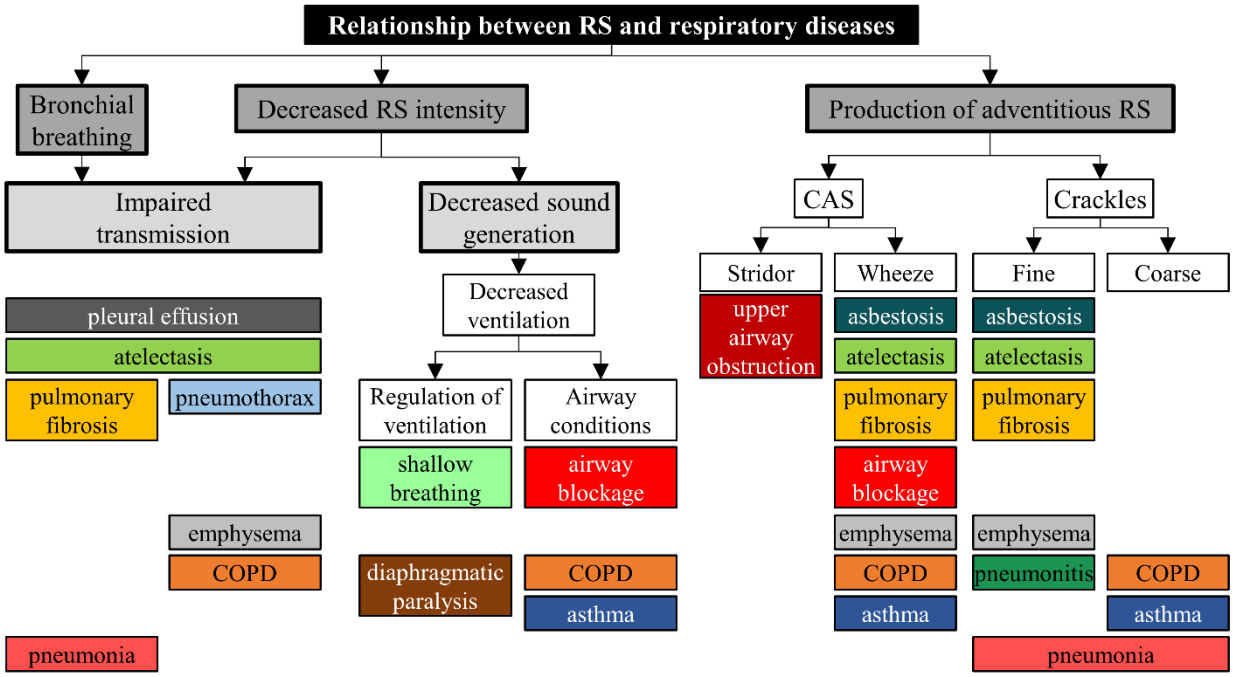


Figure 7. Relationship between RS and respiratory diseases. CAS, continuous adventitious sounds; COPD, chronic obstructive pulmonary disease; RS, respiratory sounds.

1.3.1. Decrease in RS intensity

RS intensity can be decreased due to a reduction in sound generation or impaired transmission, or both. A reduced sound generation implies a decreased ventilation and airflow limitation, since RS production depends on turbulent airflows through the airways. Further, decreased ventilation can be due either to disruption of ventilation control or to airway conditions [14], [15].

Shallow breathing, and consequent decreased ventilation, often happens in persons with decreased levels of consciousness (e.g. due to injury to the central nervous system or drug overdose). Chest pain can also be a cause of decreased ventilation, since pain can limit the breathing depth. Additionally, damage to the peripheral nerves that activate respiratory muscles can also cause decreased ventilation. For example, injury to the phrenic nerves results in diaphragmatic paralysis, which in turn results in diminished air entering the lungs and thus decreased RS.

Airway conditions include AO by a foreign body or tumour and obstructive pulmonary diseases (OPDs), such as asthma or COPD [27], [49]. AO limits airflow, thereby directly effecting generation of RS, which are reduced over the obstructed area and more distal areas. The reduction in RS intensity can be reversible or permanent, depending on whether the obstruction is reversible or not.

RS transmission can be impaired anywhere between the site of RS production and the chest surface. An impaired transmission implies either a change in the mechanical properties of a medium or an impedance mismatch between two media, which occurs when RS are transmitted through media with very different acoustic properties, such as density [14], [15]. Damaged lung tissue in emphysema or increased air trapped in hyperinflated lungs (e.g. in COPD) change the acoustical properties of the lung parenchyma and lead to a generalised reduction of lung sound intensity. Obesity, with an excess of adipose tissue and a resulting increase in chest wall thickness, also decreases lung sound intensity at the chest wall surface. Further, in pneumothorax or pleural effusion, the presence of air or fluid accumulation within the pleural space causes an impedance mismatch between the normal lung parenchyma and the pleural cavity, resulting in diminished lung sound intensity over the affected area [50].

In some cases, lung sound intensity can be abnormally increased over the chest wall surface. Patients with pneumonia and lung consolidation, atelectasis, or fibrosis may have abnormally loud RS (bronchial breathing) over the affected area. These conditions increase the lung tissue density, improving the impedance match between the lung tissue and the chest wall and enhancing the transmission of RS to the chest wall surface [15], [24], [36].

1.3.2. Production of adventitious RS

Several airway and lung parenchyma disorders are involved in the production of adventitious RS. In general, CAS are more related to obstructive airway diseases than to lung parenchyma disorders, although the latter can also produce CAS. All types of CAS share the same mechanism of generation, which revolves around airway narrowing and airflow limitation. Forgacs [24] reported that the correct model of the production of wheezes "*is a simple uncoupled reed, represented by a bronchus narrowed to the point of closure, whose opposite walls oscillate between the closed and barely open positions*". Subsequently, this theory was corroborated by Grotberg and Gavriely [51], who developed a theoretical model of wheeze generation using collapsible tubes, showing that wheezes are generated by oscillation of narrowed airways whose walls are in apposition.

Wheezes only appear above a critical flow, which depends on the mechanical properties of airways, including the airway diameter and the airway wall thickness, stiffness, elastance, and longitudinal tension [27], [51]. These mechanical properties also affect the pitch of wheezes, which increases with narrower airways, thinner airway walls, or stiffer airway walls. However, wheeze pitch is not affected by the length or diameter of airways [15], [24], [27]. Further, although wheezes are always associated with airflow limitation, the opposite is not true—airflow limitation can occur without wheezes [15], [51]. For instance, patients can have severe AO without wheezes if airflow is below the critical flow to produce wheezes. In these cases, normal RS intensity is highly reduced or even absent, which is known as silent lung [15].

Distinguishing between wheeze types can be important for diagnosis. Localized monophonic wheezes are usually related to a local obstruction (for instance, by a foreign body or tumour), although single or

multiple monophonic wheezes are also heard in patients with asthma, which can lead to a misdiagnosis [15]. Moreover, squawks (a particular type of wheeze) can be heard in lung parenchyma disorders (e.g. asbestosis, pneumonia, or pulmonary fibrosis). These are either single or multiple short monophonic wheezes that appear during inspiration over the base of the lungs and are associated with late inspiratory crackles [24], [52], [53].

Polyphonic wheezes are a common symptom of COPD and usually appear during expiration [24]. In contrast to multiple monophonic wheezes, which start and end at different times, the different frequency components of a polyphonic wheeze start and end simultaneously. Polyphonic wheezes are produced by simultaneous compression of different airways, hence the presence of different frequency components. This type of wheeze can also be heard in healthy people during forced expiratory manoeuvres. However, when heard during normal breathing, polyphonic wheezes are a clear sign of AO.

Stridors are produced by turbulent airflow passing through narrowed upper airways (caused for example by vocal cord dysfunctions, a foreign body, or a tumour). They are louder than wheezes and are more clearly heard on inspiration and over the neck. However, when recorded over the lungs, these high-pitched tracheal wheezes can be confused with asthma wheezes, highlighting the importance of recording tracheal sounds [15], [54], [55].

While all types of CAS share the same mechanism of generation, fine and coarse crackles are produced by two well-differentiated mechanisms. Fine crackles are generated by the sudden reopening of peripheral airways that were abnormally closed due to deflation of dependent lung regions in several diseases affecting the lung parenchyma, such as pulmonary fibrosis, asbestosis, or pneumonitis [20], [56]. In deflated lung regions, the airways remain unusually closed until late inspiration, causing a large difference between the atmospheric pressure and the intra-alveolar pressure. When pressure overcomes the airway resistance, peripheral airways suddenly reopen and the subsequent explosive equalization of pressure produces crackles [24]. Fine crackles, heard during mid-to-late inspiration, are not transmitted to the mouth, or affected by cough but are altered by changes in body position [15]. Fine expiratory crackles can also be due to the sudden closing of peripheral airways during expiration [56].

Coarse crackles are usually heard at the beginning of inspiration and throughout expiration and are due to the airflow going through airways that open and close intermittently. They can be related to secretions, may coexist with fine crackles in pneumonia, and, together with wheezes, are a common sign of OPDs, such as COPD or asthma [20], [24]. Further, coarse crackles can be heard over any lung region and are transmitted to the mouth, affected by cough, and altered by changes in body position [15].

Crackles can also appear in healthy people during deep inspiration following a maximum expiration. However, these crackles disappear with coughing and have no clinical relevance [14].

1.4. RS analysis

Since Rene Laënnec invented the stethoscope in 1816 [57], lung auscultation by physicians has been an essential part of cursory physical examinations, providing relevant clinical information about the respiratory system structure and functioning in a quick and easy way. However, traditional manual auscultation suffers from lack of objectivity, since it highly depends on the skills and experience of the physicians involved.

During the past three decades, advances in computer and sensor technologies have made the objective analysis of RS more practical, with increasing interest in its clinical application for diagnostic purposes. In 2000, the European Respiratory Society published guidelines for Computerised Respiratory Sound Analysis (CORSA), with the purpose of standardising both the recording and signal processing of RS [58]–[60]. Although the CORSA project described the auscultation points, type of sensors, acquisition and pre-processing guidelines, and basic signal processing techniques, a variety of approaches to RS recording and analysis have been proposed since then [23].

The next subsections describe the most relevant findings reported in previous studies regarding the technical aspects of the analysis of normal RS and CAS.

1.4.1. Relationship between normal RS intensity and airflow

As described in Section 1.2.2, the intensity of airflow turbulence and normal RS increases with airflow. The relationship between normal RS intensity and airflow has been the subject of numerous studies. Early studies described a linear relationship between the amplitude of normal RS recorded at the chest surface and the airflow recorded at the mouth [61], [62]. Leblanc et al. [61] focused on inspiratory sounds, whose intensity linearly increased with airflows varying between 1 and 4 L/s. A linear relationship between the mean absolute amplitude of normal RS and airflow was reported by Kraman [62] in four healthy participants breathing at airflows between 1.4 and 4 L/s. However, he also found a non-zero slope between airflow and the amplitude of normal RS divided by airflow, which suggested a nonlinear relationship.

Subsequent studies corroborated the nonlinear relationship between normal RS intensity and airflow [63]–[65]. The best model to describe this relationship is a power law (linear relationship in a logarithmic scale):

$$I \propto F^{\alpha} \tag{1}$$

$$\log I \propto \alpha \log F \tag{2}$$

In (1) and (2), I is RS intensity, F is airflow, and α is a constant. Different α values, which determine the power relationship, have been reported. A quadratic function ($\alpha = 2$) was found to be the best relation between the root mean square (RMS) envelope of RS and airflow for both inspiration and expiration [63]. However, α values below 2 were reported in other studies using the RS spectrum in order to

calculate RS intensity [64], [65]. There were small differences between the values of α found in these studies (1.66 ± 0.35 in [64] compared to 1.89 ± 0.57 in [65]) due to the use of different frequency bands (100–1000 Hz in [64] compared to 150–450 Hz in [65]) of the RS spectrum and the small sample size in both studies.

A third-power relationship has been reported between the RS energy at peak inspiration and airflow using the VRI system [66], although this relationship did not affect the distribution of RS energy in normalized acoustic maps, especially at airflows above 1 L/s.

The close relationship between RS and airflow led other studies to focus on flow estimation from RS, especially from tracheal sounds. Que et al. [67] estimated respiratory flow from the envelope of tracheal sounds, obtained using the Hilbert transform. The estimation of ventilation parameters from the estimated airflow was called phonospirometry. Yadollahi and Moussavi [68] used a linear model to estimate flow using the entropy of tracheal sounds in the frequency band of 75–600 Hz. Other studies have been based on the RS–airflow relationship for developing a respiratory phase detection method using only tracheal sounds [69] or by analysing variations in upper airway physiology from wakefulness to sleep in persons with obstructive sleep apnoea [70].

The aforementioned studies focused on analysing the relationship between RS and airflow, thus demonstrating that RS are highly dependent on airflow. However, in most previous studies that focused on analysing RS to assess respiratory diseases, RS were recorded without controlling for airflow, with maintained airflows, or with forced expiratory manoeuvres (Sections 1.6.2.1 and 1.6.2.2). Therefore, further studies addressing airflow-dependent analyses of RS are required.

1.4.2. Detection and characterisation of CAS

As described in Section 1.3.2, CAS are a common symptom of various respiratory diseases and are a clear sign of AO and airflow limitation. In order to objectify auscultation, several previous studies have been focused on the automatic detection and analysis of wheezes, although the reported techniques can also be used to detect other types of CAS. Early studies used the power spectrum to detect and analyse wheezes [71]–[79]. Since wheezes have a sinusoidal-like waveform, they are represented in the power spectrum as peaks centred at the fundamental frequencies of CAS and clearly higher than the baseline level (Figure 8).

Besides the power spectrum, time-frequency distributions (TFDs) have been the most commonly used methods to detect and analyse wheezes, especially those based on Fourier analysis, such as spectrogram, and wavelet analysis. TFDs show that the energy of wheezes concentrates around their pitch for as long as they last (Figure 9).

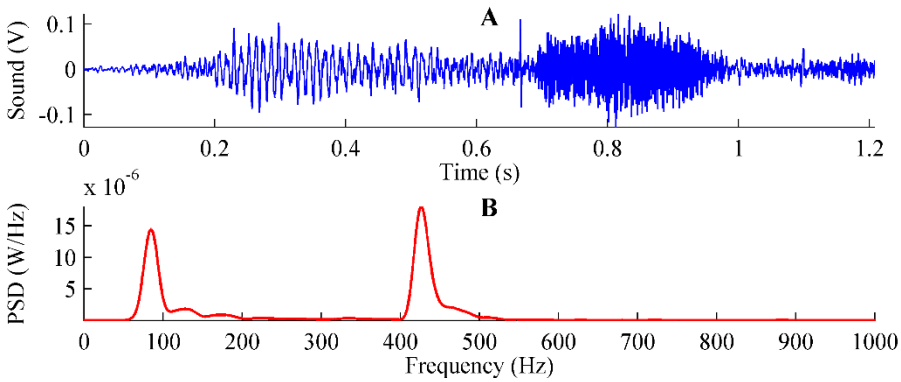


Figure 8. Power spectrum of an RS signal containing CAS. (A) RS signal containing two CAS. (B) Power spectral density (PSD). CAS appear in the PSD as peaks centred at around 85 Hz and 430 Hz.

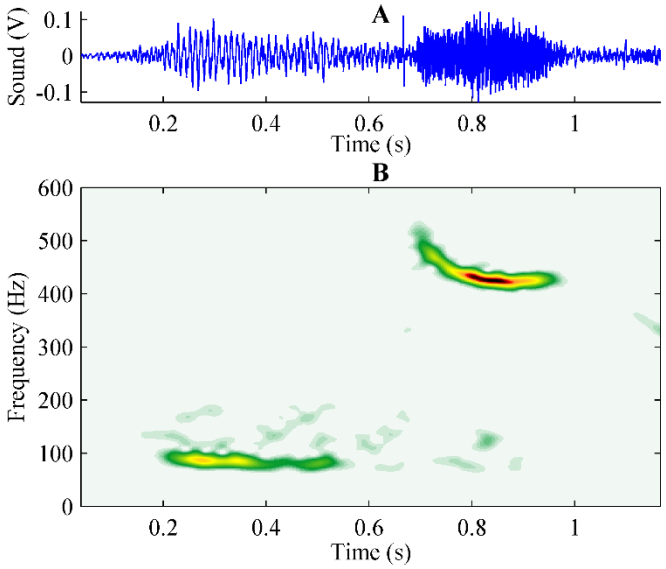


Figure 9. Spectrogram of an RS signal containing CAS. (A) RS signal containing two CAS. (B) Spectrogram. CAS appear in the spectrogram as ridges of energy around 85 Hz and 430 Hz.

Spectrogram—the square module of the short-time Fourier transform—has been the most used TFD for wheeze analysis [80]–[87]. Different techniques have been applied to spectrogram to identify wheezes. Some studies used a set of thresholds to detect local maximums that were likely to be wheezing peaks [80], [81], [84], [85]. Thereafter, the detected peaks were grouped using several temporal and spectral continuity criteria until wheezes were defined. Other studies treated spectrogram as an image and used digital image processing techniques to detect wheezes [82], [83], [87]. Zhang et al. [86] calculated the Shannon entropy of the RS spectrogram to distinguish between the power distribution pattern of normal RS and that of wheezes.

Analogous to spectrogram, scalogram—the square module of the continuous wavelet transform—was proposed by Taplidou et al. [88] to analyse wheezes. Additionally, more advanced TFDs have also been proposed for wheeze detection and analysis. Taplidou and Hadjileontiadis [89] combined the continuous wavelet transform with third-order statistics to calculate the evolutionary wavelet bicoherence, which was used to quantify the nonlinear characteristics of wheezes in the time-bi-frequency domain. Jin et al. [90] proposed a new TFD for RS analysis called temporal-spectral dominance spectrogram. Although

this TFD is calculated from the short-time Fourier transform, like spectrogram, it has higher resolution and enhances the identification of wheezes.

Besides TFDs, other methodologies have been proposed for classifying RS as normal or abnormal RS [91]–[94] or as wheezing or non-wheezing [95]–[97]. These studies have used standard pattern recognition methodology, which comprises three basic steps: feature extraction, dimensionality reduction or feature selection, and pattern classification. Classifying RS does not necessarily imply using time-frequency analysis techniques, such as TFDs, but other techniques can be used to extract features that allow different types of RS to be distinguished. Indeed, in previous studies, feature extraction was performed using wavelet coefficients [91], multiscale principal component analysis in Fourier domain [93], morphological complexities of RS (i.e. entropy, kurtosis, or skewness) [94], [97], or Mel frequency cepstral coefficients [95], [96]. These techniques allow CAS to be distinguished from other RS but do not allow CAS to be characterised. Characterising CAS requires the duration, mean frequency, and intensity of CAS to be calculated. To do this, time-frequency analysis techniques are required.

Although spectrogram has been widely used to characterise CAS, its poor resolution and its low energy concentration could prevent it from delimiting CAS accurately, especially weak CAS that overlap with normal RS in the time-frequency plane. In general, TFDs derived from Fourier or wavelet transforms have low energy concentration, which makes the detection of CAS highly dependent on amplitude criteria. Moreover, Fourier and wavelet-based techniques require *a priori* knowledge of the signal characteristics to correctly choose the analysis parameters. Further, analysing CAS using TFDs is more complex, since the entire time-frequency plane has to be processed to search for the ridges described by CAS (Figure 9). In this thesis, we propose a new approach to CAS analysis in order to overcome these issues in previous techniques (Section 1.5.4).

1.5. Proposed approach to RS analysis

There are four key aspects to be considered in RS analysis, as briefly outlined in the previous sections: the heterogeneous distribution of RS (Section 1.2.3), the airflow dependence of RS (Sections 1.2.2 and 1.4.1), the different characteristics of RS types (Section 1.2.1), and the effect of respiratory diseases on RS (Section 1.3). An approach to RS analysis that considers all these aspects is expected to highly contribute to improving both the study of RS and their applicability in clinical practice.

In this thesis, we addressed several technical issues to propose a new approach to RS analysis, taking all the aforementioned aspects into consideration. We began this thesis by analysing normal RS intensity based on the multichannel RS recording (Section 1.5.1), the progressive respiratory manoeuvre (Section 1.5.2), and the RS intensity–airflow curves (Section 1.5.3). We then developed a novel approach to CAS detection and characterisation based on the Hilbert-Huang transform (HHT) (Section 1.5.4). Finally, we proposed two clinical applications of our approach to RS analysis, as described in the next Sections 1.6.1.2 and 1.6.2.3.

1.5.1. Multichannel recording of RS

When recorded at the chest surface, RS contain regional information about pulmonary ventilation and the structure of the respiratory system. Using a single channel to record RS could fail to detect local changes in RS distribution due to localised AO or lung disorders. Thus, multichannel recordings using several sensors should be implemented to allow RS to be analysed at several points over the chest surface.

Our proposed approach to RS analysis includes a multichannel recording of RS using five contact microphones placed at the trachea, at the posterior base of the lungs, and near the posterior upper lobe of the lungs (Figure 10). High-frequency components of RS are better transmitted through the airways than through the lung parenchyma and are more clearly recorded over the trachea. Therefore, recording tracheal sounds provides distinct but complementary information to that provided by lung sounds [20], [39]. Moreover, the proposed four points for lung sound recording were the most used locations in previous studies and allow the distribution of RS to be analysed both horizontally and vertically, covering the commonly reported lung regions [35], [40], [41], [73], [98].

1.5.2. Airflow-dependent analysis of RS

Most previous studies analysed RS recorded without controlling airflow, at maintained airflows, or during forced expiratory manoeuvres. However, it is widely accepted that both normal RS and CAS highly depend on airflow. Therefore, it is important to perform airflow-dependent analyses of RS using well-controlled respiratory manoeuvres.

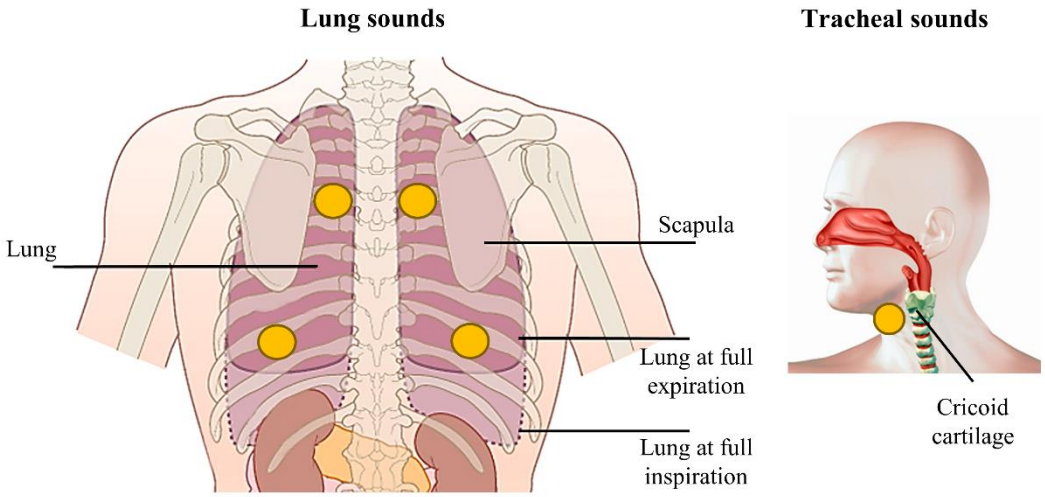


Figure 10. Recording points for lung and tracheal sounds.

This thesis proposes to perform a progressive respiratory manoeuvre with variable airflow for recording RS. For this, a person starts to breath normally, progressively breathes faster and deeper until the deepest breaths possible, and returns to normal breathing (Figure 11). In this way, both normal RS and CAS can be analysed depending on the total airflow range of each person.

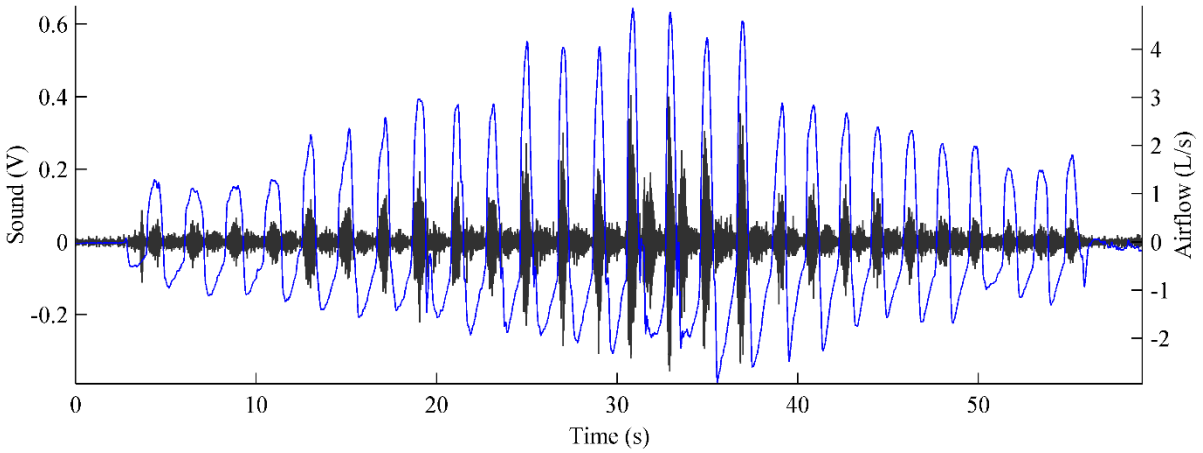


Figure 11. Progressive respiratory manoeuvre. Airflow signal (blue) and RS signal (grey) recorded during a progressive respiratory manoeuvre. RS were recorded at the posterior base of the right lung.

1.5.3. Analysis of normal RS based on the RS intensity–airflow curves

We propose calculating the RS intensity–airflow curves to analyse normal RS intensity. Thus, normal RS intensity is calculated as the area under the curve of the Welch's power spectral density estimate [99], within a particular frequency band (3).

$$I = \frac{f_m}{NFFT} \sum_{f=f_1}^{f_2} PSD(f) \tag{3}$$

In (3), I is RS intensity, f_m is the sampling rate of the recorded RS signal, $NFFT$ is the number of points used to calculate the fast Fourier transform (FFT), f_1 and f_2 are the frequency boundaries of the frequency band of interest, and PSD is the Welch's power spectral density estimate of the RS signal. After calculating the intensity of the RS signals corresponding to the respiratory cycles of a progressive

manoeuvre, each respiratory cycle is characterised with the parameters of the RS intensity and the peak flow. In this way, the RS intensity–airflow curves are obtained for each person.

1.5.4. Analysis of CAS based on HHT

1.5.4.1. The need for time-frequency analysis

We aimed to analyse not only normal RS but also CAS for this thesis. This dual approach to RS analysis required the development of a robust classifier that allowed CAS to be differentiated from normal RS. Further, CAS analysis requires going further than simple CAS identification—also the basic CAS characteristics, including duration, mean frequency, and intensity, must be calculated. Therefore, it is necessary to use a technique that allows CAS to be analysed simultaneously in both time and frequency domains.

TFD analysis is the most straightforward methodology in time-frequency analysis (Figure 12). TFDs represent the distribution of the energy of a signal throughout the time-frequency plane [100]. Therefore, TFDs can be used to simultaneously obtain information about the temporal and spectral content of CAS.

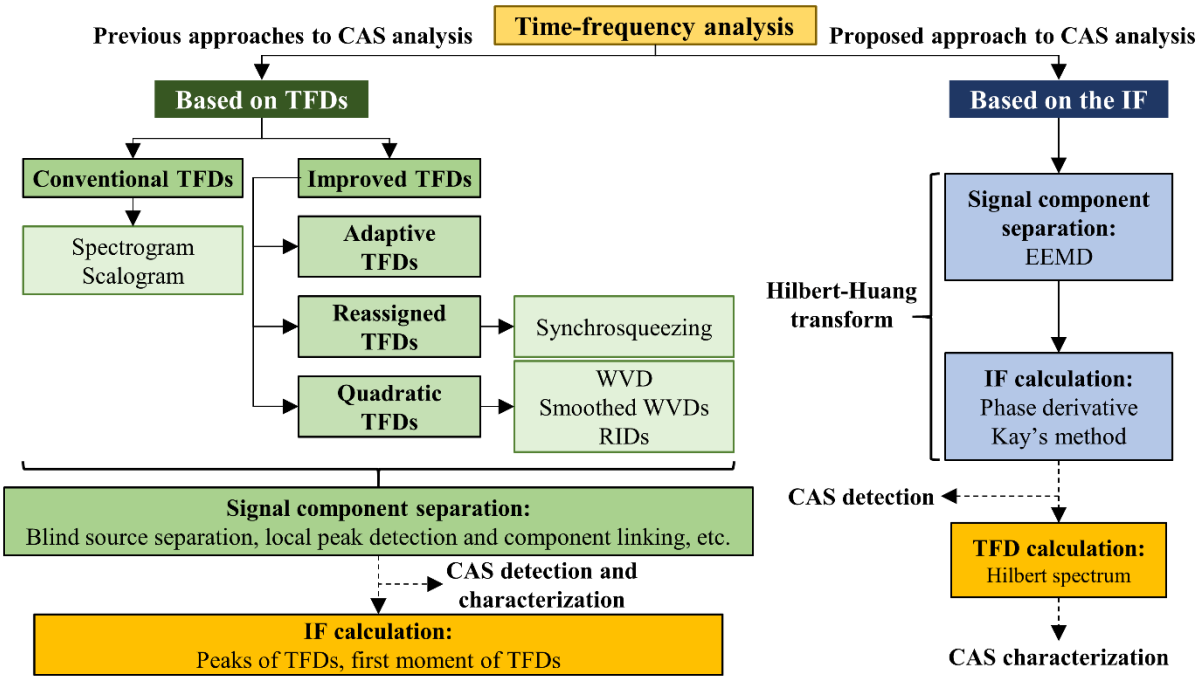


Figure 12. Proposed approach to CAS analysis based on HHT as compared to previous approaches based on TFDs. CAS, continuous adventitious sounds; EEMD, ensemble empirical mode decomposition; IF, instantaneous frequency; TFD, time-frequency distribution; WVD, Wigner-Ville distribution; RID, reduced interference distribution.

As described in Section 1.4.2, spectrogram was used in most previous studies that were focused on CAS analysis. For a real signal, $s(t)$, the spectrogram, $S(t,f)$, is defined as the squared magnitude of the short-time Fourier transform of the signal $s(t)$:

$$S(t,f) = \left| \int_{-\infty}^{\infty} s(\tau) w(\tau - t) e^{-j2\pi f \tau} d\tau \right|^2 \quad (4)$$

In (4), $w(\tau-t)$ is a real window centred at time $\tau = t$. A major limitation of spectrogram is that its resolution depends on the length of the window $w(\tau)$. Short windows provide good time resolution, but poor frequency resolution. In contrast, long windows provide good frequency resolution, but poor time resolution. Therefore, spectrogram compromises the two resolutions, which cannot be maximised at the same time, in accordance to the uncertainty principle [100]. On the other hand, the signal energy in spectrogram is scattered over the entire time-frequency plane, and spectrogram therefore has low energy concentration. Both the poor resolution and the low energy concentration could prevent spectrogram from delimiting CAS accurately, especially weak CAS with low intensity that overlap with the scattered energy of normal RS in the time-frequency plane.

Besides spectrogram, a few other techniques have been used for analysing CAS (Section 1.4.2), which are mainly based on the continuous wavelet transform [88], [89] and the temporal-spectral dominance spectrogram proposed by Jin et al. [90]. The continuous wavelet transform suffers from the same problems as spectrogram. The time and frequency resolutions of the continuous wavelet transform depend on frequency. At high frequencies, the time resolution is high but the frequency resolution is low, whereas at low frequencies, the frequency resolution is high but the time resolution is low. However, in each case, the two resolutions are affected by the uncertainty principle. In fact, any TFDs derived from integration-based transforms, such as Fourier or wavelet transforms, are affected by the uncertainty principle. Moreover, the conventional Fourier and wavelet-based techniques require *a priori* knowledge of the signal characteristics to correctly choose the analysis parameters.

The ideal TFD would have perfect resolution in both time and frequency domains and a high energy concentration. Different techniques, such as the adaptive short-time Fourier transform that uses a variable window length adapted to signal characteristics, have been proposed with the aim of increasing the resolution and energy concentration of conventional TFDs [101], [102]. A common approach to enhancing TFDs includes reassignment techniques [103], [104], whose objective is to reduce the energy spread in TFDs. Reassignment techniques are based on the reallocation of the signal energy in the time-frequency plane. A reassigned version of a TFD is obtained by moving its values from their original locations to the centre of gravity of their energy contributions, thus increasing energy concentration and producing a better localisation of signal components. A special type of reassignment technique is the synchrosqueezing transform, which allows energy concentration to be increased and modes of multicomponent signals to be retrieved [105].

1.5.4.2. The concept of instantaneous frequency and HHT

One of the most important parameters of time-frequency analysis is instantaneous frequency (IF), which—as its name indicates—consists of having values of the frequency content of a signal at each time instant [106], [107]. This parameter is extremely useful for the analysis of non-stationary signals, whose spectral content varies over time.

In the field of communication theory [108], Gabor defined the IF, $f(t)$, of a real signal, $s(t)$, from its corresponding analytic signal, $z(t)$, as follows:

$$z(t) = s(t) + jH[s(t)] = a(t)e^{j\theta(t)} \tag{5}$$

$$f(t) = \frac{1}{2\pi}\frac{d\theta(t)}{dt} \tag{6}$$

In (5), $H[\cdot]$ is the Hilbert transform, $a(t)$ is the absolute value of $z(t)$ and the instantaneous envelope (IE) of $s(t)$, and $\theta(t)$ is the phase of $z(t)$. The equation (6) yields frequency values at each time instant, so that the obtained IF sequence has maximum time resolution and provides information about the temporal evolution of the frequency content of a signal, regardless of its energy.

The concept of IF led to the development of a TFD that represents the true IF law of monocomponent signals in the time-frequency plane. In other words, the ideal TFD of a monocomponent signal would contain a single frequency value at each time instant. Under this premise, the Wigner-Ville distribution (WVD) [106], $W_z(t,f)$, was defined as:

$$W_z(t,f) = \int_{-\infty}^{\infty} z\left(t+\frac{\tau}{2}\right) z^*\left(t-\frac{\tau}{2}\right) e^{-j2\pi f\tau} d\tau \tag{7}$$

In (7), $z(t)$ is the analytic signal of a real signal. One of the most important properties of the WVD is that it yields the true IF values of monocomponent linear FM signals through its first moment with respect to frequency:

$$f(t) = \frac{1}{2\pi}\frac{d\theta(t)}{dt} = \frac{\int_{-\infty}^{\infty} fW_z(t,f)df}{\int_{-\infty}^{\infty} W_z(t,f)df} \tag{8}$$

However, a major drawback of WVD is that, when applied to multicomponent signals, it produces cross-terms. Due to its quadratic form, WVD produces spurious components that appear between the true components of a signal in the time-frequency plane. To reduce cross-terms, a number of quadratic TFDs, including the reduced interference distributions [109], have been proposed as smoothed versions of WVD. In fact, any quadratic TFD, $\rho_z(t,f)$, can be expressed as a filtered WVD using a specific time-lag kernel filter, $G(t,\tau)$, as follows:

$$\rho_z(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(t-u,\tau) z\left(u+\frac{\tau}{2}\right) z^*\left(u-\frac{\tau}{2}\right) e^{-j2\pi f\tau} du d\tau \tag{9}$$

The various existing quadratic TFDs propose different tradeoffs between energy concentration, resolution, and cross-term reduction [100]. In general, quadratic TFDs have higher resolution and energy concentration around the IF than spectrogram or scalogram, which allows the components of a signal to be identified more easily. However, cross-terms are a major drawback of quadratic TFDs, since they make it more difficult to identify the true components of a signal. When applied to RS signals, quadratic TFDs could lead to the detection of false CAS components.

Despite the advantages of reassignment techniques and quadratic TFDs, none of these techniques have been used in the field of RS analysis, especially for analysing CAS. Besides the extra computational cost of reassigned and quadratic TFDs, analysing CAS directly from TFDs implies working in a three-dimensional space, in which differentiating CAS from other RS is more complex, since the entire time-frequency plane has to be processed to search for the ridges described by CAS (Figure 9). Moreover, the commonly used method based on local peak detection and component linking over the spectrogram is highly dependent on amplitude criteria.

In this thesis, we now propose a new approach to CAS analysis based on the IF of RS signals (Figure 12). This approach includes detecting CAS with the assumption that the IF of an RS signal can reveal the difference between the deterministic nature (quasi-periodic) of CAS and the random nature of normal RS. In other words, the assumption is that IF values concentrated around the pitch of a CAS are obtained for as long as a CAS lasts within an RS signal, whereas more dispersed IF values are obtained when only normal RS are present. However, calculating IF only makes sense for monocomponent signals—that is, signals with a unique frequency component at each time instant. Therefore, estimating IF in multicomponent signals, such as RS, requires an extra step to separate out the different signal components.

Several IF estimation methods for multicomponent signals have been proposed using TFDs [110], [111]. Nevertheless, these methods require calculating TFD and then separating the different components of a signal in the time-frequency plane prior to estimating IF for each component (Figure 12). In the case of CAS analysis, using a TFD-based IF estimation method would imply detecting CAS prior to IF estimation and then using the IF estimates to detect CAS, which is paradoxical. Therefore, it makes sense to use the IF estimation method proposed in (6) rather than those based on TFDs.

As (6) yields a single frequency value at each time instant, it can be only applied to monocomponent signals. However, the approach we propose to analyse CAS uses HHT, which allows IF in multicomponent signals to be estimated without involving TFDs [112] (Figure 12). Prior to IF calculation, HHT uses empirical mode decomposition (EMD) to decompose a multicomponent signal into a set of narrowband components, called intrinsic mode functions (IMFs), for which IF can be calculated as in (6) [113]. Given a multicomponent real signal, $s(t)$, EMD allows the IMFs of $s(t)$ to be obtained by a sifting process that involves the following steps:

1) Identify of all the local extrema (both maxima and minima) of $s(t)$ and connect them using cubic spline interpolation, to obtain the upper and lower envelopes
2) Calculate the local mean of the two envelopes, $m(t)$, and subtract it from $s(t)$, to obtain $h(t)$
3) Repeat the first two steps on $h(t)$ until $m(t)$ is close to zero under certain criterion, usually based on the standard deviation of $m(t)$. Then, $h(t)$ is considered to be the first IMF, $c_1(t)$
4) Calculate the residue $r_1(t)$ by subtracting $c_1(t)$ from $s(t)$

5) Repeat steps from 1 to 4 on $r_1(t)$ until the obtained residue, $r_n(t)$, is a monotonic function and no further IMF can be extracted. Then, the original signal $s(t)$ can be expressed as:

$$s(t) = \sum_{i=1}^{n} c_i(t) + r_n(t) \quad (10)$$

Having decomposed a signal by EMD (Figure 13), the IF and IE sequences can be calculated from the analytic signal of each IMF as in (5) and (6) (Figure 14). After that, $s(t)$ can be expressed as:

$$s(t) = \sum_{i=1}^{n} a_i(t) \cos(\int 2\pi f_i(t) dt) + r_n(t) \quad (11)$$

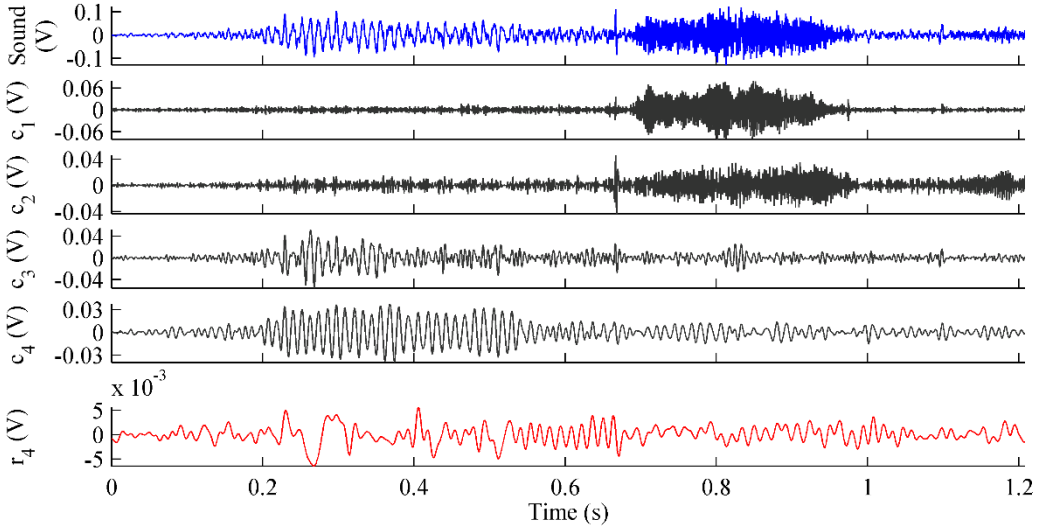


Figure 13. RS decomposition by ensemble empirical mode decomposition (EEMD). IMFs (c_1-c_4) and residue (r_4) of an RS signal that contains two CAS. The IMFs were calculated by EEMD, which improves on the performance of EMD.

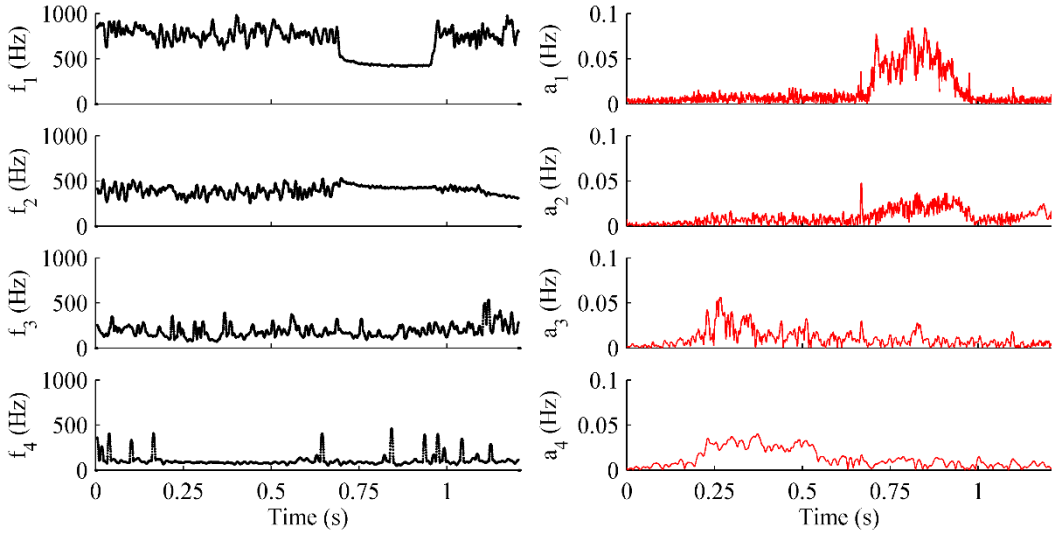


Figure 14. IF (f_1-f_4) and IE (a_1-a_4) sequences of the RS signal shown in Figure 13.

In (11), $a_i(t)$ and $f_i(t)$ are the IE and IF of the *i-th* IMF, respectively. Alternatively, the $a_i(t)$ and $f_i(t)$ sequences can be rearranged into an array to obtain an IF-based TFD called the Hilbert spectrum (HS) [114] (Figure 15D).

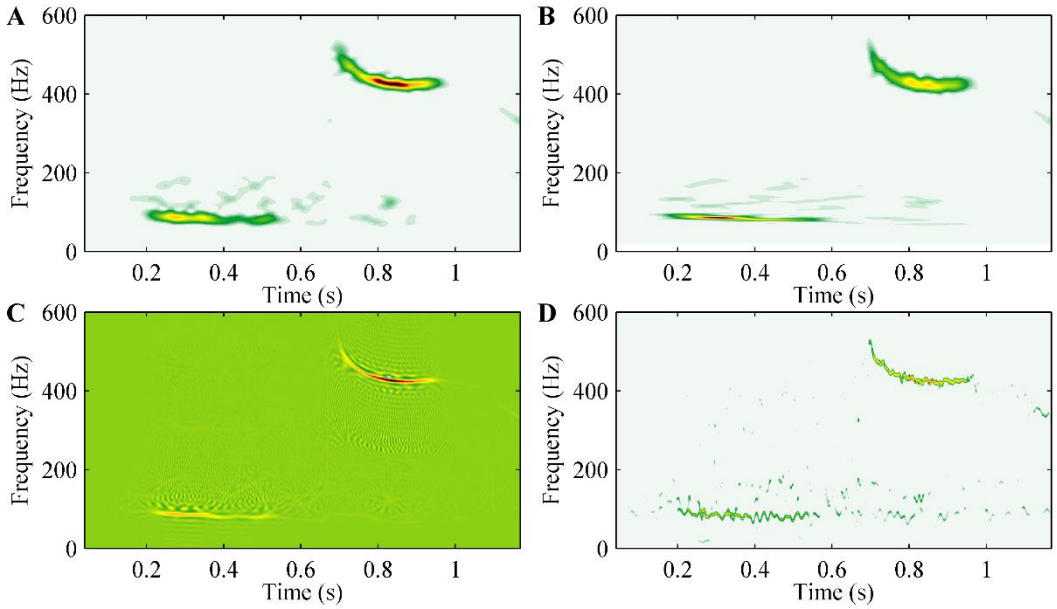


Figure 15. TFDs for RS analysis. Spectrogram (A), scalogram (B), WVD (C), and HS (D) of the RS signal shown in Figure 13.

Despite the many favourable properties of HHT [114], the original EMD method suffers from the mode mixing effect [112]. Ideally, each IMF calculated by EMD would contain a few different frequency components of a multicomponent signal. However, the mode mixing effect of EMD causes some components to appear within different IMFs, such that some IMFs contain widely different components. Although the mode mixing effect of EMD poses a problem for detecting CAS from the IF sequences of RS signals, the original EMD was used in an HHT-based approach to CAS analysis previously proposed by Reyes et al. [115], [116].

For the work in this thesis, we implemented an improved version of the original EMD, called ensemble EMD (EEMD), as the first step of HHT [117], [118]. EEMD consists of the iterative application of the original EMD to a signal plus different realizations of filtered white noise. The final IMFs are calculated as the mean of those resulting from each iteration. This method is based on the filter bank property of the original EMD when applied to white noise [119], [120]. By adding white noise to a multicomponent signal, EMD automatically separates its different components based on the reference scales set by white noise, thus avoiding the mode mixing effect.

Here, we originally used the proposed EEMD-based HHT to estimate IF and IE in RS signals containing normal RS and CAS (presented in Chapter 3 [121]). Afterwards, we developed a new method to automatically differentiate between CAS and normal RS, based on the different IF dispersion patterns of the two types of RS, as described above. A set of features extracted from the IF and IE sequences of the RS signals was used to train and to test a classifier, using support vector machines [122], [123].

In parallel to developing the aforementioned RS classifier, we evaluated in-depth the performance of the EEMD-based HHT for RS analysis (Chapter 4 [124]). For this, several technical difficulties of HHT had to be addressed and then overcome. First, we analysed the mode mixing effect of the original EMD

in RS signals. We then determined the performance for RS of the two most well-established and frequently-used solutions for mode mixing, the EEMD and the noise-assisted multivariate EMD [125], [126], which allowed us to propose numerous quantitative parameters to evaluate the performance of the two methods. Further, we tested and compared different IF estimation methods, including those based on the phase derivative of the analytic signal [127] and one based on the Teager energy operator [107], [128]. Kay's IF estimator [127] was chosen for proposing an EEMD–Kay-based HS. Additionally, we developed an automatic CAS segmentation and characterisation algorithm using the proposed HS. Finally, we compared the performance of spectrogram and the proposed HS for calculating the duration and pitch of CAS.

