ANNÉE 2013

UNIVERSITÉ DE RENNES 1

THÈSE / UNIVERSITÉ DE RENNES 1
sous le sceau de l'Université Européenne de Bretagne

pour le grade de

DOCTEUR DE L'UNIVERSITÉ DE RENNES 1

Mention : Traitement du signal et télécommunications

Ecole doctorale MATISSE

présentée par

Xavier Navarro

préparée à l'unité de recherche LTSI - INSERM U1099
Laboratoire Traitement du Signal et de l'Image
Université de Rennes 1

Analysis of cerebral and respiratory activity in neonatal intensive care units for the assessment of maturation and infection in the early premature infant

Thèse soutenue à Rennes
le 22 octobre 2013

devant le jury composé de :

Anne HUMEAU-HEURTIER
PU, Université d'Angers / Rapporteur

Leif SÖRNMO
Pr, Université de Lund / Rapporteur

Pablo LAGUNA
Pr, Université de Saragosse / Examinateur

Patrick PLADYS
PU-PH, Université de Rennes 1 / Examinateur

Fabienne PORÉE
MCU, Université de Rennes 1 / Encadrant

Guy CARRAULT
PU, Université de Rennes 1 / Directeur de thèse

Acknowledgments

This Ph.D. thesis has represented four years of work at Laboratoire Traitement du Signal et de l'Image (LTSI) of Rennes after the six months of master's thesis I started in March 2009. Since my first day in the LTSI, I felt at home thanks to the warm welcome of its director, Prof. Lotfi Senhadji.

First and foremost, I wish to express my sincere gratitude to my mentors, Prof. Guy Carrault and Dr. Fabienne Porée for giving me the privilege to do my Ph.D. thesis in the LTSI. Thanks to their constant support, advice, encouragement and to their efforts to provide all the means during my stay in Rennes, this work came into fruition.

I also would like to thank Prof. Anne Humeau-Heurtier, from Université d'Angers, Prof. Leif Sörnmo, from Lund's University for accepting reviewing this work and for their proficient comments improving the quality of this work. I also would like to thank Prof. Pablo Laguna from Universidad de Zaragoza, for kindly taking part in the thesis' evaluation committee and for their constructive comments during the defense.

I am very thankful to M.D. Alain Beuchée, from the CHU of Rennes, for his priceless assistence in the clinical aspects of this thesis but also to his remarkable competences in signal processing that made to work with him very motivating. Prof. Patrick Pladys and M.D. Tiphaine Mialet-Marty contributed actively as well to this thesis, providing helpful advice and valuable knowledge in neonatology. I am also grateful to M.D Camille Arberet, M.D Mathieu Kuchenbuch and Anne-Marie Lamour for their participation in the clinical part.

The help and advice provided by several researchers in the LTSI deserve an special mention. These are Dr. Amar Kachenoura, Dr. Nathalie Costet, Dr. Alfredo Hernández, Dr. Jean-Jacques Bellanger, Dr. Miguel Altuve and David Ojeda.

Likewise, I would like to acknowledge the scientists from other establishments from whom I exchanged relevant discussions and generously shared their codes: Dr. Pierre Borgnat, from ENS Lyon, Prof. Eugene N. Bruce, from the University of Kentucky and Cheolwoo Park, from the University of Georgia.

Finally, but not less importantly, I greatly appreciated the support and encouragement from the LTSI staff and colleagues that led to many unforgettable friendships.

Contents

List of Figures

List of Tables

List of acronyms

ADIF Averaged sum of the absolute values of amplitudes

aEEG Amplitude-integrated electroencephalogram

AIT Algorithmic information theory

AMI Average mutual information

ApEn Approximate entropy

AS Active sleep

AOP Apnea of prematurity

AW Active wakefulness

BSS Blind Source Separation

CA Conceptional age

CCA Canonical correlation analysis

CEDAC Combination of EEG decomposition and artifact cancellation

CEEEMDAN Complete ensemble empirical mode decomposition with adaptive noise

CHU *Centre hospitalier universitaire*

CL Classification

CPAP Continuous positive airway pressure

CWT Continuous wavelet transform

DFA Detrended fluctuation analysis

DFT Discrete Fourier transform

DP Data preprocessing

DWT Discrete wavelet transform

ECG Electrocardiogram

EDIN *Échelle douleur inconfort nouveau-né*

EEG Electroencephalogram

EEMD Ensemble empirical mode decomposition

EMD Empirical mode decomposition

EMG Electromyogram

EOG Electrooculogram

EOS Early onset sepsis

FARIMA Fractional autoregressive-integrated moving average models

fBm Fractional Brownian motions

FE Feature extraction

FFT Fast Fourier transform

fGn Fractional Gaussian noise

FIR Finite impulsional response filters

fLn Fractional lognormal noise

FN Number of false negatives

FP Number of false positives

Fp1,Fp2 Left and right fronto-temporal electrodes

FT Fourier transform

GA Gestational age

GKA Gaussian kernel algorithm

HAS Hurst-adjusted surrogates

HF High Frequency

HHT Hilbert-Huang transform

HRV Heart rate variability

HT Hilbert transform

iAAFT Iterated amplitude adjusted Fourier transformed surrogates

IBI Interburst intervals

IC Independent component

ICA independent component analysis

IIR Infinite impulsional response filters

IMF Intrinsic modal function

INSERM *Institut national de santé et de recherche médicale*

INTEM Intelligent monitoring system developed in the LTSI

IQR Inter-quartile range

IS Indeterminate sleep

IT Information Theory

LF Low Frequency

LFN Low frequency noise

LLE Largest Lyapunov exponent

LOOCV Leave-one-out cross-validation

LOS Late onset sepsis

LR Logistic regression

LRD Long-range dependence

LSSD Least squares based on standard deviation method

LTSI *Laboratoire de Traitement du Signal et de l'Image*

LZ Lempel-Ziv complexity

MLE Maximum likelihood estimate

Mm Maximum - minimum operator

NI Nonlinear index

NICU Neonatal intensive care unit

NIDCAP Newborn individualized developmental care and assessment program

NL Noise limit

NLEO Nonlinear energy operator

NREM Sleep states without rapid eye movements

PMA Post-menstrual age

QS Quiet sleep

QW Quiet wakefulness

REM Sleep state with rapid eye movements

RMSE Root mean squared error

ROC Receiver operating characteristic curve

RR Cardiac cycle duration

RSS Random-shuffled surrogates

RV Respiratory variability

SAT Spontaneous activity transients

SEPIA *Surveillance, Explication et Prévention de l'Insuffisance cardiaque et de l'Apnée-bradycardie*

SG Savitzky-Golay filter

Sn Sensitivity

SNR Signal-to-Noise Ratio

Sp Specificity

SpEn Sample entropy

SRD Short-range dependence

STFT Short-time Fourier transform

SSS Small-shuffled surrogates

SVM Support vector machines

T3,T4 Left and right temporal electrodes

TN Number of true negatives

TP Number of true positives

VLBW Very low-birth-weight infants

VLF Very low frequency

Résumé

L'incidence des naissances prématurés augmente progressivement dans le monde et est devenue un enjeu global de santé. Les répercussions entraînent non seulement une augmentation de la mortalité et la morbidité pendant l'enfance, mais également quelquefois des incapacités à l'âge adulte. Les avancées récentes dans le domaine du traitement du signal et son intégration dans les unités de soins intensifs néonatales (USIN) ont permis le développement de marqueurs non-invasifs, rapides et fiables du suivi du bébé prématuré. A titre d'exemple, l'évaluation de l'activité cardiaque et de sa variabilité au cours du temps a permis la mise au point d'outils efficaces pour la détection et la prédiction de l'apnée-bradycardie. Cependant, la disponibilité d'index objectifs permettant le diagnostic précoce d'infection généralisée (Sepsis) ou encore quantifiant la maturité restent encore aujourd'hui une question ouverte en néonatalité où des réponses sont attendues.

L'objectif de ce travail était d'étudier plusieurs modalités d'observation enregistrés en USIN de manière continue, en particulier l'électroencéphalogramme et la respiration. En ce qui concerne L'ECG, De nombreux travaux ont été menés par le Laboratoire Traitement du Signal et de l'Image (LTSI) dans le contexte de la détection de sepsis. Ceci explique que l'ECG n'a pas été abordé explicitement dans le cadre de ce travail. En revanche, le signal de respiration et l'EEG n'avaient jamais été auparavant analysés.

Pour l'EEG, son recueil en USIN se fonde sur un nombre réduit d'électrodes et implique de fait des enregistrements fortement bruités. Une partie de ce travail a donc été de minimiser les bruits superposés à l'EEG compte tenu des particularités de ces signaux contenant des formes très immatures, une forte non-stationnarité caractérisée par un tracé alternant transitoire de basses fréquences (bouffées EEG) et des périodes d'inactivité. Une approche originale, associant des techniques de décomposition de signal et d'annulation de bruit, a été donc proposée et étudiée en profondeur au Chapitre III. Les décompositions de l'EEG s'appuient sur la transformée discrète d'ondelettes (TDO), la décomposition modale empirique (DME) et la DME ensembliste avec bruit adaptatif (DMEEBA). L'annulation du bruit ECG exploite le filtrage adaptatif (FA) ou la correction par soustraction de l'interférence cardiaque moyenne (SIM).

Nous avons ainsi pu montrer, tant sur des simulations réalistes que sur des signaux réels, l'intérêt et la supériorité de la combinaison de la DMEEBA et le FA pour minimiser la distorsion de ces signaux EEG. Lorsque les signaux contiennent du bruit cardiaque et de basse fréquence, cette dernière association permet en effet d'augmenter le taux de suppression de bruit jusqu'au 50% par rapport à l'utilisation d'un filtrage passe-bande classique associé au FA.

Débarrassés du bruit, nous avons proposé dans le Chapitre IV un nouveau détecteur d'activités EEG basées

sur la régression logistique (LR) de plusieurs paramètres. La LR permet la classification en bouffées et en intervalles inter-bouffées (IIB). Un étude comparative avec un détecteur de seuil classique, utilisé classiquement en néonatologie, a montré la supériorité de notre approche, présentant des performances en sensibilité de 96.1% (contre 91.4%) et de 95.9% en spécificité (contre 93.6%). En parallèle, nous avons également mené une étude sérieuse en comparant les performances d'annotation de bouffées de plusieurs experts. Nous avons ainsi souligné que le détecteur proposé présente un taux de concordance très satisfaisant avec trois experts cliniques différents.

Sur le plan clinique, nous nous sommes intéressés à la maturité du bébé prématuré ex-utero. L'objectif final étant à terme de proposer aux cliniciens des indices quantifiant cette maturité. Sur la base de données à notre disposition, nous avons pu montrer que la durée des bouffées et leurs fréquences d'apparition permettraient de quantifier la maturité.

Nous nous sommes également intéressés à l'impact de la vaccination sur l'activité cérébrale du nouveau-né. Cette analyse venait en complément d'une étude portant sur la variabilité cardiaque sur la même cohorte qui avait montré une augmentation des incidences cardio-respiratoires (notamment bradycardies et désaturations) aux données post-vaccinales. L'analyse quantitative de l'EEG, des hypnogrammes et des bouffées sur la base de données à notre disposition montrent que l'immunisation perturbe très faiblement le sommeil avec cependant une redistribution des énergies dans le spectre de l'EEG. En particulier, on a constaté une diminution de la puissance dans la bande delta inférieur (0.5 - 1.5 Hz) et une augmentation notable de la puissance dans la bande thêta inférieur (5.1 - 8 Hz).

Pour la deuxième modalité abordée dans ce mémoire, la respiration, il importe de rappeler que ce signal n'avait jamais été analysé par notre équipe de recherche. Il a été étudié sous sa forme temporelle mais aussi sous sa forme de séries de variabilité de cycle respiratoire, identiquement aux études de l'analyse de la variabilité cardiaque. Un détecteur de cycles respiratoires a donc été mis au point dans le Chapitre V. Il repose sur un débruitage préalable, afin de rejeter les artefacts de mouvements très fréquents en USIN, et un classifieur. Une nouvelle fois, ce classifieur se fonde sur la régression logistique et l'apprentissage à partir d'annotations d'experts. Les performances mesurées sur les bases de données à notre disposition montrent une sensibilité et une spécificité autour de 86% pour la réjection des artefacts, une valeur positive prédictive et une sensibilité de 97.7% pour la détection des cycles respiratoires.

Appliquée à la quantification de la maturation, nous avons souligné que le nombre d'apnées, et leur durée, la durée du cycle respiratoire ainsi que les puissances dans les bandes de basse fréquence -obtenues à partir d'une analyse spectrale des signaux de variabilité- varient avec la maturation du bébé. Pour la détection en ligne d'infection, nous avons constaté que l'écart type de la durée du cycle respiratoire et la durée des apnées s'allongeaient avec l'infection.

Le Chapitre VI s'est intéressé à l'application de méthodes non linéaires tels que la dimension de corrélation, l'entropie d'échantillonnage et la titration numérique de bruit pour l'étude la maturation. Dans un premier temps, sur des séries respiratoires régulières de courte durée, nous avons vérifié l'existence de non-linéarité. Dans un second temps, vis-à-vis de critères maturatifs, cette propriété évolue d'une manière irrégulière tendant à diminuer. Ces résultats contredisent des études similaires.

L'ultime chapitre –centré sur l'analyse d'une autre propriété non-linéaire, la dépendance long terme (DLT)–, avait pour objectif d'utiliser l'exposant de Hurst (H) pour la détection d'infection à partir de séries de variabilité respiratoire. Afin de mener à bien cette étude, nous avons recherché le meilleur estimateur l'exposant de Hurst parmi trois : basé sur la transformé d'ondelettes (DWT), la "Detrended Fluctuation Analysis" (DFA) et le "Least Squares based on Standard Deviation" (LSSD). Une analyse de performances sur signaux simulés à H prescrit a été conduite. Le simulateur exploite soit des données de substitution (surrogates), appelé "Hurst-adjusted surrogates" (HAS), soit un modèle stochastique à bruit fractionnaire Gaussien. Outre le contôle du H, l'intérêt est de générer des signaux qui présentent des caractéristiques très similaires aux patterns respiratoires néonataux : Régulier, périodique et erratique. Les simulations ont montré que, pour le problème posé, les estimations d'H issues de la DFA et du LSSD s'avèrent plus robustes (biais et variance d'estimation plus faibles).

L'application de l'exposant de Hurst comme indicateur potentiel d'infection a été ensuite exploré. La population de nouveau-nés infectés témoigne une diminution significative de sa valeur d'H -mettant en exergue une dégradation de l'organisation fractale- par rapport à la population saine.

Ce travail a également conduit plusieurs plateformes logicielles et déjà exploitables par les cliniciens en recherche. La première concerne l'annotation des bursts, la deuxième celle des cycles respiratoires et des périodes artefactées. Ces stations représentent en effet un point de départ fondamental pour mener une étude à grande échelle sur la discordance entre plusieurs annotateurs. Comme déjà mentionné, les PHRC INTEM et CARESS-PREMI vont nous permettre de constituer des bases de données uniques sur plusieurs centres et ainsi évaluer les pratiques en tenant compte des effets centres.

Les suites à donner à ce travail sont multiples. En ce qui concerne le traitement du signal EEG, il importe de constituer une base de données horizontal permettant de valider l'ensemble des résultats obtenus. Ces derniers ne peuvent être à ce jour que considérés comme préliminaires dans la mesure où la base de données étudiée est réduite. En profitant des méthodes de classification, il sera alors pertinent d'analyser l'ensemble des grapho-éléments constituant l'EEG néonatale pour mener une étude de la maturation.

En ce qui concerne la seconde modalité, la respiration, il convient également de constituer une base de données horizontale élargie. Sur un plan plus méthodologique et pour le détecteur de cycles respiratoire, une comparaison avec d'autres méthodes de classification, telles que les séparateurs de vaste marge ou réseaux de neurones, serait necessaire afin d'envisager l'amélioration des performances actuelles.

Concernant les differents paramètres issus des outils linéaires et non-linéaires, la conception d'une analyse statistique multivariée permettrait vraisemblablement d'améliorier les performances la détection d'infection. Il est en effet naturel dans ce contexte d'exploiter nos résultats antérieurs fondés sur la variabilité cardiaque. Cette perspective fait explicitement partie du protocole hospitalier de recherche clinique CARESS-PREMI qui a la volonté de valider l'ensemble des outils fondés sur la variabilité cardiaque et respiratoire pour la détection d'infection. Ce protocole, coordonné par notre partenaire clinique, implique trois CHU (Rennes, Lille et Angers) et represente autant d'elements positifs et favorables pour l'exploitation clinique auprès du bébé prématuré des solutions décrites dans ce mémoire.

Introduction

The incidence of preterm births is increasing consistently worldwide and has become a global health concern, having repercussions not only on the increased mortality and morbidity rates during infancy, but also on the potential disabilities in adulthood [1]. Apnea-bradycardia and early onset sepsis are, without a doubt, two major issues of prematurity as they are manifested more frequently than in full-term populations [2].

Thanks to the introduction of neonatal intensive care units a few decades ago and to the development of computer-assisted monitoring systems, diagnostic and treatments can start without delay, improving the prognosis of these premature babies [3]. Recently, more sophisticated tools based upon physiological signals have emerged, leading to the vast research area with the "dream" of providing fast and non-invasive diagnostic indicators.

To succeed in such a purpose, the advances in signal processing applications and their integration in neonatal intensive care units are a real challenge. For instance, the assessment of the cardiac activity has been especially prolific, providing effective detection tools to prevent apnea-bradycardia [4, 5].

On the contrary, the electroencephalogram (EEG) and the respiratory trace have been less explored. This dissertation is focused on these signals and proposes issues for the early diagnosis of septicemic infection, the availability of objective maturity indexes or the influence of immunization.

The knowledge of the signals acquired in the NICU and their physiological particularities are essential and are described in a first chapter. Preterm's EEG, characterized by a highly non-stationary activity and immature breathing, having a variety of patterns conditioned by apnea [6, 7], are studied in detail. In Chapter II, a literature review focusing on the study of physiological signals from preterm infants is presented. The previous research related to these topics done by the Laboratoire Traitement du Signal et de l'Image (LTSI), in collaboration with the University Hospital (CHU) of Rennes, is also introduced, as well as a description of the two clinical databases recorded at the CHU since the last years.

The second part is devoted to the EEG, obtained from a simplified four-electrode system, and characterized by nonstationary patterns such as tracé alternant and slow delta waves. These signals are often artifacted and have remarkably different characteristics from adults', hence an important effort will be done to preprocess them properly. In view of the multiple and complex noise sources and knowing that preserving the original waveform is crucial, a new denoising framework combining EEG decomposition and noise cancellation is introduced in Chapter III. An ensemble of tests compares different combinations of these methods, including a recent improved version of empirical mode decomposition [8], to find the most effective solution.

Afterwards, an automatic detector based in logistic regression is designed in Chapter IV to classify accurately burst/inter-burst intervals (IBIs) in EEG. Two clinical applications, the influence of immunization on the brain

activity and the investigation of some maturative indexes based upon the EEG, are also described.

In the last part of this work, breathing signals, obtained from abdominal movements, are exploited both in the temporal form and as respiratory variability (RV) signals. An artifact-rejection algorithm and an automatic cycle detector are designed for this purpose in Chapter V. Maturation and sepsis diagnosis are first studied by using linear signal processing methods. Next, Chapter VI carries out several analyses in breathing signals to quantify their nonlinear content, including classical chaos-based methods and noise titration [9–11]. These nonlinear tools are then explored to quantify the degree of maturation in premature newborns.

The long-range dependence property [12], estimated by the Hurst exponent (H), is studied in Chapter VII. A comparison of several estimation techniques of H is first performed, testing their robustness by evaluating the errors in artificial RV signals with prescribed H. To obtain signals with similar characteristics than the typical immature breathing patterns, a modified fractional Gaussian model and a newly-proposed surrogate technique are designed. The last part of the chapter goes back to a clinical application, the interest of the Hurst exponent as a real-time index to detect late onset sepsis.

In view of to the very challenging tools designed in this dissertation, several conclusions and perspectives are finally drawn.

Bibliography

[1] W. H. Organization. March of dimes, PMNCH, save the children, WHO. Born too soon: the global action report on preterm birth. Technical report (2012).

[2] J. M. Abu-Shaweesh and R. J. Martin. *Neonatal apnea: What's new?* Pediatric Pulmonology, Pediatric Pulmonology **43**(10), 937–944 October (2008).

[3] D. K. Richardson, J. E. Gray, S. L. Gortmaker, D. A. Goldmann, D. M. Pursley, and M. C. McCormick. *Declining severity adjusted mortality: Evidence of improving neonatal intensive care.* Pediatrics **102**(4), 893–899 October (1998).

[4] G. Pravisani, A. Beuchée, L. Mainardi, and G. Carrault. Short term prediction of severe bradycardia in premature newborns. In *Computers in Cardiology, 2003*, pages 725 – 728 September (2003).

[5] F. Portet, F. Gao, J. Hunter, and S. Sripada. *Evaluation of on-line bradycardia boundary detectors from neonatal clinical data.* Conference proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference **2007**, 3288–3291 (2007).

[6] J. S. Hahn, H. Monyer, and B. R. Tharp. *Interburst interval measurements in the EEGs of premature infants with normal neurological outcome.* Electroencephalography and clinical neurophysiology **73**(5), 410–418 November (1989).

[7] M.-F. Vecchierini, M. André, and A. d'Allest. *Normal EEG of premature infants born between 24 and 30 weeks gestational age: Terminology, definitions and maturation aspects.* Neurophysiologie Clinique/Clinical Neurophysiology **37**(5), 311–323 (2007).

[8] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin. A complete ensemble empirical mode decomposition with adaptive noise. In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4144–4147. IEEE May (2011).

[9] J. S. Richman and J. R. Moorman. *Physiological time-series analysis using approximate entropy and sample entropy.* American Journal of Physiology - Heart and Circulatory Physiology **278**(6), H2039 –H2049 June (2000).

[10] A. Lempel and J. Ziv. *On the complexity of finite sequences.* IEEE Transactions on Information Theory **22**(1), 75–81 (1976).

[11] C.-S. Poon and M. Barahona. *Titration of chaos with added noise.* Proceedings of the National Academy of Sciences **98**(13), 7107 –7112 June (2001).

[12] B. B. Mandelbrot. *The Fractal Geometry of Nature.* W.H. Freeman (1983).

Part I

Clinical context

Chapter I

Prematurity

The first part of this thesis introduces the premature birth, provides clinical definitions and outlines the problem statement. More particularly, this chapter aims at introducing the physiological and clinical problems associated to prematurity, as well as providing a brief review of the extensive research contributing to improve the overall preterm's quality of life.

1 Definitions

In humans, prematurity refers to the broad category of neonates born between 24 and 37 weeks after the last menstrual period, in contrast to full-term birth, in which infants are born between 37 and 42 weeks (See Figure I.1).

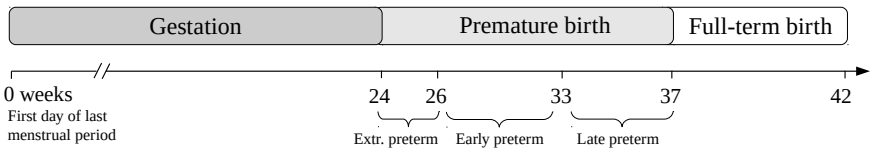

Figure I.1 – Birth chronogram according to the American Academy of Pediatrics [1].

Newborns undergo rapid changes as they are born. Stating the significance of the degree of prematurity their ages can be defined, as recommended by the standard terminology of the American Academy of Pediatrics [1], in gestational, chronological or post-conceptional terms (See Figure I.2). Likewise, different categories of prematurity are defined regarding at the post-menstrual age (PMA):

Extremely preterm infant, born before 26 weeks PMA.

Early preterm infant, born between 26 and 32 weeks PMA.

Late or moderate preterm infant, born between 33 and 37 weeks PMA.

Newborns weighing less than 1 500 grammes at birth, regardless of their postmenstrual age, are also referred as very low-birth-weight (VLBW) infants. This group constitutes the 10-12% of premature births and is at high risk for significant mortality and morbidity [2].

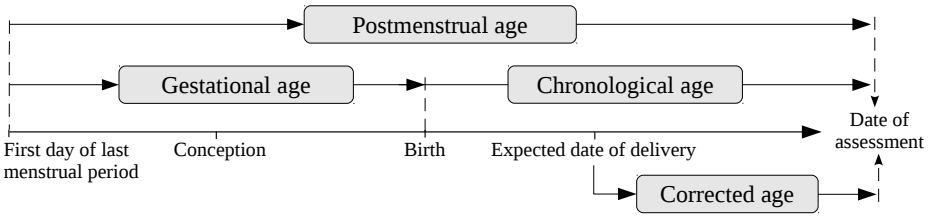

Figure I.2 – Definition of ages in prematurity according to the American Academy of Pediatrics [1].

Nowadays, more than 55 000 infants are born prematurely every year in France, representing approximately the 7% of births [3], and more than 400 000 in the United States [4]. Worldwide, it is estimated that preterm births in 2010 amounted to about 15 millions and, only in three of the 184 censored countries, the preterm birth rate decreased between years 1990 and 2010 [5].

Preterm birth is the main cause of neonatal mortality and morbidity, increased as the younger the infant is born: while only half of extremely preterm infants survive, the expectancy of survival rises to more than 90% between 27 and 28 weeks PMA [6]. Although the favorable outcome has improved since the implantation of neonatal intensive care units (NICUs) in the early 70's, prematurity also tends to increase every year [5]. The causes for this phenomenon are not completely known, but approximately the 60% of premature births are provoked by medical complications during pregnancy and the remaining 40% are spontaneous [7].

Obviously, prematurity is a major global concern and, as a recent report of the World Health Organization outlined [2], more initiatives need to be proposed both in family planning and strategic investments in research and innovation to prevent and minimize the problem.

2 Particularities of preterm infants

During the third trimester of pregnancy (gestation weeks 29 to 40) the fetus undergoes important changes in the brain and lungs, the last organs to develop. The brain increases rapidly in size and establishes connections between nerve cells and the lungs are provided with surfactant while the blood-air barrier is established [8]. Hence, during the first weeks of extrauterine life, breathing and sleep-wake patterns in preterm infants are the most remarkable external manifestations of the immaturity of the lungs and the brain, respectively.

In comparison to their full-term counterparts, these manifestations, procured by the measure of physiological signals, have some particularities and reflect their maturation. Understanding the observed signals and sleep-wake organization in neonatal intensive care units is, then, crucial in the assessment of the infants evolution and in the implementation of individualized developmental programs.

2.1 Cerebral activity

Early brain development

During fetal development, the brain is subject to rapid and complex changes. Within the six weeks following conception, the neural tube (basic brain and spinal cord templates) is established. Before 24 weeks post-menstrual age, a massive proliferation of neurons and glial cells takes place, followed by a period called pattern formation

in which neurons migrate and axons grow to form the framework of intercellular wiring. At early preterm phases (25-33 weeks PMA) the development of the brain is characterized by a remarkable complex activity involving neuronal maturation and connectivity in the frontal, somatosensory, visual and auditory cortex. After the 34th week, changes in sensory processing at the cortical level seem to occur, as suggests the overall advancement in the organization of cortical pathways [9]. In the term neonatal phase, the cortical circuitry reorganizes and there is a rapid development of spines and synapses [10].

The earlier is the interruption of gestation before the term phase, the more vulnerable the premature brain becomes. Actually, a decreased cephalic volume [11, 12] and histological differences [9, 13] can be observed when compared to term infants' brain, but more importantly, preterm birth is associated with cerebral structural alterations. Although major tissue lesions are relatively rare, there are common alterations in preterm infants at term-equivalent age such as white matter disease, abnormal cortical and cerebellum development and other morphological abnormalities. Often subtle, they require computer-assisted imaging techniques to be evaluated [14].

In neurological examinations –tests evaluating muscular tone– premature infants at term ages are more hyperexcitable and tend to have less flexor tone in the limbs [15], and electroencephalograms are described more frequently by characteristic discontinuous patterns [16, 17]. But from the functional point of view, the true effect of developing the brain at preterm ages is not completely known. It is well documented that the extrauterine period can accelerate some neurophysiologic behavior (EEG organization and sleep), but at the same time, the cardiorespiratory function presents more immature patterns than full-term matched cohorts. This contradiction seems to be a reflect of the adaptive measures of the immature brain facing the environmental stress [18, 19].

Sleep-wake organization

The organization of sleep and waking patterns in the neonatal period has been considered a relevant indicator of the neuro-physiological maturation and underlying functioning of the brain and can be used to predict developmental outcome [16, 20, 21]. Moreover, sleep has implications on many physiological responses, such as the cardiovascular and respiratory systems [22, 23].

The behavior of newborns was first analyzed by Prechtl and Beintena in 1967 [24], who demonstrated the predictable occurrence of the observed physiological rhythms (circadian rhythms), subsequently called behavioral states (described in Table I.1).

State 1	Eyes closed, regular breathing, no movement
State 2	Eyes closed, irregular breathing, no gross movements
State 3	Eyes open, no gross movements
State 4	Eyes open, gross movements, no crying
State 5	Eyes open or closed, crying

Table I.1 – Grading of infant's states according Pretchl & Beintena [25].

Later, variables from the polysomnography such as respiration, heart rate, EEG and eye movements were included to this classification to become the well-known quiet, active sleep and wakefulness used to describe neonatal sleep-wake patterns during the first six months in infancy [26]. Although these sleep states exist in fetuses and preterm infants, their earliest age of appearance is still controversial [27, 28] and hence no general

consensus has been reached among neonatologists. Depending on the degree of maturity, the sleep-wake behavior can be classified according to the following states [29], albeit different versions exist:

Quiet sleep (QS): The breathing and heart ratio are predominantly regular. Eye or corporal movements are not observed. EEG is basically discontinuous.

Indeterminate sleep (IS): Contains patterns from both quiet and active sleep. Some authors do not differentiate indeterminate from active sleep.

Active sleep (AS): The heartbeat and respiratory rhythms are irregular. Sporadic movements occur and rapid eye movements (REM) are present intermittently. EEG activity is continuous and mixed.

Quiet wakefulness (QW): Infants appear alert with the eyes open and a quiet motor activity.

Active wakefulness (AW): Infants appear irritated or crying. QW patterns may be included.

The graphical representation of sleep states as a function of time is known as hypnogram (see an example in Figure I.3).

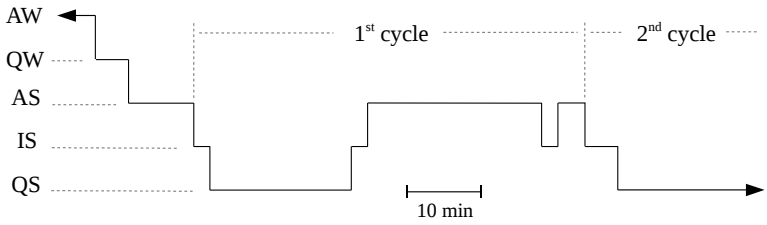

Figure I.3 – Example of hypnogram from a 35 weeks CA newborn, where the baby felt asleep in a short active sleep period. The first sleep cycle was measured from the end of the post-waking AS to the end of the next AS period (following a QS sleep period) [29].

Normal infants undergo a rapid evolution of the organization of the sleep states during the preterm period. Quiet sleep is rare at the earliest preterm ages but it increases progressively, whereas indeterminate sleep decreases in favor of an augmentation of active sleep (see Figure I.4) [29]. The amount of waking states also increases during this period, continuing beyond term ages [30].

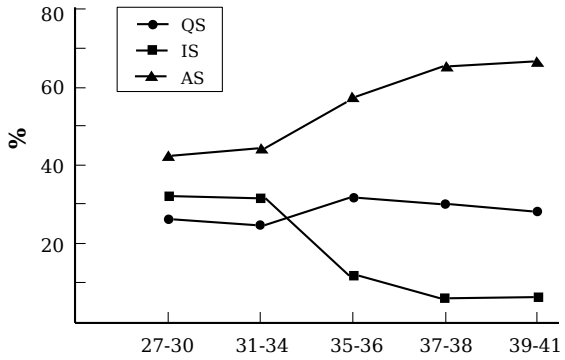

Figure I.4 – Evolution of the percentage of sleep states from 27 to 41 weeks of conceptional age in neurologically normal infants [29]. IS decreases progressively in favour of more differentiated sleep states (QS and AS).

2.2 Cardiorespiratory system

The still immature brain and underdeveloped lungs in the preterm infant are, without a doubt, the responsible of the instable manner they breathe [31, 32]. Actually, the main signatures of preterm's breathing are apneas or pauses in ventilation that may be accompanied by bradycardia, a decrease of the heart beat rhythm (described more extensively in section 3.2). The incidence of apnea is inversely related to the gestational age at birth and concerns the vast majority of very preterm and VLBW infants [33, 34].

The exact mechanisms underlying apnea have not been completely elucidated, but presumably, apnea is a consequence of the immaturity of the respiratory control centers. A fundamental source of breathing pattern generation is the pre-Bötzinger complex, located in the brain stem as a part of the ventral respiratory column, where the activity of pacemaker neurons results in the rhythmogenesis [35]. Other control centers apparently involved in apnea are chemoreceptors, due to their immature (exaggerated) responses to hypoxia and hypercapnia [36, 37].

On the other hand, the genesis of bradycardia –closely related to apnea– remains also unclear, but it seems to be strongly correlated with a decrease of oxygen saturation, suggesting a dependence on the hypoxic stimulation on the carotid body chemoreceptors [38].

Apnea-bradycardia may appear spontaneously, attributing this phenomena to the mere condition of prematurity but they can also be provoked or become more severe when infection, hypoxemia, thermal instability, cerebral pathology or metabolic problems is present [37, 39].

3 Risks due to prematurity

3.1 General risks

Premature newborns have many physiologic challenges when adapting to the extrauterine environment. During gestation, the fetus has all the needs covered into a safe and comfortable environment. Oxygen, food and antibodies are supplied by mother's placenta, the temperature is constant and external effects (gravity, light, sound) are imperceptible. When infants are born without completing the gestational cycle, they are not prepared to live in such a hostile environment and, the higher is the degree of prematurity, the more severe are their health problems. The most common are:

Cardiorespiratory: They are related to the immaturity of the lungs and the respiratory control system, such as respiratory distress syndrome, bronchopulmonary dysplasia and apnea-bradycardia.

Immunological: The still incomplete immunological system makes the preterm infants vulnerable to virus and fungi. Sepsis and pneumonia are common diseases.

Neurological: Cerebral palsy, intraventricular hemorrhage and other problems are consequences of their immature brain.

Thermical: The relatively large surface area compared to the infant's volume and the low portion of body fat makes difficult to maintain corporal temperature.

Digestive: The gastrointestinal system is not fully developed, so breast milk or formula cannot be digested adequately. Other problems, like the inability to coordinate sucking and swallowing and the gastroesophageal reflux are also present in the very preterm infant.

Metabolic: The immaturity of organs such as pancreas or liver may provoke hypoglycemia and jaundice.

Premature infants also have a greater risk of morbidity in the long term, including neurodevelopmental disabilities and behavioral sequelae in the transition into the adulthood [40]. In a recent survey, it was found that up to a 40% of 5-year-old children born before 33 weeks PMA showed cognitive deficits and or some degree of neuromotor deficits [41]. Moreover, with regard to the hospitalization in the neonatal care units, apnea of prematurity (AOP) and the high risk of sepsis still remain the major clinical problems.

3.2 Apnea of prematurity

AOP is defined as a cessation of the ventilatory activity during at least 15 seconds, or less than 15 seconds if this pause is accompanied by oxygen blood desaturation [42]. The respiratory pause may be central, obstructive, or mixed.

Central apnea: There is a complete absence of respiratory effort.

Obstructive apnea: Pharyngeal muscles, that maintain upper airways patency, reduce their tone collapsing the walls of the pharynx, obstructing the air flow.

Mixed: It presents characteristics of both apneas, i.e. an initial respiratory effort non sustained or the absence of effort, followed by the onset of a breathing effort.

Episodes of central apnea below 15-20 seconds can be normal at all ages, but a respiratory pause is abnormal if it is prolonged (20 seconds or more) or associated with cyanosis, hypotonia, or bradycardia. In this case, apnea is often reported as severe or pathologic [43].

Bradycardia is a decrease below the 30% in the baseline of the heart rate, normally related to apneas and/or oxygen desaturations. An episode of bradycardia may begin within 1.5 to 2 seconds of the onset of apnea and it can be produced indirectly, by the under-stimulation of the carotid body chemoreceptors, or directly, by the effect of hypoxia on the heart.

Even though the above definition of AOP is generally accepted to describe severe apnea, there is no consensus about the duration of nonpathologic or moderate apnea, i.e. brief respiratory pauses of less than 10 seconds. Typically, these are not associated with bradycardia or hypoxemia and can occur in conjunction with startles, movement or during feeding [44]. Among respiratory electrophysiologists is commonly accepted that an episode of moderate apnea can be considered if one of the two following situations are accomplished:

1. There is a cessation of breathing equivalent at least to three consecutive respiratory cycles.
2. There is a cessation of breathing lasting three times the average of a complete respiratory cycle.

In most instances, apnea can be resolved by cutaneous stimulation, so that the infant resumes breathing. However, the most severe episodes might require resuscitation with oxygen by bag and mask. Apneic episodes of 45 or more seconds lead to mottling, cyanosis, hypotonia and unresponsiveness to stimulation [13, 45]. The

administration of pharmacological agents from the family of methylxanthines (such as theophylline and caffeine) is nowadays the most widespread practice to reduce apnea frequency [46, 47]. This therapy can be assisted by continuous positive airway pressure (CPAP) when drug treatment alone is unresponsive, although it is ineffective to treat central apnea [48, 49]. In critically ill neonates, mechanical ventilation can also be applied.

Recurrent apnea-bradycardia might be a threat for the infants neurodevelopment because cerebral blood volume decreases, compromising the oxygenation of the cerebral tissue and provoking a depression of electroencephalographic activity [50, 51]. In the long term, a poor neuromotor prognosis at 3 years is related to the repetition of these episodes [52]. Hence, the continuous monitoring of breathing and cardiac frequencies are of crucial importance to an early intervention and avoid or palliate the associated risks of AOP.

3.3 Neonatal sepsis

Sepsis or septicemia is a generalized infection induced by bacteria, virus or fungi which affects the blood circulation. The underdeveloped immune system of preterm infants –during the last weeks of pregnancy the fetus is provided with antibodies– predisposes them to microbial infection in the first days of life. The risk to acquire infection is inversely related to gestational age, therefore preterm infants are more exposed to sepsis than term infants and, as a consequence, the morbidity and mortality attributed to this cause are increased.

Neonatal septicemia may be difficult to diagnose because of the multiplicity of associated risk factors and clinical manifestations, that may include respiratory distress, feeding intolerance, unstable temperature and cardiovascular depression [53, 54]. From a clinical point of view, two different kinds of sepsis with distinct etiologies and outcomes are defined: Early onset sepsis and Late onset sepsis.

Early onset sepsis (EOS): Acquired in the intrauterine period, it manifests in the first 48-72 hours after birth. The incidence is one to eight cases per 1000 live births, with the highest rates in VLBW preterm infants, but mortality directly related to this infection is low. The main causative organisms are bacteria, both Gram-positive and Gram-negative and there are several predispositions, including obstetric and infant factors [55].

Late onset sepsis (LOS): It is usually acquired at hospital and clinically evident more than 72 hours after birth. More than 50% of the neonates born at less than 25 weeks of gestation have LOS versus less than 10% of those born after 32 weeks. The usual vectors to transmit this category of infection are catheters and endotracheal intubation, being the vast majority (70%) of first episode late-onset infections caused by Gram-positive organisms. Infants developing LOS have a significantly prolonged hospital stay (mean length of stay: 79 vs 60 days) and are significantly more likely to die than those non infected (18% vs 7%) [56].

Complications of prematurity correlated with an increased rate of sepsis include patent ductus arteriosus, prolonged ventilation, prolonged intravascular access, bronchopulmonary dysplasia and necrotizing enterocolitis [56].

Since blood culture is necessary to evaluate LOS, the presumption of sickness requires prompt laboratory analysis and to start subsequently an antibiotic therapy. However, the haematological and biochemical markers that have been described require invasive procedures that cannot be frequently repeated and have low predictive value in the early phase of infection [57]. This is why, alternative, fast, noninvasive markers are desirable.

4 Neonatal intensive care units

Since their general establishment in the seventies, neonatal intensive care units have been evidenced to decline substantially neonatal mortality and have become an essential practice in perinatal medicine [58]. During hospitalization at NICU, sick term and premature infants receive specialized medical and nursing care to allow optimal neurodevelopment and outcome.

The central element of a NICU is the incubator (see an illustration in Figure I.5), a bassinet enclosed in plastic to keep the infant within an optimal temperature and humidity conditions, a minimal exposure to germs and external noise. The functionality of the incubator depends on the special needs of each newborn, but usually includes the following elements:

Breathing support: To avoid respiratory distress syndrome and reduce apneas, oxygen is supplemented by head hood or nasal cannula, CPAP or in case of severe sickness, mechanical ventilation.

Feeding and drug administration: An oral or nasogastric tube is used to provide to the infant nutrients and fluids, complemented if necessary by intravenous, umbilical or arterial catheters, administrating as well the medication.

Monitoring of vital signs: Cardiac activity, blood pressure and breathing are continuously monitored to detect apnea-bradycardia episodes.

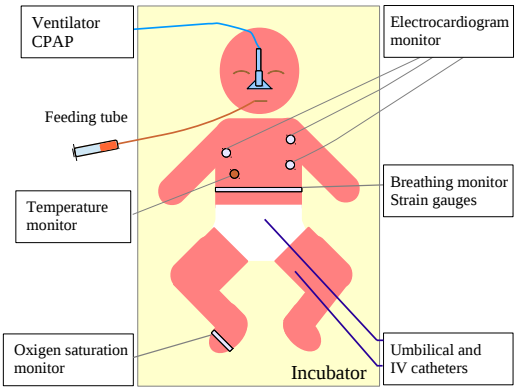

Figure I.5 – Illustration of some of the common elements in the neonatal intensive care unit.

4.1 Monitoring equipment

Monitoring devices at NICU allows the clinical staff to continuously supervise the status of the infant and start immediate nursing actions if alarms are triggered. These alarms can be configured to provide an alert when any of the vital signs being monitored crosses a given limit.

Cardiorespiratory monitor: It is attached to sensors on the infant's body and provides constant read-out of the heart and breathing rate, arterial or central venous pressure and other information. Recent monitors contain embedded computer systems capable of performing sophisticated functions: filtering out false alarms, recording, processing and analysis of data.

Temperature monitor: A temperature probe attached to the skin surface measures the infant's body temperature and activates the incubator warmer if necessary.

Pulse oximetry: This device measures the blood oxygen saturation and pulse rate noninvasively through a photodetector. It is used in most NICUs as a critical detector for desaturation and hypoxemia events.

Brain activity: Although the electroencephalographic activity is not, in general, monitored routinely, in some cases it is judged necessary to evaluate possible brain damage. Recording techniques such as the classical EEG montage using a few electrodes or the amplitude-integrated EEG (aEEG) allows long-time and continuous follow-up to detect easily seizures or depressed activity [59].

4.2 Cares and comfort at NICU

Some of the ordinary cares in the NICU include the vigilance of alarms, drug administration, feeding, calming the infants and communicate with parents. The most common causes of alarm are oxygen desaturation, bradycardia and apnea [60, 61], situations in which an immediate intervention is crucial to avoid severe harm to the patient. The estimated delay from the alarm triggering and the therapy application is 33 seconds in average, whereas the mean duration of the intervention to resolve the apnea-bradycardia episode, normally a slight tactile stimulation, is about 13 seconds [62].

Premature infants may also experience pain or discomfort during certain interventions at NICUs. Contrary to what was believed a decade ago, there are evidences that preterm babies respond to painful and noxious stimuli [63]. Moreover, the repetition of pain may have long-term adverse behavioral and physiological effects [64, 65]. The awareness of this fact contributed to ameliorate modern NICUs practices, where there is an increasing effort to minimize the stress of the baby and manage invasive procedures with an appropriate analgesia [66]. With this aim, new care approaches such as the Newborn Individualized Developmental Care and Assessment Program (NIDCAP) [67] provides a framework to improve the quality of life of both parents and baby.

In France, pain in preterm infants is commonly assessed with EDIN (*Échelle Douleur Inconfort Nouveau-Né*, or neonatal pain and discomfort scale) [68]. It consists on the sum of five indicators, each one scored from 0 (maximum level of comfort) to 3 (pain or discomfort maximum) according to nurses observations, so that the scale ranges from 0 to 15. These indicators include the facial activity, body movements, the quality of sleep, the quality of contact with nurses and the consolability. High scores (>10) denote a high level of discomfort and the consequential application of analgesia to relief pain.

4.3 Follow-up of maturation and discharge

Historically, preterm infants were discharged to home only when they achieved a certain weight, typically 2000 g. However, it has been shown that earlier discharge is possible without adverse health effects when preterm infants are evaluated by both medical stability and maturational criteria rather than body weight [69]. The three major criteria are:

1. Ability to feed by mouth to support appropriate growth.
2. Independent thermoregulation to maintain normal body temperature in a home environment.
3. Sufficiently mature respiratory control, considered if apnea of prematurity is resolved.

These competencies are achieved by most preterm infants between 36 and 37 weeks postmenstrual age [70], but discharge can be delayed if maturation of respiratory control is judged insufficient. Infants born at earlier gestation ages or at VLBW often suffer from more complicated medical courses and chronicity of neonatal illnesses, which tends to prolong the discharge predictions. Therefore, hospitalization should be continued until the cessation of severe apnea plus some days (margin of safety), occasionally requiring the use of telemonitoring at home [69, 71].

Stay at NICU for healthy premature infants varies widely in length among hospitals. It has been speculated that this variation relies in part from differences in monitoring, mechanical ventilation, hyperalimentation and decision about the timing of discharge [72, 73]. These differences may have implications not only for the health of infants, but also for the costs of care [74]. Certainly, prematurity is associated with high hospital costs. In the United States, it implies the 12% of total live births but consumes approximately 60% of the health budget allocated to neonatal services [75]. In France, the costs of prematurity in 2010 were estimated to 1.5 billion €, without taking into account long term cares [76].

5 Description of NICU signals

In previous sections, the origin and consequences of the immature brain activity and cardiorespiratory responses have been introduced. Here, the characteristics of perterm's signals acquired at NICU are studied in more details, focusing in their morphology and temporal evolution during prematurity. Since they are the basis of most of this work, EEG and breathing signals are described more extensively. The electrocardiogram, with a secondary role, is briefly introduced.