1.6. Applications of the proposed approach to RS analysis

To put into practice the proposed approach to RS analysis (Section 1.5), we conducted two studies in the Pulmonary Function Testing Laboratory of the Germans Trias i Pujol University Hospital (HUGTiP) in Badalona, Spain, in the course of this thesis. Patients with either unilateral phrenic paralysis (UPP) or asthma were recruited for performing pulmonary function tests and participating in the RS recording and analysis study. These two respiratory diseases affect the regional distribution of ventilation, and therefore RS, in different ways. UPP reduces regional ventilation of the lung on the paralysed side, which is more evident at the base of the lung [129], [130]. However, asthma may reduce regional ventilation at any point throughout the lungs, since AO can be either localized or diffused throughout the tracheobronchial tree [131]–[134]. Further, ventilation heterogeneity in asthma varies with the administration of bronchodilators [135]–[137]. Moreover, AO in asthma may be accompanied by the presence of CAS.

The two studies shared two points of the proposed approach to RS analysis: the multichannel recording of RS and the progressive respiratory manoeuvre that allows an airflow-dependent RS analysis. However, the RS recording protocol was customised for each study, as described in the next sections (1.6.1.2 and 1.6.2.3).

1.6.1. Assessment of UPP by RS analysis

Diaphragmatic paralysis consists of the loss of diaphragm motor activity due to either compression or sectioning of the phrenic nerve that activates the diaphragm (Figure 16) [138], [139]. UPP occurs when only one hemidiaphragm is affected, and it is caused mainly by nerve compression by tumours [138], nerve injury due to thoracic surgery or trauma [140], infections, or neurological disease. The effects of UPP are mild and variable and include shortness of breath and exercise limitation [139], [141].

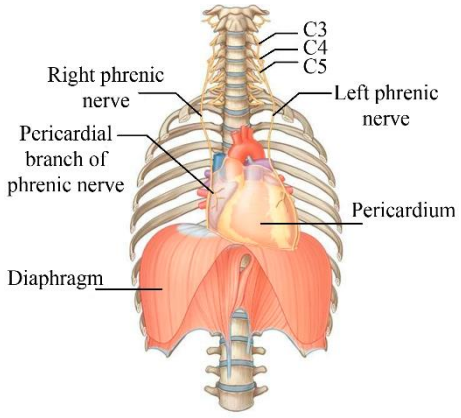


Figure 16. Motor innervation of the diaphragm by the phrenic nerves.

1.6.1.1. Previous work on the assessment of diaphragmatic paralysis

The techniques traditionally used to diagnose diaphragmatic paralysis are chest radiography, fluoroscopy, ultrasonography, and phrenic nerve stimulation [138], [139], [142]. Chest radiography is the simplest technique. In UPP, an elevated hemidiaphragm is usually seen on the affected side in chest

radiographs. However, in addition to requiring ionising radiation, chest radiography has low specificity and low positive predictive value for diagnosing UPP [143].

One of the most commonly used techniques for assessing diaphragm function is fluoroscopy [144], which allows diaphragm mobility to be assessed. In this technique, a continuous source of X-rays is used to produce real-time motion imaging of the diaphragm. Therefore, it is an invasive procedure, whose potential risks must be balanced with patient benefits. Ultrasonography can be used as an alternative to fluoroscopy [145]–[147]. Ultrasounds provide dynamic images of the diaphragm in a non-invasive way. However, ultrasonography is operator dependent and requires significant expertise.

Electrical or magnetic stimulation of the phrenic nerve has also been widely used to assess diaphragm function [148]–[158]. Several measurements of diaphragm strength following phrenic nerve stimulation have been proposed, such as transdiaphragmatic pressure, maximal static inspiratory mouth pressure, and sniff nasal inspiratory pressure [149]–[151], [154], [156]. Electromyography using either oesophageal electrodes [152], [153] or surface electrodes [155], [157], [158] has also been used to assess diaphragm function. In this case, the most relevant parameters are the mean amplitude of the compound muscle action potential and the phrenic nerve conduction time. However, phrenic nerve stimulation is an invasive technique. Further, most of these techniques require transnasal placement of either pressure catheters for measuring transdiaphragmatic pressure or oesophageal electrodes for measuring diaphragm electromyogram.

Recently, more advanced techniques for evaluating diaphragm function have been proposed, such as high-resolution computed tomography [142], [159] or an optoelectronic plethysmography system [160]. The former uses ionising radiation, and the latter is not easy to perform and is cost-prohibitive with reduced availability.

Lung function tests are usually performed to corroborate the diagnosis of diaphragm dysfunction, since they are easily available and non-invasive. Spirometric parameters, especially the total lung capacity, are usually decreased in patients with UPP, depending on the degree of weakness [139], [149], [161], [162].

An alternative way of measuring ventilation is to analyse RS. Tejman-Yarden et al. [163] used RS analysis to detect selective lung ventilation in intubated patients. In healthy people, slight differences of a few dB in the lung sound intensity between the lungs have been reported in many previous studies [36], [40]–[42].

RS analysis has not been previously used to assess UPP, as we propose in this thesis. A few previous studies have reported decreased lung sounds on auscultation over the affected side in patients with UPP [138], [164], [165]. However, these studies did not present quantitative lung sound measurements.

1.6.1.2. Proposed approach to assessing UPP by RS analysis

In this thesis, we propose to analyse RS intensity–airflow curves as an indirect measurement of pulmonary function in patients with UPP. At the beginning of this thesis, we launched a new RS recording protocol to assess UPP in the Pulmonary Function Testing Laboratory of the HUGTiP, based on a clinical protocol previously described by Fiz et al. [40].

Two piezoelectric contact microphones were placed on the surface of the patients' back, at the base of the lungs (Figure 17). All patients were asked to perform a progressive respiratory manoeuvre with variable airflow in a sitting position. In this study, we propose that analysing and comparing the RS intensity–airflow curves of the lungs provides powerful information for diagnosing UPP. Full details of the study are provided in Chapter 2 [166].

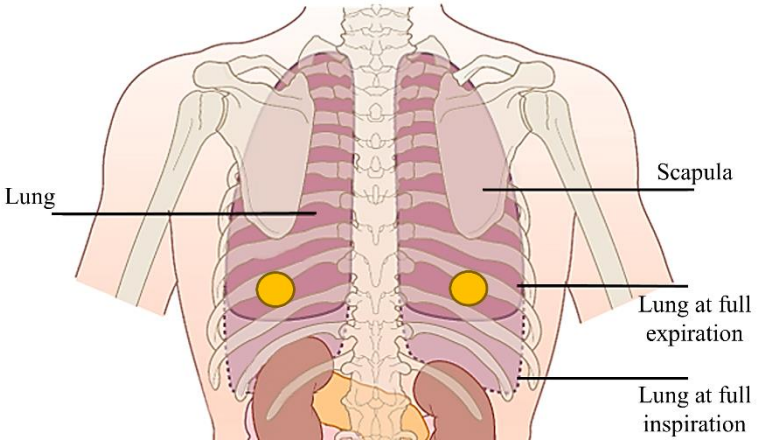


Figure 17. RS recording points to assess UPP.

1.6.2. Assessment of OPDs by RS analysis

As described in Section 1.3, airway conditions, such as OPDs, directly affect the generation of RS, as these depend on the airflow patterns within the airways. Airflow limitation in OPDs has two major effects on RS: the intensity of normal lung sounds is decreased over the obstructed areas, while CAS are generated. The analysis of both normal RS and CAS to assess OPDs has been the topic of extensive research, as described in Sections 1.6.2.1 and 1.6.2.2. However, as explained in the next section (1.6.2.3), some methodological issues in previous studies could have limited the potential of RS for assessing OPDs. In the work carried out for this thesis, we have addressed and overcome these issues, allowing us to propose a novel approach to RS analysis.

1.6.2.1. Previous work on normal RS and OPDs

Many previous studies have been focused on quantifying the effect of OPDs on normal RS. Pardee et al. [167] and Bohadana et al. [49] performed manual auscultation over six points of the chest surface in patients with different degrees of AO and scored the perceived RS intensity. The former found a strong correlation between RS intensity and FEV_1, whereas the latter reported close relationships between RS intensity and several indices of airflow obstruction, such as maximal expiratory flow at 50% of vital

capacity (VC), FEV_1, or the FEV_1/VC ratio. Both studies concluded not only that reduced RS intensity was a clear sign of airflow obstruction, but also that RS intensity is insensitive to mild degrees of airflow obstruction. The decreased RS intensity reported in these studies was closely related to the decreased inspiratory flows of the patients.

During the '90s, many studies used bronchoprovocation tests to analyse the effects of airway narrowing on RS [98], [168]–[172]. In the absence of wheezes, Bohadana et al. [168] reported that RS intensity, measured as the RMS of the inspiratory sound signals, decreased with bronchoconstriction, and that this change was almost completely reverted in most patients after inhaling a bronchodilator. Since airflow was neither recorded nor controlled, reductions in RS intensity were supposed to be due to decreased inspiratory airflows during bronchoconstriction. However, subsequent studies reported that, when compared at similar airflows, RS intensity increased during induced airway narrowing in both healthy participants and patients with asthma [169], [170]. Moreover, at similar levels of induced bronchoconstriction, the increase in RS intensity was higher in patients with asthma than in healthy participants. The explanation for these findings was that, at similar mouth airflows, the velocity of airflow within the airways is higher during bronchoconstriction than at baseline. The increased velocity of airflow enhances turbulence, thus generating louder RS.

Other bronchoprovocation studies reported that induced airway narrowing caused significant changes in the RS power spectrum. Malmberg et al. [169] and Anderson et al. [171] found a strong correlation between the median frequency of the RS power spectrum and changes in FEV_1 induced by histamine. They found that airway narrowing caused the RS median frequency to increase, making it a clear indicator of AO and airflow limitation. Pasterkamp et al. [98] revealed that airway narrowing was accompanied by a decrease in power at low frequencies during inspiration and an increase in power at high frequencies during expiration. Habukawa et al. [172] showed that the highest frequency of inspiratory and expiratory RS increased during methacholine inhalation challenge, and that this corresponded with a decrease in the forced expiratory parameters.

At baseline, the median frequency of RS was found to be significantly different between patients with asthma and both healthy participants and patients with COPD [173]. After bronchodilation, the centroid frequency of the spectrum of tracheal sounds was found to clearly decrease in patients with asthma but to show no significant changes in healthy participants [77]. These differences between both groups were found at different airflow ranges, of between 0.4 and 1.2 L/s during forced spirometry manoeuvres. More recently, Habukawa et al. [174] developed a new method to assess the control level of asthma based on the analysis of lung and tracheal sounds. They calculated the acoustic transfer characteristics of the pulmonary system using the cross-spectrum between tracheal and chest wall sounds. They then calculated two indices—the chest wall sound index from the transfer characteristics, and the tracheal sound index. Based on the two-dimensional diagram of these indices, they developed the breath sound

index, which differs significantly between children with asthma and healthy children and between well-controlled and not well-controlled asthmatic children.

During the past decade, the VRI system has led to a number of studies focusing on the effect of OPDs on the regional distribution of RS intensity. Inhomogeneity in ventilation and asynchronous airflow during asthma exacerbations were found to cause significant asynchrony between the expiratory vibration energy peaks of the lungs, and correspondingly, the asynchrony reverted towards the normal synchronous pattern after clinical improvement of the patients [175]. Inhomogeneity of VRI images was also observed in patients with asthma and COPD, who had abnormally loud and prolonged RS during expiration due to airflow obstruction [176].

Patients with asthma could also be distinguished from patients with COPD by qualitatively evaluating the dynamic VRI images and analysing the quantitative data of synchronization between the lungs during bronchodilation tests [177].

Using the VRI system to calculate the ratio of lung sound intensity between the lower and upper lungs revealed a significantly decreased ratio in patients with COPD as compared to healthy participants, most likely due to hyperinflation and redistribution of lung sounds to the upper lung areas [178]. The same technique was also used to evaluate the bronchodilator effect in patients with COPD, for which the ratios significantly increased after treatment [179].

In summary, normal RS analysis has been mainly based on the calculation of the following parameters:

- *Intensity*, calculated as either the RMS of RS signals in time domain or as the area under the power spectrum in a particular frequency band. Different frequency bands were used in previous studies
- *Frequency parameters*, including median frequency, centroid frequency, and highest frequency
- *Synchrony* between the intensity peaks of the lungs using dynamic VRI images

1.6.2.2. Previous work on CAS and OPDs

As described in Section 1.3.2, CAS are related to obstructive airway diseases and are an unequivocal sign of airflow limitation. Accordingly, many previous studies have focused on analysing CAS features and their relation to pulmonary function in OPDs. Table 1 contains detailed information about some of the most relevant studies on CAS analysis in OPDs. The main findings can be summarised as follows:

- *Number of wheezes*: a higher number of wheezes is associated with positive BDR and positive bronchoconstriction tests. The number of wheezes usually decreases after bronchodilation, with a greater change in patients with OPDs than in healthy people. A higher number of wheezes is associated with lower FEV_1 and lower peak expiratory flow. The absence of pulmonary obstruction usually correlates with no detected wheezes
- *Features of wheezes*: biphasic wheezes are associated with lower peak expiratory flows than monophasic wheezes. Longer, higher-pitched, and louder wheezes are associated with lower peak expiratory flows. Both duration and pitch of wheezes vary with bronchodilation

Table 1. Analysis of CAS in OPDs

Authors	*Sensors / Recording Points*	*Subjects*	*Manoeuvres / Tests*	*Type of RS*	*Methodology / Extracted Features*	*Significant Findings*
Marini et al. (1979)[180]	Manual auscultation / several areas anteriorly and posteriorly	83 patients with chronic airflow obstruction	Deep unforced breathing and forced expiration / BDR	Wheezes	Regional scoring / frequency and intensity: from 0 (no wheezing) to 3 (loud wheezing in every expiration)	- Unforced wheezes, 48 patients; forced wheezes, 80 patients - Positive BDR more prevalent in patients with unforced wheezing - Moderate positive correlations between the unforced wheezing score and both the degree of obstruction and BDR
Shim and Williams (1983)[181]	Manual auscultation / several areas anteriorly, laterally, and posteriorly	93 patients with asthma	Quiet breathing	Wheezes	Regional grading / duration, pitch, and intensity (minimal, moderate, or severe)	- PEFR patients with wheezing < PEFR patients without wheezing - Expiratory wheezing often accompanied by inspiratory wheezing - PEFR in biphasic wheezing < PEFR expiratory wheezing alone - Wheezing of high pitch, moderate to severe intensity, and spanning the entire phase of the breath associated with lower PEFR
Baughman and Loudon (1984)[71]	Electronic stethoscope / 4 areas posteriorly right/left apex/bases	20 patients with asthma	Tidal breathing / BDR	Wheezes	Successive FFT, window: 250 ms, 60% overlap / T_w/T_{tot}, frequency, and intensity	- Moderate negative correlation between T_w/T_{tot} and FEV_1 - T_w/T_{tot} (from 86% to 31%) and pitch (from 440 Hz to 298 Hz) reduced with medication
Fenton et al. (1985)[39]	2 contact accelerometers / right upper chest anteriorly and neck	5 asthmatic children and 2 controls	Normal breathing / BDR	Wheezes	Successive FFT / identification of peaks above 200 Hz with a power 15-times greater than average	- Wheezing strongly depended on airflow - Wheezing followed the changes in FEV_1 - The trachea was the best location for wheeze analysis
Baughman and Loudon (1989)[54]	Electronic stethoscope / anterior trachea and right upper chest anteriorly	5 patients with extrathoracic obstruction, 25 asthma patients, and 7 controls	Normal breathing	Stridors	Successive FFT, window: 100 ms, 50% overlap / identification of peaks above 200 Hz, greater than 3-times the baseline signal, and lasting more than 200 ms	- Stridor was associated with extrathoracic obstruction and was similar to wheezing - Timing and location of stridors (mainly inspiratory, over the neck) was significantly different from wheezes (mainly expiratory, over the chest)
Schreur et al. (1994)[72]	3 piezoelectric microphones / right chest anteriorly and posteriorly	9 patients with asthma and 8 healthy controls	Quiet breathing up to 1.5 L/s and forced manoeuvres	Normal and wheezes	Successive FFT at different airflows / identification of peaks above 150 Hz, greater than 3-times the baseline signal. LSI from power spectra, quartile frequencies, and extent of wheezing	- At low airflows, expiratory LSI in patients with asthma < expiratory LSI in controls - During quiet expiration, quartile frequencies in asthmatics > quartile frequencies in controls - Change of quartile frequencies with flow in asthmatics > change of quartile frequencies with flow in controls
Gavriely et al. (1994)[73]	4 piezoelectric sensors / trachea, right chest anteriorly, and right/left bases posteriorly	493 active workers	Maintained airflows around 1 L/s	Normal and abnormal	Averaged power spectra by FFT and linear regression / identification of abnormal spectra with peaks of power	- Combination of spirometry and RS analysis increased the sensitivity from 71% to 87% detection of pulmonary diseases
Rietveld et al. (1994)[74]	1 electret microphone / trachea	28 asthmatic children	Quiet breathing / histamine challenge	CAS	Audition of the RS signals and visualization of the power spectra / 5 different RS patterns, including different CAS	- The presence of one of the RS patterns during unforced breathing predicted a fall in FEV_1 of > 20% in 21 subjects
Bohadana et al. (1995)[75]	Electronic stethoscope / right base posteriorly	38 patients	Fast and deep inspirations / carbachol airway challenge	Normal and wheezes	TEWA and Fourier spectrum for wheeze detection and RMS for calculating RS intensity	- Wheezes detected after bronchoconstriction more often in positive patients (10/21) than in negative patients (1/17) - RS intensity decreased in the 11 non-wheezing positive patients but not in the negative patients
Shreur et al. (1996)[76]	3 piezoelectric microphones / right chest anteriorly and posteriorly	8 patients with asthma	Quiet breathing and forced	Normal and wheezes	Successive FFT at different airflows / identification of peaks above 150 Hz, greater than 3-times the baseline signal.	- LSI, quartile frequencies, and extent of wheezing were higher during the bronchoconstriction phases

			manoeuvres / allergen challenge		LSI from power spectra, quartile frequencies, and extent of wheezing	
Fiz et al. (2002)[182]	1 contact microphone / trachea	16 asthma patients, 6 COPD patients, and 15 healthy participants	Forced expiration / BDR	Wheezes	Spectrogram / mean frequency, number of wheezes, % of wheezing time, and change in the number of wheezes	- More wheezes recorded in patients than in control participants - Mean frequency of wheezes in control participants > mean frequency of wheezes in patients - Change in the number of wheezes in patients > change in the number of wheezes in control participants
Kraman et al. (2002)[55]	1 electret microphone / trachea	1 patient with tracheal obstruction	Normal breathing	Stridors	Spectrogram	- Tracheal obstruction could be detected through the analysis of tracheal sounds
Mazic et al. (2003)[183]	1 electret microphone and accelerometers / trachea and right base posteriorly	7 asthmatic children	Normal breathing and forced breathing	Wheezes	Successive spectra / duration and pitch of wheezes	- Wheezing was detected in 70% of patients during asthmatic seizures - No wheezing was detected in the absence of pulmonary obstruction
Bentur et al. (2003)[78]	5 piezoelectric contact sensors / trachea, right/left axillae, and right/left bases posteriorly	12 asthmatic and 7 healthy children	Normal breathing / 6 weeks treatment	Wheezes	FFT-based algorithm for wheeze detection / % of wheezing time to breathing time and extent of wheezing for the total night	- The extent of wheezing decreased after 6 weeks of treatment, and FEV_1 improved
H-Corbera et al. (2004)[81]	1 piezoelectric sensor / trachea	16 patients with asthma and 15 healthy people	Forced expiration / BDR	Wheezes	Spectrogram / mean frequency, number of wheezes, % of wheezing time, and change in the number of wheezes	- Differences in all parameters between both groups - Change in the number of wheezes in patients > change in the number of wheezes in control participants
Sánchez et al. (2005)[79]	2 sensors / right/left bases posteriorly	22 infants with acute bronchiolitis	Normal breathing / BDR	Normal and wheezes	Successive FFT at flows around 0.1 L/s / quartile frequencies and identification of wheezes from peaks in the power spectra and auditory verification	- Wheezes, 11 patients; wheezes and crackles, 11 patients - Positive relationship between positive BDR and an increase in power at low frequencies after medication - Positive correlation between wheezing and the increase in the power spectra measured by the quartile frequencies
Fiz et al. (2006)[184]	1 microphone / trachea	20 patients with asthma and 14 healthy people	Forced expiration / BDR	Wheezes	Spectrogram / mean frequency, number of wheezes, % of wheezing time, and change in the number of wheezes	- Number of wheezes and change in the number of wheezes higher in obstructed patients than in stable patients and controls - Mean frequency of wheezes higher in control participants
Taplidou and L.J.H. (2010)[185]	5 electret microphones / trachea, right/left axillae, and right/left bases posteriorly	10 COPD and 11 asthma patients	Normal breathing at flows ≤ 1.5 L/s	Wheezes	Instantaneous wavelet bicoherence / 23 nonlinear features of wheezes	- 22 out of 23 features showed significant difference between the COPD and asthma patients
Oliveira et al. (2013)[186]	Digital stethoscope / several areas anteriorly, laterally, and posteriorly	6 subjects with LRTI	Normal breathing / 3 weeks treatment	Wheezes	Spectrogram / number, duration, and mean frequency of wheezes	- Strong negative correlations (post-treatment) between FEV_1 and duration of wheezes at the lateral right region, and between FVC and duration of wheezes at the posterior right region
Fischer et al. (2016)[187]	2 piezoelectric contact sensors / manubrium and left axillary line	110 infants	Normal breathing / Lung function tests	Wheezes	FFT-based algorithm for wheeze detection / duration of wheezes	- Wheezing detected in 65% of infants (39% with inspiratory wheezing and 48% with expiratory wheezing) - Airway resistance increased in all infants with wheezes and correlated with the duration of expiratory wheezes

BDR, bronchodilator response; CAS, continuous adventitious sounds; COPD, chronic obstructive pulmonary disease; FEV_1, forced expiratory volume in 1 second; FFT, fast Fourier transform; FVC, forced vital capacity; LRTI, lower respiratory tract infection; LSI, lung sound intensity; PEFR, peak expiratory flow rate; RMS, root mean square; RS, respiratory sounds; TEWA, time expanded waveform analysis; T_w/T_{tot}, proportion of the respiratory cycle occupied by a wheeze.

1.6.2.3. Proposed approach to assessing OPDs by RS analysis: application to BDR assessment in asthma

After reviewing previous studies on RS analysis for assessing OPDs, the following issues were found:

- A number of different recording locations have been used in previous studies. Using a large number of channels increases the complexity of recording and analysing RS, due to the large amount of data, and it does not necessarily yield better outcomes. On the other hand, using only one microphone can lead to failure to detect changes in regional ventilation
- Typically, RS signals are recorded either at maintained airflows or during forced expiratory manoeuvres. However, both normal RS and CAS are highly dependent on airflow
- Previous studies focused either on normal RS or on CAS. Co-analysis of the two types of RS was addressed only in a few studies, most of which were early studies that used inaccurate techniques for detecting and characterising CAS
- Unlike CAS duration and frequency, their intensity has rarely been analysed. In fact, some studies only analysed the number of CAS

For this thesis, we propose a new integrated approach to RS analysis in OPDs. This proposed approach includes a multichannel recording of RS using five contact microphones (Section 1.5.1) and performing progressive respiratory manoeuvres with variable airflow (Section 1.5.2). We propose a dual approach to RS analysis, by analysing both normal RS intensity and CAS through the RS intensity–airflow curves (Section 1.5.3) and HHT (Section 1.5.4), respectively. This dual approach allows the two effects of OPDs on RS to be assessed. Moreover, the proposed HHT-based methodology allows CAS to be fully characterised with respect to duration, mean frequency, and intensity.

Based on the aforementioned approach, a study was launched in the Pulmonary Function Testing Laboratory of the HUGTiP for assessing BDR in patients with asthma (Figure 18).

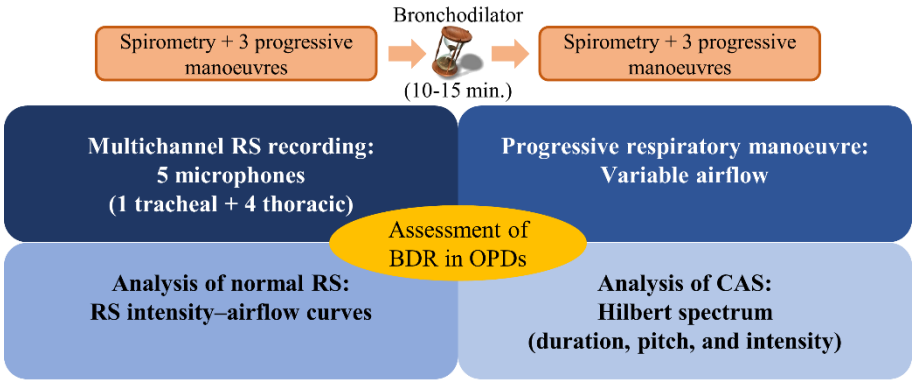


Figure 18. Proposed approach to assessing BDR in OPDs by RS analysis. BDR, bronchodilator response; OPDs, obstructive pulmonary diseases; RS, respiratory sounds.

According to the definition from the Global Initiative for Asthma: "*Asthma is a heterogeneous disease, usually characterized by chronic airway inflammation. It is defined by the history of respiratory*

symptoms such as wheeze, shortness of breath, chest tightness and cough that vary over time and in intensity, together with variable expiratory airflow limitation" [6].

Several factors must be considered for an asthma diagnosis:

- A detailed clinical history
- A complete physical examination to identify characteristic respiratory symptoms
- Pulmonary function tests to assess airflow limitation

Lung auscultation is usually performed to identify wheezes in asthma. However, traditional manual auscultation highly depends on the subjectivity of the physicians involved. In this sense, using RS analysis allows wheezes, and other types of CAS, to be identified in more objective and practical manner. On the other hand, spirometry is a simple and useful technique for assessing airflow limitation. When airflow obstruction is suspected based on spirometric parameters, a BDR test is usually performed to support or reject an asthma diagnosis. However, as described in Section 1.1.3, spirometry and the BDR test are still controversial even today, and any diagnosis of asthma should not be made based only on spirometric parameters.

A combination of spirometry and RS analysis can increase the sensitivity of pulmonary disease detection [73]. Therefore, for this thesis, we used both spirometry and RS analysis to measure BDR in patients with asthma (Figure 18). In the study conducted at HUGTiP, RS were recorded in patients with asthma as they performed six progressive manoeuvres with variable airflow, three of which were pre-bronchodilation and three, post-bronchodilation. Moreover, BDR was measured in all participants by spirometry. To date, RS signals and spirometric data have been obtained from a total of 50 patients with asthma and 10 control participants.

A preliminary study showing the potential of the proposed approach to assessing BDR in patients with asthma based on the analysis of normal RS and CAS was presented at the XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013 (MEDICON 2013) and published as a book chapter in 2014 [188]. Specifically, we presented a case study of three adult patients with asthma who had different baseline spirometric values and BDR. Further details are provided in Chapter 5.

Before applying the proposed approach to assessing BDR (Figure 18) to the whole database of patients with asthma, we carried out a prior study with 10 patients with asthma and 5 control participants from the database to evaluate the CAS analysis part of the approach. Specifically, we evaluated how well our CAS characterisation algorithm (proposed in Chapters 3 [121] and 4 [124]) works in RS signals recorded in the clinical environment, and the potential of CAS analysis for assessing BDR. Besides CAS, RS signals contain normal RS and background noises from the clinical environment, which may appear in the HS as ridges similar to those described by CAS (Figure 15) and could lead to an overestimation of CAS. To avoid analysing false CAS, we trained and validated a classifier to distinguish signal

components corresponding to CAS from those corresponding to other sounds. We also studied how the analysis of CAS features, including number, duration, pitch, and intensity, could improve the stratification of BDR levels. We have now completed this study and (at the time of this writing) have submitted it to PLOS ONE (see publications derived from this thesis).

1.7. Hypotheses

We based this current work on a number of hypotheses that we posed regarding different technical and clinical aspects of RS analysis. We posed several technical hypotheses regarding the analysis of normal RS using the RS intensity–airflow curves and the analysis of CAS based on HHT and HS.

We also posed several clinical hypothesis regarding the analysis of RS for assessing pulmonary function in UPP and BDR in asthma.

Analysis of normal RS

- Normal RS intensity is affected by airflow limitation, and this effect may be more noticeable at high airflows
- The RS intensity–airflow curves may contribute to improving the assessment of changes and asymmetries in normal RS intensity due to respiratory diseases

Analysis of CAS

- Using EEMD may allow RS signals to be decomposed into narrowband components, thereby avoiding the mode mixing effect of EMD and allowing physically meaningful IF to be calculated
- IF dispersion might be markedly decreased within RS signals for the duration of CAS, such that CAS could be detected from changes in IF dispersion
- HS may provide high energy concentrations around the ridges described by CAS components, which would make HS less dependent on amplitude criteria for CAS characterisation
- CAS may be more accurately determined with HS than with spectrogram

Assessment of pulmonary function in patients with UPP

- There might be significant differences between the RS intensity–airflow curves at the posterior base of the lungs in patients with UPP, and the intensity in the affected side might be significantly lower than the intensity in the healthy side

Assessment of BDR in patients with asthma

- The RS intensity–airflow curves may change significantly at one or several locations over the chest surface after bronchodilator administration in patients with asthma
- The distribution, number, and features of CAS may vary significantly following bronchodilator administration in patients with asthma
- The number of CAS and their loudness may be higher over obstructed areas. However, normal RS intensity may be lower over obstructed areas
- Both the number of CAS at baseline, and the change in the number of CAS after bronchodilator administration, may vary with airflow in patients with asthma

1.8. Objectives

1.8.1. Main objective

Here, we aimed to develop a new approach to analysing normal RS and CAS, to create a more sensitive and alternative tool for assessing the pulmonary function in terms of acoustic parameters. The proposed approach includes multichannel RS recording, a novel and useful progressive respiratory manoeuvre with variable airflow, and the use of advanced digital signal processing techniques to better characterise and interpret RS. This approach to RS analysis is more objective than traditional auscultation and provides distinct, but complementary, information to that provided by spirometry.

This thesis therefore was divided into two major approaches: normal RS analysis and CAS analysis. Moreover, the dual proposed approach to RS analysis was carried out by the assessment of both pulmonary function in patients with UPP and BDR in patients with asthma. Some common specific objectives are proposed regarding the RS recording protocol. Further, specific objectives are proposed separately for the analysis of normal RS and CAS, as well as for the two clinical applications of the proposed approach to RS analysis.

1.8.2. Specific objectives

RS recording protocol

- Design a protocol for the multichannel recording of RS and respiratory airflow based on the progressive respiratory manoeuvre and considering the number and location of sensors as a key aspect
- Create a database of RS and airflow signals recorded in patients with asthma and patients with UPP. To do this, we aimed to launch two RS recording protocols in the Pulmonary Function Testing Laboratory of the HUGTiP

Analysis of normal RS

- Analyse the relationship between normal RS intensity—calculated as the mean power within a particular frequency band of the PSD—and airflow through the RS intensity–airflow curves, and apply linear regression models

Analysis of CAS

- Propose an HHT adapted to the characteristics of RS signals, using EEMD and several IF estimation methods
- Evaluate the performance of HHT for the time-frequency analysis of RS and analyse its advantages with respect to other time-frequency analysis techniques
- Develop an algorithm for the automatic detection of CAS based on the IF and IE sequences of RS
- Characterise CAS from HS using number, duration, mean frequency, and intensity

- Compare the performance of HS for characterising CAS with that of traditional techniques, such as spectrogram

Assessment of pulmonary function in patients with UPP

- Analyse the differences between the RS intensity–airflow curves of the lungs as an alternative method for detecting UPP

Assessment of BDR in patients with asthma

- Analyse the changes in the RS intensity–airflow curves at several points over the chest surface following bronchodilator administration
- Analyse the changes in the number of CAS and their features following bronchodilator administration
- Analyse the relationship between acoustic parameters and spirometric parameters

Chapter 2: Detecting unilateral phrenic paralysis by acoustic respiratory analysis

Title: Detecting unilateral phrenic paralysis by acoustic respiratory analysis

Authors: J. A. Fiz, R. Jané, M. Lozano, R. Gómez, J. Ruiz

Journal: PLOS ONE, 2014

OPEN ACCESS Freely available online

PLOS ONE

Detecting Unilateral Phrenic Paralysis by Acoustic Respiratory Analysis

José Antonio Fiz[1,2,3], Raimon Jané[2,3,4]*, Manuel Lozano[2,5], Rosa Gómez[1], Juan Ruiz[1]

1 Pneumology Service, Germans Trias i Pujol University Hospital, Badalona, Spain, **2** Institute for Bioengineering of Catalonia (IBEC), Barcelona, Spain, **3** Biomedical Research Networking Center in Bioengineering, Biomaterials, and Nanomedicine (CIBER-BBN), Barcelona, Spain, **4** Dept. ESAII, Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, **5** Innovation Group, Health Sciences Research Institute of the Germans Trias I Pujol Foundation (IGTP), Badalona, Spain

Abstract

The consequences of phrenic nerve paralysis vary from a considerable reduction in respiratory function to an apparently normal state. Acoustic analysis of lung sound intensity (LSI) could be an indirect non-invasive measurement of respiratory muscle function, comparing activity on the two sides of the thoracic cage. Lung sounds and airflow were recorded in ten males with unilateral phrenic paralysis and ten healthy subjects (5 men/5 women), during progressive increasing airflow maneuvers. Subjects were in sitting position and two acoustic sensors were placed on their back, on the left and right sides. LSI was determined from 1.2 to 2.4 L/s between 70 and 2000 Hz. LSI was significantly greater on the normal (19.3±4.0 dB) than the affected (5.7±3.5 dB) side in all patients ($p = 0.0002$), differences ranging from 9.9 to 21.3 dB (13.5±3.5 dB). In the healthy subjects, the LSI was similar on both left (15.1±6.3 dB) and right (17.4±5.7 dB) sides ($p = 0.2730$), differences ranging from 0.4 to 4.6 dB (2.3±1.6 dB). There was a positive linear relationship between the LSI and the airflow, with clear differences between the slope of patients (about 5 dB/L/s) and healthy subjects (about 10 dB/L/s). Furthermore, the LSI from the affected side of patients was close to the background noise level, at low airflows. As the airflow increases, the LSI from the affected side did also increase, but never reached the levels seen in healthy subjects. Moreover, the difference in LSI between healthy and paralyzed sides was higher in patients with lower FEV_1 (%). The acoustic analysis of LSI is a relevant non-invasive technique to assess respiratory function. This method could reinforce the reliability of the diagnosis of unilateral phrenic paralysis, as well as the monitoring of these patients.

Citation: Fiz JA, Jané R, Lozano M, Gómez R, Ruiz J (2014) Detecting Unilateral Phrenic Paralysis by Acoustic Respiratory Analysis. PLoS ONE 9(4): e93595. doi:10.1371/journal.pone.0093595

Editor: Thomas Penzel, Charité - Universitätsmedizin Berlin, Germany

Received August 8, 2013; **Accepted** March 7, 2014; **Published** April 9, 2014

Funding: This work was made possible thanks to a collaboration agreement, between IBEC and IGTP, to create a joint research Unit, and it was supported in part by the Spanish Ministry of Economy and Competitiveness under grant TEC2010-21703-C03-01. No additional external funding received for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: rjane@ibecbarcelona.eu

Introduction

There are several causes of diaphragmatic dysfunction that can affect one or both muscles. The decrease in or cessation of motor activity can be caused by compression or section of the phrenic nerve in certain segments of the spinal cord [1]. The consequences of diaphragm dysfunction vary from the most serious cases of bilateral lesions that can require mechanical ventilation, to the mildest unilateral lesions that may to some extent impair breathing and in consequence exercise capacity [2,3].

Diaphragm dysfunction due to phrenic paralysis has been studied with various techniques including x-ray, fluoroscopy, ultrasonography, and external or internal stimuli of the diaphragm. These techniques provide information regarding the position and mobility of the diaphragm muscle [4–7], but do not predict the degree of respiratory dysfunction [8].

On the contrary, breathing function can be measured by routine spirometry [9,10]. Recently, Sokolowska et al. measured variations in breathing patterns in animals with bilateral phrenic paralysis, confirming that the measurement of breathing parameters could be an appropriate method to monitor this diaphragm dysfunction [11]. However, in cases of unilateral paralysis, spirometric function may be normal.

An alternative useful method to monitor breathing function is the measurement of pulmonary sounds [12–16]. In fact, it is known that airflow is correlated with lung sound intensity (LSI) [17], including in pulmonary conditions with restrictive ventilatory function [16].

Our hypothesis in the present study was that in patients with unilateral phrenic paralysis, the LSI on inspiration would be lower on the affected side than the healthy side. If this hypothesis were to be confirmed, measurements of LSI comparing the two sides could be useful to diagnose conditions associated with restricted thoracic mobility [6], as well as to monitor the response to specific physiotherapy treatments targeting the respiratory muscle.

Materials and Methods

Ethics Statement

The study was conducted in the Respiratory Function Laboratory at HUGTIP, since February 2011 to December 2013, and approved by the Human Research and Ethics Committee of the hospital. All participants gave written informed consent, following the World Medical Association's Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects.

Study subjects

Patients with unilateral phrenic paralysis [18], who were previously diagnosed in the Department of Internal Medicine at Germans Trias i Pujol University Hospital (HUGTIP), were considered eligible for this study. All patients underwent chest radiography and computed tomography scanning of the chest, which reveal elevated hemidiaphragm on the affected side. Moreover, according to their medical history, most of the patients had previous thoracic or surgical trauma as the major cause of diaphragmatic paralysis. Only patient ID 2 had an unknown etiology. However, all patients related some level of functional dyspnea.

On the other hand, controls were selected from healthy subjects who had never been diagnosed of phrenic paralysis and had normal baseline spirometric values. According to these inclusion criteria, ten men with unilateral phrenic paralysis in a stable condition and ten controls (five men/five women) were included in the study for pulmonary function test and the acoustic respiratory analysis.

Pulmonary function and lung sound testing

At baseline, lung function was measured by spirometry (Hyp'Air Compact, Medisoft). Table 1 shows baseline spirometric results from each subject. Measurements were obtained in accordance with established guidelines [19], and results compared to reference values [20].

After this previous test, each subject was coached to progressively increase the airflow from shallow breathing to the deepest breaths they were able to, reaching 1.2 to 2.4 L/s [14]. Lastly, at the end of the respiratory test, subjects were asked to hold their breath for a few seconds in order to estimate background noise intensity (BNI). One recording of a total 120 seconds was obtained from each subject in a sitting position. Respiratory flow and sounds were acquired simultaneously during the test.

Lung sounds and respiratory airflow measurements

Respiratory sounds were recorded using two contact microphones (TSD108, Biopac Systems, Inc.) with a frequency response of 35–3500 Hz. Microphones were positioned on the surface of the back, at each side of the spinal cord and 3 cm below the bottom tip of the shoulder blades. They were attached to the skin with double-sided adhesive discs, in a noninvasive way. In addition, respiratory airflow was recorded with a pneumotachograph (TSD107B, Biopac Systems, Inc.). Subjects wore a nose clip and breathed through the mouthpiece of the instrument.

Airflow and sound signals were amplified and filtered by hardware, before analog-to-digital conversion and acquisition. On the one hand, high- and low-pass filters with cut-off frequencies of 10 and 5000 Hz, respectively, were applied to respiratory sound signals, and they were amplified by a factor of 200. On the other hand, low-pass filter with a cut-off frequency of 10 Hz was applied to the airflow signal, and this was amplified by a factor of 1000. Then, both sound and flow signals were recorded at a sample rate of 12500 Hz using a 16-bit analog-to-digital converter (MP150, Biopac Systems, Inc.). Since this study is only focused on normal pulmonary sounds, whose bandwidth of interest is below 2000 Hz, respiratory sound signals were digitally filtered using a combination of 8th order Butterworth high- and low-pass filters with cut-off frequencies of 70 and 2000 Hz, respectively.

Lung sound analysis

Respiratory sound signals were automatically segmented by extracting respiratory phases from the airflow signal. Respiratory cycles in which the flow reached at least 0.35 L/s were considered valid cycles. In order to avoid false detections caused by background noise, two thresholds of 0.2 and 4 seconds were established for minimum and maximum durations of breathing phases, respectively, according to time duration of normal respiratory cycles. In addition, a threshold of 0.5 seconds was fixed for the maximum time interval between the end of inspiration and the beginning of the corresponding expiration. All cycles not meeting these criteria were rejected. The final dataset for each subject was formed by audio-visual selection of pairs of sound signals, one from each side, from the same inspiratory cycles, avoiding artifacts such as those from swallowing or rubbing.

Each inspiratory sound cycle was firstly classified according to the maximum airflow reached. For that purpose, the airflow scale was divided into intervals of 0.2 L/s, from 1.2 L/s upwards. Furthermore, only inspiratory sound segments corresponding to the top airflow interval, whose duration is at least 20% of cycle length, of each inspiratory cycle were used for assessing the LSI.

The LSI was calculated as the mean power, in the frequency band from 70 to 2000 Hz, obtained from the power spectral density (PSD) of each inspiratory sound segment, according to the following expression:

$$LSI = \frac{f_m}{NFFT} \sum_{f=70}^{2000} PSD(f)$$

where f_m is the sample rate, and $NFFT$ is the number of points for the fast Fourier transform (FFT). Just as in some previous studies, which were focused on the intensity of respiratory sounds [14,21,22] the PSD was calculated using Welch's periodogram, with a Hanning window of 1000 data samples (80 ms), a 50% overlap between adjacent segments, and 1024 points for the FFT. The same method was applied to apnea segments from both left and right sides, in order to calculate the mean background noise intensity ($\overline{\mathrm{BNI}}$). The resultant LSI values from all inspiratory sound segments were expressed in dB with respect to this $\overline{\mathrm{BNI}}$. Having calculated the LSI, each subject was characterized by the relationship between the LSI and the airflow on both left and right sides. In addition, the LSI was averaged over the airflow range 1.2–2.4 L/s, in order to obtain a mean LSI for each side. Normality in the mean LSIs of both sides, from patients and healthy subjects, as well as in their differences was tested with a Lilliefors test. Since we did not know the parameters of the hypothesized distributions, and those parameters must be estimated from the data sample, the Lilliefors test was preferable. On the other hand, the statistical differences were tested between: 1) the mean LSIs of both sides, and 2) the differences in the mean LSIs of both sides from patients and healthy subjects. Since normality could not be assumed in all cases, and the sample size ($n = 20$) was small, a non-parametric test, such as the Wilcoxon rank sum test, was used to check for statistical differences.

Results

Lung sound intensity in unilateral phrenic paralysis

Acoustic and spirometric parameters were analyzed in patients and healthy subjects. As shown in Table 1, eight patients had left side paralysis (ID 2–7, 9, and 10) and two patients had right side paralysis (ID 1 and 8). Regardless of the side affected, all patients had lower FVC ($57\pm9\%$) and FEV_1 ($57\pm10\%$) values than healthy subjects, in whom the percentages were $94\pm11\%$ and $93\pm7\%$, respectively.

Table 1. Subject Characteristics*.

Phrenic Nerve Paralysis Patients

ID	Paralyzed Side	Age (yr)	Sex	BMI (kg/m^2)	FVC L (% of predicted)	FEV_1 L (% of predicted)	FEV_1/FVC (%)
1	Right	46	Male	32.7	2.86 (56)	2.11 (54)	74.0
2	Left	52	Male	27.0	3.80 (70)	2.75 (66)	72.0
3	Left	38	Male	29.4	2.81 (55)	2.10 (51)	74.7
4	Left	70	Male	30.5	2.11 (56)	1.32 (49)	63.0
5	Left	41	Male	33.0	2.59 (57)	2.07 (57)	79.9
6	Left	74	Male	35.0	1.56 (40)	1.21 (42)	77.2
7	Left	54	Male	28.6	2.13 (50)	1.76 (53)	82.5
8	Right	70	Male	28.4	2.44 (53)	1.95 (57)	79.6
9	Left	59	Male	24.2	3.17 (69)	2.79 (80)	87.9
10	Left	69	Male	30.9	2.44 (64)	1.68 (62)	69.0
Mean		57		30.0	(57)	(57)	76.0
SD		13		3.2	(9)	(10)	7.1

Healthy Subjects

ID	Paralyzed Side	Age (yr)	Sex	BMI (kg/m^2)	FVC L (% of predicted)	FEV_1 L (% of predicted)	FEV_1/FVC (%)
1	None	28	Male	25.3	5.56 (97)	4.21 (90)	75.7
2	None	31	Male	22.3	4.50 (82)	3.85 (89)	86.0
3	None	48	Female	22.4	2.56 (76)	2.38 (92)	93.0
4	None	36	Male	24.4	5.38 (114)	4.05 (105)	75.3
5	None	24	Female	22.4	4.38 (92)	3.66 (94)	83.4
6	None	35	Female	23.0	4.45 (105)	3.63 (106)	81.5
7	None	31	Female	23.7	3.23 (91)	2.80 (95)	86.9
8	None	49	Female	26.5	3.25 (89)	2.66 (92)	82.0
9	None	22	Male	24.5	5.42 (91)	4.03 (82)	74.4
10	None	44	Male	24.4	5.43 (99)	3.84 (89)	70.8
Mean		35		23.9	(94)	(93)	80.9
SD		10		1.4	(11)	(7)	6.8

* BMI = body mass index; FVC = forced vital capacity; FEV_1 = forced expiratory volume in 1 second. Anthropometric and spirometric values of ten patients with a unilateral phrenic paralysis and ten healthy subjects are shown. References for percentages of predicted normal values are from Roca et al. [20]. FEV_1 and FVC in percentages were lower in all patients than in healthy subjects.

doi:10.1371/journal.pone.0093595.t001

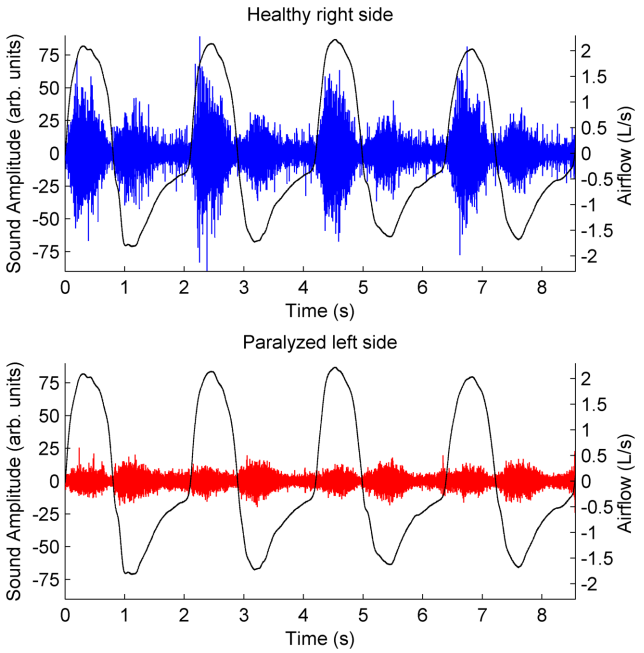


Figure 1. Airflow and lung sound signals. Airflow signal (black) and the corresponding lung sound signals, in arbitrary units, for both right (blue) and left (red) sides, in a patient with left side phrenic paralysis (ID 4). Sound amplitudes from the left side were lower than those from the healthy right side.
doi:10.1371/journal.pone.0093595.g001

With regard to lung sounds, the signal amplitude was much lower on the paralyzed side than the healthy side, in patients with unilateral phrenic nerve paralysis, as shown in the example from Figure 1. It contains the lung sound and the airflow signals from a patient with left phrenic paralysis (ID 4). Accordingly, the magnitude of the PSDs from both sides, and the consequent signal powers, are quite different, as shown in Figure 2. It exhibits the PSDs from two inspiratory sound segments, one from each side, of an inspiratory cycle from patient ID 4. As shown, the PSD of the right sound segment (healthy side) is a long way from the PSD of the right background noise segment, in all the frequency range. On the other hand, the PSD of the left sound segment (affected side) is slightly above the left background noise. As a result, the LSI calculated from the PSD of the healthy side is much larger than the affected side.

The aforementioned pattern was confirmed by comparison of the acoustic parameters in all patients with unilateral phrenic paralysis (Table 2). Calculation of the $\overline{\text{BNI}}$ from the BNI for left and right sides allowed us to express the mean LSI from each side in dB with respect to the same reference value. In addition, the mean LSI was calculated for both sides from the same set of cycles.