5.1 Preterm's electroencephalogram

EEG, having multiple appearances and interpretations depending on the electrode position, is without a doubt, the most complex signal in the NICU. The majority of works describing the characteristics of preterm EEG include channels from the full 10-20 system of electrodes or a restricted 10-20 system of nine scalp electrodes [77]. The reduced four-electrode system is almost exclusively employed for continued monitoring.

Discontinuous patterns

The main characteristic of the immature EEG is the discontinuity of the background activity. The intermittency of the endogenous activity seems to be crucial at early stages of development since it is the result of establishing and maintaining the connectivity of cortical networks.

The background activity in the preterm EEG corresponds to relatively inactive brain periods having amplitudes typically below 10 μV, although values up to 30 μV can be considered by some authors [78–81] depending on the number of channels, on the gestational age of the newborn or if it concerns analogical or digital recordings. Background activity can also be referred as hypoactivity periods and has to be distinguished from burst suppression, a pathological hypoactivity after major cerebral insults in older subjects.

Spontaneous activity transients (SAT) [82, 83] or EEG bursts refer to a short and physiological event arising from the background activity. A considerable amount of the preterm EEG activity is confined to these spontaneous, slow activity transients characterized by a large voltage deflection in several frequency bands, covering the whole frequency spectrum of the preterm EEG (0.3 - 30 Hz).

Inactivity periods between bursts are typically referred as interburst intervals (IBI) and, depending on their alternating patterns, they can be classified differently [84]:

- *Tracé continu* or continuous tracing: It is characterized by the persistence of theta and/or delta activity (> 25 μV) periods during at least 1 min (Figure I.6).

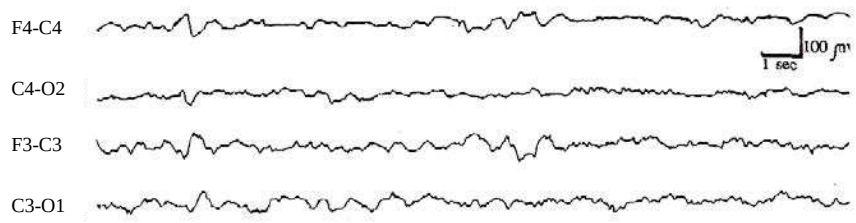

Figure I.6 – Example of *tracé continu* from a 40 weeks PMA infant in active sleep, with continuous mixed delta plus theta activity [85].

- *Tracé discontinu* or discontinuous tracing: It implies a sequence of physiological EEG bursts and hypoactivity periods variable in duration but exceeding 3 seconds. The portion of inactive periods should be at least the 50% in one minute (Figure I.7). Sometimes, the term used by some authors is *tracé semi-discontinu* or semidiscontinuous tracing.

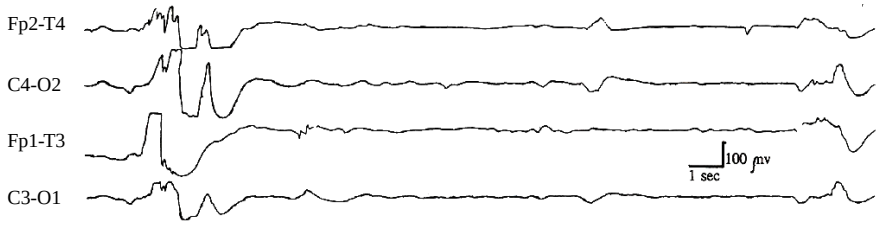

Figure I.7 – Example of *tracé discontinu* from a 27 weeks PMA infant in quiet sleep, with slow delta and theta bursts and very low (< 20 μV) background activity [85].

- *Tracé alternant*: The alternation of bursts of delta activity (75 to 100 μV) with lower voltage (25 to 50 μV) theta and delta background activity. The bursts appear about every 4-5 seconds and last 2-4 seconds (Figure I.8).

Symmetry, synchrony and characteristic graphoelements

The inter-hemispherical symmetry is described by the equality between the EEG figures from one hemisphere from the other. The term asymmetry is employed when there is, at least, a 50% of difference in the level of both hemispheres [86]. On the other hand, it can be observed that the physiological EEG figures do not appear simultaneously on the hemispheres. Hence, bilateral asynchrony refers to a delay of at least 1.5 seconds between both sides, but there is no general consensus about this value [78, 87]. While infants less than 30 weeks PMA

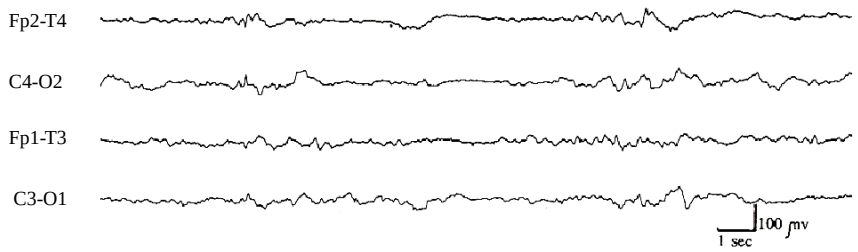

Figure I.8 – Example of *tracé alternant* in a 38 weeks PMA infant, with delta bursts of higher amplitude alternating with lower voltage theta and delta background activity [85].

exhibit almost complete synchrony, asynchrony is present in 30% of bursts at 31-32 weeks, decreasing to 20% by 33-34 weeks and disappearing about 37 weeks PMA [88, 89].

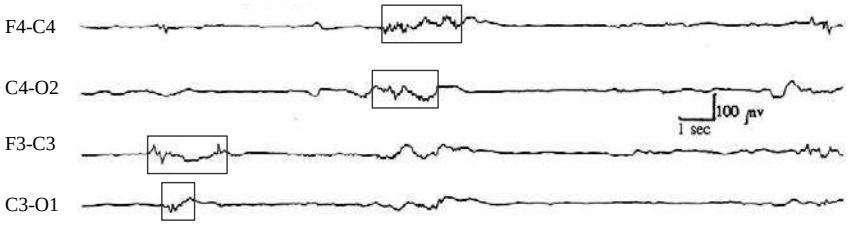

Figure I.9 – Discontinuous EEG from a 31 weeks PMA infant in quiet sleep, having delta brushes of rapid activity distributed asymmetrically (rectangles)[85].

Besides the organization of bursts and background activity, some characteristic waveforms are almost exclusive to the preterm EEG. They may include:

- Delta brushes: Also named *beta-delta complexes*, it constitutes one of the most relevant characteristic EEG graphoelements during prematurity. It consists on delta waves (0.3 to 1.5 Hz) of 50-250 μV with superimposed faster frequencies from 8 to 20 Hz (see Figure I.9) [90]. They are related to the development of the somatosensory cortex and their presence may be important for a normal maturation [91].

- Biphasic frontal sharp waves (*encoches frontales*) are sporadic sharp waves occurring simultaneously and symmetrically in the anterior head areas (See Figure I.10). They may appear before 35 weeks PMA but at this age they are maximally present [92].

- Anterior slow dysrhythmia is a normal pattern in premature infants, consisting on bursts of delta activity ($50-100\mu V$) lasting a few seconds on frontal regions [93] and usually associated with *encoches frontales*. They are present in all behavioral states although an increase of slow dysrhythmia may indicate encephalopathy [94].

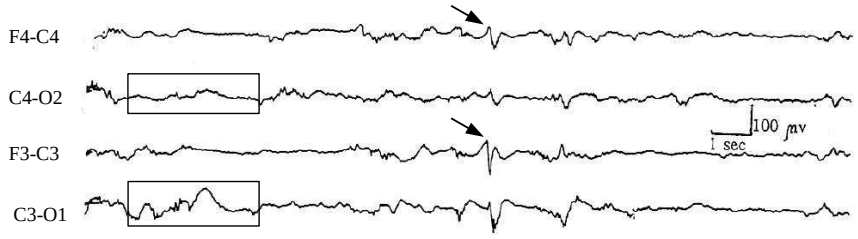

Figure I.10 – Semidiscontinous EEG tracing in a 36 weeks PMA infant in quiet sleep. A pair of *encoches frontales* are identified with arrows and asymmetric bursts are contained into rectangles [85].

Maturation of EEG and abnormalities

During the brain development of the preterm infant, the EEG patterns change in a predictable manner. The period between early and late prematurity (26 to 33 weeks PMA) coincides with a particularly fast development phase of the EEG, in which it is predominantly discontinuous. During this time, sensory systems are not completely functional but as brain wiring increases, the maturation is reflected in the organization of these discontinuities. Indeed, as it can be observed in Figure I.11, they show a significant relationship with PMA, shortening hypoactivity gradually with increasing maturation [95] until close to term ages, where *tracé alternant* appears.

In parallel, the background activity increases progressively in activity and frequency, resulting in a continuous EEG trace at term ages [84]. The main frequency of the EEG migrates from slow delta (0.3 - 1 Hz) to the fast delta and theta bands (3 - 4 Hz) and more complex figures appear as frontal transients. The SATs usually disappear by 40-42 weeks PMA.

Reactivity to stimuli and changes in behavioral states can be detected as early as 27 weeks, yet the sleep-wake cycles are not completely differentiated until the 31th week.

If neonatal EEG patterns suggest a postmenstrual age at least two weeks less than the actual age, the EEG is considered dysmature [96].

Abnormalities in preterm infants cover a broad spectrum of EEGs, making their classification a difficult task and, in some cases, arbitrary. One possibility for classification is to distinguish between abnormalities by their particular morphology and/or frequency and/or location and to categorize separately patterns associated to neonatal convulsions.

Many factors can perturb the normal discontinuous patterns in EEG and the sleep-wake cycling [97]. To enumerate just a few, it could be included drug administration, congenital malformations, cerebral insults, intraventricular hemorrhage and hipoxic ischemic encephalopaty (see Figure I.12). Low-voltage background activity ($< 20\ \mu V$) can indicate moderate encephalopathies and lower values ($< 5\ \mu V$) with intermittent short burts (burst suppression) are observed after severe asphyxia. The most extreme cases, prolonged periods of very low background activity during more than half an hour, are known as electro-cerebral silence, with fatal consequences for the infant's survival.

With different severity degrees, all these phenomena are correlated with a less favourable prognosis for cerebral development in the long term [94, 97].

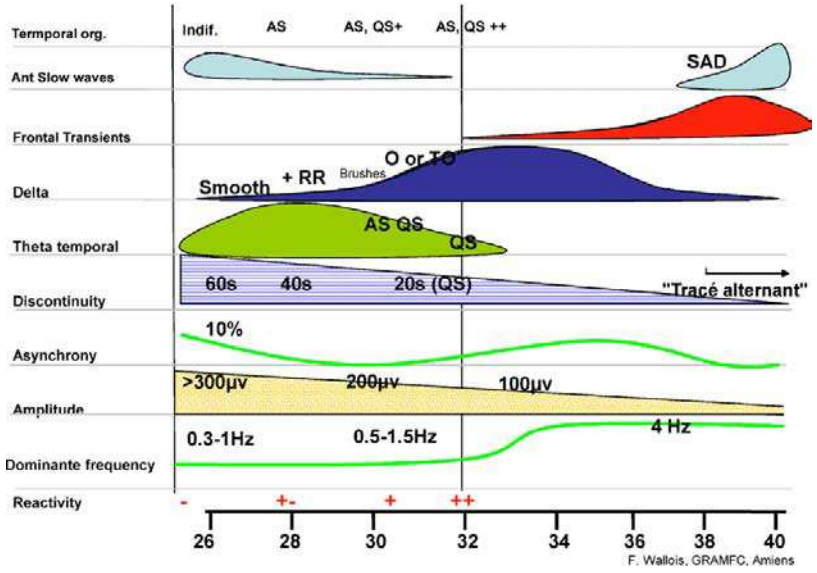

Figure I.11 – Evolution of the characteristics of preterm EEG with post-menstrual age. Legend: RR Rapid rhythms; TO: Temporo-occipital; QS, AS : Quiet and active sleep; SAD: slow anterior dysrhythmia.

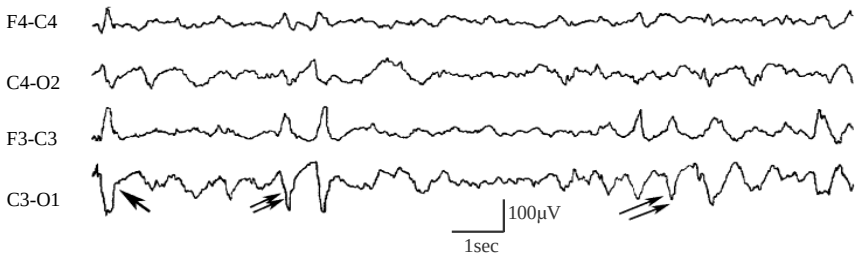

Figure I.12 – EEG from a 39 weeks PMA infant who suffered hypoxic ischemic encephalopathy. Arrows indicate epileptiform abnormalities (positive slow rolandic waves), alone (simple) and in bursts (double) [98].

5.2 Respiration

Breathing signals reflect the temporal evolution of some parameters describing the respiratory activity, such as the air pressure, volume or voltage generated by the muscles. It can be monitored in several ways, both invasive and non-invasive: measuring the airflow (pneumotachometry, spirometry), registering the electrical activity of the muscles with electrodes (plethysmography or even indirectly by extracting the signal from an electrocardiogram) or using movement sensors to measure the thorax distention, the technique employed to obtain our breathing signals. Another type of signal commonly used to study the dynamics of breathing is the respiratory variability (RV), formed by the succession of the breathing cycle times. Both signals, processed and analyzed in this work, are introduced in the following lines.

Immature breathing patterns

Preterm infants breathing is much faster than children's and adults' –having a normal breathing rate about 40 times each minute– but the major hallmark of these signals is, as underlined before, apnea of prematurity. Regardless of their origin, sighs and apneas have an important influence on the patterns that can be observed in breathing signals from preterm infants. Typically, they are classified in three categories:

Periodic breathing: The alternation of pauses lasting a few seconds followed by several rapid and shallow breaths of less than 30 seconds, associated to an increased sensitivity of the peripheral chemoreceptors [99]. Although it is more prevalent among preterm than full-term newborns, periodic breathing gradually resolves during infancy [100]. An example is given in Figure I.13-a.

Irregular or erratic breathing: Moderate and long apneas are frequently mixed, the respiratory rhythm and tidal volume are instable (see Figure I.13-b). The longest episodes of apnea are observed almost exclusively among infants having the most irregular respiratory rhythms [33]

Regular breathing: Ventilation is continuous at constant rate without interruptions by apnea, deep breaths or gross movements (see Figure I.13-c). It seems that newborns having higher levels of arterial O_2 are more likely to breathe regularly, whereas lower levels correspond to infants breathing periodically [101].

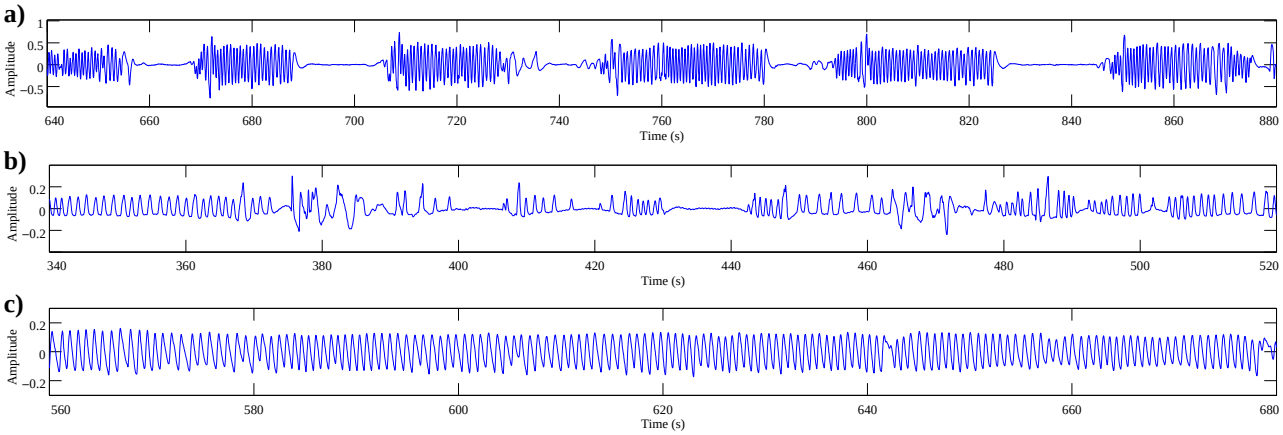

Figure I.13 – Examples of the three typical neonatal respiratory patterns. a) Periodic breathing, b) Erratic breathing and c) Regular breathing.

Respiratory variability

Respiratory variability or inter-breath interval signals are formed by collecting the time durations in all cycles of a breathing trace (see Figure I.14-a). Since the fluctuations in the abdominal trace are proportional to the variations of the tidal volume, a breathing cycle comprises the elapsed time between successive minima, called total time (t_{tot}). Inspiration (t_i) and expiration (t_e) times correspond to the first and second half-cycles delimited by the maximum (see Figure I.14-c).

Note that RV signals are constituted by a discrete sequence of time intervals, $X_i = t_{tot}(i)$, where index i is not proportional to time. They can also be resampled to obtain time-dependent signals, but introducing error due to interpolation.

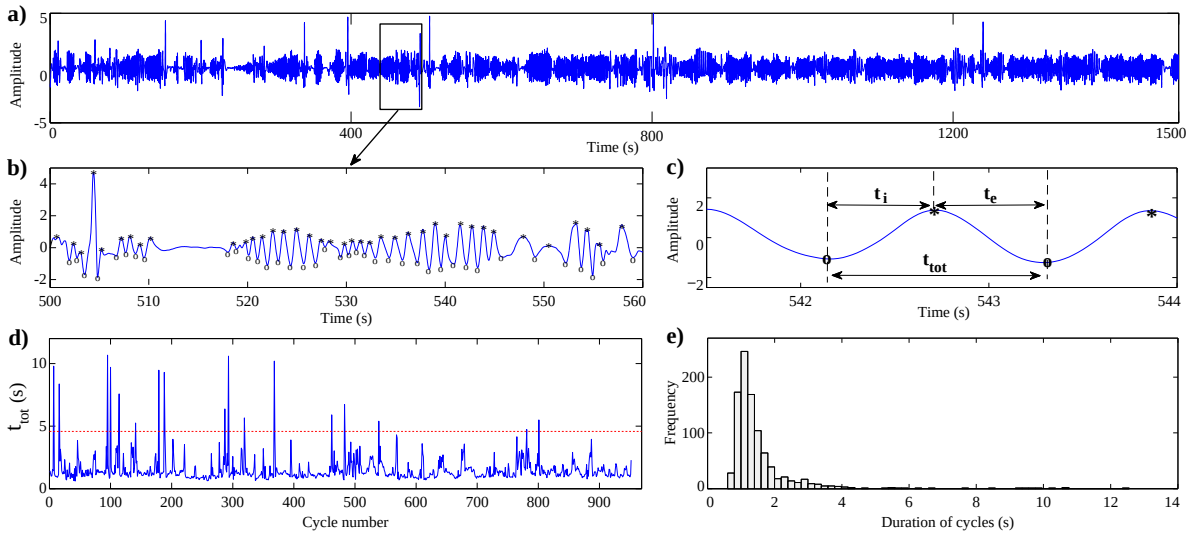

Figure I.14 – a) Example of a pre-processed respiratory signal corresponding to a patient born after 32 weeks of gestation. b) One-minute excerpt showing apnea and irregular breaths, with detected minima (o) and maxima (*). c) Definition of inspiration, expiration and total times within a cycle. d) Resulting RV signal with the horizontal line marking the apnea threshold ($3 * \overline{t_{tot}}$). e) Empirical distribution of the RV signal.

5.3 Electrocardiogram

The electrocardiogram is a basic signal for cardiac monitoring and continues to be the most extensively used tool to diagnose cardiovascular pathologies. The ECG measures the electrical activity originated in the heart by the depolarization and polarization of the myocardial cells. It is acquired by means of a set of electrodes placed on the chest surface.

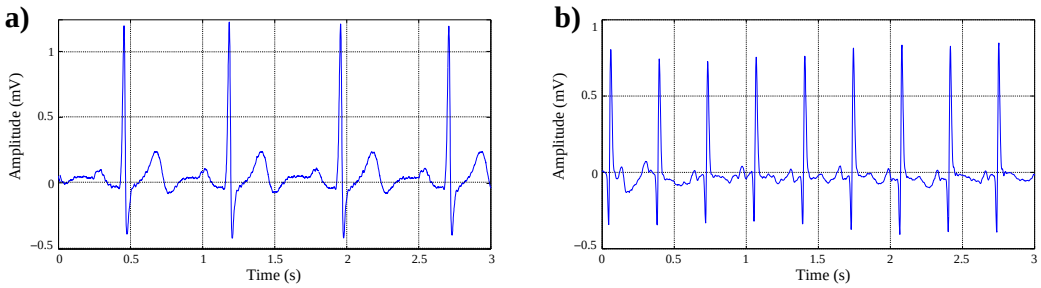

Figure I.15 – Examples of ECGs from a 20 years old healthy adult (a) and from a 32 weeks PMA healthy preterm infant (b) [102].

Newborns heart rate, and more precisely, the duration of PR intervals and QRS complexes, is substantially different from adults' (see Figure I.15). Due to the physiological evolution of the heart, during the firsts weeks of life, the normal cardiac frequency is about 90-180 beats per minute [103].

In the present work, the ECG is not analyzed because it has already extensively studied in other works in our research team (detailed in next section). However, as it appears as an artefact into the EEG, it will be employed in the denoising framework proposed in Chapter III.

6 Conclusion

Preterm birth is an increasing phenomenon worldwide, threatening the survival of the newborns and their subsequent quality of life. The higher incidences of apnea-bradycardia and early onset sepsis compared to full-term populations are, still, major challenges in neonatology.

Since its general implantation a few decades ago, the neonatal intensive care unit has played a crucial role to ameliorate the prognosis of these infants, but at the expenses of a considerable clinical effort. Consequently, the development of computer-assisted algorithms integrated in monitoring systems have been leading the research in this field to help the clinical staff improving diagnostic and starting treatments without delay.

To succeed in such a purpose, the knowledge of the signals acquired in the NICU and their physiological particularities is essential. Preterm's EEG, characterized by a highly non-stationary activity and immature breathing, having a variety of patterns conditioned by apnea, have been elucidated in this chapter to lay down the foundations and assumptions of this work.

Bibliography

[1] W. A. Engle. *Age terminology during the perinatal period.* Pediatrics **114**(5), 1362–1364 November (2004).

[2] W. H. Organization. March of dimes, PMNCH, save the children, WHO. Born too soon: the global action report on preterm birth. Technical report (2012).

[3] F. Venditttelli, O. Rivière, C. Crenn-Hébert, O. Claris, V. Tessier, D. Pinquier, F. Teurnier, J. Lansac, and B. Maria. *Réseau sentinelle audipog 2004–2005. partie 1 : résultats des principaux indicateurs périnatals.* Gynécologie Obstétrique & Fertilité **36**(11), 1091–1100 November (2008).

[4] J. A. Martin, B. E. Hamilton, P. D. Sutton, S. J. Ventura, T. J. Mathews, S. Kirmeyer, and M. J. K. Osterman. *Births: final data for 2010.* National vital statistics reports **61**(1), 1–72 (2012).

[5] H. Blencowe, S. Cousens, M. Z. Oestergaard, D. Chou, A.-B. Moller, R. Narwal, A. Adler, C. Vera Garcia, S. Rohde, L. Say, and J. E. Lawn. *National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications.* Lancet **379**(9832), 2162–2172 June (2012).

[6] J. Tucker and W. McGuire. *Epidemiology of preterm birth.* BMJ : British Medical Journal **329**(7467), 675–678 September (2004).

[7] R. L. Goldenberg, J. F. Culhane, J. D. Iams, and R. Romero. *Epidemiology and causes of preterm birth.* Lancet **371**(9606), 75–84 January (2008).

[8] D. R. Harding and A. D. Bocking. *Fetal Growth and Development.* Cambridge University Press (2001).

[9] I. Kostovic and N. Jovanov-Milosevic. *The development of cerebral connections during the first 20–45 weeks' gestation.* Seminars in Fetal and Neonatal Medicine **11**(6), 415–422 December (2006).

[10] P. R. Huttenlocher and A. S. Dabholkar. *Regional differences in synaptogenesis in human cerebral cortex.* The Journal of comparative neurology **387**(2), 167–178 October (1997).

[11] T. E. Inder, S. J. Wells, N. B. Mogridge, C. Spencer, and J. J. Volpe. *Defining the nature of the cerebral abnormalities in the premature infant: a qualitative magnetic resonance imaging study.* The Journal of pediatrics **143**(2), 171–179 August (2003).

[12] T. E. Inder, S. K. Warfield, H. Wang, P. S. Hüppi, and J. J. Volpe. *Abnormal cerebral structure is present at term in premature infants.* Pediatrics **115**(2), 286–294 February (2005).

[13] J. Kattwinkel, H. S. Nearman, A. A. Fanaroff, P. G. Katona, and M. H. Klaus. *Apnea of prematurity: Comparative therapeutic effects of cutaneous stimulation and nasal continuous positive airway pressure.* The Journal of Pediatrics **86**(4), 588–592 April (1975).

[14] S. J. Counsell and J. P. Boardman. *Differential brain growth in the infant born preterm: Current knowledge and future developments from brain imaging.* Seminars in Fetal and Neonatal Medicine **10**(5), 403–410 October (2005).

[15] E. Mercuri, A. Guzzetta, S. Laroche, D. Ricci, I. vanHaastert, A. Simpson, R. Luciano, C. Bleakley, M. F. Frisone, L. Haataja, G. Tortorolo, F. Guzzetta, L. de Vries, F. Cowan, and L. Dubowitz. *Neurologic examination of preterm infants at term age: Comparison with term infants.* The Journal of Pediatrics **142**(6), 647–655 June (2003).

[16] T. F. Anders, M. A. Keener, and H. Kraemer. *Sleep-wake state organization, neonatal assessment and development in premature infants during the first year of life. II.* Sleep **8**(3), 193–206 (1985).

[17] M.-F. Vecchierini, M. André, and A. d'Allest. *Normal EEG of premature infants born between 24 and 30 weeks gestational age: Terminology, definitions and maturation aspects.* Neurophysiologie Clinique/Clinical Neurophysiology **37**(5), 311–323 (2007).

[18] M. S. Scher. *Neurophysiological assessment of brain function and maturation. II. a measure of brain dysmaturity in healthy preterm neonates.* Pediatric neurology **16**(4), 287–295 May (1997).

[19] M. S. Scher, D. A. Steppe, D. G. Salerno, M. E. Beggarly, and D. L. Banks. *Temperature differences during sleep between fullterm and preterm neonates at matched post-conceptional ages.* Clinical neurophysiology: official journal of the International Federation of Clinical Neurophysiology **114**(1), 17–22 January (2003).

[20] D. H. Davis and E. B. Thoman. *Behavioral states of premature infants: implications for neural and behavioral development.* Developmental psychobiology **20**(1), 25–38 January (1987).

[21] D. Holditch-Davis and L. J. Edwards. *Temporal organization of sleep-wake states in preterm infants.* Developmental Psychobiology **33**(3), 257–269 November (1998).

[22] D. Holditch-Davis, L. J. Edwards, and M. C. Wigger. *Pathologic apnea and brief respiratory pauses in preterm infants: relation to sleep state.* Nursing research **43**(5), 293–300 October (1994).

[23] B. C. Galland, R. M. Hayman, B. J. Taylor, D. P. Bolton, R. M. Sayers, and S. M. Williams. *Factors affecting heart rate variability and heart rate responses to tilting in infants aged 1 and 3 months.* Pediatric research **48**(3), 360–368 September (2000).

[24] H. F. Prechtl. *The behavioural states of the newborn infant (a review).* Brain research **76**(2), 185–212 August (1974).

[25] L. M. S. Dubowitz, V. Dubowitz, and E. Mercuri. *The Neurological Assessment of the Preterm and Full-term Newborn Infant.* Cambridge University Press February (1999).

[26] J. Dittrichová, V. Brichacek, K. Paul, and M. Tautermannova. *The structure of infant behavior: an analysis of sleep and waking in the first months of life.* Review of Child Development Research pages 6:73–100 (1982).

[27] L. Curzi-Dascalova, P. Peirano, and F. Morel-Kahn. *Development of sleep states in normal premature and full-term newborns.* Developmental psychobiology **21**(5), 431–444 July (1988).

[28] M. Mirmiran, Y. G. Maas, and R. L. Ariagno. *Development of fetal and neonatal sleep and circadian rhythms.* Sleep Medicine Reviews **7**(4), 321–334 August (2003).

[29] L. Curzi-Dascalova. *Développement du sommeil et des fonctions sous contrôle du système nerveux autonome chez le nouveau-né prématuré et à terme.* Archives de Pédiatrie **2**(3), 255–262 March (1995).

[30] D. Holditch-Davis, M. Scher, T. Schwartz, and D. Hudson–Barr. *Sleeping and waking state development in preterm infants.* Early Human Development **80**(1), 43–64 October (2004).

[31] M. Durand, L. A. Cabal, F. Gonzalez, S. Georgie, C. Barberis, T. Hoppenbrouwers, and J. E. Hodgman. *Ventilatory control and carbon dioxide response in preterm infants with idiopathic apnea.* American journal of diseases of children (1960) **139**(7), 717–720 July (1985).

[32] R. J. Martin, M. J. Miller, and W. A. Carlo. *Pathogenesis of apnea in preterm infants.* The Journal of pediatrics **109**(5), 733–741 November (1986).

[33] H. C. Miller, F. C. Behrle, and N. W. Smull. *Severe apnea and irregular respiratory rhythms among premature infants. A clinical and laboratory study.* Pediatrics **23**(4), 676–685 April (1959).

[34] E. R. Alden, T. Mandelkorn, D. E. Woodrum, R. P. Wennberg, C. R. Parks, and W. A. Hodson. *Morbidity and mortality of infants weighing less than 1,000 grams in an intensive care nursery.* Pediatrics **50**(1), 40–49 July (1972).

[35] J. C. Rekling and J. L. Feldman. *PreBötzinger complex and pacemaker neurons: hypothesized site and kernel for respiratory rhythm generation.* Annual review of physiology **60**, 385–405 (1998).

[36] R. J. Martin and A. A. Fanaroff. *Neonatal apnea, bradycardia, or desaturation: Does it matter?* The Journal of Pediatrics **132**(5), 758–759 May (1998).

[37] J. M. Abu-Shaweesh and R. J. Martin. *Neonatal apnea: What's new?* Pediatric Pulmonology, Pediatric Pulmonology **43**(10), 937–944 October (2008).

[38] D. J. Henderson-Smart, M. C. Butcher-Puech, and D. A. Edwards. *Incidence and mechanism of bradycardia during apnoea in preterm infants.* Archives of Disease in Childhood **61**(3), 227–232 March (1986).

[39] C. F. Poets, V. A. Stebbens, M. P. Samuels, and D. P. Southall. *The relationship between bradycardia, apnea, and hypoxemia in preterm infants.* Pediatric research **34**(2), 144–147 August (1993).

[40] S. Saigal and L. W. Doyle. *An overview of mortality and sequelae of preterm birth from infancy to adulthood.* The Lancet **371**(9608), 261–269 (2008).

[41] B. Larroque, P.-Y. Ancel, S. Marret, L. Marchand, M. André, C. Arnaud, V. Pierrat, J.-C. Rozé, J. Messer, G. Thiriez, A. Burguet, J.-C. Picaud, G. Bréart, and M. Kaminski. *Neurodevelopmental disabilities and special care of 5-year-old children born before 33 weeks of gestation (the EPIPAGE study): a longitudinal cohort study.* Lancet **371**(9615), 813–820 March (2008).

[42] C. o. F. a. Newborn. *Apnea, sudden infant death syndrome, and home monitoring.* Pediatrics **111**(4), 914–917 April (2003).

[43] K. P. Dunne and T. G. Matthews. *Infantile apnoea and home monitoring.* British Medical Journal (Clinical research ed.) **295**(6597), 553 August (1987).

[44] B. T. Thach and A. R. Stark. *Spontaneous neck flexion and airway obstruction during apneic spells in preterm infants.* The Journal of pediatrics **94**(2), 275–281 February (1979).

[45] W. J. R. Daily, M. Klaus, H. Belton, and P. Meyer. *Apnea in premature infants: Monitoring, incidence, heart rate changes, and an effect of environmental temperature.* Pediatrics **43**(4), 510–518 April (1969).

[46] J. V. Aranda and T. Turmen. *Methylxanthines in apnea of prematurity.* Clinics in perinatology **6**(1), 87–108 March (1979).

[47] D. J. Henderson-Smart and P. Steer. *Methylxanthine treatment for apnea in preterm infants.* Cochrane database of systematic reviews (Online) (3), CD000140 (2001).

[48] M. J. Miller, W. A. Carlo, and R. J. Martin. *Continuous positive airway pressure selectively reduces obstructive apnea in preterm infants.* The Journal of Pediatrics **106**(1), 91–94 January (1985).

[49] B. Schmidt, R. S. Roberts, P. Davis, L. W. Doyle, K. J. Barrington, A. Ohlsson, A. Solimano, and W. Tin. *Caffeine therapy for apnea of prematurity.* New England Journal of Medicine **354**(20), 2112–2121 (2006).

[50] V. T. Ramaekers, P. Casaer, and H. Daniels. *Cerebral hyperperfusion following episodes of bradycardia in the preterm infant.* Early Human Development **34**(3), 199–208 October (1993).

[51] G. Pichler, B. Urlesberger, and W. M ller. *Impact of bradycardia on cerebral oxygenation and cerebral blood volume during apnoea in preterm infants.* Physiological Measurement **24**(3), 671–680 August (2003).

[52] A. Janvier, M. Khairy, A. Kokkotis, C. Cormier, D. Messmer, and K. J. Barrington. *Apnea is associated with neurodevelopmental impairment in very low birth weight infants.* Journal of perinatology: official journal of the California Perinatal Association **24**(12), 763–768 December (2004).

[53] J. S. Gerdes. *Diagnosis and management of bacterial infections in the neonate.* Pediatric clinics of North America **51**(4), 939–959, viii–ix August (2004).

[54] W. McGuire, L. Clerihew, and P. W. Fowlie. *Infection in the preterm infant.* BMJ : British Medical Journal **329**(7477), 1277–1280 November (2004).

[55] B. J. Stoll, T. Gordon, S. B. Korones, S. Shankaran, J. E. Tyson, C. R. Bauer, A. A. Fanaroff, J. A. Lemons, E. F. Donovan, W. Oh, D. K. Stevenson, R. A. Ehrenkranz, L.-A. Papile, J. Verter, and L. L. Wright. *Early-onset sepsis in very low birth weight neonates: A report from the national institute of child health and human development neonatal research network.* The Journal of Pediatrics **129**(1), 72–80 July (1996).

[56] B. J. Stoll, N. Hansen, A. A. Fanaroff, L. L. Wright, W. A. Carlo, R. A. Ehrenkranz, J. A. Lemons, E. F. Donovan, A. R. Stark, J. E. Tyson, W. Oh, C. R. Bauer, S. B. Korones, S. Shankaran, A. R. Laptook, D. K. Stevenson, L.-A. Papile, and W. K. Poole. *Late-onset sepsis in very low birth weight neonates: the experience of the NICHD neonatal research network.* Pediatrics **110**(2 Pt 1), 285–291 August (2002).

[57] A. Malik, C. P. S. Hui, R. A. Pennie, and H. Kirpalani. *Beyond the complete blood cell count and c-reactive protein: a systematic review of modern diagnostic tests for neonatal sepsis.* Archives of Pediatrics & Adolescent Medicine **157**(6), 511–516 June (2003).

[58] D. K. Richardson, J. E. Gray, S. L. Gortmaker, D. A. Goldmann, D. M. Pursley, and M. C. McCormick. *Declining severity adjusted mortality: Evidence of improving neonatal intensive care.* Pediatrics **102**(4), 893–899 October (1998).

[59] L. Hellström-Westas. *Comparison between tape-recorded and amplitude-integrated EEG monitoring in sick newborn infants.* Acta Pædiatrica **81**(10), 812–819 (1992).

[60] S. T. Lawless. *Crying wolf: false alarms in a pediatric intensive care unit.* Critical care medicine **22**(6), 981–985 June (1994).

[61] Y. Bitan, J. Meyer, D. Shinar, and E. Zmora. *Nurses' reactions to alarms in a neonatal intensive care unit.* Cogn. Technol. Work **6**(4), 239–246 November (2004).

[62] R. Pichardo, J. S. Adam, E. Rosow, J. Bronzino, and L. Eisenfeld. *Vibrotactile stimulation system to treat apnea of prematurity.* Biomedical Instrumentation & Technology **37**(1), 34–40 January (2003).

[63] K. J. S. Anand. *Pain, plasticity, and premature birth: a prescription for permanent suffering?* Nature Medicine **6**(9), 971–973 (2000).

[64] K. J. Anand, V. Coskun, K. V. Thrivikraman, C. B. Nemeroff, and P. M. Plotsky. *Long-term behavioral effects of repetitive pain in neonatal rat pups.* Physiology & behavior **66**(4), 627–637 June (1999).

[65] A. T. Bhutta, M. A. Cleves, P. H. Casey, M. M. Cradock, and K. J. S. Anand. *Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis.* JAMA: the journal of the American Medical Association **288**(6), 728–737 August (2002).

[66] American Academy of Pediatrics. *Prevention and management of pain and stress in the neonate. Committee on Fetus and Newborn. Committee on Drugs. Section on Anesthesiology. Section on Surgery. Canadian Paediatric Society. Fetus and Newborn Committee.* Pediatrics **105**(2), 454–461 February (2000).

[67] H. Als. *Newborn individualized developmental care and assessment program (NIDCAP): new frontier for neonatal and perinatal medicine.* Journal of Neonatal-Perinatal Medicine **2**(3), 135–147 January (2009).

[68] T. Debillon, V. Zupan, N. Ravault, J. Magny, M. Dehan, and H. ABU-SAAD. *Development and initial validation of the EDIN scale, a new tool for assessing prolonged pain in preterm infants.* Archives of Disease in Childhood. Fetal and Neonatal Edition **85**(1), F36–F41 July (2001).

[69] C. o. F. a. Newborn. *Hospital discharge of the high-risk neonate.* Pediatrics **122**(5), 1119–1126 November (2008).

[70] P. J. Powell, C. V. Powell, S. Hollis, and M. J. Robinson. *When will my baby go home?* Archives of disease in childhood **67**(10 Spec No), 1214–1216 October (1992).

[71] S. A. Lorch, L. Srinivasan, and G. J. Escobar. *Epidemiology of apnea and bradycardia resolution in premature infants.* Pediatrics **128**(2), e366–e373 August (2011).

[72] E. C. Eichenwald, M. Blackwell, J. S. Lloyd, T. Tran, R. E. Wilker, and D. K. Richardson. *Inter-neonatal intensive care unit variation in discharge timing: influence of apnea and feeding management.* Pediatrics **108**(4), 928–933 October (2001).

[73] M. C. McCormick, G. J. Escobar, Z. Zheng, and D. K. Richardson. *Place of birth and variations in management of late preterm ("near-term") infants.* Seminars in perinatology **30**(1), 44–47 February (2006).

[74] W. M. Gilbert, T. S. Nesbitt, and B. Danielsen. *The cost of prematurity: quantification by gestational age and birth weight.* Obstetrics and gynecology **102**(3), 488–492 September (2003).

[75] N.-H. Morken. *Preterm birth: new data on a global health priority.* The Lancet **379**(9832), 2128–2130 June (2012).

[76] D. de la Recherche, des Etudes, de l'Evaluation et des Statistiques du Ministère de la Santé. La situation périnatale en France en 2010. Technical Report Etudes et Résultats num. 775 (2011).

[77] A. W. De Weerd, P. A. Despland, and P. Plouin. *Neonatal EEG. the international federation of clinical neurophysiology.* Electroencephalography and clinical neurophysiology. Supplement **52**, 149–157 (1999).

[78] C. M. Anderson, F. Torres, and A. Faoro. *The EEG of the early premature.* Electroencephalography and clinical neurophysiology **60**(2), 95–105 February (1985).

[79] G. I. Benda, R. C. Engel, and Y. P. Zhang. *Prolonged inactive phases during the discontinuous pattern of prematurity in the electroencephalogram of very-low-birthweight infants.* Electroencephalography and clinical neurophysiology **72**(3), 189–197 March (1989).

[80] E. Biagioni, L. Bartalena, A. Boldrini, G. Cioni, S. Giancola, and A. E. Ipata. *Background EEG activity in preterm infants: correlation of outcome with selected maturational features.* Electroencephalography and clinical neurophysiology **91**(3), 154–162 September (1994).

[81] M. Hayakawa, A. Okumura, F. Hayakawa, K. Watanabe, M. Ohshiro, Y. Kato, R. Takahashi, and N. Tauchi. *Background electroencephalographic (EEG) activities of very preterm infants born at less than 27 weeks gestation: a study on the degree of continuity.* Archives of Disease in Childhood. Fetal and Neonatal Edition **84**(3), F163–F167 May (2001).

[82] S. Vanhatalo, J. M. Palva, S. Andersson, C. Rivera, J. Voipio, and K. Kaila. *Slow endogenous activity transients and developmental expression of K+-Cl- cotransporter 2 in the immature human cortex.* European Journal of Neuroscience **22**(11), 2799–2804 (2005).

[83] S. Vanhatalo and K. Kaila. *Development of neonatal EEG activity: from phenomenology to physiology.* Seminars in fetal & neonatal medicine **11**(6), 471–478 December (2006).

[84] M. D. Lamblin, M. André, M. J. Challamel, L. Curzi-Dascalova, A. M. d'Allest, E. De Giovanni, F. Moussalli-Salefranque, Y. Navelet, P. Plouin, M. F. Radvanyi-Bouvet, D. Samson-Dollfus, and M. F. Vecchierini-Blineau. *Electroencephalography of the premature and term newborn. maturational aspects and glossary.* Clinical Neurophysiology **29**(2), 123–219 April (1999).

[85] L. Curzi-Dascalova and M. Mirmiran. *Manuel des techniques d'enregistrement et d'analyse des stades de sommeil et de veille chez le prématuré et le nouveau-né à terme.* Les Editions INSERM, Paris, France (1996).

[86] R. R. Clancy, H. J. Chung, and J. P. Temple. *Neonatal Electroencephalography.* Elsevier (1993).

[87] K. Aso, M. S. Scher, and M. A. Barmada. *Neonatal electroencephalography and neuropathology.* Journal of clinical neurophysiology: official publication of the American Electroencephalographic Society **6**(2), 103–123 April (1989).

[88] C. T. Lombroso. *Neonatal polygraphy in full-term and premature infants: a review of normal and abnormal findings.* Journal of clinical neurophysiology: official publication of the American Electroencephalographic Society **2**(2), 105–155 April (1985).

[89] B. R. Tharp. *Electrophysiological brain maturation in premature infants: an historical perspective.* Journal of clinical neurophysiology: official publication of the American Electroencephalographic Society **7**(3), 302–314 July (1990).

[90] C. Dreyfus-Brisac. *Sleep ontogenesis in early human prematurity from 24 to 27 weeks of conceptional age.* Developmental Psychobiology **1**(3), 162–169 (1968).

[91] R. Khazipov, A. Sirota, X. Leinekugel, G. L. Holmes, Y. Ben-Ari, and G. Buzsáki. *Early motor activity drives spindle bursts in the developing somatosensory cortex.* Nature **432**(7018), 758–761 December (2004).

[92] C. Dreyfus-Brisac. *Ontogenesis of sleep in human prematures after 32 weeks of conceptional age.* Developmental Psychobiology **3**(2), 91–121 (1970).

[93] M. Tolonen, J. Palva, S. Andersson, and S. Vanhatalo. *Development of the spontaneous activity transients and ongoing cortical activity in human preterm babies.* Neuroscience **145**(3), 997–1006 March (2007).

[94] G. L. Holmes and C. T. Lombroso. *Prognostic value of background patterns in the neonatal EEG.* Journal of clinical neurophysiology: official publication of the American Electroencephalographic Society **10**(3), 323–352 July (1993).

[95] E. Biagioni, M. Frisone, S. Laroche, B. Kapetanakis, D. Ricci, M. Adeyi-Obe, H. Lewis, N. Kennea, G. Cioni, F. Cowan, M. Rutherford, D. Azzopardi, and E. Mercuri. *Maturation of cerebral electrical activity and development of cortical folding in young very preterm infants.* Clinical Neurophysiology **118**(1), 53–59 January (2007).

[96] C. T. Lombroso. *Neurophysiological observations in diseased newborns.* Biological psychiatry **10**(5), 527–558 October (1975).

[97] K. Watanabe, S. Miyazaki, K. Hara, and S. Hakamada. *Behavioral state cycles, background EEGs and prognosis of newborns with perinatal hypoxia.* Electroencephalography and clinical neurophysiology **49**(5-6), 618–625 September (1980).

[98] D. Selton, M. André, and J. Hascoët. *EEG et accident vasculaire cérébral ischémique du nouveau-né à terme.* Neurophysiologie Clinique/Clinical Neurophysiology **33**(3), 120–129 June (2003).

[99] K. J. Barrington and N. N. Finer. *Periodic breathing and apnea in preterm infants.* Pediatric Research **27**(2), 118–121 (1990).

[100] H. Rigatto and J. P. Brady. *Periodic breathing and apnea in preterm infants. i. evidence for hypoventilation possibly due to central respiratory depression.* Pediatrics **50**(2), 202–218 August (1972).

[101] A. Al-Matary, I. Kutbi, M. Qurashi, M. Khalil, R. Alvaro, K. Kwiatkowski, D. Cates, and H. Rigatto. *Increased peripheral chemoreceptor activity may be critical in destabilizing breathing in neonates.* Seminars in perinatology **28**(4), 264–272 August (2004).

[102] M. Altuve. *Détection multivariée des épisodes d'apnée-bradycardie chez le prématuré par modèles semi-Markovien cachés.* Thèse de Doctorat, Université de Rennes 1 (2011).

[103] G. Q. Sharieff and S. O. Rao. *The pediatric ECG.* Emergency medicine clinics of North America **24**(1), 195–208, vii–viii February (2006).

Chapter II

Problem statement and aims of the study

Unquestionably, the improvement of the technologies in neonatal monitoring equipment, allowing the real-time vital signs vigilance and the design of alarms, has ameliorated substantially the quality of preterm infants health in the last two decades. But also, it has derived concurrently in a more sophisticated off-line analysis of physiological signals, leading to the research of new, fast and non-invasive indicators with a potential use as diagnostic and predictive tools.

As the major concerns in preterm infants are apnea of prematurity and sepsis, these issues have aroused a vast number of research works. On the other hand, the study of early EEG patterns continues to be a common procedure to evaluate brain maturation, outcome prediction, epilepsy and abnormalities. Sleep, distress and pain are also attracting the attention of scientists as the comfort of the premature newborn seems to play an important role for the correct neurodevelopment at such stages of life [1, 2].

In this chapter we point out the most remarkable research with regard to the study of physiological signals from preterm infants at NICU, among them the previous works of the SEPIA (*Surveillance, Explication et Prévention de l'Insuffisance cardiaque et de l'Apnée-bradycardie*) team from the LTSI-INSERM (University of Rennes 1), working in close collaboration with the University Hospital of Rennes (CHU). The database employed in this study, as well as the main objectives respecting the clinical problem, are also described.

1 State of the art

1.1 Apnea-bradycardia detection

Many works have been devoted to uncover the possible physiopathological causes of apnea-bradycardia and their consequences and to find the best therapy (see [3] and [4] for complete reviews). But during the last two decades, the clinical management of AOP at neonatal intensive care units has not changed much: breathing assistance and/or administration of caffeine when the risk of AOP is established, and cutaneous stimulation or resuscitation when a severe episode of apnea is declared. In this sense, technological advancements are providing tools helping the clinical staff, like vibrotactile skin stimulation, designed to arouse the infant if an apnea is detected [5, 6].

Recent research is orientated to the enhancement of AOP detections, not only to minimize the false alarms but also in prediction so that they can be triggered before the recrudescence of the apnea and hence, nursing actions can be started without delay [7–9].

The SEPIA team of LTSI has been engaged in the last years with the investigation of apnea-bradycardia and the design of support technologies for NICUs. A multidisciplinary team works from the assessment of animal

models in order to shed some light on the ontogeny of apnea of prematurity [10–12], in innovative algorithms to an early detection of bradycardia [13–16] and the integration of the latter to a patented intelligent monitoring system (INTEM) [17, 18].

A. Beuchée has been leading a vast study of the heart rate variability to observe perturbations in the cardiovascular regulation, testing and proposing analysis tools from animal experiments to human preterm data [19, 20]. His research found that complexity in the cardiac rhythm reflects the autonomic nervous system [13] and can be used to characterize populations having AOP. He proposed as well an algorithm to the early detection of bradycardia and the design of an automatic kinesthetic stimulator [21]. More recently, M. Altuve contributed to this investigation with a multivariate analysis from ECG data obtained in the NICU [16], proposed an improved R-complex detector for INTEM [22] and developed a method to predict bradycardia online from the dynamics of HRV applying hidden semi-Markovian models [15] to RR series.

1.2 Early detection of infection and effects of immunization

The diagnosis of late onset sepsis in premature infants remains difficult because clinical signs are subtle and nonspecific and none of the laboratory tests have high predictive accuracy. However, it seems clear that recurrent and severe spontaneous apneas, bradycardias and irregularities in the heart beat ratio may itself reveal systemic infection [23, 24]. Nonlinear indexes like entropy and long-range correlations have been proved to be decreased in septicemic infants [25, 26]. Therefore, the analysis of cardiac and breathing activity has emerged as a promising diagnostic tool.

The SEPIA team has also been involved with the early diagnosis of sepsis. Since 2003, a database designed to this purpose, is being constructed with NICU signals from preterm infants born at CHU Rennes. To the present day, the number of exploitable records is about 70, many of them used in the present study. The analysis of RR series from this database not only corroborated the above findings [27], but also provided new insights in the relationships between HRV and spontaneous respiration as potential indicator of sepsis, using time-frequency techniques developed by G. Carrault [28] and several indexes from linear correlation, coherence and nonlinear regression [29]. Infected newborns were found to have a lower degree of cardiorespiratory coupling than the healthy population.

Concerning immunization of very preterm infants, little is known about the consequences of vaccination on their cardiovascular system. This motivated the work of T. Mialet-Marty at CHU Rennes, where a specific protocol and data collection was designed to evaluate changes in heart rate and respiratory variability signals before and after vaccination [30]. The database of this study, composed by 31 very preterm infants (<32 weeks GA and >7 weeks chronological age) constitutes the second source of data of this thesis. Among the main findings, the post-vaccinal group had higher incidences of AOP and bradycardia, suggesting an adaptation of the cardiovascular system facing disturbances induced by the inflammatory response to immunization [31].

1.3 Preterm's brain activity and sleep

It is now generally accepted that the behavioral observations together with the analysis of EEG signals are the main sources of information to assess the preterm's neurological status and maturation. In effect, observing maturational criteria, experts can assess EEG maturity with 2 weeks of accuracy with respect the newborn's stated age. A mismatch between the EEG patterns and preterm's physiological age alerts clinicians about possible neurological problems [32].

Multichannel EEG has been used traditionally by clinicians to analyze the main patterns related to the brain maturity, but it requires to place multiple electrodes over the small scalp surface of preterm infants. In practice, preparing this montages is not only laborious for the clinical staff, but also uncomfortable for the newborn, disturbing sleep and provoking frequent muscle artifacts. In general, the use of a smaller number of channels ameliorates the sleep quality in newborns, reducing the amount of movement artifacts. For instance, Holthausen et al. [33] used central-temporal leads (C3-T3 and C4-T4) to obtain dysmaturity scores, stating that these channels suffice to represent maturity-related patterns with a small amount of noise. Employing the same montage but using a Bayesian approach, Jakaite et al. [34] assessed automatically the brain maturity of newborns aged 36 weeks and older with similar results than visual expert's observations from polysomnograms.

The temporal organization of the inter-burst intervals in early EEG patterns follows a continuous evolution from preterm to post-term ages [35, 36], hence the automatic classification of EEG activity, using both temporal and spectral characteristics, has called the attention of many researchers [37, 38].

Sleep is also an important milestone of the infant's maturation. Nowadays, its study is still based in the observations by well-trained clinicians and, in spite of the help of polysomnographic signals, scoring sleep states is a time-consuming task with a non-negligible level of inter-rater discrepancy. As polysomnography consists in complex montages with many electrodes and can result intrusive, there is an increasing interest in minimizing the number of channels and physiological signals. In full-term infants, a large number of works proposed automatic sleep-wake classification with relatively good performances (60 to 80%) comparing to "gold standards" (manual scorings), employing only cardiac and breathing measures [39], ECG and actigraphy [40], and even on the heart rate [41] or body movements [42] alone. The automatic sleep classification in preterm infants seems to be much more investigated at term ages than during the preterm period, presumably due to the ambiguity of certain states and their rapid evolution. Works dealing with premature babies at term ages (40 weeks PMA), state that classification criteria using EEG differs significantly from full-term cohorts and performances are slightly decreased [43, 44]. In preterm infants, studying sleep-wake cyclicity from the EEG results a more feasible task, opening the possibility to monitor the brain maturity in NICUs [45].

The LTSI, engaged with the design of state-of-the-art monitoring equipment, has also been implicated in the design of home devices with a reduced number of electrodes to score sleep in adults. In these environments, noise corruption and the lack of traditionally-employed signals can be a difficulty to score reliably the different states. Using a minimal configuration with four electrodes (the bilateral fronto-parietal pair), F. Porée proposed a blind source separation solution to obtain EMG, EOG and ECG signals and showed that subsequent sleep stage classification provided good results [46].

2 Databases

The data processed and analyzed in this work has been collected at the neonatology service of the University Hospital of Rennes (CHU) over different periods. Parents were informed and written consent was obtained. Studies involving the PHYSIDEV database were approved by the local ethics committee (03/05-445) and those involving the VACCIN database were approved by the committee on protection of individuals (CPP Ouest 6-598). The characteristics of the respective populations, registering conditions and the different types of recorded signals are described in the following lines.

2.1 PHYSIDEV database

The main database for the study of apnea-bradycardia and sepsis in the SEPIA team is constituted by 76 premature infants having less than 33 weeks PMA, recorded from June 2003 to the present day through different projects: COREC and PHRC-INTEM. The newborns present more than one bradycardia per hour and/or the need for bag-and-mask resuscitation. Exclusion criteria are: ongoing inflammatory response, medication known to influence the autonomic nervous system except caffeine, intratracheal respiratory support, intracerebral lesion or malformation

Recordings were performed by a Powerlab system (ADInstruments, Oxfordshire, UK) at 400 Hz sampling rate during 1 to 3 hours and contained, at least, the following signals:

1. Electrocardiogram: Composed by a three-lead system placed on the infant's chest.
2. Respiration: Uncalibrated signals from piezoelectric sensors, which produce voltage in proportion to compressive or tensile mechanical strain.

Eventually, these signals were complemented by an EEG, EOG, breathing flow or manual annotations (arousals, agitation, etc.), but they have not been used because they were missing in many of the selected patients.

The recording protocol consisted on placing the infants in bassinets, positioned on their side, wrapped in a single blanket roll and loosely covered by another. The babies were also prevented from excessive or inappropriate sensory aspects of the environment, such as noise or light.

A sub-cohort from this database, that will be referred as SEPSIS, has been formed by 32 patients, 16 healthy and 16 with positive septicemic infection, paired by age-related, weight and sex criteria so that comparisons can be done in similar individuals.

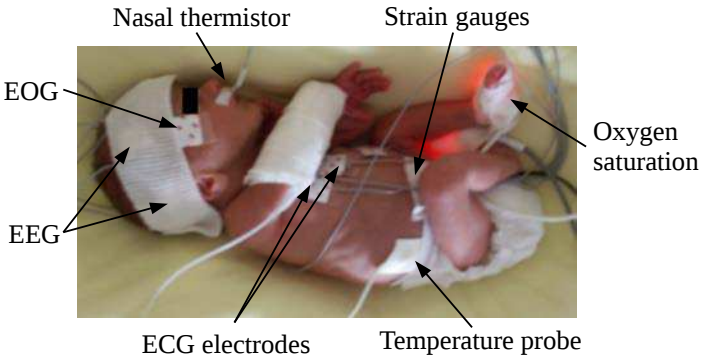

Figure II.1 – Image of the recording conditions at NICU for the PHYSIDEV database with arrows describing the electrodes. The infant is placed in a cocoon, trying to minimize external perturbations.

2.2 VACCIN database

This database is constituted by very preterm infants born between 26.9 and 29 weeks of gestation and having at least seven weeks of postnatal age at the moment of vaccination, (i.e a corrected age from 36 to 41 weeks) between September 2008 and February 2009. Magnetic resonance imaging was normal for all the infants and

EDIN score was not modified between the 24-hours period before and the 48-hours period after vaccine. Infants were excluded of the study if they had active infections, cardiac dysrhythmia, gastroesophageal-reflux requiring treatment, severe intra-cerebral lesions, or if they were currently treated by doxapram, caffeine, intratracheal ventilation or CPAP.

Signals. The polysomnographical recordings, ensured by a Galiléo-NT system, were composed by the following elements:

1. Electrocardiogram: Three electrodes on the infant's chest.
2. Respiration: Piezoelectric sensor Pneumotrace II placed in the abdomen.
3. O_2 saturation: Provided by a Masimo® monitor.
4. Electroencephalogram: Four electrodes placed on the scalp at frontal and parietal levels (Fp1, Fp2, T3 and T4, see Figure II.2) plus a reference.
5. Manual annotations: Sleep-wake states classified visually by nurses according to the standard criterion in neonatology (QS, AS, IS, QW, AW).

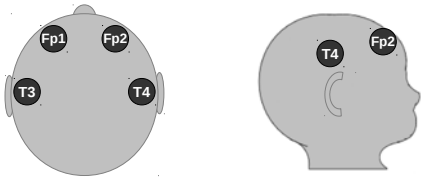

Figure II.2 – Electrode positions for the minimal EEG montage in the VACCIN database.

The protocol consisted on a polygraphic recording in the day before immunization (A period) for three hours, repeated between 12 and 18 hours after (B period, see Figure II.3) during two hours. In parallel, recording monitors were used continuously from 24 hours before to 48 hours after immunization. The environment was standardized, seating infants in their side-position, avoiding any sound or tactile stimuli.

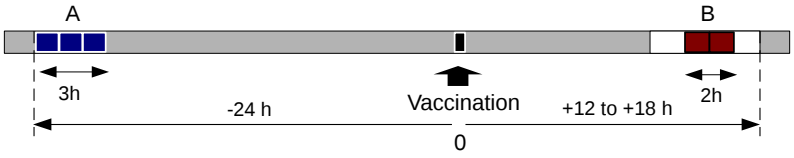

Figure II.3 – Illustration of the protocol of the VACCIN experiment.

3 Aims and hypotheses of the study

As introduced before, the works developed at LTSI explored extensively the cardiac signals and, to a lesser extent, the cardiorespiratory interactions, to prevent and treat apnea of prematurity in advanced monitoring

systems. But a deep analysis of breathing activity is still lacking. In the same manner, the analysis of the effects of immunization in preterm populations was limited to the cardiorespiratory system, without employing the EEG signals recorded during the experiment.

The present work essentially exploits EEG and respiratory signals, from their acquisition to their analysis at different levels, with the aim at determining if they can be used as complementary indicators for neonatal intensive care units. In this context, two main aspects will be stressed: the assessment of maturation and the early detection of sepsis. We underline the fact that ECG signals are not employed to develop our work, not only because its prior extensive analysis at LTSI, but also to assess the contribution of the respiratory or EEG signals alone to our objectives.

The interest of quantifying the maturity by means of physiological signals is not only to help in the decision of discharge, but also to detect possible maturation alterations and start early interventions that may positively influence the neonates' prognoses. On the other hand, determining as early as possible the onset of sepsis would indeed improve the effectiveness of treatments, reduce risks of death and save hospital costs.

The different signals available in this study can be analyzed from many approaches and interpretations, requiring certain assumptions about the nature of data. For instance, regarding physiological sources (cerebral or breathing activity), the time-dependency of data (variability signals or time series) or their length (long term versus short term) will be related to a specific methodology. Keeping that in mind, we outline here the research proposed in this work with respect to each database.

3.1 Analysis of cerebral activity from a minimal EEG montage

As some works already pointed out, analyzing EEG should take advantage of the relative power of the delta band and discontinuous patterns, which seems to be the most useful and repeatable measure for continuous long-term monitoring [47, 48]. However, they also noted that effective, automatic artifact rejection software needs to be developed before continuous quantitative EEG monitoring can be used in the neonatal intensive care environment.

A substantial part of this dissertation focuses on EEGs from a double pair montage (acquired in the VACCIN experiment), taking into account the reduced amount of information, their sensitivity to artifacts and the particularities of the preterm's brain activity. Once the signals are ready to be exploited, some processing tools will be applied to study several aspects of clinical interest.

Denoising EEG

If EEG is probably the most sensitive signal to be monitored, the fact that it is acquired in NICU environments is an additional aggravating factor. Recordings are usually unsupervised and a small number of electrodes are placed without a cap, so their sensitivity to noise are considerably higher than in standard systems. However, the smaller amount of information and the possibility of continuum acquisitions in reduced montages, could have interesting applications if EEG signals are properly filtered.

Since a reduced number of electrodes makes difficult the application of commonly-used strategies such as blind source separation, other single-channel strategies will be tackled. Besides to conventional filtering, signal decomposition could be convenient for immature EEG patterns –characterized by slow waves and discontinuities– and contribute to the noise removal effectively. Cardiac artifacts and low frequency noise, ubiquitously present in the NICU, must be eliminated with minimal losses in the true EEG content.

With this purpose, some tests will analyze extensively the association of classical and more recent decomposition techniques with canceling tools to find the best solution.

Analysis of preterm EEG

In spite of losing relevant spatial information, a bipolar EEG montage should provide essential features for the assessment of the background activity and the evolution of the transient patterns.

Inter-burst intervals will be then characterized in the EEG according to the sleep stages annotated in the VACCIN database. Maturational aspects, such as the age and weight, will be confronted to EEG features. If our minimal EEG montage was set up in the adequate positions and correctly filtered, it will be able to provide reliable information about maturation and be easily integrated in neonatal monitoring systems.

A second purpose of the analysis of EEG will be to investigate the eventual effects of vaccination on the cerebral activity. Knowing that some changes occurred on the cardiorespiratory system, we hypothesize that changes in the central nervous system might be perceived as well.

Considerations about sleep

Although the sleep-wake states were scored conforming to the standard criterion for neonatology, the analysis performed in this work considers the IS as part of AS (named subsequently Non-quiet sleep, NQS) for the following reasons:

1. To simplify the study of sleep in the dichotomy QS/NQS. In accordance to the above objectives, the quantity of QS relative to NQS will be analyzed to assess maturation, but also to find possible markers of comfort at NICU.

2. Some authors argue that it is preferable to distinguish between these two major sleep states in early infancy [49, 50], whereas IS is a minor state that was derived from combined behavioral and polygraphic criteria.

3. To gain consistency in the classifications. Since IS is largely subject to inter-operator variability, errors due to an inappropriate classification may be reduced. Although IS contains elements of QS, the latter can be identified easily by its complete absence of limb movements and a very regular respiration.

Applying this rule, we will study the eventual changes in sleep before and after vaccination, as well as the proportions of sleep stages with regard to the patient's age.

3.2 Study of the immature breathing

This work will exploit the respiratory signal in two forms: the temporal trace provided by piezoelectric sensors and the variability of the respiratory rhythm. The abdominal motion has been already used to study noninvasively the respiratory control and sleep-disordered breathing [51, 52]. It measures essentially the movement of the abdomen and can be employed to monitor central apnea –the absence of respiratory effort– but, on the contrary, obstructive apnea cannot be monitored.

As breathing signals are often artifacted, an automatic denoising solution and a cycle detection algorithm taking into consideration the particularities of preterm's immature patterns will be designed. Besides their utility to detect apnea (implemented in bedside monitors), clean signals could also be exploited in monitoring systems to assess the patients maturation and the risk of being infected.

Assessment of maturation

It is well known that immature responses in the preterms' breathing evolve during infancy. The decrease of the number of apneas and the progressive reduction of both periodic and erratic patterns are some of the evidences demonstrating an increase of maturation.

However, there are other measures that might serve as complementary tools to quantify the degree of maturation. The oscillators of the respiratory control are in continuous evolution, increasing their complexity as the respiratory neurons increase their connectivity.

We believe that in certain conditions, when external perturbations are minimized and respiration is smooth and regular, the breathing signal obtained by abdominal movements may suffice to study the evolution of the oscillators complexity. Taking advantage of the large choice of nonlinear methods, we will verify whether this hypothesis is confirmed.

Nonlinear methods will be also applied in variability signals, more robust to artifacts because they only contain information about the respiratory rhythm, to compare their potential to quantify maturation.

Detection of infection

As previously introduced, insufficient variability in cardiorespiratory rhythms could result from an inability to adapt to environmental challenges, provoked by sickness or aging. Consequently, respiratory variability signals may provide additional information to the current knowledge about the early detection of sepsis from RR signals.

A study of the fractal properties, quantified by the Hurst exponent (H), will be carried out on these signals to investigate the possible differences between healthy and infected patients. However, the computation of H is still unexplored in neonatal breathing, so a thorough performance analysis will be done previously to determine the best estimation algorithm.

4 Thesis outline

After the introduction to the clinical context and to the problem statement by the present part, the contribution of this work is developed in the next two parts, each focused on the exploited signal: electroencephalogram in Part II and breathing in Part III. A synthesis of the contents of each part can be seen in the flowcharts in Figure II.4.

Part II provides the background information about artifacts in neonatal EEG in order to define properly the filtering strategies. A newly-proposed denoising framework based on empirical mode decomposition and adaptive filter as well as a comparative analysis with alternative techniques are included here. Then, a new burst detection algorithm for preterm infants is designed and validated thanks to a gold standard elaborated from the decisions of three clinicians of the CHU Rennes. Finally, after adopting the optimal filtering and classification tools, several tests that could be useful in the clinical context, including the assessment of maturation and post-vaccine effects, are examined.

Part III deals with breathing signals in a similar manner: first, pre-processing tools are adequately designed to obtain signals as clean as possible. An automatic algorithm based on expert visual annotations served to reject artifact movements, obtaining thereafter variability signals by a cycle detector. These time series are

then exploited with nonlinear tools, i.e. chaos-based and long-range dependence estimation methods. The latter concerns specifically the Hurst exponent, which is evaluated through several estimation techniques by formalizing breathing models for preterm infants. Nonlinear tools are also applied for a clinical interest: to find useful indicators of maturity and septicemic infection.

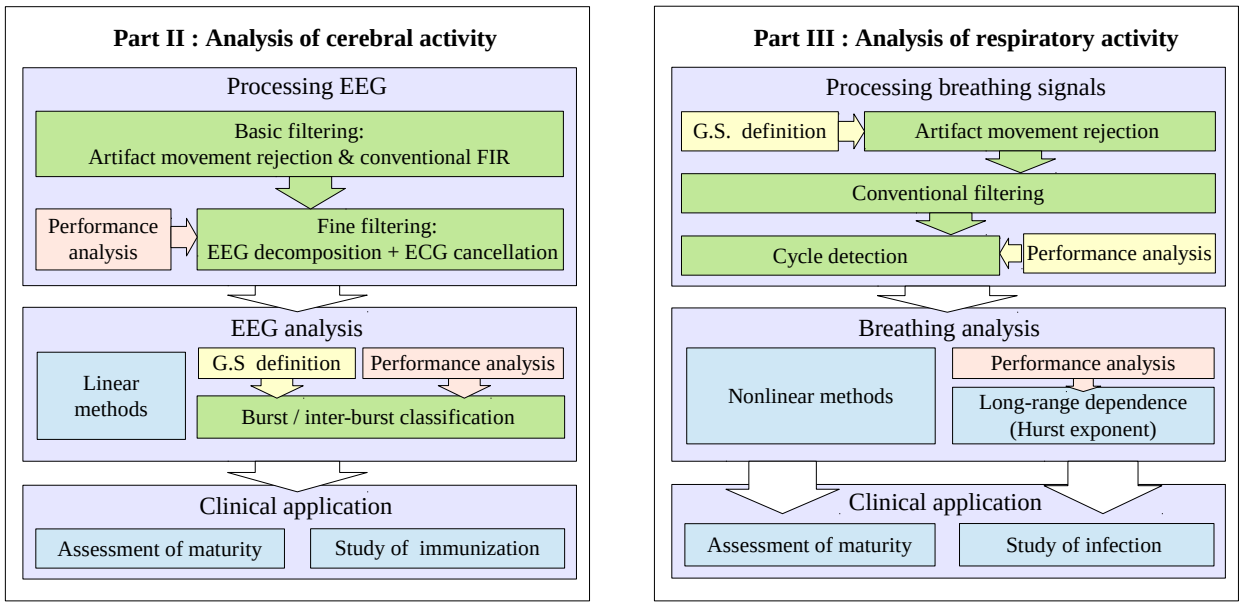

Figure II.4 – Thesis outline illustration. Methods are enclosed in different colors depending on their application. Green: processes that can be automated; yellow: manual procedures involving expert's intervention; orange: simulations and performance analyses; blue: exploration of diverse methodologies.

Bibliography

[1] C. C. Johnston and B. J. Stevens. *Experience in a neonatal intensive care unit affects pain response.* Pediatrics **98**(5), 925–930 November (1996).

[2] K. A. Allen. *Promoting and protecting infant sleep.* Advances in Neonatal Care **12**(5), 288–291 October (2012).

[3] C. F. Poets. *Apnea of prematurity: What can observational studies tell us about pathophysiology?* Sleep Medicine **11**(7), 701–707 August (2010).

[4] J. M. Abu-Shaweesh and R. J. Martin. *Neonatal apnea: What's new?* Pediatric Pulmonology, Pediatric Pulmonology **43**(10), 937–944 October (2008).

[5] L. I. Eisenfeld. Vibrotactile stimulator system for detecting and interrupting apnea in infants, September (1996). U.S. Classification: 600/534; 600/595 International Classification: A61B 500.

[6] R. Pichardo, J. S. Adam, E. Rosow, J. Bronzino, and L. Eisenfeld. *Vibrotactile stimulation system to treat apnea of prematurity.* Biomedical Instrumentation & Technology **37**(1), 34–40 January (2003).

[7] G. Pravisani, A. Beuchée, L. Mainardi, and G. Carrault. Short term prediction of severe bradycardia in premature newborns. In *Computers in Cardiology, 2003*, pages 725 – 728 September (2003).

[8] F. Portet, F. Gao, J. Hunter, and S. Sripada. *Evaluation of on-line bradycardia boundary detectors from neonatal clinical data.* Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society, EMBS'07 **2007**, 3288–3291 (2007).

[9] S. Y. Belal, A. J. Emmerson, and P. C. W. Beatty. *Automatic detection of apnoea of prematurity.* Physiological Measurement **32**(5), 523–542 May (2011).

[10] P. Pladys, J. Arsenault, P. Reix, J. Rouillard Lafond, F. Moreau-Bussière, and J.-P. Praud. *Influence of prematurity on postnatal maturation of heart rate and arterial pressure responses to hypoxia in lambs.* Neonatology **93**(3), 197–205 (2008).

[11] V. Le Rolle, A. I. Hernandez, G. Carrault, N. Samson, and J.-P. Praud. *A model of ventilation used to interpret newborn lamb respiratory signals.* Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society, EMBS'08 **2008**, 4945–4948 (2008).

[12] H. Patural, M. St-Hilaire, V. Pichot, A. Beuchée, N. Samson, C. Duvareille, and J.-P. Praud. *Postnatal autonomic activity in the preterm lamb.* Research in veterinary science **89**(2), 242–249 October (2010).

[13] A. Beuchée. *Interêt de l'analyse de la variabilité du rythme cardiaque en néonatologie.* Thèse de Doctorat, Université de Rennes 1 (2005).

[14] J. Cruz, A. Hernandez, S. Wong, G. Carrault, and A. Beuchée. Algorithm fusion for the early detection of apnea-bradycardia in preterm infants. In *Computers in Cardiology, 2006*, pages 473 –476 September (2006).

[15] M. Altuve, G. Carrault, J. Cruz, A. Beuchée, P. Pladys, and A. I. Hernandez. *Multivariate ECG analysis for apnoea bradycardia detection and characterisation in preterm infants.* International Journal of Biomedical Engineering and Technology **5**(2-3), 247–265 (2011).

[16] M. Altuve. *Détection multivariée des épisodes d'apnée-bradycardie chez le prématuré par modèles semi-Markovien cachés.* Thèse de Doctorat, Université de Rennes 1 (2011).

[17] A. Hernández. Dispositif intégré auto-adaptatif de surveillance et de stimulation sur base de signaux physiologiques, par exemple sur base de fréquence cardiaque, procédé et programme d'ordinateur correspondants.

[18] J. Cruz. *Sistema multiagentes de monitoreo aplicado a cuidados neonatales.* Thèse de Doctorat, Universidad Simon Bolivar, Venezuela (2007).

[19] A. Beuchée, P. Pladys, L. Senhadji, P. Bétrémieux, and F. Carré. *Beat-to-beat blood pressure variability and patent ductus arteriosus in ventilated, premature infants.* Pflügers Archiv: European journal of physiology **446**(2), 154–160 May (2003).

[20] A. Beuchée, A. Hernandez, P. Pladys, and G. Carrault. *Stimulateur kinesthésique automatisé asservi à la détection d'apnées-bradycardies chez le nouveau-né prématuré.* IRBM **28**(3-4), 124–130 September (2007).

[21] A. Beuchée, E. Nsegbe, M. St Hilaire, G. Carrault, B. Branger, P. Pladys, and J.-P. Praud. *Prolonged dynamic changes in autonomic heart rate modulation induced by acid laryngeal stimulation in non-sedated lambs.* Neonatology **91**(2), 83–91 (2007).

[22] M. Altuve, G. Carrault, J. Cruz, A. Beuchée, P. Pladys, and A. Hernandez. Analysis of the QRS complex for apnea-bradycardia characterization in preterm infants. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2009. EMBC 2009*, pages 946–949. IEEE September (2009).

[23] M. P. Griffin and J. R. Moorman. *Toward the early diagnosis of neonatal sepsis and sepsis-like illness using novel heart rate analysis.* Pediatrics **107**(1), 97–104 January (2001).

[24] H. Cao, D. E. Lake, M. P. Griffin, and J. R. Moorman. *Increased nonstationarity of neonatal heart rate before the clinical diagnosis of sepsis.* Annals of Biomedical Engineering **32**(2), 233–244 February (2004).

[25] M. P. Griffin, T. M. O'Shea, E. A. Bissonette, F. E. Harrell, D. E. Lake, and J. R. Moorman. *Abnormal heart rate characteristics preceding neonatal sepsis and sepsis-like illness.* Pediatric Research **53**(6), 920–926 (2003).

[26] J. R. Moorman, J. B. Delos, A. A. Flower, H. Cao, B. P. Kovatchev, J. S. Richman, and D. E. Lake. *Cardiovascular oscillations at the bedside: early diagnosis of neonatal sepsis using heart rate characteristics monitoring.* Physiological measurement **32**(11), 1821–1832 November (2011).

[27] A. Beuchée, G. Carrault, J. Y. Bansard, E. Boutaric, P. Bétrémieux, and P. Pladys. *Uncorrelated randomness of the heart rate is associated with sepsis in sick premature infants.* Neonatology **96**(2), 109–114 (2009).

[28] G. Carrault, A. Beuchée, P. Pladys, L. Senhadji, and A. Hernandez. Time-frequency relationships between heart rate and respiration: A diagnosis tool for late onset sepsis in sick premature infants. In *Computers in Cardiology, 2009*, pages 369–372. IEEE September (2009).

[29] R. Loforte, G. Carrault, L. Mainardi, and A. Beuchée. Heart rate and respiration relationships as a diagnostic tool for late onset sepsis in sick preterm infants. In *Computers in Cardiology, 2006*, pages 737–740. IEEE September (2006).

[30] T. Mialet-Marty. Effets de la vaccination du grand prématuré sur la régulation des rythmes cardio-respiratoires. DES de pédiatrie Université de Rennes 1, Faculté de Médecine April (2009).

[31] T. Mialet-Marty, A. Beuchée, W. Ben Jmaa, N. N'Guyen, X. Navarro, F. Porée, A. M. Nuyt, and P. Pladys. *Possible predictors of cardiorespiratory events after immunization in preterm neonates.* Neonatology **104**(2), 151–155 (2013).

[32] E. M. Mizrahi, R. Hrachovy, and A. Kellaway. *Atlas of neonatal electroencephalography.* Lippincott Williams & Wilkins, Philadelphia (2004).

[33] K. Holthausen, O. Breidbach, B. Scheidt, and J. Frenzel. *Brain dysmaturity index for automatic detection of high-risk infants.* Pediatric Neurology **22**(3), 187–191 March (2000).

[34] L. Jakaite, V. Schetinin, and C. Maple. *Bayesian assessment of newborn brain maturity from two-channel sleep electroencephalograms.* Computational and Mathematical Methods in Medicine **2012** March (2012).

[35] J. S. Hahn, H. Monyer, and B. R. Tharp. *Interburst interval measurements in the EEGs of premature infants with normal neurological outcome.* Electroencephalography and clinical neurophysiology **73**(5), 410–418 November (1989).

[36] M.-F. Vecchierini, M. André, and A. d'Allest. *Normal EEG of premature infants born between 24 and 30 weeks gestational age: Terminology, definitions and maturation aspects.* Neurophysiologie Clinique/Clinical Neurophysiology **37**(5), 311–323 (2007).

[37] J. Löfhede, M. Thordstein, N. Löfgren, A. Flisberg, M. Rosa-Zurera, I. Kjellmer, and K. Lindecrantz. *Automatic classification of background EEG activity in healthy and sick neonates.* Journal of Neural Engineering **7**(1), 016007 February (2010).

[38] K. Palmu, S. Wikström, E. Hippeläinen, G. Boylan, L. Hellström-Westas, and S. Vanhatalo. *Detection of 'EEG bursts' in the early preterm EEG: visual vs. automated detection.* Clinical Neurophysiology **121**(7), 1015–1022 July (2010).

[39] R. M. Harper, V. L. Schechtman, and K. A. Kluge. *Machine classification of infant sleep state using cardiorespiratory measures.* Electroencephalography and Clinical Neurophysiology **67**(4), 379–387 October (1987).

[40] A. T. Lewicke, E. S. Sazonov, and S. A. C. Schuckers. *Sleep-wake identification in infants: heart rate variability compared to actigraphy.* Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society, EMBS'04 **1**, 442–445 (2004).

[41] G. P. Nason, T. Sapatinas, and A. Sawczenko. *Wavelet Packet Modelling of Infant Sleep State Using Heart Rate Data.* (2001).

[42] E. Sazonov, N. Sazonova, S. Schuckers, and M. Neuman. *Activity-based sleep-wake identification in infants.* Physiological measurement **25**(5), 1291–1304 October (2004).

[43] A. Piryatinska, G. Terdik, W. A. Woyczynski, K. A. Loparo, M. S. Scher, and A. Zlotnik. *Automated detection of neonate EEG sleep stages.* Computer Methods and Programs in Biomedicine **95**(1), 31–46 July (2009).

[44] L. Fraiwan, K. Lweesy, N. Khasawneh, M. Fraiwan, H. Wenz, and H. Dickhaus. *Time frequency analysis for automated sleep stage identification in fullterm and preterm neonates.* Journal of Medical Systems **35**(4), 693–702 August (2011).

[45] K. Palmu, T. Kirjavainen, S. Stjerna, T. Salokivi, and S. Vanhatalo. *Sleep wake cycling in early preterm infants: Comparison of polysomnographic recordings with a novel EEG-based index.* Clinical Neurophysiology **124**(9), 1807–1814 September (2013).

[46] F. Porée, A. Kachenoura, H. Gauvrit, C. Morvan, G. Carrault, and L. Senhadji. *Blind source separation for ambulatory sleep recording.* IEEE Transactions on Information Technology in Biomedicine **10**(2), 293–301 April (2006).

[47] S. Victor, R. E. Appleton, M. Beirne, A. G. Marson, and A. M. Weindling. *Spectral analysis of electroencephalography in premature newborn infants: Normal ranges.* Pediatric Research **57**(3), 336–341 March (2005).

[48] L. Hellström-Westas and I. Rosén. *Continuous brain-function monitoring: State of the art in clinical practice.* Seminars in Fetal and Neonatal Medicine **11**(6), 503–511 December (2006).

[49] C. Dreyfus-Brisac. *Sleep ontogenesis in early human prematurity from 24 to 27 weeks of conceptional age.* Developmental Psychobiology **1**(3), 162–169 (1968).

[50] T. F. Anders and P. Weinstein. *Sleep and its disorders in infants and children: A review.* Pediatrics **50**(2), 312–324 August (1972). PMID: 4339857.

[51] E. Tabachnik, N. Muller, B. Toye, and H. Levison. *Measurement of ventilation in children using the respiratory inductive plethysmograph.* The Journal of pediatrics **99**(6), 895–899 December (1981).

[52] H. R. Gribbin. *Using body surface movements to study breathing.* Journal of medical engineering & technology **7**(5), 217–223 October (1983).

Part II

Analysis of cerebral activity

Chapter III

Denoising EEG by signal decomposition and artifact cancellation

The electroencephalography enables brain functions to be investigated by monitoring synaptic potentials from the scalp surface, constituting an important source of information to assess the neurological status and maturation of preterm infants. However, the EEG acquired in the NICU lacks the noise protection of standard recordings and hence, signals are often corrupted by different types of artifacts, posing serious problems for their correct interpretation. Therefore, a specific denoising strategy must be found before tackling the EEG analysis proposed further on.

Concerned with this purpose, this chapter first introduces the most common sources of noise in preterm EEG and then shows the application of classic filtering methods. Furthermore, given that some noise sources require more sophisticated filtering actions, other state-of-the-art approaches are described, tested and compared with a newly-proposed framework combining EEG decomposition and noise cancellation.

1 Noise in preterm EEG

During the recordings of EEG at NICU, if technicians are present, either the source of noise is identified and resolved, either any information to help the noise removal off-line is annotated to its subsequent analysis. Otherwise, if the signals are acquired without the presence of clinical staff, neurophysiologists –or software for automated analysis– confront a very challenging step before performing the interpretation of the EEG: the recognition of noise. Indeed, some noise sources can mimic true brain-generated waveforms, specially if recordings concern premature newborns less than 30 weeks PCA [1]. Due to their differing amplitude and frequency content in the background activity, identifying artifacts in preterm infants varies from that in adult's EEG. In the latter, fast EEG waves can be masked by EMG activity whereas in the preterm EEG fast rhythms are not frequently observed. On the other hand, respiratory or cardiac artifacts are more difficult to eliminate in premature newborns, because of their closer frequencies to physiological activities.

The most important sources of noise can be classified in three categories: environment, patient-instrumentation interface and interferences from other electrophysiological signals. Whatever the artifacts are, they are superimposed on background activities.

1.1 Environmental noise

In neonatal intensive care units a large number of instruments to monitor and care the preterm infants coexist together. Electrical interferences induced by the external currents of, for instance, ultrasound instrumentation, artificial ventilators, pumps for artificial intravascular infusions, appears in all the channels of the EEG with the same level if the electrode impedances are equal. Capacitive-induced potentials in the electrode wires can also be induced by other devices as radio transmitters or mobile phones. Interferences of the power line at 50 Hz (Fig III.1) are

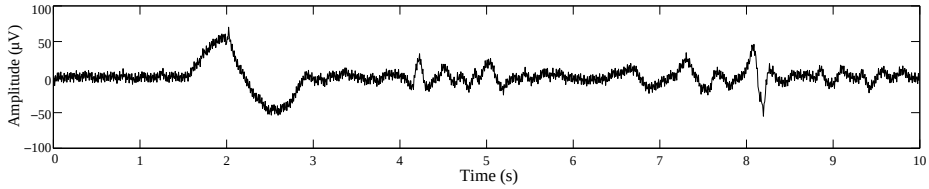

Figure III.1 – Example of power line noise in a patient of 37 weeks PMA. The trace appears thicker due to the superimposition of the 50 Hz signal.

1.2 Patient-instrumentation interface noise

The interface between the infant's skin and the instrument is the surface of the electrode, so an unstable contact can result in a sudden change of the impedance causing extraneous potentials that may appear as spike-like waves. They can be present in repetitive or single patterns in the concerned channel with an abrupt initial phase followed by a gradual return to the baseline. This kind of artifacts are also known as electrode pop-up (Fig III.2-a).

The variation of the skin conductances due, for instance, to changes on the sweat level, plays also an important role in this category of artifacts. If the individual's body temperature or stress level increases, sweat may appear in older preterm infants and consequently the impedance between the inside and outside the skin increases. Then, the electrical potential also changes, creating a very large artifact (often several hundred microvolts).

Head movements produced by respiration can push the electrodes against the bed so conductances may be modified. The provoked artifacts are usually slow and large movements similar to the breathing trace superimposed to the EEG (Fig III.2-b). Sharp artifacts can be also observed when other random limb or body movements occur. Finally, manipulations, comforting and feeding of the infant are another important source of alteration of the electrode interface.

1.3 Noncerebral electrophysiological interferences

The electrical activity generated by the heart is a major source of artifacts and can appear in channels from one or both sides depending on body and head position. Since it appears continuously during relative long periods of time and its waveform is similar to the QRS complex and synchronized with the ECG channel, it can be easily identified (Figure III.3-a).

Many movements of the infant producing electrical potentials can be reflected in the EEG sporadically. These can include breathing, glossokinetic (tongue movements), laryngeal or pharyngeal activity (swallows) and

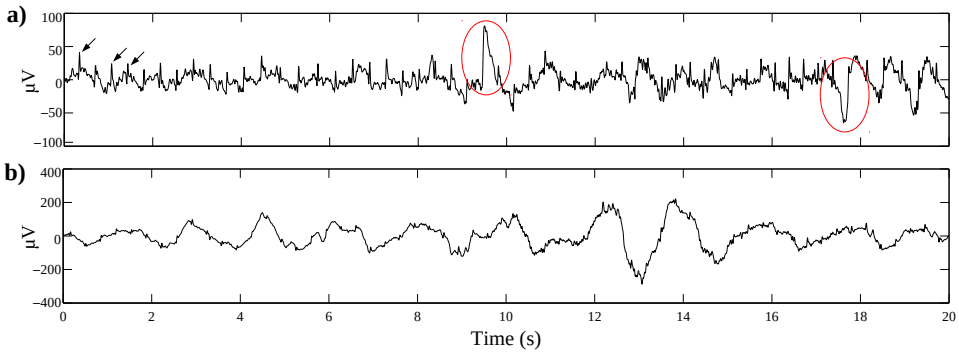

Figure III.2 – Examples of noise in the patient-instrumentation interface. a) EEG from a 37 weeks PMA infants having electrode pop-up (red circles) and ECG artifacts (arrows). b) EEG contaminated with breathing artifacts (low-frequency waves) and also a weak ECG noise component in a 38 weeks PMA infant.

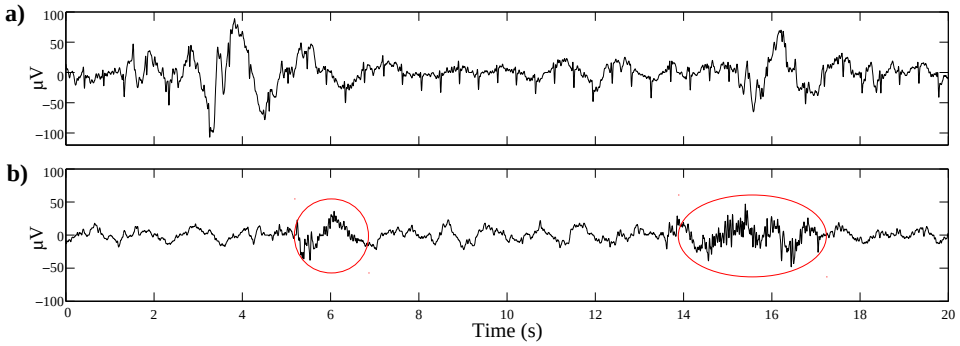

Figure III.3 – Examples of non-cephalic noise in the patient-instrumentation interface. a) ECG noise (periodic sharp spikes). b) EMG noise (high frequency waves in circles).

electromyographic noise produced by jaw tremor and other facial muscles (Figure III.3-b). The ocular activity, mainly caused by eye deviations, also belong to this category. The knowledge of the sleep stages is crucial to their detection since most of the mentioned movements only are manifested in active sleep and wakefulness stages. And reciprocally, the presence of certain artifacts can help to the sleep state identification.

1.4 Artifact removal strategies

Although numerous publications are devoted to the EEG artifact problem in adults, children and infants, very few works attempt to characterize specifically this issue in premature newborns. Even if some assumptions and approaches are valid for both full-term and preterm infants, the still immature brain activity of the latter should be kept in mind.

Both the nature of artifacts and the particularities of preterm EEG recorded in the NICU require a specific denoising strategy. Whether the noise is continuous, sporadic or associated to some frequencies in the spectrum, some factors are to be considered. Sporadic artifacts (related mostly to sudden perturbations in the electrodes, saturation of the amplification devices) usually implies the rejection of the corrupted EEG segment. For long-term noises, a spectral analysis can be a first procedure to identify them in frequency. While spectrally independent artifacts can be easily removed with conventional filtering techniques, those overlapping the EEG bands (see Figure III.4) require most sophisticated filtering tools.

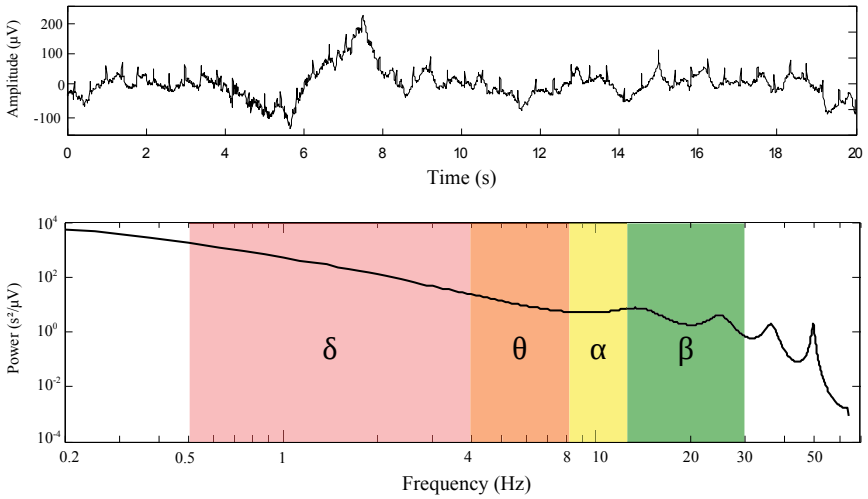

Figure III.4 – Example of a raw EEG epoch and its power spectrum. Low frequency and 50 Hz power line noise can be noticed in the spectrum, but ECG artifacts (between 4 and 40 Hz approximately) are masked by the θ, α and β bands.

For this purpose, we propose the methodology (depicted in Figure III.5), composed essentially by two stages:

1. A basic preprocessing stage, detecting and rejecting the artifacts related to gross body movements and then band-pass filtering the EEG with conventional techniques to eliminate the baseline drift and high frequency noise.

2. A more accurate denoising stage, the combination of EEG decomposition and artifact cancellation, subsequently called CEDAC. ECG artifacts –canceled if necessary using the recorded reference in the NICU– and low frequency noise (LFN) can be removed by several techniques.

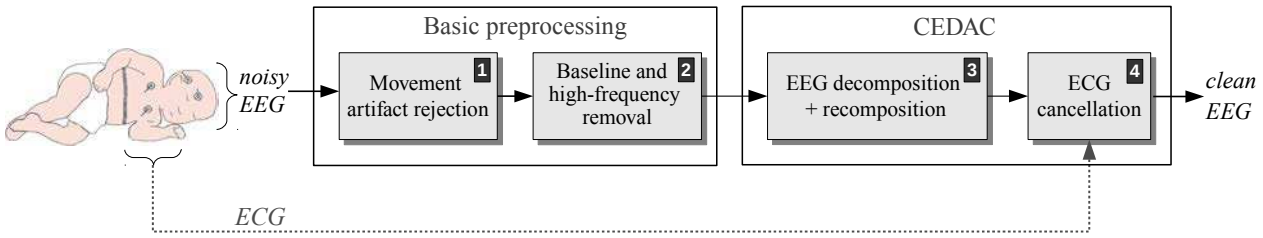

Figure III.5 – Basic block diagram of the different filtering steps applied to the EEG from the NICU. The noisy EEG in our database is composed by four channels, simultaneously recorded with an ECG channel and used as a reference if cardiac artifacts need to be canceled. The functions of the numbered blocks are described in Table III.1.

The automated noise removal proposed above can be summarized in Table III.1, where noise sources susceptible to be removed are described and located in the block diagram of Figure III.5.