In healthy subjects, the mean LSI was much higher than the $\overline{\text{BNI}}$ on both left (15.1 ± 6.2 dB) and right (17.4 ± 5.7 dB) sides. However, patients had mean LSIs only a few dBs above the $\overline{\text{BNI}}$ on the affected side (5.7 ± 3.5 dB) while their mean LSIs on the healthy side (19.3 ± 4.0 dB) were not significantly different from the values measured in the healthy participants. To show this trend clearly, we calculated the difference between the mean LSI of each side.

Figure 3 shows the mean LSI, for each side, as a function of airflow level in all patients, and all healthy subjects. On the one hand, considerable differences, of more than 13 dB, can be seen between the LSI from the affected and healthy sides. On the other hand, differences in LSI between the sides are less than 3 dBs in healthy subjects. It should be noted that the LSI from the affected sides are close to the $\overline{\text{BNI}}$ (0 dB) at low airflows. As the airflow increases, the LSI from the affected sides does also increase, but never reaches the levels seen in healthy subjects.

Furthermore, Figure 3 shows a clear linear relationship between the LSI and the airflow level. This sound-flow relationship has been reported in some previous studies [23–25], and it usually follows a power law. In a logarithmic scale (dB), this relationship can be formulated by a linear equation:

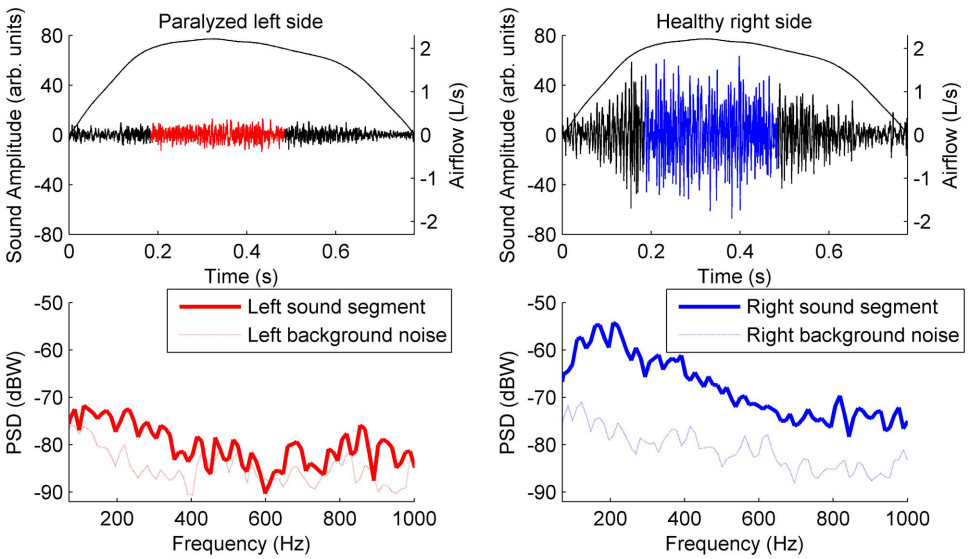


Figure 2. Power spectral density of lung sounds from a patient. Airflow (L/s), lung sound signals (arbitrary units), and the corresponding power spectral densities (dBW), for both sides of an inspiratory cycle from a patient with left side phrenic paralysis (ID 4). Solid and dotted lines in the PSDs correspond to the central sound segments and the background noise segments from both sides, respectively.
doi:10.1371/journal.pone.0093595.g002

Table 2. Acoustic Characteristics*.

Phrenic Nerve Paralysis Patients

ID	$BNI_{Healthy}$ (μW)	$BNI_{Paralyzed}$ (μW)	BNI_{Mean} (μW)	$LSI_{Healthy}$ (dB)*	$LSI_{Paralyzed}$ (dB)*	$LSI_{Healthy}$ - $LSI_{Paralyzed}$ (dB)	# Left and Right Cycles
1	15.5	7.8	11.6	16.6	6.2	10.4	30
2	8.1	7.9	8.0	13.6	3.7	9.9	28
3	2.4	1.8	2.1	21.2	4.8	16.3	30
4	16.0	3.2	9.6	16.5	3.5	12.9	30
5	7.8	5.5	6.7	13.4	2.0	11.4	18
6	15.2	3.6	9.4	21.5	0.2	21.3	18
7	4.7	3.3	4.0	24.2	10.2	13.9	34
8	9.5	6.8	8.1	23.9	7.8	16.1	21
9	28.0	14.2	21.1	21.8	11.0	10.8	18
10	14.4	9.5	11.9	20.1	7.9	12.2	27
Mean	12.2	6.4	9.3	19.3	5.7	13.5	25
SD	7.3	3.7	5.2	4.0	3.5	3.5	6

Healthy Subjects

ID	BNI_{Left} (μW)	BNI_{Right} (μW)	BNI_{Mean} (μW)	LSI_{Left} (dB)*	LSI_{Right} (dB)*	LSI_{Right} – LSI_{Left} (dB)	# Left and Right Cycles
1	1.2	4.7	2.9	12.8	15.8	3.0	40
2	9.0	3.5	6.2	8.4	8.9	0.5	44
3	20.2	26.7	23.4	23.8	24.9	1.1	16
4	0.8	1.9	1.4	13.3	17.9	4.5	30
5	1.7	3.9	2.8	13.5	17.1	3.6	28
6	1.5	25.7	13.6	11.0	13.2	2.2	42
7	6.6	5.2	5.9	26.6	27.0	0.4	26
8	1.1	2.2	1.6	20.5	22.1	1.6	47
9	1.8	4.7	3.3	10.9	12.5	1.6	14
10	1.5	3.0	2.2	10.3	15.0	4.6	22
Mean	4.5	8.2	6.3	15.1	17.4	2.3	31
SD	6.2	9.6	7.0	6.3	5.7	1.6	12

* With respect to mean background noise intensity (BNI_{Mean}); BNI = background noise intensity; LSI: lung sound intensity; #: number of cycles analyzed. Acoustic characteristics are shown for ten patients with a unilateral phrenic paralysis and ten healthy subjects. Absolute differences in mean inspiratory LSI ($|LSI_{Left}-LSI_{Right}|$) were higher in patients than in healthy subjects.

doi:10.1371/journal.pone.0093595.t002

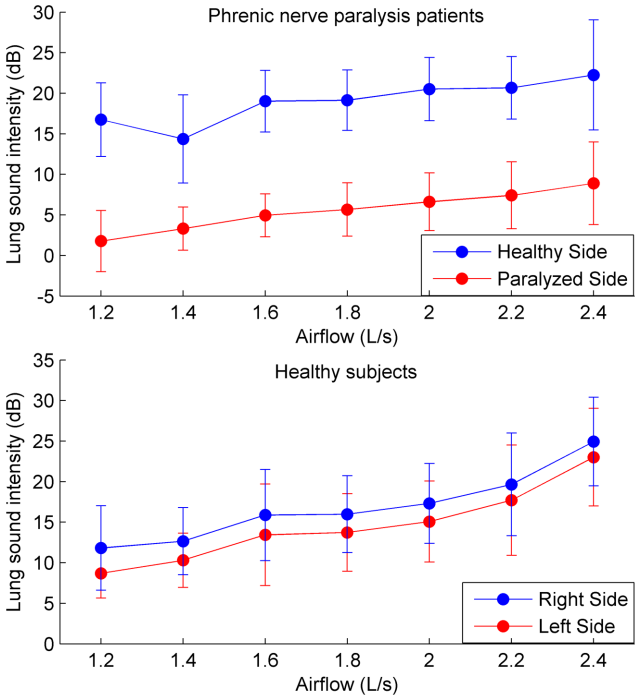


Figure 3. Comparison of inspiratory LSI-Airflow relationship between both hemithoraxes in patients and healthy subjects. Mean inspiratory LSI (dB) as a function of airflow (L/s), from the ten patients and the ten healthy subjects. All values are the mean ± SD.
doi:10.1371/journal.pone.0093595.g003

$$LSI(dB) = m * Airflow + b$$

where m is the slope of the line, and b is the y-intercept. As shown in Table 3, all LSI-airflow relationships from Figure 3 can be properly expressed by a linear equation. Moreover, there is a clear difference between the slope of healthy subjects (around 5 dB/L/s) and patients (around 10 dB/L/s), independently of the analyzed side.

The mean inspiratory LSI from both sides of patients and healthy subjects has been statistically analyzed, as shown in Figure 4. The null hypothesis that the mean LSIs were normally distributed was accepted as much for both healthy and paralyzed sides in patients ($p = 0.4135$ and 0.9436, respectively), as for the right side in healthy subjects ($p = 0.5790$). However, the null hypothesis was rejected for the left side in healthy subjects ($p = 0.0104$).

The Wilcoxon rank sum test showed that the mean LSIs of healthy and paralyzed sides in all patients were statistically different ($p = 0.0002$). On the contrary, the difference between the mean LSIs of right and left sides in all healthy subjects was not statistically significant ($p = 0.2730$).

Table 3. Linear regression parameters*.

	Healthy subjects		Patients	
	Right side	**Left side**	**Paralyzed side**	**Healthy side**
R^2	0.91	0.93	0.98	0.79
Slope (dB/L/s)	9.78	10.61	5.58	5.48

* Corresponding to graphs from Figure 3.
doi:10.1371/journal.pone.0093595.t003

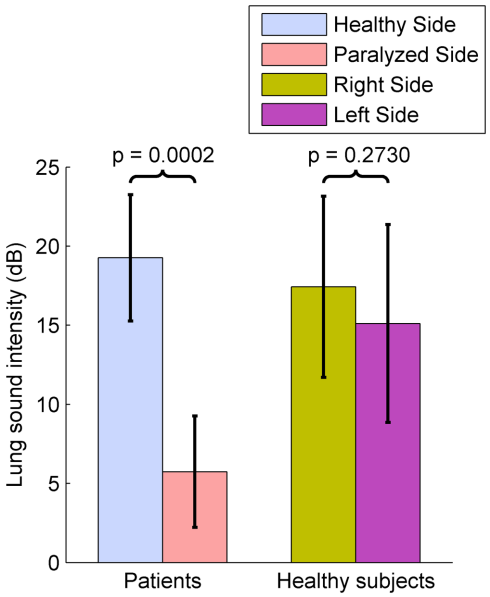


Figure 4. Comparison of mean inspiratory LSI between both hemithoraxes in patients and healthy subjects. Mean inspiratory LSI (dB) in healthy and paralyzed sides (ten patients), and in right and left sides (ten healthy subjects). The mean LSI from patients is significantly higher in healthy side than paralyzed side ($p = 0.0002$). On the contrary, there are not significant differences between mean LSI from both hemithoraxes in healthy subjects ($p = 0.2730$).
doi:10.1371/journal.pone.0093595.g004

Lung sound intensity differences and FEV_1 relationship

Figure 5-A shows the absolute value of the differences between the mean LSIs of both sides ($|LSI_{Left}\text{-}LSI_{right}|$). In this case, the null hypothesis that the differences were normally distributed was accepted as much in patients ($p = 0.6078$), as in healthy subjects ($p = 0.4693$).

The Wilcoxon rank sum test showed that the differences between the mean LSIs of both sides were statistically significant in both groups ($p = 0.0002$). Moreover, it was found that there was a clear cut off around 6–8 dB which distinguished patients from healthy subjects.

Figure 5-B illustrates the relationship between the mean LSI difference and FEV_1, showing high differences in the LSI and low FEV_1 in patients with phrenic nerve paralysis. Moreover, in patients there is an inverse relationship between the two parameters, namely the lower the FEV_1, the higher the mean LSI difference. In contrast, healthy subjects have low mean LSI differences, and there is no any clear relationship between these LSI differences and the corresponding FEV_1.

Discussion and Conclusions

Our study shows that patients with unilateral phrenic nerve paralysis have a lower inspiratory sound intensity on the affected side than the healthy one. We did not analyze expiratory sounds due to the lower values of expiratory intensity with respect to inspiratory values at isoflows [14]. This study illustrates the potential of lung acoustic analysis for the diagnosis and management of these patients.

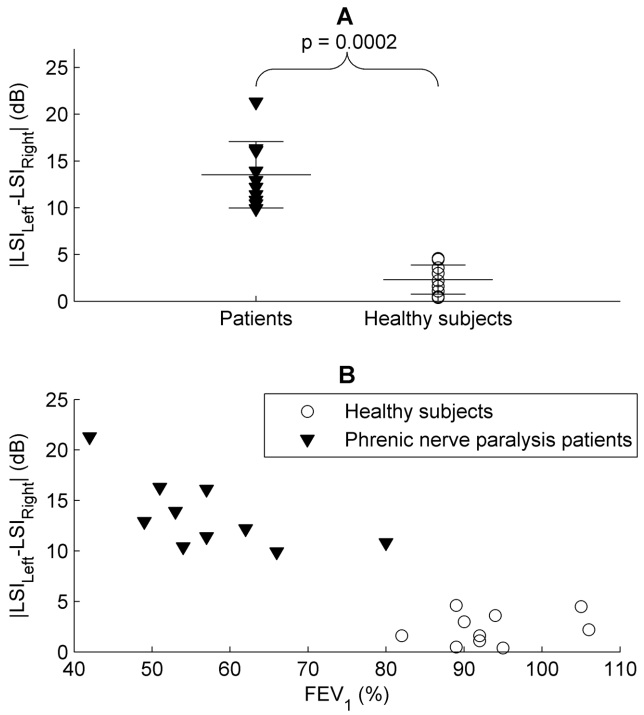


Figure 5. LSI differences and FEV_1 relationship in patients and healthy subjects. A: Mean inspiratory LSI difference (dB) between both hemithoraxes, in ten patients with phrenic nerve paralysis and ten healthy subjects. The LSI difference was higher in patients than healthy subjects ($p = 0.0002$). Solid lines indicate the mean and SD for each group. B: Mean inspiratory LSI difference as a function of FEV_1 in patients and healthy subjects.
doi:10.1371/journal.pone.0093595.g005

Respiratory sounds are an alternative method to measure both pulmonary [16] and diaphragmatic function. Some previous studies reported decreased breath sound on the affected side in patients with unilateral phrenic paralysis [18,26,27], but they were assessed by traditional auscultation. However, there are no references about quantitative analysis of respiratory sounds for the diagnosis of these patients. In what is related to laterality of respiratory sounds, they have been used to distinguish between bilateral and unilateral lung ventilation in intubated patients [28]. Nevertheless, many studies have analyzed the differences between the LSI of both sides in healthy subjects [14,29,30,31], thus reporting slight differences of a few dB. In any case, sound analysis can detect differences in airflow entering the two sides of the thoracic cage in diseases that affect respiratory ventilation, and our study demonstrates this for the case of unilateral phrenic nerve paralysis. Consistently, we found a clear cut-off in the mean differences of LSI between the two sides in healthy subjects and patients.

In addition to unilateral phrenic paralysis, it has been recognized by other authors that lung sound analysis is also a very useful technique to study many others pulmonary diseases [21,22,32,33].

When the diaphragm is paralyzed, it does not have an influence on expansion of the homolateral lung and breathing is maintained by accessory muscles such as those of the chest wall. The movement of the paralyzed hemidiaphragm is determined by the balance between the change in pleural pressure and the shortening of the healthy hemidiaphragm. This is manifested by a cranial displacement of the ipsilateral hemidiaphragm and a small caudal displacement of the contralateral hemidiaphragm [6]. Such a retraction is ineffective for respiration and has been related to patient dyspnea [34].

It has been suggested that the airflow to dependent areas of the lung is directed by the diaphragm and non-dependent areas by the intercostal muscles [35]. The gas flow to the dependent areas of the paralytic side would therefore be lower than that to the healthy side.

In this study, there was considerably less airflow entering the dependent areas of the pathological side, measured in an indirect way by the quantification of the LSI. Specifically, the LSI of the affected side was close to the level of the background noise for a low airflow rate, as seen in Figure 3, while the signal from the healthy side remained a long way from the background noise at all measured flow rates. Although the BNIs from both sides are slightly different, it is not relevant for the results of this study, since the $\overline{BNI}$ is used as the unique reference value in order to express the LSI in dB.

In addition, pulmonary perfusion is redistributed from the base toward the apex in these patients [36]. The result of this pathological situation is that the work of breathing (measured in terms of oxygen uptake) is increased, which suggests that intercostal muscle breathing is less efficient than diaphragm breathing [37]. Spirometric changes have been widely commented on in the literature. In our study, spirometry values of patients with phrenic paralysis were low with respect to normal reference values, as has been found previously in other studies [9,38].

With respect to traditional techniques to diagnose the unilateral phrenic paralysis, they include: x-ray imaging, fluoroscopy, ultrasonography, and phrenic nerve stimulation [18,39]. Usually, unilateral phrenic paralysis is diagnosed by a combination of these techniques, since none of them is totally concluding by itself. Of these techniques, x-ray imaging is the simplest and it has some obvious limitations: it uses ionizing radiation, and it does not allow us to assess the diaphragm or the pulmonary function of patients. Moreover, in unilateral diaphragmatic paralysis, the sensitivity of plain chest radiograph is as high as 90%, whereas its specificity is, however, low (44%) [39].

Fluoroscopy and the external or internal stimuli of the diaphragm allow evaluating the diaphragm mobility [4]. However, fluoroscopy also makes use of x-rays to obtain dynamic images of the diaphragm, and both fluoroscopy and stimuli of the diaphragm are invasive techniques. Moreover, none of these methods provides information about the pulmonary function.

Ultrasonography is an alternative non-invasive technique to assess the diaphragmatic function [5,40], since it works on ultrasounds. Nevertheless, just as the aforementioned techniques, ultrasounds do not provide any data about the pulmonary function of patients. Moreover, ultrasonography is operator dependent and requires significant expertise [39].

Recently, a new non-invasive method has been proposed to measure the movements of the thoracic wall [41]. This new method makes use of a motion analysis system, which is called optoelectronic plethysmography. It was used to estimate the total rib cage volume, as well as its changes in both healthy and paralyzed sides.

In this study, the potential of acoustic respiratory analysis for detecting unilateral phrenic paralysis has been clearly shown. Despite a relatively small population has been analyzed, the results from 20 subjects (10 patients and 10 healthy subjects) reinforce the reliability of the proposed method. On the other hand, in the database, there is a slight difference in the male-female ratio between patients and healthy subjects, but gender is not a relevant factor in the analysis of normal lung sound intensity [42].

However, further studies will be needed to clinically validate this technique as a new complementary tool for phrenic paralysis diagnosis.

In conclusion, measurement of LSI can provide quantitative information about the extent of impairment of respiratory function in patients with unilateral phrenic nerve paralysis. In these patients, LSI is an indirect measure of the airflow that enters the lungs, this being lower on the affected side due to inefficient diaphragmatic muscle function. This technique represent a step forward in the diagnostic procedure of unilateral phrenic nerve paralysis, since it has some advantages with respect to current techniques: non-invasiveness, objectivity, simplicity, easiness and cost. The acoustic respiratory analysis, in conjunction with spirometry, could reinforce the reliability of the diagnosis of unilateral phrenic paralysis.

Regarding the future use of the method, its major application is the non-invasive assessment of respiratory function, providing objective information of the affected side. Therefore, the method offers the capability for long-term monitoring of recovery in respiratory function in patients who undergo physical therapy [43]. These patients are regularly monitored in order to check whether the physical therapy is improving their pulmonary function in the affected side or not. In this context, the advantages of the proposed technique gain relevance since several and repeated tests are required for the long-term monitoring of these patients.

Acknowledgments

All authors would like to thank all the team in the Pulmonary Function Testing Laboratory at Germans Trias i Pujol University Hospital, for their collaboration in the patient recruitment as well as the performing of the spirometry tests.

Author Contributions

Conceived and designed the experiments: JAF RJ ML JR. Performed the experiments: JAF ML RG. Analyzed the data: JAF RJ ML. Contributed reagents/materials/analysis tools: JAF RJ ML. Wrote the paper: JAF RJ ML. Interpreted results of experiments: JAF RJ ML.

References

1. DeVita MA, Robinson LR, Rehder J, Hattler B, Cohen C (1993) Incidence and natural history of phrenic neuropathy occurring during open heart surgery. Chest 103(3): 850–856.
2. Brochard L, Harf A, Lorino H, Lemaire F (1989) Inspiratory pressure support prevents diaphragmatic fatigue during weaning from mechanical ventilation. Am Rev Respir Dis 139(2): 513–521.
3. Freilich S, Janssen JC, Polkey MI (2011) An unusual case of diaphragm paralysis. Thorax 66(2): 133.
4. Yi LC, Nascimento OA, Jardim JR (2011) Reliability of an analysis method for measuring diaphragm excursion by means of direct visualization with videofluoroscopy. Arch Bronconeumol 47(6): 310–314.
5. Kim SH, Na S, Choi JS, Na SH, Shin S, et al. (2010) An evaluation of diaphragmatic movement by M-code sonography as a predictor of pulmonary dysfunction after upper abdominal surgery. Anesth Analq 110(5): 1349–1354.
6. Scillia P, Cappello M, De Troyer A (2004) Determinants of diaphragm motion in unilateral diaphragmatic paralysis. J Appl Physiol 96(1): 96–100.
7. Watson AC, Hughes PD, Louise Harris M, Hart N, Ware RJ, et al. (2001) Measurement of twitch transdiaphragmatic, esophageal, and endotracheal tube pressure with bilateral anterolateral magnetic phrenic nerve stimulation in patients in the intensive care unit. Crit Care Med 29(7): 1325–1331.
8. Chetta A, Rehman AK, Moxham J, Carr DH, Polkey MI (2005) Chest radiography cannot predict diaphragm function. Respir Med 99(1): 39–44.
9. Lisboa C, Paré PD, Pertuzé J, Contreras G, Moreno R, et al. (1986) Inspiratory muscle function in unilateral diaphragmatic paralysis. Am Rev Respir Dis 134(3): 488–492.
10. Epstein SW, Vanderlinden RG, Man SF, Hyland RH, Lenkei SC, et al. (1979) Lung function in diaphragm pacing. Can Med Assoc J 120(11): 1360–1368.
11. Sokolowska B, Józwik A, Pokorski M (2003) A fuzzy-classifier system to distinguish respiratory patterns evolving after diaphragm paralysis in the cat. Jpn J Physiol 53(4): 301–307.
12. Pasterkamp H, Patel S, Wodicka GR (1997) Asymmetry of respiratory sounds and thoracic transmission. Med Biol Eng Comput 35(2): 103–106.
13. Hult P, Wranne B, Ask P (2000) A bioacoustic method for timing of the different phases of the breathing cycle and monitoring of breathing frequency. Med Eng Phys 22(6): 425–433.
14. Fiz JA, Gnitecki J, Kraman SS, Wodicka GR, Pasterkamp H (2008) Effect of body position on lung sounds in healthy young men. Chest 133(3): 729–736.
15. Alshaer H, Fernie GR, Bradley TD (2011) Monitoring of breathing phases using a bioacoustic method in healthy awake subjects. J Clin Monit Comput 25(5): 285–294.
16. Morice RC, Jimenez CA, Eapen GA, Mehran RJ, Keus L, et al. (2010) Using quantitative breath sound measurements to predict lung function following resection. J Cardiothorac Surg 5: 81.
17. Shykoff BE, Ploysongsang Y, Chang HK (1988) Airflow and normal lung sounds. Am Rev Respir Dis 137(4): 872–876.
18. Gibson GJ (1989) Diaphragmatic paresis: pathophysiology, clinical features, and investigation. Thorax 44(11): 960–970.
19. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, et al. (2005) Standardisation of spirometry. Eur Respir J 26(2): 319–338.
20. Roca J, Burgos F, Sunyer J, Saez M, Chinn S, et al. (1998) References values for forced spirometry. Group of the European Community Respiratory Health Survey. Eur Respir J 11(6): 1354–1362.
21. Oud M, Dooijes EH, Van Der Zee JS (2000) Asthmatic airways obstruction assessment based on detailed analysis of respiratory sound spectra. IEEE Trans Biomed Eng 47(11): 1450–5.
22. Montazeri A, Giannouli E, Moussavi Z (2012) Assessment of obstructive sleep apnea and its severity during wakefulness. Ann Biomed Eng 40(4): 916–24.
23. Hossain I, Moussavi Z (2004) Finding the lung sound-flow relationship in normal and asthmatic subjects. Conf Proc IEEE Eng Med Biol Soc 5: 3852–5.
24. Golabbakhsh M, Moussavi Z (2004) Relationship between airflow and frequency-based features of tracheal respiratory sound. Can Conf Electr Comput Eng 2: 751–754.
25. Yadollahi A, Montazeri A, Azarbarzin A, Moussavi Z (2013) Respiratory flow-sound relationship during both wakefulness and sleep and its variation in relation to sleep apnea. Ann Biomed Eng 41(3): 537–46.
26. Dernaika TA, Younis WG, Carlile PV (2008) Spontaneous recovery in idiopathic unilateral diaphragmatic paralysis. Respir Care 53(3): 351–4.
27. Habib GS (2012) Unilateral diaphragm paralysis following vaccination. J Med Cases 3(3): 164–166.
28. Tejman-Yarden S, Lederman D, Eilig I, Zlotnik A, Weksler N, et al. (2006) Acoustic monitoring of double-lumen ventilated lungs for the detection of selective unilateral lung ventilation. Anesth Analg 103(6): 1489–1493.
29. Kompis M, Pasterkamp H, Oh Y, Wodicka GR (1997) Distribution of inspiratory and expiratory respiratory sound intensity on the surface of the human thorax. Conf Proc IEEE Eng Med Biol Soc 5: 2047–2050.
30. Torres-Jiménez A, Charleston-Villalobos S, González-Camarena R, Chi-Lem G, Aljama-Corrales T (2008) Asymmetry in lung sound intensities detected by respiratory acoustic thoracic imaging (RATHI) and clinical pulmonary auscultation. Conf Proc IEEE Eng Med Biol Soc 2008: 4797–4800.
31. Dellinger RP, Parrillo JE, Kushnir A, Rossi M, Kushnir I (2008) Dynamic visualization of lung sounds with a vibration response device: a case series. Respiration 75(1): 60–72.
32. Pasterkamp H, Kraman SS, Wodicka GR (1997) Respiratory sounds. Advances beyond the stethoscope. Am J Respir Crit Care Med 156(3 Pt 1): 974–987.
33. Gavriely N, Nissan M, Cugell DW, Rubin AH (1994) Respiratory health screening using pulmonary function tests and lung sounds analysis. Eur Respir J 7(1): 35–42.
34. Gharagozloo F, McReynolds SD, Snyder L (1995) Thoracoscopic plication of the diaphragm. Surg Endosc 9(11): 1204–1206.
35. Roussos CS, Fixley M, Genest J, Cosio M, Kelly S, et al. (1977) Voluntary factors influencing the distribution of inspired gas. Am Rev Respir Dis 116(3): 457–467.
36. Amis TC, Ciofetta G, Hughes JM, Loh L (1980) Regional lung function in bilateral diaphragmatic paralysis. Clin Sci (Lond) 59(6): 485–492.
37. DiMarco AF, Connors AF Jr, Kowalski KE (2004) Gas exchange during separate diaphragm and intercostal muscle breathing. J Appl Physiol 96(6): 2120–2124.
38. Xu WD, Gu YD, Lu JB, Yu C, Zhang CG, et al. (2005) Pulmonary function after complete unilateral phrenic nerve transection. J Neurosurg 103(3): 464–467.
39. McCool FD, Tzelepis GE (2012) Dysfunction of the diaphragm. N Engl J Med 366(10): 932–42.
40. Balaji S, Kunovsky P, Sullivan I (1990) Ultrasound in the diagnosis of diaphragmatic paralysis after operation for congenital heart disease. Br Heart J 64(1): 20–2.

41. Boudarham J, Pradon D, Prigent H, Falaize L, Durand MC, et al. (2013) Optoelectronic plethysmography as an alternative method for the diagnosis of unilateral diaphragmatic weakness. Chest 144(3): 887–95.
42. Gross V, Dittmar A, Penzel T, Schüttler F, von Wichert P (2000) The relationship between normal lung sounds, age, and gender. Am J Respir Crit Care Med 162(3): 905–909.
43. Gayan-Ramirez G, Gosselin N, Troosters T, Bruyninckx F, Gosselink R, et al. (2008) Functional recovery of diaphragm paralysis: a long-term follow-up study. Respir Med 102(5): 690–698.

Chapter 3: Automatic differentiation of normal and continuous adventitious respiratory sounds using ensemble empirical mode decomposition and instantaneous frequency

Title: Automatic differentiation of normal and continuous adventitious respiratory sounds using ensemble empirical mode decomposition and instantaneous frequency

Authors: M. Lozano, J. A. Fiz, R. Jané

Journal: IEEE Journal of Biomedical and Health Informatics, 2016

Automatic Differentiation of Normal and Continuous Adventitious Respiratory Sounds Using Ensemble Empirical Mode Decomposition and Instantaneous Frequency

Manuel Lozano*, José Antonio Fiz, and Raimon Jané, *Senior Member, IEEE*

***Abstract*—Differentiating normal from adventitious respiratory sounds (RS) is a major challenge in the diagnosis of pulmonary diseases. Particularly, continuous adventitious sounds (CAS) are of clinical interest because they reflect the severity of certain diseases. This study presents a new classifier that automatically distinguishes normal sounds from CAS. It is based on the multi-scale analysis of instantaneous frequency (IF) and envelope (IE) calculated after ensemble empirical mode decomposition (EEMD). These techniques have two major advantages over previous techniques: high temporal resolution is achieved by calculating IF-IE and a priori knowledge of signal characteristics is not required for EEMD. The classifier is based on the fact that the IF dispersion of RS signals markedly decreases when CAS appear in respiratory cycles. Therefore, CAS were detected by using a moving window to calculate the dispersion of IF sequences. The study dataset contained 1494 RS segments extracted from 870 inspiratory cycles recorded from 30 patients with asthma. All cycles and their RS segments were previously classified as containing normal sounds or CAS by a highly experienced physician to obtain a gold standard classification. A support vector machine classifier was trained and tested using an iterative procedure in which the dataset was randomly divided into training (65%) and testing (35%) sets inside a loop. The SVM classifier was also tested on 4592 simulated CAS cycles. High total accuracy was obtained with both recorded (94.6% ± 0.3%) and simulated (92.8% ± 3.6%) signals. We conclude that the proposed method is promising for RS analysis and classification.**

***Index Terms*—Asthma, continuous adventitious sounds (CAS), ensemble empirical mode decomposition (EEMD), instantaneous frequency (IF), respiratory sounds (RS), wheezes.**

I. Introduction

PULMONARY auscultation is an essential technique even in cursory physical examinations. Respiratory sounds (RS) contain relevant information about the structure and function of the respiratory system. Therefore, the skills of understanding and interpreting RS are crucial. However, physician subjectivity is a substantial problem in this field and it may lead to an incorrect diagnosis.

RS are generally classified as normal or abnormal [1], [2]. Normal RS originate in healthy lungs, as a consequence of the airflow through the airways during normal breathing. On the other hand, abnormal RS appear in certain pulmonary disorders and they are further classified into discontinuous adventitious sounds and continuous adventitious sounds (CAS). Discontinuous adventitious sounds are explosive and transient sounds that are short in duration but contain a wide range of frequencies. In contrast, CAS typically last more than 100 ms and are characterized by quasi-periodic waveforms with a fundamental frequency of over 100 Hz [1], [2].

Identifying abnormal sounds is a key step in the computerized analysis of RS [3]. From a clinical point of view, CAS are important because they are present in various respiratory diseases involving some degree of airway obstruction, such as asthma or COPD [1]. These diseases produce changes in the mechanical properties of the airways, and these changes explain the presence of CAS [4]. Further, there is a relationship between the number of CAS and the severity of airway obstruction in asthmatic patients, severe airway obstruction being related to more CAS [5], [6]. Moreover, the characteristics of CAS, such as mean frequency, duration, intensity, and respiratory phase in which they appear, are also related to the severity of airway obstruction [7], [8]. Therefore, detecting and analyzing CAS may help to assess the severity of obstructive pulmonary diseases, such as asthma, in which measuring the extent of obstruction and its reversibility is important for reaching a correct diagnosis [9]. However, estimating the number of CAS and their characteristics is quite difficult by audition alone. Consequently, there is need for an efficient algorithm to identify and distinguish CAS from

This work was supported in part by the Spanish Ministry of Economy and Competitiveness under grant TEC2010-21703-C03-01.

M. Lozano is with the Innovation Group at Health Sciences Research Institute of the Germans Trias i Pujol Foundation (IGTP), Ctra. de Can Ruti, Camí de les escoles, s/n, 08916, Badalona, Spain. He is also with the Institute for Bioengineering of Catalonia (IBEC) (e-mail: mlozano@igtp.cat, mlozano@ibecbarcelona.eu).

J.A. Fiz is with the Pulmonology Service at Germans Trias i Pujol University Hospital, Ctra. de Canyet, s/n, 08916, Badalona, Spain. He is also with the IBEC and the Biomedical Research Networking Center in Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN) (e-mail: jafiz@msn.com).

R. Jané is with the Department of Automatic Control (ESAII) at the Universitat Politècnica de Catalunya (UPC). He is also with the IBEC and the CIBER-BBN, Baldiri Reixac, 4, Tower I, 9th floor, 08028, Barcelona, Spain (e-mail: rjane@ibecbarcelona.eu).

normal RS in a straightforward and reliable way.

Several previous studies have proposed a range of RS classification schemes focused exclusively on distinguishing normal from abnormal RS [10]-[13]. Those studies followed the standard methodology of pattern recognition, which comprises feature extraction, dimensionality reduction, and pattern classification. Features were extracted using either wavelet coefficients [10], morphological complexities (lacunarity, sample entropy, kurtosis, and skewness) [11], multi-scale principal component analysis in Fourier domain [12], or hidden Markov models [13]. The sounds analyzed included many types of adventitious RS, not only CAS. Accordingly, the techniques used did not take into account certain features that are relevant to the accurate identification of CAS, such as temporal features. Further, some other studies followed similar methodologies even though they were focused specifically on wheeze and non-wheeze epoch classification within RS signals. For example, such studies have been based on the combination of Mel frequency cepstral coefficients and Gaussian mixture models [14], [15], or linear analysis using the Fisher discriminant method and the Neyman-Pearson test after extraction of features [16].

In contrast to the aforementioned techniques, the analysis of time-frequency distributions is the most common and straightforward method for CAS identification. The combined analysis of time, frequency, and energy features provides extra information about duration, pitch variation, and magnitude, strengthening the identification of CAS. In this context, many studies have been focused on the identification of wheezing-episodes from spectrograms [17]-[23], by applying temporal and spectral continuity criteria to previously detected peaks. Alternatively, digital image processing techniques have also been applied to the spectrogram for wheeze detection [24], [25], and the scalogram, calculated using the wavelet transform, has been used to classify RS [26]. Recently, more advanced TFDs have been proposed for the detection and analysis of CAS, involving either the combination of the wavelet decomposition with third-order spectral features (instantaneous bispectrum/bicoherence) [27], or the derivation of a temporal-spectral dominance spectrogram from the short-time Fourier transform [28].

Despite the widespread use of Fourier and wavelet-based techniques, they have some limitations. First, they are non-adaptive techniques, which are limited by the fact that a priori knowledge of the signal characteristics is required for the correct choice of fixed analysis parameters. Second, as is well known, the Fourier transform, like any integration-based transform, is governed by the uncertainty principle, which limits the combined time-frequency resolution and, in turn, the accuracy of TFDs.

Besides the technical aspects, many previous studies used RS signals recorded either at constant airflows or during forced expiratory maneuvers. However, CAS may appear at different airflow levels which, in turn, affect the properties of CAS.

Empirical mode decomposition (EMD) was proposed as an alternative and suitable tool for the analysis of nonlinear and non-stationary signals, such as RS [29]. Since EMD is an adaptive and direct decomposition technique [29], [30], a priori knowledge of the signal characteristics is not required for the choice of analysis parameters. Moreover, EMD allows a signal to be decomposed into a set of components for which the instantaneous frequency (IF) and envelope (IE) can be defined at any point. Therefore, high temporal and spectral resolutions are achieved [31], [32].

In order to take advantage of the EMD-IF properties, we had previously developed a new method for RS analysis based on customized IF and IE estimation by EMD, adapted to RS signal characteristics [33]. In that previous study, we showed that the IF distribution changed when CAS appeared in a respiratory cycle [33]. Nevertheless, we found that the EMD had a mode mixing effect when applied to RS from some inspiratory cycles, which resulted in poor separation of frequency scales.

Building on the results of our previous work, in the present study, we propose a new method to automatically identify CAS and distinguish them from normal RS. The proposed algorithm makes use of ensemble empirical mode decomposition (EEMD) for IF estimation, which improves on the performance of EMD [34]. As we reported in [33], the IF dispersion markedly decreases when CAS appear in respiratory cycles. This property concerns not only wheezes, but any CAS or musical sounds [1] that are characterized by a dominant fundamental frequency. Accordingly, the present study uses this property to detect CAS at a wide range of frequencies. For that purpose, a set of thresholds was applied to the IF dispersion of RS from each inspiratory cycle, delimiting RS segments with a lower IF dispersion. Then, a specific set of features was extracted from the IF and IE sequences to characterize each delimited RS segment within each inspiratory cycle. These features were used to train and test our RS segment classifier, which was a support vector machine (SVM) classifier. Finally, the whole set of inspiratory cycles were classified as containing CAS or only normal RS, depending on whether they contained at least one CAS segment or not. The proposed technique was tested on recorded RS from asthmatic patients, who performed a variable airflow respiratory maneuver to provide CAS at a wide range of airflow levels. This technique was also tested on simulated CAS to analyze the effect of SNR on the performance of the proposed classifier.

II. Materials And Methods

A. Data Acquisition

Recorded RS were obtained at the Pulmonary Function Testing Laboratory, Germans Trias i Pujol University Hospital, Badalona, Spain. All recordings were acquired from patients who had previously been diagnosed with asthma in the Pulmonology Service at this hospital. First, lung function was measured by spirometry (Hyp'Air Compact, Medisoft). Then, four piezoelectric contact microphones (TSD108, Biopac Systems, Inc.) were placed on the surface of the patient's back, on each side of the spinal cord, at the base and near the upper lobe of the right/left lung. An additional microphone was placed on the surface of the neck, over the right side of the

TABLE I
ANTHROPOMETRIC AND SPIROMETRIC CHARACTERISTICS

Feature/Set	Total
Number of subjects	30
Age (yr)	45 ± 14
BMI (kg/m^2)	25 ± 4
FVC (% of predicted)	87 ± 15
FEV_1 (% of predicted)	80 ± 17

BMI = body mass index; FVC = forced vital capacity; FEV_1 = forced expiratory volume in 1 second.

trachea, at the level of the cricoid cartilage. All sensors were attached to the skin using adhesive rings. The airflow signals were recorded simultaneously with the sound signals, using a pneumotachograph (with an integrated differential pressure transducer, TSD107B, Biopac Systems, Inc.). All signals were sampled at 12 500 samples/s, using a 16-bit analogue-to-digital converter (MP150, Biopac Systems, Inc.). After digitalization, the sound signals were band-pass filtered (70-2 000 Hz) and the respiratory phases were automatically detected using the airflow signal as the reference signal.

Patients were asked to perform a respiratory maneuver consisting of progressively varying their airflow levels, ranging from shallow breathing to the deepest breaths they were able to take. We used this maneuver because CAS do not only appear at high airflows, but also at moderate and low airflows. Further, at the beginning and at the end of the respiratory maneuver, patients were asked to hold their breath for a few seconds. Corresponding signal segments were used to evaluate background noise and heart sound interference.

After cycle segmentation, the minimum and the maximum peak flow values of the inspiratory phases delimited the airflow range of each patient. In order to allow comparisons of signals from these different airflow ranges, each one was divided into four proportional intervals or quartiles, from Q1 to Q4, grouping the peak flow values of the corresponding patient in ascending order, from lowest to highest, respectively.

B. Study Dataset

For this study, RS signals were recorded from 30 asthmatic subjects, whose main anthropometric and spirometric characteristics are summarized in Table I. After cycle segmentation, the study dataset was formed by RS from the inspiratory phase measured with the four microphones on the back, where auscultation is typically performed. We manually selected the inspiratory cycles from the dataset by audiovisual inspection, rejecting those inspiratory cycles containing artifacts (clicks, bursts, or scratches), surrounding noises (background talking or swallowing), or strong heart sound interference (this signal normally being much weaker than RS on the surface of the back). In the end, a total of 870 inspiratory cycles from different airflow quartiles formed the study dataset, including 385 with normal sounds and 485 with CAS (Table II).

A total of 1494 RS segments were extracted from the inspiratory cycles in this dataset. All inspiratory cycles and their corresponding RS segments were manually classified by a highly experienced physician who, by audiovisual inspection, differentiated those with audible CAS from those with only normal sounds. In this way, we obtained the target classification labels for both the inspiratory cycles and the RS segments. Besides using his personal experience and knowledge in RS analysis, the physician followed the criteria explained in a recent publication about RS [1], and the CORSA guidelines of the European Respiratory Society [2], [3]. According to these guidelines, CAS are defined as musical sounds, heard on inspiration, expiration, or both, with a duration of over 100 ms, and a fundamental frequency above 100 Hz. This definition of CAS includes both wheezes and ronchi. The main difference between them is that rhonchi have lower pitch than wheezes. Based on this definition, different types of CAS were included in the dataset, including some polyphonic CAS. In any case, the goal of this study is not to distinguish between different types of CAS, but to automatically differentiate CAS from normal RS.

TABLE II
DISTRIBUTION OF THE RECORDED RS

	Inspiratory cycles					RS segments
	Total	Q1	Q2	Q3	Q4	Total
Total	870	163	213	254	240	1494
Normal	385	105	94	109	77	861
CAS	485	58	119	145	163	633

C. Ensemble Empirical Mode Decomposition of RS

The proposed algorithm for RS classification is based on the fact that RS segments containing CAS have lower IF dispersion than RS segments containing normal sounds in a respiratory cycle, as we reported in [33]. Consequently, the first step of our RS classification scheme is the estimation of the IF of the RS for each inspiratory cycle in the dataset. However, estimating the IF only makes sense for mono-component or, at least, narrowband signals. As RS are multi-component signals, we need to decompose them before being able to calculate a physically meaningful IF. In this context, we previously described in [33] that the EMD allows RS signals to be decomposed into a set of components, called intrinsic mode functions (IMFs), for which the IF can be defined at any point. While, ideally, each IMF would contain different frequency components of the RS signal, we found that EMD had a mode mixing effect when applied to some RS signals containing CAS. Due to this mode mixing effect, some frequency components of CAS appeared in different IMFs, leading to poor separation of frequency scales.

Recently, EEMD was proposed as an improvement to EMD to overcome this mode mixing effect [34]. EEMD consists of the iterative application of the original EMD to a signal plus white noise. The idea behind EEMD is that EMD acts as an adaptive dyadic filter bank when applied to white noise, thereby separating the different frequency scales of a signal naturally without any a priori subjective criteria [34]. Moreover, another previous study showed that the computational cost of the EEMD was reduced by replacing white noise by band-limited noise [35], just covering the bandwidth of the signal to be decomposed.

In the present study, we developed a customized EEMD algorithm adapted to RS signals, for use prior to IF calculations. The basic steps involved in our customized EEMD algorithm are described below:

1) Generate a low-pass filtered white noise signal. Since the fundamental frequency of CAS is typically below 1000 Hz [1], the cut-off frequency of the low-pass filter was empirically set to 1150 Hz.
2) Add this band-limited noise signal to the RS signal from an inspiratory cycle at an SNR of 0 dB.
3) Decompose the resulting noisy signal into noisy IMFs 1-4 using the original EMD algorithm reported by Rilling and Flandrin [36], [37].
4) Repeat steps one to three 100 times.
5) Calculate the final IMFs as the mean of those resulting from the iteration process.

Fig. 1 illustrates the EEMD method applied to RS from an inspiratory cycle containing a wheeze. IMFs 1-4 were obtained in decreasing order of frequency, with IMF 1 including the highest frequency components in the RS. In this case, the main components of the CAS were within IMF 2. We found that EEMD provided an efficient way to separate the different frequency scales for computerized RS analysis. Moreover, since the frequency range of interest for CAS detection is from 100 Hz upwards, IMFs 1-4 proved sufficient to cover this range.

D. Instantaneous Frequency of RS

As defined by Gabor [38], the IF can be calculated from the phase of a complex signal by taking the time derivative. The most practical way to obtain a unique complex signal from a real one is to construct the analytic signal.

In this study, we calculated an analytic signal, $z_i(t)$, from each IMF, $s_i(t)$, as in (1).

$$z_i(t) = s_i(t) + jH[s_i(t)] = a_i(t)e^{j\Phi_i(t)} \quad (1)$$

where $H[\cdot]$ is the Hilbert transform, $a_i(t)$ is the absolute value of $z_i(t)$, and $\Phi_i(t)$ is the phase of $z_i(t)$. Once we had obtained the analytic signal from each IMF, $IF_i(t)$ and $IE_i(t)$ could be defined as in (2) and (3), by the phase derivative and envelope of the analytic signal, respectively.

$$IF_i(t) = \frac{1}{2\pi}\frac{d\Phi_i(t)}{dt} \quad (2)$$

$$IE_i(t) = |z_i(t)| = a_i(t) \quad (3)$$

A major challenge in calculating the IF for RS is to obtain a physically meaningful frequency, which means not having negative values. In relation to this issue, (2) would only yield true positive values for the IF if $s_i(t)$ were to be a mono-component signal, the phase signal of which is monotonically increasing point-by-point. However, IMFs are not truly mono-component signals and, therefore, their phase signals could produce unusual negative IF values.

In order to ensure real positive IF values, we developed a smoothing function, which consisted of applying a shape-preserving piecewise cubic interpolation to segments where the phase was not increasing.

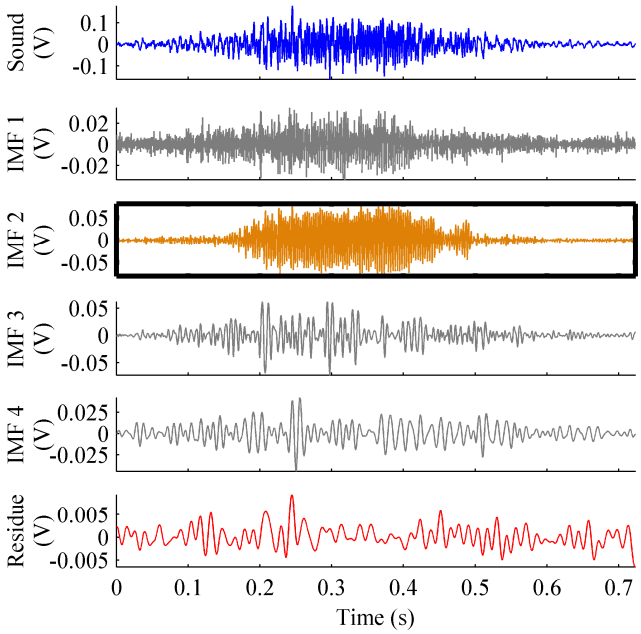


Fig. 1. EEMD applied to RS from an inspiratory cycle containing a wheeze. IMFs 1-4 cover all the frequency range of interest for CAS detection. The major components of the wheeze are within IMF 2 (solid line box).

The complete algorithm to calculate IF and IE comprises the following steps:

1) Calculate IMFs 1-4 by means of EEMD from RS of each inspiratory cycle.
2) Calculate the analytic signal, $z_i(t)$, using the Hilbert transform for each IMF.
3) Obtain the phase signal, $\Phi_i(t)$, for each $z_i(t)$.
4) Smooth the phase signal using the smoothing function, which prevents IF from taking negative values.
5) Calculate IF_i and IE_i by differentiating the smoothed $\Phi_i(t)$ and taking the absolute value of $z_i(t)$, respectively. The time derivative is calculated using a 5-point least squares polynomial approximation [39].

Fig. 2 shows an example of IF sequences obtained from two inspiratory cycles, one containing a wheeze (a) and another with normal RS (b). The IF dispersion was low for as long as the CAS lasted, as shown in IF 2 in Fig. 2 (a). In contrast, IFs from normal RS had uniform dispersions throughout the inspiratory cycle, as shown in Fig. 2 (b).

E. RS Classification

1) Extraction of Candidate CAS Segments

Our RS classifier works at two levels: the classification of RS segments within an inspiratory cycle, and the classification of the entire inspiratory cycle. The algorithm extracts RS segments that are candidates for being CAS segments within an inspiratory cycle. Then, feature extraction is applied to each extracted RS segment. These features are used to classify each RS segment, as a normal or CAS segment, using an SVM classifier. Finally, the whole inspiratory cycle is classified as containing CAS or normal sounds, depending on the classification of its RS segments (see Fig. 3).

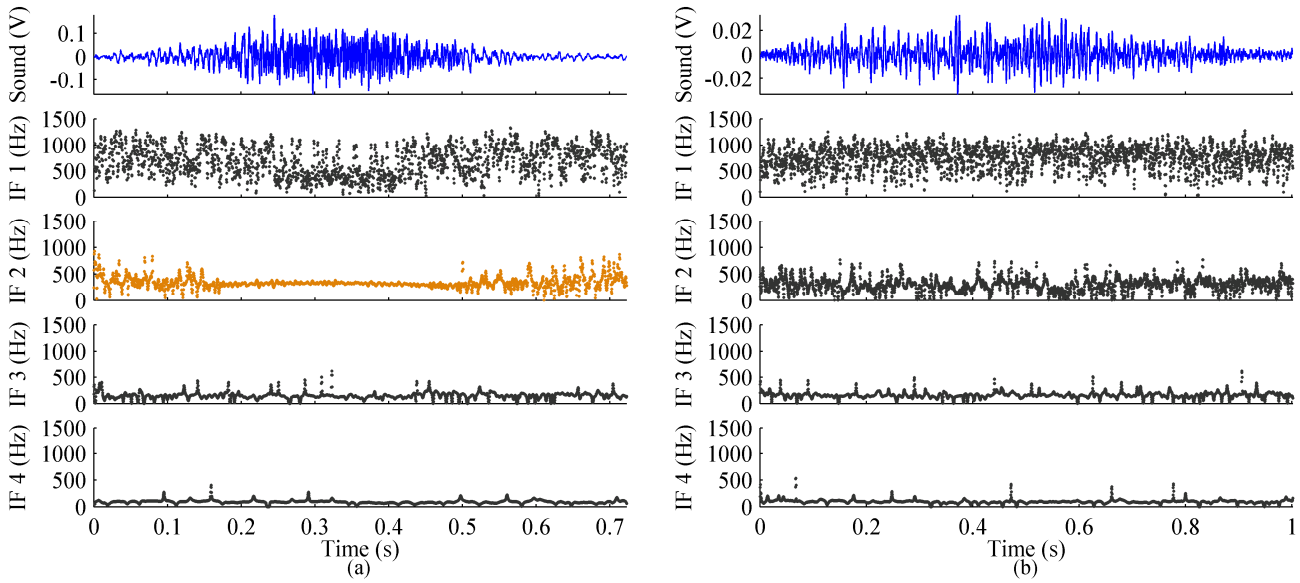


Fig. 2. Instantaneous frequencies (IF 1-IF 4) for IMFs 1-4 from the CAS cycle in Fig. 1 (a), and from one inspiratory cycle with normal RS (b). As shown in (a), a segment with a markedly lower IF dispersion appears within IF 2, and this corresponds to the location of the wheeze. In contrast, IFs from normal RS (b) are highly variable throughout the entire inspiratory cycle.

Given that CAS segments have lower IF dispersion than segments containing normal RS, candidate CAS segments can be detected using a moving window that slides over the entire IF sequences to calculate the IF dispersion (σ_{IFi}). In this study, we used a window length of 40 ms (125 samples at 3 125 samples/second) with 92% overlap (115 samples). Both parameters were determined empirically. The window length (40 ms) is less than half of the minimum CAS duration (100 ms). Shorter window lengths would produce highly variable σ_{IFi} sequences, due to unusual spurious IF values. On the other hand, longer window lengths would produce only slight variations in the σ_{IFi} sequences, making it more difficult to detect CAS segments. The 92% overlap was chosen to increase the number of data points in the σ_{IFi} sequences. If the overlap was small, CAS segments would produce only a few data points with low IF dispersion, whereas, by increasing the overlap between adjacent windows, we obtained more data points improving the accuracy of the detection of CAS segments.

After calculation of the σ_{IFi} sequences, a thresholding method was used to detect candidate CAS segments. Thresholds Th1, Th2, and Th3 represent different levels of IF dispersion (Fig. 3) and these were also determined empirically. Threshold Th1 was the mean of an entire σ_{IFi} sequence, that is, the mean IF dispersion throughout an entire inspiratory cycle. It was chosen as the reference value to detect CAS segments. The IF dispersion of normal RS segments was expected to be above Th1, whereas the IF dispersion of CAS segments was expected to be below Th1. However, using only Th1 was a weak criterion, which might lead to false CAS segment detections. To avoid false detections, we defined two more restrictive thresholds (Th2 and Th3). Th2 was the upper limit for the IF dispersion of a candidate CAS segment. That is, all segments with an IF dispersion below Th2 were potentially CAS segments. Furthermore, threshold Th3 was used to distinguish those segments with high probability of being CAS segments, due to their very low IF dispersion.

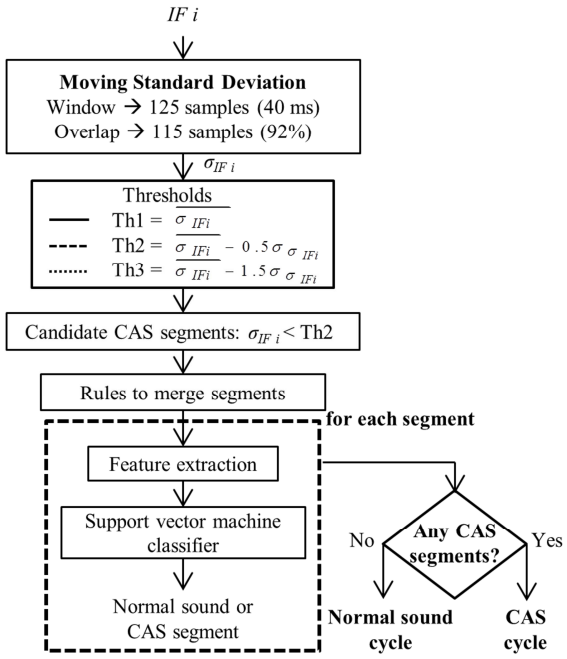


Fig. 3. Flowchart of the proposed RS classification algorithm.

Fig. 4 shows how IF can track not only constant monophonic CAS (a), but also variable monophonic (b) and polyphonic (c) CAS. As shown, the σ_{IFi} sequences are below Th2 throughout the presence of CAS.

2) Feature Extraction

Every candidate CAS segment was labeled as one of the

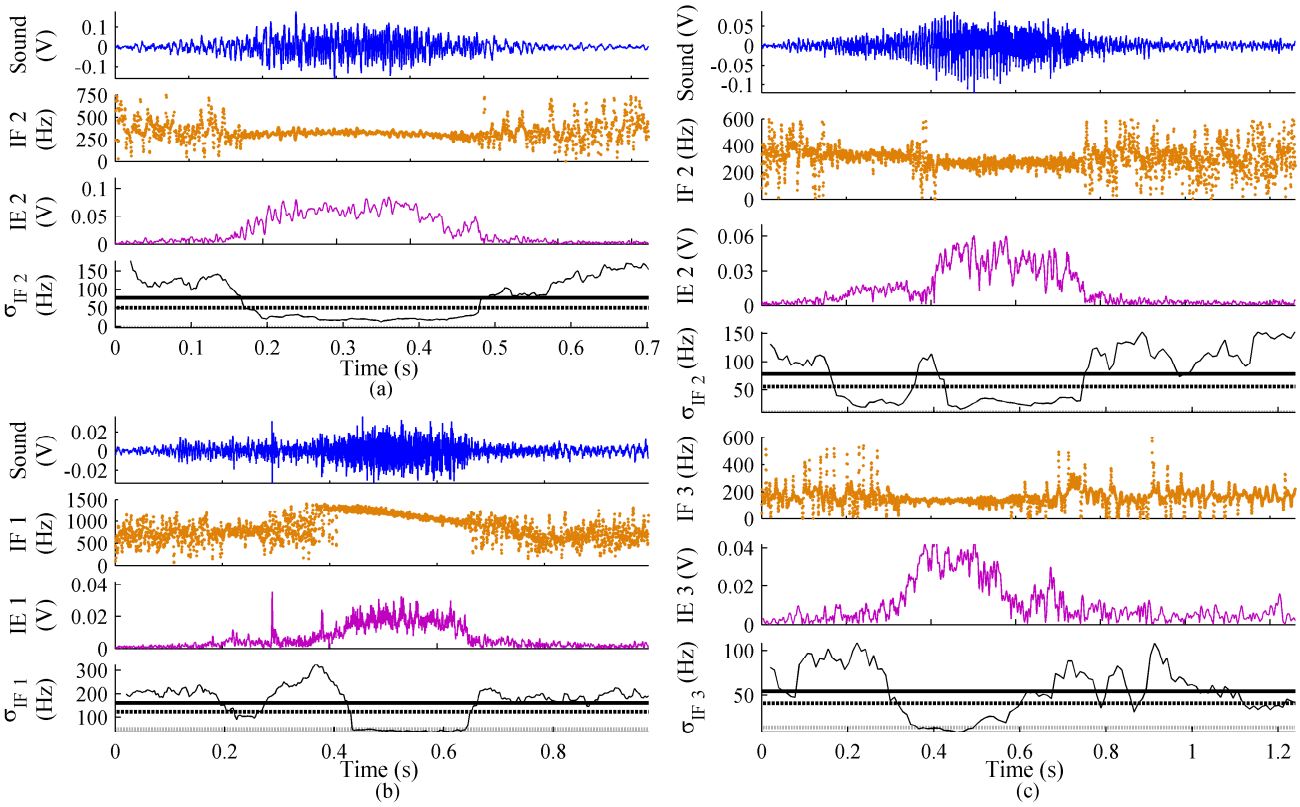


Fig. 4. Instantaneous frequency (IF i), instantaneous envelope (IE i), and IF dispersion sequence ($\sigma_{IF\,i}$) from: (a) one CAS cycle containing a constant monophonic wheeze, (b) one CAS cycle containing a variable monophonic wheeze with frequency sweeping, and (c) one CAS cycle containing a polyphonic wheeze, which main components are at mean frequencies of 327 Hz (IF 2), 266 Hz (IF 2), and 134 Hz (IF 3).

following types according to its minimum σ_{IFi} and duration, d_1 (Fig. 5-a):

Type 1: $min(\sigma_{IF}) \le Th3$, $d_1 \ge 100$ *ms*

Type 2: $min(\sigma_{IF}) \le Th3$, $d_1 < 100$ *ms*

Type 3: $min(\sigma_{IF}) > Th3$, $d_1 \ge 100$ *ms*

Type 4: $min(\sigma_{IF}) > Th3$, $d_1 < 100$ *ms*

As different RS segments might belong to the same CAS, a grouping algorithm was used to merge RS segments, based on maximum IF dispersion and time elapsed between two adjacent RS segments. Moreover, since the merged segments might have different labels, a priority order was established to set the label for the resulting compound segment, according to the following sequence:

Priority order: *Type 1* > *Type 2* > *Type 3* > *Type 4*

Each compound segment was labeled as the same type as its highest priority sub-segment.

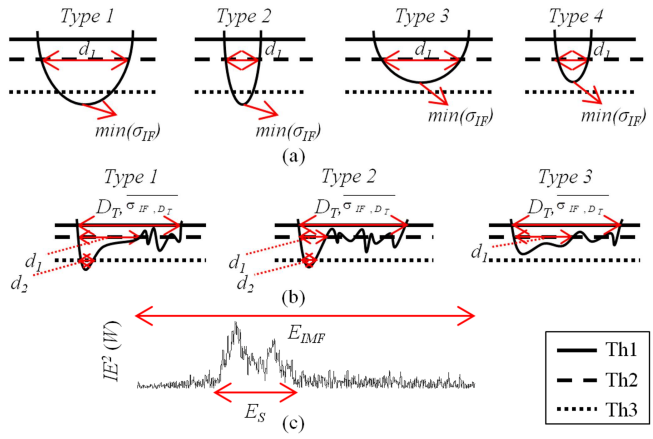


Fig. 5. Characteristic parameters of candidate CAS segments, extracted from the IF dispersion (a), (b), and the instantaneous envelope (IE) (c).

In the final classification stage, only type 1-3 RS segments were taken into account (Fig. 5-b). Moreover, according to the standard definition of CAS [1]-[3], only RS segments whose total duration (D_T) were greater than or equal to 100 ms were considered candidate CAS segments. The following features were defined, as in (4), (5), and (6), to characterize the segments (Fig. 5-b, and 5-c).

$$ER_1 = 100 * \frac{E_S}{E_{IMF}},\ ER_2 = 100 * \frac{E_S}{E_T},\ ETR = \frac{ER_2}{D_T} \quad (4)$$

$$DR = \frac{d_2}{d_1} \quad (5)$$

$$Q = \frac{\overline{\sigma_{IF,D_T}}}{\overline{\sigma_{IF}} - 0.75\sigma_{\sigma_{IF}}} \quad (6)$$

where E_{IMF} is the sum of all terms of the squared IE of an IMF, E_S is the sub-interval of E_{IMF} corresponding to the segment location, E_T is the sum of E_{IMF1}, E_{IMF2}, E_{IMF3}, and E_{IMF4}, and $\overline{\sigma_{IF,D_T}}$ is the mean IF dispersion over the length of the segment. These five features correspond to measurements that depend either on the energy (ER_1, ER_2, and ETR) or on the IF dispersion (DR and Q) of RS segments.

3) RS Segment Classification using Support Vector Machines

The proposed SVM classifier works at the segment level of our classification scheme (Fig. 3). That is, the input to the SVM classifier was a numerical data vector containing features ER_1, ER_2, *ETR*, *DR*, and *Q*, and the IMF number (from 1 to 4) of an RS segment. In order to train and test our SVM classifier, we used an iterative procedure, as illustrated in Fig. 6.

First, the inspiratory cycles from the full dataset were divided into two subsets, one for training (559 cycles) and another for testing (311 cycles). Due to the high between-subject variability of the RS, the subsets were formed independently and hence included inspiratory cycles from different asthmatic patients. Then, 921 RS segments extracted from the training inspiratory cycles were used to find the optimum parameters for our SVM classifier (the regularization parameter, C, and the kernel parameter, σ). For that purpose, we used a 5x5 grid with increasing sequences of C (from 1 to 3) and σ (from 0.1 to 1). Each combination of parameters was the starting point for finding a local minimum of the 10-fold cross-validation loss function. The parameters which produced the lowest cross-validation loss were chosen as the optimum parameters (C_{opt} = 2.72 and σ_{opt} = 1.61).

Having found the optimum SVM parameters, we employed them to train an SVM classifier using the RS segments from the training inspiratory cycles. The resulting SVM classifier was validated using the RS segments from both training and testing inspiratory cycles and the target labels from the manual RS segment classification. The performance of the SVM classifier was evaluated in terms of accuracy, sensitivity, specificity, and positive predictive value.

Since the results obtained might depend on the initial grouping of the inspiratory cycles, we repeated the training and testing steps using different partitions of the data. For that purpose, the inspiratory cycles from the dataset were randomly divided into training (65%) and testing (35%) sets inside a loop. The random partitions were done in such a way that each subset contained a proportionate share of the normal and CAS Q1-Q4 inspiratory cycles from the full dataset (Table II). After each random partition, the RS segments from the resulting training and testing inspiratory cycles were used to train and test an SVM classifier with the optimum parameters (C_{opt} and σ_{opt}). The total performance of the SVM classifier was calculated as the mean and the standard deviation of the classification results of each iteration.

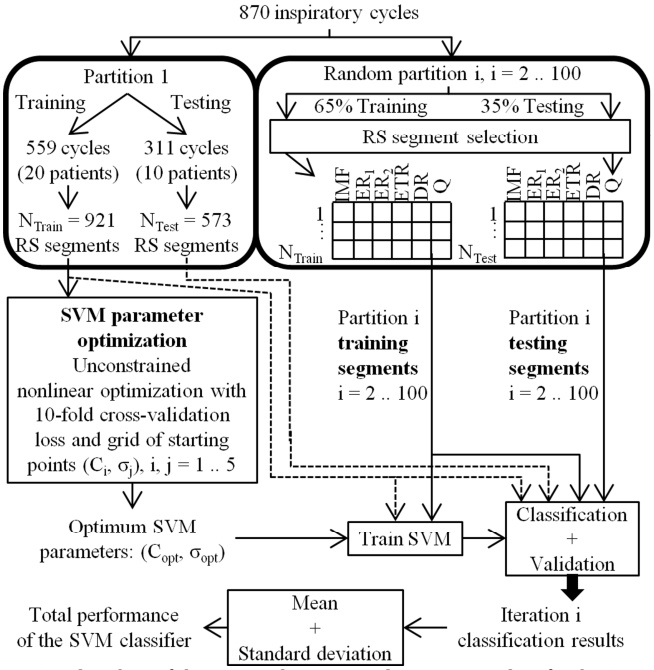


Fig. 6. Flowchart of the proposed training and testing procedure for the SVM classifier.

4) Inspiratory Cycle Classification

Considering the RS segment classification obtained in each iteration of the previous stage, each entire inspiratory cycle was classified as containing CAS if any of its RS segments was classified as a CAS segment. Otherwise, the entire inspiratory cycle was classified as containing normal sounds.

The performance of the SVM classifier was also evaluated at cycle level, by comparing the inspiratory cycle classification obtained with the classifier to the manual inspiratory cycle classification.

We selected the SVM with the best performance, among all iterations, for the classification of simulated CAS cycles in the next Section II.F.

F. Simulation of CAS Cycles

This section describes a procedure to study the effect of duration and SNR of simulated CAS on the performance of the classifier. For that purpose, the classifier was tested on 4592 simulated CAS cycles. The procedure to obtain the simulated CAS cycles was as follows:

1) Eight CAS segments were manually extracted from different inspiratory cycles with CAS from the study dataset.
2) Either time contraction or dilation was applied to those eight CAS segments in order to obtain 41 new CAS segments, $n_i(t)$, thus covering time durations from 100 ms to 500 ms, in increments of 10 ms.
3) Each one of those 41 CAS segments was added to 7 inspiratory cycles with normal sounds, $s_k(t)$, as in (7), at 16 SNRs, SNR_j, from -5 dB to 10 dB, in increments of 1 dB. As a result, a total of 656 simulated CAS cycles, $w_{ijk}(t)$, were obtained for each original normal sound cycle, $s_k(t)$.

$$w_{ijk}(t) = s_k(t) + n_{ijk}(t), i = 1..41, j = 1..16, k = 1..7 \quad (7)$$

In (7), *i* is the CAS segment subscript, *j* is the SNR subscript, and *k* is the subscript for the inspiratory cycle with normal sounds. In this study, the SNR was defined as in (8).

$$SNR_j = 10\,log(\frac{P_{n_{jk}}}{P_{S_k}}) \rightarrow P_{n_{jk}} = P_{S_k}\,10^{\frac{SNR_j}{10}} \quad (8)$$

$$n_{ijk}(t) = \sqrt{\frac{P_{n_{jk}}}{P_{n_i}}}\,n_i(t) \quad (9)$$

where $P_{S_k}(t)$ is the power of $s_k(t)$, $n_{ijk}(t)$ is the $n_i(t)$ CAS segment added to the $s_k(t)$ inspiratory cycle at SNR_j, $P_{n_{jk}}(t)$ is the power of $n_{ijk}(t)$, and P_{n_i} is the power of $n_i(t)$.

In this way, we obtained a set of simulated CAS cycles that allowed us to explore the effectiveness of the proposed algorithm in terms of two basic parameters of CAS: duration and intensity. The influence of these parameters on each simulated CAS cycle was calculated as the percentage increase in mean power with respect to the original cycle with normal sounds, $s_k(t)$, given by expression (10).

$$\Delta P_{ijk} = 100\frac{P_{w_{ijk}}}{P_{S_k}} \quad (10)$$

This new parameter allowed us to analyze the classification results of the simulated CAS cycles for different ranges of power increase.

III. Results

A. Recorded RS Classification

The performance of the SVM classifier at segment level, obtained after the training and testing described in Section II.E.3, is summarized in Table III. As shown, high accuracy (94.0% ± 0.8%) and sensitivity (92.8% ± 1.7%) were achieved with the test set for classifying recorded RS segments.

The performance of the SVM classifier at cycle level is reported in Table IV. As shown, our classifier achieved high total accuracy (94.6% ± 0.3%) and sensitivity (94.2% ± 0.4%) for classifying recorded RS from inspiratory cycles as normal sounds or CAS. Among all iterations, the best SVM classifier had an accuracy of 95.1%, a sensitivity of 94.2%, a specificity of 96.1%, and a positive predictive value of 96.8%, at cycle level. That SVM classifier was used to obtain the classification results for the simulated CAS cycles described in the next Section III.B.

TABLE III
Classification of Recorded RS Segments

Parameter/Set	Total	Training	Test
Accuracy (%)	94.9 ± 0.3	95.4 ± 0.4	94.0 ± 0.8
Sensitivity (%)	93.9 ± 0.6	94.5 ± 0.6	92.8 ± 1.7
Specificity (%)	95.6 ± 0.6	96.1 ± 0.6	94.8 ± 1.4
Positive Predictive Value (%)	94.1 ± 0.7	94.7 ± 0.8	93.0 ± 1.7

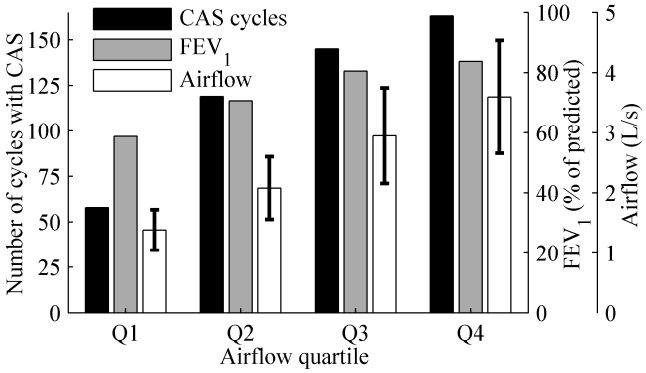


Fig. 7. Distribution of CAS cycles from the dataset among airflow quartiles (black). Mean FEV_1 of patients who provided CAS cycles at each airflow quartile (grey). Mean and standard deviation of the airflow levels included in each airflow quartile (white).

Furthermore, we evaluated the performance of our classifier for different airflow levels. Figure 7 shows the distribution of the CAS cycles from the dataset (485 cycles) among the four quartiles. Moreover, the mean FEV_1 value was calculated from the asthmatic subjects who provided CAS in each quartile. As shown, cycles containing CAS not only appear at high airflows (Q3 and Q4), but also at low and moderate airflows (Q1 and Q2). It is noteworthy that patients who provided CAS at low airflows had lower FEV_1 values than those who only provided CAS at high airflows. In fact, there is a direct relationship between the two variables. Patients with severe asthma (with low FEV_1 values) may produce CAS as much at low airflows as at high airflows. However, those with mild asthma only generate CAS at high airflow levels. Independently of the airflow quartile, our classifier performed very well in all cases (see Table IV).

TABLE IV
Classification of Recorded RS from Inspiratory Cycles

TOTAL	**Total**	Q1	Q2	Q3	Q4
Accuracy (%)	**94.6 ± 0.3**	95.3 ± 0.7	94.5 ± 0.6	94.1 ± 0.6	94.7 ± 0.6
Sensitivity (%)	**94.2 ± 0.4**	92.3 ± 1.2	94.7 ± 0.5	94.5 ± 0.8	94.4 ± 0.5
Specificity (%)	**95.0 ± 0.9**	97.0 ± 0.9	94.2 ± 1.2	93.6 ± 1.5	95.4 ± 1.6
Positive predictive value (%)	**96.0 ± 0.7**	94.4 ± 1.6	95.4 ± 0.9	95.1 ± 1.1	97.7 ± 0.8
TRAINING	**Total**	Q1	Q2	Q3	Q4
Accuracy (%)	**95.0 ± 0.6**	95.5 ± 1.3	94.8 ± 1.3	94.7 ± 1.2	95.2 ± 1.1
Sensitivity (%)	**94.6 ± 0.7**	92.5 ± 2.7	94.9 ± 1.4	94.8 ± 1.3	94.9 ± 1.3
Specificity (%)	**95.6 ± 1.0**	97.2 ± 1.3	94.7 ± 1.9	94.5 ± 2.2	96.0 ± 2.4
Positive predictive value (%)	**96.4 ± 0.8**	94.9 ± 2.3	95.8 ± 1.5	95.8 ± 1.6	98.0 ± 1.1
TESTING	**Total**	Q1	Q2	Q3	Q4
Accuracy (%)	**93.7 ± 1.2**	94.9 ± 2.3	93.8 ± 2.2	92.9 ± 2.1	93.8 ± 2.3
Sensitivity (%)	**93.6 ± 1.9**	91.9 ± 5.1	94.2 ± 2.8	93.8 ± 3.1	93.5 ± 2.7
Specificity (%)	**93.9 ± 2.0**	96.5 ± 2.9	93.2 ± 3.5	91.8 ± 4.1	94.3 ± 4.0
Positive predictive value (%)	**95.1 ± 1.5**	93.8 ± 4.9	94.7 ± 2.6	94.0 ± 2.8	97.2 ± 1.9

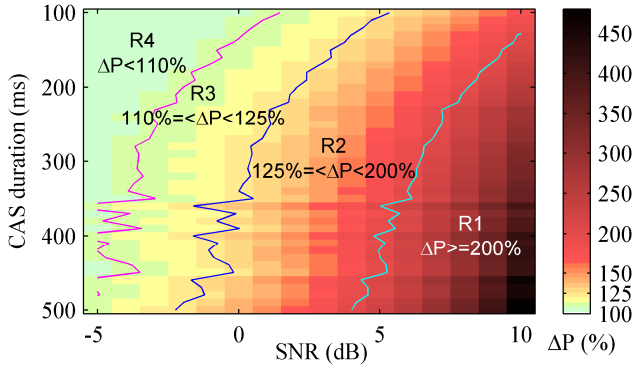


Fig. 8. Percentage increase of mean power (ΔP_{ijk}) of simulated CAS cycles, $w_{ijk}(t)$, with respect to their original normal sound cycles, $s_k(t)$, as a function of SNR and duration of CAS segments.

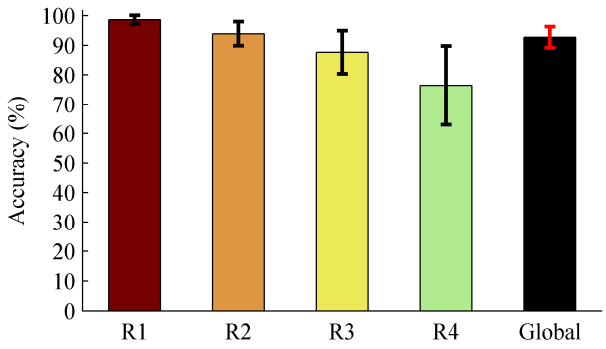


Fig. 9. Accuracy in simulated CAS cycle detection for regions R1-R4, which represent different SNR conditions (see Fig. 8).

B. *Effect of SNR on Detection of Simulated CAS Cycles*

This section describes the results of the sensitivity analysis described in Section II.F.

Fig. 8 shows the results after obtaining 656 simulated CAS cycles from one inspiratory cycle with normal sounds, $s_k(t)$. The power increase (ΔP_{ijk}) parameter was used to evaluate the effect of adding CAS segments to an inspiratory cycle with normal sounds, with different durations and different SNRs. Three thresholds were set for power increases of 200%, 125%, and 110%, defining the boundaries for four regions of interest (R1-R4), plotted in Fig. 8.

The same procedure (described in Section II.F) was applied to six more inspiratory cycles with normal sounds, $s_k(t)$. Then, the classifier was applied to the total set of 4592 simulated CAS cycles. As shown in Fig. 9, the proposed classification algorithm achieved high accuracy for detecting cycles containing CAS in regions R1 (98.7% ± 1.4%) and R2 (93.9% ± 4.1%), while the accuracy was markedly lower in region R4 (76.4% ± 13.4%). However, region R4 corresponds to very low SNR values. This region included weak CAS, which had very low energy and only slightly increased the mean power of the RS signals ($\Delta P \leq 110\%$). Although there is no standard criterion for amplitude in the definition of CAS, considering the recorded CAS in the study dataset, regions R1-R2 represented more realistic SNR values. For this reason, the number of simulated CAS cycles progressively decreased from region R1 to region R4. In any case, the overall accuracy, including all regions, was 92.8% ± 3.6%. This high accuracy demonstrates that the proposed method offers high performance under both low and high SNR conditions. In fact, high accuracy (87.7% ± 7.3%) was achieved in region R3, which still represents an unfavorable scenario, in which the SNR may be below 0 dB.

IV. Discussion And Conclusions

Asthma is characterized by a series of variable symptoms, including airflow limitation, shortness of breath, cough, and the presence of CAS [9]. Although these symptoms vary over time and CAS are not always present in asthma, when present, such sounds clearly indicate airway obstruction [4]. Therefore, detecting and analyzing CAS can provide some additional information about the pulmonary function of these patients.

This study demonstrates that the analysis of RS in terms of EEMD-based IFs provides an efficient, simple, and robust method for CAS detection and RS classification.

First, we have shown the viability of estimating the IF by EEMD. The definition of the IF of multi-component signals, such as RS, has been a subject of controversy, since strictly speaking the concept of IF is only meaningful for mono-component signals. In this work, we found that EEMD allowed us to decompose RS into IMFs for which a physically meaningful IF was defined point-by-point. EEMD represented a key step in the multi-scale analysis of the IF in RS, as it determined the quality of the IF estimation and thus the performance of our RS classifier.

Indeed, we opted to use EEMD as it improves on the original EMD method, which had a mode mixing effect when applied to RS from some inspiratory cycles [33]. After comparing several methods addressing that mode mixing effect [34], [40]-[42], we identified some clear advantages of EEMD. The IMFs from EEMD had smaller frequency overlaps than IMFs from other methods. Due to properties similar to a dyadic filter bank, EEMD achieved a much better separation of different frequency scales, which reduced the mode mixing effect and improved the IF estimate. Moreover, thanks to the better separation of frequency scales, EEMD allowed us to analyze only IMFs 1-4 to cover the frequency range of interest, which goes from 100 Hz upwards for CAS detection. However, other methods required us to analyze at least 5 IMFs for the same purpose.

Two parameters determined the effectiveness of the EEMD: noise amplitude and number of iterations. If the added noise amplitude was too small (high SNR), there were few changes in the maxima of the original sound signal and, therefore, the benefits of EEMD were not evident. On the other hand, using high amplitude noise (low SNR) produced slightly noisy IMF components, but provided a reference scale distribution to enhance EMD and avoid the mode mixing effect. Furthermore, the residual noise level was able to be reduced by increasing the number of iterations. Indeed, a key issue was finding a balance between the SNR and number of iterations in order to minimize the mode mixing effect and obtain an acceptable residual noise level.

As RS are random in nature, for this study, we added band-limited noise to RS signals from inspiratory cycles at an SNR of 0 dB, in order to perceive the benefits of EEMD. Despite this low SNR, applying 100 iterations to RS from each inspiratory cycle proved sufficient to obtain slightly noisy IMFs that allowed us to calculate a meaningful IF in a moderate time. Using fewer iterations would result in a higher residual noise level, whereas using more iterations would substantially increase the computation time. Furthermore, the slight residual noise was not a problem for our application, since we did not need a perfect decomposition to reconstruct the original signal from the IMFs, but rather we wanted to use them for IF estimates.

Although there are several methods of IF calculation [32], we found that the Hilbert method produced more accurate IF values than other methods, such as the Teager-Huang method, which produced more dispersed values.

We have presented a set of features, which were extracted from the IF and IE sequences, to train and test an SVM classifier for RS segments. For that purpose, a moving window was applied to the IF sequence from each IMF to calculate the dispersion, since we found that RS segments containing CAS were characterized by markedly lower IF dispersion [33]. We used the standard deviation just as a tool for measuring dispersion within each IF sequence. Then, we applied a set of thresholds (Fig. 3) to the IF dispersion in order to detect candidate CAS segments. Since these thresholds were proportional to the IF dispersion of each IMF, they adapted to the characteristics of each signal. The extracted IF dispersion and energy features (Fig. 5-b and 5-c) allowed us to achieve high sensitivity (94.2%) and accuracy (94.6%) in CAS detection and RS classification.

As shown in Table V, the performance of our method is similar to or better than that reported for many previous CAS detection methods. However, these different approaches cannot be directly compared because there are some important differences between them, which affect the classification results. The main differences are in the following: the position and number of sensors (tracheal or lung sounds, monochannel or multichannel recordings), the respiratory maneuver performed (constant airflow, forced expiratory maneuver, or variable airflow), the respiratory disease of patients (asthma, COPD, or others), the size of the dataset, the types of CAS included in the dataset, or the parameters for evaluating the performance of the classifiers.

Despite the aforementioned differences, our approach has some objective advantages over previous approaches. We have shown the ability of our method to detect CAS derived from a wide range of airflow levels. Thanks to the progressive respiratory maneuver, which was performed by the asthmatic patients, the dataset of this study included CAS appearing at low airflows as well as that appearing at high airflows. This is a strength compared to previous studies for which RS were recorded at constant airflows or during forced expiratory maneuvers. We have shown that CAS may appear under different respiratory conditions, and that this has no effect on the performance of our classifier.

TABLE V
PERFORMANCE COMPARISON OF CAS DETECTION ALGORITHMS

Features	Year	Method	Performance (%)
EEMD + IF	2014	Proposed method	Real: 94.6 (Acc), 94.2 (S), 95.0 (Sp) Sim: 92.8 (Acc)
Spectrogram	2004	Homs [18]	71 < (S) < 100
Spectrogram	2005	Hsueh [24]	89 < (S) & (Sp)
Spectrogram	2006	Lin [25]	96.7 (S), 90.9 (Sp)
Spectrogram	2007	Taplidou [19]	95.5 (S), 93.7 (Sp)
Spectrogram	2008	Jain [20]	84 (S), 86 (Sp)
Spectrogram	2009	Riella [21]	84.8 (Acc)
MFCC + GMM	2007	Chien [14]	90 < (Acc)
MFCC + GMM	2009	Bahoura [15]	94.6 (S), 91.9 (Sp)
Linear analysis	2009	Aydore [16]	93.5 (Acc)
Dominance spectrogram	2011	Jin [28]	92.4 (Acc)
FT + DWT + PCA	2012	Xie [12]	97.3 (Acc)

Acc: accuracy, S: sensitivity, Sp: specificity, MFCC: Mel frequency cepstral coefficients, GMM: Gaussian mixture models, FT: Fourier transform, DWT: discrete wavelet transform, PCA: principal component analysis.

We have also performed a detailed sensitivity analysis of the influence of the SNR (amplitude of CAS) on the performance of the proposed classifier. We applied our classifier to simulated CAS cycles with different SNRs (see Fig. 8). As shown in Fig. 9, an overall accuracy of 92.8% was achieved. Although the performance of the RS classifier diminished from region R1 to R4, the accuracy was high (87.7% in region R3) even at low SNRs. These results reinforce the reliability of our method, in that they demonstrate that this technique works properly under both high and low SNR conditions.

In contrast to previous approaches for RS analysis, such as those based on spectrograms or wavelet transformations, EEMD-IF used in the first stage of our classification scheme is an adaptive technique, which does not require a priori knowledge of the RS signal characteristics. That is, IF and IE sequences are calculated without choosing fixed analysis parameters and regardless of the type of RS and their temporal or spectral characteristics. On the other hand, IF and IE sequences allow us to work independently in either a time-frequency or a time-energy domain. For this reason, we were able to use dispersion-based criteria on IF sequences, this representing a novel and straightforward technique for detecting CAS segments. Moreover, IF and IE sequences are defined point-by-point, thus providing very high time resolution. In fact, EEMD-IF provides precise IF values, which represent the frequency content of RS signals at each time instant.

Thanks to the aforementioned properties of EEMD-IF, on which the proposed method relies, this is a suitable technique for RS classification, and also represents an alternative way of accurately analyzing RS signals. Consequently, unlike some techniques previously used for CAS detection, EEMD-IF could be used not only for detecting CAS, but also for characterizing these sounds. Although this proposed EEMD-IF technique has not been widely applied to RS analysis, various EMD-based approaches have already been used in previous studies either for detection and analysis of discontinuous adventitious sounds [43], [44], or for detection and separation of heart sounds from lung sounds [45]-[47].

In this study, we have focused on inspiratory cycles because inspiratory sounds are much louder than expiratory sounds on the back, where we recorded the analyzed RS. Nevertheless, this classifier could be effectively applied to RS from expiratory cycles and even from the full respiratory cycle. What is more, although this study is focused on RS from the surface of the back, the proposed classifier could be also applied to tracheal sounds. These factors do not have influence on the hypothesis of our method. That is, it depends on the variations of the IF dispersion inside a cycle, regardless of whether it is an inspiratory, an expiratory, or a full cycle.

The use of our RS classifier is a prior step for a detailed analysis of RS. Specifically, easier CAS detection will facilitate new research into the analysis and characterization of these sounds. On the other hand, the identification of normal inspiratory sounds, without CAS, is the basis for future analysis of normal inspiratory sound intensity. This classifier could be the first stage of a more complex system for the analysis of RS. Such a system could serve as much for the diagnosis of patients with obstructive respiratory diseases as for their long-term monitoring. The idea is that this type of system could be used in a routine way together with the spirometry. The combined information from these techniques could increase the reliability in the diagnostic assessment of these patients.

ACKNOWLEDGMENT

This work was made possible thanks to a collaboration agreement, between IBEC and IGTP, to create a joint research Unit. All authors would like to thank all the team in the Pulmonary Function Testing Laboratory at Germans Trias i Pujol University Hospital, for their collaboration in the patient recruitment as well as the performing of the spirometry tests.

REFERENCES

[1] A. Bohadana, G. Izbicki, and S.S. Kraman, "Fundamentals of lung auscultation," *N. Engl. J. Med.*, vol. 370, pp. 744–751, Feb. 2014.

[2] A.R.A. Sovijärvi, L.P. Malmberg, G. Charbonneau, J. Vanderschoot, F. Dalmasso, C. Sacco, M. Rossi, and J.E. Earis, "Characteristics of breath sounds and adventitious respiratory sounds," *Eur. Respir. Rev.*, vol. 10, pp. 591–596, 2000.

[3] A.R.A. Sovijärvi, F. Dalmasso, J. Vanderschoot, L.P. Malmberg, G. Righini, and S.A.T. Stoneman, "Definitions of terms for applications of respiratory sounds," *Eur. Respir. Rev.*, vol. 10, pp. 597–610, 2000.

[4] N. Meslier, G. Charbonneau, and J.L. Racineux, "Wheezes," *Eur. Respir. J.*, vol. 8, pp. 1942–1948, Nov. 1995.

[5] J.J. Marini, D.J. Pierson, L.D. Hudson, and S. Lakshminarayan, "The significance of wheezing in chronic airflow obstruction," *Am. Rev. Respir. Dis.*, vol. 120, pp. 1069–1072, Nov. 1979.

[6] J.A. Fiz, R. Jané, A. Homs, J. Izquierdo, M.A. García, and J. Morera, "Detection of wheezing during maximal forced exhalation in patients with obstructed airways," *Chest*, vol. 122, pp. 186–191, Jul. 2002.

[7] C.S. Shim and M.H. Williams, "Relationship of wheezing to the severity of obstruction in asthma," *Arch. Intern. Med.*, vol. 143, pp. 890–892, May 1983.

[8] J.A. Fiz, R. Jané, J. Izquierdo, A. Homs, M.A. García, R.Gómez, E. Monso, and J. Morera, "Analysis of forced wheezes in asthma patients," *Respiration*, vol. 73, pp. 55–60, Aug. 2006.

[9] E.D. Bateman *et al*., "Global strategy for asthma management and prevention: GINA executive summary," *Eur. Respir. J*., vol. 31, 143–78, Jan. 2008.

[10] A. Kandaswamy, C.S. Kumar, R.P. Ramanathan, S. Jayaraman, and N. Malmurugan, "Neural classification of lung sounds using wavelet coefficients," *Comput. Biol. Med.*, vol. 34, pp. 523–537, Sep. 2004.

[11] A. Mondal, P. Bhattacharya, and G. Saha, "Detection of lungs status using morphological complexities of respiratory sounds," *Sci. World J.*, vol. 2014, 9 pages, Feb. 2014.

[12] S. Xie, F. Jin, S. Krishnan, and F. Sattar, "Signal feature extraction by multi-scale PCA and its application to respiratory sound classification," *Med. Biol. Eng. Comput.*, vol. 50, pp. 759–768, Jul. 2012.

[13] S. Matsunaga, K. Yamauchi, M. Yamashita, and S. Miyahara, "Classification between normal and abnormal respiratory sounds based on maximum likelihood approach," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Taipei, 2009, pp. 517–520.

[14] J.C. Chien, H.D. Wu, F.C. Chong, and C.I. Li, "Wheeze detection using cepstral analysis in Gaussian mixture models," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, Lyon, 2007, pp. 3168–3171.

[15] M. Bahoura, "Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes," *Comput. Biol. Med.*, vol. 39, pp. 824–843, Sep. 2009.

[16] S. Aydore, I. Sen, Y.P. Kahya, and M. Mihcak, "Classification of respiratory signals by linear analysis," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, Minneapolis, 2009, pp. 2617–2620.

[17] S.A. Taplidou, L.J. Hadjileontiadis, T. Penzel, V. Gross, and S.M. Panas, "WED: An efficient wheezing-episode detector based on breath sounds spectrogram analysis," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, Cancun, 2003, pp. 2531–2534.

[18] A. Homs-Corbera, J.A. Fiz, J. Morera, and R. Jané, "Time-frequency detection and analysis of wheezes during forced exhalation," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 1, pp. 182–186, Jan. 2004.

[19] S.A. Taplidou and L.J. Hadjileontiadis, "Wheeze detection based on time-frequency analysis of breath sounds," *Comput. Biol. Med.*, vol. 37, pp. 1073–1083, Aug. 2007.

[20] A. Jain and J. Vepa, "Lung sound analysis for wheeze episode detection," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, Vancouver, 2008, pp. 2582–2585.

[21] R.J. Riella, P. Nohama, and J.M. Maia, "Method for automatic detection of wheezing in lung sounds," *Braz. J. Med. Biol. Res.*, vol. 42, pp. 674–684, Jul. 2009.

[22] J. Zhang, W. Ser, J. Yu, and T.T. Zhang, "A novel wheeze detection method for wearable monitoring systems," in *Int. Symp. Intell. Ubiquitous Comput. Educ.*, Chengdu, 2009, pp. 331–334.

[23] A. Oliveira, C. Pinho, J. Dinis, D. Oliveira, and A. Marques, "Automatic wheeze detection and lung function evaluation: a preliminary study," in *Proc. Int. Conf. Health Inform.*, Barcelona, 2013, pp. 323–326.

[24] M.L. Hsueh, J.C. Chien, F.C. Chang, H.D. Wu, and F.C. Chong, "Respiratory wheeze detection system," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, Shanghai, 2005, pp. 7553–7559.

[25] B.S. Lin, B.S. Lin, H.D. Wu, F.C. Chong, and S.J. Chen, "Wheeze recognition based on 2D bilateral filtering of spectrogram," *Biomed. Eng. Appl. Basis Commun.*, vol. 18, pp. 128–137, Jun. 2006.

[26] S.A. Taplidou *et al*., "On applying continuous wavelet transform in wheeze analysis," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, San Francisco, 2004, pp. 3832–3835.

[27] S.A. Taplidou and L.J. Hadjileontiadis, "Analysis of wheezes using wavelet higher order spectral features," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 7, pp. 1596–1610, Jul. 2010.

[28] F. Jin, S.S. Krishnan, and F. Sattar, "Adventitious sounds identification and extraction using temporal-spectral dominance-based features," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 11, pp. 3078–3087, Nov. 2011.

[29] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, and H.H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. A*, vol. 454, pp. 903–995, Mar. 1998.

[30] N.E. Huang, X. Chen, M.T. Lo, and Z. Wu, "On Hilbert spectral representation: a true time-frequency representation for nonlinear and nonstationary data," *Adv. Adapt. Data Anal.*, vol. 3, pp. 63–93, Apr. 2011.

[31] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal–Part 1: Fundamentals," *Proc. IEEE*, vol. 80, no. 4, pp. 520–538, Apr. 1992.

[32] N.E. Huang, Z. Wu, S.R. Long, K.C. Arnold, X. Chen, and K. Blank, "On instantaneous frequency," *Adv. Adapt. Data Anal.*, vol. 1, pp. 177–229, Apr. 2009.

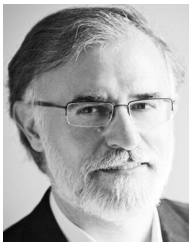
[33] M. Lozano, J.A. Fiz, and R. Jané, "Estimation of instantaneous frequency from empirical mode decomposition on respiratory sounds analysis," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, Osaka, 2013, pp. 981–984.
[34] Z. Wu and N.E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, pp. 1–41, Jan. 2009.
[35] J. Zhang, R. Yan, R.X. Gao, and Z. Feng, "Performance enhancement of ensemble empirical mode decomposition," *Mech. Syst. Signal Process.*, vol. 24, pp. 2104–2123, Oct. 2010.
[36] G. Rilling, P. Flandrin, and P. Gonçalvès, "On empirical mode decomposition and its algorithms," in *Proc. IEEE/EURASIP Workshop Nonlinear Signal Image Process.*, Grado, 2003.
[37] http://perso.ens-lyon.fr/patrick.flandrin/emd.html
[38] D. Gabor, "Theory of communication," *J. IEE–Part III: Radio Commun. Eng.*, vol. 93, no. 26, pp. 429–457, Nov. 1946.
[39] W.J. Tompkins, "Finite impulse response filters," in *Biomedical Digital Signal Processing*, New Jersey: Prentice-Hall, 1993, pp. 111–116.
[40] N. Rehman and D.P. Mandic, "Multivariate empirical mode decomposition," *Proc. R. Soc. A*, vol. 466, pp. 1291–1302, 2010.
[41] N. Rehman and D.P. Mandic, "Filter bank property of multivariate empirical mode decomposition," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2421–2426, May 2011.
[42] H. Hong, X. Wang, and Z. Tao, "Local integral mean-based sifting for empirical mode decomposition," *IEEE Signal Process. Lett.*, vol. 16, no. 10, pp. 841–844, Oct. 2009.
[43] L.J. Hadjileontiadis, "Empirical mode decomposition and fractal dimension filter. A novel technique for denoising explosive lung sounds," *IEEE Eng. Med. Biol. Mag.*, vol. 26, no. 1, pp. 30–39, 2007.
[44] B.A. Reyes, S. Charleston-Villalobos, R. Gonzalez-Camarena, and T. Aljama-Corrales, "Analysis of discontinuous adventitious lung sounds by Hilbert-Huang spectrum," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, Vancouver, 2008, pp. 3620–3623.
[45] S. Charleston-Villalobos, L.F. Dominguez-Robert, R. Gonzalez-Camarena, and A.T. Aljama-Corrales, "Heart sounds interference cancellation in lung sounds," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, New York, 2006, pp. 1694–1697.
[46] C. Papadaniil and L.J. Hadjileontiadis, "Efficient heart sound segmentation and extraction using ensemble empirical mode decomposition and kurtosis features," *IEEE J. Biomed. Health Inform.*, vol. 18, no. 4, pp. 1138–1152, Jul. 2014.
[47] L. ChingShun, W.A. Tanumihardja, and S. HongHui, "Lung-heart sound separation using noise assisted multivariate empirical mode decomposition," in *Int. Symp. Intell. Signal Process. Commun. Syst.*, Naha, 2013, pp. 726–730.

Manuel Lozano was born in Barcelona, Spain, in 1983. He received his B.Eng. degree in Telecommunications Engineering and M.S.E. in Biomedical Engineering from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 2008 and 2011, respectively.