Noise source	Description	Solution	Block #
Body & limbs movements	Large voltage transients	Cut damaged excerpt	1
Power line & NICU instrumentation	50 Hz (or more) sinus-like interferences	Lowpass filter	2
Baseline drift	Very low frequency or continuous component	Highpass filter	2
Muscle activity	Rapid oscillations (40-70 Hz)	Lowpass filter	2
Eye motion	Slow transients (< 0.5 Hz) in active sleep	EEG decomposition + rejection of concerned components	3
Breathing	Slow wave synchronized with breathing	EEG decomposition + rejection of concerned components	3
Heartbeats	Periodic QRS complexes	Cancellation with ECG reference	4

Table III.1 – Summary of the different noise sources and the proposed solutions to remove them.

2 Basic preprocessing

This section covers two solutions generally applied in EEG preprocessing. Their application is a priori straightforward because the noise concerned is out of the frequencies of interest and involves high-amplitude excursions easily identifiable.

2.1 Rejection of movement artifacts

Outlier values in the EEG, dues for instance to sporadic, abrupt artifacts of external origin or to saturation in the amplifiers, can be detected by voltage thresholds. Once identified, the affected region is usually cut, being its limits rejoined. This procedure is also known as threshold artifact rejection.

The detector implemented in this work is based on a simple criterion proposed by Durka et al. [2] for adult polysomnographic recordings with good concordances with experts visual detections. The threshold, Th_r, is calculated as follows:

$$Th_r = m_{0.5} + k_r \sigma_{25}, \tag{III.1}$$

where $m_{0.5}$ is the median of the EEG, k_r is a factor between 2 and 10 and σ_{25} the standard deviation of the entire record neglecting the 25 % of the tails from each side of the distribution.

In our signals, the excluded extremes were set to the 10% of the tails instead of the 25% for adults to avoid inappropriate rejections. This more flexible criterion is justified because preterm EEG distributions are more skewed due to the predominance of patterns like *tracé discontinu* and *tracé alternant*.

Artifacted epochs in the raw EEG were identified by simply finding the absolute values within a 1-second sliding window exceeding the threshold Th_r, as illustrated in Figure III.6.

2.2 Baseline and high frequency noise rejection

When artifacts and EEG do not share the same spectral content, a conventional band-pass filtering can be performed. Any activity under δ band (< 0.5 Hz), such as the baseline drift or slow movements, can be eliminated using a high-pass filter (HPF). Likewise, a low-pass filter (LPF) can be used to remove high frequency noise over the β band (30 Hz). Since linear phase is needed to avoid signal distortion, finite impulsional response (FIR)

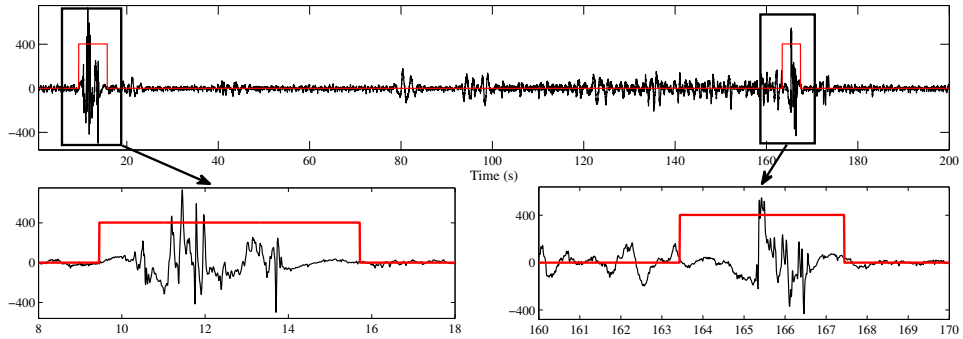

Figure III.6 – Example of artifact rejection in the raw EEG. Red steps constitute selected periods for rejection.

digital filters are preferred to infinite impulsional response filters (IIR) [3]. FIR filters have a greater stability and less complex design, but their major drawback is that they can be computationally expensive to implement and may require a long transient to meet some design parameters (see scheme in Figure III.7).

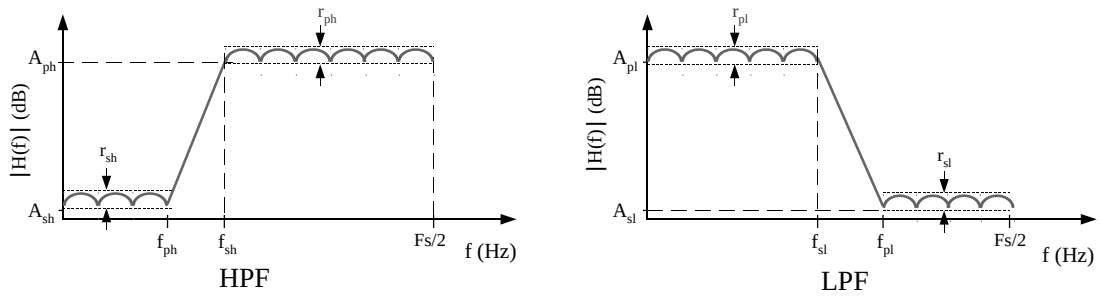

Figure III.7 – Filter specifications diagrams for a high-pass filter (left) and a low-pass filter (right). A_s, f_s and r_s are the attenuation, the limit frequency and ripple of the stopband, A_p, f_p and r_p are the counterpart parameters for the passband.

If the minimum stopband attenuation (A_s), the passband ripple (r_p) and the transition band ($f_p - f_s$) are constrained, a steeper slope will be attained if the filter order increases. To illustrate this effect, the order for some specification parameters in the case of a HPF is given in Table III.2. The Kaiser window design method [4] has been employed due to the small ripple obtained in the passband.

A_{sh} (dB)	r_p (dB)	f_{ph} (Hz)	f_{sh} (Hz)	Order	Delay (s)
50	0.25	0.45	0.50	7500	29.29
40	0.25	0.45	0.50	5718	22.30
40	0.25	0.40	0.50	2860	11.17
40	0.25	0.30	0.50	1432	5.59
40	0.25	0.20	0.50	956	3.73
30	0.25	0.20	0.50	656	2.55
25	1	0.15	0.55	382	1.49

Table III.2 – Order and delay values for several high-pass specifications using Kaiser window design method at a sampling frequency of 128 Hz.

Removing selectively the low frequency noise without attenuating delta waves is not straightforward. Very narrow transition widths are necessary at 0.5 Hz, implying higher filter orders and consequently, longer delays in

the transient response. In real-time applications, the efficiency of low frequency noise removal using a FIR filter can be compromised if short delays are needed. Conscious of its limitations, we opted to use the fastest solution, the 382-order FIR filter detailed in the above table. However, as it will be explained later on, the low frequency noise can be rejected by unmixing EEG in subcomponents, improving the performance of FIR filters. The tests proposed in Section 5 compare the use of the HPF against the noise rejection by means of EEG decomposition.

On the other hand, the interest of low-pass filtering the EEG is to eliminate any noncephalic high frequency, mainly EMG activity and 50 Hz power line interferences. In the particular case of preterm infants, any frequency content over 30 Hz can be removed safely because fast EEG waves are rare, appearing from 38 weeks PMA. Consequently, constraints are less demanding in the design of the LPF. Fixing $f_{sl} = 30$ Hz and $f_{pl} = 35$ Hz and ensuring a good attenuation in the stopband with $A_{sl} = 60$ dB, an order 78 FIR filter attains the requirements.

3 Combination of EEG decomposition and artifact cancellation

The present section proposes the combination of some techniques for EEG decomposition and ECG/LFN cancellation with the aim at verifying to what extent the final quality of the EEG is enhanced compared to the use of cancellation techniques and high-pass filters alone. First, we review the recent literature regarding advanced denoising solutions, stressing the application of signal decomposition methods, combined or not with other techniques. Then, we describe the different components and steps to denoise the EEG in the CEDAC framework.

3.1 State of the art

As introduced before, the EEG is a biosignal with multiple sources of artifacts. In recent years, a number of methods based on removing undesired EEG components, under the hypothesis that those components are related to noise sources, have become very popular. In these methods the signal is decomposed in several subcomponents according to some criteria and assumptions, permitting its reconstruction without information loss when their corresponding inverse transformations are applied.

In this concern, two different approaches must be considered. The first decomposition strategy relies on the separation of the different sources that produce by mixture the observed signal (Blind Source Separation, BBS) and the second one performs the separation regarding the oscillatory or spectral properties of the signal. The last approach, where discrete wavelet transform (DWT) and empirical mode decomposition (EMD) are included, can be very effective if the artifacts are closely associated to certain frequencies.

From the nineties, many authors have developed frameworks in which short time Fourier or DWT were performed prior to adaptive filters (See for instance [5] and [6]). This methodology, also referred as sub-band adaptive filtering, decomposes the input signal into multiple parallel channels by a filter bank to increase convergence ratios and ameliorate the signal tracking with less complex sub-filters. In the particular case of polysomnography, many works perform a DWT decomposition on corrupted EEG to increase the effectiveness of a subsequent noise canceling or other specific techniques. To cite a few examples, Kumar et al. [7] used adaptive filters to cancel effectively EOG on a decomposed EEG, Browne et al. [8] utilized DWT to compensate the inter-trial variability in EEG event-related potential datasets and Olkkonen et al. [9] canceled noise in EEG subcomponents using subspace methods.

Concerning empirical mode decomposition, its combination with other techniques to denoise or characterize

underlying processes in biomedical signals is a very up-to-date trend. McKeown et al. [10] were the first to combine independent component analysis (ICA) and EMD to investigate electromyography and EEG coherence. Lindsen and Bhattacharya [11] proposed a new method for removing eye blink artifacts from the EEG by first finding the independent components (ICs), then performing EMD over ICs containing blink-related information and finally recovering selectively the EMD unmixed components containing EEG activity. EMD has been also used to extract ECG artifacts, as for instance, the work of Taelman et al. [12] in which ensemble EMD (EEMD) and ICA were applied in EMG data showing the efficacy of combining both methods.

The present work explores the potential of the family of EMD methods combined with adaptive filters as a solution to denoise the EEG, comparing their performance with alternative decomposition and cancellation techniques.

Note that in the above lines we omitted a literature review of the blind source separation applied to EEG signals. Indeed, the use of this approach in our particular signals did not yield satisfactory results and thus, is not included in this chapter. For additional information, in Appendix A we describe briefly this method and the performed tests, arguing the reasons why BSS was discarded.

3.2 Description of the proposed method

The principle of our proposal to denoise effectively EEG signals –contaminated by ECG and other noise sources– is supported by the advantages of decomposing the EEG in a natural manner using efficient algorithms. More generally, we hypothesize that EEG decomposition:

- Improves the performance of the ECG canceling techniques if they deal directly with those subcomponents containing the cardiac artifacts.
- Reduces distortion due to ECG cancellation if the non-contaminated EEG subcomponents remain aside.
- Removes efficiently high and low frequency noise (non-overlapping with true EEG) by rejecting the corresponding EEG subcomponents. This may avoid the use of band-pass filters, whose design implies the risk of attenuating the signal of interest.

To find out the best solution according to the above presumptions, the following decomposition methods will be employed:

1. The decomposition by multiresolution analysis performed by DWT (Section 4.1).
2. The EMD classical algorithm, described in Section 4.1.
3. The complete ensemble EMD with adaptive noise (CEEMDAN), a recent improved version of EMD (Section 4.1).

They will be combined with two noise cancellation techniques: the well-known adaptive filtering (AF) and the ensemble average subtraction (EAS). The denoising procedure carried out by the CEDAC solution is illustrated in Figure III.8 and described by the following steps:

1. Perform EEG decomposition and obtain M independent subcomponents (ISs). They correspond to intrinsic modal functions (IMFs) if EMD/CEEMDAN is applied (see example in Figure III.9), or sub-bands if DWT is performed.

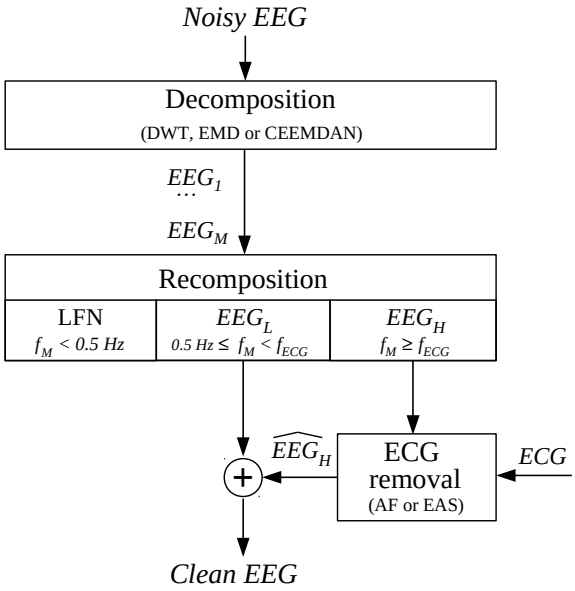

Figure III.8 – Block diagram to compare different options in the CEDAC framework.

2. Estimate the main frequency (f_M) of each IS from its power spectral density.
3. Find f_{ECG}, the frequency splitting the ECG spectrum area in 5% on the left and 95% on the right, so that the range [0 - f_{ECG}] is considered to have a non significant portion of the ECG energy.
4. Construct EEG_H, a component formed by the addition of the ISs having $f_M > f_{ECG}$.
5. Likewise, obtain EEG_L, the component with the addition of the remaining ISs, carrying δ-waves (0.5 - 4 Hz).
6. Reject LFN, the noisy ISs ($f_M < 0.5$ Hz) obtained from the decomposition.
7. Remove ECG in EEG_H using the recorded ECG as a reference in the ECG removal block and obtain the cleaned component, $\widehat{EEG}_H$. To this end, ensemble average subtraction or adaptive filtering can be performed.
8. Reconstruct the clean EEG by the addition of EEG_L and $\widehat{EEG}_H$.

4 Methods used in CEDAC

4.1 EEG decomposition

Wavelet decomposition

The wavelet theory provides a unified framework for a number of techniques which had been developed independently for various signal processing and mathematic applications, as for example, multiresolution signal analysis, subband coding and wavelet series expansions [13]. The Wavelet Transform (WT) is of interest for

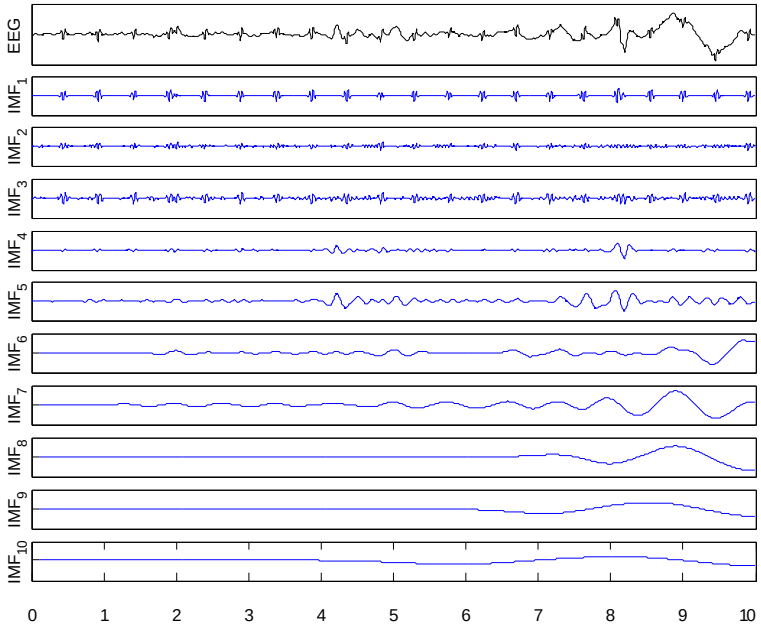

Figure III.9 – Ten-second excerpt of noisy EEG (upper plot) decomposed with EMD. The sum of IMF_1 to IMF_4 ($f_M > f_{ECG}$) constitutes EEG_H.

the analysis of non-stationary signals, because it provides an alternative to the classical Short-Time Fourier Transform (STFT) or Gabor transform. The latter was introduced to get information about frequency changes in a signal $x(n)$ at a time location τ and is computed by applying the Fourier Transform (FT) over short segments of the signal using a shifting window, $g^*(t-\tau)$:

$$STFT(\tau, f) = \int x(n) g^*(t-\tau) e^{-2j\pi f t} dt \tag{III.2}$$

The STFT maps the signal into a time-frequency plane (τ, f) at a constant resolution because the splitting window is fixed over the entire plane.

In contrast, the continuous wavelet transform (CWT) uses short windows at high frequencies and long windows at low frequencies. The CWT is thus performed by introducing the wavelet window, $\psi_{a,b}(n)$:

$$CWT_x(a, b) = \int x(n) \psi_{a,b}(n) dt \tag{III.3}$$

where

$$\psi_{a,b}(n) = \frac{1}{\sqrt{a}} \psi \left(\frac{t-b}{a} \right) \tag{III.4}$$

with a and b positive real numbers. a denotes the scaling factor (the frequency counterpart of STFT) and b refers to the time translation or location (τ).

The wavelet window is the scaled and the translated version of the basic or mother wavelet, $\psi(n)$, a function with a specific frequency response to convolve with $x(n)$. A wide choice of functions have been designed for wavelet transform analysis, as for instance, the Harr, Morlet and Daubechies [14].

The **discrete wavelet transform** is achieved using digital filtering techniques, passing the signal through

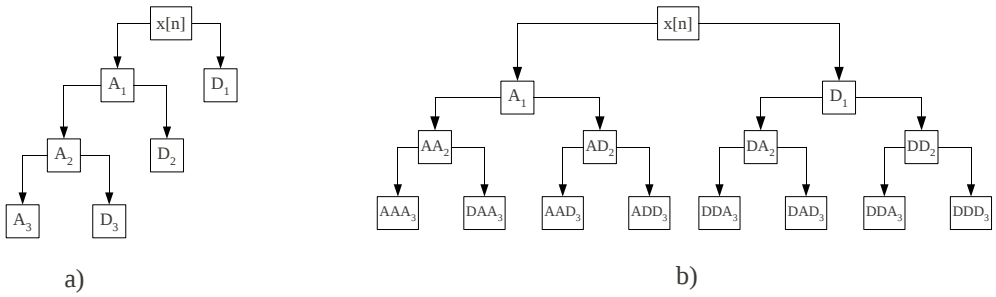

Figure III.10 – Obtaining the approximation and detailed components by a) standard wavelet transform and b) wavelet packet transform.

different cuttoff frequencies at different scales. In order to compute DWT, discrete values of the scale and location parameters are taken: given j, k integer numbers, $a = a_0^j$ and $b = ka_0^j b_0$ ($b_0 > 0$):

$$DWT_{j,k}(x) = a_0^{-\frac{j}{2}} \int x(n)\ \psi(a_0^{-j}t - kb_0)\ dt \tag{III.5}$$

When $a_0 = 2$ and $b_0 = 1$ the discrete wavelet transform can be used for multiresolution analysis since it is generated from a pair of quadrature mirror filters whose FIR representation is:

$$g(k) = (-1)^k h(1-k), \tag{III.6}$$

where $g(k)$ is a high-pass filter related to the so-called scaling function (ϕ) and $h(k)$ is a low-pass filter related to the mother wavelet (ψ):

$$\phi(x) = \sum_k h(k)\sqrt{2}\phi(2x-k) \tag{III.7}$$

$$\psi(x) = \sum_k g(k)\sqrt{2}\phi(2x-k). \tag{III.8}$$

The outputs of the quadrature filters, H_l and G_l are, respectively, low-pass and high-pass filtered versions of the discrete signal $x(n)$:

$$H_l = \sum_k h(k-2l)x(k) \tag{III.9}$$

$$G_l = \sum_k g(k-2l)x(k). \tag{III.10}$$

Applying these operators, two subbands downsampling by two the Nyquist frequency (Fn) of the signal are obtained. The output corresponding to H_l, known as the approximation component, contains the lower frequencies ($0 - Fn/2$) and the output given by G_l, called detailed decomposition, carries the upper frequencies ($Fn/2 - Fn$). The standard wavelet transform applies the low-pass result recursively to the next wavelet step, until the component cannot be decimated by two any more. Since the subbands obtained with this scheme are limited to the wavelet bases that increases by a power of two with each step (see Figure III.10-a), a more complete representation can be achieved applying the transform to both the low-pass and the high-pass. This combination of bases is called wavelet packet transform (see Figure III.10-b).

The reconstruction of the signal is achieved by following the inverse procedure without loss of information. This technique is identified as the inverse discrete wavelet transform (IDWT) and differs from the DWT in the order that it implies upsampling and filtering of the approximation and detailed subband components.

Empirical Mode Decomposition

The EMD is a relatively new time series decomposition method with an excellent time-frequency resolution [15]. In virtue of EMD, the signal under study can be represented as a sum of intrinsic components with different frequency content, the IMFs. This technique is very convenient to decompose neurosignals, allowing to effectively discriminate, for instance, the gamma activity of field potentials under stimulus [16]. It has been also utilized in cardiology [17, 18] and in some respiratory applications [19].

The EMD is a nonlinear, non-stationary time series analysis method developed by Huang et al. [15]. While the Fourier analysis combine fixed sine and cosine waves as the basis functions, in the EMD they are derived directly from the time series itself.

The EMD performs a sifting process resulting in a set of components, the intrinsic modal functions with a time scale intrinsic to the time series. The instantaneous frequency of this components is defined by the time lapse between successive extrema and can be easily obtained by the Hilbert transform (HT). The HT allows to represent the amplitude and the instantaneous frequency in a three-dimensional plot. Combination of EMD and Hilbert transform (Hilbert-Huang transform, HHT) provides a powerful analysis technique, more precise in time-frequency than Fourier-based analysis.

During the sifting process the time series must satisfy two conditions in order to be an IMF:

1. The number of extrema and the number of zero crossings are either equal or differ at most by one in the whole dataset.

2. The mean of its upper and lower envelopes must be close to zero according to some criterion.

Given these requirements, the sifting of the time series $X(n)$ can be performed. It consists in an iterative process:

1. Identify $x(n)$ local minima and local maxima (initially, $x(n)$ is set to $X(n)$) Then interpolate these points via cubic splines, forming an upper envelope, $e_u(n)$ and lower envelope, $e_l(n)$. The two envelopes must contain all the points, wrapping $x(n)$.

2. The mean of the two envelopes $(m_1(n) = (e_l(n) + e_u(n))/2)$ is calculated and then subtracted to $x(n)$, obtaining $h_1(n)$. If h_1 does not accomplish the conditions to be an IMF, the iteration continues. At the second iteration h_1 will be set as the signal $x(n)$, $h_1 - m_{11} = h_{11}$. This can be repeated k times, $(h_{1(k-1)} - m_{1k} = h_{1k})$ until the first IMF $(c_1 = h_{1k})$ is found.

3. The IMF obtained is subtracted to the initial data $X(n)$ to obtain the residue. The residue is then treated as a new time series and begins the sifting process from the first step.

The procedure continues until the last residue has no turning points. This residual component have the lowest frequency content (continuous voltage) of the time series.

In the present work, we performed the EMD on EEG signals employing an efficient algorithm implementation in Matlab given by Rilling et al. [20]. The Rilling algorithm introduces two thresholds in the second criterion to

consider a IMF, θ_1 and θ_2. They aim at guaranteeing globally small fluctuations in the mean while taking into account locally large excursions. Defining the mode amplitude, $a(n) = (e_u(n) - e_l(n))/2$ and the evaluation function, $\sigma(n) =\mid m(n)/a(n) \mid$ and the tolerance $\alpha \in [0, 1]$, the author proposes the following stopping criteria in the sifting process:

- $\sigma(n) < \theta_1$ during the portion (1- α) of the time of $X(n)$.
- $\sigma(n) < \theta_2$ during the rest of the portion of the signal.
- The difference between the number of zero crossings and extrema must be less or equal to 1.

It is suggested the use of the default values of $\theta_1 = 0.05$, $\theta_2 = 10\theta_1$, $\alpha = 0.05$ to prevent the sifting process from over-iteration, and hence to avoid an over-decomposition of the signal.

CEEMDAN

In contrast to kernel based approaches (like the wavelet decomposition) where the components correspond to a predetermined filters or basis functions, EMD obtains the modes adaptively, without doing any assumptions about the nature of oscillations. This allows a more optimal, compact representation, with modes containing frequencies naturally present in the signal, but decomposition of noisy signals may result in the corruption of modes, i.e. the presence of one oscillation in different IMFs (see Figure III.11).

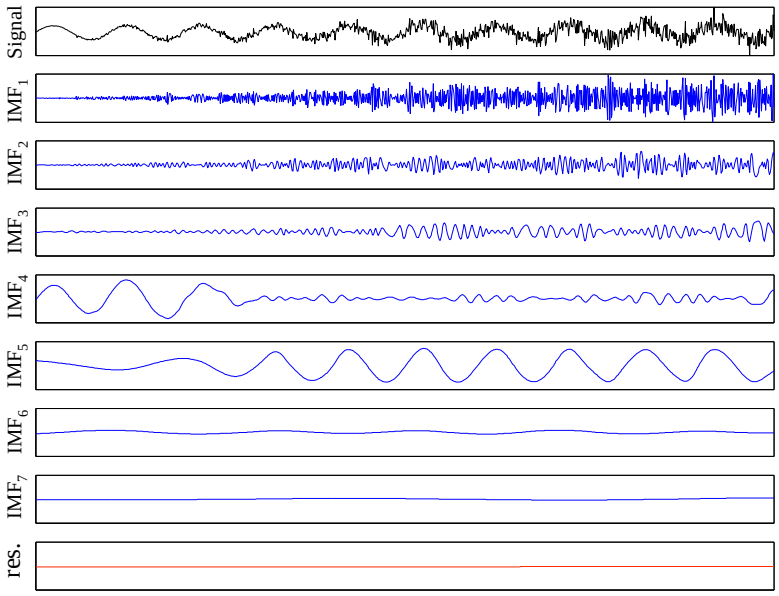

Figure III.11 – Illustration of mode mixing. A contaminated sinus with increasing white noise has been decomposed by EMD. As it can be observed, IMFs 3 to 5 contain mixed frequency tones [21].

Ensemble empirical mode decomposition (EEMD) was introduced by Wu and Huang [22] to reduce the corruption of unmixed modes. In EEMD, the decomposition is repeated I times by adding white noise to

the signal:

$$s^i(n) = s(n) + w_i(n), \quad i = 1, .., I, \tag{III.11}$$

where $w_i(n)$ are different realizations of white noise with variance ε. Then, the final oscillatory modes are obtained by averaging:

$$\overline{IMF}_k = \frac{1}{I} \sum_{i=1}^{I} IMF_k^i(n). \tag{III.12}$$

This procedure improves the quality of the separation (see Figure III.12), but at the expenses of a high computational cost and without the warranty of a perfect reconstruction of the signal because a different number of IMFs may be obtained for each iteration.

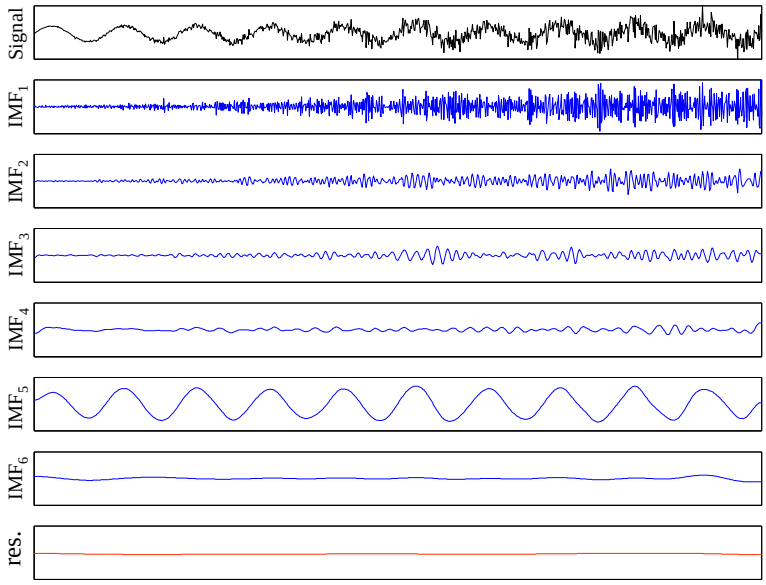

Figure III.12 – The same signal shown in Figure III.11 decomposed by EEMD with 100 realizations of noise. Unlike Figure III.11, the IMFs here contain pure tones without mode mixing [21].

The complete EEMD with adaptive noise (CEEMDAN) was proposed recently [23] to ameliorate the spectral separation of modes and reduce computational time (see Figure III.13). In EEMD each $s^i(n)$ is decomposed independently and I residuals are obtained. The CEEMDAN computes the first decomposition as EEMD, so that a unique first residue is obtained:

$$r_1(n) = s(n) - \overline{IMF}_1(n). \tag{III.13}$$

Then, EMD is performed over a set of $r_1(n)$ plus different noise realizations, to obtain $\overline{IMF_2}(n)$ by averaging. The next residue is $r_2(n) = r_1(n) - \overline{IMF_2'}(n)$ and so on, until the stopping criterion is achieved.

Let $E_j(\cdot)$ be an operator producing by EMD the j-th IMF and $w_i(n)$ a white noise realization with standard deviation ε, CEEDMAN can be described by the following algorithm:

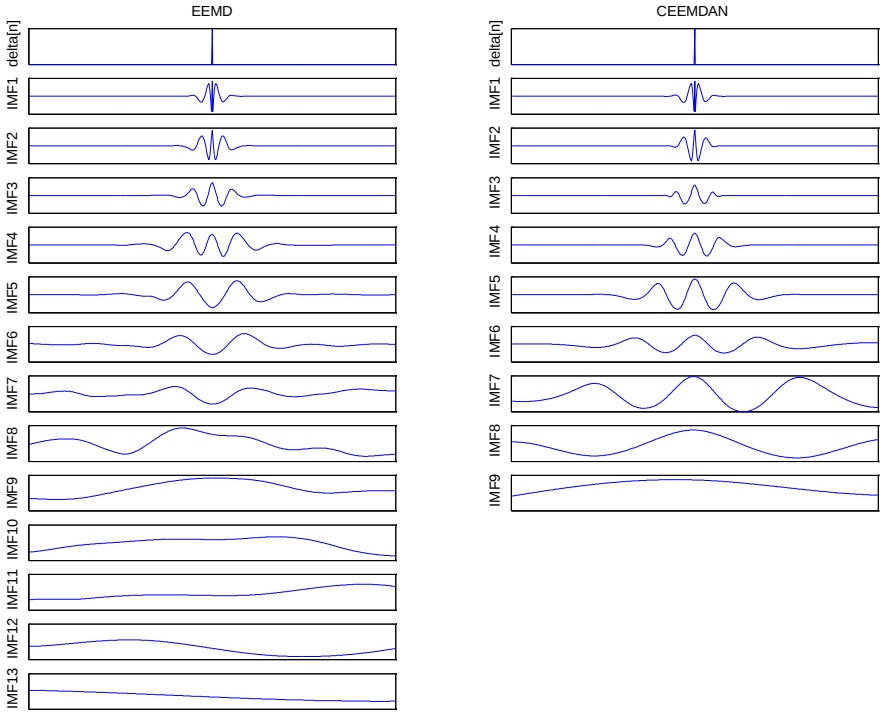

Figure III.13 – Comparison of the decomposition of a delta signal by EEMD (left) and CEEMDAN (right) [21]. As it can be observed, the superior performance of CEEDMAN results in a smaller number of IMFs.

1. Perform I realizations of EMD on $x(n) + \varepsilon_0 w^i(n)$ $(i = 1, 2, ..I)$ to obtain their first modes and compute:

$$\overline{IMF}_1 = \frac{1}{I}\sum_{i=1}^{I} IMF_1^i. \tag{III.14}$$

2. Calculate the first residue at the first stage $(k = 1)$, obtaining $r_1(n)$ as in Equation III.13

3. Perform the decomposition on I realizations of $r_1(n) + \varepsilon_2 E_2(w^i(n))$ to obtain their first mode and define second mode:

$$\overline{IMF}_2 = \frac{1}{I}\sum_{i=1}^{I} E_1(r_1(n) + \varepsilon_2 E_2(w^i(n))) \tag{III.15}$$

4. Calculate the k-th residue for $k = 2, 3, ..K$:

$$r_k(n) = r_{(k-1)}(n) - \overline{IMF}_k(n). \tag{III.16}$$

5. Perform the decomposition on $r_k(n) + \varepsilon_k E_k(w^i(n))$, $i = 1, ..I$ until their first EMD mode and define the $(k+1)$-th mode as:

$$\overline{IMF}_{(k+1)} = \frac{1}{I}\sum_{i=1}^{I} E_1(r_k(n) + \varepsilon_{k+1} E_{k+1}(w^i(n))). \tag{III.17}$$

6. Repeat the procedure from step 4 for next CEEMDAN mode k.

Steps 4 to 6 are performed until the residue obtained cannot be decomposed any more (i.e. it has not turning points). The final residue satisfies:

$$R(n) = x(n) - \sum_{k=1}^{K} \overline{IMF}_k \tag{III.18}$$

and the signal decomposed signal can be expressed as:

$$x(n) = \sum_{k=1}^{K} \overline{IMF}_k + R(n), \tag{III.19}$$

with K the number of CEEMDAN modes.

4.2 ECG cancellation

Cardiac artifacts remain a major concern among EEG researchers because they can mimic physiological brain activity and pathological patterns (epileptic spikes, for example) [24, 25]. They can be originated by the propagation of the QRS electrical activity to the scalp (so that they are similar to the ECG waveform), but also if the electrode is placed over a pulsating vessel, meaning that the artifacts waveform can be considerably different from the ECG. If cardiac artifacts appear in the EEG periodically, they can be eliminated by synchronization analysis using an ECG reference, either the original ECG simultaneously recorded or a signal generated from the detection of QRS complexes in EEG. In this concern, two methods taking advantage of the synchronization between the reference signal and the artifacted EEG are described in the following lines: the ensemble average subtraction and the adaptive filtering.

Ensemble average subtraction

EAS was proposed by Nakamura and Shibasaki [26] to eliminate ECG interferences by subtracting the averaged waveform of QRS peaks contaminating the EEG. This technique can use either a single ECG channel as a reference [27] or QRS detection techniques when the ECG is unavailable [28]. It is computed in five steps:

1. Generation of a reference to the R peaks: the trigger signals can be obtained from the recorded ECG channel or by detecting QRS peaks in the EEG.

2. Exact positioning of interferences: since the exact locations of the QRS instants into the EEG does not necessarily match those from the recorded ECG, the trigger pulses are repositioned by looking for the maxima in a surrounding window of 15% the inter-beat time.

3. Averaging: EEG raw signals are segmented into windows starting 200 ms prior to the one triggering point and ending 200 ms before the next one. The average of these windows provide an estimate of the ECG artifact.

4. Synchronized repetition: the average artifact is chained and synchronized with the trigger pulse of the ECG.

5. Subtraction: the previously estimated artifact signal is subtracted to the raw EEG.

Adaptive filtering

Adaptive filtering has been extensively used as a strategy to remove ECG artifacts in EEG signals. The typical configuration for this procedure is shown in Figure III.14. The input signal, $x(n)$, is used as a reference by a filter H(Z) to generate an output, $y(n)$, capable to cancel the artifacts of the desired signal, $d(n)$ [29]. In our case, $x(n)$ is the recorded ECG and $d(n)$ the contaminated EEG.

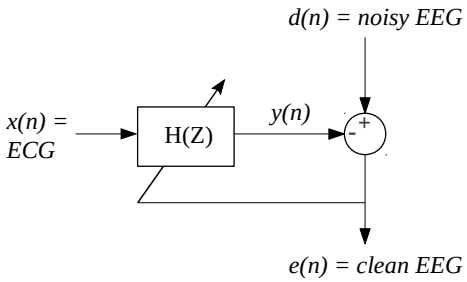

Figure III.14 – Basic diagram of an adaptive filter. The reference signal, $x(n)$, is supplied by an independent ECG channel recorded in the NICU. The desired signal, $d(n)$, is the noisy EEG.

The adaptive process involves minimization of a cost function, which is used to determine the filter coefficients. Usually, linear FIR filters are employed for their stability and simplicity. They can be represented as:

$$y(n) = \sum_{k=0}^{L} w_k x(n-k), \tag{III.20}$$

where $x(n)$ and $y(n)$ are the before-mentioned input and output of the filter, L is the order and w_k are the filter coefficients. The error signal, $e(n)$, defined as difference between $y(n)$ and $d(n)$, is fed back into H(Z) to modify iteratively its coefficients and hence, to adapt the output to the artifacts contained in the desired signal. In stationary environments, the filter should converge to the Wiener solution and in non-stationary conditions, the coefficients will change with time according to the signal variation, converging to an optimum filter. The transfer function is then adjusted according to an optimizing algorithm.

The **least mean square** (LMS) algorithm is a very used, simple and robust criterion based on the minimization of the squared error $e^2(n)$:

$$e^2(n) = d^2(n) - 2d(n)\sum_{k=0}^{L} w_k x(n-k) + \left[\sum_{k=0}^{L} w_k x(n-k)\right]^2, \tag{III.21}$$

and whose squared error expectation for N samples is given by:

$$\xi = E[e^2(n)] = \sum_{k=0}^{N} e^2(k). \tag{III.22}$$

A number of proposed algorithms can minimize the squared error by different optimization techniques, as for instance, methods based on gradient descent [30].

The **recursive least squares** (RLS) is another well-known algorithm for the same purpose. In contrast to the LMS algorithm, the RLS algorithm uses information from all past input samples (and not only from the current tap-input samples) to estimate the transfer function. To decrease the influence of input samples from the far past, a weighting factor for the influence of each sample is used. This weighting factor λ is introduced in the cost function, so that the resulting expectation for $e^2(n)$ is given by:

$$\xi(n) = \sum_{k=0}^{N} \lambda^{n-k} |e^2(k)| + \delta\lambda^n ||\mathbf{w}(n)||^2, \tag{III.23}$$

where λ $(0 < \lambda \leq 1)$ is known as the forgetting factor, which gives more importance to recent samples as the closer is its value to one. δ is the regularization factor, positive and real. Then, RLS updates the filter coefficients by means of:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + e(n)\mathbf{k}(n), \tag{III.24}$$

$$\mathbf{k}(n) = \frac{\mathbf{P}(n)\mathbf{u}(n)}{\lambda + \mathbf{u}^T(n)\mathbf{P}(n)\mathbf{u}(n)}, \tag{III.25}$$

with $\mathbf{u}(n) = [e(n), e(n-1), ...e(n-L+1)]^T$. $\mathbf{P}(n)$ is the inverse correlation matrix of the input signal. Initially, $\mathbf{w}(0) = 0$ and $\mathbf{P}(0) = \delta^{-1}\mathbf{I}$, with $\mathbf{I}$ the identity matrix. Next, $\mathbf{P}(n)$ is recursively updated by performing:

$$\mathbf{P}(n) = \lambda^{-1}\mathbf{P}(n-1) - \lambda^{-1}\mathbf{k}(n)\mathbf{u}^T(n)\mathbf{P}(n-1). \tag{III.26}$$

The choice of RLS allows the filter to track the complexity of EEG and effectively remove ECG artifacts. It has a faster convergence and a better adaptation to rapidly varying environments than the classic LMS algorithm [31].

5 Validation strategy

This section explains how the efficacy of the different denoising combinations are validated. The generation of artificial data and the organization of the tests to compare the results are first given and then, some preliminary tests necessary to set up the CEDAC framework are presented.

5.1 Generation of artificially contaminated data from real signals

Contaminated signals were generated from a selection of sixty 20-second excerpts in sleep from the VACCIN database. Several channels and patients at different post-menstrual ages were considered to ensure the inclusion of different neonatal EEG patterns. After verifying the nonexistence of ECG artifacts by visual inspection, original EEGs (sampled at 512 Hz) were band-pass filtered (0.4 - 35 Hz) to eliminate baseline, EMG and high frequency noise and then subsampled to 128 Hz.

In real-life recordings, cardiac artifacts and ECG reference exhibit often delays and different waveforms, thus we generated the interferences by applying linear transformations to the recorded ECG lead (a 21th-order FIR filter with random coefficients). Then, its amplitude was properly modified to obtain signal to noise ratios

(SNRs) from -5 to 10 dB in steps of 5 according to the following equation:

$$SNR = 10 \log \frac{P_{EEG}}{P_{noise}} \tag{III.27}$$

where P_{EEG} and P_{noise} are the power of a original EEG excerpt and the power of the added artifacts, respectively.

Since EEG signals acquired in monitoring environments can be contaminated with low frequency noise, we constructed two groups to simulate realistic conditions:

1. Group 1: Thirty clean EEG excerpts with added ECG noise from -5 to 10 dB.

2. Group 2: Thirty clean EEG excerpts with added ECG at 5dB SNR plus LFN at variable SNR. The latter consisted in one cycle of a sinusoidal wave, very close to the limit of the delta band, fixed randomly between 0.3 and 0.5 Hz. The power of the wave was set to obtain SNRs (relative to the original signal) ranging from -10 to 10 dB in steps of 5 dB.

Some examples of artificially contaminated EEG signals are shown in Figure III.15.

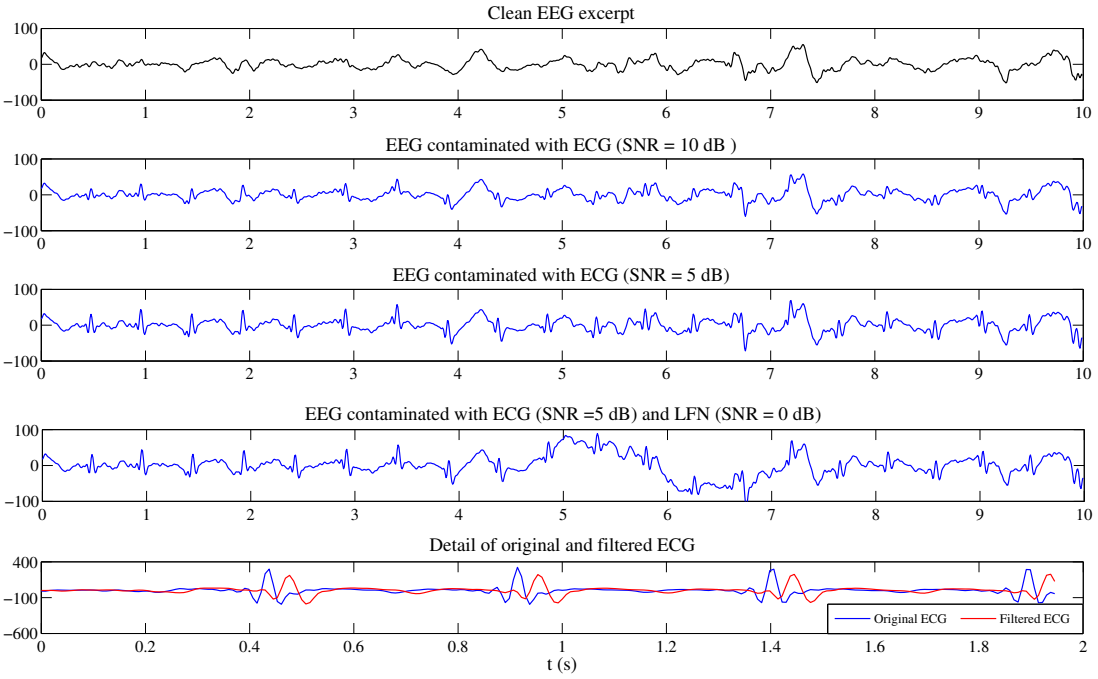

Figure III.15 – Example of the generation of artificially contaminated EEGs by the filtered ECG and LFN. Using the first signal (clean EEG), two examples of Group 1 (second and third signals) have been formed. The fourth signal (a Group 2 example) is formed by the addition of the third signal and LFN at 0 dB. The lower plot shows the originally recorded ECG and the reference $x(n)$ used to test the AF, obtained by applying the 21th-order FIR filter.

5.2 Organization of tests

Once the groups of contaminated signals were created, they were denoised by the different combinations of CEDAC. In parallel, they were cleaned by EAS and AF methods alone (without unmixing previously the noisy EEG) to compare the eventual improvement of introducing decomposition.

Concerning tests on Group 2, a high-pass filtering is added when EAS and AF are evaluated alone. In effect, the prior elimination of LFN in this case is necessary to avoid performance losses in the ECG cancellation. On the other hand, in CEDAC this is not necessary because low frequency noise is systematically rejected in its design. A summary of all tested combinations is outlined in Figure III.16.

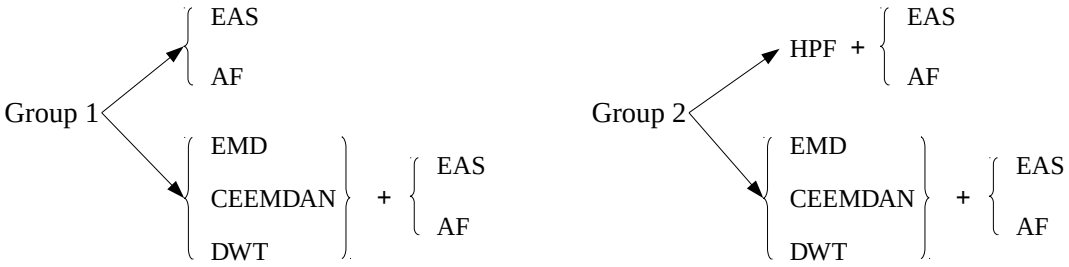

Figure III.16 – Diagram of the different tested combinations to clean artificially contaminated EEGs.

5.3 Preliminary tests and parameter tuning

In order to set up the CEDAC framework, several tests using the artificially contaminated signals from Group 1 and 2 were carried out to find the optimal parameters of the different blocks. These tests concern the adaptive filter and EMD-based decompositions.

Order of adaptive filters

An optimal order of the AF, or the length of coefficients for the RLS algorithm, is necessary to avoid an incomplete ECG removal (when the order chosen is too small), or a slow tracking of the EEG (when the filter order is too large). We therefore designed a quick test to verify to what extent the low frequencies (mainly delta waves) present in EEG modify the performance of the adaptive filter and which would be the optimal order in our data. The two set of signals employed in this test were:

a) The sixty original, clean excerpts selected previously contaminated by ECG noise (generated as in Section 5.1) between 0 and 15 dB.

b) The same original signals, processed by a 656-order FIR filter (described in Table III.2) to remove frequencies below 5 Hz. Then, they were contaminated by ECG at the same levels.

The two sets were afterwards denoised by the AF, that used the recorded ECG as reference signal, trying several filter orders ($L = 8, 9..40$) and forgetting factors close to 1 yet depending on L ($\lambda = 1 - \frac{1}{10L}$) as suggested by Eleftheriou and Falconer [31].

Once sets a) and b) were cleaned, the quality of the ECG removal was measured by the mean square error (MSE):

$$MSE(\widehat{EEG}) = \frac{1}{N}\sum_{i=1}^{N}(EEG - \widehat{EEG})^2, \tag{III.28}$$

where $\widehat{EEG}$ is the signal whose ECG noise has been canceled by the AF and EEG is the original, clean excerpt.