He is currently working towards a Ph.D. degree in Biomedical Engineering at UPC. Since 2011, he has been working as a research technician in the Innovation Mixed Unit at the Health Sciences Research Institute of the Germans Trias i Pujol Foundation (IGTP), Badalona, Spain, and the Institute for Bioengineering of Catalonia (IBEC), Barcelona, Spain. His research interests include biomedical signal processing and multichannel respiratory sound analysis in respiratory diseases.

José Antonio Fiz was born in Pamplona, Spain, in 1953. He received his Diploma in Medicine from the University of Barcelona, Barcelona, Spain, in 1978. He specialized in Pulmonology at the University of Valencia and obtained his M.S.E. in Bioengineering from the UPC, Barcelona, Spain, in 1994.

Since 1991, he has been working in the respiratory physiology and sleep laboratories at the Germans Trias i Pujol University Hospital, Badalona, Spain. In 2008, he joined the IBEC and the Biomedical Research Networking center in Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN). His research interests include respiratory sound analysis, sleep apnea, neural networks applied to clinical data, and clinical time series analysis.



Raimon Jané (M'91, SM'14) received his Ph.D. from the UPC, Barcelona, Spain, in 1989.

He is currently Director of Research in the Department of Automatic Control (ESAII), UPC, and the Scientific Group Leader of the Biomedical Signal Processing and Interpretation Group, IBEC, Barcelona. Since 2008, he has been the Principal Investigator of the Biomedical Signals and Systems (SISBIO) Group of the CIBER-BBN. He is Professor of the Master's and the Coordinator of the Ph.D. programs in Biomedical Engineering. His research interests include multi-modal and multi-scale biomedical signal processing in cardiorespiratory and brain diseases.

Prof. Jané was a recipient of the Barcelona City Prize, from the Barcelona Council, in the area of technology research, in 2005. Currently, he is the President of the Spanish Society of Biomedical Engineering (SEIB).

Chapter 4: Performance evaluation of the Hilbert–Huang transform for respiratory sound analysis and its application to continuous adventitious sound characterization

Title: Performance evaluation of the Hilbert–Huang transform for respiratory sound analysis and its application to continuous adventitious sound characterization

Authors: M. Lozano, J. A. Fiz, R. Jané

Journal: Signal Processing, 2016

Performance evaluation of the Hilbert-Huang transform for respiratory sound analysis and its application to continuous adventitious sound characterization

Manuel Lozano [a,c,*], José Antonio Fiz [a,b,d], Raimon Jané [a,d,e]

[a] Institute for Bioengineering of Catalonia (IBEC), Baldiri Reixac, 4, Tower I, 9th floor, 08028, Barcelona, Spain
[b] Pulmonology Service at Germans Trias i Pujol University Hospital, Ctra. de Canyet, s/n, 08916, Badalona, Spain
[c] Health Sciences Research Institute of the Germans Trias i Pujol Foundation (IGTP), Ctra. de Can Ruti, Camí de les escoles, s/n, 08916, Badalona, Spain
[d] Biomedical Research Networking Center in Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN)
[e] Department of Automatic Control (ESAII), Universitat Politècnica de Catalunya (UPC), Barcelona, Spain

Keywords:
Hilbert–Huang transform
Ensemble empirical mode decomposition
Instantaneous frequency
Respiratory sounds
Continuous adventitious sounds

ABSTRACT

The use of the Hilbert-Huang transform in the analysis of biomedical signals has increased during the past few years, but its use for respiratory sound (RS) analysis is still limited. The technique includes two steps: empirical mode decomposition (EMD) and instantaneous frequency (IF) estimation. Although the mode mixing (MM) problem of EMD has been widely discussed, this technique continues to be used in many RS analysis algorithms.

In this study, we analyzed the MM effect in RS signals recorded from 30 asthmatic patients, and studied the performance of ensemble EMD (EEMD) and noise-assisted multivariate EMD (NA-MEMD) as means for preventing this effect. We propose quantitative parameters for measuring the size, reduction of MM, and residual noise level of each method. These parameters showed that EEMD is a good solution for MM, thus outperforming NA-MEMD. After testing different IF estimators, we propose Kay's method to calculate an EEMD-Kay-based Hilbert spectrum that offers high energy concentrations and high time and high frequency resolutions. We also propose an algorithm for the automatic characterization of continuous adventitious sounds (CAS). The tests performed showed that the proposed EEMD-Kay-based Hilbert spectrum makes it possible to determine CAS more precisely than other conventional time-frequency techniques.

1. Introduction

Respiratory sounds (RS) are multicomponent, nonlinear, and non-stationary signals. In general, RS signals are comprised of normal RS and may contain superimposed abnormal RS, such as continuous (CAS) and discontinuous (DAS) adventitious sounds, as well as different types of noise, such as clicks, background talking, or heart sounds. Normal RS are random in nature, whereas CAS are quasi-periodic waveforms with a duration of more than 80-100 ms and a fundamental frequency of over 100 Hz, and DAS are transient and short sounds (around 20 ms), with high frequency components (above 300 Hz) [1,2]. Therefore, RS are complex signals made up of a set of components, each one having different time-frequency features.

Due to the different and variable characteristics of RS, time-frequency distributions (TFDs) have become the

*Corresponding author at: Institute for Bioengineering of Catalonia (IBEC), Baldiri Reixac, 4, Tower I, 9th floor, 08028 Barcelona, Spain. Tel.: þ34 695119250.
E-mail addresses: **mlozanogar@gmail.com (M.Lozano), jafiz@msn.com (J.A.Fiz), rjane@ibecbarcelona.eu (R.Jané).**

the most commonly used and straightforward techniques for RS characterization. In CAS analysis, spectrogram has been the most widely used TFD [3-7], despite its poor and window-dependent resolution. Nevertheless, more advanced TFDs have recently been proposed for CAS analysis, either through combining wavelet decomposition with third order spectra features [8], or by deriving a temporal-spectral dominance spectrogram from the short-time Fourier transform [9].

As opposed to CAS analysis, DAS analysis requires TFDs with higher time resolution than spectrogram. Wavelet-based techniques, such as scalogram, have been widely used for DAS detection [10-12]. In addition, DAS have also been analyzed by means of nonlinear techniques, such as kurtosis and fractal dimension as measures of gaussianity and complexity, respectively [13-15].

Besides Fourier and wavelet-based techniques, one of the most relevant parameters of time-frequency analysis for nonlinear and non-stationary signals, such as RS, is the instantaneous frequency (IF), which consists of the frequency content of a signal at each time instant [16]. The concept of IF has led to the definition of TFDs that highly concentrate the energy of a signal along its IF, which makes it possible to identify signal components more precisely. Several IF estimators have been proposed, such as the phase derivative of the analytic signal associated with a real signal [17,18], zero-crossing [19], or adaptive IF estimators based on data modeling [20-22]. However, the most common IF estimators are based on TFDs [16,23], which give IF estimates with lower variance.

Quadratic TFDs, such as the Wigner-Ville distribution (WVD), were defined based on the IF concept with the aim of improving the resolution and concentration of energy of spectrogram [16]. However, a major drawback of the WVD is the presence of cross-terms, which complicate IF estimation, especially in multicomponent signals. In order to reduce cross-terms, other quadratic TFDs have been proposed as filtered versions of the WVD, using different time-frequency smoothing kernels, such as the smoothed pseudo-WVD [24] or the reduced interference distributions [25], which reduce cross-terms while maintaining high resolution.

In addition to these smoothing approaches for cross-term reduction, other techniques have been proposed to increase the signal energy concentration and resolution of different TFDs. For example, the adaptive short-time Fourier transform [26,27] uses a variable window length adapted to signal characteristics in order to improve the resolution of spectrogram. Moreover, reassignment techniques [28,29] are alternative approaches for the enhancement of TFDs, especially the synchrosqueezing transform [30], which allows mode retrieval in multicomponent signals.

Beyond the calculation of the aforementioned TFDs for IF estimation, strictly speaking, estimating the IF only makes sense for monocomponent or narrowband signals [18]. For that reason, estimating IF from the peaks of TFDs in multicomponent signals requires an extra step to extract and isolate different components before IF estimation methods can be applied to each component. For this purpose, a conventional approach consists of segmenting the TFD of a multicomponent signal using image processing techniques, including local peak detection and component linking [31] and blind source separation [32], among others.

The Hilbert-Huang transform (the HHT) [33,34] has been proposed as a new adaptive technique for the analysis of nonlinear and non-stationary signals. The technique consists of combining empirical mode decomposition (EMD) and Hilbert spectral analysis to obtain an alternative TFD of a signal, called the Hilbert spectrum (the HS), as a function of its IF and instantaneous amplitude (IA).

The HHT has some advantages over TFD-based IF estimation methods, which is why it was chosen for RS analysis in this paper. Since EMD is an adaptive and direct decomposition technique, it makes it possible to retrieve the modes of a multicomponent signal, called intrinsic mode functions (IMFs), without any a priori knowledge of the signal characteristics. In addition, HHT-based IF estimation is performed by means of differentiation; therefore, the HHT does not suffer from the uncertainty principle and simultaneously provides both high time and high frequency resolutions. Moreover, since IF and IA sequences are separately calculated for each component, we can work independently in either a time-frequency or a time-energy domain, without having to process an entire TFD. Furthermore, although the properties of the HHT have led to its application to a number of biomedical signals [35-38], it has rarely been used for RS analysis, as there are only a few studies, mainly focusing on DAS detection [39-41]. However, we found in our previous studies that the HHT also performed well in CAS detection [42,43], which inspired us to analyze its performance for CAS characterization in depth and explore its advantages over spectrogram, which has traditionally been the most commonly used technique for this purpose.

Another reason for which this study was carried out was that most proposed HHT-based methods for RS analysis [39-42] used the original EMD, which has a mode mixing (MM) effect. The MM effect consists either of an IMF containing components of widely different frequencies or of a signal component appearing in different IMFs [44]. Due to this MM, we found that EMD, when applied to some RS signals, resulted in poor separation of RS signal components [42]. Nevertheless, the original EMD has been used in other RS analysis approaches [45-48]. Among the proposed solutions for MM, the ensemble EMD (EEMD) [44,49] and the noise-assisted multivariate EMD (NA-MEMD) [50] are some of the most well-established and widely used methods, but they have rarely been applied to RS analysis [43,51]. Moreover, the implementation and performance of these methods depend on each application and a detailed analysis of the MM effect and the performance of EEMD and NA-MEMD in RS signals is lacking.

The aim of this study is to provide an in-depth evaluation of the performance of the HHT for RS analysis, which led us to calculate the HS with high resolution as an alternative representation to improve on the performance of spectrogram, especially for CAS characterization. The study is divided into two parts. First, we analyze the MM effect of EMD in recorded RS signals and evaluate the performance of EEMD and NA-MEMD

to reduce this effect using a number of quantitative parameters (section 4). Second, we evaluate the performance of three different IF estimators to obtain a suitable EEMD-Kay-based HS for CAS characterization (section 5), and we propose a new method for the automatic segmentation and characterization of CAS based on the HS processing (section 6). This algorithm was tested using a set of synthetic and recorded CAS signals, which allowed us to compare the performance of the HS and spectrogram.

2. Study dataset

2.1. Recorded RS signals

Recorded RS were obtained from the Pulmonary Function Testing Laboratory of Germans Trias i Pujol University Hospital in Badalona, Spain. All recordings were acquired from 30 patients with asthma. Four piezoelectric contact microphones (TSD108, Biopac Systems, Inc.) were placed on the surface of the patients' backs, on either side of the spinal cord, at the base and near the upper lobe of the right/left lung. All sensors were attached to the skin using adhesive rings. Airflow signals were recorded simultaneously with sound signals using a pneumotachograph (TSD107B, Biopac Systems, Inc.). All signals were sampled at 12,500 samples/s, using a 16-bit analogue-to-digital converter (MP150, Biopac Systems, Inc.). After digitalization, the sound signals were decimated by a factor of 4 to 3,125 samples/s and the respiratory phases were automatically detected using the airflow signal as the reference signal. After cycle segmentation, we selected RS from 636 inspiratory phases, with 353 normal sounds signals and 283 CAS signals, including both monophonic and polyphonic CAS.

2.2. Synthetic signals

In order to test the performance of different IF estimators and the proposed CAS characterization algorithm, we generated several synthetic CAS signals. Monophonic CAS signals with slightly variable IFs were modeled as sinusoid frequency modulated signals:

$$c_{1,k}(t) = sin[2\pi f_c(k)t + 0.6\,sin(2\pi 15t)] \quad (1)$$

$$f_{1,k}(t) = \frac{1}{2\pi}\frac{\partial \Phi_{1,k}}{\partial t} = \frac{1}{2\pi}\frac{\partial(2\pi f_c(k)t + 0.6\,sin(2\pi 15t))}{\partial t} = f_c(k) + 9\,cos(2\pi 15t) \quad (2)$$

$$t = 0\,..\,0.3\ s, k = 1\,..\,4, f_c = [80,150,250,550]$$

where $f_{1,k}$ are the IF sequences of signals $c_{1,k}(t)$ (Fig. 1-a). Monophonic CAS signals with frequency sweeping were modeled as linear frequency modulated signals:

$$c_{2,k}(t) = \sin[2\pi f_c(k)t + 2\pi\mu(k)t^2] \quad (3)$$

$$f_{2,k}(t) = \frac{1}{2\pi}\frac{\partial \Phi_{2,k}}{\partial t} = \frac{1}{2\pi}\frac{\partial(2\pi f_c(k)t + \pi\mu(k)t^2)}{\partial t} = f_c(k) + 2\mu(k)t \quad (4)$$

$$t = 0\,..\,0.25\ s, k = 1\,..\,4, f_c = [80,150,250,550],$$
$$\mu = [100,150,300,500]$$

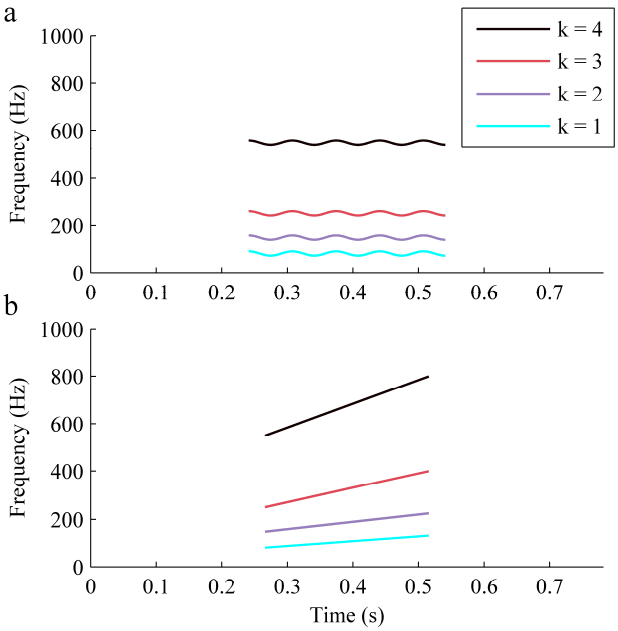


Fig. 1. **Theoretical IF of synthetic CAS.** IF laws of monophonic CAS signals with slightly variable IF, $c_{1,k}(t)$ (a), and monophonic CAS signals with frequency sweeping, $c_{2,k}(t)$ (b).

where $f_{2,k}$ are the IF sequences of signals $c_{2,k}(t)$ (Fig. 1-b).

Finally, polyphonic CAS signals were formed by combining $c_{1,k}(t)$ and $c_{2,k}(t)$ as follows:

$$c_{3,k}(t) = \begin{cases} c_{1,k}(t) & t \in [0,0.025] \\ c_{1,k}(t) + c_{2,k+1}(t) & t \in (0.025,0.275] \\ c_{1,k}(t) & t \in (0.275,0.3] \end{cases} \quad (5)$$

$$t = 0\,..\,0.3\ s, k = 1\,..\,3$$

We obtained 11 different synthetic CAS signals in total. Each synthetic CAS signal was added to a recorded RS signal containing normal RS at different signal-to-noise ratios (SNRs), thus simulating real CAS that superimposed on normal RS. Since normal RS usually have a sharp energy drop at about 200-250 Hz [1,2], synthetic CAS signals containing components below 200 Hz ($c_{1,k}(t)$, $c_{2,k}(t)$, and $c_{3,k}(t)$ for k=1,2), which overlap with normal RS, were added at SNRs from -4 dB to 12 dB, in increments of 2 dB. However, synthetic CAS signals containing components above 200 Hz ($c_{1,k}(t)$, $c_{2,k}(t)$, and $c_{3,k}(t)$ for k>2) were added at SNRs from -8 dB to 12 dB, in increments of 2 dB. As a result, a total of 109 synthetic CAS signals were obtained, including 80 monophonic and 29 polyphonic CAS signals.

3. Overview of the HHT

The HHT consists of two steps, EMD and the Hilbert transform. The central step of the HHT is EMD, which decomposes a multicomponent signal $s(t)$ into a set of zero mean narrowband components (IMFs), for which meaningful IF and IA can be calculated at any point by

means of the Hilbert transform. The main advantage of EMD is that it is a direct and adaptive decomposition technique, which extracts each IMF directly from the original signal by means of a sifting process [33]. As a result of this process, the signal $s(t)$ can be expressed as a linear combination of its components as follows:

$$s(t) = \sum_{i=1}^{n} IMF_i(t) + r_n(t) \quad (6)$$

where n is the number of extracted IMFs and $r_n(t)$ is the residue of $s(t)$. Having decomposed $s(t)$ by EMD, IF and IA can be calculated by the phase derivative and envelope of the analytic signal of each IMF and, therefore, $s(t)$ can be expressed as a function of its IF and IA as follows:

$$s(t) = \sum_{i=1}^{n} a_i(t) \cos\left(\int 2\pi f_i(t) dt\right) + r_n(t) \quad (7)$$

where $f_i(t)$ and $a_i(t)$ are the IF and IA of the *i-th* IMF, respectively. Building on expression (7), we can rearrange IF and IA in a three-dimensional TFD of the amplitude, the HS.

4. Evaluation of the EMD step of the HHT in RS signals

4.1. The MM effect of EMD in RS signals

Ideally, each IMF of a multicomponent signal would contain a few different frequency components of the signal. However, due to the MM effect of EMD, some components may appear within different IMFs, thus leading to some IMFs containing components of widely different frequencies.

Assessing the MM effect in multicomponent random signals, such as RS, is a complex task, since there is no a priori knowledge of the signal component characteristics. Nevertheless, this effect can be clearly observed in CAS signals because they are sinusoidal-like waveforms with well-defined fundamental frequencies. With the aim of illustrating the MM effect in RS signals, we applied EMD to a recorded polyphonic CAS signal with two overlapped components at around 140 Hz and 275 Hz. The resulting IMFs, which were obtained in decreasing order of frequency, are shown in Fig. 2.

The MM effect clearly occurs for both components. The highest frequency component appears in IMFs 1 and 2, whereas the lowest frequency component appears within IMFs 2 and 3. We calculated the power spectral density (PSD) of each IMF using Welch's periodogram with a Hanning window of 80 ms, 40 ms overlap, and 1,024 points for the fast Fourier transform. In order to make different PSDs comparable, we divided them by their respective maximum value (Fig. 3).

As shown in Fig. 3, MM is evident because IMF 2 includes two widely separated frequency components, which correspond to the CAS components. Moreover, the CAS component at around 140 Hz is included within both IMF 2 and IMF 3, whose PSDs overlap to a great extent. Due to this MM effect, the obtained IMFs do not ensure that the application of the Hilbert transform would yield physically meaningful IF estimates.

4.2. EEMD and NA-MEMD as solutions for mode mixing

Over the past few years, many studies have focused on solving the MM effect of EMD. Although several solutions have been proposed, EEMD and NA-MEMD are the most well established and widely used methods. These methods are examples of noise-assisted techniques, which use the benefits of noise in data analysis.

The MM effect occurs when some frequency scales are missing in the original signal. In this case, envelopes calculated during the sifting process are influenced by the extrema of widely different frequency components. However, when applied to white noise, which has scales uniformly distributed across the entire time-frequency plane, EMD acts as an adaptive dyadic filter bank [52,53]. Accordingly, when white noise is added to a multicomponent signal, all signal components with different frequencies are automatically separated by the reference scales set by white noise.

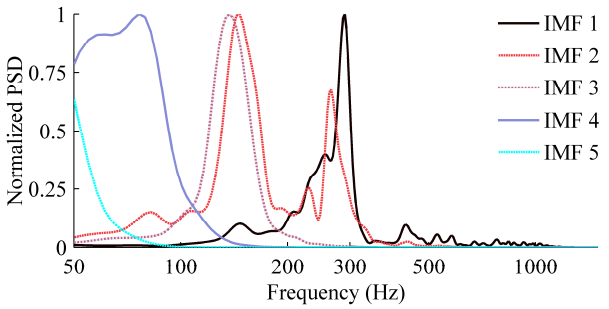


Fig. 3. The MM effect in the frequency domain. PSDs of IMFs shown in Fig. 2.

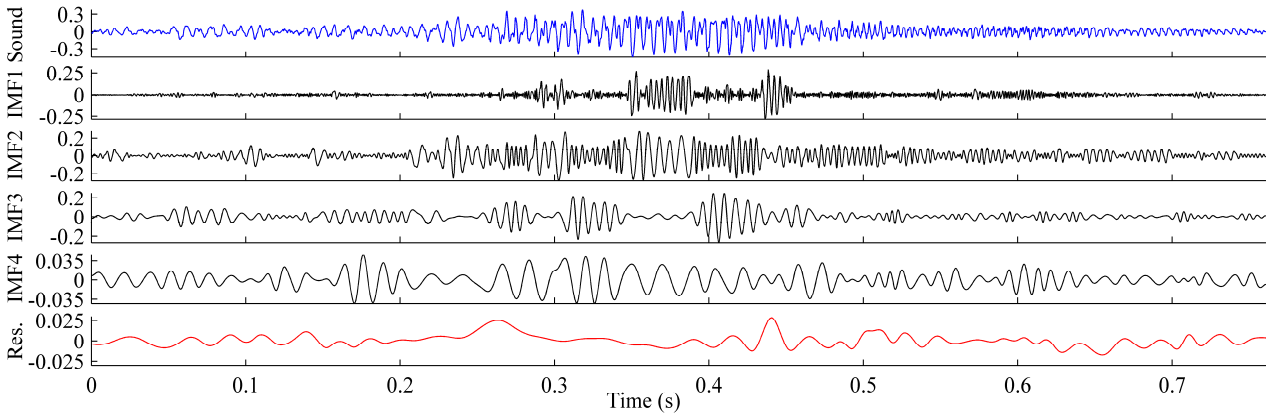


Fig. 2. The MM effect of EMD. IMFs obtained by means of EMD on an RS signal with polyphonic CAS.

Based on the aforementioned principle, EEMD was first proposed by Wu and Huang in 2009 [44]. The method consists of the iterative application of the original EMD to a signal plus multiple realizations of white noise. The final IMFs are calculated as the mean of those resulting from each iteration. Although the resulting IMFs contain a residual noise level, it can be almost totally cancelled using an appropriate number of iterations.

The MEMD method was later proposed by Rehman and Mandic [54], initially as an extension of EMD for multivariate signals. Like EMD, MEMD has a dyadic filter bank property on white noise [50]. Based on this property, NA-MEMD was proposed to avoid MM in multivariate signals. The idea behind NA-MEMD consists of adding extra channels containing different realizations of white noise to the original signal, and then decomposing the resulting multivariate signal by means of MEMD. This method can also be applied to univariate signals. In this case, only those IMFs obtained for the first channel (original signal) are retrieved.

To provide an example of the performance of noise-assisted techniques in avoiding the MM effect, we have applied the EEMD method to the polyphonic CAS signal shown in the previous section, 4.1, using 100 iterations and noise added at an SNR of 0 dB. The resulting IMFs are shown in Fig. 4.

As shown in Fig. 4, the EEMD method manages to reduce MM, as different frequency components are separated in different IMFs. This separation of frequency components is better observed in the PSDs of the resulting IMFs, shown in Fig. 5.

Comparing these PSDs with those shown in Fig. 3, it is clear that noise has forced the frequency components to be uniformly distributed along the whole frequency range, thus separating widely different frequency components into different IMFs.

Although both EEMD and NA-MEMD manage to reduce MM, there are some differences in the performance of the two methods, which are analyzed in depth in the following section.

4.3. Performance assessment of EMD, EEMD, and NA-MEMD in recorded RS signals

With the aim of comparing the performance of EMD, EEMD, and NA-MEMD in RS signal decomposition, we applied these methods to the 636 RS signals recorded as described in section 2.1, which included normal RS and CAS.

The choice of the EEMD and NA-MEMD parameters highly depends on the type of signal to be analyzed. Therefore, we followed some basic instructions, as described in [44] and [55], to choose the analysis parameters of each method. Input parameters for EEMD include the number of iterations and the SNR for the added noise. As explained in section 4.2, the residual noise level (n_{res}) of the obtained IMFs can be reduced by increasing the number of iterations. Usually, a few hundred iterations are enough to significantly reduce n_{res}. In fact, n_{res} decreases following the rule $n_{res} = n/\sqrt{N}$, where n is the amplitude of the added noise and N is the number of iterations [44]. Based on this rule, we decided to use square numbers for the number of iterations and a wide range of SNRs.

With regard to NA-MEMD, input parameters include the number of noisy extra channels, the amplitude of the added noise, and the number of directions used in the MEMD process. At least two noisy extra channels should be used and, as a rule of thumb, the minimum number of directions should be twice the number of data channels [55]. Therefore, for EEMD and NA-MEMD, all possible combinations of the following parameters were tested for each RS signal:

- EEMD → number of iterations: 1, 2, 4, 16, 25, 36, 64, 100, 225, and 400; SNRs: -9, -6, -3, 0, 3, 6, 9, 15, and 21.
- NA-MEMD → number of extra channels: 2, 3, and 4; number of directions: 8, 16, 32, and 64; SNRs: -9, -6, -3, 0, 3, 6, 9, 15, and 21.

We programmed the EEMD algorithm using the original EMD algorithm reported by Rilling and Flandrin [56,57].

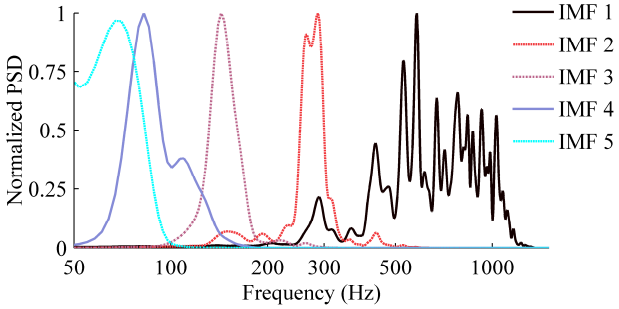


Fig. 5. **Separation of frequency components by EEMD.** PSDs of IMFs shown in Fig. 4.

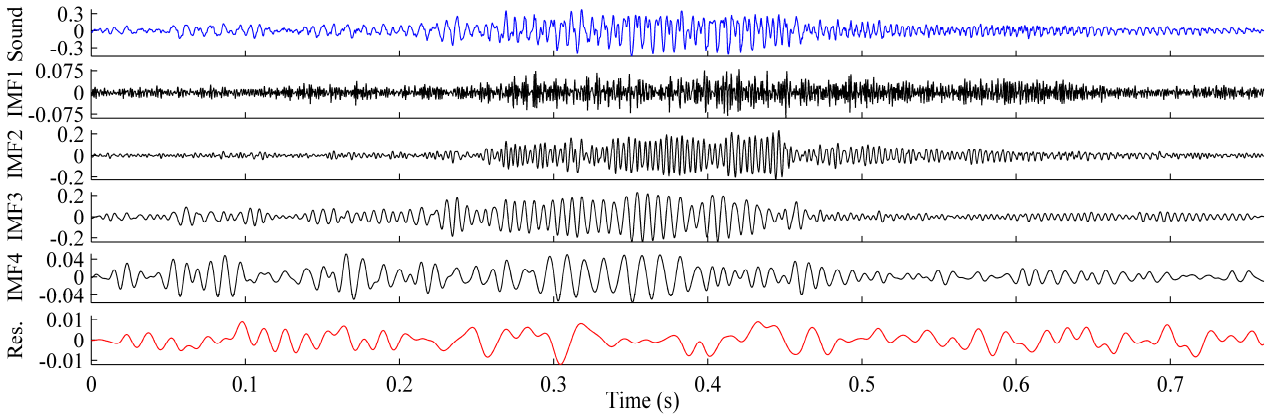


Fig. 4. **Avoiding MM by EEMD.** IMFs obtained by EEMD of the recorded polyphonic CAS signal shown in Fig. 2.

For the NA-MEMD method, we used the Matlab code provided by Mandic [55], which applies the MEMD algorithm reported by Rehman and Mandic [54].

The results were evaluated by means of six quantitative parameters that allowed us to choose the most suitable method for RS signal decomposition. The proposed parameters were divided into three groups depending on the measured feature.

4.3.1. Size and processing time

The first parameter used to compare the performance of EMD, EEMD and NA-MEMD was the total number of IMFs (N_{IMF}) resulting from the decomposition of each RS signal. Moreover, since the frequency range of interest for RS analysis goes from 70 Hz onwards, we also calculated the number of IMFs whose central frequency (f_c), measured from the PSD, was greater than 70 Hz ($N_{IMF\text{-}FR}$).

The results shown in Fig. 6 indicate that the EEMD method (solid lines in Fig. 6-a) provides lower IMFs (about 8-10 IMFs) than the NA-MEMD method, which produces between 14 and 16 IMFs (solid lines in Fig. 6-b). Similar to EEMD, the mean N_{IMF} of the original EMD for all RS signals was 9.1 IMFs.

Figure 6 also shows that the mean $N_{IMF\text{-}FR}$ for EEMD was around 5 IMFs at most (dotted lines in Fig. 6-a), which means that EEMD produces between 3 and 5 IMFs (difference between N_{IMF} and $N_{IMF\text{-}FR}$) outside the frequency range of interest for RS analysis. Nevertheless, the mean $N_{IMF\text{-}FR}$ for NA-MEMD was around 7 IMFs (dotted lines in Fig. 6-b), which indicates that this method produces more redundant IMFs (around 7-9 IMFs) at low frequencies, which are irrelevant for RS analysis.

As for the high frequencies, although CAS may appear at up to 1,000 Hz, the frequency range of normal RS barely exceeds 250 Hz [2], so having many IMFs covering high frequencies generates redundancy. In this sense, we analyzed f_c of the first four IMFs generated by EEMD and NA-MEMD along all RS signals (Fig. 7). As shown, NA-MEMD usually produces 4 IMFs above 250 Hz, whereas EEMD only produces 2 IMFs in the same frequency range. Therefore, NA-MEMD also produces more redundant IMFs at high frequencies than EEMD.

In addition to N_{IMF} and $N_{IMF\text{-}FR}$, we also calculated the mean decomposition computing time (D_{CT}) along all RS signals for each method. In other words, we looked at how long it takes for each method to decompose an RS signal into IMFs. All simulations were run in a server with Windows Server 2008 R2 Enterprise installed, an Intel® Xeon® processor E7340 at 2.40 GHz with 4 kernels, and 88 GB of usable RAM. Results shown in Fig. 8 indicate that EEMD is much faster than NA-MEMD.

Despite the fact that D_{CT} increases exponentially with the number of iterations in EEMD, a few seconds is long enough to decompose RS signals using a few hundred iterations, which are sufficient to reduce MM, as explained in the next section, 4.3.2. However, NA-MEMD is a time-consuming method. While the amount of MM decreases with an increase in number of directions (see section 4.3.2), D_{CT} increases in the same way. Therefore, D_{CT} required to substantially reduce the MM effect is too high in comparison with EEMD.

4.3.2. Reduction of MM

In section 4.1 we showed that MM causes frequency overlap between PSDs of different IMFs. Based on this fact, we propose the following parameter to measure the amount of MM, based on frequency overlap (FO) between pairs of IMFs:

$$FO_{i,j}(\%) = 100\frac{max([f_{c2i}, f_{c8i}] \cap [f_{c2j}, f_{c8j}]) - min([f_{c2i}, f_{c8i}] \cap [f_{c2j}, f_{c8j}])}{min\{IDR60_i, IDR60_j\}} \quad (8)$$

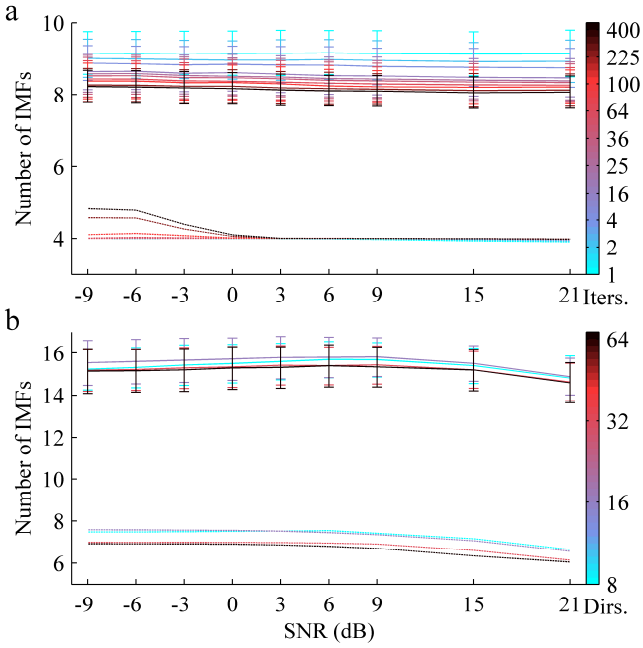


Fig. 6. **Size of EEMD and NA-MEMD.** N_{IMF} (solid lines) and $N_{IMF\text{-}FR}$ (dotted lines) for EEMD (a) and NA-MEMD (b). All values are the mean and standard deviation along all RS signals. For NA-MEMD (b), all values are also averaged along the number of extra channels.

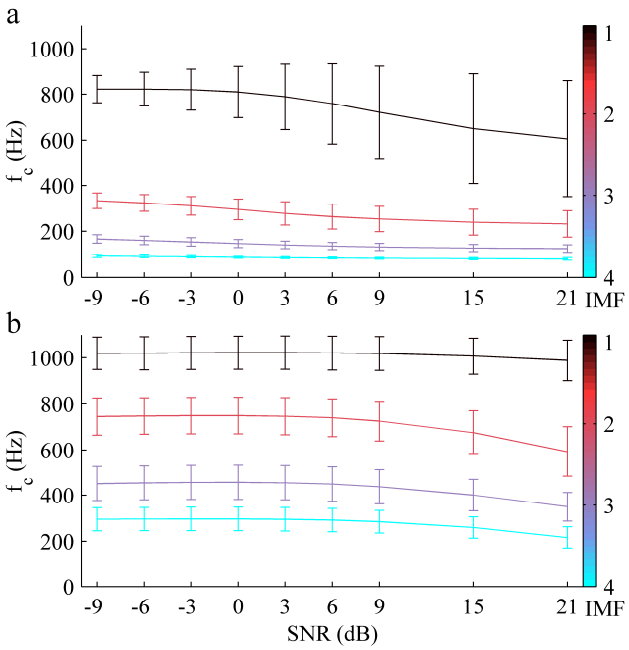


Fig. 7. **Tracking of high frequencies by EEMD and NA-MEMD.** Central frequency (f_c) of IMFs 1-4 for EEMD (a) and NA-MEMD (b). All values are the mean and standard deviation along all RS signals. For EEMD (a), all values are also averaged along the number of iterations. For NA-MEMD (b), all values are also averaged along the number of extra channels and the number of directions.

where f_{c2i}, f_{c8i}, and $IDR60_i$ are frequency parameters measured from the PSD of the *i-th* IMF. Specifically, f_{c2} and f_{c8} are the frequencies at which 20% and 80% of the energy of an IMF are reached, respectively. The *IDR60* parameter is calculated as the difference between f_{c8} and f_{c2}.

We calculated the mean $FO_{i,j}$ along all RS signals, for EMD, EEMD, and NA-MEMD, and for different pairs of IMFs. The mean $FO_{i,j}$ for the original EMD was 22.5% between IMFs 1-2, 22.0% between IMFs 2-3, and 21.3% between IMFs 3-4. The results for EEMD are shown in Fig. 9, which illustrates that $FO_{i,j}$, and hence the amount of MM, mainly depends on the SNR when the number of iterations is greater than or equal to 16.

Since $FO_{i,j}$ decreases as the SNR decreases, it would seem that using the lowest SNR is the best solution for reducing MM. However, as explained in the next section, 4.3.3, using very low SNRs increases the residual noise level.

Figure 10 shows the results for NA-MEMD. Since this method produces more IMFs than the EEMD method, we calculated $FO_{i,j}$ between pairs of IMFs from IMF 1 to IMF 5.

As the figure shows, $FO_{i,j}$ depends on both the SNR and the number of directions. In general, the amount of MM decreases with a decrease in SNR and an increase in the number of directions (Figs. 10-b,c,d). However, $FO_{1,2}$ increases with a decrease in SNR, especially for a low number of directions (Fig. 10-a). As explained in section 4.3.1, NA-MEMD tends to generate too many IMFs at high frequencies. If the SNR is too low, noise components cause IMFs to be uniformly distributed along all frequencies, which forces IMFs covering high frequencies to be cramped, and $FO_{1,2}$ increases. On the contrary, if the SNR is high, the effect of noise is negligible, which allows IMFs 1 and 2 to be more widely separated. In any case, NA-MEMD needs a high number of directions to achieve results similar to those of EEMD, and this greatly increases the D_{CT} (see section 4.3.1).

4.3.3. Residual noise level

A major challenge when working with EEMD and NA-MEMD is minimizing the residual noise level in the resulting IMFs. In order to quantify this residual noise level, we propose the following parameters:

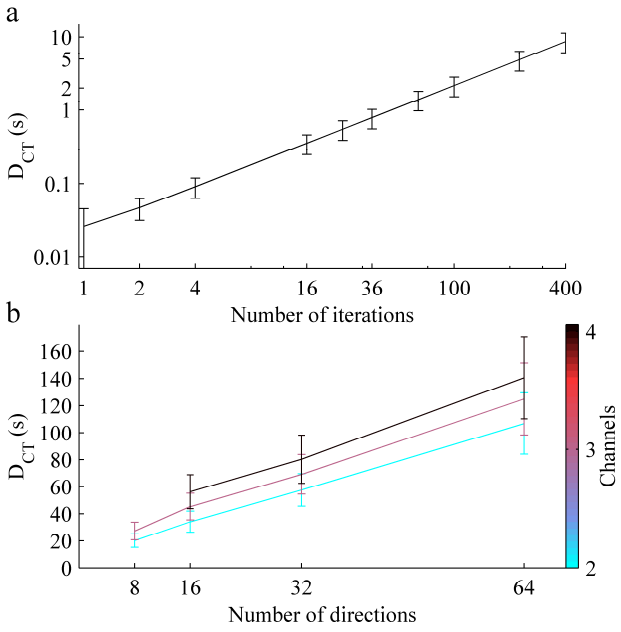


Fig. 8. **Processing time, D_{CT}, for EEMD (a) and NA-MEMD (b).** All values are the mean and standard deviation along all RS signals and SNRs.

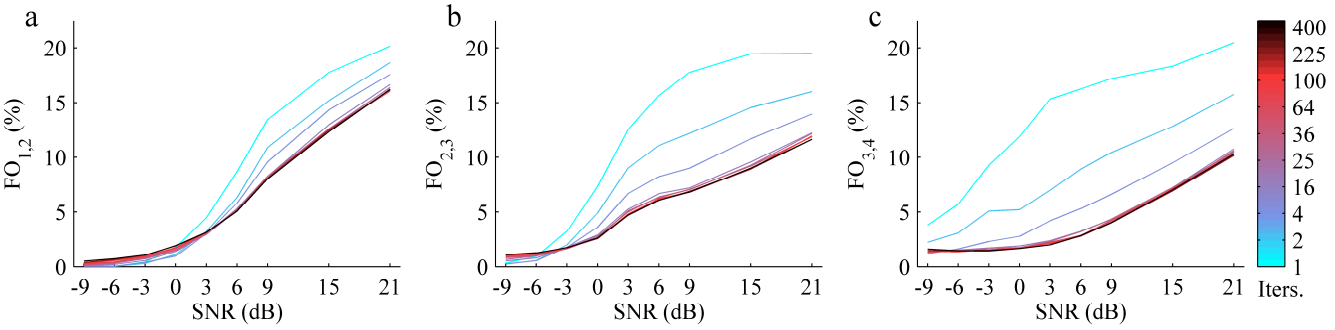


Fig. 9. **MM reduction by EEMD.** Frequency overlap (*FO*) between IMFs 1-2 (a), IMFs 2-3 (b), and IMFs 3-4 (c) for the EEMD method. All values are the mean along all RS signals.

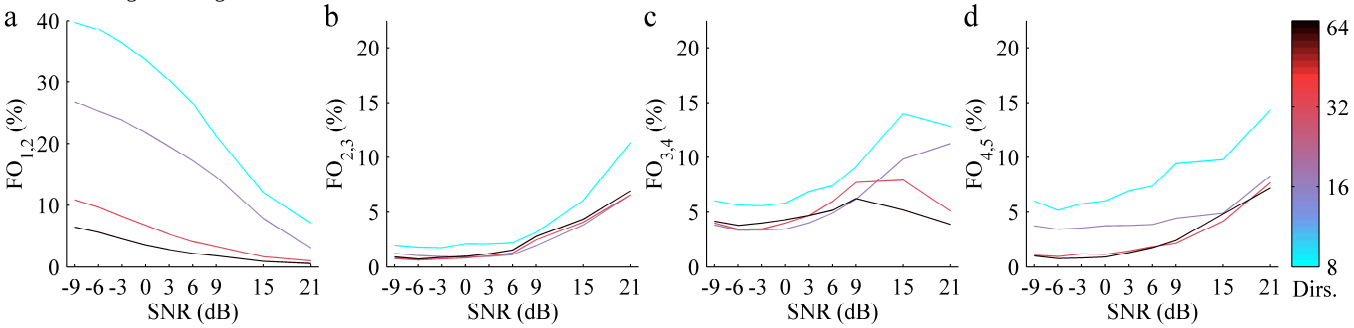


Fig. 10. **MM reduction by NA-MEMD.** Frequency overlap (*FO*) between IMFs 1-2 (a), IMFs 2-3 (b), IMFs 3-4 (c), and IMFs 4-5 (d) for the NA-MEMD method. All values are the mean along all RS signals and the number of extra channels.

- *CC*: cross-correlation at zero lag between the PSD of the original signal (PSD_o) and the PSD of the reconstructed signal (PSD_{rec}).

$$CC = \hat{R}_{PSD_o PSD_{rec}}(0) = \frac{\sum_n PSD_o(n) PSD_{rec}(n)}{\sqrt{\sum_n PSD_o(n)^2 \sum_n PSD_{rec}(n)^2}} \quad (9)$$

- *PSDR*: ratio of the absolute error between PSD_{rec} and PSD_o versus PSD_o.

$$PSDR(\%) = 100 \frac{\sum_n |PSD_{rec}(n) - PSD_o(n)|}{\sum_n PSD_o(n)} \quad (10)$$

where n is the number of points used for the fast Fourier transform. The reconstructed signals were calculated as the direct sum of the corresponding IMFs and residues.

Ideally, *PSDR* and *CC* would be 0 and 1, respectively, if the reconstructed signal were exactly equal to the original signal. This is the case of EMD and NA-MEMD, which provide a perfect reconstruction of the original signal. However, EEMD causes a slight error in the reconstructed signal due to the use of white Gaussian noise in the decomposition process.

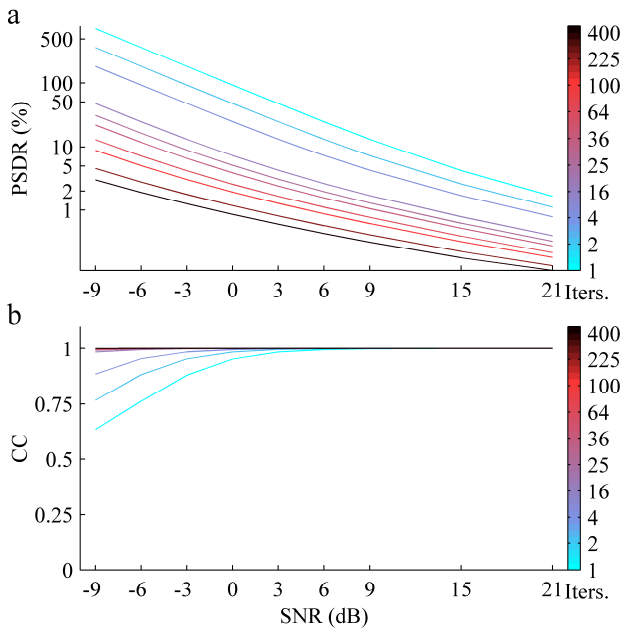


Fig. 11. **Measures of residual noise level.** *PSDR* (a) and *CC* (b) for EEMD, as a function of SNR and number of iterations. All values are the mean along all RS signals.

This error depends on both the amplitude of the white noise and the number of iterations, as shown in Fig. 11.

As shown, *CC* reaches its maximum using 16 iterations or more, independently of the SNR. However, *PSDR* highly depends on the two parameters. Nevertheless, by applying a few hundred iterations, we obtained an acceptable *PSDR* (below 3%) for a wide range of SNRs. Therefore, we can assume that the residual noise level is not a major drawback of EEMD in RS signals provided that the SNR and number of iterations are correctly chosen.

4.4. Selection of parameters in EEMD

Based on the previous results, we decided that EEMD was better than NA-MEMD for RS signal decomposition, since EEMD produced fewer redundant IMFs, managed to reduce the MM effect to a greater extent, and was faster than NA-MEMD. However, in order for EEMD to perform at its fullest potential, we had to fix its parameters (SNR and number of iterations) so that MM was reduced as much as possible and the residual noise level was not significant. To this end, we analyzed the $FO_{i,j}$ and the *PSDR* parameters together for different pairs of IMFs, as shown in Fig. 12.

Values of the intersection points between both parameters for each number of iterations are shown in Table 1.

Assuming 3% as an acceptable upper limit for $FO_{i,j}$ and *PSDR*, an SNR below 1 dB and more than 64 iterations should be used for EEMD. We decided to fix the SNR at 0 dB and use 100 iterations to decompose RS signals by means of EEMD. The $FO_{i,j}$ did not decrease significantly, neither by using more iterations nor by decreasing the SNR. However, increasing the number of iterations or decreasing the SNR greatly increased the D_{CT} and the *PSDR*, respectively.

5. IF-IA estimation and the HS in RS signals

Having decomposed an RS signal into IMFs, the next step in calculating the HS is IF-IA estimation. The use of HHT involves estimating IF as the phase derivative of the analytic signal of each IMF. It is the most intuitive and direct way to define IF of a real signal [17]. For the case of a length-M IMF, the analytic signal, $z_i(n)$, is defined as follows:

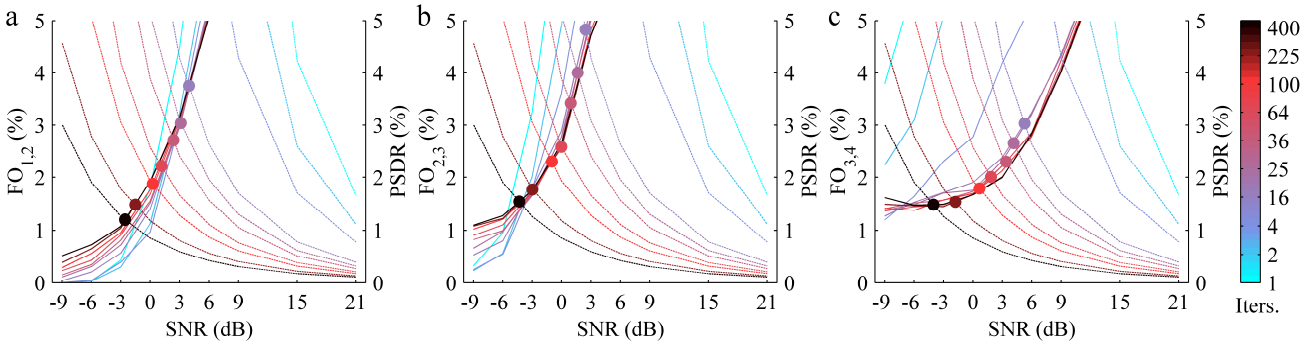


Fig. 12. **Performance parameters of EEMD.** Combination of measures of the amount of MM (*FO*) (solid lines) and residual noise level (*PSDR*) (dotted lines) for IMFs 1-2 (a), 2-3 (b), and 3-4 (c) obtained by EEMD. All values are the mean along all RS signals.

Table 1. Performance of EEMD for RS signal decomposition as a function of the number of iterations and SNR.

	Iterations	1	2	4	16	25	36	64	100	225	400
IMFs 1-2	SNR (dB)	8.7	7.7	6.3	3.6	2.8	2.1	1.0	0.1	-1.7	-2.7
	$FO_{1,2}$/*PSDR* (%)	14.1	9.8	7.0	3.9	3.2	2.9	2.3	1.9	1.5	1.2
IMFs 2-3	SNR (dB)	8.0	6.7	5.4	2.2	1.3	0.6	-0.5	-1.6	-3.7	-5.2
	$FO_{2,3}$/*PSDR* (%)	17.1	11.7	8.4	5.1	4.2	3.6	2.8	2.5	2.0	1.7
IMFs 3-4	SNR (dB)	7.9	7.5	6.9	4.7	3.5	2.7	1.1	-0.3	-2.8	-5.0
	$FO_{3,4}$/*PSDR* (%)	17.4	10.2	6.4	3.3	2.9	2.5	2.2	2.0	1.7	1.7

FO: frequency overlap parameter; *PSDR*: residual noise level parameter.

$$z_i(n) = IMF_i(n) + jH[IMF_i(n)] = a_i(n)exp[j\Phi_i(n)], n = 1..M \quad (11)$$

where $H[]$ is the Hilbert transform, $a_i(n)$ is the IA, and $\Phi_i(n)$ is the phase of $z_i(n)$. Having calculated $\Phi_i(n)$ for each *i-th* IMF, the next issue is how to address the phase derivative in discrete time. The most common approach is to use finite impulse response derivative filters. This is the case of our first IF estimation method, which is a five-point least squares polynomial derivative (LSPD) approximation [58]:

$$f_i(n) = \frac{f_m}{2\pi}\sum_{k=0}^{4} b_k \Phi_i(n-k), b = [b_0, b_1, b_2, b_3, b_4] = \frac{1}{10}[2, 1, 0, -1, -2], n = 1..M \quad (12)$$

where f_m is the sample frequency. A major problem of this estimator is that it has very high variance. However, low variance estimators are preferable for calculating the HS, since it ought to be an accurate time-frequency representation in which the signal energy is as concentrated as possible around the IF.

In order to reduce variance, Kay proposed a weighted phase difference estimator [17]. This method consists of calculating the IF estimate by a weighted averaging of a sequence of phase difference measurements, as follows:

$$f_i\left(n + \left\lfloor\frac{N}{2}\right\rfloor\right) = \frac{f_m}{2\pi}\sum_{k=0}^{N-2} w(k)[\Phi_i(n+k+2) - \Phi_i(n+k+1)], n = 1..M-N \quad (13)$$

$$w(k) = \frac{\frac{3}{2}N}{N^2-1}\left(1 - \left(\frac{k-\left(\frac{N}{2}-1\right)}{N/2}\right)^2\right) \quad (14)$$

where $w(k)$ is the length-N averaging window. The larger the window size, the smaller the variance will be. After testing different window sizes, we propose an averaging window of 32 samples.

Together with the aforementioned methods, in this study we also tested an alternative approach for IF estimation based on the Teager energy operator (TEO) [59]. This method has very low computational complexity and is very straightforward, as IF and IA are directly calculated from the IMF signal as follows:

$$\psi[IMF(n)] = IMF^2(n) - IMF(n-1)IMF(n+1) \quad (15)$$

$$f_i(n) \approx \frac{f_m}{2\pi}\sqrt{\frac{\psi[IMF_i(n+2) - IMF_i(n+1)]}{\psi[IMF_i(n+2)]}}, n = 1..M-3 \quad (16)$$

$$|a_i(n)| \approx \frac{\psi[IMF_i(n+2)]}{\sqrt{\psi[IMF_i(n+2) - IMF_i(n+1)]}}, n = 1..M-3 \quad (17)$$

where $\psi[]$ is the TEO.

After calculating IF and IA for each IMF, the HS can be directly obtained by constructing a two-dimensional array with the accumulation of all of the values of the IA sequences at the positions determined by the corresponding IF values and time instants. Since time instants can be determined within the resolution of the sampling period, and IFs can be precise at any number below the Nyquist frequency, the HS can have high time and high frequency resolutions. Both resolutions depend on the bin size selected for each dimension. In this study, we defined time and frequency bin sizes of $1/f_m$ seconds (0.32 ms for f_m = 3125 Hz) and 0.5 Hz, respectively. The resulting HS consisted of a matrix of 3125 rows and M columns, where M was the number of samples of the analyzed RS signal. For display reasons, a smoothing filter is usually applied to the HS. In this study, we used a 20-sample Gaussian filter with a standard deviation of five samples.

Figure 13 shows the spectrogram and three HSs, whose IFs were calculated by means of the aforementioned IF estimators, of a synthetic polyphonic CAS, $c_{3,3}(t)$, added to a normal RS signal at an SNR of 6 dB (section 2.2). The spectrogram was calculated using a 250-sample length Hanning window, with 240 overlapping samples, and 2,048 points for the fast Fourier transform.

It is clearly noticeable that the five-point LSPD approximation (Fig. 13-c) and the TEO-based method (Fig. 13-d) have very high variance. However, Kay's method (Fig. 13-b) greatly reduces variance and provides an accurate HS in which the energy of the signal is highly concentrated around its IF. Furthermore, this EEMD-Kay-based HS has higher energy concentration and resolution than spectrogram. The performance of these two representations for CAS characterization is evaluated in the next section using a larger dataset.

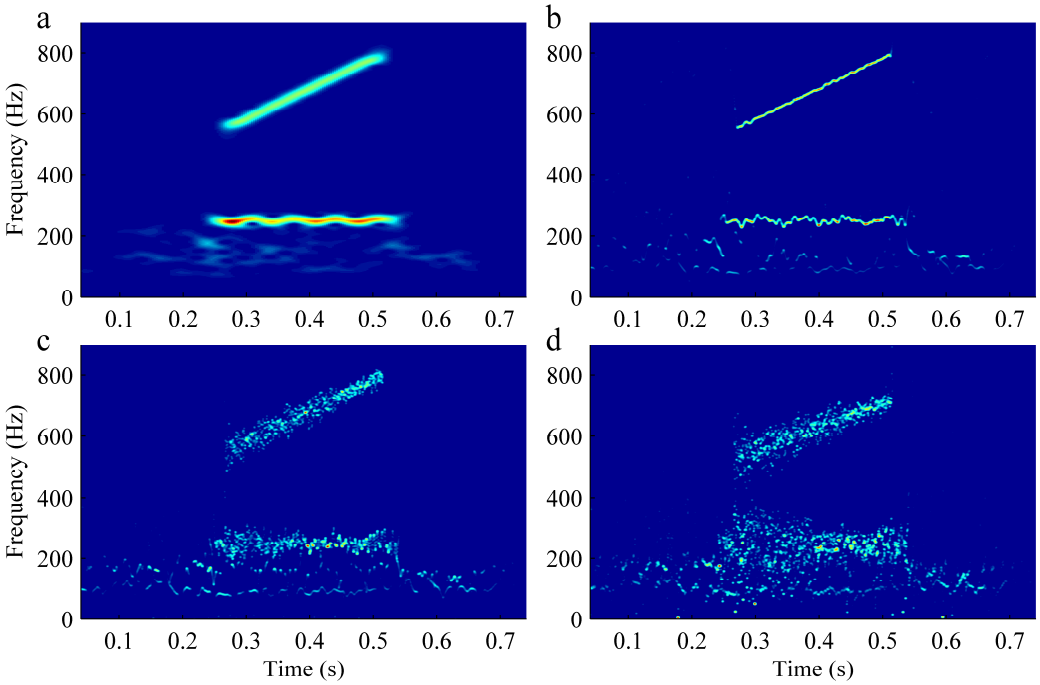


Fig. 13. **Energy concentration of different TFDs.** Spectrogram (a) and the HSs with IFs calculated using Kay's method (b), five-point LSPD approximation (c), and TEO-based method (d) of a polyphonic CAS.

6. Performance assessment of the HS for CAS characterization

6.1. Processing of the HS for CAS segmentation

In this section, we propose a new method for the automatic segmentation and characterization of CAS based on HS processing. The proposed algorithm is based on the region growing methodology and consists of three parts: detection of analysis areas in the HS, selection of seed points for local region growing, and region linking.

In light of the results from the previous sections 4 and 5, we decided to calculate the HS using EEMD with 100 iterations and 0 dB for the SNR, Kay's method with a 32-sample length window for IF estimation, time and frequency bin sizes of 0.32 ms and 0.5 Hz, respectively, and a 20-sample Gaussian smoothing filter with a standard deviation of five samples.

6.1.1. Detection of analysis areas in the HS

The first part of the CAS segmentation algorithm consisted of the CAS detection algorithm proposed in our previous study [43]. That algorithm detected the segments within an RS signal that were more likely to contain CAS based on the hypothesis that the IF dispersion markedly decreases when CAS appear in an RS signal. For that purpose, IF and IA sequences were calculated from the IMFs of an RS signal that had been previously decomposed by EEMD. In that previous study, IF was calculated using the five-point LSPD approximation, which had high variance and emphasized the differences in IF dispersion between those segments of an RS signal containing CAS and those containing normal RS. So, since IF and IA sequences allowed us to work independently in either a time-frequency or a time-energy domain, we were able to use simple dispersion-based criteria on IF sequences to delimit RS signal segments with a lower IF dispersion. Each delimited RS signal segment was characterized by means of a specific set of features extracted from the IF and IA, including the mean and standard deviation IF, among other features. Those features were used to classify each delimited RS signal segment as containing CAS or normal RS using a support vector machine classifier.

In the present study, we used the outputs from the aforementioned algorithm (classification, mean IF, and standard deviation IF) to mark out an analysis area in the HS for each delimited segment of an RS signal. Each analysis area was centered on the mean IF, had a frequency width of twice the standard deviation IF, and was delimited by the first and the last time instant of the delimited RS signal segment. Figure 14 shows some examples of analysis areas detected from two different synthetic CAS signals. Analysis areas with green edges (CAS areas) correspond to segments classified as containing CAS, whereas analysis areas with red edges (normal areas) correspond to segments classified as containing normal RS.

Although some areas of an RS signal may be misclassified, CAS areas are more likely to contain either an entire or a part of a CAS component than normal areas. Moreover, we considered CAS areas not overlapping with normal RS to be more likely to actually contain CAS components than CAS areas overlapping with normal RS. According to these criteria, three types of analysis areas were defined:

- High-pitched CAS areas: CAS areas whose mean IF was above 250 Hz.
- Low-pitched CAS areas: CAS areas whose mean IF was below 250 Hz.
- Normal areas

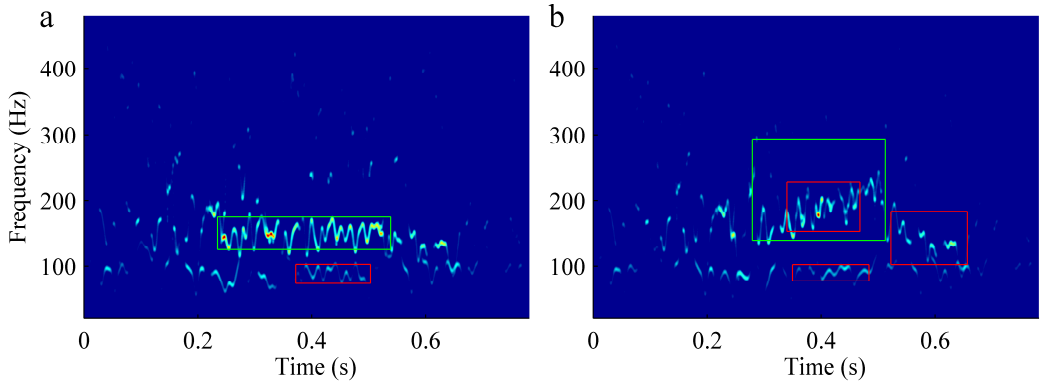


Fig. 14. **Detection of analysis areas in the HS.** Detected analysis areas for synthetic CAS signals $c_{1,2}(t)$ (a) and $c_{2,2}(t)$ (b) added to a normal RS signal at an SNR of -2 dB and -4 dB, respectively, as described in section 2.2.

This classification of analysis areas was used in section 6.1.3 for applying different thresholds when linking regions detected from each type of area. Previously, all areas were considered for seed point searching and local region growing, as explained in the next section.

6.1.2. Seed point search and local region growing

Ideally, CAS components are represented in the HS as ridges describing the IF where signal energy concentrates (see Fig. 13-b and Fig. 14). These ridges are composed of several linked regions, which, in turn, are formed by a set of connected points. Therefore, having detected an analysis area in the HS of an RS signal, the next step was to detect regions with a high concentration of energy around this area.

A determining parameter in this CAS segmentation algorithm was the point amplitude, which was associated with the signal energy at a certain IF. Since different RS signals had different signal energies, we first normalized the HS by dividing it by its maximum. Then, we determined an amplitude threshold to reject those points corresponding to background noise or having very low amplitude. After analyzing histograms from several HSs, we decided to consider only those points with an amplitude exceeding 0.05.

A region growing algorithm was applied to each detected analysis area, in which regions were grown from seed points by adding neighboring points that met a particular inclusion criterion. The first seed point was the point with the highest amplitude inside the analysis area. The first region was then grown by adding neighboring points that met the following criterion:

$$\left|f_{np} - \bar{f}_{CAS}\right| \leq \beta_1 \sigma_{f,area}$$

where f_{np} was the frequency of the neighboring point, $\bar{f}_{CAS}$ was the mean frequency along all points already included in the first region, $\sigma_{f,area}$ was the standard deviation frequency along all points of the HS within the analysis area, and β_1 was a scale factor, which was empirically set to 3. When the growth of the first region stopped, another seed point was sought and a new region was grown. Each new seed point was the point with the highest amplitude not yet belonging to any region and sought between the following frequency boundaries:

$$\max\{\bar{f}_{CAS} - \beta_2 \sigma_{f,CAS}, f_{min,area}\} \leq f_{np} \leq \min\{\bar{f}_{CAS} + \beta_2 \sigma_{f,CAS}, f_{max,area}\}$$

where $\sigma_{f,CAS}$ was the standard deviation frequency along all points already included in a region, $f_{min,area}$ and $f_{max,area}$ were the minimum and the maximum frequencies of the analysis area, respectively, and β_2 was a scale factor. This factor was first set to 3 when searching for new seed points between the temporal boundaries of the analysis area. When no new seed points were found within these temporal boundaries, β_2 was set to 2 to search for new seed points outside the temporal boundaries of the analysis area.

Each new region was grown by adding neighboring points that met either the following criterion:

$$\left|f_{np} - \bar{f}_{CAS}\right| \leq \beta_1 \sigma_{f,CAS}$$

when the analysis area corresponded to a segment extracted from either IMF 1 or IMF 2, or the following criterion:

$$\left|f_{np} - \bar{f}_{CAS}\right| \leq \max\{\beta_1 \sigma_{f,CAS}, \beta_1 \sigma_{f,area}\}$$

when the analysis area corresponded to a segment extracted from either IMF 3 or IMF 4. The region growing process was continued until no new seed points were found.

6.1.3. Region linking

The last step in the CAS segmentation algorithm was to retain only those regions that guaranteed the temporal and frequency continuity of the CAS component. We first rejected any CAS component not containing at least one region longer than 20 ms for high-pitched CAS areas, 50 ms for low-pitched CAS areas, or 80 ms for normal areas. In this way, we prevented the detection of false CAS components that might result from linking many short regions corresponding to background noise. Then, assuming that the longest region inside the analysis area truly belonged to the CAS component, adjacent regions were progressively checked from the nearest to the farthest in both directions along the time axis. Three parameters were calculated for each *i-th* region to measure its proximity to regions already retained as part of the CAS component:

- $t_{i\text{-}cas}$ and $f_{i\text{-}cas}$: temporal and frequency distances between the two nearest points among those of the *i-th* region and those of all regions already retained.
- $\Delta fmean_{i\text{-}cas}$: difference between the mean frequency of the *i-th* region and the mean frequency of the nearest 20-millisecond length segment along the regions already retained.

We considered the nearest regions to be more likely to belong to the CAS component than the farthest regions. Accordingly, we defined three ranges for $t_{i\text{-}cas}$ whose boundaries were determined by thresholds *th1*, *th2*, and *th3* as follows:

$$\text{Range j: } tht_{j-1} < t_{i-cas} \leq tht_j \text{ , } j = 1..3, tht0 = 0$$

Regions belonging to ranges 3 and 1 had the most and the least restrictive conditions, respectively, for retention. The *i-th* region belonging to the *j-th* range was retained if it met one of the following criteria:

$$d_i \geq thdur_j \;\&\; \Delta fmean_{i-cas} \leq thf \;\&\; f_{i-cas} \leq thf \quad (18)$$

$$t_{i-cas} \leq 100\ ms \;\&\; d_i \geq 125\ ms \quad (19)$$

where threshold $thdur_j$ was the required minimum length of the *i-th* region (d_i) belonging to the *j-th* range, and threshold *thf* was the limit for the frequency parameters $f_{i\text{-}cas}$ and $\Delta fmean_{i\text{-}cas}$ and guaranteed the frequency continuity of the CAS component. We made all the thresholds dependent on the type of analysis area (see section 6.1.1) and whether the *i-th* region was inside or outside the analysis area. In this study, thresholds *tht1*, *tht2*, *tht3*, *thdur1*, *thdur2*, *thdur3*, and *thf* were empirically set to the values shown in Table 2. These values were fixed after analyzing many HSs from recorded RS signals.

All regions not belonging to some range and not meeting any of the criteria defined in (18) and (19) were rejected. After checking all of the regions, the regions retained formed the segmented CAS component. Finally, according to the definition of CAS [1,2], we rejected any CAS component shorter than 100 ms.

Each segmented CAS component was characterized by means of the most relevant parameters from a clinical point of view:

- Duration (D): difference between the last and the first point along the time axis.
  Weighted mean frequency (Fmean) of all points, whose amplitudes in the HS were the weights.

6.2. Characterization of simulated CAS signals

In this section, we applied the proposed CAS segmentation algorithm to the HS and spectrogram of the 109 synthetic CAS signals described in section 2.2. Since our algorithm was designed to be applied to the proposed EEMD-Kay-based HS, we had to adjust some parameters for its use with spectrogram, which was calculated using a 250-sample length Hanning window, with 240 overlapping samples, and 2,048 points for the fast Fourier transform. Specifically, we increased the amplitude threshold from 0.05 to 0.1 and β_1 (see section 6.1.2) from 3 to 4.

Together with *D* and F_{mean}, we calculated the following two parameters for each segmented CAS component as a means of measuring the concentration of both TFDs:

- σ_F: average value of the point by point weighted standard deviation frequency.
- $\sigma_{F\text{-}FI}$: average value of the point by point weighted standard deviation frequency in relation to real IF values, which were defined in (2) and (4) (section 2.2).

These parameters allowed us to compare the performance of the proposed HS and spectrogram for RS analysis, at different SNRs. Figure 15 shows the absolute value of the differences between *D* (Err_D) and F_{mean} (Err_{Fmean}) calculated using the CAS segmentation algorithm and their real values, and the concentration measures (σ_F and $\sigma_{F\text{-}FI}$) for the HS and spectrogram. Statistical differences between these parameters of the two TFDs were evaluated using a one-sided Wilcoxon signed rank test at the 5% significance level.

As shown in Fig. 15-a, there were two SNR ranges in which Err_D was similar for both TFDs. One of those SNR ranges included synthetic CAS signals added to a normal RS signal at SNRs greater than or equal to 0 dB, which represented situations where, due to their high amplitude, CAS components could be easily detected by the two methods. Similarly, the performance of both TFDs was also similar within the SNR range from -8 dB to -6 dB, which included only synthetic CAS signals containing components above 200 Hz (see section 2.2). Despite the low SNR, those CAS components did not overlap with normal RS and, therefore, could be segmented more easily by both TFDs. The mean Err_D in all of those cases was 18.2 ± 31.6 ms for the HS and 19 ± 21.6 ms for the spectrogram. Compared to the duration of synthetic CAS signals (either 250 ms or 300 ms), both mean Err_D were acceptable, as they represented between 6% and 8% of synthetic CAS duration.

Table 2. Thresholds for region linking

Type of analysis area	*tht1* (ms)	*thdur1* (ms)	*tht2* (ms)	*thdur2* (ms)	*tht3* (ms)	*thdur3* (ms)	*thf* (Hz)
High-pitched CAS areas, inside	20	none	30	10	100	50	25
High-pitched CAS areas, outside	17.5	5	30	25	100	50	25
Low-pitched CAS areas, inside	17.5	5	30	25	50	100	25
Low-pitched CAS areas, outside	10	15	22.5	35	50	100	25
Normal areas, inside	10	15	22.5	35	22.5	none	20
Normal areas, outside	5	20	10	55	10	none	20

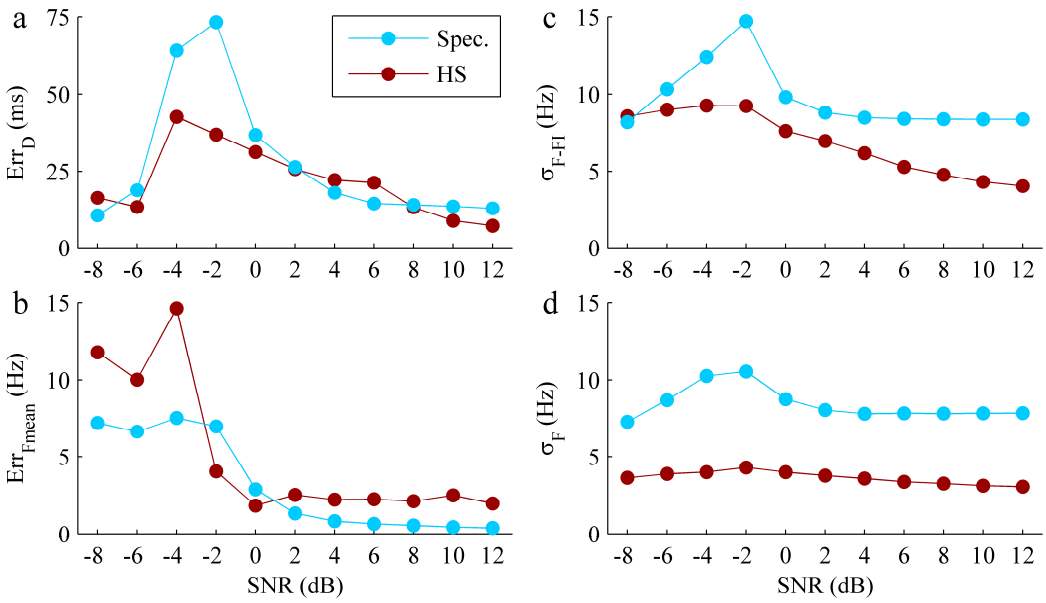


Fig. 15. **Characterization of synthetic CAS signals.** Absolute value of the differences between calculated and real D (Err_D) (a) and F_{mean} (Err_{Fmean}) (b), $\sigma_{F\text{-}FI}$ (c), and σ_F (d), for the HS and spectrogram, as a function of SNR. All values are the mean along RS signals of each SNR value.

However, the advantages of the HS became clear in the SNR range from -4 dB to -2 dB, which represented an unfavorable scenario for synthetic CAS signals containing components below 200 Hz, due to their overlap with normal RS. In those cases, thanks to the high temporal resolution and energy concentration of the HS, it allowed the temporal boundaries of CAS components to be more accurately determined (Err_D = 39.7 ± 46.5 ms). However, the poor resolution and the scattered energy of the spectrogram prevented it from delimiting CAS accurately (Err_D = 68.9 ± 84.6 ms). These differences between Err_D of the two TFDs were statistically significant (p = 0.0018), which indicated that Err_D was higher for the spectrogram than for the HS. Relative to synthetic CAS duration, mean Err_D was between 13% and 16% for the HS and between 23% and 28% for the spectrogram.

As an example, Figs. 16 and 17 show the CAS segmentation of two synthetic monophonic CAS signals added to a normal RS signal at an SNR of -2 dB and -4 dB, respectively, as described in section 2.2.

White rectangles in the figures above (Figs. 16-c,d and 17-c,d) show the boundaries of the theoretical IFs. As shown, it is more difficult to distinguish between CAS components and normal RS in the spectrogram than in the HS, where the boundaries of the CAS component can be detected more accurately. However, some normal RS components are detected as part of the CAS components in the spectrogram. The performance of spectrogram could be improved by increasing the amplitude threshold. However, it would be more difficult to detect weak CAS components, as explained in the next section 6.3.

With regard to Err_{Fmean} (Fig. 15-b), although it was higher in the HS (4.1 ± 9.7 Hz) than in the spectrogram (2.6 ± 6.1 Hz), both Err_{Fmean} were low in comparison with Err_D, which is more critical. However, there were clear and significant differences (p << 0.0001) between the frequency dispersion (σ_F and $\sigma_{F\text{-}FI}$) of both TFDs, as shown in Figs. 15-c and 15-d, which means that the HS has higher energy concentration than spectrogram.

6.3. *Characterization of recorded CAS signals*

In this section, we applied the proposed CAS segmentation algorithm to the HS and spectrogram of the 283 CAS signals recorded as described in section 2.1. In this case, as quantitative measurements of Err_D and Err_{Fmean} could not be obtained, we calculated the mean and standard deviation of D, F_{mean}, and σ_F along all segmented CAS components (Table 3).

There were no relevant differences between D and F_{mean} measured from the two TFDs. Even so, these parameters were absolute measures and were not representative of the performance of the two TFDs in recorded CAS segmentation, unlike Err_D and Err_{Fmean} described in the previous section 6.2. Nevertheless, there were significant differences between the σ_F of both TFDs (p << 0.0001), which again demonstrates that the HS has higher energy concentration than spectrogram.

The advantages of the HS over spectrogram in recorded CAS segmentation, especially weak CAS with low energy, are clearly illustrated in Figs. 18, 19, and 20.

The figures above show that the spectrogram failed to entirely extract some CAS components with low amplitude, especially in CAS signals that contained several CAS components with quite different amplitudes. This performance was contrary to that shown in the examples with synthetic CAS (Fig. 16 and Fig. 17), in which the spectrogram detected some normal RS components as CAS components. However, the same algorithm and thresholds were used in both cases. This means that spectrogram is more dependent on the amplitude threshold. However, since the HS achieves more energy concentration along the CAS components, it is less dependent on this parameter of the CAS segmentation algorithm.

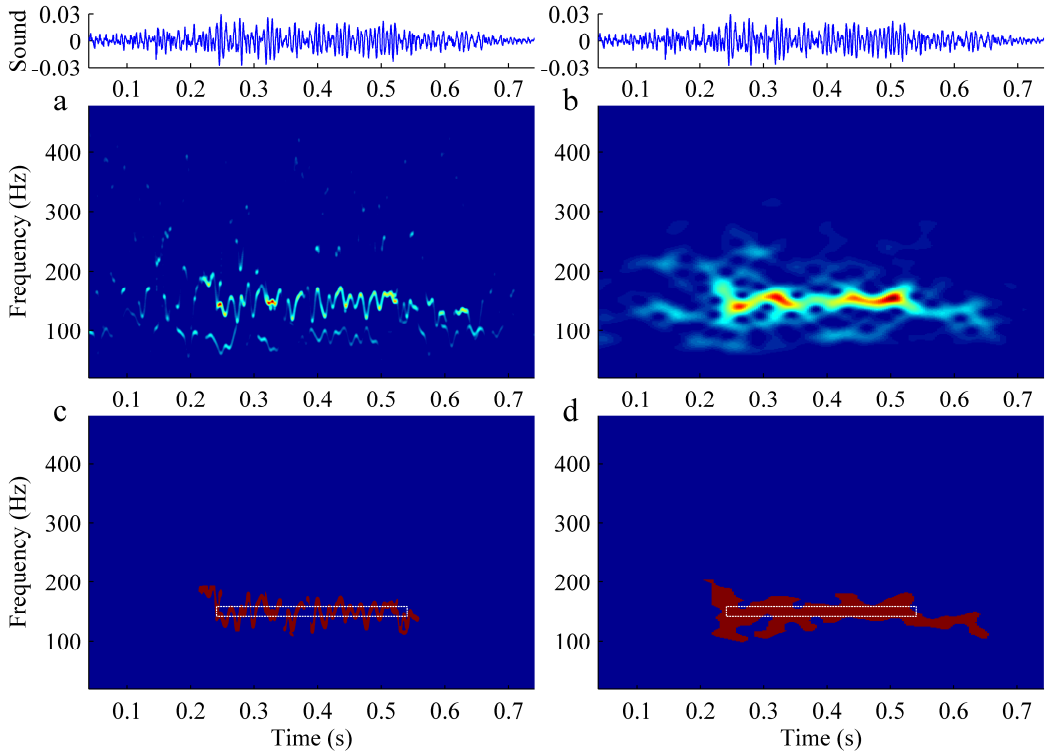


Fig. 16. **Segmentation of synthetic CAS signals (I).** HS (a) and spectrogram (b) of a synthetic monophonic CAS signal. Segmented CAS component using the HS (c) and spectrogram (d).

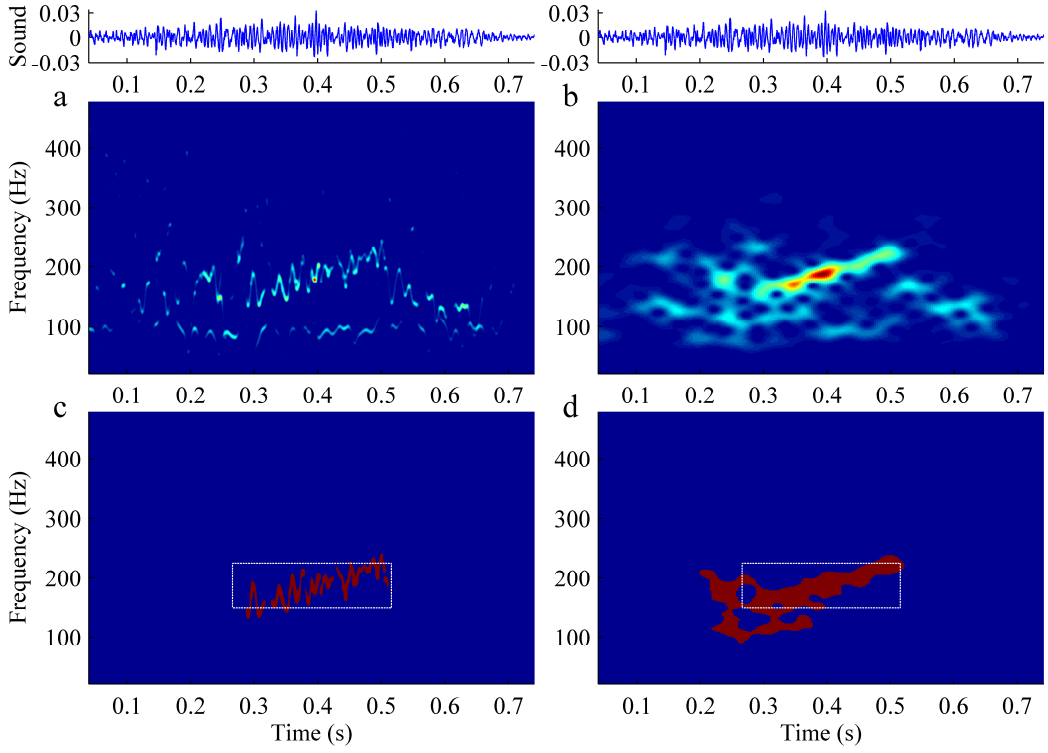


Fig. 17. **Segmentation of synthetic CAS signals (II).** HS (a) and spectrogram (b) of a synthetic monophonic CAS signal with frequency sweeping. Segmented CAS component using the HS (c) and spectrogram (d).

7. Discussion and conclusions

In this study, we conducted a comprehensive evaluation of the performance of the HHT for RS analysis, which led us to propose an EEMD-Kay-based HS that performed very well for CAS characterization. In comparison with spectrogram, which is the most widely used technique for CAS analysis, the HS detected CAS components more precisely, especially those at low SNR that overlap with normal RS.

The most critical stage of the HHT is EMD, due to its MM effect, which causes poor separation of frequency scales. We have gone into detail about the MM effect of EMD in RS signals and the performance of EEMD and NA-MEMD to solve MM. We propose a number of parameters to quantify the size, reduction of MM, and residual noise level of each method. The results after applying EEMD and NA-MEMD to recorded RS signals showed that EEMD is more concise than NA-MEMD, as EEMD produces fewer redundant IMFs and is faster than NA-MEMD. Moreover, EEMD reduces the MM effect more effectively than NA-MEMD.

Table 3. Characteristic parameters of recorded CAS signals

	Spectrogram	HS
D (ms)	299.7 ± 132.6	302.8 ± 122.9
F_{mean} (Hz)	287.0 ± 136.0	285.6 ± 135.0
σ_F (Hz)	7.2 ± 2.6	4.2 ± 1.2

D: duration; F_{mean}: mean frequency; σ_F: frequency dispersion

Although EEMD has already been used in previous studies as part of the HHT [60-62], the procedure for the correct choice of its parameters (amplitude of the added noise and number of iterations) is still unclear, and they must be adjusted to the characteristics of different signals. In this regard, our proposed parameters could be used to assess MM and select EEMD parameters in other applications.

The IF is estimated in the HHT by means of the phase derivative of the analytic signal of IMFs. However, a major drawback of this IF estimation method is the high variance of the IF estimates. Kay's IF estimator proved to be a direct and straightforward method that significantly reduced that variance, which is a desirable property for the purpose of obtaining an accurate HS with high concentrations of signal energy. In fact, we propose the use of an EEMD-Kay-based HS as an alternative and precise time-frequency representation of RS signals.

The main advantage of the proposed HHT-based approach over other TFD-based approaches for RS analysis is the high temporal and high frequency resolution of the HS. Since IF is calculated by differentiation, time resolution can be as high as that determined by the sampling rate. Moreover, the frequency resolution in the HS does not depend on the data length as Fourier-based or wavelet-based techniques, but rather it is determined by the bin size selected. Furthermore, the EEMD allows us to separate different signal components prior to IF and IA calculation without having to process an entire TFD.

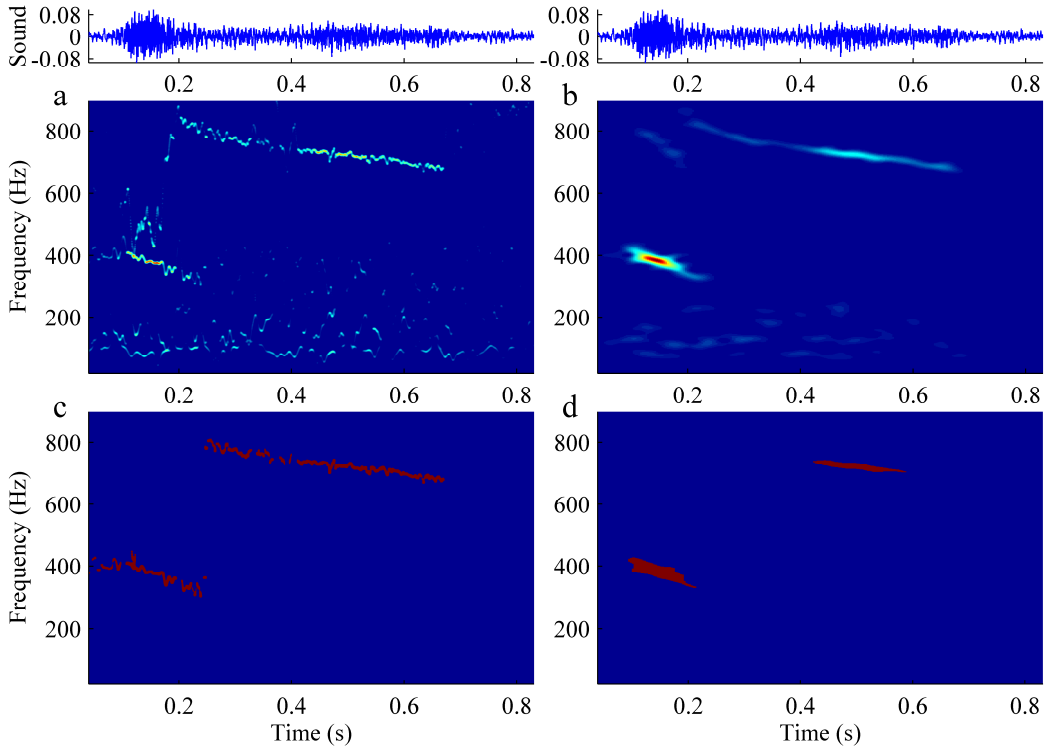


Fig. 18. **Segmentation of recorded CAS signals (I).** HS (a) and spectrogram (b) of a recorded CAS signal that contains two CAS components. Segmented CAS components using the HS (c) and spectrogram (d).

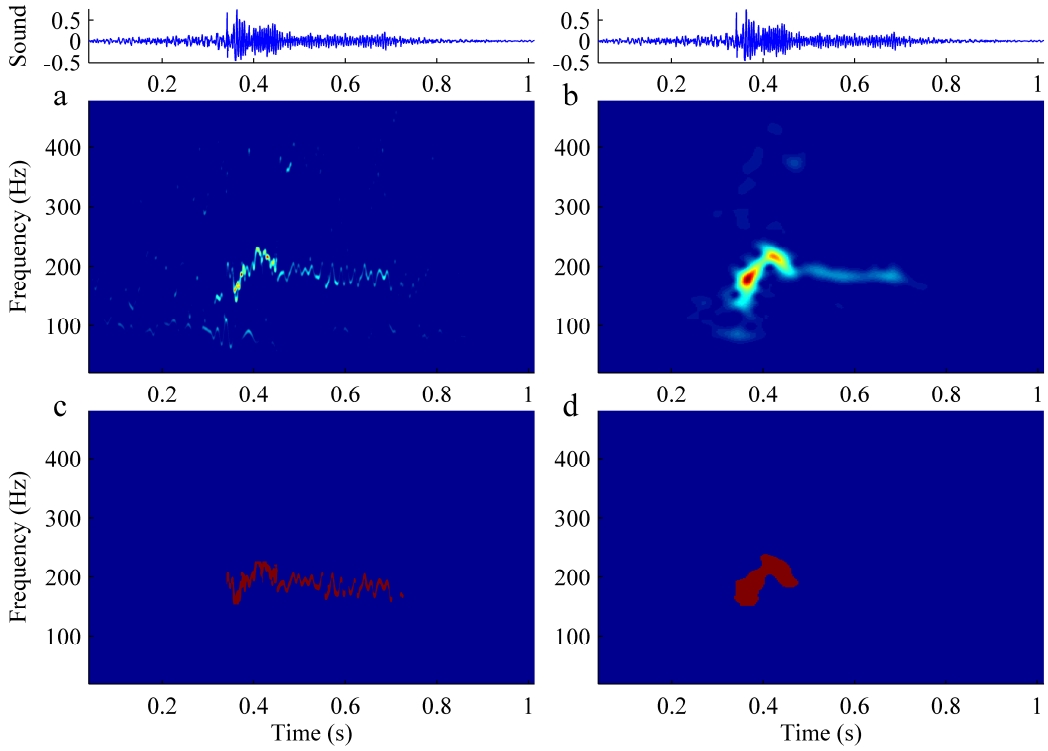


Fig. 19. **Segmentation of recorded CAS signals (II).** HS (a) and spectrogram (b) of a recorded monophonic CAS signal. Segmented CAS component using the HS (c) and spectrogram (d).

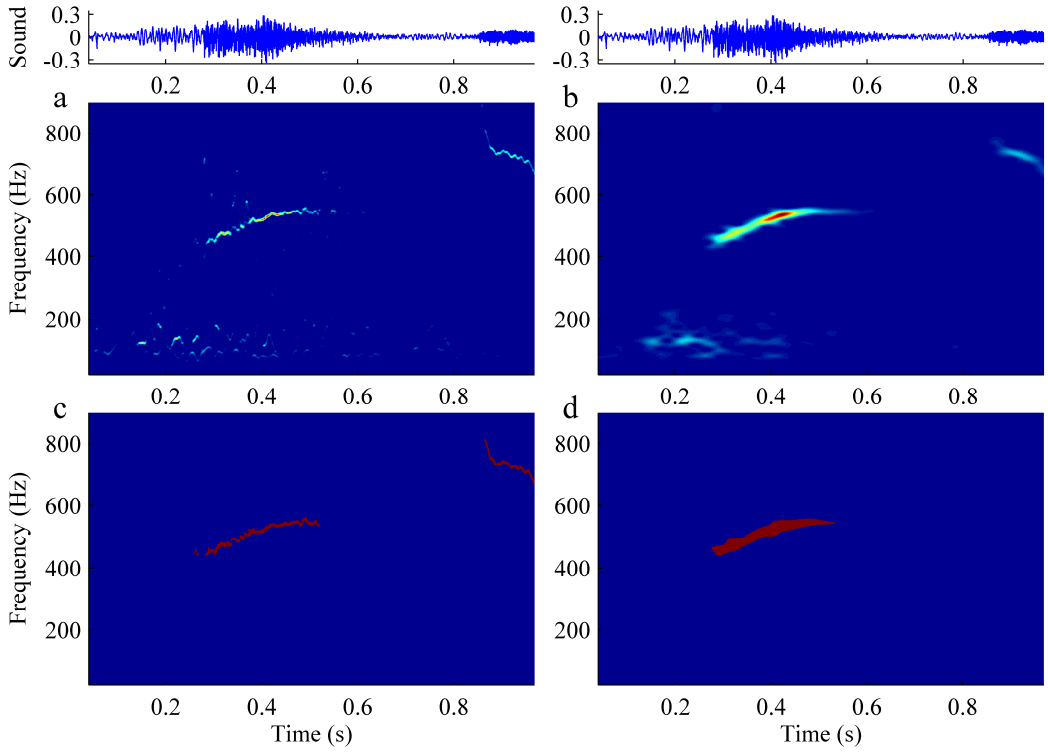


Fig. 20. **Segmentation of recorded CAS signals (III).** HS (a) and spectrogram (b) of a recorded CAS signal that contains two CAS components. Segmented CAS components using the HS (c) and spectrogram (d).

Accordingly, the HHT does not only allow working in a time-frequency plane, like other TFDs, but also analyzing IF and IA sequences separately.

Taking advantage of properties of the HS, we propose a new method for the automatic segmentation and characterization of CAS. A key point of our method is the

CAS detector (section 6.1.1) that we previously proposed in [43]. Using this CAS detector as the first step of our CAS segmentation algorithm, we can locate CAS components in the time-frequency plane, which facilitates their subsequent segmentation using a region growing methodology together with a set of region linking criteria. The main advantage of the CAS detector is that it is based solely on a number of IF criteria, which makes our CAS segmentation algorithm less dependent on amplitude criteria.

We applied the proposed CAS segmentation algorithm to the HS and spectrogram of two sets of synthetic and recorded CAS signals to compare the performance of the two TFDs. The results showed that the HS has some clear advantages over spectrogram. The resolution of spectrogram is limited due to the use of a finite length analysis window. Furthermore, the signal energy in spectrogram is more scattered across the entire time-frequency plane, which blurs the boundaries of some CAS components, especially in CAS components that overlap with normal RS at low SNRs. Therefore, the performance of spectrogram is more dependent on the amplitude thresholds. However, the HS achieves higher energy concentrations around the ridges described by CAS components, which makes the HS less dependent on amplitude criteria. This property, together with the high resolution in time and frequency domains, allows the duration (D) and mean frequency (F_{mean}) of CAS components to be more accurately determined using the HS than spectrogram.

If we consider that the mean D and mean F_{mean} obtained from recorded CAS signals (Table 3 in section 6.3) are representative of real CAS in asthmatic patients, the mean Err_D and mean Err_{Fmean} obtained from synthetic CAS signals (Fig. 15 in section 6.2) were quite acceptable. Both the HS and spectrogram had high performance in measuring F_{mean}, since their mean Err_{Fmean} (below 15 Hz) represented less than 5% of the mean F_{mean} of real CAS (around 290 Hz). With respect to D, the HS performed significantly better than spectrogram, especially for low SNRs, where their mean Err_D reached around 13% (39.7 ms) and 23% (68.9 ms), respectively, of the mean D of real CAS (around 300 ms). In addition to this difference in Err_D at low SNR, we showed some examples of weak CAS that were properly segmented by the HS but were either wrongly segmented or not detected by spectrogram. From a clinical point of view, spectrogram could lead to an underestimation of the presence of CAS and their features, which are directly related to the severity of airflow obstruction [1,2] in patients with obstructive pulmonary diseases.

This study implies a step forward in the analysis of CAS, as most previous approaches for CAS analysis [3-7,9] mainly focused on differentiating CAS from normal RS, but not on analyzing CAS features, such as duration and mean frequency, which are the most relevant clinical parameters. Moreover, some of those studies used spectrogram for CAS detection and extraction. However, our algorithm characterizes CAS more accurately in both time and frequency domains thanks to the properties of the HS, whose performance has been thoroughly tested on both synthetic and real CAS signals.

Our proposed version of the HHT based on EEMD and Kay's IF estimator is a promising tool for the analysis of RS signals. Due to its high resolution, the proposed HS is a suitable TFD to analyze not only CAS signals, but also shorter RS signal components, such as DAS. This methodology, including the CAS characterization algorithm could be included within a more complex RS analysis system that facilitates long-term monitoring and improves reliability in the diagnosis of obstructive pulmonary diseases.

Acknowledgements

All of the authors would like to thank the entire team at the Pulmonary Function Testing Laboratory of Germans Trias i Pujol University Hospital for their collaboration in recording RS from asthmatic patients. We also thank the joint research unit between the Institute for Bioengineering of Catalonia and Health Sciences Research Institute of the Germans Trias i Pujol Foundation. Finally, special thanks to B. Boashash for the TFSA Package (*Time-Frequency Signal Analysis Package*, The University of Queensland, UQ Centre for Clinical Research, 2010), which was used for IF estimation using Kay's method.

References

[1] A. Bohadana, G. Izbicki, and S.S. Kraman, "Fundamentals of lung auscultation," *N. Engl. J. Med.*, vol. 370, pp. 744–751, Feb. 2014.

[2] A.R.A. Sovijärvi, L.P. Malmberg, G. Charbonneau, J. Vanderschoot, F. Dalmasso, C. Sacco, M. Rossi, and J.E. Earis, "Characteristics of breath sounds and adventitious respiratory sounds," *Eur. Respir. Rev.*, vol. 10, pp. 591–596, 2000.

[3] A. Homs-Corbera, J.A. Fiz, J. Morera, and R. Jané, "Time-frequency detection and analysis of wheezes during forced exhalation," *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 182–186, Jan. 2004.

[4] B.S. Lin, B.S. Lin, H.D. Wu, F.C. Chong, and S.J. Chen, "Wheeze recognition based on 2D bilateral filtering of spectrogram," *Biomed. Eng. Appl. Basis Commun.*, vol. 18, pp. 128–137, Jun. 2006.

[5] S.A. Taplidou and L.J. Hadjileontiadis, "Wheeze detection based on time-frequency analysis of breath sounds," *Comput. Biol. Med.*, vol. 37, pp. 1073–1083, Aug. 2007.

[6] R.J. Riella, P. Nohama, and J.M. Maia, "Method for automatic detection of wheezing in lung sounds," *Braz. J. Med. Biol. Res.*, vol. 42, pp. 674–684, Jul. 2009.

[7] A. Oliveira, C. Pinho, J. Dinis, D. Oliveira, and A. Marques, "Automatic wheeze detection and lung function evaluation: a preliminary study," in *Proc. Int. Conf. Health Inform.*, Barcelona, 2013, pp. 323–326.