The average of all MSEs, represented in Figure III.17, shows that denoising b) yield lower orders (minimum in $L = 16$) than denoising set a), with a minimal order of 19. More importantly, the quality of ECG cancellation was considerably better in b), as evidenced by the lower MSEs.

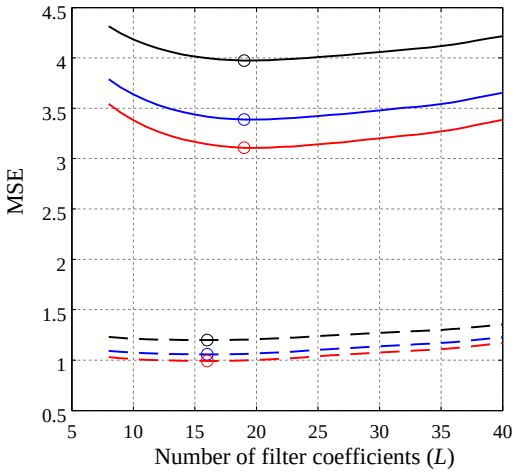

Figure III.17 – Average MSEs of denoising original EEG excerpts plus ECG noise at several SNRs as a function of the number of coefficients, L. They correspond to the set a), represented in solid lines. In dashed lines, the results of denoising set b), original EEGs high-pass filtered and then contaminated with ECG. The lower MSEs obtained in last set demonstrate that avoiding low frequency components in the EEG improves the performance of the AF. The different noise levels are: 0dB (black lines), 5dB (blue lines) and 15 dB (red lines). The lowest filter orders are marked in circles.

In conclusion, this test showed that the efficacy of adaptive filters is expected to better perform if low frequency components are removed. Hence, the AF utilized in the CEDAC framework would perform optimally processing only the highest EEG components (EEG_H) with $L = 16$ coefficients and $\lambda = 0.9994$.

Mode mixing of IMFs

As introduced before, decomposing by EMD has the risk of producing mode mixing, therefore, it is necessary to know whether the existence of this phenomenon affects the quality of the overall denoising process.

A fast way to verify mode mixing is checking the stationarity of the first mode [32]. This test was applied on signals from Group 1 and 2, resulting in the detection of mode mixing in the vast majority of decompositions. However, after analyzing the IMFs, most of the cases concerned the two highest oscillatory modes, with low frequencies constituting EEG_L unrelated to mode mixing. Consequently, if this phenomenon mixes frequencies from only one side of the separation determined by f_{ECG}, it should not be an inconvenient because the addition of the highest IMFs yields EEG_H uncorrupted and the canceling block performs the ECG suppression normally.

To this extent, the major concern should be if any of the IMFs contains energy on both sides of f_{ECG}, independently if mode mixing occurs. Two different cases need to be analyzed:

1. Modes having $f_M \geq f_{ECG}$ (added to form EEG_H) contain some low-frequency content below f_{ECG}. Since they are selected to be processed by the AF, these low-frequencies could lead to underperform the ECG cancellation.

2. Modes having $0.5 \leq f_M < f_{ECG}$ contain some high-frequency content above f_{ECG}, so that the ECG noise

could be partially installed in EEG_L. Since this component is not concerned by the AF block, the artifacts would be incompletely removed.

We subsequently refer to this effect as critical mixing, I or II if it describes the first or second case.

A simple method was then designed to quantify critical mixing by estimating the energy of frequencies exceeding the limit established by f_{ECG}. Taking advantage of the purity of the oscillatory tones in IMFs, the instantaneous frequencies of each cycle can be estimated easily by calculating the distance between maxima or minima and the associated instantaneous energy by integrating the signal between these limits (see Figure III.18).

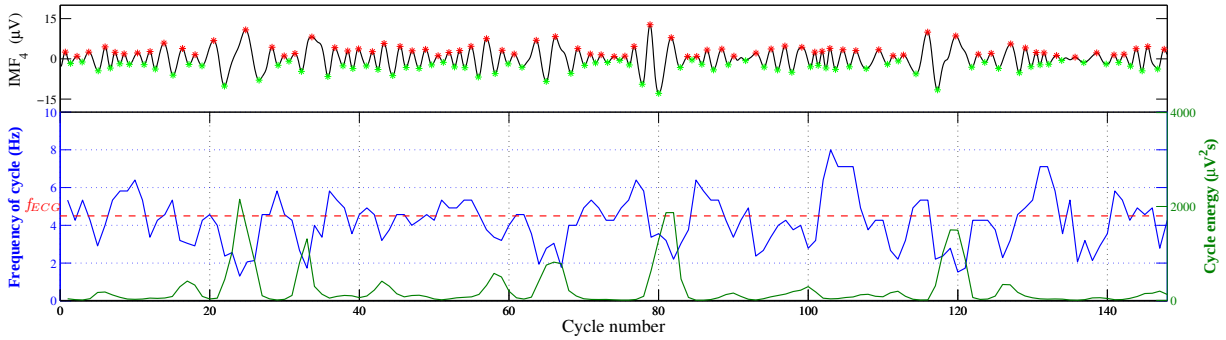

Figure III.18 – Estimation of the instantaneous frequency (blue lines) and energy (green lines) in IMFs. In this example, the fourth IMF (upper signal) of a contaminated EEG from Group 2 has been selected. The detected maxima and minima (red and green asterisks) allowed to estimate the period and energy by half-cycles, represented in the lower plot. The red dashed line is the frequency from which the ECG artifacts are present, f_{ECG}. As IMF_4 has $f_M < f_{ECG}$, the energies on frequencies exceeding f_{ECG} contribute to critical mixing II.

Once all instantaneous frequencies of IMFs are estimated (Figure III.19), the sum of individual cycle energies exceeding f_{ECG} is computed for critical mixes I and II, then it is divided by the sum of all energies in every IMF to obtain the relative portion of the EEG having critical mixing. This allows to evaluate the quality of the separation by applying different parameters to the EMD, in particular the thresholds of the sifting process (θ_1 and θ_2) and the tolerance α. In Figure III.20, the variation of these two parameters are represented as a function of critical mixing and the number of IMFs. The percentage of energy crossing the f_{ECG} boundaries varies irregularly versus θ_1, but the number of produced IMFs decreases inversely. In any case, the values minimizing both critical mixtures considering EEGs with ECG and LFN noise are about $\theta_1 = 0.45$ and $\alpha = 0.02$.

A second test evaluated critical mixing employing the CEEMDAN decomposition and varying the parameters I (number of realizations) and ϵ (standard deviation of added white noise). The starting point of tests was based on Wu et al. [22] recommendations for EEMD, who suggest to perform a few hundreds of EMD realizations and to set $\epsilon = 0.2$ to obtain good separations, but taking into account that if high frequencies predominate, ϵ may be smaller and when the data is dominated by low-frequency signals, it may be increased. The tests (see Figure III.21) show that critical mixing is reduced in relation to EMD decompositions and the number of IMFs, close to 11 in average, is less variable against the variation of parameters. The optimal values according the plots are $I = 300$ and $\varepsilon = 0.45$.

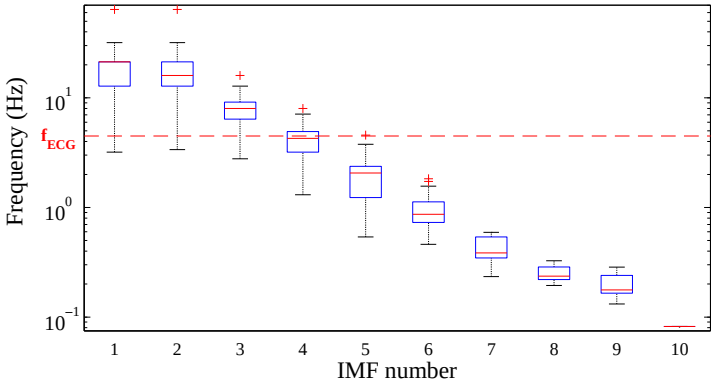

Figure III.19 – Instantaneous frequencies of all IMFs. The two first IMFs clearly exhibit mode mixing, but since it occurs beyond the limit of f_{ECG}, it has no consequences on the denoising process. The energies carried by frequencies in the lower whiskers of IMF_1, IMF_2 and IMF_3 crossing the red line constitute critical mixing I. Likewise, the energies in the frequencies crossing f_{ECG} in IMF_4 and IMF_5 determine critical mixing II.

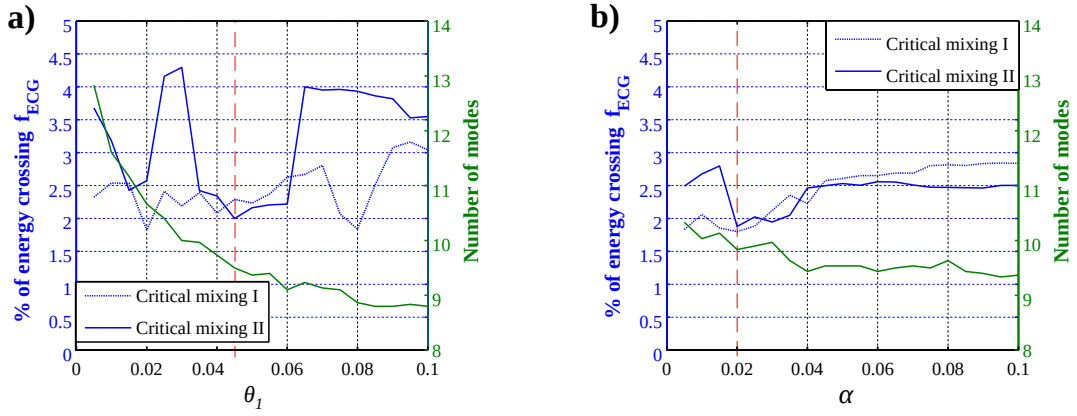

Figure III.20 – Critical mixing as a function of several parameters of the EMD. a) Threshold of the sifting process (setting $\theta_2 = 10\theta_1$ as suggested by [20]) and b) tolerance. The plots show the mean values over the 30 noisy EEG signals from Group 2. Red dashed vertical lines indicate the values minimizing critical mixing.

6 Results

In this section, the results from cleaning the noisy EEGs employing different combinations (outlined in Figure III.16) are compared, first using artificially contaminated data and then using real data.

To perform adequate decompositions by EMD and CEEMDAN, we used the optimal parameters found in previous section. The DWT decomposition was performed using 6th order Daubechies wavelets, a choice with satisfactory results in EEG signals [33], obtaining 8 levels of successive details (sub-bands with the frequency intervals at 32-64, 16-32, 8-16, 4-8, 2-4, 1-2, 0.5-1 and <0.5 Hz). Concerning the ECG cancellation, the RLS algorithm was implemented with an order-16 filter and $\lambda = 0.9994$ according to the previous tests. EAS was computed in 20-second windows, finding the QRS instants in ECG with a R-peak detector based on Pan & Tompkins algorithm [34] adapted to minimize the false detections in preterm infants [35].

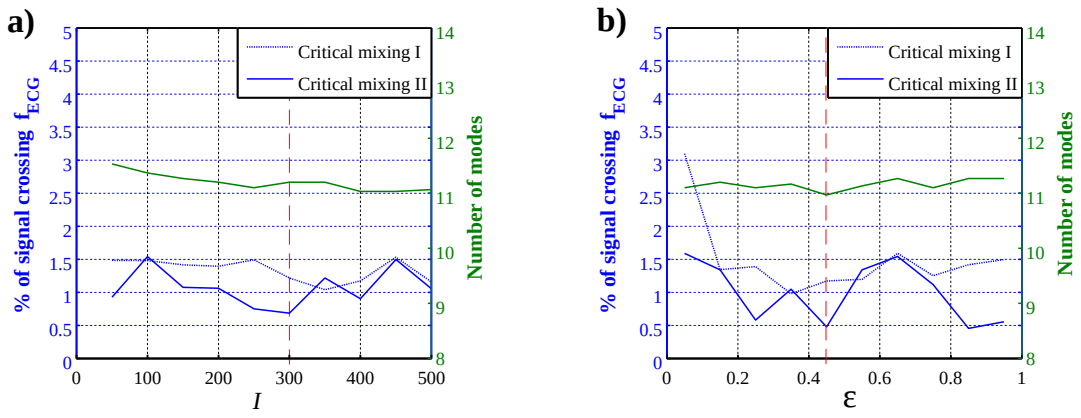

Figure III.21 – Critical mixing as a function of several parameters of the CEEMDAN. a) Number of realizations (I) and b) noise level (ε) to assist the decomposition process. The plots show the mean values over the 30 noisy EEG signals from Group 2. Red dashed vertical lines indicate the values minimizing critical mixing.

6.1 Tests on artificial signals

Since the original, clean data is known, the quality of each CEDAC combination has been measured in terms of root mean squared error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[EEG(i) - \widehat{EEG}(i)]^2} \tag{III.29}$$

where EEG is the original signal without artifacts, $\widehat{EEG}$ is the result of removing noise on the contaminated excerpt and N is the number of samples of the signals.

ECG removal (Group 1)

In a first instance, the EEG signals contaminated only by ECG (Group 1) were denoised. The first noticeable effect of introducing EEG decomposition and separation of artifact-related components is the enhancement of both AF and EAS (see Figure III.22).

The improvements of RMSE, taking as a reference the results of denoising without EEG decomposition become more significant as the level of the ECG noise decreases. High SNR levels are then more favorable to perform EEG signal decomposition, probably due to smaller distortion introduced to the final reconstructed signal when only the noisy part of the EEG is processed by the cancelers. On the contrary, in low SNRs the denoising levels of all methods converge because useful information is masked by noise and the decomposition becomes less effective.

Comparing the repercussion of the decomposition techniques in the denoised signals, CEEMDAN has the best results closely followed by EMD and finally DWT. In any of the three cases, no significant differences are observed between the use of adaptive filtering and ensemble average subtraction.

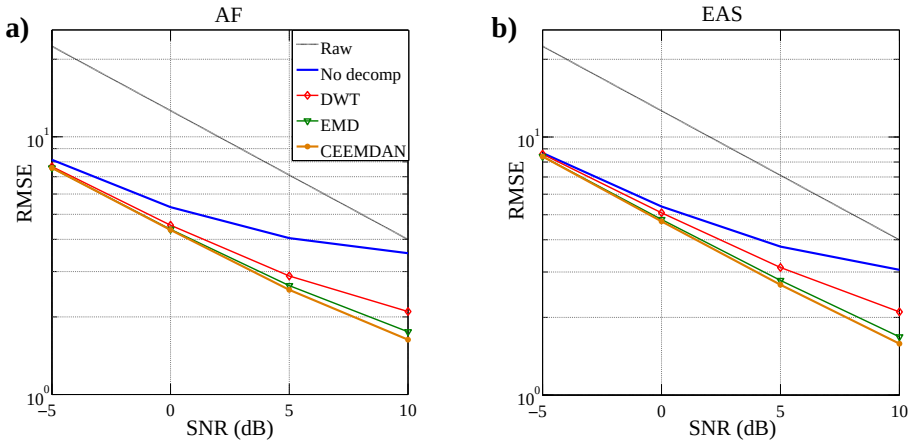

Figure III.22 – RMSEs of cleaning noisy signals in Group 1 (ECG noise at several SNRs). a) Using the adaptive filter for canceling cardiac artifacts and b) using ensemble average subtraction. The dotted lines (labeled *Raw*) correspond to the RMSE between the noisy signals and the original uncorrupted EEG and the blue lines (labeled *No decomp*) to those of original and cleaned signals only by AF or EAS. As it can be observed, the improvement of introducing EEG decomposition is particularly advantageous when SNRs are high.

Removal of both ECG and low frequency noise (Group 2)

To evaluate the denoising framework in a scenario where ECG and low frequency noise contaminate the ECG, we employed the signals from Group 2.

These tests also evaluated whether removing LFN by EEG decomposition results more advantageous than using conventional filtering. In tests without decomposition, raw signals were processed by a FIR high-pass filter with a cutoff frequency of 0.5 Hz before applying the ECG noise suppressor. As linear phase FIR filters have a constant group delay, no delay distortion is introduced in the output, but as described in Section 2.2, a very high order is needed to obtain a sharp transition band. The 382-order filter specified in Table III.2 produced a delay of approximately 1.5 seconds and was not be able to eliminate completely the low-frequency noise without partly attenuating the EEG lower bands. IIR filters could be employed to avoid long processing delays in real-time applications, but sacrificing distortion.

As it can be observed in Figure III.23, AF and EAS obtain better RMSEs when DWT, EMD or CEEMDAN are used. Due to the proximity of the cutoff frequency with the δ content and to the transition width of the FIR filter, the HPF associated with AF or EAS are the least performing solutions specially in low to moderate noise environments.

On the other hand, if a neat EEG decomposition is achieved, modes whose main frequency is under 0.5 Hz should not contain energy of higher oscillatory modes and thus, a more efficient low frequency noise removal can be expected. However, our tests reveal that the quality of the separated modes (i.e. its spectral independence) deteriorates with the highest signal-to-noise ratios and consequently, the efficacy of the denoising process decreases at the same level of HPF tests. Cleaning EEGs with LFN at 10 dB SNR results more advantageous using a high-pass filter than introducing DWT or EMD. Only denoising with CEEMDAN, thanks to its superior decomposition, performs better in all the cases.

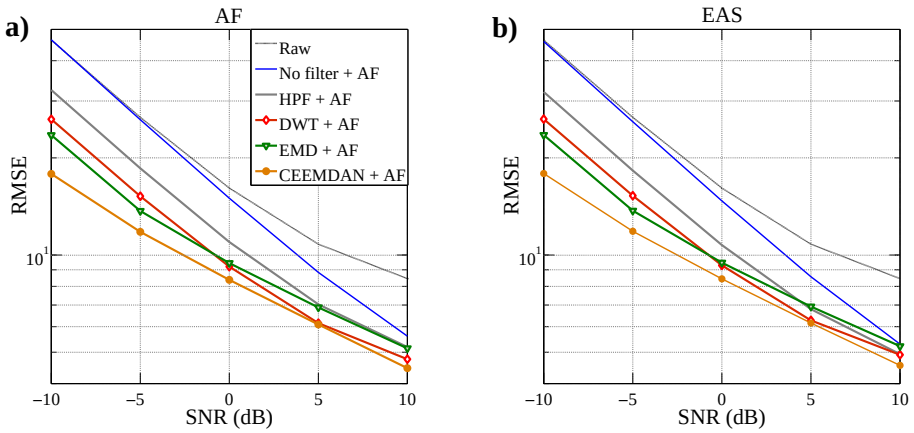

Figure III.23 – RMSEs of denoised signals from second group (ECG contamination at SNR=5dB and variable LFN) using the adaptive filter (a) and EAS (b).

6.2 Results in real data

To evaluate the efficacy of our denoising solution on real data, we first performed a basic preprocessing –using the two first blocks of the diagram described in Figure III.5– of all EEGs in the VACCIN database. Then, twenty excerpts selected randomly were denoised by the CEDAC framework using the different options to compare the efficacy of each combination.

EEG preprocessing and selection

The 31 infants constituting the VACCIN database, totaling 124 EEG records, were first preprocessed to cut the gross movement artifacts and remove other disturbances (Table III.3 summarizes rejection rates). The high-pass filter was not applied if EEG decomposition was done in CEDAC.

	A	B	Total
Recorded EEG	372 h	248 h	620 h
Rejected EEG	15.9 h (4.3%)	11.4 h (4.6%)	27.3 h (4.4%)

Table III.3 – Summary of the results of rejecting movement artifacts by thresholding. The total length of EEGs and the rejected excerpts are given for recordings before (A) and after vaccine (B). This process lead to reject 27.3 over a total of 620 h.

A 20-second sliding window looked for ECG peaks in EEG using the same algorithm employed by ensemble average subtraction (see 4.2), finding out that cardiac noise was present (to greater or a lesser degree) in the majority of the patients, having in average their 65.6% ± 30.3 of excerpts contaminated in A and the 71.9% ± 30.8 in B. Therefore, the excerpt selection only considered these noise concerned periods and ensured that different EEG channels and patients were included.

Evaluation of noise removal quality

For real EEG, the artifact-free or true EEG is unknown, hence the efficacy of denoising methods is usually studied subjectively, based on visual inspection of the cleaned signals [36, 37]. However, to have a quantitative

idea of the eliminated artifacts, Puthusserypady [38] et al. proposed the ratio of the power of the removed artifacts to the cleaned EEG, so that the higher the ratio, the better the performance of the method:

$$R = \frac{\sum_{i=1}^{N}[EEG'(i) - \widehat{EEG}(i)]^2}{\sum_{i=1}^{N}\widehat{EEG^2}(i)} \tag{III.30}$$

where EEG' and $\widehat{EEG}$ are the real EEG before and after the denoising process, respectively.

Since R gives information of the whole denoising process –comprising low frequency and cardiac noise–, in a similar manner we computed this ratio in small windows surrounding the QRS-peak instants (50 ms) on the noisy excerpt and on the cleaned one. This metric, R_{qrs}, is conceived to give a quantitative idea of how attenuated are the heartbeat artifacts.

R and R_{qrs} were computed on the selected excerpts (see Table III.4), contaminated with ECG artifacts and low frequency noise. As it can be noticed, the overall noise removal given by R is greater using CEEDMAN, although no significant differences were found with respect to the other two decomposing techniques employed, DWT and EMD. Regarding only the ECG removal (R_{qrs}), CEEDMAN continues to be the most effective combined both with AF or EAS. This table also reveals that decomposition techniques are not always advantageous to suppress QRS artifacts. This is the case of the EMD, that obtained the lowest R_{qrs} values probably for an incomplete separation of EEG_H before the ECG cancellation takes place. Knowing the percentages of critical mixing could be a possible way to corroborate this low performance. Finally, it can be also stated that adaptive filtering is more effective than ensemble average subtraction in all the combinations used for our tests, so it should be preferred in the implementation of the CEDAC framework.

Decomposition	ECG cancellation	median (R)	mean ± std (R)	median (R_{qrs})	mean ± std (R_{qrs})
None (HPF)	AF	$1.30^{a,c}$	4.95 ± 13.6	2.65^{e}	10.0 ± 12.3
DWT	AF	1.92	2.77 ± 2.98	7.34	7.89 ± 6.77
EMD	AF	1.48	6.88 ± 20.3	2.55	6.34 ± 6.60
CEEMDAN	AF	$\mathbf{2.52}^{a,b}$	**18.7 ± 62.4**	**7.76**	**9.95 ± 9.72**
None (HPF)	EAS	$1.22^{b,d}$	5.58 ± 15.4	$2.02^{e,f}$	3.42 ± 4.01
DWT	EAS	1.85	5.67 ± 12.8	1.96	3.48 ± 4.96
EMD	EAS	1.39	7.31 ± 21.3	1.31	2.95 ± 4.36
CEEMDAN	EAS	$\mathbf{2.47}^{c,d}$	**18.4 ± 62.1**	$\mathbf{2.12}^{f}$	**3.58 ± 4.75**

Table III.4 – Values for the ratios of removed artifacts, R (LFN+ECG) and R_{qrs} (ECG). An statistical analysis using Mann-Whitney U-test compared the ratios of denoising methods without decomposition (HPF + AF or EAS) and those using decomposition. Pairs having the same super-indexed letters have statistical significant differences (p-value < 0.05). In bold, best results.

7 Conclusion

In this chapter, the numerous noise sources contaminating the EEG in the NICU have been assessed to find an appropriate denoising technique. While some basic strategies, such as discarding movement-related epochs or band-pass filtering can be applied systematically, more specific techniques should be applied if certain noise sources appear discontinuously in some EEG channels.

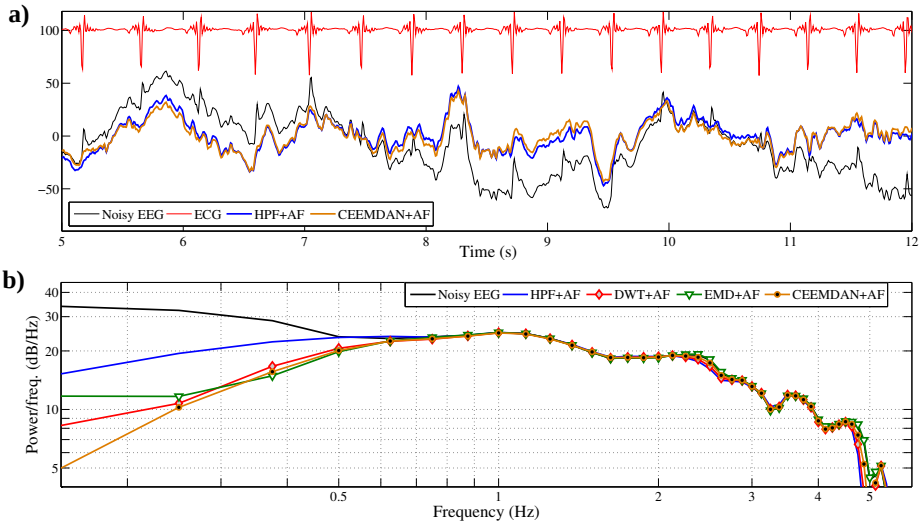

Figure III.24 – a) Real contaminated EEG excerpt recorded at NICU and the results of denoising it using HPF+FA and CEEMDAN+AF. b) Detail of the power spectrum between 0.1 and 5 Hz, where it can be observed that decomposition techniques remove low frequency noise more efficiently than a high-pass filter.

To perform an effective cancellation of cardiac artifacts we designed a well-adapted solution for NICU environments, the CEDAC framework. Since the availability of a few EEG channels and an ECG reference is a limitation for the blind source separation approach, we considered a single-channel processing method combining EEG decomposition and ECG noise cancellation. The interest here of separating the EEG in subcomponents is double-fold. Firstly, it allows to set aside the low frequency content to increase the performance of adaptive filters and secondly, it can adequately reject the low frequency noise content. In effect, very slow waves occur naturally in the preterm infant's EEG, making it vulnerable to slow artifacts as movements or sweat potentials. In our proposal, these artifacts can be automatically rejected avoiding high-pass filters and preserving at maximum the waveform of the delta content.

An ensemble of tests using simulated data from real preterm infants has been designed to compare the efficacy of the different options of CEDAC. They demonstrated that applying EEG decomposition ameliorates in general the QRS artifact suppression of adaptive filtering or ensemble average subtraction. Indeed, processing separately the components containing ECG noise is advantageous because the signal to be denoised has a smaller portion of the original EEG and thus, distortion is reduced in the final reconstructed signal. Furthermore, less complex and faster adaptive filters can be implemented when the noisy signal has a smaller bandwidth.

Mode mixing, the well-known limitation of EMD-based decompositions, has been thoroughly studied in this chapter. However this phenomenon does not compromise directly the quality of denoising because IMFs are added in two main components (EEG_L and EEG_H), reassembled once the ECG cancellation is performed. We found out that IMFs having oscillations on both sides of the ECG artifact frequency could be the main cause of an incomplete noise removal. The two possible scenarios, named critical mixing I and II, have been properly identified and quantified for EMD and CEEMDAN, permitting, in turn, to chose the appropriate decomposition parameters in our signals. In average, critical mixing produced by CEEMDAN was approximately the half of those obtained by the classical EMD.

Tests in simulated and real data revealed the association CEEMDAN+AF as the most effective one, however

it is the most time-consuming algorithm -in which EMD is executed several hundred of times to obtain good results- could be inconvenient for real-time applications. DWT in combination with AF would be, then, a trade-off between speed and performance.

Effective in EEG sleep patterns, CEDAC frameworks could be also useful in neonatal polysomnography systems and clinical tools dealing with EEG quantitative analysis (as sleep stage classification or seizure detection), where a selective artifact removal with a minimal information loss is required.

Finally, some limitations concerning the automation of the noise removal in the NICU should be pointed out. There are some sporadic artifacts still unresolved that would require a more complex denoising strategy, such as short transients dues to electrode pop-up or glossokinetic activity. A possible solution in this case could be to design a classification algorithm based in EEG features allowing to discriminate excerpts containing the mentioned artifacts. On the other hand, eye movements can produce oscillations in fronto-parietal channels during active sleep, overlapping the delta band and misleading information about true delta bursts. This could be solved by adding additional EOG electrodes to the montage, which would serve as a reference to extract ocular noise in EEG by, for instance, and adaptive filter [39]. Nevertheless, with the actual configuration it would be preferable to exploit temporal electrode positions –less sensitive to eye movements– if active sleep occurs.

Bibliography

[1] F. Wallois, M.-F. Vecchierini, C. Héberlé, and E. Walls-Esquivel. *Electroencephalography (EEG) recording techniques and artifact detection in early premature babies.* Clinical neurophysiology **37**(3), 149–161 July (2007).

[2] P. Durka, H. Klekowicz, K. Blinowska, W. Szelenberger, and S. Niemcewicz. *A simple system for detection of EEG artifacts in polysomnographic recordings.* IEEE Transactions on Biomedical Engineering **50**(4), 526 –528 April (2003).

[3] A. Oppenheim and R. W. Schafer. *Digital Signal Processing.* Prentice-hall edition (1975).

[4] J. F. Kaiser. *Nonrecursive digital filter design using the - sinh window function.* Proc. 1974 IEEE Symp. Circuits and Systems pages 20–23 April (1974).

[5] N. Erdol and F. Basbug. *Wavelet transform based adaptive filters: analysis and new results.* IEEE Transactions on Signal Processing **44**(9), 2163 –2171 September (1996).

[6] S. Hosur and A. Tewfik. *Wavelet transform domain adaptive FIR filtering.* IEEE Transactions on Signal Processing **45**(3), 617 –630 March (1997).

[7] P. Kumar, R. Arumuganathan, K. Sivakumar, and C. Vimal. Removal of artifacts from EEG signals using adaptive filter through wavelet transform. In *9th International Conference on Signal Processing, 2008. ICSP 2008*, pages 2138 –2141 October (2008).

[8] M. Browne and T. R. Cutmore. *Adaptive wavelet filtering for analysis of event-related potentials from the electro-encephalogram.* Medical & biological engineering & computing **38**(6), 645–652 November (2000).

[9] H. Olkkonen, P. Pesola, J. Olkkonen, A. Valjakka, and L. Tuomisto. *EEG noise cancellation by a subspace method based on wavelet decomposition.* Medical science monitor: international medical journal of experimental and clinical research **8**(11), MT199–204 November (2002).

[10] M. J. McKeown, R. Saab, and R. Abu-Gharbieh. A combined independent component analysis (ICA)/ empirical mode decomposition (EMD) method to infer corticomuscular coupling. In *2nd International IEEE EMBS Conference on Neural Engineering, 2005. Conference Proceedings*, pages 679–682. IEEE March (2005).

[11] J. P. Lindsen and J. Bhattacharya. *Correction of blink artifacts using independent component analysis and empirical mode decomposition.* Psychophysiology **47**(5), 955–960 September (2010).

[12] J. Taelman, B. Mijovic, S. Van Huffel, S. Devuyst, and T. Dutoit. ECG artifact removal from surface EMG signals by combining empirical mode decomposition and independent component analysis. In *Proc. of the International Conference on Bio-inspired Systems and Signal Processing, subconference of the 4th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2011)* (2010).

[13] O. Rioul and M. Vetterli. *Wavelets and signal processing.* Signal Processing Magazine, IEEE **8**(4), 14 –38 October (1991).

[14] I. Daubechies. *Ten Lectures on Wavelets.* SIAM: Society for Industrial and Applied Mathematics, 1 edition May (1992).

[15] 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.* Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences **454**(1971), 903 –995 March (1998).

[16] H. Liang, S. L. Bressler, E. A. Buffalo, R. Desimone, and P. Fries. *Empirical mode decomposition of field potentials from macaque v4 in visual spatial attention.* Biological cybernetics **92**(6), 380–392 June (2005).

[17] J. C. Echeverría, J. A. Crowe, M. S. Woolfson, and B. R. Hayes-Gill. *Application of empirical mode decomposition to heart rate variability analysis.* Medical & biological engineering & computing **39**(4), 471–479 July (2001).

[18] S. Ari and G. Saha. *Classification of heart sounds using empirical mode decomposition based features.* International Journal of Medical Engineering and Informatics **1**(1), 91 – 108 (2008).

[19] A. Torres, J. A. Fiz, R. Jané, J. B. Galdiz, J. Gea, and J. Morera. *Application of the empirical mode decomposition method to the analysis of respiratory mechanomyographic signals.* Conference proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference **2007**, 1566–1569 (2007).

[20] G. Rilling, P. Flandrin, and P. Gonçalvès. On empirical mode decomposition and its algorithms. In *Proceedings of the 6th IEEE/EURASIP Workshop on Nonlinear Signal and Image Processing (NSIP '03), Grado, Italy* (2003).

[21] P. Flandrin. Empirical mode decomposition, presentation from the ENS Lyon (france). http://perso.ens-lyon.fr/patrick.flandrin/, (2012).

[22] Z. Wu and N. E. Huang. *ENSEMBLE EMPIRICAL MODE DECOMPOSITION: a NOISE-ASSISTED DATA ANALYSIS METHOD.* Advances in Adaptive Data Analysis **01**(01), 1 (2009).

[23] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin. A complete ensemble empirical mode decomposition with adaptive noise. In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4144–4147. IEEE May (2011).

[24] P. Anderer, S. Roberts, A. Schlögl, G. Gruber, G. Klösch, W. Herrmann, P. Rappelsberger, O. Filz, M. J. Barbanoj, G. Dorffner, and B. Saletu. *Artifact processing in computerized analysis of sleep EEG – a review.* Neuropsychobiology **40**(3), 150–157 (1999).

[25] P. J. V. d. Berg and G. J. F. Brekelmans. *Rhythmic heart beat artefacts mimicking an ictal EEG pattern with abundant interictal epileptiform activity in both mother and child.* Epileptic Disorders **13**(2), 177–180 June (2011).

[26] M. Nakamura and H. Shibasaki. *Elimination of EKG artifacts from EEG records: a new method of non-cephalic referential EEG recording.* Electroencephalography and Clinical Neurophysiology **66**(1), 89–92 January (1987).

[27] K. Harke, A. Schlögl, P. Anderer, and G. Pfurtscheller. *Cardiac field artifact in sleep EEG.* (1999).

[28] H.-J. Park, D.-U. Jeong, and K.-S. Park. *Automated detection and elimination of periodic ECG artifacts in EEG using the energy interval histogram method.* IEEE Transactions on Biomedical Engineering **49**(12), 1526–1533 December (2002).

[29] S. Haykin. *Adaptive Filter Theory.* Pearson Education, 4 edition (2002).

[30] J. A. Snyman. *Practical Mathematical Optimization - An Introduction to Basic Optimization Theory and Classical.* (2005).

[31] E. Eleftheriou and D. Falconer. *Tracking properties and steady-state performance of RLS adaptive filter algorithms.* IEEE Transactions on Acoustics, Speech and Signal Processing **34**(5), 1097–1110 (1986).

[32] J. Terrien, C. Marque, and B. Karlsson. *Automatic detection of mode mixing in empirical mode decomposition using non-stationarity detection: application to selecting IMFs of interest and denoising.* EURASIP Journal on Advances in Signal Processing **2011**(1), 37 August (2011).

[33] J. Turnbull, K. Loparo, M. Johnson, and M. Scher. *Automated detection of tracé alternant during sleep in healthy full-term neonates using discrete wavelet transform.* Clinical Neurophysiology **112**(10), 1893–1900 October (2001).

[34] J. Pan and W. J. Tompkins. *A real-time QRS detection algorithm.* IEEE Transactions on Bio-Medical Engineering **32**(3), 230–236 March (1985).

[35] M. Altuve, G. Carrault, J. Cruz, A. Beuchée, P. Pladys, and A. Hernandez. Analysis of the QRS complex for apnea-bradycardia characterization in preterm infants. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2009. EMBC 2009*, pages 946–949. IEEE September (2009).

[36] K. H. Ting, P. C. W. Fung, C. Q. Chang, and F. H. Y. Chan. *Automatic correction of artifact from single-trial event-related potentials by blind source separation using second order statistics only.* Medical engineering & physics **28**(8), 780–794 October (2006).

[37] S. Devuyst, T. Dutoit, P. Stenuit, M. Kerkhofs, and E. Stanus. Removal of ECG artifacts from EEG using a modified independent component analysis approach. In *30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2008. EMBS 2008*, pages 5204–5207. IEEE August (2008).

[38] S. Puthusserypady and T. Ratnarajah. *Robust adaptive techniques for minimization of EOG artefacts from EEG signals.* Signal Processing **86**(9), 2351–2363 September (2006).

[39] D. P. He, G. Wilson, and C. Russell. *Removal of ocular artifacts from electro-encephalogram by adaptive filtering.* Medical and Biological Engineering and Computing **42**(3), 407–412 May (2004).

Chapter IV

Automatic detection of EEG bursts

Because the EEG activity in preterm infants is characterized by changing discontinuous patterns –alternating quiescence periods with slow, high voltage transients or bursts–, many authors have been interested in its study. Previous works stated a decrease of the inter-burst intervals with increasing age in healthy preterm infants [1, 2] and underlined the usefulness of assessing the EEG background activity to provide valuable information for prognostication and maturation.

Nowadays, the recognition of IBI, as many other analyses concerning the monitored EEG, is still performed manually by clinicians and therefore, an automated, continuous recognition of the EEG bursts in the NICU could contribute effectively to this purpose, saving time costs and disagreement between raters in such a subjective task. However, the implementation of an automated burst detection confronts the lack of a unified definition of burst and hence, the validation of detection algorithms are limited to 'gold standards', datasets with known classification from visual recognitions by clinicians.

Employing the EEG signals preprocessed previously, the present chapter focuses on automatic EEG burst detection and its potential applications. After reviewing the existing literature about EEG burst detection, the foundations of this kind of classification, based on supervised learning, are then laid emphasizing the two developed binary classifiers. These will be evaluated in a learning set by our gold standard –created by three clinicians from the CHU of Rennes– so the best option will be employed to detect bursts in all VACCIN infants. Finally, thanks to the information given by the hypnograms, the clean EEGs from the last chapter and the burst data obtained here, we explore two possible clinical applications: the study of post-immunization effects and the assessment of maturation.

1 State of the art

Several algorithms have already been proposed in the literature to detect EEG bursts. Simple thresholding solutions, using the well-known nonlinear energy operator (NLEO) [3], have been applied, for example, to detect burst suppression during anesthesia in adults [4]. Concerning full-term newborns, Löfhede et al. [5], compared the efficacy of different burst suppression classifiers from five EEG features and concluded that support vector machines (SVM) perform slightly better than artificial neural networks and Fisher's linear discriminant. Less complex methods, as the assessment of the instantaneous variance of aEEG during short periods [6], have also been used to investigate the relationship between *tracé alternant* slow-wave bursts and the heart ratio in preterm infants of about 36 weeks PMA. More recently, Palmu et al. [7] proposed an automatic detector of EEG bursts for extremely and very preterm infants after three days of birth (24 to 30 weeks PMA). Their detector

performed classifications by comparing the feature value yielded by NLEO with a threshold (classification by thresholding). They found the highest agreement ratios between raters (up to 86%) and the best precision in automatic detections (up to 93%) within the population having the highest degree of prematurity, most likely due to the easier interpretation of bursts in the earliest *tracé discontinu*.

In a similar way, this chapter aims at assessing the raters concordance versus the automatic detection, but employing more mature EEG patterns, containing *tracé alternant* and/or *tracé discontinu* from older preterm infants from 36 to 39 weeks PMA. We first study the disagreement between three EEG readers and thereafter we establish our gold standard, which will be used to test the performance of two EEG burst classifiers, the first one inspired on Palmu's thresholding algorithm and the second one based on logistic regression.

2 Automatic detection of bursts

This section describes the two employed methods to detect EEG bursts. To understand binary classifiers, we first provide some basic definitions, the terminology and the formalization of each approach. Then, both classifiers are positioned in the particular context of the burst detection framework proposed here.

2.1 Introduction to binary classifiers

Automatic detection in clinical applications is, in general, based on supervised learning. It is a field of study whose purpose is to classify patterns (or instances) into a set of categories (or labels) employing classification models that have been constructed (learned) from exemplary data. These pre-classified data, obtained manually by experts, is also known as gold standard.

Commonly, automatic detection involves three steps:

Data preprocessing (DP): The raw signal may be filtered and scaled to avoid bias in the classification and reduce noise.

Feature extraction (FE): A set of variables, describing quantitative characteristics of data, is derived (such as statistical properties or frequency content, etc.).

Classification (CL): Data is categorized according to a classification model which makes a decision as a function of input features.

In this chapter, the methodology for automatic detection is needed to classify EEG in 'burst' and 'inter-burst' periods, a dichotomy that can be solved by a binary classification model such as, for instance, decision trees, neural networks, support vector machines or logistic regression.

In two class or binary classification problems, data is formed by the pair $\{(\mathbf{X}_i, Y_i) : i = 1...N\}$, with the set of labels $Y_i \in \{1, 0\}$ and the set of features $\mathbf{X_i} = [X_{1,i}, X_{2,i}, .., X_{d,i}]$, $\mathbf{X_i} \in \mathbb{R}^d$ where N is the number of training examples and d the number of features. Y is considered as a Bernoulli random variable whose class 1 probability is $\pi = P(Y = 1)$ and the class 0 probability is $1 - \pi = P(Y = 0)$.

Having knowledge of the conditional distribution of $\mathbf{X}$, $P(\mathbf{X}|Y = 1)$, the aim of classification is to predict $P(Y = 1|\mathbf{X})$, i.e. future outcomes. π can be then interpreted as a function of the instances, $\pi(\mathbf{X}) = P(Y = 1|\mathbf{X})$. This leads to two possible manners of tackling the classification problem of interest in this work:

1. $\pi(\mathbf{X})$ maps the zones where $Y = 1$ is more likely and is used to classify future outcomes. In this approach, we can find thresholding classification.

2. $\pi(\mathbf{X})$ is approximated by fitting a model, obtaining the model parameters from the instances. The logistic regression belongs to this category.

In binary classification, class membership is normally decided by an activation function, $f : \mathbb{R} \to \{0, 1\}$, which attaches to a set on unseen instances $\mathbf{X}$ the predicted labels, $\widehat{Y}$. The loss function or prediction error, defined as $L(f) = P(f(\pi) \neq Y)$, is then minimized according to the Bayes rule to obtain the appropriate model parameters (a threshold or a logistic function, in our case).

2.2 Thresholding-based classifiers

Definition

Thresholding is a simple solution that can be employed when a single random variable X is mapped according to two distributions, associated to each class. This is the case, for example, of voltage amplitudes in EEG bursts or amplitudes of motion-artifacted periods. π and $1 - \pi$ are known, and new instances are classified by partitioning X in two disjoint regions by a threshold T (e.g a voltage). Consequently, the activation function:

$$f(X) = \begin{cases} 1 & \text{if } X > T \\ 0 & \text{if } X \leq T \end{cases} \tag{IV.1}$$

The threshold is chosen according to some optimality criterion, usually evaluating the misclassification rate, or inversely, the agreement rate, AR. Being y and $\widehat{y}$ a single label and a single prediction for a given example, AR is defined as:

$$AR = \frac{1}{l} \sum_{k=1}^{l} y_k \widehat{y}_k + (1 - y_k)(1 - \widehat{y}_k) \tag{IV.2}$$

where l is the total number of labels in Y.

Burst detection by thresholding

The thresholding-based framework chosen to perform the burst detection is very similar to the proposal of Palmu et al. [7, 8], but introducing slight modifications to adapt the detector to our case and to study the influence of three different features. Briefly, in Palmu's algorithm the EEG was first processed by a passband filter with low cutoff frequency f_{LC} and high cutoff frequency f_{HC}. Next, the NLEO operator was computed in WIN-second averaging windows so that values over a threshold SAT_{THR} yielded a first classification, corrected later by eliminating bursts below a minimal predefined duration. By means of an optimization process, the mentioned parameters were adjusted to obtain a maximum agreement with the gold standard.

In our detection framework (see Figure IV.1), we replaced the passband filter by a low-pass filter with cutoff frequency f_L to attenuate irrelevant frequencies –the *tracé alternant* is predominantly composed by delta and theta waves–. A high-pass step can be avoided because the CEDAC framework ensures EEGs without low frequency noise. Next, an activity function Φ is applied on the filtered EEG (EEG') to obtain a feature signal X.

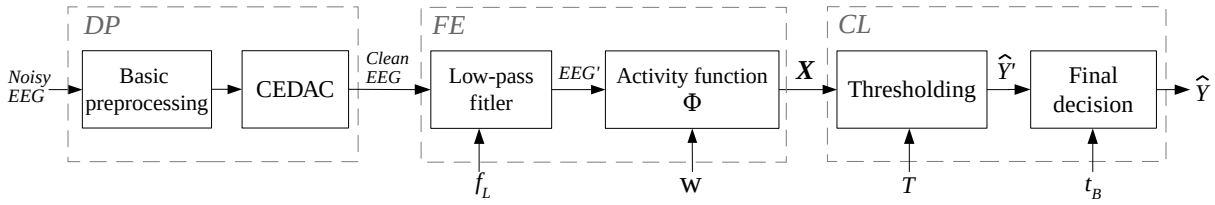

Figure IV.1 – Diagram of the detection algorithm, identifying the three main blocks. In DP (data preprocessing) the noisy signal is properly cleaned the denoising framework developed in last chapter. In FE (feature extraction), the EEG is filtered again to emphasize the bursts, then characterized by an activity function to obtain the feature signal X. Finally, in CL (classification) the decision threshold T performs a first burst/inter-burst classification, later corrected as a function of the minimal duration of bursts in the gold standard, t_B.

Unlike the original algorithm, we introduced the possibility of testing three different features by changing Φ. They were computed in sliding windows of W seconds by the following operators:

1. Mm: The difference between the maximal and the minimal values of the window. The activity function results in:
$$\Phi_{Mm} = \max_{k=1,..l}\{EEG(k), EEG(k) \geq 0\} - \min_{k=1,..l}\{EEG(k), EEG(k) < 0\} \tag{IV.3}$$
where l is the number of samples in W seconds.

2. ADIF: The sum of absolute signal values contained in W divided by its sample number [9]. It produced an activity function:
$$\Phi_{ADIF} = \frac{1}{FsW}\sum_{k=1}^{l}|EEG(i)| \tag{IV.4}$$
where Fs is the sampling frequency.

3. NLEO: Nonlinear energy operator. We employed the version used by Särkelä et al. [4], which reflects both frequency and amplitude in the signal at each sample:
$$NLEO(k) = EEG(k)EEG(k-3) - EEG(k-1)EEG(k-2) \tag{IV.5}$$
The NLEO operator in a window W is then computed as:
$$\Phi_{NLEO} = \frac{1}{FsW}\sum_{k=1}^{l}|NLEO(k)| \tag{IV.6}$$

Feature signals (X_{Mm}, X_{ADIF} or X_{NLEO}) are subsequently compared to their corresponding burst threshold (T) in the primary classification block, which yields the primary classification signal, $\widehat{Y}$. So, if $x > T$ there is a positive identification of burst ($\widehat{Y}' = 1$) and otherwise an inter-burst detection ($\widehat{Y}' = 0$). The final decisions are taken in next block depending on the minimal length of bursts, t_B, obtained from the gold standard. As in Palmu's algorithm, bursts not exceeding t_B are discarded in the final classification signal, $\widehat{Y}$. An example showing the signals involved in this procedure can be seen in Figure IV.2.

To obtain the values for parameters W and T, an optimization process needs to be run. The other detection parameters (f_L and t_B) can be fixed from the prior knowledge of data, in particular the frequencies characterizing

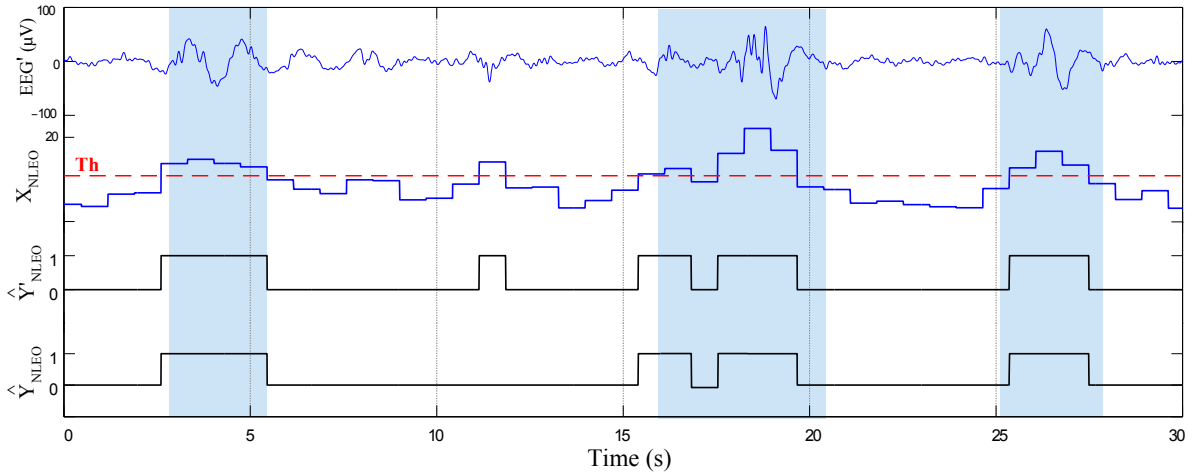

Figure IV.2 – Example of step-by-step burst detection using NLEO. EEG' is the low-pass filtered EEG and X_{NLEO} the activity signal from which the threshold T determines a first binary classification ($\widehat{Y}'$). Finally, detection is improved by removing short bursts ($\widehat{Y}$). For comparison, the shadowed areas delimit the gold standard bursts.

the EEG and the distribution of bursts in the gold standard. More information about optimization and tuning of parameters can be found in Section 4.2.

2.3 Logistic regression-based classifiers

Definition

Logistic regression (LR) is a type of regression analysis derived from the linear case. In linear regression, a function $g : \mathbb{R}^d \to \mathbb{R}$ related to the probabilities π_i, is expressed as a linear combination of $\mathbf{X_i}$:

$$g(\mathbf{X_i}) = w_0 + w_1 X_{1,i}, .., w_d X_{d,i} = w_0 + \boldsymbol{w} \cdot \mathbf{X_i}, \tag{IV.7}$$

where $\boldsymbol{w} = [w_1, .., w_d]$ is the vector of regression coefficients and w_0 is the intercept coefficient. In logistic regression, π_i is expressed through a function called logit, of the form:

$$logit(\pi_i) = \ln\left(\frac{\pi_i}{1-\pi_i}\right) = w_0 + \boldsymbol{w} \cdot \mathbf{X_i}. \tag{IV.8}$$

The inverse of the above expression, called logistic function, is expressed as:

$$logit^{-1}(\pi_i) = \frac{1}{1 + e^{-(w_0 + \boldsymbol{w}\mathbf{X_i})}} = g(\mathbf{X_i}, \boldsymbol{w}). \tag{IV.9}$$

An important characteristic of the logistic function is that it is bounded between 0 and 1, and thus, it can be used directly to model the probabilities of the possible outcomes:

$$P(Y = 1|\boldsymbol{w}, \mathbf{X_i}) = g(\mathbf{X_i}, \boldsymbol{w}). \tag{IV.10}$$

Given the set of N examples $\{\mathbf{X_i}, Y_i\}$, the learning process aims at finding the best $\boldsymbol{w}$, which is to maximize

the conditional probabilities $P(Y_i|\mathbf{X_i}, \boldsymbol{w})$. This can be achieved by maximizing the log likelihood function:

$$L(\boldsymbol{w}) = \sum_{i=1}^{N} Y_i \ln P(Y_i = 1|\mathbf{X_i}, \boldsymbol{w}) + (1 - Y_i) \ln P(Y_i = 0|\mathbf{X_i}, \boldsymbol{w}). \tag{IV.11}$$

The maximum likelihood estimate (MLE) gives the optimal coefficients, $\widehat{\boldsymbol{w}}$, employed to estimate the probabilities, $\widehat{\pi}_i$. It is found by a numerical optimization, a process that starts with a set of initial values and iterate according to a hill-climbing algorithm such as Newton-Raphson's method [10].

Several approaches exist to assess the goodness of the fitted model. The principle of some of them is to estimate two models and compare the fits removing predictor variables from one of the two models. By doing this, the model having less variables (the most restrictive) fits less well, showing a lower log likelihood. The likelihood ratio test, for instance, compares the log likelihoods of the two models and if this difference is statistically significant, then the less restrictive model is chosen as it fits the data significantly better.

The Wald test approximates the likelihood ratio test but only requires estimating one model. Essentially, it tests the null hypothesis that a set of parameters is equal to some value, zero if some variable is removed. If the test fails to reject the null hypothesis, then removing the variables from the model does not harm the fit of that model significantly, because a predictor with a coefficient that is very small relative to its standard error does not generally contribute substantially predict the dependent variable [11]. The contribution of a given coefficient $\widehat{w_j}$ to the model can be quantified by the z-statistic:

$$z = \frac{\widehat{w_j}^2}{SE^2}, \tag{IV.12}$$

where SE is the standard error (an estimation of the standard deviation) of the estimated parameter.

The decision of class membership given by the activation function is bounded by a cut-off value c, such that $f(\widehat{\pi}_i) > c$ assigns the predictive output value, $\widehat{y}$, to class 1, and $f(\widehat{\pi}_i) \leq c$ assigns $\widehat{y}$ to class 0. By default $c = 0.5$, but it can be adjusted to certain specificity or sensitivity specifications.

Detection by logistic regression

The proposed LR detector is also composed by three essential classification blocks as in the thresholding diagram. But as it can be observed in Figure IV.3, only DP is implemented by the same functions.

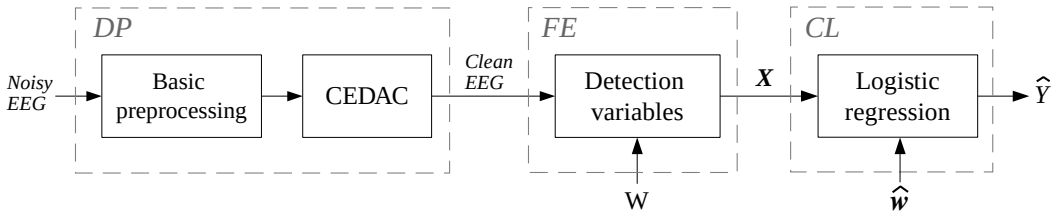

Figure IV.3 – Diagram of the LR detection algorithm, identifying the three main blocks as in Figure IV.1. In this classifier there is no need to add a low-pass filter in FE nor a final decision function in CL.

Unlike the previous approach, the LR classifier does not need any extra low-pass filtering in FE because it obtains the features directly from the clean EEG. In the above diagram, 'Detection variables' block yields the

set of features, nine continuous variables computed in windows of W seconds, containing $n_s = WFs$ samples. They have been chosen arbitrarily, trying to retain different characteristics of the EEG. They can be divided in:

- Amplitude features:
  1. Mm: Computation of the Mm operator in the current window.
  2. Mm_{-1}: Mm value in the previous window.
  3. Mm_{+1}: Mm value in the following window.
  4. DM: Maximum of absolute values of the derivate in W:

$$DM = \max_{k=1,..n_s} \{|EEG(k) - EEG(k-1)|\}. \tag{IV.13}$$

- Statistical features:
  5. Sd: Standard deviation of W.
  6. Kt: Kurtosis of W.
- Energy features:
  7. NL: Computation of the NLEO operator in the current window.
  8. RMS: The root mean square value in W, computed as:

$$RMS = \sqrt{\frac{1}{n_s}\sum_{k=1}^{n_s} EEG^2(k)}. \tag{IV.14}$$

- Frequency features:
  9. MF: Main frequency in the EEG segment, defined as:

$$MF = \frac{1}{n_s}\sum_{k=1}^{n_s} |EEG(k) - EEG(k-1)|. \tag{IV.15}$$

Finally, the CL block applies the logistic function (using the regression coefficients, $\widehat{\boldsymbol{w}}$, estimated in the learning process) to the set of features $\mathbf{X}$, yielding the classification signal, $\widehat{Y}$.

As performed in thresholding, the size of the window computing the features has been optimized manually, i.e. varying W between two limits and retaining the value maximizing the agreements with the gold standard. On the other hand, obtaining $\widehat{\boldsymbol{w}}$ given W and the set of features, is optimized by a numerical algorithm finding the MLE. In principle, the features can be chosen arbitrarily but their contribution to the classification model need to be studied, removing less significant variables to avoid over-parametrization. Further details about the set up of detection parameters can be found in Section 4.2.

3 Evaluation methodology

Usually, bursts are defined as EEG activity exceeding some voltage threshold during a variable lapse of time in all channels. Nevertheless, voltage criteria can limit the interpretation of bursts because amplitudes can be

modified by inappropriate filtering or data recording, even if in practice expert's eye can compensate this effect. Moreover, other factors can influence visual detection, for instance, if the identification is based on hypo-activity periods rather than in the bursts limits, or if the morphology of bursts are considered [12]. All these factors can bring up questions such as how reliable are manual interpretations and if burst classifications could gain consistency by introducing automatic detectors. Therefore, the interest of this experiment is:

1. To quantify the clinicians subjectivity by knowing the inter- and intra-rater agreements.
2. To establish a gold standard serving as a reference for automatic detection.
3. To compare the performance of automatic versus visual detections.

3.1 Study design

Manual marks were carried out on ten EEG examples ($N = 10$) of five minutes selected from different infants of the VACCIN database. EEGs were selected in quiet sleep periods (according to the hypnograms) to ensure the inclusion of discontinuous patterns. The newborns had post-menstrual ages from 36.1 to 39.7 weeks (more information is provided in Table IV.1).

Num.	Patient	GA	PMA	PNA	Birth weight	Day weight	Sex
1	011208LT01	27,86	37,71	69	1420	2600	1
2	020709RC01	27,14	37,14	74	880	1965	2
3	021109ZB01	29,43	38,00	59	1510	2595	1
4	040510GC01	28,57	37,00	59	1040	2145	2
5	080210RL01	30,29	38,00	54	1760	3050	2
6	100609SE01	32,71	39,71	49	1330	2800	1
7	190110LL01	25,86	36,14	71	820	2150	1
8	210109LC01	30,29	38,57	58	1150	2600	1
9	280109BY01	28,00	38,00	66	750	1975	1
10	290610GH01	26,71	38,71	84	755	2690	2

Table IV.1 – List of selected infants, describing gestational (GA) and post-menstrual (PMA) in weeks, post-natal age (PNA) in days, the weights of birth and current day in grams and sex (1 male; 2 female).

The four EEG channels (Fp1,Fp2,T3,T4) were previously filtered to ensure the nonexistence of artifacts and were displayed simultaneously to the reader as the pairs Fp1-T3 and Fp2-T4 in 20-second windows. The ECG channel and the respiratory trace were also available to the raters.

Three neonatologists from the CHU of Rennes (named A, B and C) were recruited to evaluate the EEGs. They judged if a burst was present of absent, by just marking the limits of the burst on the Fp1-T3 lead. The raters were also asked to replicate the interpretations in a different day.

The only criterion was to consider a burst as a clear EEG event containing slow activity (< 2 Hz) with superimposed faster waves, being their limits the time instants where a deviation from the baseline can be observed. Any voltage or frequency criteria was avoided purposely for two main reasons:

1. Bursts are always arbitrary to EEG readers, so strict impositions could influence their decisions in the experiment.
2. To allow a certain flexibility ('eye normalization') if the EEG amplitude suffers inter-individual variations or if signal preprocessing modified the original record.

To provide the raters an easy to use interface to mark the EEG bursts, a Matlab program implemented under the GUI environment has been designed (see Figure IV.4). In this software, the evaluation of a record is done in four steps: 1) Enter the identification code of the rater (A,B or C), 2) Select one of the ten EEG records, 3) Mark the begin and end of a bursts in each window until the end of the record. The interface includes a sliding bar to evaluate the quality of the record (left position meaning that the EEG is not interpretable and the opposite meaning an optimal quality) but it was not employed in the present study.

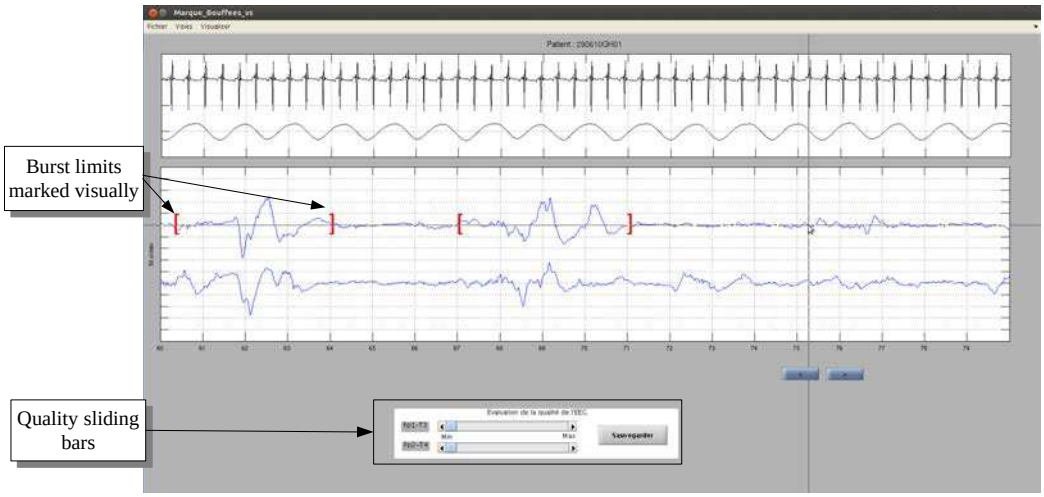

Figure IV.4 – Appearance of the burst detection interface during the visual marking procedure. Burst limits are enclosed in red brackets. At the bottom, two sliding bars evaluate subjectively the quality of the EEG.

In the above procedure, continuous time series (the EEG signals) are converted to discrete series so that the presence of bursts (class 1) is tabulated by discrete sequences of 1's and the absence of them (class 0) by 0's. Thus, each manual detection yields a binary array $Y_{r,i} \in \{1, 0\}$ of length l, where r represents the rater's code, $i = \{1, 2\}$ is the repetition number and l the length of the evaluated EEG. For example, the first manual detection for clinician A is noted $Y_{A,1}$.

3.2 Measures of agreement

There are numerous statistical measures to evaluate the reliability of manual scorings, depending if scores are continuous or categorical data. In our study, the raters had the freedom of choosing the onset and the end of a burst, so labels need to be discretized according to a sample rate ($Fs = 128$ Hz). Hence, the EEG is not labeled by fixed time segments, but split up into 1 and 0 sequences of unequal size. According to an essay of Norman et al. [13], continuous labels lead commonly utilized statistics for inter-rater (Fleiss' kappa) and intra-rater agreements (Cohen's kappa) to meaningful bias. These measures are intended to express to which extent the agreements are not produced by chance but discretized labels –that may be highly serially dependent–, violate the independence assumptions of the kappa statistic.

An appropriate measure of reliability would be to compute the overall agreement rate, AR_o, proposed by Fleiss [14], also adopted by the before-mentioned work of Palmu et al. [8]. AR_o is obtained by computing the ratio of the number of agreements between R raters over the number of possible agreements, expressed formally

as:

$$AR_o = \frac{1}{SR(R-1)} \sum_{i=1}^{S} nb_i(nb_i - 1), \tag{IV.16}$$

where S is the total number of labeled samples and nb_i is the number of times sample i has been classified as a burst ($nb_i \in \{0, 1, ..R\}$).

The above equation was used for obtaining the overall inter-rater agreement between the three raters ($R = 3$), between a pair of raters ($R = 2$) and for intra-rater agreements ($R = 2$, using the two replications). In practice, we took advantage of binary marks to compute the agreements using logical functions (see Section 4.1).

3.3 Gold standard

The gold standard constitutes a solid clinical criterion that supplies the reference labels to solve the classification problem. In the particular case of EEG bursts, it can be defined as the unanimous decisions between operators' marks [8].

The binary array containing the gold standard, Y, was obtained from the final marks for each operator, Y_r. The latter is constructed by including the bursts of the two replicates as illustrated in Figure IV.5.

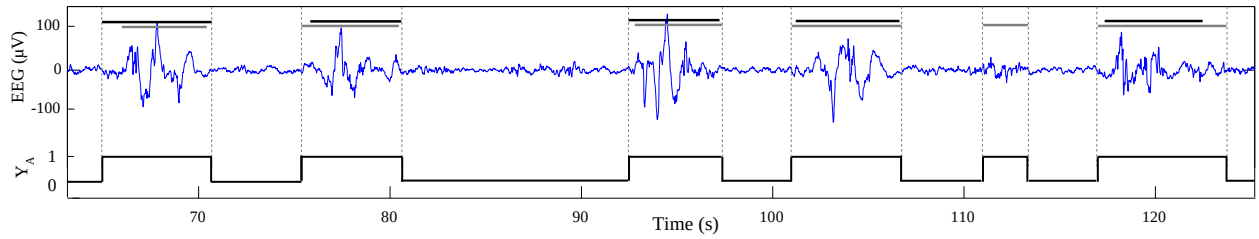

Figure IV.5 – Illustration of how the final marks for a given rater are generated. Upper black and grey lines are the first and second replications, respectively and the square signal on the bottom is the Y_r signal for rater A, Y_A.

The reason to adopt this criterion rather than its alternative (an inter-burst inclusive version) was that it seemed more consistent after inspecting the two repetitions, i.e. bursts in a hypothetic third replication would be more likely to fall within the limits merging the first an second burst. Expressing Y_r formally, we have:

$$Y_r = y_r(k) = \begin{cases} 0 & \text{if } y_{r,1}(k) + y_{r,2}(k) = 0 \\ 1 & \text{otherwise} \end{cases} \tag{IV.17}$$

for $k = 1...l$ and $r = \{A, B, C\}$. Note that the above expression is equivalent as the logical operation OR, thus $Y_r = \text{OR}(Y_{r,1}, Y_{r,2})$.

Given that the gold standard incorporates unanimous decisions, bursts are only validated if the final labels of all raters agree –time intervals where all of them identified a burst– and analogy, inter-bursts were validated in the intervals where all raters did not recognize a burst. Otherwise, intervals without agreement are not considered. Thus, Y contains sequences of 1's in bursts, 0's in inter-bursts and empty values (ø) in disagreement zones (see Figure IV.6).

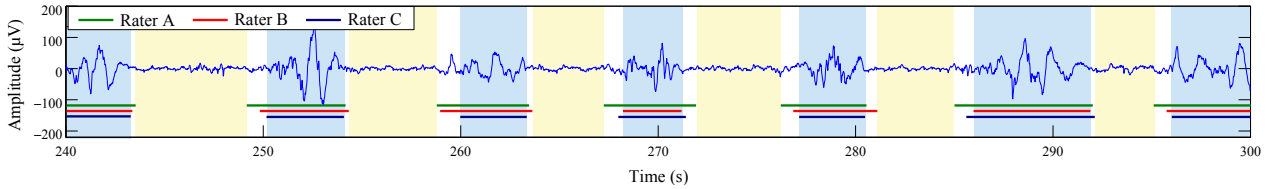

Figure IV.6 – Example of one-minute EEG signal with the raters labels (Y_A, Y_B, Y_C, see legend). Only the consensual marks are taken into account to establish the gold standard (yellow and blue areas). Areas in white represent disagreeing zones, not considered to construct the gold standard.

Following the notation used in Equation IV.17:

$$Y = y(k) = \begin{cases} 1 & \text{if } \prod_r y_r(k) = 1 \\ 0 & \text{if } \sum_r y_r(k) = 0 \\ \varnothing & \text{otherwise} \end{cases} \tag{IV.18}$$

with $k = 1...l$. Translating this criterion to logical operators, we employed the AND function for bursts and the OR function for inter-bursts.

3.4 Measures of performance

To judge the quality of the burst detectors, automatic marks need to be compared to the gold standard. When a binary classifier predicts the category of an observation correctly ($\widehat{y} = y$), the agreement can be either in class 1 (True positive, TP) or in class 0 (True negative, TN). Incorrect predictions ($\widehat{y} \neq y$) can result in labeling the true class 1 as 0 (False negative, FN) or conversely, in labeling the true class 0 as 1 (False positive, FP). From these four possible combinations, several measures of performance are derived. In particular, we are interested in the following:

- **Sensitivity**, Sn, also called true positive rate, is the probability of correct detection of positives (e.g to detect a burst when a burst is present). It is found by the expression:

$$Sn = \frac{TP}{TP + FN} \tag{IV.19}$$

- **Specificity**, Sp, or true negative rate, measures the proportion of negatives which are correctly identified as such (e.g the probability of not detecting bursts when they are absent):

$$Sp = \frac{TN}{TN + FP} \tag{IV.20}$$

Nevertheless, the sensitivities and specificities calculated from the observed data are overly optimistic. This is why a validation procedure called leave-one-out cross-validation (LOOCV) has been utilized. In LOOCV, the performance of a classifier is examined in datasets that did not take part in the training process, so unbiased evaluations of classification rules can be performed. This involves four steps:

1. Leave out a single sample from the set.

2. Build the model with the remaining data ($N-1$ examples).

3. Test the omitted sample with the learned model.

4. Repeat the above steps until each sample has been omitted and tested once.

Once this test is completed, N receiver operating characteristic (ROC) curves can be plotted. These curves represent a sensitivity/specificity pair corresponding to a particular decision threshold.

In practice, Sn and Sp will be computed bearing in mind that disagreement zones in Y have to be excluded. Figure IV.7 illustrates an example of gold standard and the corresponding automatic detection, differentiating the zones to measure the performance of classification.

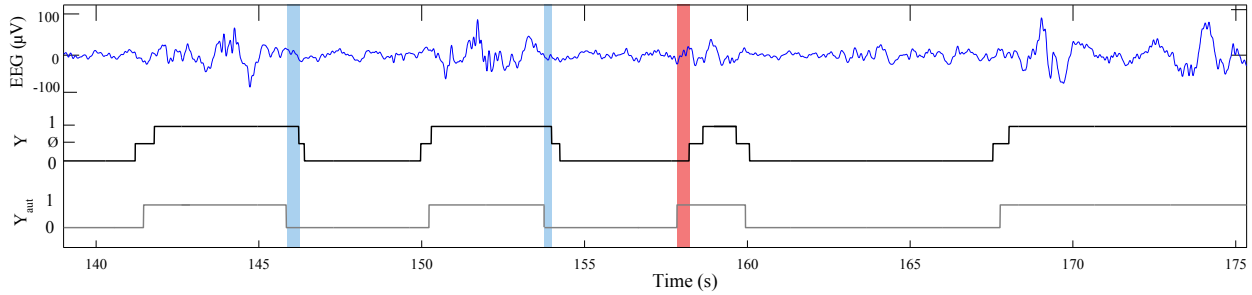

Figure IV.7 – Comparison of Y with automatic detection, Y_{aut}. Zones in blue and red represent false negative and false positive detections, respectively. Level ø in Y (disagreeing zones) is not considered to compute performance measures.

4 Results

4.1 Raters agreement

Intra- and inter-rater agreements have been computed by applying Equation IV.2 to the binary arrays representing operator's marks. Intra-rater agreements are equivalent to *not exclusive or* ($!XOR$), because regarding the first and second marks only identical decisions ($!XOR(1,1) = !XOR(0,0) = 1$) are accounted. Hence, for rater A:

$$AR_A = \frac{1}{S}\sum_{k=1}^{S} !XOR(y_{A,1}(k), y_{A,2}(k)), \tag{IV.21}$$

where S is the number of labeled samples of the binary signals. AR_B and AR_C were computed equivalently.

In a similar manner, inter-rater agreements between two operators consider equal decisions. For instance, for A and B:

$$AR_{AB} = \frac{1}{S}\sum_{k=1}^{S} !XOR(y_A(k), y_B(k)). \tag{IV.22}$$

The overall inter-rater agreement was obtained from the three raters, hence the number of actual agreements between all raters over the total number of possible agreements have to be considered [14]. Still using binary

functions, this results in the following expression:

$$AR_{ABC} = \frac{1}{3S}\sum_{k=1}^{S} !XOR(y_A(k), y_B(k)) + !XOR(y_A(k), y_C(k)) + !XOR(y_B(k), y_C(k)). \quad \text{(IV.23)}$$

Table IV.2 summarizes the agreements analyzed in this experiment. Comparing results with similar works, we obtained a lower overall inter-rater agreement ($AR_{ABC} = 78.7\%$) than Palmu et al. [8], who found 81% accordance in very preterm and 86% in extremely preterm infants.

	Intra-rater			Inter-rater (paired)			Inter-rater (all)
Num.	AR_A	AR_B	AR_C	AR_{AB}	AR_{AC}	AR_{BC}	AR_{ABC}
1	88.8	83.8	89.0	83.7	83.8	90.4	86.0
2	84.3	80.0	85.2	76.2	70.6	82.4	76.5
3	92.6	93.4	85.8	84.2	88.3	86.8	86.5
4	88.1	95.2	87.9	77.9	89.6	82.2	83.3
5	87.2	88.4	89.0	68.3	76.5	81.4	75.4
6	78.0	75.7	82.6	74.0	70.0	77.5	73.9
7	89.5	89.7	90.1	82.2	81.8	90.7	84.9
8	80.1	72.7	79.3	66.7	65.5	68.7	67.0
9	84.0	93.6	87.6	66.2	76.1	83.1	75.1
10	84.5	85.8	84.5	84.5	74.6	75.9	78.4
Mean	**85.7**	**85.9**	**86.1**	76.5	77.8	81.9	**78.7**

Table IV.2 – Agreements (%) between first and second visual marks in each rater (intra-rater agreements) and between different raters (inter-rater agreements). As evidenced by the highest intra-rater agreement (86.1%), operator C is the most self-consistent. Operators B and C, with $AR_{BC} = 81.9$ have the closest labeling criterion.

The decreased inter-rater agreements in more developed preterm infants can be explained by the increased maturation of EEG, which implies augmented background activity and hence discrepancy between raters is more plausible. Since our cohort is constituted by extremely to moderate preterm infants evaluated at about 38 weeks PMA, a 78.7% agreement does not seem to be incongruous.

4.2 Selection and optimization of detection parameters

In automatic classification, the detection parameters need to be learned so that detections lead to a maximum agreement to the gold standard. In the threshold detector, this iterative process has been optimized by modifying the parameters within certain limits decided previously. In a different manner, the optimization in the LR detector finds numerically the model which maximizes a likelihood function.

Optimization in thresholding

As introduced above, the optimization process varies a set of detection parameters and evaluates, at each iteration, the agreement with the gold standard. Since the boundaries for these parameters have been chosen deliberately, it could be said that this procedure optimizes "manually".

For each thresholding operator (Mm, ADIF, NLEO), the optimal values for parameters f_L, W, T and t_B need to be found. However not all of them required an optimization:

- f_L: The low-pass filter (Butterworth, order 5) was set to attenuate the EEG activity above the theta band. Simulations with $f_L = 8$ Hz provided the best results.

- t_B: It was set to 1 second because shorter bursts were not observed in the gold standard.

On the contrary, the following parameters were resolved by optimizing the values between given limits:

- W: The sliding window had to be large enough (to avoid an excessive resolution that could split long bursts) yet not overmuch (to avoid including low EEG activity into bursts periods), so it was set from 0.2 to 3.2 seconds in steps of 0.1.
- T: Its range depended on the employed operator. For Mm it was fixed from 30 to 60 μV in steps of 1 μV. ADIF from 5 to 15 and NLEO from 2 to 12, both with an increasing factor of 1/3. In all cases, we performed a normalization to compensate eventual baseline changes by extracting to current thresholds the minimal value of T in the past minute [8].

The agreement rates were found by the following expression, equivalent to Equation IV.2:

$$AR_Y = \frac{1}{l}\sum_{k=1}^{l}!XOR(\widehat{y}(k), y(k)). \tag{IV.24}$$

where $\widehat{y}(k)$ refers to the automatic detection and $y(k)$ to the gold standard bursts.

The three detectors were then applied to the learning set (the ten selected EEGs) and the mean agreements, $\overline{AR}_Y$, were obtained from the average of AR_Y over the ten datasets at each simulation step. The maximal mean agreement was obtained by the ADIF operator, a 93.0%, setting W to 0.90 seconds and T to 6 μV. Mm with $\overline{AR}_Y = 92.8\%$ ($W = 1.10$ s, $T = 34\mu V$) and NLEO with $\overline{AR}_Y = 91.3\%$ ($W = 0.90$ s, $T = 3.33$) yielded slightly

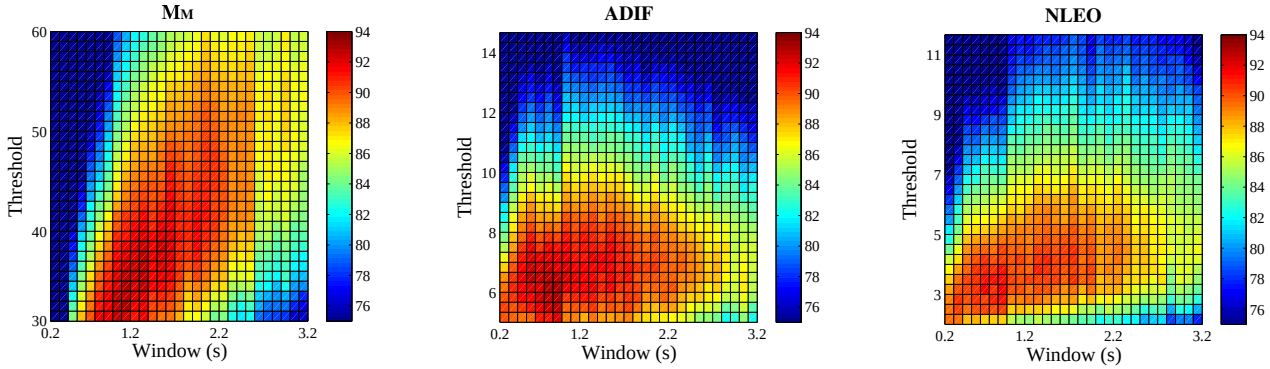

Figure IV.8 – Mean agreements using all datasets with the gold standard for each detection algorithm. Dark red zones correspond to optimal detection values. The modified parameters were the sliding window length (W) and the detection threshold (T).

The combination of the values yielding maximal mean agreements have been used to perform later the tests of performance (see Section 4.3) by leave-one-out cross validation.

Parameter tuning in logistic regression

In a first instance, the logistic function of the classifier was fitted using the labels of the $N = 10$ examples and the nine initial features, modifying only the length of the window. Apart from W, there was no need to optimize

manually any other parameter because a numerical function based in Newton-Raphson algorithm perform the optimization by finding the MLE. In this particular case, we employed the $glmfit$ function included in the Statistics toolbox of Matlab.

Another advantage of LR classifiers is that the estimated function yields directly the class probability and the default cut-off ($c = 0.5$) is implicitly the value maximizing the agreement with the gold standard.

Hence, the manual optimization required to fit an ensemble of logistic regression models, varying the length of detection windows from $W = 0.3$ s to 2.3 s in steps of 0.05. We found the average agreement rates with all examples (see Figure IV.9) and retained the LR model maximizing $\overline{AR}_Y$ (96.3%), corresponding to $W = 0.8$ s.

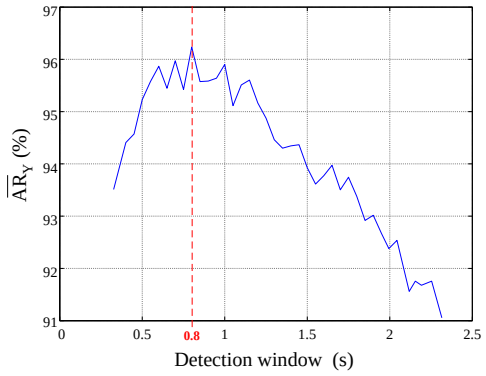

Figure IV.9 – Average agreements with the gold standard ($\overline{AR}_Y$) of the detections performed by the logistic regression at several window lengths.

The estimated coefficients ($\widehat{w}$) for the optimal window are depicted in Table IV.3, where the contribution of each coefficient can be analyzed through the p-value of the Wald test. The kurtosis (Kt), with a p-value over the significance level (0.05), does not contribute significantly to the regression. Dismissing this variable, our LR model is finally constituted by eight features.

Dep var →	Intercept	Mm	Mm_{-1}	Mm_{+1}	DM	Sd	Kt	NL	RMS	MF
$\widehat{w}$	-6.2084	0.0064	0.0293	0.0166	-0.1243	-0.1475	0.0011	-0.0127	0.2567	1.5850
SE	0.0316	0.0013	0.0002	0.0002	0.0025	0.0050	0.0074	0.0006	0.0033	0.0188
z	196.28	5.0690	143.13	93.625	49.104	29.589	0.1430	22.547	77.855	84.378
p	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	0.8863	< 0.001	< 0.001	< 0.001

Table IV.3 – Estimated coefficients ($\widehat{w}$) and standard error (SE) of the fitted LR model. The z-statistics and p-values of the Wald test are also given. As it can be observed, Kt has a very small contribution to the model (small z), so its removal does not harm substantially the model (p=0.8863).

4.3 Performance of automatic detections

The averaged agreement ratios obtained in last subsection suggest that the LR-based detector is the most performing, followed by the thresholding detectors (ADIF, Mm and finally NLEO). But to have a more complete and unbiased notion of the algorithms accuracy, we performed leave-one-out cross-validation on the ten signals composing the learning set.

This resulted in ten ROC curves, obtained with the optimal W and modifying T in Mm, ADIF and NLEO or modifying the cut-off value in LR (see Figure IV.10).

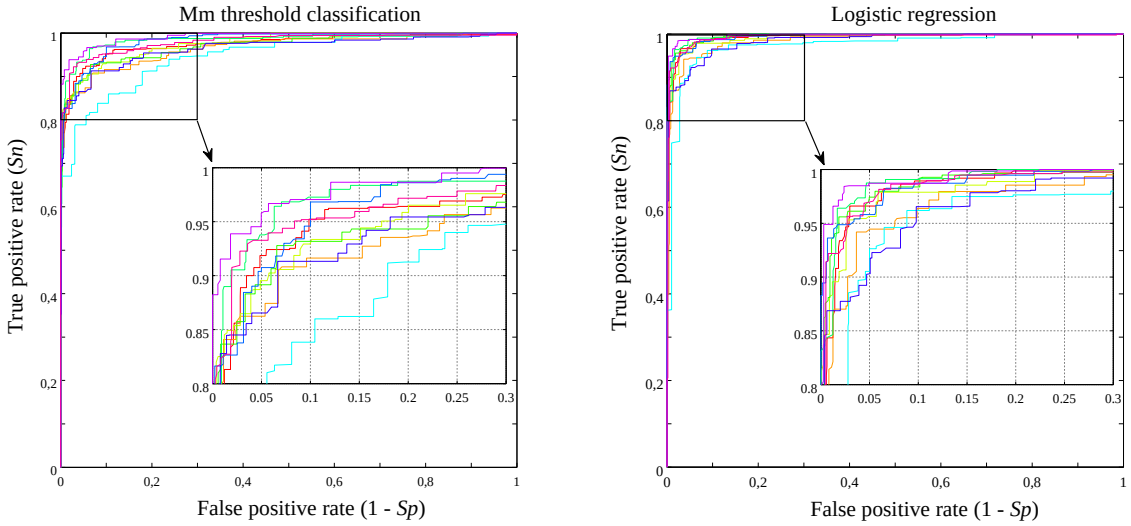

Figure IV.10 – ROC curves of the ten datasets. Each point on a ROC curve represents a sensitivity/specificity pair corresponding to a particular decision threshold. In left panel, the detection threshold T is modified for the Mm detector. In right panel the cutoff value c is modified for the LR detector. Curves having closer points to the upper left corner have better detections.

The corresponding areas under curve (Figure IV.11d) and optimal operating points were found to calculate sensitivities (Figure IV.11a) and specificities (Figure IV.11b). Additionally, the agreements with the gold standard (Figure IV.11c) using the average parameters of the learning set, were also compared.

Statistical tests evidenced a significantly superior performance of LR over the three thresholding versions regarding sensitivity, 96.11% ± 1.52 (mean ± std. dev) against 91.43% ± 3.43 in Mm, the best thresholding representative. The differences in AUCs were also relevant, as LR obtained a remarkable mean value (0.990 ± 0.007), significantly greater than Mm's (0.973 ± 0.013). The other performance measures, although non significantly different, were slightly greater in LR, as indicated by the mean of Sp (95.94% ± 3.10) or AR (93.2%± 4.62) against Mm (93.64% ± 6.47 and 87.90% ± 4.39, respectively). Note that the agreement rates obtained here by LOOCV decreased with respect those presented in last subsection, biased due to the inclusion of all datasets to find the optimal detection parameters.

Given that the best performances correspond to the LR detector, this method will be used to classify bursts automatically and compare the results with rater's criterion.

4.4 Comparison of visual and automatic detections

To quantify how different are manual and automatic labeled bursts, we computed the agreement between the three raters' marks and the detections given by the LR detector (see Table IV.4). It is noteworthy saying that two of the three possible comparisons (LR-B and LR-C) obtained higher percentages than inter-rater agreements. Only LR-A pair, with 70.5% accordance, was below AR_{AB} (76.5%), AR_{AC} (77.8%) and AR_{BC} (81.9%).

Furthermore, to evaluate the characteristics of the bursts detected by each rater and the LR algorithm, we computed the following measures, referred to as discontinuity parameters:

- Number of bursts

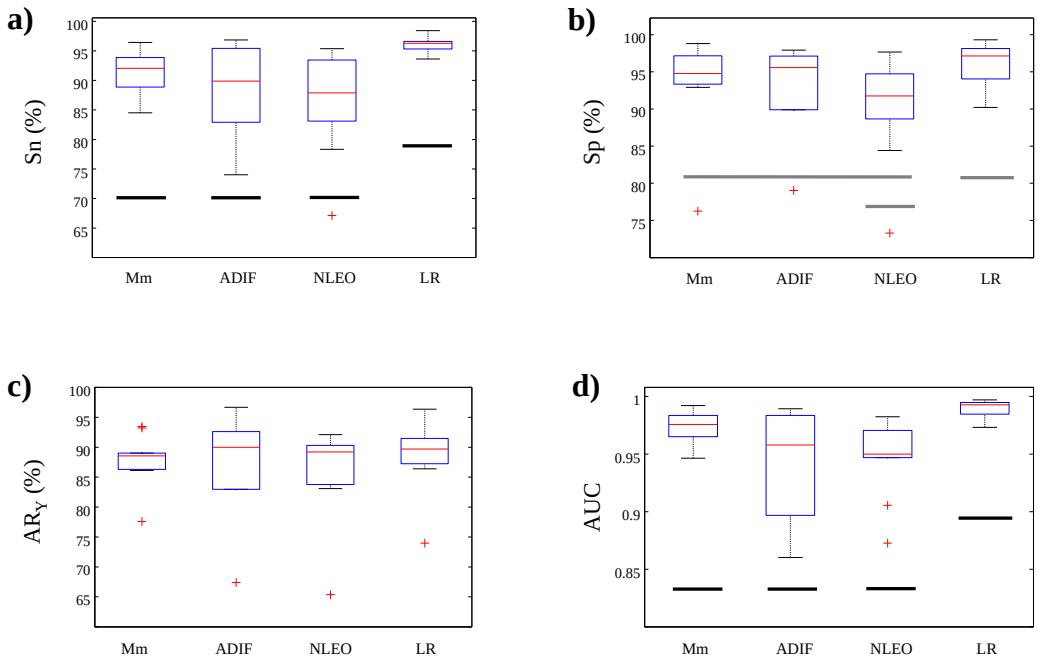

Figure IV.11 – Performance measures of LOOCV for simple thresholding detections (Mm, ADIF, NLEO) and logistic regression (LV). Horizontal lines of different heights below boxes represent significantly different rates running a Mann-Whitney U test (black lines $p < 0.01$ and grey lines $p < 0.05$) among different classifiers. Boxes having horizontal lines of the same height are not statistically different.

Num.	AR_{LR-A}	AR_{LR-B}	AR_{LR-C}
1	82.1	89.2	90.0
2	63.7	83.7	78.4
3	74.9	87.5	77.9
4	72.1	90.0	76.1
5	77.7	83.2	86.8
6	61.7	67.8	73.0
7	71.6	87.8	86.4
8	67.6	71.7	78.9
9	59.6	88.5	81.0
10	74.1	78.9	83.8
Mean	**70.5**	**82.8**	**81.2**

Table IV.4 – Pair-wise agreements (%) between the three rater's marks and LR detections. The mean values for AR_{LR-B} and AR_{LR-B} surpass the inter-rater agreements described in Table IV.2.

- Mean duration of bursts
- Percentage of bursts.

These measures are relative to the length of the records, five minutes in the ten examples of the learning set.

As can be observed in Table IV.5, the values corresponding to automated detections are not far from those defined by the raters.

Rater	Number of bursts	Mean burst duration (s)	Burst %
A	27.5 ± 4.12	7.31 ± 2.46	64.41 ± 13.4
B	25.6 ± 3.86	5.18 ± 1.71	43.46 ± 13.0
C	33.5 ± 6.96	4.19 ± 0.83	45.69 ± 7.60
LR	34.5 ± 5.25	4.62 ± 0.77	54.13 ± 15.5

Table IV.5 – Characteristics of bursts according the different raters (A, B, C) and automatic detection by logistic regression (LR). A statistical analysis is shown in Table IV.6.

In effect, a statistical analysis performed by a Mann-Whitney U test revealed significant differences (see Table IV.6) both among the raters and pairing raters with automatic detections. For example, it can be stated that differences in classifications between rater A and rater B are in the same level than differences between rater A and LR. In view of these findings, our LR-based detector can be applied to the entire database for subsequent analyses, as it will be done in next section.

	Number of bursts	Mean burst duration (s)	Bursts %
A vs. B	0.289	**0.045**	**0.005**
A vs. C	0.088	**0.002**	**0.007**
B vs. C	**0.011**	0.241	0.473
LR vs. A	**0.010**	**0.005**	0.089
LR vs. B	**0.002**	0.970	0.241
LR vs. C	0.879	0.212	0.241

Table IV.6 – p-values from a Mann-Whitney U test comparing the different burst characteristics according the different raters and automatic detection by LR. In bold, variables having significant differences ($p < 0.05$).

5 Clinical application

So far, specific pre-processing and detection tools adapted for immature and noisy EEGs have been developed. In this section, they will be exploited in all VACCIN patients in order to investigate clinical applications from the analysis of brain activity, including EEG signals and the manually scored hypnograms. The availability of information about the age and weight of the patients before and after vaccination allows to line up the possible applications:

1. To assess the post-immunization effects. EEGs and sleep behavior before immunization (A records) will be confronted to post-immunization information.
2. The analysis of maturation. EEGs and sleep behavior will be studied according age and weight criteria.

Both aspects will be assessed basically by statistical and spectral methods.

5.1 Post-immunization study

As introduced in Chapter II, the VACCIN database was motivated by the work of Mialet-Marty et al. [15] to study the influence of the first immunization on cardio-respiratory stability in very preterm infants. The objectives were to describe if the immunization induced changes in heart and respiratory rate variability and to test a potential association between pre-immunization profiles and post-immunization cardio-respiratory events (apnea-bradycardia). In general, vaccination increased the incidence of events lasting more than ten seconds (bradycardia-desaturation) with mild changes in HRV and no changes in RV. It was also concluded that infants with a specific profile before the administration of the vaccine increased the risk of cardio-respiratory events. In particular, those having a sympathetic predominance in the heart rate control (high LF/HF ratio), a decreased HRV entropy and a more immature respiratory rhythm control (evidenced by more frequent bradycardia-desaturation episodes).

Keeping these findings in mind, the goal of the following analysis is to find eventual alterations in sleep and EEG patterns due to the immunization.

Analysis of sleep

Before performing the characterization of EEG and burst patterns, we classified the signals according to the sleep states. This division is a key point for subsequent analysis because, as introduced in Chapter I, sleep determines the spectral and temporal organization of the EEG. As argued in the same chapter, indeterminate and active sleep will be merged into the non-quiet sleep stage (NQS).

Hypnograms were available for 29 infants, but only those exhibiting quiet sleep both before and after vaccination (18 from the 29 having hypnograms) were taken into account to find percentages of spent time in each state (see Figure IV.12).

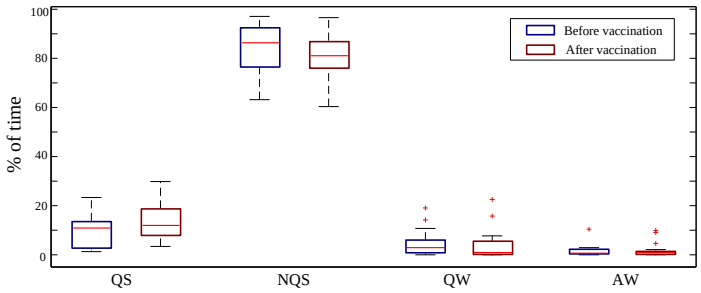

Figure IV.12 – % of spent time in each state obtained from infants having quiet sleep in A and B. Legend: QS, quiet sleep; NQS, non-quiet sleep; QW, quiet wakefulness; AW, active wakefulness. A Mann-Whitney U test did not find statistically significant differences between A and B.