[8] S.A. Taplidou and L.J. Hadjileontiadis, "Analysis of wheezes using wavelet higher order spectral features," *IEEE Trans. Biomed. Eng.*, vol. 57, pp. 1596–1610, Jul. 2010.

[9] F. Jin, S.S. Krishnan, and F. Sattar, "Adventitious sounds identification and extraction using temporal-spectral dominance-based features," *IEEE Trans. Biomed. Eng.*, vol. 58, pp. 3078–3087, Nov. 2011.

[10] L.J. Hadjileontiadis and S.M. Panas, "Separation of discontinuous adventitious sounds from vesicular sounds using a wavelet-based filter," *IEEE Trans. Biomed. Eng.*, vol. 44, pp. 1269–1281, 1997.

[11] F. Ayari, A.T. Alouani, and M. Ksouri, "Wavelets: an efficient tool for lung sounds analysis," in *IEEE Int. Conf. Comput. Syst. Appl.*, Doha, 2008, pp. 875–878.

[12] G. Serbes, C.O. Sakar, Y.P. Kahya, and N. Aydin, "Feature extraction using time-frequency/scale analysis and ensemble of feature sets for crackle detection," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, Boston, 2011, pp. 3314–3317.

[13] I.T. Rekanos and L.J. Hadjileontiadis, "An iterative kurtosis-based technique for the detection of nonstationary bioacoustic signals," *Signal Process.*, vol. 86, pp. 3787–3795, 2006.

[14] L.J. Hadjileontiadis and I.T. Rekanos, "Detection of explosive lung and bowel sounds by means of fractal dimension," *IEEE Signal Process. Lett.*, vol. 10, pp. 311–314, 2003.

[15] S. Charleston-Villalobos, G. Dorantes-Méndez, R. González-Camarena, G. Chi-Lem, J.G. Carrillo, and T. Aljama-Corrales, "Acoustic thoracic image of crackle sounds using linear and nonlinear processing techniques," *Med. Biol. Eng. Comput.*, vol. 49, pp. 15–24, 2011.

[16] B. Boashash, "Time frequency signal analysis and processing. A comprehensive reference," Elsevier, First Edition, 2003, chapter 7, 770 pages. (9780080443355, 9780080543055).

[17] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal-part 2: algorithms and applications," *Proc. IEEE*, vol. 80, pp. 540–568, 1992.

[18] N.E. Huang, Z. Wu, S.R. Long, K.C. Arnold, X. Chen, and K. Blank, "On instantaneous frequency," *Adv. Adapt. Data Anal.*, vol. 1, pp. 177–229, Apr. 2009.

[19] P. O'Shea and B. Boashash, "Some robust instantaneous frequency estimation techniques with application to non-stationary transient detection," in *Proc. Eur. Conf. Signal Process.*, Barcelona, 1990, pp. 165–168.

[20] L.J. Griffiths, "Rapid measurement of digital instantaneous frequency," *IEEE Trans. Acoust.*, vol. 23, pp. 207–222, 1975.

[21] J.M. Cioffi and T. Kailath, "Fast, recursive-least-squares transversal filters for adaptive filtering," *IEEE Trans. Acoust.*, vol. 32, pp. 304–337, 1984.

[22] K.C. Sharman and B. Friedlander, "Time-varying autoregressive modeling of a class of nonstationary signals," in *IEEE Int. Conf. Acoust. Speech Signal Process.*, San Diego, 1984, pp. 227–230.

[23] L. Stankovic, I. Djurovic, S. Stankovic, M. Simeunovic, S. Djukanovic, and M. Dakovic, "Instantaneous frequency in time-frequency analysis: enhanced concepts and performance of estimation algorithms," *Dig. Signal Process.*, vol. 35, pp. 1–13, 2014.

[24] P. Flandrin and W. Martin, "A general class of estimators for the Wigner-Ville spectrum of non-stationary processes," in *Systems Analysis and Optimization of Systems, Lecture Notes in Control and Information Sciences*. Berlin, Vienna, New York Springer-Verlag, 1984, pp. 15–23.

[25] J. Jeong and W.J. Williams, "Kernel design for reduced interference distributions," *IEEE Trans. Signal Process.*, vol. 40, pp. 402–412, 1992.

[26] S.C. Sekhar and T.V. Sreenivas, "Adaptive spectrogram vs. adaptive pseudo-Wigner-Ville distribution for instantaneous frequency estimation," *Signal Process.*, vol. 83, pp. 1529–1543, 2003.

[27] J. Zhong and Y. Huang, "Time-frequency representation based on an adaptive short-time Fourier transform," *IEEE Trans. Signal Process.*, vol. 58, pp. 5118–5128, 2010.

[28] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," vol. 43, pp. 1068–1089, 1995.

[29] G.K. Nilsen, "Recursive time-frequency reassignment," *IEEE Trans. Signal Process.*, vol. 57, pp. 3283–3287, 2009.

[30] F. Auger, P. Flandrin, Y.-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, and H.-T. Wu, "Time-frequency reassignment and synchrosqueezing: an overview," *IEEE Signal Process. Mag.*, vol. 30, pp. 32–41, 2013.

[31] L. Rankine, M. Mesbah, and B. Boashash, "IF estimation for multicomponent signals using image processing techniques in the time-frequency domain," *Signal Process.*, vol. 87, pp. 1234–1250, 2007.

[32] P. Comon and C. Jutten, "Handbook of blind source separation. Independent component analysis and applications," Academic Press, 2010.

[33] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, and H.H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. A*, vol. 454, pp. 903–995, 1998.

[34] N.E. Huang, X. Chen, M.T. Lo, and Z. Wu, "On Hilbert spectral representation: a true time-frequency representation for nonlinear and nonstationary data," *Adv. Adapt. Data Anal.*, vol. 3, pp. 63–93, 2011.

[35] L. Mayeta, A. López, and A.F. Ruiz, "Evaluation of the Hilbert-Huang transform for myoelectric pattern classification: towards a method to detect movement intention," in *Panam. Health Care Exch.*, Medellin, 2013, pp. 1–6.

[36] A. Mert and A. Akan, "Hilbert-Huang transform based hierarchical clustering for EEG denoising," in *Proc. Eur. Conf. Signal Process.*, Marrakech, 2013, pp. 1–5.

[37] J. Yan and L. Lu, "Improved Hilbert-huang transform based weak signal detection methodology and its application on incipient fault diagnosis and ECG signal analysis," *Signal Process.*, vol. 98, pp. 74–87, 2014.

[38] Z. Zhidong, Z. Zhijin, and C. Yuquan, "Time-frequency analysis of heart sound based on HHT," in *Proc. IEEE Int. Conf. Commun. Circuits Syst.*, Hong Kong, 2005, pp. 926–929.

[39] L. Zhenzhen and D. Minghui, "HHT based lung sound crackle detection and classification," in *Proc. Int. Symp. Intell. Signal Process. Commun. Syst.*, 2005, pp. 385–388.

[40] B.A. Reyes, S. Charleston-Villalobos, R. Gonzalez-Camarena, and T. Aljama-Corrales, "Analysis of discontinuous adventitious lung sounds by Hilbert-Huang spectrum," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, Vancouver, 2008, pp. 3620–3623.

[41] B.A. Reyes, S. Charleston-Villalobos, R. González-Camarena, and T. Aljama-Corrales, "Assessment of time-frequency representation techniques for thoracic sounds analysis," *Comput. Methods Programs Biomed.*, vol. 114, pp. 276–290, 2014.

[42] M. Lozano, J.A. Fiz, and R. Jané, "Estimation of instantaneous frequency from empirical mode decomposition on respiratory sounds analysis," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, Osaka, 2013, pp. 981–984.

[43] M. Lozano, J.A. Fiz, and R. Jané, "Automatic differentiation of normal and continuous adventitious respiratory sounds using ensemble empirical mode decomposition and instantaneous frequency," *IEEE J. Biomed. Health Inform.*, 2015 [Epub ahead of print].

[44] Z. Wu and N.E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, pp. 1–41, Jan. 2009.

[45] L.J. Hadjileontiadis, "Empirical mode decomposition and fractal dimension filter. A novel technique for denoising explosive lung sounds," *IEEE Eng. Med. Biol. Mag.*, vol. 26, no. 1, pp. 30–39, 2007.

[46] S. Liu, R.X. Gao, D. John, J. Staudenmayer, and P. Freedson, "Tissue artifact removal from respiratory signals based on empirical mode decomposition," *Ann. Biomed. Eng.*, vol. 41, pp. 1003–1015, 2013.

[47] S. Charleston-Villalobos, R. Gonzalez-Camarena, G. Chi-Lem, and T. Aljama-Corrales, "Crackle sounds analysis by empirical mode decomposition," *IEEE Eng. Med. Biol. Mag.*, vol. 26, no. 1, pp. 40–47, 2007.

[48] S. Charleston-Villalobos, L.F. Dominguez-Robert, R. Gonzalez-Camarena, and A.T. Aljama-Corrales, "Heart sounds interference cancellation in lung sounds," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, New York, 2006, pp. 1694–1697.

[49] J. Zhang, R. Yan, R.X. Gao, and Z. Feng, "Performance enhancement of ensemble empirical mode decomposition," *Mech. Syst. Signal Process.*, vol. 24, pp. 2104–2123, Oct. 2010.

[50] N. Rehman and D.P. Mandic, "Filter bank property of multivariate empirical mode decomposition," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2421–2426, May 2011.

[51] L. ChingShun, W.A. Tanumihardja, and S. HongHui, "Lung-heart sound separation using noise assisted multivariate empirical mode decomposition," in *Int. Symp. Intell. Signal Process. Commun. Syst.*, Naha, 2013, pp. 726–730.

[52] Z. Wu and N.E. Huang, "A study of the characteristics of white noise using the empirical mode decomposition method," *Proc. R. Soc. Lond. A*, vol. 460, pp. 1597–1611, 2004.

[53] P. Flandrin, G. Rilling, and P. Gonçalvés, "Empirical mode decomposition as a filter bank," *IEEE Signal Process. Lett.*, vol. 11, pp. 112–114, 2004.

[54] N. Rehman and D.P. Mandic, "Multivariate empirical mode decomposition," *Proc. R. Soc. Lond. A*, vol. 466, pp. 1291–1302, 2010.

[55] [Online]. 2014 Available: http://www.commsp.ee.ic.ac.uk/~mandic/research/emd.htm.

[56] G. Rilling, P. Flandrin, and P. Gonçalvès, "On empirical mode decomposition and its algorithms," in *Proc. IEEE/EURASIP Workshop Nonlinear Signal Image Process.*, Grado, 2003.

[57] [Online]. 2012 Available: http://perso.ens-lyon.fr/patrick.flandrin/emd.html.

[58] W.J. Tompkins, "Biomedical Digital Signal Processing," Prentice-Hall, 1993, chapter 5, pp. 111–116.

[59] P. Maragos, J.F. Kaiser, and T.F. Quatieri, "Energy separation in signal modulations with application to speech analysis," *IEEE Trans. Signal Process.*, vol. 41, pp. 3024–3051, 1993.

[60] H. Li, Y. Wang, and Y. Ma, "Ensemble empirical mode decomposition and Hilbert-Huang transform applied to bearing fault diagnosis," in *IEEE Int. Congr. Image Signal Process.*, Yantai, 2010, pp. 3413–3417.

[61] W. Zhang, X. Zhang, and Y. Sun, "Based on EEMD-HHT marginal spectrum of speech emotion recognition," in *IEEE Int. Conf. Comput. Meas. Control Sens. Netw.*, Taiyuan, 2012, pp. 91–94.

[62] Y. Amirat, M. Benbouzid, T. Wang, and S. Turri, "Performance analysis of an EEMD-based Hilbert Huang transform as a bearing failure detector in wind turbines," in *Int. Conf. Green Energy*, Sfax, 2014, pp. 193–198.

Chapter 5: Analysis of normal and continuous adventitious sounds for the assessment of asthma

Title: Analysis of normal and continuous adventitious sounds for the assessment of asthma

Authors: M. Lozano, J. A. Fiz, R. Jané

Book chapter: XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013

Analysis of Normal and Continuous Adventitious Sounds for the Assessment of Asthma

M. Lozano[1,4], J.A. Fiz[2,4,5] and R. Jané[3,4,5]

[1] Innovation Group at Health Sciences Research Institute of the Germans Trias i Pujol Foundation (IGTP), Badalona, Spain
[2] Pneumology Service at the Germans Trias i Pujol University Hospital, Badalona, Spain
[3] Department of Automatic Control (ESAII), Universitat Politècnica de Catalunya (UPC), Barcelona, Spain
[4] Institute for Bioengineering of Catalonia (IBEC), Barcelona, Spain
[5] Biomedical Research Networking Center in Bioengineering, Biomaterials, and Nanomedicine (CIBER-BBN), Spain

Abstract— **Assessment of asthma is a difficult procedure which is based on the correlation of multiple factors. A major component in the diagnosis of asthma is the assessment of BD response, which is performed by traditional spirometry. In this context, the analysis of respiratory sounds (RS) provides relevant and complementary information about the function of the respiratory system. In particular, continuous adventitious sounds (CAS), such as wheezes, contribute to assess the severity of patients with obstructive diseases. On the other hand, the intensity of normal RS is dependent on airflow level and, therefore, it changes depending on the level of obstruction. This study proposes a new approach to RS analysis for the assessment of asthmatic patients, by combining the quantification of CAS and the analysis of the changes in the normal sound intensity-airflow relationship. According to results obtained from three patients with different characteristics, the proposed technique seems more sensitive and promising for the assessment of asthma.**

Keywords— **asthma, bronchodilator response, continuous adventitious sound, respiratory sound intensity, wheezes.**

I. Introduction

Asthma is a complex respiratory disorder that results in a variable, recurring, and often reversible airflow obstruction [1]. Physicians have difficulties in diagnosing asthma, since they have to correlate several aspects: the medical history, a thorough physical examination, and pulmonary function test results. In this context, the bronchodilator (BD) response is a standard pulmonary function test used to control and assess the severity of asthma. It is based on spirometric measurements before and after the administration of a BD. Usually, an improvement in forced expiratory volume in one second (FEV_1) of greater than or equal to 12% is considered to be significant [2, 3]. However, some recent studies have demonstrated that using spirometric criteria alone is inadequate for the diagnosis of asthma [4].

Respiratory sounds (RS) are helpful in understanding the function of the respiratory system. They are classified as normal or adventitious sounds. Due to their clinical interest, many technical studies have tried to detect and characterize continuous adventitious sounds (CAS) [5-8], such as wheezes. CAS are characterized by a pitch of over 100 Hz that lasts more than 100 ms [9], and they are key indicators for assessing the severity of asthma [1]. On the other hand, some other studies have tried to understand the origin of normal RS and their intensity pattern distribution [10-12]. Although results from all previous studies have contributed to the characterization and understanding of RS, there is lack of clinical use and application of these techniques.

In this study, we propose a new approach to the analysis of RS for the assessment of asthmatic patients, by combining the quantification of CAS and the analysis of the normal sound intensity-airflow relationship. A few previous studies have focused on the evaluation of asthma by RS analysis [13-15]. Nevertheless, some were performed on infants and they were based on manual detection of wheezes and their characterization at a fix airflow level or during forced breathing. On the other hand, in [15] they focused on changes in the spatial distribution of breath sound intensity by analyzing dynamic images. Our technique has two major advantages: the automatic differentiation and quantifying of respiratory cycles either with normal sounds or CAS [16], and the analysis of normal RS intensity as a function of airflow level.

II. Methods

A. Signal acquisition

RS signals were recorded from asthmatic patients in a sitting position at the Pulmonary Function Testing Laboratory, Germans Trias i Pujol University Hospital, Badalona, Spain. Three piezoelectric contact microphones (TSD108, Biopac, Inc.) were attached to the skin using adhesive rings: two of them on the back at 3 cm below the left/right shoulder blade, and one over the right side of the trachea. Moreover, respiratory airflow signal was recorded using a pneumotachograph (TSD107B, Biopac, Inc.). Each patient was coached to progressively increase the airflow, from shallow breathing to the deepest breaths they could. All signals were sampled at 12500 samples/second using a 16-

bit analogue-to-digital converter. After acquisition, RS signals were band-pass filtered using a combination of 8th order Butterworth low-pass and high-pass filters (70 – 2000 Hz). We show a case study with three adult asthmatic patients with different baseline spirometric values and BD response (Table 1). For each patient, we have quantified the percentage of respiratory cycles with CAS at baseline and after BD, for both left and right sides. In addition, we have analyzed the relationship between RS intensity and airflow, before and after BD.

Table 1 Characteristics of asthmatic patients

ID	Total cycles Pre-BD		Total cycles Post-BD		Age	Sex	BMI (Kg/m^2)	FEV_1 (%)	ΔFEV_1 (%)
	Left	Right	Left	Right					
1	46	47	53	53	50	F	24.44	47	26
2	68	66	76	83	60	M	27.08	100	1
3	59	49	53	50	19	M	19.28	59	6

B. Segmentation of RS signals

After signal acquisition, respiratory phases were obtained using the airflow signal as the reference for automatic sound signal segmentation. Since airflow is positive during inspiration and negative during expiration, respiratory phases were marked off by means of a robust zero crossing detector. In order to avoid detection of false endpoints, only cycles in which the airflow reached at least 0.35 L/s were considered valid cycles. Moreover, two thresholds of 0.2 and 4 seconds were established for minimum and maximum durations of respiratory phases, respectively, according to time duration of normal respiratory cycles. In addition, a threshold of 0.5 seconds was fixed for the maximum time interval between the end of inspiration and the beginning of the corresponding expiration. Two final datasets, pre-BD and post-BD, were obtained for each patient and each side. Each dataset was formed by audio-visual selection of sound signals from the inspiratory cycles, avoiding artifacts such as those from speaking, swallowing, coughing or rubbing. We have focused on inspiratory sounds, which are much louder than expiratory sounds on the back, where we recorded the sound signals.

C. CAS detection and inspiratory cycle classification

The first step in the analysis of RS is to differentiate respiratory cycles with normal sounds from those with CAS. For that purpose, we made use of an automatic RS classifier, which we had previously developed based on the analysis and feature extraction from instantaneous frequency (IF) of sound signal from each respiratory cycle [16]. Prior to the IF calculation, sound signals were decomposed into narrow-band components by ensemble empirical mode decomposition (EEMD). The core of the proposed classifier is the fact that IF remains almost constant when a CAS signal is within a respiratory cycle. By using this classifier, we quantified the number of CAS cycles from both pre-BD and post-BD datasets, for each patient and from both left and right sides. Then, we evaluated whether the percentage of CAS cycles had significantly increased or decreased with the administration of the BD.

D. RS intensity – respiratory airflow graphs

In addition to evaluate the changes in the percentage of CAS cycles pre/post-BD, we also analyzed the changes in normal sound intensity. Sound intensity was calculated for inspiratory sound segments corresponding to the top 20% of airflow from each normal sound cycle, from both pre-BD and post-BD datasets. It was defined as the mean power obtained from the power spectral density (PSD) in the frequency band 75-600 Hz. PSD was calculated using Welch's periodogram, with a Hanning window of 80 ms, a 50% overlap between adjacent segments, and 2048 points for the Fast Fourier Transform. Then, each normal sound cycle was determined by its intensity and the maximum airflow reached. Since each patient was characterized by the relationship between normal sound intensity and airflow on both left and right sides, we analyzed the changes in these graphs in order to evaluate the BD effect.

III. Results

In this section we show results from applying the previous techniques to patients shown in Table 1. Firstly, results from respiratory cycle classification and CAS cycle quantification are shown in Fig. 1.

Secondly, we calculated RS intensity from normal sound cycles as a function of the airflow, pre-BD and post-BD, for both left and right sides (Fig. 2).

The patient with severe asthma (ID 1) had 35% (left) and 30% (right) of CAS cycles before BD, as shown in Fig. 1. It is in agreement with her low baseline FEV_1 (47%), which shows that she had a severe bronchial obstruction. After BD, she had not CAS cycles, which is in agreement with her increased FEV_1 (ΔFEV_1 = 26%). Moreover, normal sound intensity significantly increased after BD, in both sides, as can be appreciated from the polynomial regression lines shown in Fig. 2-A.

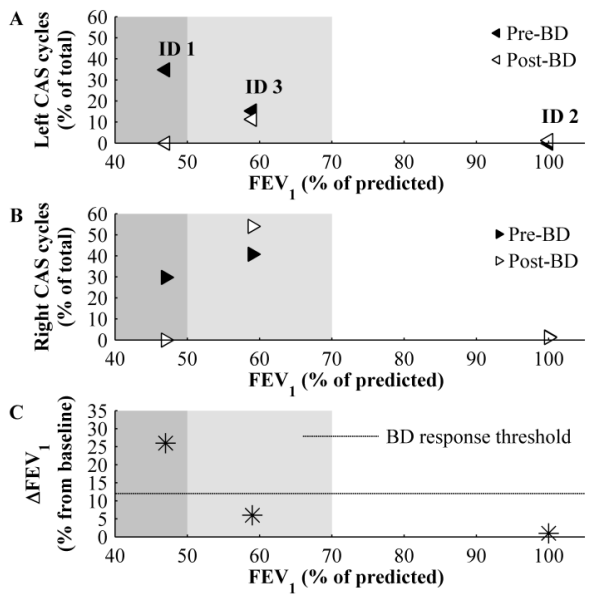


Fig. 1 Percentage of CAS cycles in both left and right sides, at baseline (pre-BD) and after the BD (post-BD) (A, B). Change in FEV_1 after BD, as a per cent from baseline, and threshold for a significant BD response (12%) (C). Three degrees of severity were defined based on baseline FEV_1: mild ($x \geq 70\%$), moderate ($50\% \leq x < 70\%$), and severe ($x < 50\%$).

Contrary to ID1, patient with mild asthma (ID 2) had a normal baseline FEV_1 (100%), and a low number of CAS cycles (<1.5%) before and after BD, as shown in Fig. 1. Moreover, he was a non-responder to BD (ΔFEV_1 = 1%), which agree with a very low increase of normal sound intensity, in both sides (Fig. 2-B). This is reflected in very close polynomial regression lines, pre and post-BD.

In addition to the previous extreme cases, we analyzed an intermediate case with moderate asthma (ID 3). He was a slight responder to BD, since he had a ΔFEV_1 of 6%. However, he had a baseline FEV_1 of 59%, which is in the range of moderate-severe asthma. It agrees with his 15% (left) and 41% (right) of CAS cycles at baseline. After BD, the percentages of CAS cycles were 11% (left) and 54% (right), which were maintained high values. Although his BD response was not significant, lines from the polynomial regression in Fig. 2-C show that he had a relevant increase in normal sound intensity, and it was clearer at high airflows.

IV. Discussion

According to the aforementioned results, the baseline FEV_1 is related to the number of CAS cycles. Those patients with a low baseline FEV_1 have high probability of having wheezes, and vice versa.

On the other hand, the changes in the number of CAS and normal sound intensity are more related to the BD response. A positive BD response indicates that the bronchial obstruction has significantly decreased and, therefore, there are less CAS cycles and higher normal sound intensity.

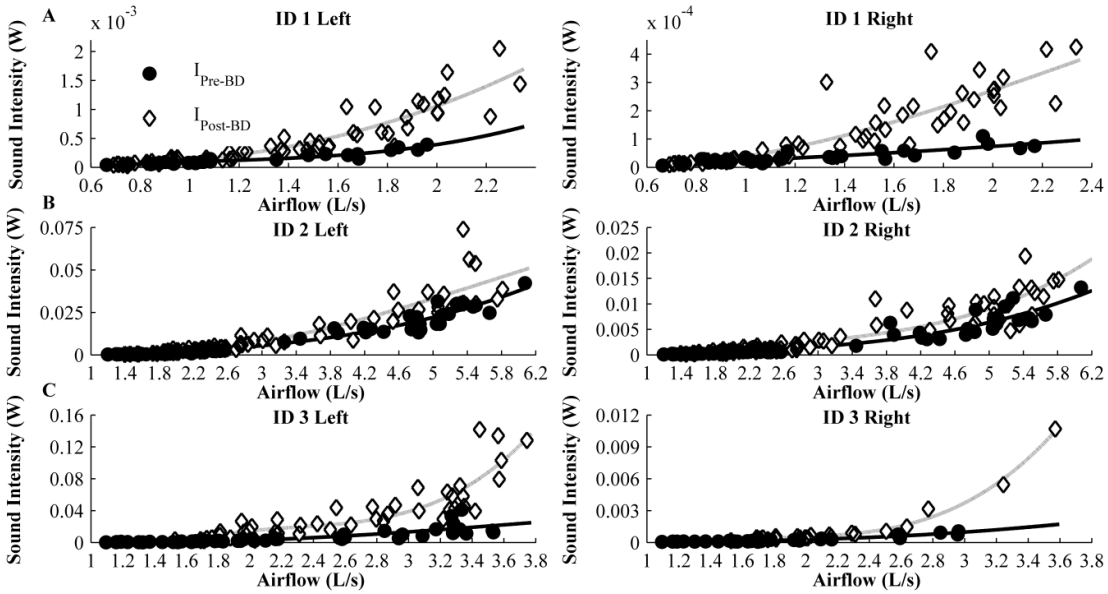


Fig 2. Comparison between normal sound intensity and airflow before and after the administration of a BD. For each patient, results from both left and right sides are shown. Lines black and grey show the 3rd order polynomial regression models that fit pre-BD and post-BD intensities as functions of airflow, respectively.

On the contrary, non-responders do not have many changes in their bronchial tree, which implies having few changes in the number of CAS cycles and low increase of their normal sound intensity.

Moreover, we have shown that for a patient with a non-significant BD response (below 12%), the analysis of normal sound intensity is a more sensitive technique, since significant increases can be detected after BD. Furthermore, these increases are more relevant at high airflow levels.

With respect to the proposed technique, it has some advantages: the exploration and evaluation of the pulmonary function at different airflow levels, which are reached during the progressive respiratory maneuver, and the use of two recording channels that allow us to obtain information about the laterality of normal RS and CAS.

V. Conclusions

In this study we have shown that the combined analysis of normal RS and CAS provides a promising approach to characterize asthmatic patients. It is a simple and non-invasive technique, which seems more sensitive to changes in the pulmonary function. Therefore, it could be a complementary tool in the diagnostic procedure for asthma. However, this work is a case-study with three patients and therefore a higher number of varied cases are required in order to draw final conclusions.

Acknowledgment

This work was made possible thanks to a collaboration agreement between IBEC and IGTP, and it was supported in part by the Spanish Ministry of Economy and Competitiveness under grant TEC2010-21703-C03-01. All authors would like to thank Rosa Gómez, from the pulmonary function testing laboratory at Germans Trias i Pujol University Hospital, for its collaboration in the patient recruitment.

References

1. National Heart, Lung, and Blood Institute (2007) National asthma education and prevention program. Expert panel report 3: Guidelines for the diagnosis and management of asthma. http://www.ncbi.nlm.nih.gov/books/NBK7232/
2. Reddel HK, Taylor DR, Bateman ED et al. (2009) An official American Thoracic Society/European Respiratory Society statement: asthma control and exacerbations: standardizing endpoints for clinical asthma trials and clinical practice. Am J Respir Crit Care Med 180:59–99
3. Pellegrino R, Viegi G, Brusasco V et al. (2005) Interpretative strategies for lung function tests. Eur Respir J 26:948–968
4. Gjevre JA, Hurst TS, Taylor-Gjevre RM, Cockcroft DW (2006) The American Thoracic Society's spirometric criteria alone is inadequate in asthma diagnosis. Can Respir J 13:433–437
5. Bahoura M (2009) Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes. Comput Biol Med 39:824–843
6. Homs-Corbera A, Fiz JA, Morera J, Jané R (2004) Time-frequency detection and analysis of wheezes during forced exhalation. IEEE Trans Biomed Eng 51:182–186
7. Taplidou SA, Hadjileontiadis LJ (2010) Analysis of wheezes using wavelet higher order spectral features. IEEE Trans Biomed Eng 57:1596–1610
8. Xie S, Jin F, Krishnan S, Sattar F (2012) Signal feature extraction by multi-scale PCA and its application to respiratory sound classification. Med Biol Eng Comput 50:759–768
9. Sovijärvi ARA, Dalmasso F, Vanderschoot J et al. (2000) Definitions of terms for applications of respiratory sounds. Eur Respir Rev 10:597–610
10. Murphy R (2007) Computerized multichannel lung sound analysis. Development of acoustic instruments for diagnosis and management of medical conditions. IEEE Eng Med Biol Mag 26:16–19
11. Pasterkamp H, Consunji-Araneta R, Oh Y, Holbrow J (1997) Chest surface mapping of lung sounds during methacholine challenge. Pediatr Pulmonol 23:21–30
12. Torres-Jimenez A, Charleston-Villalobos S, Gonzalez-Camarena R, Chi-Lem G, Aljama-Corrales T (2008) Asymmetry in lung sound intensities detected by respiratory acoustic thoracic imaging (RATHI) and clinical pulmonary auscultation. Conf Proc IEEE Eng Med Biol Soc 2008:4797–4800
13. Sánchez I, Vizcaya C, García D, Campos E (2005) Response to bronchodilator in infants with bronchiolitis can be predicted from wheeze characteristics. Respirology 10:603–608
14. Mazic I, Sovilj S, Magjarevic R (2003) Analysis of respiratory sounds in asthmatic infants. Meas Sci Rev 3:9–12
15. Guntupalli KK, Reddy RM, Loutfi RH et al. (2008) Evaluation of obstructive lung disease with vibration response imaging. J Asthma 45:923–930
16. Lozano M, Fiz JA, Jané R (2013) Estimation of instantaneous frequency from empirical mode decomposition on respiratory sounds analysis. In: Proceedings of the 35th International Conference of the IEEE Engineering in Medicine and Biology Society, Osaka (in press)

Corresponding author:

Author: Manuel Lozano
Institute: Institut for Bioengineering of Catalonia
Street: Baldiri Reixac, 4, Torre I, 9th floor, 08028
City: Barcelona
Country: Spain
Email: mlozano@ibecbarcelona.eu

Chapter 6: Discussion

In this study, we propose a new integrated approach to RS analysis that includes a multichannel recording of RS using a maximum of five microphones and a progressive respiratory manoeuvre, and the analysis of both normal RS and CAS through the RS intensity–airflow curves and HHT, respectively.

We also propose the application of our approach to assess pulmonary function, based on acoustic parameters, for patients with either UPP or asthma.

Although a detailed discussion and conclusions section was included in each of the published articles (Chapters 2–5), we will briefly overview and discuss the main findings and contributions of these studies in this section.

Multichannel RS recording protocol

We proposed a multichannel recording of RS with a maximum of five contact microphones—one on the neck for recording tracheal sounds, and four on the back for recording lung sounds (Figure 10). The CORSA guidelines highlighted that tracheal and lung sounds contain distinct, but complementary, information that may be useful in assessing respiratory diseases [31]. These guidelines also recommended using at least the trachea and the right and left posterior base of the lungs for recording RS [58]. Although the basic configuration proposed in the CORSA guidelines may be suitable for some applications, such as assessing UPP as in this thesis, this configuration does not allow the distribution of lung sounds to be analysed in the vertical plane of the chest surface. Moreover, different configurations regarding the number and location of RS recording points have been reported in the literature, most of which do not meet these requirements. This underscores the lack of a common methodology for recording RS.

The VRI system has also been used in several previous studies for analysing the distribution of RS intensity over the chest surface. It has been reported that the measurements of regional ventilation distribution provided by the VRI system are comparable to those obtained by electrical impedance tomography [189]. However, although the dynamic grey-scale images provided by the VRI system are useful for the qualitative assessment of regional ventilation distribution, this system only provides quantitative data for two or three regions of each lung (i.e., upper, middle, lower). Indeed, in the literature, most studies have focused on some or all of the following four lung regions: left upper lung, left lower lung, right upper lung, and right lower lung [35], [40], [41], [73], [98].

Airflow-dependent analysis of both normal RS and CAS

The proposed approach to RS analysis also includes the performance of a progressive respiratory manoeuvre with variable airflow that allows both normal RS and CAS to be analysed with respect to airflow (Figure 11). This is a real advantage as compared to previous studies, in which RS were recorded

at constant airflows or during forced expiratory manoeuvres. However, it has been widely reported that normal RS intensity is highly airflow-dependent (Section 1.4.1). Accordingly, we propose to use the RS intensity–airflow curves for the analysis of normal RS intensity.

Regarding CAS, it has been reported that forced expiratory wheezes have low specificity for the clinical diagnosis of asthma, since many wheezes may also appear in healthy people during forced expiratory manoeuvres [72], [180]. Further, in the case of CAS, an airflow-dependent analysis is important because CAS only appear above a critical flow, and this depends on the mechanical properties of airways, which in turn vary between people (Section 1.3.2). Therefore, recording RS at constant airflows may lead to CAS underestimation. In this sense, the proposed progressive respiratory manoeuvre allows CAS to be analysed over a wide range of airflows, from normal to forced breathing.

Analysis of CAS based on HHT

In this thesis, we propose a customised HHT adapted to RS signal characteristics as a practical and direct way of calculating IF in RS signals (Section 1.5.4.2). We also propose a comprehensive evaluation of the performance of HHT for RS analysis, and especially for CAS characterisation. The main reason for analysing CAS based on IF is that characterising CAS implying calculating both the duration and the pitch of CAS, such that RS signals containing CAS must be analysed in both time and frequency domains. In this sense, the IF parameter provides information about the frequency content of a signal at each time instant—that is, with maximum time resolution.

Since calculating IF is feasible only for monocomponent or narrowband signals, RS signals must be decomposed into narrowband components before IF can be calculated for each component. To do this, HHT makes use of EMD (Section 1.5.4.2). However, it has been shown that EMD has a mode mixing effect that may cause poor separation of frequency scales in some RS signals (Chapter 4), thus leading to inaccurate IF estimates.

We propose using EEMD as the first step of HHT, prior to IF estimation, to avoid the mode mixing effect of EMD in RS signals. The effectiveness of EEMD in reducing mode mixing is determined by two parameters: the amplitude of the added noise and the number of iterations (Chapter 4). Indeed, EEMD compromises between the two parameters, so that the mode mixing effect is minimised and the residual noise level is low. In this regard, we propose two parameters to allow the amount of mode mixing ($FO_{i,j}$) and the residual noise level (*PSDR*) to be measured (Chapter 4); the $FO_{i,j}$ parameter, which measures the frequency overlap between the PSDs of two different IMFs; and the *PSDR* parameter, which measures the residual noise level in the IMFs by comparing the PSD of the reconstructed RS signal—the direct sum of the IMFs and the residue—to the PSD of the original RS signal. Using these parameters, the amplitude of the added noise and the number of iterations of EEMD could be adjusted, so that $FO_{i,j}$ and *PSDR* were minimised at the same time. These parameters represent a step forward in the use of EEMD, since the procedure for the correct choice of the amplitude of the

added noise and the number of iterations is unclear in the literature, although an EEMD-based HHT has already been used in other applications [190], [191].

After decomposing an RS signal into IMFs by EEMD, the IF and IE sequences can be easily estimated by the phase derivative and absolute value of the analytic signal of each IMF. We propose different methods for calculating the phase derivative in this thesis.

As described in Chapter 3, a five-point least-squares polynomial derivative approximation was initially proposed for calculating the IF. This method yielded IF estimates with high variance, which contributed to emphasizing the differences in IF dispersion between RS signal segments containing CAS and those containing normal RS, as described in Section 1.5.4. By taking advantage of this property of IF, we were able to detect those segments within RS signals that were more likely to contain CAS, by applying a set of thresholds to the dispersion of IF sequences. Each detected RS signal segment was characterized using a set of features extracted from the IF and IE sequences. Those features were then used to train and test a support vector machine classifier, which succeeded in achieving high sensitivity (93.9%) and accuracy (94.9%) in differentiating RS signal segments containing CAS from those containing normal RS.

After differentiating CAS from normal RS within RS signals, our next step was to characterise CAS. In this thesis, we propose a new method for the automatic segmentation and characterisation of CAS from HS (Chapter 4). HS is obtained by rearranging the IF and IE sequences of a signal into an array. However, the high variance of the IF estimates used to detect CAS within RS signals was a major drawback for obtaining an accurate HS with high energy concentration, which is a desirable property of TFDs. Therefore, the IF sequences of RS signals were recalculated using the Kay's IF estimator, which greatly reduces the variance of the IF estimates. Even so, the IF features used for distinguishing CAS from normal RS (in Chapter 3) were used to locate CAS components in the time-frequency plane of the HS (in Chapter 4), which facilitated the subsequent segmentation of CAS by using a region growing algorithm together with a set of region linking criteria.

Using the proposed EEMD–Kay-based HS, we were able to fully characterise CAS with regard to duration, mean frequency, and intensity. Moreover, since this improved version of HS has high temporal resolution and high frequency resolution, and achieves high concentrations of energy, it allows CAS to be more accurately detected and characterised mainly based on IF criteria. This provides a stark contrast to traditional techniques, such as spectrogram, which are highly dependent on amplitude criteria and fail to detect some CAS with low intensity (Chapter 4).

Assessment of pulmonary function in patients with UPP

In patients with UPP (Chapter 2), the RS intensity–airflow curves have been proposed as an indirect way of detecting differences in airflow entering the lungs. As reported in the literature, reduction in regional ventilation of the lung on the paralysed side is more evident at the base of the lung in UPP

[129], [130]. Accordingly, we have shown that two microphones placed at the right and left posterior base of the lungs proved sufficient to analyse the differences in regional ventilation of the lungs in patients with UPP.

We analysed the RS signals of 10 patients with UPP and 10 healthy participants for airflows ranging from 1.2 to 2.4 L/s. We found that lung sound intensity—calculated as the area under the PSD of inspiratory lung sounds in the frequency band of 70–2000 Hz—was significantly reduced on the affected side for all airflows due to inefficient diaphragmatic muscle function. We expressed the lung sound intensity in dB using as a reference value the intensity of background noise, which was calculated for each microphone from the apnoea segments. A clear cut-off around 7 dB was found in the mean differences of lung sound intensity between the two sides in healthy participants and patients with UPP. Moreover, we found that the RS intensity–airflow curves can be properly expressed by a linear equation in both patients and healthy participants. Further, there was a clear difference between the slopes of healthy participants (around 5 dB/L/s) and patients (around 10 dB/L/s), independent of the analysed side.

Besides lung sound intensity, spirometric parameters were also measured in patients with UPP, as these are low with respect to normal reference values, as described in previous studies [139], [149], [161], [162].

Assessment of BDR in patients with asthma

Recording respiratory symptoms, such as wheezing, has been reported to contribute to improving the interpretation of pulmonary function tests, such as spirometry, for asthma [7]. Accordingly, in this thesis, we propose the analysis of both normal RS and CAS to assess BDR in patients with asthma. RS signals were recorded, and BDR was assessed by spirometry, in 50 patients with asthma and 10 healthy participants.

In a preliminary study (Chapter 5), we analysed the RS signals of 3 patients with asthma. Only two of the five proposed locations for recording RS were used, and the number—but not the features—of CAS was analysed. Despite these limitations, the study demonstrated the potential of RS analysis for assessing BDR in asthma, since appreciable increases in the RS intensity–airflow curves following bronchodilator administration could be detected in a patient with negative BDR in FEV_1. Furthermore, these increases were more relevant at high airflow levels.

We have also carried out a study to evaluate the potential of CAS analysis for assessing BDR in 10 patients with asthma and 5 healthy participants. Note that as we have recently submitted this study to PLOS ONE (see publications derived from this thesis), it is not been included in the list of published articles. In this recent study, we determined that analysing CAS features, including number, duration, pitch, and intensity, could improve the stratification of BDR levels, particularly for patients with asthma with negative BDR in spirometric parameters. In fact, we were able to detect appreciable changes in the

number of CAS in three of five patients with negative BDR in FEV_1, revealing changes in AO. Appreciable changes in the number of CAS were also found in all patients with positive BDR in FEV_1, whereas CAS were absent in the control participants. Further, duration, mean frequency, and intensity of CAS significantly changed following bronchodilator administration for two of five patients with negative BDR in FEV_1. The number of CAS increased with airflow, indicating that the greater the airflow, the higher the possibility of CAS appearing. Even so, CAS appeared over a wide range of airflow levels. Moreover, most significant changes in CAS features after bronchodilation occurred at medium-high airflows.

In our recent study, we have also proposed a classifier for distinguishing CAS from other sounds segmented from HS. This classifier reached high accuracy (97.7%), sensibility (93.1%), and specificity (98.5%), making our approach more robust with respect to background noises in the clinical environment.

Chapter 7: Conclusions

In this thesis, we have proposed a novel, comprehensive, and non-invasive approach to RS analysis, taking into consideration the following four key aspects: the heterogeneous distribution of RS (using multichannel RS recording), the airflow dependence of RS (using the progressive respiratory manoeuvre and the RS intensity–airflow curves), the different characteristics of RS types (analysing CAS based on HHT and IF criteria), and the effect of respiratory diseases on RS (analysing both normal RS and CAS). To the best of our knowledge, an approach that considers all these aspects of RS analysis has not been proposed to date. In this sense, our approach to RS analysis is a major novelty of this thesis.

Two clinical applications of our approach to RS analysis have been proposed: the assessment of pulmonary function in patients with UPP, and the assessment of BDR in patients with asthma. The quantitative analysis of RS has not been previously used in the assessment of UPP. Regarding BDR in asthma, previous approaches did not consider all the aforementioned aspects together.

The following main conclusions have been drawn from the published articles included in this thesis:

RS recoding protocol

- The configuration for recording RS proposed in this thesis meets the CORSA guidelines [58] and includes the minimum number of sensors required for analysing the distribution of lung sounds both laterally and vertically, without requiring a large number of sensors and the resulting larger amount of data
- Two RS recording protocols were launched in the Pulmonary Function Testing Laboratory of the HUGTiP. Patients with either UPP or asthma were recruited during the course of this thesis, so that two databases were created:
    - 10 patients and 10 healthy participants, who took part in the UPP study
    - 50 patients and 10 healthy participants, who took part in the asthma study

Analysis of normal RS

- The proposed progressive respiratory manoeuvre allows the characteristic RS intensity–airflow curves of a subject to be calculated for its entire airflow range
- There is a linear relationship between airflow and normal RS intensity calculated in the frequency band of 70–2000 Hz of the PSD, and expressed in dB with respect to the background noise level
- The RS intensity–airflow curves contribute to improving the understanding of the mechanics behind RS generation and how structural and functional respiratory defects affect RS

Analysis of CAS

- An EEMD-based HHT has been proposed to decompose RS signals into IMFs, for which a physically meaningful IF sequences can be calculated point-by-point, thereby avoiding the mode mixing effect of the original EMD. The IF sequences provide information about the frequency content of RS signals at each time instant
- A number of quantitative parameters have been proposed to quantify the size, reduction of mode mixing, and residual noise level of EEMD in RS signals. These parameters can be used to select the EEMD parameters, not only for RS analysis but also for other applications
- In the proposed EEMD-based HHT, IF is calculated via phase differentiation of the analytic signal of the IMFs obtained by EEMD. Therefore, HHT does not suffer from the uncertainty principle and provides both high time resolution and high frequency resolution at the same time
- Since the IE and IF sequences are calculated separately for each IMF, HHT allows working independently either in a time-frequency or in a time-energy domain, without having to process an entire TFD
- IF dispersion markedly decreases within RS signals for the duration of CAS. Therefore, an algorithm for the automatic differentiation between RS signal segments containing CAS and those containing normal RS has been proposed mainly based on IF dispersion criteria
- An EEMD–Kay-based HS has been proposed as an alternative and precise time-frequency representation for CAS analysis
- The proposed EEMD–Kay-based HS provides a high concentration of RS energy around IF in the time-frequency plane, which makes HS less dependent on amplitude criteria and facilitates CAS identification. This property, together with the high resolution in time and frequency domains, allows CAS to be more accurately determined with HS than with spectrogram, especially for weak CAS that overlap with normal RS in the time-frequency plane

Assessment of pulmonary function in patients with UPP

- Lung sound intensity is significantly reduced on the affected side for all measured airflows in patients with UPP. Therefore, in these patients, lung sound intensity provides quantitative information about the extent of impairment of respiratory function
- The analysis of lung sound intensity offers a method for long-term monitoring of respiratory function recovery in patients undergoing physical therapy, in a non-obtrusive and cost-effective manner

Assessment of BDR in patients with asthma

- The proposed progressive respiratory manoeuvre, which is easier to perform than the forced expiratory manoeuvre, allows the analysis of both CAS that appear naturally at low airflows and CAS that appear in a more forced way at high airflows

- The combined analysis of normal RS and CAS provides a promising approach to assessing BDR in patients with asthma. It is a simple and non-invasive technique, which is more sensitive to local changes in AO not detectable by spirometry, thus increasing the sensitivity of the BDR assessment

7.1. Future work

Despite major advances in sensors and RS analysis, the field still is suffering from a lack of a standard methodology for recording and processing RS signals. Accordingly, these techniques are still used only limitedly in clinical practice.

The next steps for the proposed approach to RS analysis require information from further studies in large populations, to clinically validate this technique as a new complementary tool for assessing pulmonary function in respiratory diseases. In particular, relevant clinical conclusions could be drawn by applying the proposed approach to assessing BDR to the whole database of 50 patients with asthma and 10 healthy participants.

In CAS analysis, the differences between pure low-pitched wheezes and rhonchus (snoring character), between monophonic harmonic wheezes and polyphonic wheezes, or between squawk and high-pitched wheezes, are still unclear [17]. Therefore, further studies should address in detail the characterisation of different types of CAS.

Besides the number and features of CAS and normal RS intensity, the spatial distribution of the two types of RS should be further analysed, by taking advantage of the multichannel RS recording using five microphones as developed here. The distribution of normal RS intensity and CAS over the chest surface could provide information about location of AO and changes in regional ventilation following bronchodilator administration in patients with asthma.

The effect of body position on RS in patients with UPP should be also analysed. Although RS have been analysed in the sitting position in these patients, extending RS analysis to different positions may help to better understand the effect of UPP on respiratory mechanics and pulmonary ventilation.

Thus, the proposed approach to RS analysis provides a sensitive tool for obtaining complementary, objective, and quantitative information about pulmonary function in respiratory diseases. Together with spirometry, RS analysis has some direct clinical applications, as discussed in this thesis. This proposed approach could be included in a system of RS analysis that could serve not only for diagnosis of patients with respiratory diseases but also for their long-term monitoring, without risk to the patient. This type of system could be used in the daily clinical routine, complementing the spirometry. The combined information from these techniques would be predicted to increase the reliability of diagnosis and drastically improve management of patients with respiratory diseases.

Bibliography

[1] J. B. West, *Respiratory Physiology: The Essentials*, 8th ed. Wolters Kluwer Health: Lippincott Williams & Wilkins, 2008.

[2] V. C. Scanlon and T. Sanders, *Essentials of Anatomy and Physiology*, 5th ed. F. A. Davis Company, 2006.

[3] C. L. VanPutte, J. L. Regan, and A. F. Russo, *Seeley's Anatomy & Physiology*, 10th ed. McGraw-Hill, 2013.

[4] T. R. Gildea and K. McCarthy, "Pulmonary function testing," in *Current Clinical Medicine*, 2nd ed. Saunders Elsevier, 2010, ch. 172, pp. 1084–1095.

[5] J. Roca, F. Burgos, J. Sunyer, M. Saez, S. Chinn, J. M. Antó, R. Rodríguez-Roisin, P. H. Quanjer, D. Nowak, and P. Burney, "References values for forced spirometry. Group of the European Community Respiratory Health Survey," *Eur. Respir. J.*, vol. 11, no. 6, pp. 1354–1362, Jun. 1998.

[6] Global Initiativa for Asthma. Global Strategy for Asthma Management and Prevention. 2015. Available from: www.ginasthma.org

[7] R. Pellegrino, G. Viegi, V. Brusasco, R. O. Crapo, F. Burgos, R. Casaburi, A. Coates, C. P. Van der Grinten, P. Gustafsson, J. Hankinson, R. Jensen, D. C. Johnson, N. MacIntyre, R. McKay, M. R. Miller, D. Navajas, O. F. Pedersen, and J. Wanger, "Interpretative strategies for lung function tests," *Eur. Respir. J.*, vol. 26, no. 5, pp. 948–968, Nov. 2005.