The greatest amount of time corresponded to NQS and only a small portion is reserved to QS and wakefulness states. It must be said that the percentages of QS are remarkably different from those published in the literature. Curzi-Dascalova found approximately the 30% of the sleep time in QS among 35 to 40 weeks' PMA infants [16], and similar results were reported by Hoppenbrouwers in a larger study [17].

The small portion of quiet sleep in our database should not be attributed to the experiment protocol but rather to the strict application of QS classification criteria. In this sense, if a quiet sleep-like epoch presented body movements or eye opening, it was classified as indeterminate sleep. Therefore, we consider these percentages valid for our analysis as far as our definition of quiet sleep is restricted to the characteristics of deep non-REM and slow wave sleep, also referred as stage 4 NREM sleep [18]. This stage, that could be considered as the kernel of quiet sleep, will be the object of interest in our study of maturity as it has been found crucial for the further cognitive development [19].

Another relevant information –even more important than the amount of time of the sleep states [20]– is the number of transitions, i.e. changes from a state to another. Typically, preterm infants have more transitions, which is a sign of immature sleep pattern [21].

We analyzed the temporal organization of our hypnograms by first computing the number of transitions concerning every stage and the probability of transition between them. Transitions were quantified by finding the number of times of each state normalized in one hour, due to the different length of records in A and B (see Figure IV.13). As it can be observed, NQS is the most changing state and the most probable in time, having a central role in sleep organization.

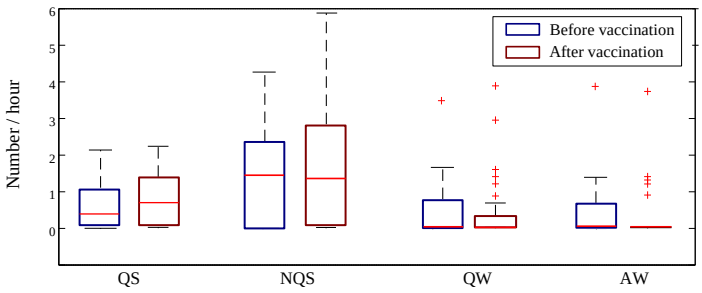

Figure IV.13 – Relative number of times for each state in one hour. A Mann-Whitney U test did not find statistically significant differences between A and B rates.

To find the probabilities of transition between states we define p_{ij}, the probability of changing from state i to state j. Since the current state i has $i-1$ target states, the expression for the individual probability pair was found by:

$$p_{ij} = \frac{k_{ij}}{\sum_{i \neq j}^{N} k_{ij}} \tag{IV.25}$$

where N is the number of different states (QS, NQS, QW, AW) and k_{ij} is the total number of transitions from the state i at time t to state j at time $t+1$.

Figure IV.14 contains the tables of probabilities concerning the 18 infants showing QS both in A and B. Both tables confirm that the deepest sleep (QS) is always preceded by REM sleep [16, 22], included into NQS states.

A statistical analysis did not found any significant differences between the most important transition probabilities (QS to NQS and vice-versa).

A

Earlier state \ Later state	QS	NQS	QW	AW
QS	--	1	0	0
NQS	0.48	--	0.17	0.35
QW	0	0.83	--	0.16
AW	0	0.38	0.62	--

B

Earlier state \ Later state	QS	NQS	QW	AW
QS	--	1	0	0
NQS	0.56	--	0.16	0.28
QW	0.05	0.84	--	0.11
AW	0	0.67	0.33	--

Figure IV.14 – Transition probabilities between states before (A) and after (B) vaccination.

Spectral analysis of EEG

In next lines we perform a basic analysis of the frequency content in the EEGs in A and B. Since the cardiac artifacts and low frequency noise have been effectively removed in previous preprocessing steps, we consider that the here processed signals contains true EEG information.

The EEG bands were defined in the following frequency ranges:

1. δ_1 : Lower delta band, 0.5 - 1.5 Hz, where the slowest bursts are included.
2. δ_2 : Upper delta band, 1.6 - 3 Hz. This interval contains faster rhythms, as delta brushes.
3. θ_1 : Lower theta band, 3.1 - 5 Hz.
4. θ_2 : Upper theta band, 5.1 - 8 Hz.
5. α : Alpha band, 8.1 - 14 Hz, the fastest rhythms of relevance in preterms.

The energies were estimated in 20-second windows using the fast Fourier transform (FFT) by finding the relative content of each frequency band against the whole spectrum. All patients having QS and NQS before and after vaccination are represented in Figure IV.15, and as expected, no statistically differences were found by looking at the entire population in A and B.

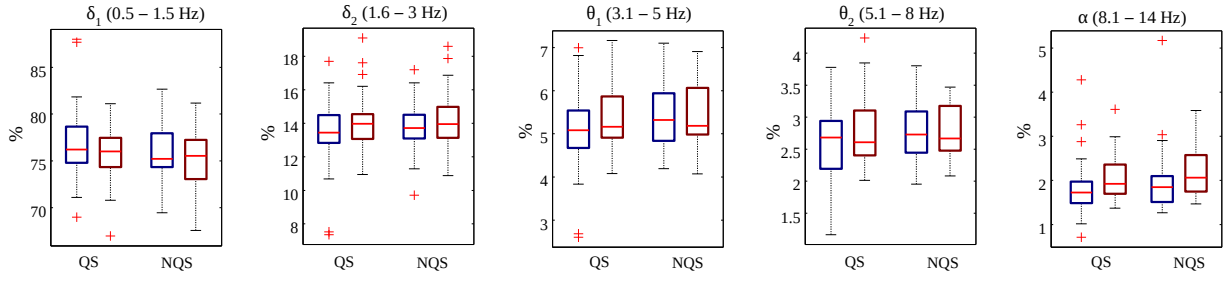

Figure IV.15 – Mean percentages of the different energy bands in EEG. The statistical analysis between percentages before (blue boxes) and after immunization (red boxes) did not show significant differences.

To find individual trends –to check if the infants tend to increase or decrease the energy of some particular band– we employed a paired difference test. Given the small number of samples and non-Gaussian distributions, we employed a Wilkoxon signed rank test.

Energy bands in QS, as well as sleep organization did not show significant differences in the tests, although many tests had a very small number of samples (in B there is a lower incidence of certain sleep states, see Figure IV.13) and consequently a weak statistical power. However, significant differences in the distribution of energies during NQS were found. In Figure IV.16 the relative percentages of the EEG bands are plotted before (left) and after (right) vaccination, evidencing a general decrease of low delta activity and an increase of the upper delta, theta and alpha activity.

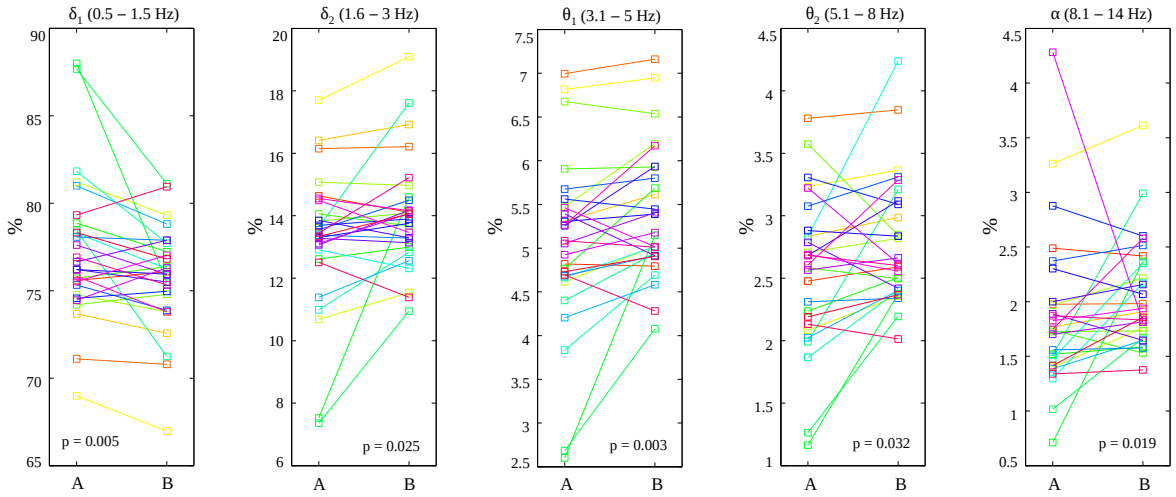

Figure IV.16 – Comparison of energy percentages in the EEG bands. The test examines 25 patients having non-quiet sleep before (A) and after (B) vaccination. p-values from a Wilkoxon signed rank test are plotted in each analyzed band. The decrease of δ_1 and the augmentation of θ_1 components in B are the most significant results from this test.

Characterization of EEG bursts

As mentioned in last section, the automatic detector based in logistic regression has been employed systematically in all EEGs to find the characteristics of inter-burst intervals. Despite the validation of the detector included only QS periods, EEGs in NQS have also been included for classification. In principle, our burst detector should be robust to both sleep states because the main difference between them is that NQS is constituted by longer transients (i.e. longer bursts and shorter IBIs), as described in Chapter I.

Bursts have been characterized through several measures averaged in each EEG pair, Fp1-T3 and Fp2-T4. Besides the discontinuity parameters used in Section 4.4 (number of bursts per minute, mean duration and percentage of bursts) we also included:

- The mean inter-burst duration
- The longest inter-burst interval
- The percentage of synchronization of bursts between both sides of the scalp

Synchronization was computed by comparing the binary sequences yielded by the burst detectors in each EEG pair, so that the percentage of synchronization was the number of equal bits divided by the length of the sequences.

All discontinuity parameters are described by the boxes in Figure IV.17, distinguishing if the EEG comes from QS or NQS, before or after vaccination.

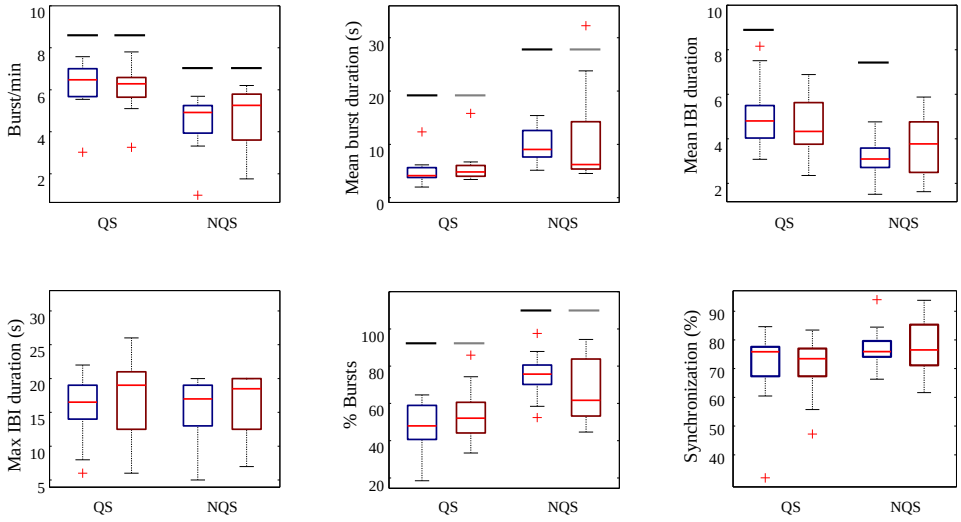

Figure IV.17 – Characteristics of burst and inter-burst intervals in newborns from the VACCIN database (blue and red boxes describe data before and after vaccination, respectively). Horizontal lines of different heights below boxes represent significantly different rates in a Mann-Whitney U Test (black lines $p < 0.005$ and grey lines $p < 0.05$).

The more discontinuous pattern of QS is evident by comparing the number of bursts per minute and the mean duration of IBIs, significantly greater than in NQS. As expected, these results are along the same lines of previously published works, for instance [23, 24]. A more detailed analysis, considering some maturative aspects of the infants, is described in next section.

Discussion

Examining the tests proposed here, it can be affirmed that immunization has no relevant –or, at least, noticeable– effects on the organization of sleep, EEG energy distribution and burst patterns. However, the decrease of δ_1 (plus increase of θ_1) after vaccination should not be underestimated, because the previous study found significant cardio-respiratory events.

In spite of the limitation of our study (the inclusion of 18 patients from 31), the relationship between an increased incidence of bradycardia and desaturation and the energy redistribution in low EEG bands might be plausible.

5.2 Study of maturation

Essentially, EEG maturation is characterized by a decrease in the component energy of the delta band during burst activity in both quiet sleep and active sleep [25]. We verified whether this trend is accomplished only in A recordings for two main reasons: to avoid possible bias due to the smaller number of patients having quiet sleep in B and to exclude possible effects dues to vaccination.

The maturative variables available (PMA, PNA, current weight and weight gain from birth) were confronted to EEG features:

- The relative content of energy in bands δ to α.
- Discontinuity parameters: Duration and synchronization of burst/inter-burst activity.
- Sleep organization: Number of transitions per hour and percentages of time for each state.

Given the limited number of records in the VACCIN database, we divided the cohort in two set of groups according the median of the maturative variables.

G_1: The more immature groups, i.e, the half of the cohort with the smallest PMA (G_1^{PMA}), PNA (G_1^{PNA}), weight (G_1^{W}) and gain of weight ($G_1^{\Delta W}$).

G_2: The less immature groups, constituted by the rest of infants with the greatest maturative variables (G_2^{PMA}, G_2^{PNA}, G_2^{W} $G_2^{\Delta W}$).

This allowed to confront the most immature infants against the less immature infants for each one of the variables, having in total four paired analysis (see Table IV.7 to see the boundaries of each group).

	PMA (weeks)	PNA (days)	W (Kg)	ΔW (Kg)
G_1	36.4 – 38.1	50 – 72	1.32 – 2.28	0.90 – 1.26
G_2	38.1 – 41.2	73 – 105	2.28 – 3.23	1.26 – 2.01

Table IV.7 – Description of the assorted groups to assess maturation: PMA (postmenstrual age), PNA (postnatal age), W (Weight at present day) and ΔW (gain of weight from birth).

Because PMA produced in general more meaningful tests, the here depicted graphs refer to this variable. More complete information about the analysis of all variables is detailed in Appendix B.

Sleep analysis

Observing some sleep organization variables in Figure IV.18, there are no significant differences by analyzing the assortment G_1/G_2 (see also Table B.1 in the Appendix). Slight decreases in the number of transitions per hour and in the portion of NQS are observed in G_2, suggesting an increase of maturation. But as mentioned before, the relative percentages of sleep in our scorings do not correspond to those documented in the literature and the results of this analysis should be interpreted with caution.

Spectral analysis of EEG

On the other hand, some differences can be found regarding the energy bands (Figure IV.19). Clearly, EEG energy is concentrated in the lowest frequency bands, with δ_1 representing about the 75% (in full-terms this portion is around 45% [26]). Regardless of the sleep state, the percentage of δ_1 energy is greater in the youngest infants, whereas older infants have a greater percentage of energy in δ_2 and θ bands.

As reported by numerous studies, this displacement of the energies to higher frequency bands is expected as the normal premature infants increase their postmenstrual age [25, 27, 28]. The other variables (see tables B.2 and B.2 in Appendix C) did not evidence statistically significant differences.

Although higher frequencies could be characterized because they appear superimposed in slower waves progressively with increasing ages [29, 30], in this work we only focus on the temporal features of the slow delta bursts.

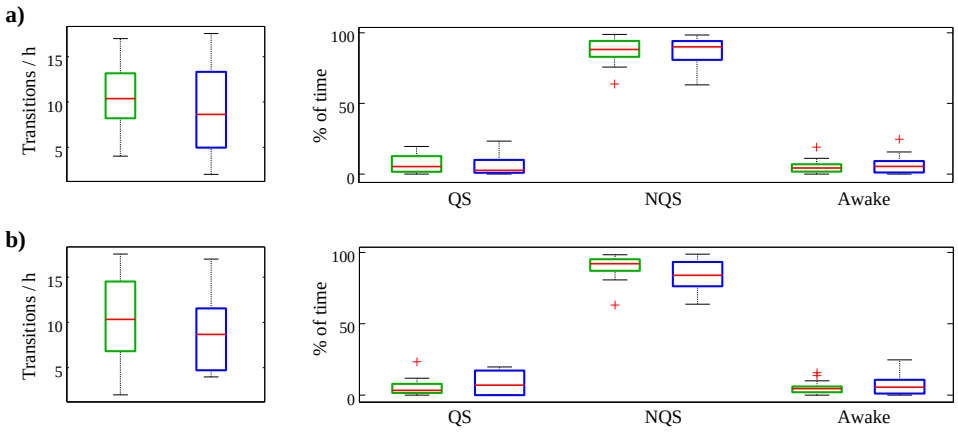

Figure IV.18 – Number of transitions per hour and percentages of sleep/wake states in both maturation groups (green boxes, G_1; blue boxes, G_2). In a) the groups concerned PMA and in b) the weight.

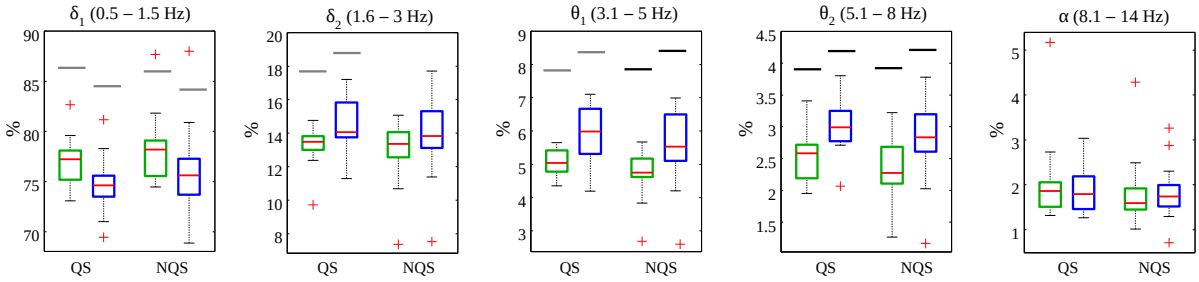

Figure IV.19 – Distribution of the energy bands within two maturation groups (green boxes, G_1^{PMA}; blue boxes, G_2^{PMA}). Boxes having lines of different height have statistically significant differences (grey lines, $p<0.05$; black lines, $p<0.01$).

Characterization of EEG bursts

With regard to discontinuity parameters, no significant differences were found comparing G_1 and G_2 although medians of the last group showed a small feeble tendencies to more mature patterns, as can be seen in Figure IV.20. Confronting our results to reference data for IBIs is not straightforward. The vast work available in the literature provides statistics describing different maturational levels and applying different definitions of bursts [1, 2, 28, 31], therefore considerable differences can be found. In any case, summarizing reference data, it is widely accepted that normal preterm infants between 35 and 39 weeks PMA rarely exhibit IBIs exceeding 20 seconds, and their mean durations range 4 to 10 seconds depending on the sleep state. Independently of the electroencephalographic device and standard used, authors also agree that IBIs reduce their durations with increasing PMA.

Another interesting element to asses maturation is to quantify the degree of bursts between hemispheres (inter-hemispheric synchrony). In healthy infants, synchronization increases with maturation during preterm ages until 40-43 weeks PMA, when bursts appear simultaneously in both sides of the brain [29, 32]. The results shown in Figure IV.20 indicate that our percentages are not far from those previously published, about 70-80% of inter-hemispheric synchrony at 37-39 weeks and about 85% at 39-40 weeks PMA.

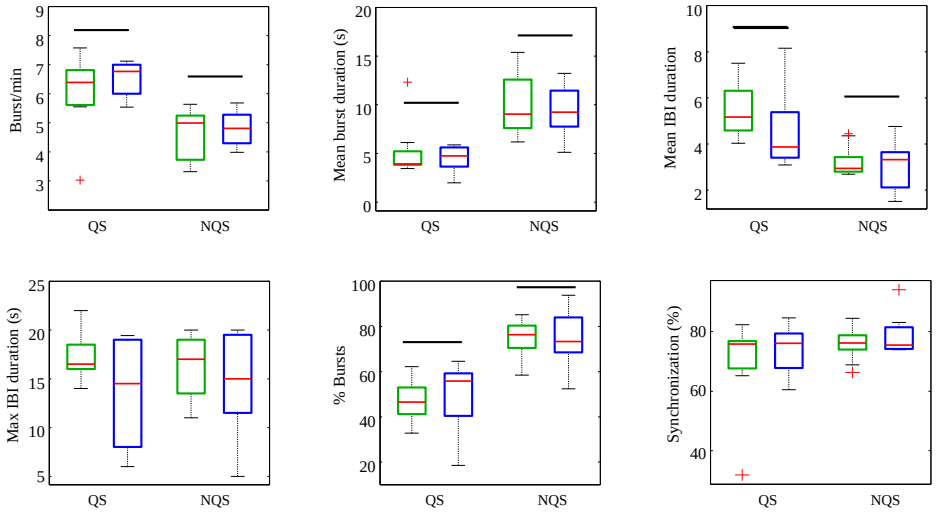

Figure IV.20 – Description of EEG discontinuity parameters before vaccination, comparing G_1^{PMA} (green boxes) and G_2^{PMA} (blue boxes). Differences are statistically significant if horizontal lines have different heights ($p<0.05$).

6 Conclusion

In summary, the classifiers based on logistic regression over-performed the threshold detectors, specially regarding sensibilities and AUCs. Concerning the three analyzed features in threshold detection, we found that the Mm operator had slightly better sensibilities and AUCs than NLEO and ADIF, although there were not statistically significant. This finding is in contradiction with previous works demonstrating the superior performance of NLEO for burst detection, a quality that introduced it to commercial software [4].

Similar works using thresholding detection in neonatal EEG found a 90% average accuracy [7] and later a 97% specificity and 95% sensitivity [8] compared to the gold standard. The inferior rates of our counterpart detector (93% specificity and 91% sensitivity in average) may be explained by the more complex EEG patterns in older preterm newborns. However, using the logistic regression classifier, performances were significantly enhanced (96% – 96% in average) and yielded the same rates as the mentioned work, demonstrating that the inclusion of several EEG features is advantageous.

Concerning the relatively low inter-rater agreements in our experiment, two main reasons can be argued:

1. The difficulty of interpreting the start and end of SATs in discontinuous patterns and *tracé alternant* in preterm infants from 36 to 39 weeks PMA. The cohort in the previously cited works was between 23 and 30 weeks PMA, a period of prematurity in which the background has still a low activity.

2. The design of the protocol. Probably the presence of two EEG leads might influence the rater to take intermediate decisions if some asymmetry or asynchrony is present.

Regardless of the gold standard quality –likely to be improved– is should be pointed out that performing automatic detections using the LR classifiers can be considered as reliable as clinicians' visual marks, besides being more repeatable (there is not self-discrepancy), fast and cost-effective.

This chapter presented also a couple of clinical applications in which the EEG from the NICU can be exploited, verifying that the use of a reduced EEG montage with two fronto-lateral leads can be useful to assess maturation and might add proofs in the study of post-vaccinal effects.

The spectral analysis of the EEG showed a significant change in the distribution of the energy bands. In effect, the group having more aged infants, a priory "more mature", revealed a lower δ_1 content whereas higher energy bands were more active in comparison with less aged infants. The study of discontinuity variables, however, did not provide significant differences although we found a coherent evolution of the IBIs, prone to shorten with maturation because of the more continuous EEG patterns appearing from 38-39 weeks PMA. It has to be underlined that the prior knowledge of the sleep state is necessary to correctly match the EEG features to maturative aspects. Actually, frontal electrodes account in NQS for some ocular movements, which may be wrongly interpreted as slow bursts. Therefore, in our experiment longer burst periods in NQS cannot be attributed only to the more continuous EEG activity but also to the eye motion, a fact that might explain why IBIs are slightly shorter than those referenced by neonatologists.

Besides the above limitation, we also remain cautious because differences found by splitting the cohort in "more immature" and "less immature" groups comprise different subjects, whereas a rigorous study of maturation should be corroborated with a horizontal database.

On the other hand, the study of the frequency bands regarding the post-immunization effects showed a diminution of the lower delta band in non-quiet sleep. This finding is not sufficient to state that vaccination perturbs significantly the cerebral activity (the discontinuity parameters and the sleep did not evidence any changes), but it could be a start point to explore more extensively, in further work, the possible effects of vaccination on the central nervous system, besides the already know effects to the autonomic nervous system.

Bibliography

[1] J. S. Hahn, H. Monyer, and B. R. Tharp. *Interburst interval measurements in the EEGs of premature infants with normal neurological outcome.* Electroencephalography and clinical neurophysiology **73**(5), 410–418 November (1989). PMID: 2479519.

[2] M. Hayakawa, A. Okumura, F. Hayakawa, K. Watanabe, M. Ohshiro, Y. Kato, R. Takahashi, and N. Tauchi. *Background electroencephalographic (EEG) activities of very preterm infants born at less than 27 weeks gestation: a study on the degree of continuity.* Archives of Disease in Childhood. Fetal and Neonatal Edition **84**(3), F163–F167 May (2001).

[3] J. Kaiser. On a simple algorithm to calculate the 'energy' of a signal. In *, 1990 International Conference on Acoustics, Speech, and Signal Processing, 1990. ICASSP-90*, pages 381–384 vol.1 (1990).

[4] M. Särkelä, S. Mustola, T. Seppänen, M. Koskinen, P. Lepola, K. Suominen, T. Juvonen, H. Tolvanen-Laakso, and V. Jäntti. *Automatic analysis and monitoring of burst suppression in anesthesia.* Journal of Clinical Monitoring and Computing **17**(2), 125–134 February (2002).

[5] J. Löfhede, N. Löfgren, M. Thordstein, A. Flisberg, I. Kjellmer, and K. Lindecrantz. *Classification of burst and suppression in the neonatal electroencephalogram.* Journal of Neural Engineering **5**(4), 402 December (2008).

[6] K. Pfurtscheller, G. R. Müller-Putz, B. Urlesberger, W. Müller, and G. Pfurtscheller. *Relationship between slow-wave EEG bursts and heart rate changes in preterm infants.* Neuroscience letters **385**(2), 126–130 September (2005).

[7] K. Palmu, S. Wikström, E. Hippeläinen, G. Boylan, L. Hellström-Westas, and S. Vanhatalo. *Detection of 'EEG bursts' in the early preterm EEG: visual vs. automated detection.* Clinical Neurophysiology **121**(7), 1015–1022 July (2010).

[8] K. Palmu, N. Stevenson, S. Wikström, L. Hellström-Westas, S. Vanhatalo, and J. M. Palva. *Optimization of an NLEO-based algorithm for automated detection of spontaneous activity transients in early preterm EEG.* Physiological measurement **31**(11), N85–93 November (2010).

[9] A. Varri. *Digital Processing of the EEG in Epilepsy.* Thèse de Doctorat, Tampere University of Technology, Finland (1988).

[10] R. I. Jennrich and S. M. Robinson. *A newton-raphson algorithm for maximum likelihood factor analysis.* Psychometrika **34**(1), 111–123 March (1969).

[11] J. Fox. *Applied Regression Analysis, Linear Models, and Related Methods.* SAGE (1997).

[12] D. Selton, M. Andre, and J. Hascoët. *Normal EEG in very premature infants: reference criteria.* Clinical Neurophysiology **111**(12), 2116–2124 December (2000).

[13] R. G. Norman and M. A. Scott. *Measurement of inter-rater agreement for transient events using monte carlo sampled permutations.* Statistics in medicine **26**(4), 931–942 February (2007).

[14] J. L. Fleiss. *Measuring nominal scale agreement among many raters.* Psychological Bulletin **76**(5), 378–382 (1971).

[15] T. Mialet-Marty, A. Beuchée, W. Ben Jmaa, N. N'Guyen, X. Navarro, F. Porée, A. M. Nuyt, and P. Pladys. *Possible predictors of cardiorespiratory events after immunization in preterm neonates.* Neonatology **104**(2), 151–155 (2013).

[16] L. Curzi-Dascalova, P. Peirano, and F. Morel-Kahn. *Development of sleep states in normal premature and full-term newborns.* Developmental psychobiology **21**(5), 431–444 July (1988).

[17] T. Hoppenbrouwers, J. E. Hodgman, D. Rybine, G. Fabrikant, M. Corwin, D. Crowell, and D. E. Weese-Mayer. *Sleep architecture in term and preterm infants beyond the neonatal period: the influence of gestational age, steroids, and ventilatory support.* Sleep **28**(11), 1428–1436 November (2005).

[18] T. Hori, Y. Sugita, E. Koga, S. Shirakawa, K. Inoue, S. Uchida, H. Kuwahara, M. Kousaka, T. Kobayashi, Y. Tsuji, M. Terashima, K. Fukuda, and N. Fukuda. *Proposed supplements and amendments to 'A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects', the rechtschaffen & kales (1968) standard.* Psychiatry and clinical neurosciences **55**(3), 305–310 June (2001).

[19] S. N. Graven and J. V. Browne. *Sleep and brain development: The critical role of sleep in fetal and early neonatal brain development.* Newborn and Infant Nursing Reviews **8**(4), 173–179 December (2008).

[20] P. Levitt. *Structural and functional maturation of the developing primate brain.* The Journal of Pediatrics **143**(4, Supplement), 35–45 October (2003).

[21] D. Holditch-Davis, M. Scher, T. Schwartz, and D. Hudson-Barr. *Sleeping and waking state development in preterm infants.* Early Human Development **80**(1), 43–64 October (2004).

[22] L. Curzi-Dascalova, P. Peirano, L. Silvestri, and G. Korn. *Organisation du sommeil des nouveau-nes prematures normaux. etude polygraphique.* Revue d'ElectroencÃĬphalographie et de Neurophysiologie Clinique **15**(3), 237–242 December (1985).

[23] J. Parmelee, A H, F. J. Schulte, Y. Akiyama, W. H. Wenner, M. A. Schultz, and E. Stern. *Maturation of EEG activity during sleep in premature infants.* Electroencephalography and clinical neurophysiology **24**(4), 319–329 April (1968).

[24] J. R. C. Conde, A. L. R. d. Hoyos, E. D. Martínez, C. G. Campo, A. M. Pérez, and A. A. Hernández Borges. *Extrauterine life duration and ontogenic EEG parameters in preterm newborns with and without major ultrasound brain lesions.* Clinical Neurophysiology **116**(12), 2796–2809 December (2005).

[25] H. Sawaguchi, T. Ogawa, T. Takano, and K. Sato. *Developmental changes in electroencephalogram for term and preterm infants using an autoregressive model.* Acta paediatrica Japonica; Overseas edition **38**(6), 580–589 December (1996).

[26] K. Paul, V. Krajca, Z. Roth, J. Melichar, and S. Petranek. *Comparison of quantitative EEG characteristics of quiet and active sleep in newborns.* Sleep Medicine **4**(6), 543–552 November (2003).

[27] F. Torres and C. Anderson. *The normal EEG of the human newborn.* Journal of clinical neurophysiology: official publication of the American Electroencephalographic Society **2**(2), 89–103 April (1985).

[28] M.-F. Vecchierini, M. André, and A. d'Allest. *Normal EEG of premature infants born between 24 and 30 weeks gestational age: Terminology, definitions and maturation aspects.* Neurophysiologie Clinique/Clinical Neurophysiology **37**(5), 311–323 (2007).

[29] G. L. Holmes and C. T. Lombroso. *Prognostic value of background patterns in the neonatal EEG.* Journal of clinical neurophysiology: official publication of the American Electroencephalographic Society **10**(3), 323–352 July (1993).

[30] K. Watanabe, S. Miyazaki, K. Hara, and S. Hakamada. *Behavioral state cycles, background EEGs and prognosis of newborns with perinatal hypoxia.* Electroencephalography and clinical neurophysiology **49**(5-6), 618–625 September (1980).

[31] E. Biagioni, L. Bartalena, A. Boldrini, G. Cioni, S. Giancola, and A. E. Ipata. *Background EEG activity in preterm infants: correlation of outcome with selected maturational features.* Electroencephalography and clinical neurophysiology **91**(3), 154–162 September (1994).

[32] C. T. Lombroso. *Quantified electrographic scales on 10 pre-term healthy newborns followed up to 40-43 weeks of conceptional age by serial polygraphic recordings.* Electroencephalography and Clinical Neurophysiology **46**(4), 460–474 April (1979).

Part III

Analysis of respiratory activity

Chapter V

Extraction of respiratory variability signals

The second part of this work is devoted to process and analyze, by means of linear and nonlinear methods, respiratory signals acquired in NICUs. As these signals, obtained by abdominal strain gauges, are uncalibrated and cannot be used to study the air flow, they will be exploited as time-series expressing the shape of breaths or as variability series describing the respiratory rhythm. The latter are less informative signals -only respiratory rate is carried- but hold interesting properties analyzed extensively later on.

This chapter describes the procedure to obtain RV signals from noisy breathing signals, including the preprocessing before the detection of respiratory cycles, of crucial importance to obtain quality RV signals. The whole process, outlined in Figure V.1 consists first in rejecting artifacted epochs involving gross body movements in raw signals by an automatic detector based on logistic regression, and secondly filtering the signal in two steps (a band-pass and a smoothing filter) leading to clean breathing signals. Finally, denoised data is processed to find the breath intervals using a cycle detector.

Figure V.1 – Block diagram proposed to denoise the raw breathing signals and obtain variability series.

After obtaining the clean respiratory series, they are analyzed using basic linear methods to investigate their clinical utility. A first analysis, taking advantage of the available information about age and weight, assesses the infant's maturity using breathing signals from the PHYSIDEV and VACCIN databases. Using a sub-cohort pairing infected and healthy infants, the second analysis aims at finding differences between them from the study of several breathing parameters.

1 Preprocessing of breathing

1.1 Artifact rejection

A common artifact detection criteria is to study the statistical distribution of the energy or root-mean square (RMS) of the breathing signals because in general, gross body movements induce higher amplitudes on the strain

gauges. Motto et al. [1] applied this feature in 45 weeks PMA full term infants breathing traces (abdominal and thoracic), employing a thresholding detector optimized by a Neyman-Pearson approach [2] that attained 89% for sensitivity and 88% for specificity. However, an energy-based threshold could be in some cases too restrictive due to the effect of deep breaths and impedance changes in the amplitudes.

On the other hand, an artifacted component could account for an unexpected transient event or for a background activity, like muscle activity or noise. Thus, in view of the noisy environments our breathing signals come from, an alternative criterion to detect the artifacts could be to measure the randomness of the traces by means of the entropy, as Mammone et al. [3] did in EEG signals by means of ICA and Renyi's entropy. Nevertheless, the erratic breathing patterns typical in preterm infants (see Figure V.2) could be an inconvenient for an entropy detector, and lead the classifier to false positive detections.

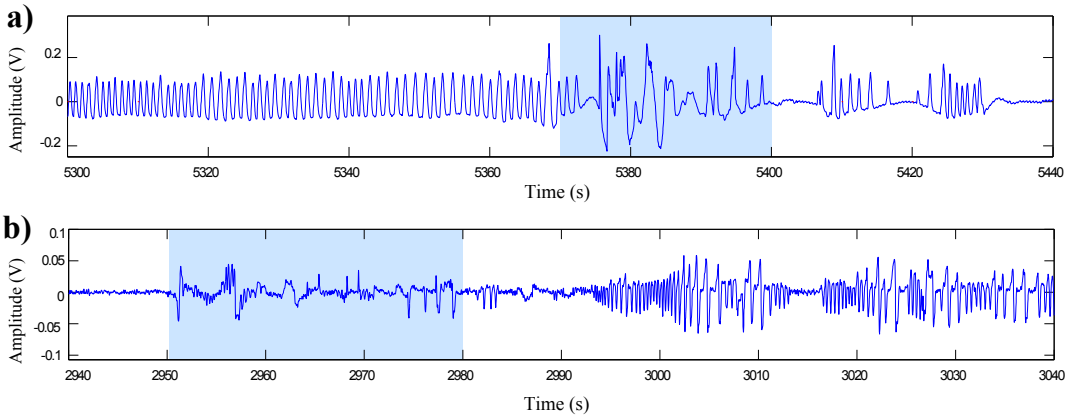

Figure V.2 – Unprocessed respiratory signals with artifacts (shadow zones) marked by a clinician. a) Example of a regular pattern turning into periodic after the artifacts. b) Example of a heavily contaminated signal exhibiting erratic pattern. In view of the similarities between artifacts in regular patterns and erratic patterns, one of the challenges of our detector will be to discriminate correctly the artifacts.

It is clear then, that a binary multi-variable model dealing with several breathing features is necessary. Although we are aware of the numerous existing solutions, such as the SVM, we employ again the classifier based on logistic regression to discriminate between clean and noisy excerpts.

Protocol

As done with the burst detector in Chapter IV, reference labels need to be first created by a clinician. To such purpose, we selected five unprocessed respiratory signals ($N = 5$ examples) exhibiting different breathing patterns, to be marked manually in a visual interface specially designed for the experiment. The selection was constituted by preterm infants (aged 31.0±1.6 weeks and weighting 1061±299 grams) from the PHYSIDEV database, providing 14.35 hours of recordings. The signals, originally sampled to 400 Hz (Fs), were resampled to 64 Hz (Fr) after eliminating the frequency content above $Fr/2$ to avoid aliasing.

The labeling procedure consisted in marking intervals of ten seconds ($W = 10$ s) as clean (class 0) or artifacted (class 1) in a screen where the ECG (if available) was also shown to help the observer to make the decision (see Figure V.3). Other basic options, as changing the scale of the signals and visualize the signals in different time ranges, were also added to the interface functionality.

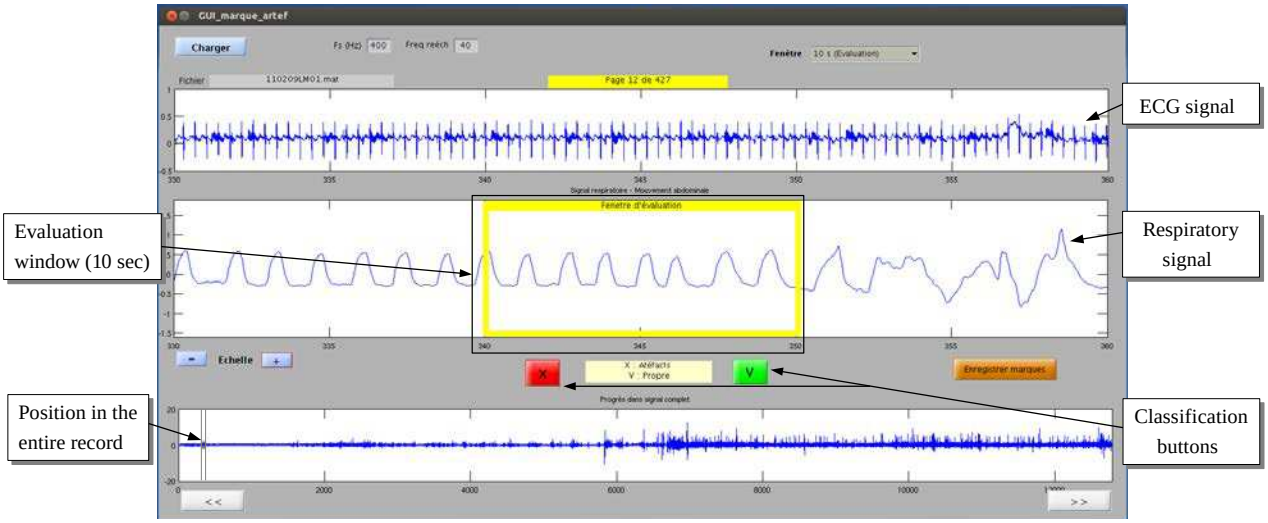

Figure V.3 – Screen capture of the artifact marking interface.

The artifact marking interface produced binary signals, $Y_{i,j}$, $i = 1, \dots N$ and $j = 1, .., n$, where n is the number of 10-second marks in each respiratory signal. The five labeled signals summed up 5167 marks, with the 11.8% of them classified as artifacts according to the rater.

Logistic regression classifier

Knowing from Chapter IV that a detector based on logistic regression provided good results in binary classification problems, the same method has been applied here. z From the raw breathing signal we obtained features based both in time and frequency domains, computed on ten-second excerpts ($WFr = 640$ samples), that we subsequently call $resp$, synchronized with the labels.

The set of instances, $\mathbf{X} \in \mathbb{R}^d$, composed by $d = 14$ features, can be divided in:

- Time-domain: The absolute difference between the maximum and minimum value (Mm), the mean of $resp$ divided by its maximum value ($\overline{resp}/M$), the standard deviation (Sd) and the kurtosis (Kt), the maximum value of the first derivate of the excerpt ($D1$) and its kurtosis ($KD1$), and the root mean squared value, RMS, computed by:

$$RMS = \sqrt{\frac{1}{WFr} \sum_{k=1}^{WFr} resp^2(k)}. \tag{V.1}$$

- Frequency-domain: the kurtosis of the power spectral density ($KPSD$) and the power comprised between the following frequency bands: 0 to 0.25 Hz ($P0\_025$), 0.25 to 0.7 Hz ($P025\_07$), 0.7 to 1.2 Hz ($P07\_12$) and 1.2 to 2 Hz ($P12\_2$). An order-30 autoregressive Burg model was used to estimate the power spectrum. We also computed the main frequency, MF, as:

$$MF = \frac{1}{WFr} \sum_{k=1}^{WFr} |resp(k) - resp(k-1)|. \tag{V.2}$$

- Entropy: The Shannon entropy in the each excerpt, $SEnt$, computed as:

$$SEnt = -\sum_{k=1}^{WFr} resp^2 \log\left(resp^2\right). \tag{V.3}$$

Other features were included to the model in a preliminary test (second derivative properties, number of zero-crossing, different frequency bands, etc.) but they did not improved significantly the model fit. Since mutual correlations occur between numerous measured variables, their number had to be reduced by applying feature selection. As done previously in Chapter IV, a Wald test was performed to retain the most contributive parameters by evaluating the z-statistic of the estimated coefficients. This resulted in reducing the initial 14 variables to seven (detailed in Table V.1).

Dep var →	Intercept	Mm	RMS	MF	$KPSD$	$P0\_025$	$P12\_2$	$SEnt$
$\widehat{w}$	1.865	-0.743	0.152	-0.009	0.968	0.030	-5.863	-7.176
SE	0.579	0.144	0.069	0.001	0.262	0.008	0.652	0.762
z	3.220	5.184	2.192	5.856	3.687	3.683	8.987	9.407
p	< 0.01	< 0.001	0.028	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Table V.1 – Results from fitting the LR model to the artifact detector. The estimated coefficients ($\widehat{w}$) corresponding to the seven selected features and the intercept, and the standard errors (SE) of the MLE are detailed. We also provide the z-statistics and p-values of the Wald test. The remaining features did not contribute significantly to the model since their p-values exceeded 0.05.

Performance of automatic artifact rejection

To see the efficacy of artifact detectors, the ROC curves (see Figure V.4a) were plotted using leave-one out cross-validation, as done in Section 4.3. In Figure V.4b the sensibilities and sensibilities are obtained using a logit cut-off value of 0.5, but to find a cost-effective point, we represented the mean of both measures as a function of c (Figure V.4c) and selected the intersection point ($c = 0.165$, $Sn = 86\%$, $Sp = 86\%$).

The less optimal values obtained with this detector (86%/86% against 89%/88% in full-terms [1]) may be due to the complexity of the classification problem in more immature breathing patterns. Indeed, the less performing ROC curve (in green) corresponds to a patient having patterns predominantly erratic (a sample can be seen in Figure V.2b) whereas the best one contains more regular patterns.

1.2 Band-pass and smoothing filters

As introduced in Chapter I, respiratory signals in preterm infants breathing spontaneously have several peculiarities. The notable influence of the technical artifacts present at NICU, changes in body position and physiological events as apnea, sights and swallows, increase the difficulty of determining appropriate filters and performing effective breath detections.

Hence, once the artifacts due to gross movements are rejected, the next step is to remove the base line and the high frequency noise. This can be performed by a simple band-pass filer, a fourth order Butterworth whose bands were set to the range 0.5–20 Hz. The higher frequency was chosen to 20 Hz because lower values may result in filtering excessively the peaks and loosing information about the shape of the breathing trace.

The resulting signal is processed by the next step, but also used to estimate the main frequency in the power

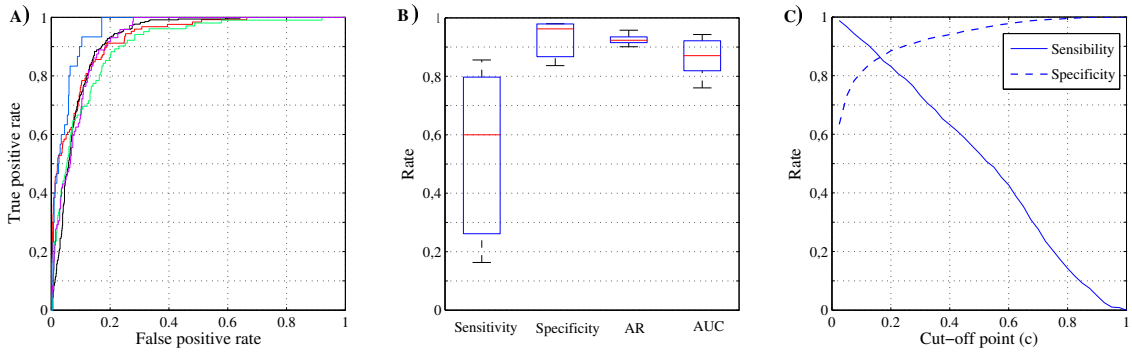

Figure V.4 – a) ROC curves from LOOCV. b) Boxplots of some measures of performance: sensitivity (Sn) and specificity (Sp) based on the default cut-off value ($c = 0.5$), agreement rate (AR) and area under curve (AUC). c) Mean of the sensitivity/specificity pair at several cut-off values.

spectrum, M_f. This information, necessary to tune the second filter, was estimated by means of a 30th order Burg autoregressive model.