[8] B. R. Celli, "The importance of spirometry in COPD and asthma: effect on approach to management," *Chest*, vol. 117, no. 2 Suppl, pp. 15S–19S, Feb. 2000.

[9] M. J. Hegewald, R. G. Townsend, J. T. Abbott, and R. O. Crapo, "Bronchodilator response in patients with normal baseline spirometry," *Respir. Care*, vol. 57, no. 10, pp. 1564–1570, Oct. 2012.

[10] D. P. Hsu, T. F. Ocampo, H. A. DiGiovanni, and E. R. Gil, "Evaluation of interpretation strategies and substantial bronchodilator response in pediatric patients with normal baseline spirometry," *Respir. Care*, vol. 58, no. 5, pp. 785–789, May. 2013.

[11] J. A. Gjevre, T. S. Hurst, R. M. Taylor-Gjevre, and D. W. Cockcroft, "The American Thoracic Society's spirometric criteria alone is inadequate in asthma diagnosis," *Can. Respir. J.*, vol. 13, no. 8, pp. 433–437, Nov-Dec. 2006.

[12] S. M. Tse, D. R. Gold, J. E. Sordillo, E. B. Hoffman, M. W. Gillman, S. L. Rifas-Shiman, A. L. Fuhlbrigge, K. G. Tantisira, S. T. Weiss, and A. A. Litonjua, "Diagnostic accuracy of the bronchodilator response in children," *J. Allergy Clin. Immunol.*, vol. 132, no. 3, pp. 554–559, Sep. 2013.

[13] D. C. Richter, J. R. Joubert, H. Nell, M. M. Schuurmans, and E. M. Irusen, "Diagnostic value of post-bronchodilator pulmonary function testing to distinguish between stable, moderate to severe COPD and asthma," *Int. J. Chron. Obstruct. Pulmon. Dis.*, vol. 3, no. 4, pp. 693–699, Dec. 2008.

[14] J. S. Coviello, *Auscultation Skills: Breath & Heart Sounds*, 5th ed. Wolters Kluwer Health: Lippincott Williams & Wilkins, 2013.

[15] A. Bohadana, G. Izbicki, and S. S. Kraman, "Fundamentals of lung auscultation," *N. Engl. J. Med.*, vol. 370, no. 8, pp. 744–751, Feb. 2014.

[16] R. Mikami, M. Murao, D. W. Cugell, J. Chretien, P. Cole, J. Meier-Sydow, R. L. Murphy, and R. G. Loudon, "International Symposium on Lung Sounds. Synopsis of proceedings," *Chest*, vol. 92, no. 2, pp. 342–345, Aug. 1987.

[17] H. Pasterkamp, P. L. Brand, M. Everard, L. Garcia-Marcos, H. Melbye, and K. N. Priftis, "Towards the standardisation of lung sound nomenclature," *Eur. Respir. J.*, vol. 47, no. 3, pp. 724–732, Mar. 2016.

[18] ATS-ACCP Ad Hoc Subcommittee, "Report on pulmonary nomenclature," *ATS News*, vol. 3, pp. 5–6, 1977.

[19] A. R. A. Sovijärvi, F. Dalmasso, J. Vanderschoot, L. P. Malmberg, G. Righini, and S. A. T. Stoneman, "Definition of terms for applications of respiratory sounds," *Eur. Respir. Rev.*, vol. 10, no. 77, pp. 597–610, 2000.

[20] A. R. A. Sovijärvi, L. P. Malmberg, G. Charbonneau, J. Vanderschoot, F. Dalmasso, C. Sacco, M. Rossi, and J. E. Earis, "Characteristics of breath sounds and adventitious respiratory sounds," *Eur. Respir. Rev.*, vol. 10, no. 77, pp. 591–596, 2000.

[21] N. Gavriely, M. Nissan, A. H. Rubin, and D. W. Cugell, "Spectral characteristics of chest-wall breath sounds in normal subjects," *Thorax*, vol. 50, no. 12, pp. 1292–1300, Dec. 1995.

[22] V. Gross, A. Dittmar, T. Penzel, F. Schüttler, and P. von Wichert, "The relationship between normal lung sounds, age, and gender," *Am. J. Respir. Crit. Care Med.*, vol. 162, no. 3 Pt 1, pp. 905–909, Sep. 2000.

[23] A. Oliveira and A. Marques, "Respiratory sounds in healthy people: a systematic review," *Respir. Med.*, vol. 108, no. 4, pp. 550–570, Apr. 2014.

[24] P. Forgacs, "The functional basis of pulmonary sounds," *Chest*, vol. 73, no. 3, pp. 399–405, Mar. 1978.

[25] H. Pasterkamp and I. Sanchez, "Effect of gas density on respiratory sounds," *Am. J. Respir. Crit. Care Med.*, vol. 153, no. 3, pp. 1087–1092, Mar. 1996.

[26] H. Pasterkamp, S. S. Kraman, and G. R. Wodicka, “Respiratory sounds. Advances beyond the stethoscope,” *Am. J. Respir. Crit. Care Med.*, vol. 156, no. 3 Pt 1, pp. 974–987, Sep. 1997.
[27] R. G. Loudon and R. L. Murphy, “Lung sounds,” in *The Lung: Scientific Foundations*, 2nd ed. Lippincott-Raven, 1997.
[28] S. S. Kraman, “Determination of the site of production of respiratory sounds by subtraction phonopneumography,” *Am. Rev. Respir. Dis.*, vol. 122, no. 2, pp. 303–309, Aug. 1980.
[29] R. G. Loudon and R. L. Murphy, “Lund sounds,” *Am. Rev. Respir. Dis.*, vol. 130, no. 4, pp. 663–673, Oct. 1984.
[30] F. Dalmay, M. T. Antonini, P. Marquet, and R. Menier, “Acoustic properties of the normal chest,” *Eur. Respir. J.*, vol. 8, no. 10, pp. 1761–1769, Oct. 1995.
[31] J. E. Earis and B. M. G. Cheetham, “Future perspectives for respiratory sound research,” *Eur. Respir. Rev.*, vol. 10, no. 77, pp. 641–646, 2000.
[32] G. R. Wodicka, K. N. Stevens, H. L. Golub, and D. C. Shannon, “Spectral characteristics of sound transmission in the human respiratory system,” *IEEE Trans. Biomed. Eng.*, vol. 37, no. 12, pp. 1130–1135, Dec. 1990.
[33] G. R. Wodicka, K. N. Stevens, H. L. Golub, E. G. Cravalho, and D. C. Shannon, “A model of acoustic transmission in the respiratory system,” *IEEE Trans. Biomed. Eng.*, vol. 36, no. 9, pp. 925–934, Sep. 1989.
[34] V. Goncharoff, J. E. Jacobs, and D. W. Cugell, “Wideband acoustic transmission of human lungs,” *Med. Biol. Eng. Comput.*, vol. 27, no. 5, pp. 513–519, Sep. 1989.
[35] S. S. Kraman and A. B. Bohadana, “Transmission to the chest of sound introduced at the mouth,” *J. Appl. Physiol.*, vol. 66, no. 1, pp. 278–281, Jan. 1989.
[36] M. Kompis, H. Pasterkamp, Y. Oh, and G. Wodicka, “Distribution of inspiratory and expiratory respiratory sound intensity on the surface of the human thorax,” in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, pp. 2047–2050, 1997.
[37] D. M. O’Donnell and S. S. Kraman, “Vesicular lung sound amplitude mapping by automated flow-gated phonopneumography,” *J. Appl. Physiol.*, vol. 53, no. 3, pp. 603–609, Oct. 1982.
[38] H. Pasterkamp, S. Patel, and G. R. Wodicka, “Asymmetry of respiratory sounds and thoracic transmission,” *Med. Biol. Eng. Comput.*, vol. 35, no. 2, pp. 103–106, Mar. 1997.
[39] T. R. Fenton, H. Pasterkamp, A. Tal, and V. Chernick, “Automated spectral characterization of wheezing in asthmatic children,” *IEEE Trans. Biomed. Eng.*, vol. 32, no. 1, pp. 50–55, Jan. 1985.
[40] J. A. Fiz, J. Gnitecki, S. S. Kraman, G. R. Wodicka, and H. Pasterkamp, “Effect of body position on lung sounds in healthy young men,” *Chest*, vol. 133, no. 3, pp. 729–736, Mar. 2008.
[41] A. Torres-Jimenez, S. Charleston-Villalobos, R. Gonzalez-Camarena, G. Chi-Lem, and T. Aljama-Corrales, “Asymmetry in lung sound intensities detected by respiratory acoustic thoracic imaging (RATHI) and clinical pulmonary auscultation,” in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, pp. 4797–4800, 2008.
[42] R. P. Dellinger, J. E. Parrillo, A. Kushnir, M. Rossi, and I. Kushnir, “Dynamic visualization of lung sounds with a vibration response device: a case series,” *Respiration*, vol. 75, no. 1, pp. 60–72, Feb. 2008.
[43] R. L. Murphy, “In defense of the stethoscope,” *Respir. Care*, vol. 53, no. 3, pp. 355–369, Mar. 2008.
[44] R. L. Murphy, “Computerized multichannel lung sound analysis,” *IEEE Eng. Med. Biol. Mag.*, vol. 26, no. 1, pp. 16–19, Jan-Feb. 2007.
[45] S. M. Akramus Salehin and T. D. Abhayapala, “Lung sound localization using array of acoustic sensors,” in *Int. Conf. Signal Process. Commun. Syst.*, pp. 1–5, 2008.
[46] S. Charleston-Villalobos, S. Cortés-Rubiano, R. González-Camarena, G. Chi-Lem, and T. Aljama-Corrales, “Respiratory acoustic thoracic imaging (RATHI): assessing deterministic interpolation techniques,” *Med. Biol. Eng. Comput.*, vol. 42, no. 5, pp. 618–626, Sep. 2004.
[47] R. González-Camarena, S. Charleston-Villalobos, A. Angeles-Olguín, and T. Aljama-Corrales, “Imaging the thoracic distribution of normal breath sounds,” *Methods Inf. Med.*, vol. 49, no. 5, pp. 443–447, Jun. 2010.
[48] M. Yigla, M. Gat, J. J. Meyer, P. J. Friedman, T. M. Maher, and J. M. Madison, “Vibration response imaging technology in healthy subjects,” *Am. J. Roentgenol.*, vol. 191, no. 3, pp. 845–852, Sep. 2008.
[49] A. B. Bohadana, R. Peslin, and H. Uffholtz, “Breath sounds in the clinical assessment of airflow obstruction,” *Thorax*, vol. 33, no. 3, pp. 345–351, Jun. 1978.
[50] R. Mor, I. Kushnir, J. J. Meyer, J. Ekstein, and I. Ben-Dov, “Breath sound distribution images of patients with pneumonia and pleural effusion,” *Respir. Care*, vol. 52, no. 12, pp. 1753–1760, Dec. 2007.
[51] J. B. Grotberg and N. Gavriely, “Flutter in collapsible tubes: a theoretical model of wheezes,” *J. Appl. Physiol.*, vol. 66, no. 5, pp. 2262–2273, May. 1989.
[52] R. Paciej, A. Vyshedskiy, D. Bana, and R. Murphy, “Squawks in pneumonia,” *Thorax*, vol. 59, no. 2, pp. 177–178, Feb. 2004.
[53] J. E. Earis, K. Marsh, M. G. Pearson, and C. M. Ogilvie, “The inspiratory ‘squawk’ in extrinsic allergic alveolitis and other pulmonary fibroses,” *Thorax*, vol. 37, no. 12, pp. 923–926, Dec. 1982.

[54] R. P. Baughman and R. G. Loudon, “Stridor: differentiation from asthma or upper airway noise,” *Am. Rev. Respir. Dis.*, vol. 139, no. 6, pp. 1407–1409, Jun. 1989.
[55] S. S. Kraman, P. Harper, H. Pasterkamp, and G. R. Wodicka, “‘Slide whistle’ breath sounds: acoustical correlates of variable tracheal obstruction,” *Physiol. Meas.*, vol. 23, no. 2, pp. 449–455, May. 2002.
[56] A. Vyshedskiy, R. M. Alhashem, R. Paciej, M. Ebril, I. Rudman, J. J. Fredberg, and R. Murphy, “Mechanism of inspiratory and expiratory crackles,” *Chest*, vol. 135, no. 1, pp. 156–164, Jan. 2009.
[57] R. T. H. Laennec, *De l’auscultation médiate, ou traité du diagnostic des maladies des poumons et du coeur: fondé principalement sur ce nouveau moyen d'exploration*. Paris: Brosson & Chaudé, 1819.
[58] M. Rossi, A. R. A. Sovijärvi, P. Piirilä, L. Vannuccini, F. Dalmasso, and J. Vanderschoot, “Environmental and subject conditions and breathing manoeuvres for respiratory sound recordings,” *Eur. Respir. Rev.*, vol. 10, no. 77, pp. 611–615, 2000.
[59] L. Vannuccini, J. E. Earis, P. Helistö, B. M. G. Cheetham, M. Rossi, A. R. A. Sovijärvi, and J. Vanderschoot, “Capturing and preprocessing of respiratory sounds,” *Eur. Respir. Rev.*, vol. 10, no. 77, pp. 616–620, 2000.
[60] B. M. G. Cheetham, G. Charbonneau, A. Giordano, P. Helistö, and J. Vanderschoot, “Digitization of data for respiratory sound recordings,” *Eur. Respir. Rev.*, vol. 10, no. 77, pp. 621–624, 2000.
[61] P. Leblanc, P. T. Macklem, and W. R. Ross, “Breath sounds and distribution of pulmonary ventilation,” *Am. Rev. Respir. Dis.*, vol. 102, no. 1, pp. 10–16, Jul. 1970.
[62] S. S. Kraman, “The relationship between airflow and lung sound amplitude in normal subjects,” *Chest*, vol. 86, no. 2, pp. 225–229, Aug. 1984.
[63] B. E. Shykoff, Y. Ploysongsang, and H. K. Chang, “Airflow and normal lung sounds,” *Am. Rev. Respir. Dis.*, vol. 137, no. 4, pp. 872–876, Apr. 1988.
[64] N. Gavriely and D. W. Cugell, “Airflow effects on amplitude and spectral content of normal breath sounds,” *J. Appl. Physiol.*, vol. 80, no. 1, pp. 5–13, Jan. 1996.
[65] I. Hossain and Z. Moussavi, “Finding the lung sound-flow relationship in normal and asthmatic subjects,” in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, pp. 3852–3855, 2004.
[66] M. Yosef, R. Langer, S. Lev, and Y. A. Glickman, “Effect of airflow rate on vibration response imaging in normal lungs,” *Open Respir. Med. J.*, vol. 3, pp. 116–122, Sep. 2009.
[67] C. L. Que, C. Kolmaga, L. G. Durand, S. M. Kelly, and P. T. Macklem, “Phonospirometry for noninvasive measurement of ventilation: methodology and preliminary results,” *J. Appl. Physiol.*, vol. 93, no. 4, pp. 1515–1526, Oct. 2002.
[68] A. Yadollahi and Z. Moussavi, “A robust method for estimating respiratory flow using tracheal sounds entropy,” *IEEE Trans. Biomed. Eng.*, vol. 53, no. 4, pp. 662–668, Apr. 2006.
[69] S. Huq and Z. Moussavi, “Acoustic breath-phase detection using tracheal breath sounds,” *Med. Biol. Eng. Comput.*, vol. 50, no. 3, pp. 297–308, Mar. 2012.
[70] A. Yadollahi, A. Montazeri, A. Azarbarzin, and Z. Moussavi, “Respiratory flow-sound relationship during both wakefulness and sleep and its variation in relation to sleep apnea,” *Ann. Biomed. Eng.*, vol. 41, no. 3, pp. 537–546, Mar. 2013.
[71] R. P. Baughman and R. G. Loudon, “Quantitation of wheezing in acute asthma,” *Chest*, vol. 86, no. 5, pp. 718–722, Nov. 1984.
[72] H. J. Schreur, J. Vanderschoot, A. H. Zwinderman, J. H. Dijkman, and P. J. Sterk, “Abnormal lung sounds in patients with asthma during episodes with normal lung function,” *Chest*, vol. 106, no. 1, pp. 91–99, Jul. 1994.
[73] N. Gavriely, M. Nissan, D. W. Cugell, and A. H. Rubin, “Respiratory health screening using pulmonary function tests and lung sound analysis,” *Eur. Respir. J.*, vol. 7, no. 1, pp. 35–42, Jan. 1994.
[74] S. Rietveld, E. H. Dooijes, L. H. Rijssenbeek-Nouwens, F. Smit, P. J. Prins, A. M. Kolk, and W. A. Everaerd, “Characteristics of wheeze during histamine-induced airways obstruction in children with asthma,” *Thorax*, vol. 50, no. 2, pp. 143–148, Feb. 1995.
[75] A. B. Bohadana, R. Peslin, H. Uffholtz, and G. Pauli, “Potential for lung sound monitoring during bronchial provocation testing,” *Thorax*, vol. 50, no. 9, pp. 955–961, Sep. 1995.
[76] H. J. Schreur, Z. Diamant, J. Vanderschoot, A. H. Zwinderman, J. H. Dijkman, and P. J. Sterk, “Lung sounds during allergen-induced asthmatic responses in patients with asthma,” *Am. J. Respir. Crit. Care Med.*, vol. 153, no. 5, pp. 1474–1480, May. 1996.
[77] J. A. Fiz, R. Jané, D. Salvatella, J. Izquierdo, L. Lores, P. Caminal, and J. Morera, “Analysis of tracheal sounds during forced exhalation in asthma patients and normal subjects: bronchodilator response effect,” *Chest*, vol. 116, no. 3, pp. 633–638, Sep. 1999.
[78] L. Bentur, R. Beck, M. Shinawi, T. Naveh, and N. Gavriely, “Wheeze monitoring in children for assessment of nocturnal asthma and response to therapy,” *Eur. Respir. J.*, vol. 21, no. 4, pp. 621–626, Apr. 2003.
[79] I. Sánchez, C. Vizcaya, D. García, and E. Campos, “Response to bronchodilator in infants with bronchiolitis can be predicted from wheeze characteristics,” *Respirology*, vol. 10, no. 5, pp. 603–608, Nov.

2005.
[80] S. A. Taplidou, L. J. Hadjileontiadis, T. Penzel, V. Gross, and S. M. Panas, “WED: An efficient wheezing-episode detector based on breath sounds spectrogram analysis,” in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, pp. 2531–2534, 2003.
[81] A. Homs-Corbera, J. A. Fiz, J. Morera, and R. Jané, “Time-frequency detection and analysis of wheezes during forced exhalation,” *IEEE Trans. Biomed. Eng.*, vol. 51, no. 1, pp. 182–186, Jan. 2004.
[82] M. L. Hsueh, J. C. Chien, F. C. Chang, H. D. Wu, and F. C. Chong, “Respiratory wheeze detection system,” in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, pp. 7553–7559, 2005.
[83] B. S. Lin, B. S. Lin, H. D. Wu, F. C. Chong, and S. J. Chen, “Wheeze recognition based on 2D bilateral filtering of spectrogram,” *Biomed. Eng. Appl. Basis Commun.*, vol. 18, no. 3, pp. 128–137, Jun. 2006.
[84] S. A. Taplidou and L. J. Hadjileontiadis, “Wheeze detection based on time-frequency analysis of breath sounds,” *Comput. Biol. Med.*, vol. 37, no. 8, pp. 1073–1083, Aug. 2007.
[85] A. Jain and J. Vepa, “Lung sound analysis for wheeze episode detection,” in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, pp. 2582–2585, 2008.
[86] J. Zhang, W. Ser, J. Yu, and T. T. Zhang, “A novel wheeze detection method for wearable monitoring systems,” in *Int. Symp. Intell. Ubiquitous Comput. Educ.*, pp. 331–334, 2009.
[87] R. J. Riella, P. Nohama, and J. M. Maia, “Method for automatic detection of wheezing in lung sounds,” *Braz. J. Med. Biol. Res.*, vol. 42, no. 7, pp. 674–684, Jul. 2009.
[88] S. A. Taplidou, L. J. Hadjileontiadis, I. K. Kitsas, K. I. Panoulas, T. Penzel, V. Gross, and S. M. Panas, “On applying continuous wavelet transform in wheeze analysis,” in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, pp. 3832–3835, 2004.
[89] S. A. Taplidou and L. J. Hadjileontiadis, “Nonlinear analysis of wheezes using wavelet bicoherence,” *Comput. Biol. Med.*, vol. 37, no. 4, pp. 563–570, Apr. 2007.
[90] F. Jin, S. S. Krishnan, and F. Sattar, “Adventitious sounds identification and extraction using temporal-spectral dominance-based features,” *IEEE Trans. Biomed. Eng.*, vol. 58, no. 11, pp. 3078–3087, Nov. 2011.
[91] A. Kandaswamy, C. S. Kumar, R. P. Ramanathan, S. Jayaraman, and N. Malmurugan, “Neural classification of lung sounds using wavelet coefficients,” *Comput. Biol. Med.*, vol. 34, no. 6, pp. 523–537, Sep. 2004.
[92] S. Matsunaga, K. Yamauchi, M. Yamashita, and S. Miyahara, “Classification between normal and abnormal respiratory sounds based on maximum likelihood approach,” in *IEEE Int. Conf. Acoust. Speech Signal Process.*, pp. 517–520, 2009.
[93] S. Xie, F. Jin, S. Krishnan, and F. Sattar, “Signal feature extraction by multi-scale PCA and its application to respiratory sound classification,” *Med. Biol. Eng. Comput.*, vol. 50, no. 7, pp. 759–768, Jul. 2012.
[94] A. Mondal, P. Bhattacharya, and G. Saha, “Detection of lungs status using morphological complexities of respiratory sounds,” *Sci. World J.*, vol. 2014, ID 182938, Feb. 2014.
[95] J. C. Chien, H. D. Wu, F. C. Chong, and C. I. Li, “Wheeze detection using cepstral analysis in Gaussian Mixture Models,” in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, pp. 3168–3171, 2007.
[96] M. Bahoura, “Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes,” *Comput. Biol. Med.*, vol. 39, no. 9, pp. 824–843, Sep. 2009.
[97] S. Aydore, I. Sen, Y. P. Kahya, and M. Mihcak, “Classification of respiratory signals by linear analysis,” in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, pp. 2617–2620, 2009.
[98] H. Pasterkamp, R. Consunji-Araneta, Y. Oh, and J. Holbrow, “Chest surface mapping of lung sounds during methacholine challenge,” *Pediatr. Pulmonol.*, vol. 23, no. 1, pp. 21–30, Jan. 1997.
[99] P. Welch, “The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms,” *IEEE Trans. Audio Electroacoust.*, vol. 15, no. 2, pp. 70–73, Jun. 1967.
[100] B. Boashash, *Time Frequency Signal Analysis and Processing. A Comprehensive Reference*, 1st ed. Elsevier, 2003.
[101] S. Chandra Sekhar and T. V. Sreenivas, “Adaptive spectrogram vs. adaptive pseudo-Wigner–Ville distribution for instantaneous frequency estimation,” *Signal Process.*, vol. 83, no. 7, pp. 1529–1543, Jul. 2003.
[102] J. Zhong and Y. Huang, “Time-frequency representation based on an adaptive short-time Fourier transform,” *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5118–5128, Oct. 2010.
[103] F. Auger and P. Flandrin, “Improving the readability of time-frequency and time-scale representations by the reassignment method,” *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1068–1089, May. 1995.
[104] G. K. Nilsen, “Recursive time-frequency reassignment,” *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3283–3287, Aug. 2009.
[105] F. Auger, P. Flandrin, Y. T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, and H. T. Wu, “Time-frequency reassignment and synchrosqueezing: an overview,” *IEEE Signal Process. Mag.*, vol. 30, no. 6, pp. 32–41, Nov. 2013.

[106] B. Boashash, “Estimating and interpreting the instantaneous frequency of a signal. I. Fundamentals,” *Proc. IEEE*, vol. 80, no. 4, pp. 520–538, Apr. 1992.
[107] N. E. Huang, Z. Wu, S. R. Long, K. C. Arnold, X. Chen, and K. Blank, “On instantaneous frequency,” *Adv. Adapt. Data Anal.*, vol. 1, no. 2, pp. 177–229, Apr. 2009.
[108] D. Gabor, “Theory of communication,” *J. Inst. Electr. Eng. - Part III Radio Commun. Eng.*, vol. 93, no. 26, pp. 429–457, Nov. 1946.
[109] J. Jeong and W. J. Williams, “Kernel design for reduced interference distributions,” *IEEE Trans. Signal Process.*, vol. 40, no. 2, pp. 402–412, Feb. 1992.
[110] L. Rankine, M. Mesbah, and B. Boashash, “IF estimation for multicomponent signals using image processing techniques in the time–frequency domain,” *Signal Process.*, vol. 87, no. 6, pp. 1234–1250, Jun. 2007.
[111] L. Stanković, I. Djurović, S. Stanković, M. Simeunović, S. Djukanović, and M. Daković, “Instantaneous frequency in time–frequency analysis: enhanced concepts and performance of estimation algorithms,” *Digit. Signal Process.*, vol. 35, pp. 1–13, Dec. 2014.
[112] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis,” *Proc. R. Soc. A Math. Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.
[113] G. Rilling, P. Flandrin, and P. Goncalves, “On empirical mode decomposition and its algorithms,” *Proc. IEEE-EURASIP Work. Nonlinear Signal Image Process.*, 2003.
[114] N. E. Huang, X. Chen, M. T. Lo, and Z. Wu, “On Hilbert spectral representation: a true time-frequency representation for nonlinear and nonstationary data,” *Adv. Adapt. Data Anal.*, vol. 3, no. 1n2, pp. 63–93, Apr. 2011.
[115] B. A. Reyes, S. Charleston-Villalobos, T. Aljama-Corrales, and R. González-Camarena, “Time-frequency representations for continuous adventitious lung sounds,” in *V Latin American Congress on Biomedical Engineering CLAIB 2011 May 16-21, 2011, Habana, Cuba*, vol. 33. Springer Berlin Heidelberg, 2013, pp. 1086–1089.
[116] B. A. Reyes, S. Charleston-Villalobos, R. González-Camarena, and T. Aljama-Corrales, “Assessment of time-frequency representation techniques for thoracic sounds analysis,” *Comput. Methods Programs Biomed.*, vol. 114, no. 3, pp. 276–290, May. 2014.
[117] Z. Wu and N. E. Huang, “Ensemble empirical mode decomposition: a noise-assisted data analysis method,” *Adv. Adapt. Data Anal.*, vol. 1, no. 1, pp. 1–41, Jan. 2009.
[118] J. Zhang, R. Yan, R. X. Gao, and Z. Feng, “Performance enhancement of ensemble empirical mode decomposition,” *Mech. Syst. Signal Process.*, vol. 24, no. 7, pp. 2104–2123, Oct. 2010.
[119] P. Flandrin, G. Rilling, and P. Goncalves, “Empirical mode decomposition as a filter bank,” *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 112–114, Feb. 2004.
[120] Z. Wu and N. E. Huang, “A study of the characteristics of white noise using the empirical mode decomposition method,” *Proc. R. Soc. A Math. Phys. Eng. Sci.*, vol. 460, no. 2046, pp. 1597–1611, Jun. 2004.
[121] M. Lozano, J. A. Fiz, and R. Jané, “Automatic differentiation of normal and continuous adventitious respiratory sounds using ensemble empirical mode decomposition and instantaneous frequency,” *IEEE J. Biomed. Health Inform.*, vol. 20, no. 2, pp. 486–497, Mar. 2016.
[122] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer-Verlag, 2006.
[123] D. L. Olson and D. Delen, *Advanced Data Mining Techniques*. Springer-Verlag, 2008.
[124] M. Lozano, J. A. Fiz, and R. Jané, “Performance evaluation of the Hilbert–Huang transform for respiratory sound analysis and its application to continuous adventitious sound characterization,” *Signal Process.*, vol. 120, pp. 99–116, Mar. 2016.
[125] N. Rehman and D. P. Mandic, “Multivariate empirical mode decomposition,” *Proc. R. Soc. A Math. Phys. Eng. Sci.*, vol. 466, pp. 1291–1302, 2010.
[126] N. Rehman and D. P. Mandic, “Filter bank property of multivariate empirical mode decomposition,” *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2421–2426, May. 2011.
[127] B. Boashash, “Estimating and interpreting the instantaneous frequency of a signal. II. Algorithms and applications,” *Proc. IEEE*, vol. 80, no. 4, pp. 540–568, Apr. 1992.
[128] P. Maragos, J. F. Kaiser, and T. F. Quatieri, “Energy separation in signal modulations with application to speech analysis,” *IEEE Trans. Signal Process.*, vol. 41, no. 10, pp. 3024–3051, Oct. 1993.
[129] J. B. Ridyard and R. M. Stewart, “Regional lung function in unilateral diaphragmatic paralysis,” *Thorax*, vol. 31, no. 4, pp. 438–442, Aug. 1976.
[130] G. Leemans, C. Van Holsbeke, J. De Backer, R. Claes, V. Verplancke, L. Vervliet, W. Vos, and W. De Backer, “The influence of diaphragm paralysis on regional ventilation,” in *Eur. Respir. Soc. Ann. Congr.*, 2014.
[131] J. Venegas, “Linking ventilation heterogeneity and airway hyperresponsiveness in asthma,” *Thorax*, vol. 62, no. 8, pp. 653–654, Aug. 2007.

[132] R. Aysola, E. E. de Lange, M. Castro, and T. A. Altes, “Demonstration of the heterogeneous distribution of asthma in the lungs using CT and hyperpolarized helium-3 MRI,” *J. Magn. Reson. Imaging*, vol. 32, no. 6, pp. 1379–1387, Dec. 2010.
[133] W. G. Teague, N. J. Tustison, and T. A. Altes, “Ventilation heterogeneity in asthma,” *J. Asthma*, vol. 51, no. 7, pp. 677–684, Sep. 2014.
[134] J. K. Lui, H. Parameswaran, M. S. Albert, and K. R. Lutchen, “Linking ventilation heterogeneity quantified via hyperpolarized 3He MRI to dynamic lung mechanics and airway hyperresponsiveness,” *PLoS One*, vol. 10, no. 11, e0142738, Nov. 2015.
[135] T. W. Chick and D. P. Nicholson, “The effects of broncholdilators on the distribution of ventilation and perfusion in asthma,” *Chest*, vol. 63, Suppl, pp. 11S–12S, Apr. 1973.
[136] A. R. Sovijärvi, L. Pöyhönen, L. Kellomäki, and A. Muittari, “Effects of acute and long-term bronchodilator treatment on regional lung function in asthma assessed with krypton-81m and technetium-99m-labelled macroaggregates,” *Thorax*, vol. 37, no. 7, pp. 516–520, Jul. 1982.
[137] S. Svenningsen, M. Kirby, D. Starr, D. Leary, A. Wheatley, G. N. Maksym, D. G. McCormack, and G. Parraga, “Hyperpolarized (3) He and (129) Xe MRI: differences in asthma before bronchodilation,” *J. Magn. Reson. Imaging*, vol. 38, no. 6, pp. 1521–1530, Dec. 2013.
[138] G. J. Gibson, “Diaphragmatic paresis: pathophysiology, clinical features, and investigation,” *Thorax*, vol. 44, no. 11, pp. 960–970, Nov. 1989.
[139] F. D. McCool and G. E. Tzelepis, “Dysfunction of the diaphragm,” *N. Engl. J. Med.*, vol. 366, no. 10, pp. 932–942, Mar. 2012.
[140] M. A. DeVita, L. R. Robinson, J. Rehder, B. Hattler, and C. Cohen, “Incidence and natural history of phrenic neuropathy occurring during open heart surgery,” *Chest*, vol. 103, no. 3, pp. 850–856, Mar. 1993.
[141] N. Hart, A. H. Nickol, D. Cramer, S. P. Ward, F. Lofaso, N. B. Pride, J. Moxham, and M. I. Polkey, “Effect of severe isolated unilateral and bilateral diaphragm weakness on exercise performance,” *Am. J. Respir. Crit. Care Med.*, vol. 165, no. 9, pp. 1265–1270, May. 2002.
[142] N. Kharma, “Dysfunction of the diaphragm: imaging as a diagnostic tool,” *Curr. Opin. Pulm. Med.*, vol. 19, no. 4, pp. 394–398, Jul. 2013.
[143] A. Chetta, A. K. Rehman, J. Moxham, D. H. Carr, and M. I. Polkey, “Chest radiography cannot predict diaphragm function,” *Respir. Med.*, vol. 99, no. 1, pp. 39–44, Jan. 2005.
[144] L. C. Yi, O. A. Nascimento, and J. R. Jardim, “Reliability of an analysis method for measuring diaphragm excursion by means of direct visualization with videofluoroscopy,” *Arch. Bronconeumol.*, vol. 47, no. 6, pp. 310–314, Jun. 2011.
[145] E. Gottesman and F. D. McCool, “Ultrasound evaluation of the paralyzed diaphragm,” *Am. J. Respir. Crit. Care Med.*, vol. 155, no. 5, pp. 1570–1574, May. 1997.
[146] E. M. Summerhill, Y. A. El-Sameed, T. J. Glidden, and F. D. McCool, “Monitoring recovery from diaphragm paralysis with ultrasound,” *Chest*, vol. 133, no. 3, pp. 737–743, Mar. 2008.
[147] S. H. Kim, S. Na, J. S. Choi, S. H. Na, S. Shin, and S. O. Koh, “An evaluation of diaphragmatic movement by M-mode sonography as a predictor of pulmonary dysfunction after upper abdominal surgery,” *Anesth. Analg.*, vol. 110, no. 5, pp. 1349–1354, May. 2010.
[148] S. W. Epstein, R. G. Vanderlinden, S. F. Man, R. H. Hyland, S. C. Lenkei, D. J. Wise, and H. Meindok, “Lung function in diaphragm pacing,” *Can. Med. Assoc. J.*, vol. 120, no. 11, pp. 1360–1368, Jun. 1979.
[149] C. M. Laroche, A. K. Mier, J. Moxham, and M. Green, “Diaphragm strength in patients with recent hemidiaphragm paralysis,” *Thorax*, vol. 43, no. 3, pp. 170–174, Mar. 1988.
[150] G. H. Mills, D. Kyroussis, C. H. Hamnegard, S. Wragg, J. Moxham, and M. Green, “Unilateral magnetic stimulation of the phrenic nerve,” *Thorax*, vol. 50, no. 11, pp. 1162–1172, Nov. 1995.
[151] M. I. Polkey, M. Green, and J. Moxham, “Measurement of respiratory muscle strength,” *Thorax*, vol. 50, no. 11, pp. 1131–1135, Nov. 1995.
[152] Y. M. Luo, L. C. Johnson, M. I. Polkey, M. L. Harris, R. A. Lyall, M. Green, and J. Moxham, “Diaphragm electromyogram measured with unilateral magnetic stimulation,” *Eur. Respir. J.*, vol. 13, no. 2, pp. 385–390, Feb. 1999.
[153] Y. M. Luo, M. L. Harris, R. A. Lyall, A. Watson, M. I. Polkey, and J. Moxham, “Assessment of diaphragm paralysis with oesophageal electromyography and unilateral magnetic phrenic nerve stimulation,” *Eur. Respir. J.*, vol. 15, no. 3, pp. 596–599, Mar. 2000.
[154] A. C. Watson, P. D. Hughes, M. Louise Harris, N. Hart, R. J. Ware, J. Wendon, M. Green, and J. Moxham, “Measurement of twitch transdiaphragmatic, esophageal, and endotracheal tube pressure with bilateral anterolateral magnetic phrenic nerve stimulation in patients in the intensive care unit,” *Crit. Care Med.*, vol. 29, no. 7, pp. 1325–1331, Jul. 2001.
[155] J. C. Glerant, N. Mustfa, W. D. Man, Y. M. Luo, G. Rafferty, M. I. Polkey, and J. Moxham, “Diaphragm electromyograms recorded from multiple surface electrodes following magnetic stimulation,” *Eur. Respir. J.*, vol. 27, no. 2, pp. 334–342, Feb. 2006.
[156] J. Steier, S. Kaul, J. Seymour, C. Jolley, G. Rafferty, W. Man, Y. M. Luo, M. Roughton, M. I. Polkey, and

J. Moxham, “The value of multiple tests of respiratory muscle strength,” *Thorax*, vol. 62, no. 11, pp. 975–980, Nov. 2007.

[157] M. Y. Chien, Y. T. Wu, and Y. J. Chang, “Assessment of diaphragm and external intercostals fatigue from surface EMG using cervical magnetic stimulation,” *Sensors*, vol. 8, no. 4, pp. 2174–2187, Mar. 2008.

[158] G. Gayan-Ramirez, N. Gosselin, T. Troosters, F. Bruyninckx, R. Gosselink, and M. Decramer, “Functional recovery of diaphragm paralysis: a long-term follow-up study,” *Respir. Med.*, vol. 102, no. 5, pp. 690–698, May. 2008.

[159] C. S. Van Holsbeke, G. Leemans, W. G. Vos, J. W. De Backer, S. C. Vinchurkar, M. Geldof, P. R. Verdonck, P. M. Parizel, P. E. Van Schil, and W. A. De Backer, “Functional respiratory imaging as a tool to personalize respiratory treatment in patients with unilateral diaphragmatic paralysis,” *Respir. Care*, vol. 59, no. 9, pp. e127–131, Sep. 2014.

[160] J. Boudarham, D. Pradon, H. Prigent, L. Falaize, M. C. Durand, H. Meric, M. Petitjean, and F. Lofaso, “Optoelectronic plethysmography as an alternative method for the diagnosis of unilateral diaphragmatic weakness,” *Chest*, vol. 144, no. 3, pp. 887–895, Sep. 2013.

[161] W. D. Xu, Y. D. Gu, J. B. Lu, C. Yu, C. G. Zhang, and J. G. Xu, “Pulmonary function after complete unilateral phrenic nerve transection,” *J. Neurosurg.*, vol. 103, no. 3, pp. 464–467, Sep. 2005.

[162] A. Mier-Jedrzejowicz, C. Brophy, J. Moxham, and M. Green, “Assessment of diaphragm weakness,” *Am. Rev. Respir. Dis.*, vol. 137, no. 4, pp. 877–883, Apr. 1988.

[163] S. Tejman-Yarden, D. Lederman, I. Eilig, A. Zlotnik, N. Weksler, A. Cohen, and G. M. Gurman, “Acoustic monitoring of double-lumen ventilated lungs for the detection of selective unilateral lung ventilation,” *Anesth. Analg.*, vol. 103, no. 6, pp. 1489–1493, Dec. 2006.

[164] T. A. Dernaika, W. G. Younis, and P. V Carlile, “Spontaneous recovery in idiopathic unilateral diaphragmatic paralysis,” *Respir. Care*, vol. 53, no. 3, pp. 351–354, Mar. 2008.

[165] G. S. Habib, “Unilateral diaphragm paralysis following vaccination,” *J. Med. Cases*, vol. 3, no. 3, pp. 164–166, May. 2012.

[166] J. A. Fiz, R. Jané, M. Lozano, R. Gómez, and J. Ruiz, “Detecting unilateral phrenic paralysis by acoustic respiratory analysis,” *PLoS One*, vol. 9, no. 4, e93595, Jan. 2014.

[167] N. E. Pardee, C. J. Martin, and E. H. Morgan, “A test of the practical value of estimating breath sound intensity. Breath sounds related to measured ventilatory function,” *Chest*, vol. 70, no. 3, pp. 341–344, Sep. 1976.

[168] A. B. Bohadana, M. C. Kopferschmitt-Kubler, and G. Pauli, “Breath sound intensity in patients with airway provocation challenge test positive by spirometry but negative for wheezing: a preliminary report,” *Respiration*, vol. 61, no. 5, pp. 274–279, Jan. 1994.

[169] L. P. Malmberg, A. R. Sovijärvi, E. Paajanen, P. Piirilä, T. Haahtela, and T. Katila, “Changes in frequency spectra of breath sounds during histamine challenge test in adult asthmatics and healthy control subjects,” *Chest*, vol. 105, no. 1, pp. 122–131, Jan. 1994.

[170] H. J. Schreur, J. Vanderschoot, A. H. Zwinderman, J. H. Dijkman, and P. J. Sterk, “The effect of methacholine-induced acute airway narrowing on lung sounds in normal and asthmatic subjects,” *Eur. Respir. J.*, vol. 8, no. 2, pp. 257–265, Feb. 1995.

[171] K. Anderson, S. Aitken, R. Carter, J. E. MacLeod, and F. Moran, “Variation of breath sound and airway caliber induced by histamine challenge,” *Am. Rev. Respir. Dis.*, vol. 141, no. 5 Pt 1, pp. 1147–1150, May. 1990.

[172] C. Habukawa, K. Murakami, H. Mochizuki, S. Takami, R. Muramatsu, H. Tadaki, S. Hagiwara, T. Mizuno, H. Arakawa, and Y. Nagasaka, “Changes in the highest frequency of breath sounds without wheezing during methacholine inhalation challenge in children,” *Respirology*, vol. 15, no. 3, pp. 485–490, Apr. 2010.

[173] L. P. Malmberg, L. Pesu, and A. R. Sovijarvi, “Significant differences in flow standardised breath sound spectra in patients with chronic obstructive pulmonary disease, stable asthma, and healthy lungs,” *Thorax*, vol. 50, no. 12, pp. 1285–1291, Dec. 1995.

[174] C. Habukawa, K. Murakami, N. Horii, M. Yamada, and Y. Nagasaka, “A new modality using breath sound analysis to evaluate the control level of asthma,” *Allergol. Int.*, vol. 62, no. 1, pp. 29–35, Mar. 2013.

[175] Z. Wang, T. Bartter, B. M. Baumann, B. M. Baugmann, B. M. Baumman, W. Abouzgheib, M. E. Chansky, and S. Jean, “Asynchrony between left and right lungs in acute asthma,” *J. Asthma*, vol. 45, no. 7, pp. 575–578, Sep. 2008.

[176] Z. Wang, S. Jean, and T. Bartter, “Lung sound analysis in the diagnosis of obstructive airway disease,” *Respiration*, vol. 77, no. 2, pp. 134–138, Jan. 2009.

[177] K. K. Guntupalli, R. M. Reddy, R. H. Loutfi, P. M. Alapat, V. D. Bandi, and N. A. Hanania, “Evaluation of obstructive lung disease with vibration response imaging,” *J. Asthma*, vol. 45, no. 10, pp. 923–930, Dec. 2008.

[178] M. Mineshita, H. Kida, H. Handa, H. Nishine, N. Furuya, S. Nobuyama, T. Inoue, S. Matsuoka, and T. Miyazawa, “The correlation between lung sound distribution and pulmonary function in COPD patients,”

PLoS One, vol. 9, no. 9, e107506, Jan. 2014.
[179] M. Mineshita, S. Matsuoka, and T. Miyazawa, "Effects of bronchodilators on regional lung sound distribution in patients with chronic obstructive pulmonary disease," *Respiration*, vol. 87, no. 1, pp. 45–53, Jan. 2014.
[180] J. J. Marini, D. J. Pierson, L. D. Hudson, and S. Lakshminarayan, "The significance of wheezing in chronic airflow obstruction," *Am. Rev. Respir. Dis.*, vol. 120, no. 5, pp. 1069–1072, Nov. 1979.
[181] C. S. Shim and M. H. Williams, "Relationship of wheezing to the severity of obstruction in asthma," *Arch. Intern. Med.*, vol. 143, no. 5, pp. 890–892, May. 1983.
[182] J. A. Fiz, R. Jané, A. Homs, J. Izquierdo, M. A. García, and J. Morera, "Detection of wheezing during maximal forced exhalation in patients with obstructed airways," *Chest*, vol. 122, no. 1, pp. 186–191, Jul. 2002.
[183] I. Mazic, S. Sovilj, and R. Magjarevic, "Analysis of respiratory sound in asthmatics infant," *Meas. Sci. Rev.*, vol. 3, no. 2, pp. 9–12, 2003.
[184] J. A. Fiz, R. Jané, J. Izquierdo, A. Homs, M. A. García, R. Gomez, E. Monso, and J. Morera, "Analysis of forced wheezes in asthma patients," *Respiration*, vol. 73, no. 1, pp. 55–60, Jan. 2006.
[185] S. A. Taplidou and L. J. Hadjileontiadis, "Analysis of wheezes using wavelet higher order spectral features," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 7, pp. 1596–1610, Jul. 2010.
[186] A. Oliveira, C. Pinho, J. Dinis, D. Oliveira, and A. Marques, "Automatic wheeze detection and lung function evaluation: a preliminary study," in *6th Int. Conf. Health Inform.*, 2013.
[187] H. S. Fischer, L. C. Puder, S. Wilitzki, J. Usemann, C. Bührer, S. Godfrey, and G. Schmalisch, "Relationship between computerized wheeze detection and lung function parameters in young infants," *Pediatr. Pulmonol.*, vol. 51, no. 4, pp. 402–410, Sep. 2016.
[188] M. Lozano, J. A. Fiz, and R. Jané, "Analysis of normal and continuous adventitious sounds for the assessment of asthma," in *XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013 SE - 243*, vol. 41. Springer International Publishing, 2014, pp. 981–984.
[189] C. Shi, S. Boehme, A. H. Bentley, E. K. Hartmann, K. U. Klein, M. Bodenstein, J. E. Baumgardner, M. David, R. Ullrich, and K. Markstaller, "Assessment of regional ventilation distribution: comparison of vibration response imaging (VRI) with electrical impedance tomography (EIT)," *PLoS One*, vol. 9, no. 1, e86638, Jan. 2014.
[190] Y. Amirat, M. Benbouzid, T. Wang, and S. Turri, "Performance analysis of an EEMD-based Hilbert Huang transform as a bearing failure detector in wind turbines," in *Int. Conf. Green Energy*, pp. 193–198, 2014.
[191] W. Zhang, X. Zhang, and Y. Sun, "Based on EEMD-HHT marginal spectrum of speech emotion recognition," in *Int. Conf. Comput. Meas. Control Sens. Netw.*, pp. 91–94, 2012.

Publications derived from this thesis

Articles indexed in the Journal Citation Reports

1. J. A. Fiz, R. Jané, M. Lozano, R. Gómez, and J. Ruiz, "Detecting unilateral phrenic paralysis by acoustic respiratory analysis," *PLoS One*, vol. 9, no. 4, e93595, Jan. 2014.
2. M. Lozano, J. A. Fiz, and R. Jané, "Automatic differentiation of normal and continuous adventitious respiratory sounds using ensemble empirical mode decomposition and instantaneous frequency," *IEEE J. Biomed. Health Inform.*, vol. 20, no. 2, pp. 486–497, Mar. 2016.
3. M. Lozano, J. A. Fiz, and R. Jané, "Performance evaluation of the Hilbert–Huang transform for respiratory sound analysis and its application to continuous adventitious sound characterization," *Signal Process.*, vol. 120, pp. 99–116, Mar. 2016.
4. M. Lozano, J. A. Fiz, C. Martínez-Rivera, A. Torrents, J. Ruiz, and R. Jané, "Novel approach to assessing bronchodilator response by continuous adventitious respiratory sound analysis," submitted to *PLoS One*.

Book chapters

5. M. Lozano, J. A. Fiz, and R. Jané, "Analysis of normal and continuous adventitious sounds for the assessment of asthma," in *XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013 SE - 243*, vol. 41. Springer International Publishing, 2014, pp. 981–984.

International conferences

6. M. Lozano, J. A. Fiz, and R. Jané, "Estimation of instantaneous frequency from empirical mode decomposition on respiratory sounds analysis," in *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, pp. 981–984, 2013.

National conferences

7. M. Lozano, J. A. Fiz, and R. Jané, "Análisis multicanal de sonidos respiratorios en acústica pulmonar: aplicación clínica en pacientes asmáticos," in *XXX Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, San Sebastián, Spain, 2012.
8. M. Lozano, J. A. Fiz, and R. Jané, "Análisis de la intensidad de los sonidos respiratorios para el diagnóstico de la parálisis frénica unilateral," in *XXXII Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, Barcelona, Spain, 2014.
9. M. Lozano, J. A. Fiz, and R. Jané, "Análisis de sonidos adventicios continuos en pacientes asmáticos mediante el espectro de Hilbert," in *XXXIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, Madrid, Spain, 2015.