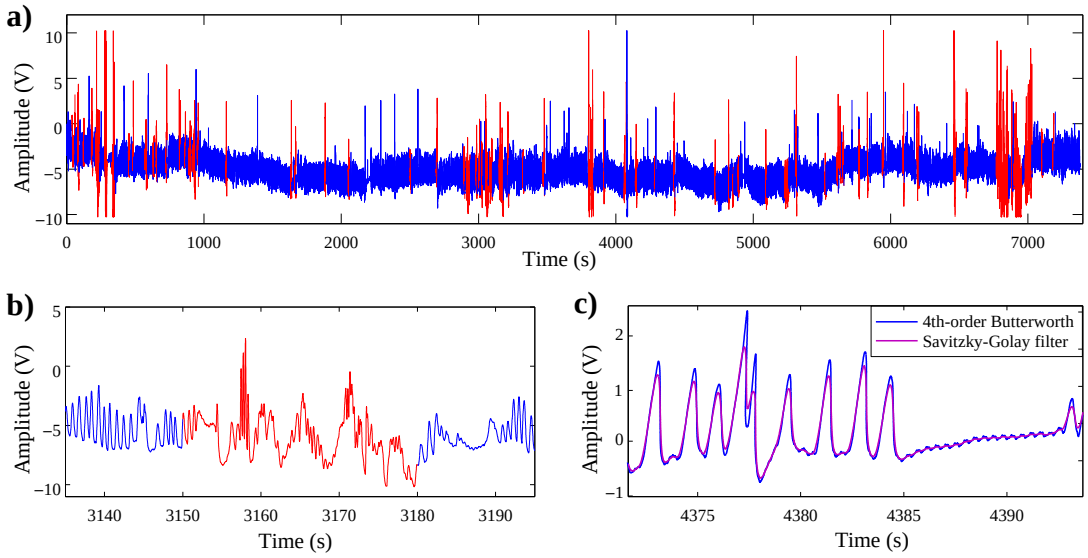

Figure V.5 – a) Example of artifact detection (traces in red) in a patient not used in the training set. As the baseline evidences, no filtering is still performed at this step. b) Detail of detected artifacts. c) Detail of the signal, after being processed by the two filters. Note that the small ripple disappears after the SG filter.

The second filter is used to smooth without distortion the signal so that cycle detection can be performed more effectively. Smoothing is necessary to reduce small ripple due, for instance, to cardiac artifacts and other noise unfiltered in the previous step. This can be achieved by Savitzky-Golay (SG) filters [4], specially recommended because they preserve the width and height of peaks of the original signal, which are usually flattened by classical moving average or FIR filters.

In SG filters, data is smoothed by applying local least-squares polynomial approximation. Applied to oversampled signals (our trace is sampled at 400 or 512 Hz whereas respiratory oscillations are mainly about

1 Hz) corrupted by noise, SG filters match appropriately the original waveform with fitted polynomial slopes reducing high-frequency noise [5].

Two parameters need to be chosen to apply this filter: the polynomial order (N) and the length of frame (M). To perform a good smoothing, it should be considered that:

- $N = 2$ takes into account the curvature of the signals and $N = 3$ the inflection points. Larger values are unnecessary and $N = 0$ produces the same effect than a moving average filter.
- M needs to be larger than N. If $M \simeq N$ the interpolation uses almost as many points as those contained in the original signal and no smoothing is produced.

After some tests, we found that an optimal filtering was obtained by fixing $N = 2$ and $M = \frac{Fs}{2M_f}$.

An example of a raw signal, first processed by the artifact detector and then by the filters, is shown in Figure V.5.

2 Respiratory cycle detection

Once the breathing signal has been cleaned and smoothed, the respiratory cycles are recognized by an automatic detector, providing the respiratory variability series. In this section, we describe the detection algorithm and examine its performance by means of manually-validated reference data.

2.1 Detection algorithm

The proposed breath-recognition algorithm takes advantage of the sinusoidal shape of the ventilatory profile to detect the beginning and end of individual breaths. Since a minimum corresponds to the minimal wall distension it is directly related to the end of exhalation, so it is considered as the start-point of a respiratory cycle. Likewise, a maximum in the respiratory trace is reached when the lungs contain the tidal volume, so the time elapsed between minima and maxima defines the inspiration time (t_i), and conversely, the time between maxima and minima yields the exhalation time, t_e (see Figure V.6).

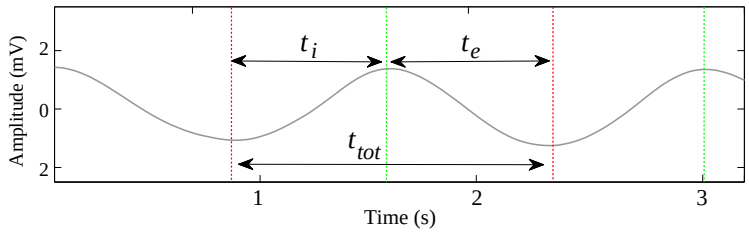

Figure V.6 – Illustration of t_i, t_e and t_{tot} times in a respiratory trace. Red and green lines constitute detected minima and maxima, respectively.

Detection algorithms exploiting large slopes or high frequent content in signals (for example, the QRS complex in ECG) employ transformations such as the signal derivative or the dyadic wavelet transform, producing a feature in which peaks can be easily detected by thresholding [6, 7]. However, breathing signals are sinusoidal-like, non-impulsive signals where most of the spectral power is located in a lower frequency range, typically from 0.5

to 2 Hz in newborns. Consequently, a simple extrema detector seems to be a better solution that the mentioned strategies.

Briefly, this detection algorithm finds all extrema, that will be validated only if the relative difference between surrounding peaks exceeds a predefined threshold, th. The pseudo-code is described in the following lines:

```
Inputs
x  := Breathing signal
th := Detection threshold
Outputs
m := Detected minima
M := Detected maxima
Begin
   i_M := Find the indexes of x such that: x[i-1] <= x[i] <= x[i+1]
   M := Find x[i_M] such that x[i_M]-th > 0
   i_m := Find the indexes of x such that: x[i-1] >= x[i] >= x[i+1]
   m := Find x[i_m] such that x[i_m]+th > 0
End.
```

The detection threshold is based on the inter-quartile range (IQR), a measure of dispersion excluding the 25% of extrema, so deep and shallow breaths are not taken into account in the detection. The performance of the IQR is evaluated in next section by adjusting its value with κ, a multiplicative coefficient.

2.2 Validation of automatic marks

For practicality and rapidity, the validation of the cycle detector was in this case different from the previous classifiers. Here, detections were first performed automatically using the parameters described before and then inspected visually.

An interface implemented with Matlab software was designed purposely to correct automatic detections by adding or removing peaks (see Figure V.7). The original, unfiltered signal was visualized in 20-second windows simultaneously with the filtered signal to help the identification of unclear cycles. Wrong detections, mainly due to small ripple or unfiltered noise (false positives) and undetected cycles (false negatives) were accounted by the program, which provided the final statistics about the performance of detection.

Seven patients from the PHYSIDEV database supplied the breathing signals to validate the cycle detector, totalizing 7231 breaths correctly annotated.

2.3 Performance of detection

Following the guidelines proposed by the Association for the Advancement of Medical Instrumentation [8], the parameters used to evaluate the performance of the breath detection algorithm were based on the traditionally-employed measures for heart beat detectors. These were sensitivity, Sn (previously defined in Chapter IV), and positive predictive value, $+P$, the proportion of positive test results that are true positives (rate of correct detections):

$$+P = \frac{TP}{TP + FP} \tag{V.4}$$

The automatic detector was applied to the labeled dataset, varying the coefficient κ from 0.05 to 1 in steps of 0.05. Since the measures of performance considered the number of well detected extrema, a peak was classified as valid if there was a match within the surrounding 10% of the cycle time with a manual label. Otherwise, it

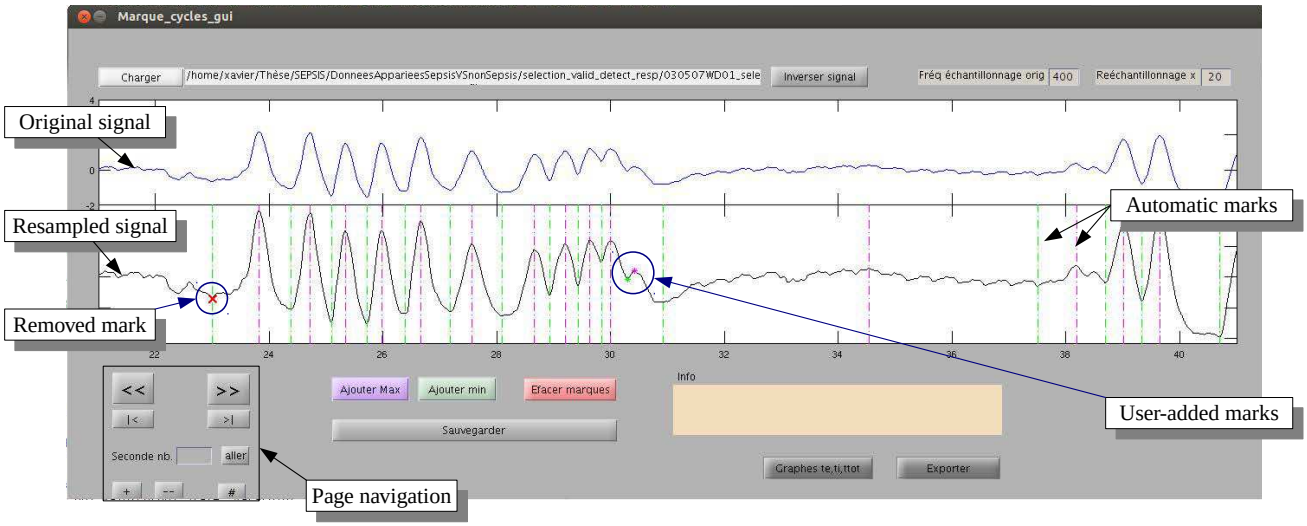

Figure V.7 – Screen capture of the cycle marking interface.

was counted as false positive. Figure V.8 shows the results of the performances for thresholds based on IQR and standard deviation. As it can be observed, the cut-off point of the curves of sensitivity and positive predictability (the best performance) corresponds to $\kappa = 0.463$ in the IQR, yielding $Sn = +P = 97.7\%$.

3 Clinical application

Once the respiratory data is properly cleaned and their corresponding variability series obtained, this section proposes a linear analysis employing standard statistical measures and a basic spectrum analysis. Usually, linear tools constitute the first resource for characterizing and understanding the irregularity of the neonatal breathing patterns. Following the common thread of this thesis, the here-proposed tests are aimed at studying the maturation and finding indicators of infection.

3.1 Study of maturation by linear analysis

Selection of data and variables

Among the breathing signals available in both VACCIN and PHYSIDEV databases, we selected the 31 patients from the first and 27 from the second to organize data in maturation groups according the different ages and weights, proceeding similarly than in Chapter IV. However, in this case the range of the variables mixing both databases was larger and allowed to split the population in four groups as shown in Table V.2.

Before analyzing respiratory data, it was preprocessed by the noise rejection algorithm and the filters. In average, the 58 signals contained a 14.8% ± 14.2 of artifacts. Then, a first comparison of the general descriptors of breathing could be done between grups. These descriptors reflect the temporal distribution and irregularity of breathing:

- Concerning cycles: The mean duration of cycle times ($\overline{t_{tot}}$) in seconds and their standard deviations (σt_{tot})

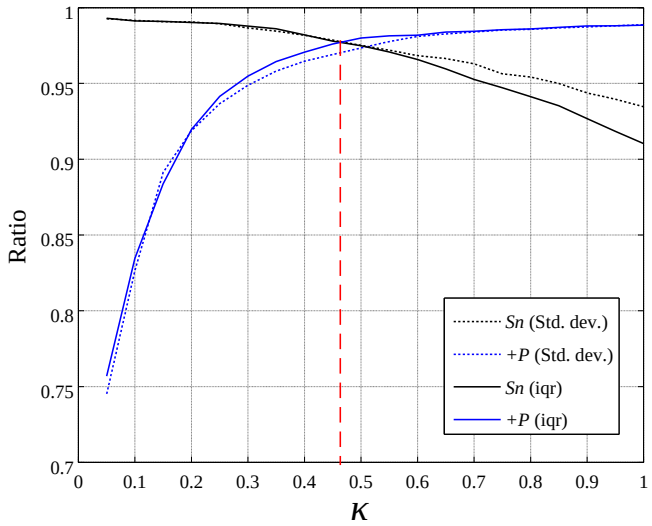

Figure V.8 – Performance plot of the cycle detector, applying the κ coefficient to two features: standard deviation (dotted lines) and IQR (solid lines). The optimal operating point, for each feature, is the intersection of the sensibility and positive predictability curves. The horizontal red dashed line indicates this point for IQR (κ=0.463).

	PMA (weeks)	PNA (days)	W (Kg)
G_1	27.7 – 30.7	3.00 – 12.2	0.69 – 1.14
G_2	30.7 – 35.3	12.2 – 42.0	1.14 – 1.57
G_3	36.4 – 38.1	50.0 – 72.2	1.64 – 2.28
G_4	38.1 – 41.2	72.0 – 105	2.28 – 3.23

Table V.2 – Description of the groups according to their PMA (postmenstrual age), PNA (postnatal age) and W (weight at present day). In the nomenclature of groups, the concerned variable is super-indexed. For instance, the most immature group according PMA is G_1^{PMA}.

to describe breathing variability. To describe the shape of breathing cycles we computed the mean ratio of inspiration and exhalation times $\overline{t_i/t_e}$.

- Concerning apnea: The average number of apneas normalized in one hour (N_{ap}), their mean duration (D_{ap}) and their standard deviation (σ_{ap}) in seconds.

Inspired by the spectral analysis of the heart beat ratio, we hypothesize that some system's control information may be obtained by investigating periodic components in the respiratory rhythm. In RV spectrum, only data about the duration of cycles is processed and bias due to amplitude variations in breaths can be avoided. Indeed, temporal signals are directly measured from abdominal movements –uncalibrated and noisy– and a spectral analysis would not provide reliable information.

Therefore, we computed the power spectra using FFT on resampled RV signals at 2 Hz (see Figure V.9), dividing the range of frequencies into four components expressed as the percentage relative to the total spectral power:

- VLF, or very low frequency: defined from 0 to 0.05 Hz.
- LF, or low frequency: from 0.05 to 0.15 Hz.

- MF, or medium frequency: from 0.15 to 0.25 Hz.
- HF, or medium frequency: from 0.25 to 0.60 Hz.

Note that the above frequencies have been chosen arbitrarily and the components are not equivalent to those typically studied for heart rate variability because of the different dynamics of breathing.

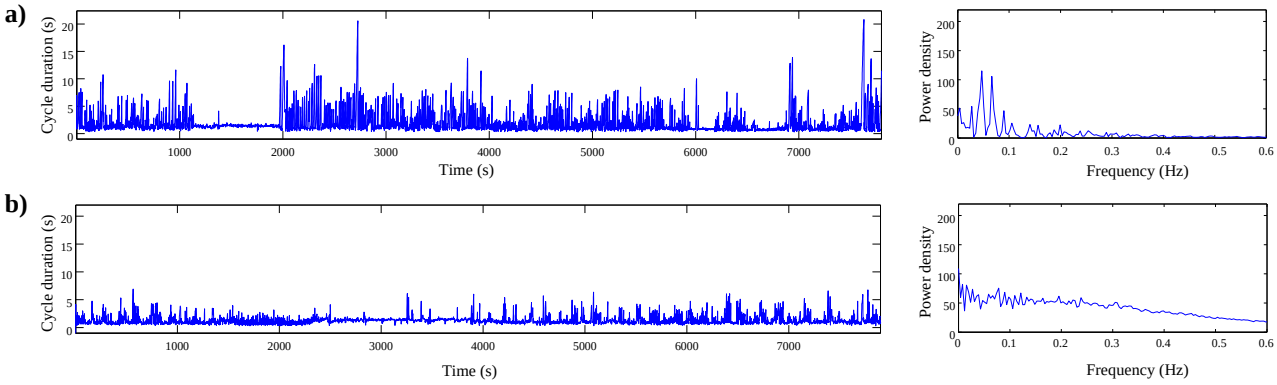

Figure V.9 – (a) RV signal having periodic breathing and its power density. Periodicities can be found in the spectrum between 0.05 and 0.1 Hz. (b) RV signal predominantly regular, characterized by a more flat spectrum.

The respiratory variables were computed in the entire recordings (two to three hours per patient) and were averaged. Due to the absence of hypnograms in the PHYSIDEV database, sleep could not be quantified in this analysis, but provided that the selected infants are healthy, we presume that their sleep patterns are in the normal ranges for their age.

Results and discussion

The results of computing the respiratory variables as a function of the four maturity groups by PMA are given in Figure V.10, where it can be stated a general increase of t_{tot} with increasing age, a decrease of the number of apneas, N_{ap}, and an overall decrease in their standard deviation σ_{ap}.

Although the statistical analysis did not reveal significant differences in all variables comparing age-coarse groups (see horizontal bars in the same figure), this trend is in accordance with reference works in the literature [9, 10].

The other assortments comparing the breathing variables between postnatal age and weight groups provided less meaningful, but a still noticeable trend towards the decrease of N_{ap} and t_{tot}. The full results are detailed in the tables from Appendix D.

The spectral analysis of RV signals showed as well slight differences between groups. The percentages of VLF decrease whereas LF and HF increase. However, finding a physiological explanation to this observation is not obvious because we did not find similar works in the literature. The closest research concerns probably the works of Aärima and Välimäki [11] and Rantonen et al. [12] in which spectral estimation of respirograms (temporal signals from abdominal movements) were performed in newborns. The latter defined frequency bands such that they encompass periodic breathing, the natural baroreflex activity and thermoregulation in the lowest frequency ranges. On the other hand, the spontaneous breathing corresponded to higher frequency bands.

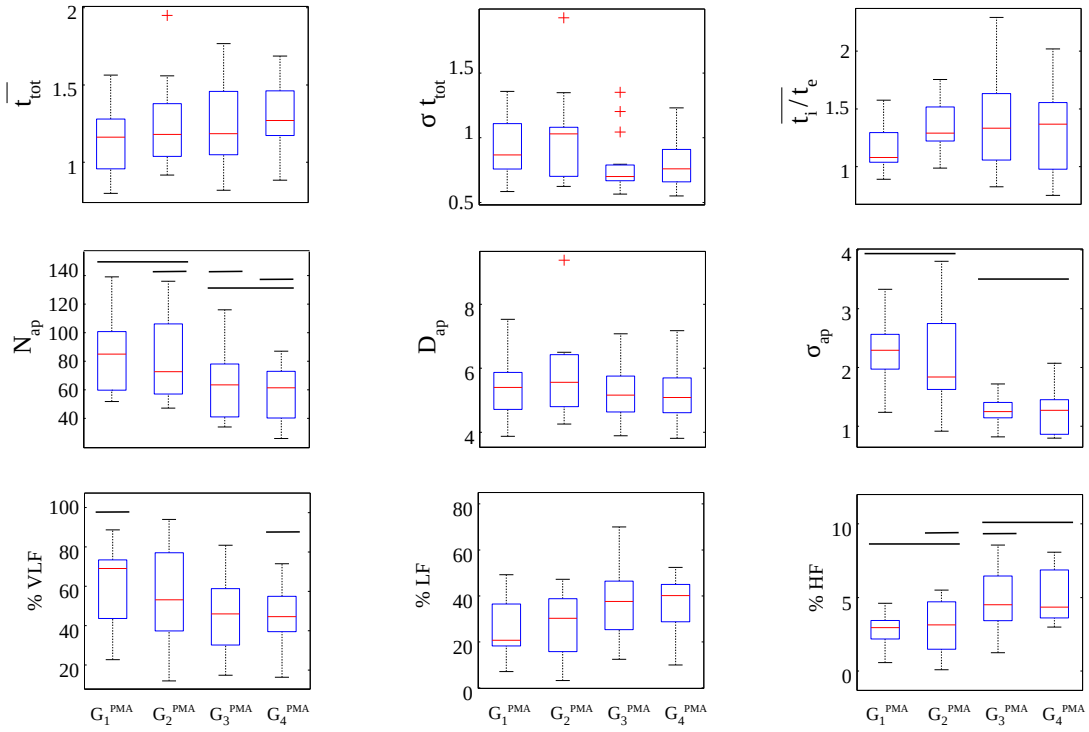

Figure V.10 – Respiratory variables for the four maturation groups according to PMA. Horizontal lines of different heights indicate significant differences in a Mann-Whitney U test.

In our case, the relative power in each component depends on both short-term and long-term variability of breaths. High short-term variability of regular, spontaneous breaths is associated to the HF component and long-term variations, such as nonstationarities due to changes in the baseline, are associated to the VLF component and might be originated by changes in thermoregulation, feeding or baroreflex activity. On the other hand, LF could be associated to periodic breathing, appearing between 0.05 and 0.1 Hz, as shows the example in Figure V.9-a).

According to Aärima and Välimäki, the ratio LF/HF in impedance respirograms is higher in preterm than in term infants, evidencing that the control of respiration changes in relation to maturity. Similarly, we can conclude that a decrease of VLF and an increase of LF and HF in RV signals could be related to an increase of maturation and may constitute an alternative indicator reflecting the development of the respiratory control.

3.2 Study of infection

Selection of data and variables

In highly immature and unstable breathing patterns, investigating abnormalities caused by a septicemic process is not straightforward. Given the well-known dependence on maturation, comparing age-equivalent sick and healthy infants is mandatory. Hence, a sub-cohort constructed from the PHYSIDEV database, named SEPSIS, contained a selection of sixteen infected (referred as *Sepsis*) and sixteen non-infected infants (*No-sepsis*) paired by age, gender and weight criteria (see Table V.3).

To find differences between *Sepsis* and *No-sepsis* groups, we studied the same linear variables presented in

	Sepsis	*No-sepsis*
Infants	16	16
PMA (weeks)	30.5 ± 1.73	30.4 ± 1.64
PNA (days)	15.6 ± 12.2	15.8 ± 10.7
Weight (Kg)	1.11 ± 0.27	1.12 ± 0.23
Duration (h)	2.14 ± 1.07	2.61 ± 0.64

Table V.3 – Description of the characteristics composing the sub-cohort SEPSIS. There are no significant differences between the age and weight of infants and duration of records.

Section 3.1.

Results and discussion

The results comparing the respiratory variables between healthy and infected patients are summarized in Table V.4. *Sepsis* patients have in average slightly longer and more variable duration of breaths, a higher number of apneas and more pronounced VLF content although statistical tests did not find significance.

	Sepsis	*No-sepsis*
t_{tot} (s)	1.49 ± 0.45	1.37 ± 0.26
σt_{tot}	**1.27 ± 0.59**	**0.84 ± 0.24***
t_i/t_e	1.17 ± 0.10	1.12 ± 0.07
N_{ap}	50.6 ± 33.7	41.1 ± 34.0
D_{ap} (s)	**8.09 ± 2.84**	**5.91 ± 1.20***
σ_{ap}	**3.48 ± 2.06**	**2.01 ± 1.01***
VLF (%)	66.30 ± 19.93	56.87 ± 19.86
LF (%)	29.98 ± 17.20	39.15 ± 17.52
HF (%)	3.32 ± 3.43	3.77 ± 2.85

Table V.4 – Results from the comparative study of *Sepsis* and *No-sepsis* populations. The variables are expressed as mean ± std. dev. Asterisks denote statistically significant differences in a Mann-Whitney U test ($p < 0.05$).

On the other hand, the standard deviation of t_{tot} was significantly higher, a fact that reveals more unstable breathing patterns in sick infants. Moreover, the mean duration of apneas, D_{ap}, as it standard deviation, σ_{ap}, are substantially longer.

Therefore, it can be stated that breathing in infected infants is, in general, more variable concerning the duration of breaths and characterized by longer episodes of apnea. Considering the relationships between apnea and bradycardia, these results are complementary to the findings related to cardiac variability, in which severe, unusual and recurrent bradycardias are documented in infected populations [13, 14].

4 Conclusion

In this chapter, we have processed the raw breathing signals acquired in the NICU to obtain clean respiratory traces and their respective RV series. Despite the probability of including corrupted segments (or excluding good segments) is non-negligible, the detector can be adjusted by modifying the cut-off point c in the LR classifier. If

the subsequent application does not tolerate noise, such as certain nonlinear analyses of the temporal signal, c should maximize sensitivity at the cost of loosing useful information. On the contrary, applications more robust to noise and needing long records, could deal with lower sensitivities.

Notwithstanding, the presented rejection algorithm has little room for improvement. Changing the labeling protocol, allowing the rater to mark freely the limits of the artifacts could be a solution, but it may result tedious. Including complementary breathing signals (thoracic movements or nasal flux) could constitute another solution to obtain more reliable data.

Because the normal development of respiratory controlling systems is associated with characteristic changes in linear and nonlinear measures, we first studied the statistical characteristics of the breathing cycles, apneas and the spectra of RV signals. As expected, most of the measures obtained from our data (duration of breaths, apneas) confirmed their evolution in accordance to maturative physiological criteria (age and weight), evidencing functional and anatomical changes in the respiratory system. But the spectral content of RV signals was, to our knowledge, unexplored. The progressive reduction of VLF and increase of HF components with age might be an additional proof of maturity ans should be further investigated in horizontal databases. The study of our healthy and infected cohort also provided some insights about the breathing characteristics during late onset sepsis. Indeed, linear parameters evidenced the existence of more irregular breathing cycles in sick infants, as well as longer apneas.

In conclusion, the results obtained by applying linear tools on our data were consistent to other reference works and should be useful to quantify preterm infants' maturity or for diagnostic purposes. But to best characterize the dynamics of breathing, it is necessary to go one step beyond and explore nonlinear methods.

Bibliography

[1] A. L. Motto, H. L. Galiana, K. A. Brown, and R. E. Kearney. *Detection of movement artifacts in respiratory inductance plethysmography performance analysis of a neyman pearson energy based detector.* 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society 2004, IEMBS'04 **1**, 49–52 September (2004).

[2] S. Kay. *Fundamentals of Statistical Signal Processing, Volume II: Detection Theory.* Prentice Hall, 1 edition February (1998).

[3] N. Mammone and F. C. Morabito. *Enhanced automatic artifact detection based on independent component analysis and renyi's entropy.* Neural Networks **21**(7), 1029–1040 September (2008).

[4] A. Savitzky and M. J. E. Golay. *Smoothing and differentiation of data by simplified least squares procedures.* Analytical Chemistry **36**, 1627–1639 (1964).

[5] R. Schafer. *What is a savitzky-golay filter? [lecture notes].* IEEE Signal Processing Magazine **28**(4), 111 –117 July (2011).

[6] W. P. Holsinger, K. M. Kempner, and M. H. Miller. *A QRS preprocessor based on digital differentiation.* IEEE Transactions on Biomedical Engineering **BME-18**(3), 212 –217 May (1971).

[7] J. Pan and W. J. Tompkins. *A real-time QRS detection algorithm.* IEEE Transactions on Bio-Medical Engineering **32**(3), 230–236 March (1985).

[8] ANSI/AAMI CE57. Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms. recommended Practice/American national standard. available at: http://www.aami.org, (1998).

[9] L. Curzi-Dascalova. *Développement du sommeil et des fonctions sous contrôle du système nerveux autonome chez le nouveau-né prématuré et à terme.* Archives de Pédiatrie **2**(3), 255–262 March (1995).

[10] D. Holditch-Davis, M. Scher, T. Schwartz, and D. Hudson-Barr. *Sleeping and waking state development in preterm infants.* Early Human Development **80**(1), 43–64 October (2004).

[11] T. Äärimaa and I. A. Välimäki. *Spectral analysis of impedance respirogram in newborn infants.* Biology of the neonate **54**(4), 188–194 (1988).

[12] T. Rantonen, J. Jalonen, J. Grönlund, K. Antila, D. Southall, and I. Välimäki. *Increased amplitude modulation of continuous respiration precedes sudden infant death syndrome: Detection by spectral estimation of respirogram.* Early Human Development **53**(1), 53–63 November (1998).

[13] M. P. Griffin, T. M. O'Shea, E. A. Bissonette, F. E. Harrell, D. E. Lake, and J. R. Moorman. *Abnormal heart rate characteristics preceding neonatal sepsis and sepsis-like illness.* Pediatric Research **53**(6), 920–926 (2003).

[14] A. Beuchée, G. Carrault, J. Y. Bansard, E. Boutaric, P. Bétrémieux, and P. Pladys. *Uncorrelated randomness of the heart rate is associated with sepsis in sick premature infants.* Neonatology **96**(2), 109–114 (2009). PMID: 19279394.

Chapter VI
Nonlinear analysis

In previous chapter, breathing signals have been properly processed to obtain exploitable temporal and variability series. The analysis of some linear measures permitted a first characterization of data with regard to maturity and diagnostic of neonatal sepsis.

Beyond these evidences, however, breathing signals may hide underlying information related to the dynamics generated by the respiratory control. Therefore, this chapter immerses itself in nonlinear methods, a vast field of analysis providing potential indicators for our research. First, we argue why the nonlinear approach could be useful in our signals and provide the definitions of the selected methods, as well as justifying their choice. Then, the breathing signals will be analyzed by the nonlinear approach to examine their usefulness in one of the clinical applications of interest in this thesis: the study of maturity.

1 Nonlinearity in biomedical signals

Nonlinear phenomena are certainly behind the genesis of cardio-respiratory rhythms and brain activity. They are determined by complex interactions of hemodynamic and electrophysiological but also by the regulation of autonomic and central nervous systems. It has been speculated that analyzing biomedical signals through methods based on nonlinear dynamics may unveil valuable information, unaccessible to linear methods, for physiological interpretation.

Indeed, the nonstationary and random appearance of physiological time-series is found to exhibit complex autocorrelation patterns, a behavior associated to nonlinear dynamics. And often, implicit to this nonlinearity, it is assumed that chaotic behavior and fractal structures exist [1]:

- Chaotic behavior implies a unstable, aperiodic behavior with deterministic, dynamical systems highly sensitive to the initial conditions.
- Fractal structures connote the repetition of shapes at different scale levels, a self similar behavior.

Despite chaos and fractals are closely related and often occur together, the analysis of each property may suggest different approaches. While chaos is a class of dynamical behavior, deals with time evolution and its underlying characteristics, fractals, in contrast, deal with geometrical patterns and quantitative ways of characterizing those patterns.

In biomedical signals, chaotic and fractal behavior is evident only to a certain extent due to the presence of noise, both internal (inherent of physiological control systems) and related to the observation. Therefore, the underlying determinism and/or self-similarity can only be estimated with a few degrees of freedom.

In this chapter, we avoid to classify nonlinear methods upon chaotic terms because, strictly speaking, this property does exist in our signals. Instead, we prefer to refer to them as methods quantifying complexity, distinguishing between the approaches proposed by the Information Theory and the Chaos Theory. In next chapter, we study extensively methods quantifying long-range dependence, the other notion of nonlinearity -the fractal properties- explored in the present work.

2 Measuring complexity

A large number of methods deal with the notion of complexity through different approaches to provide insights of the underlying dynamics and relationships between physiological control systems. In this section, we include some of the commonly used methods in the context of biomedical engineering, whenever they are adequate regarding the characteristics of our signals. The Information Theory (IT), proposed by Shannon and Weaver [2], emerged as a field to measure and process information mathematically, from a statistical and probabilistic point of view. IT considers signals as random variables, so that the complexity is interpreted as a measure of the uncertainty or unpredictability, or equivalently, as the information content of those variables. The entropy is the most employed expression of this complexity and constitutes the basis of the herewith described measures.

2.1 Mutual information

Intuitively, the mutual information (MI) of two variables is a quantification of the dependence, or the shared information, between the two variables. It can also be interpreted in probability terms, that is, by knowing the information from one variable, the mutual information provides a measure of the uncertainty about the knowledge of the other [2].

Let $X = [x_1, x_2, ...x_n]$ and $Y = [y_1, y_2, ..y_m]$ be two discrete random variables, with probability $P_X(x_i) \geq 0$ and $P_Y(y_i) \geq 0$, the mutual information is expressed formally as:

$$MI_{XY} = \sum_{i=1}^{n} \sum_{j=1}^{m} P_{XY}(x_i, y_j) \log \frac{P_{XY}(x_i, y_j)}{P_X(x_i) P_Y(y_j)}, \tag{VI.1}$$

where $P_{XY}(x_i, y_j)$ is the joint probability distribution of X and Y, and $P_X(x_i)$, $P_Y(y_j)$ their respective marginal probability distributions.

The auto-mutual information (AMI) in a random variable estimates how much –on average– the value of the variable can be predicted from values of the variable in question at preceding points. It can be defined as the MI between X and its τ delayed version, X_τ:

$$AMI_{XX} = \sum_{x_{i} \in X} \sum_{x_{i+\tau} \in X} P_{XX}(x_i, x_{i+\tau}) \log \frac{P_{XX}(x_i, x_{i+\tau})}{P_X(x_i) P_X(x_{i+\tau})}. \tag{VI.2}$$

In a completely regular signal, AMI would be the same at $\tau = 0$ and at any prediction time ($\tau > 0$) but when determinism does not exist (a random signal), AMI=0 for any $\tau > 0$. Therefore, a decrease of the AMI function with increasing τ, is related to a loss of the predictability on the information contained in X, or equivalently, to a higher complexity.

AMI has been employed in several applications, including the measure of the linear and nonlinear content in

experimental signals. In EEG, it resulted useful to study the information transmission in the human cerebral cortex [3], to detect the Alzheimer's disease [4, 5], or in monitoring the depth of anesthesia [6]. It has been also applied on breathing signals, quantifying nonlinearity in spontaneously breathing rats [7] or contributing in the construction of a sleep-disorder diagnostic apparatus [8].

2.2 Sample entropy

As introduced above, entropy refers to the degree of uncertainty of a random variable but, according to the theoretical definition of Shannon, vast amounts of data are required to estimate the entropy accurately. Experimental time-series, relatively short and noisy, need alternative algorithms to solve this limitation.

Sample Entropy (SpEn) [9] was proposed as a statistic to measure the entropy in short data-series based on Approximate Entropy (ApEn), a similar algorithm introduced earlier.

It has been largely utilized to detect abnormalities in heart rate series, because a decrease in complexity is associated to a degradation of the cardiovascular control system [9, 10]. Applied to RR series of preterm neonates, it has been found that low values of entropy are strongly associated with the risk of having neonatal sepsis [11, 12].

The calculation of ApEn [13] from discrete-series $X = [x_1, x_2, ...x_n]$ is carried out by firstly constructing sub-series of vectors (patterns) V:

$$V = [v_1, v_2, \dots v_m] = [x_n, x_{n+1}, \dots x_{n+m-1}]. \tag{VI.3}$$

Then, by defining:

1. D_{ij} as the distance between two vectors so that $D_{ij} = d[v_i, v_j]$,
2. r as a parameter establishing the tolerance so that two vectors are comparable,
3. N_r^m as the number of vectors or patterns j such that D_{ij} is equal to r,
4. $C_r^m(i)$ as the ratio $N_r^m/(N-m+1)$, and
5. $F_r^m = \frac{\sum_{N-m+1}^{i=1} ln(C_r^m(i))}{N-m+1}$

ApEn is given by:

$$ApEn_r^m = F_r^m - F_r^{m+1}. \tag{VI.4}$$

This method examines time-series for similar patterns, so the more frequent and more similar they are, the lower values of ApEn are obtained. A low value of ApEn reflects a high degree of regularity. Nevertheless, this algorithm counts each sequence as matching itself to avoid the occurrence of $ln(0)$ in the calculations. In practice, this leads to bias ApEn, to a heavily dependence of the record length and to a lack of relative consistency. Indeed, if ApEn of one dataset is higher to another one, that should remain higher for all conditions tested, but it does not.

Consequently, SpEn was introduced to reduce this bias [9] by avoiding self-matches counted in $C_r^m(i)$.

Intuitively, SpEn is a permutation of the ratio of the number of matches for the $(m+1)$-point templates (green boxes in Figure VI.1) to the number of matches for the m-point templates excluding self-matches.

More precisely, SpEn is calculated by the following equation:

$$SpEn(\tau) = -\ln \frac{\sum_{i=1}^{N-m} A_i^m(\tau, r)}{\sum_{i=1}^{N-m} B_i^m(\tau, r)}, \tag{VI.5}$$

where $A_i^m(\tau, r)$ and $B_i^m(\tau, r)$ are the probabilities that two $(m+1)$-point sequences and two m-point sequences match within the given tolerance, respectively, and τ is the delay introduced between the points.

Computing SpEn across several time delays (see Figure VI.1) allows to better capture nonlinear contributions than using a single sample delay ($\tau = 1$) as done traditionally [14].

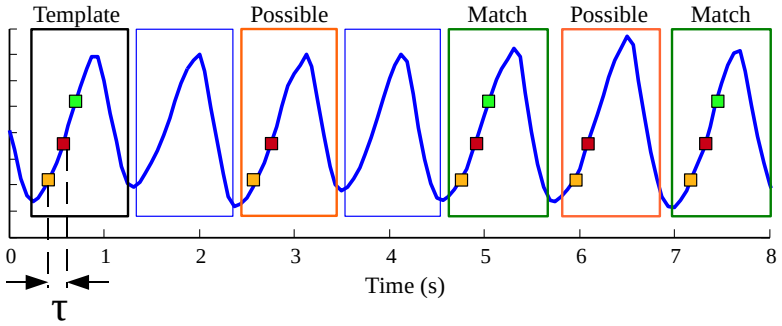

Figure VI.1 – Example of SpEn computation with a size pattern of $m = 2$. First a template is constructed (yellow and red points) introducing a delay τ. Then, this pattern is identified in the time-series (orange windows), and finally, the pattern is extended to $m + 1$. If the points are within the tolerance r (green points), the matching patterns are counted (green boxes).

While low values of entropy denote regularity or determinism in time-series – SpEn in a continuous signal is close to zero– , higher values correspond to random, unpredictable signals – SpEn ≈ 2.2 in white noise–.

2.3 Lempel-Ziv complexity

According to the Algorithmic Information Theory (AIT) [15], a subfield of the IT, the information content or complexity of a signal can also be measured by the length of its shortest description. The notion of simplicity and complexity becomes very intuitive, because signals having little changes can be described in a few words –for example a continuous DC signal– and hence, they are "simple". Conversely, for random signals there is no such short description, needing bit-by-bit literal descriptors.

The Lempel-Ziv (LZ) complexity for finite sequences was proposed as a nonparametric measure of complexity in univariate signals [16]. It is a straightforward technique that does not require long data segments to be computed.

LZ complexity expresses the number of distinct substrings and their recurrence in a sequence, so that larger values mean more complex data. A number of works have applied this measure of complexity on biomedical signals, specially EEG. It has been found that LZ is sensitive to the brain function [17], useful to quantify brain activity during anesthesia [18] and as a predictor of epileptic seizures [19]. LZ complexity might be, therefore, an useful tool to quantify the maturity of the burst patterns in our EEGs.

To compute LZ complexity the analyzed signal has to undergo a quantification, i.e. continuous values must be converted in a symbolic sequence, such as the binary coding employed for burst/inter-burst periods.

Being the binary sequence $S = s_1 s_2 \dots s_n$, with $s_i \in \{0, 1\}$ and $i = 1, \dots, n$, it has to be parsed to obtain distinct words, which will be encoded. There are several manners to perform parsing, but for simplicity we describe the scheme proposed by Cover and Thomas [20]: S is sequentially scanned and rewritten as a concatenation $w_1 w_2 \dots$ of words w_k, chosen such that $w_1 = s_1$ and w_{k+1} is the shortest word that has not appeared previously. For example, after parsing the sequence 10101110100 one obtains 1·0·10·11·101·00, so the number of words, denoted $c(n)$, is 6.

Next step is the encoding of the words of the sequence. Following the criterion given by [20], $\log_2 c(n)$ bits are needed to describe the location of the prefix for each word, and one bit to describe the last symbol. In the above example, 000 describes an empty prefix, so the entire sequence can be expressed as (000,1)(000,0)(001,0)(001,1)(011,1)(010,0).

Finally, the LZ complexity is found by the following expression:

$$C_{LZ} = c(n)\frac{\log_2 c(n) + 1}{n}. \tag{VI.6}$$

3 Measures based on chaos estimations

As introduced previously, it has not been formally demonstrated that biosignals are chaotic, but many methods based on the Chaos Theory have been used widely to study biomedical signals because chaotic behavior is observed [21]. Although these techniques can be exact when computed on dynamical systems described by equations, as for instance, the Rössler system [22], the measurement noise in experimental signals is a major problem for many of them. Here, we will focus on chaos estimation techniques traditionally used to analyze short and noisy data, the correlation dimension and the largest Lyapunov exponent (LLE).

3.1 State space reconstruction

Chaos-based methods try to provide a complete description of the high dimensional nonlinear dynamics by first unfolding time-series into a state space, where the time evolution of the system may be constrained in certain finite regions [23]. Since the system function describing the time-series dynamics is unknown, an embedding space preserving the invariant characteristics of the original attractor is needed to characterize its dimension.

The most popular approach to reconstruct the attractors is based on Taken's theorem [24], which states that a chaotic dynamical system can be reconstructed by embedding a discrete signal, s_i, $i = [1, 2...n]$ in a m-dimensional Euclidean space to create $N = n - m(1 - l)$ delay state vectors, $\mathbf{x}_i$, $i = [1, 2..N]$, such that:

$$\mathbf{x}_i = [s_i, s_{i+l}, s_{i+2l}, \dots, s_{i+(m-1)l}], \tag{VI.7}$$

where $l \in \mathbb{N}$ is called time lag and m is referred as embedding dimension. The theorem guarantees that the attractor embedded in the m-dimensional state space is completely unfolded, without any self-intersections.

It has been found that estimating the correlation dimension and LLE in finite and noisy data can result highly dependent on m and l [25, 26], thus the main concern of state space reconstruction is to determine these values properly.

The dimension m can be estimated by assessing the geometrical structure of the reconstructed attractor while embedded in increasing dimensions. Low values of m produce self-intersections, and the resulting attractor has

spatially nearby points, either due to the dynamics of the system (real neighbors) and due to self-intersections (false neighbors). This is precisely the principle of the false nearest neighbor technique [25, 26], which finds m such that no false neighbors are identified as dimensions increase.

The lag l between coordinates has to be chosen not too small because the reconstructed state space would be similar nor too large to avoid information loss. An effective and simple solution is to set the lag equal to the first minimum of the autocorrelation function [27].

3.2 Correlation dimension

Correlation dimension (D_2) is a measure to estimate the temporal behavior of chaotic attractors from measured signals. The first step consists on constructing delay state vectors, $\mathbf{x}$, and thereafter computing the correlation integral, $C_m(r)$.

The Grassberger and Procaccia algorithm [28] is a widely-used method to compute D_2. The starting point consists on defining a correlation integral:

$$C_m(r) = \lim_{N\to\infty} \frac{1}{N^2} \sum_{i\neq j}^{\infty} \Theta(r - |\mathbf{x}_i - \mathbf{x}_j|), \tag{VI.8}$$

where Θ is the Heaviside step function and $|\cdot|$ denotes the distance between a pair of points, so $C_m(r)$ is the fraction of pairs in the delay state vectors whose distance is smaller than r. $C_m(r)$ is monotonically decreasing to zero as $r \to 0$, and if the above limit exists, the correlation dimension is defined as:

$$D_2 = \lim_{r\to\infty} \frac{\log C_m(r)}{\log r}. \tag{VI.9}$$

The main inconvenient of this approach is that the presence of measurement noise in experimental signals put severe limitations on the estimation of dynamical invariants [29]. To solve this issue, Diks et al. [30] proposed the use of a kernel function $w(x) = e^{-x^2/4}$ instead of the Heaviside function, stating that the correlation dimension and correlation entropy can be found for time series with up to 20% of noise. This solution, called Gaussian kernel algorithm (GKA), has been adopted in the present work to compute D_2, employing an efficient algorithm implementation described in [31].

3.3 Numerical noise titration

The numerical noise titration was proposed by Poon and Barahona [32] to provide a robust measure of chaos intensity even in the presence of noise and with short data series. So far, a classically employed measure of chaos was the Lyapunov exponent, which focuses on the sensitivity to the initial conditions of a dynamic system. But this method fails to distinguish chaos from noise reliably unless the data series are very long and virtually free of noise [33].

Despite noise titration was also found to have some pitfalls detecting deterministic chaos confidently [34], it can still be used as an efficient tool to detect nonlinearity in time series as an alternative to the LLE.

Noise titration, applied to the ECG, allowed to characterize different cardiac pathologies in adults [35], or to study the breathing dynamics in spontaneously breathing adults [36, 37] and in mechanical-assisted ventilation [38]. It has also been employed in variability series (heart beat rate) to assess the effect of age in cardiac activity

[39].

The basis of the noise titration is to detect nonlinearity in a time-series, S_i (containing the measured signal), by comparing the one-step-ahead errors of the best linear model and the best nonlinear model, found by fitting autoregressive polynomial models (both linear and nonlinear) from S_i. This procedure is known as nonlinearity test.

Linear models are described by their order, p, (maximal temporal lag) and a set of parameters, ϕ_i, and can be expressed as:

$$S_i = \phi_0 + \phi_1 s_{i-1} + \phi_2 s_{i-2} + \ldots + \phi_p s_{i-p}. \tag{VI.10}$$

They have degree $d = 1$, and a number of terms $\eta = p + 1$, but in nonlinear models η is determined also by d. Setting a maximum number of terms η, the Volterra-Wiener algorithm produces a family of nonlinear and linear polynomial autoregressive models varying d and p to optimally fit to s_i. It has to be noted that, unlike other nonlinear detection techniques, the Volterra-Wiener algorithm does not perform surrogate-data testing, but it provides a very robust and sensitive indicator of nonlinearity in short and noisy data, thereby establishing a necessary (but not sufficient) condition for the existence of chaos [32, 40].

Once the best linear and nonlinear models are chosen according to a theoretic information criterion (Akaike), the null hypothesis, that the best model is linear, is tested by a Mann-Whitney statistic. If the resulting p value is under 0.01, then it is considered that there are nonlinear dynamics between two samples delayed in time.

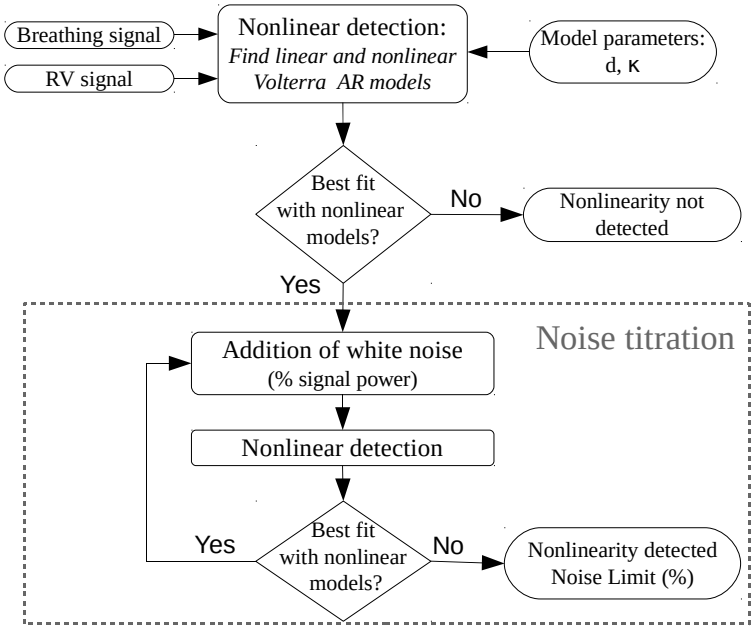

Figure VI.2 – Illustration of the noise titration technique.

If s_i passes the nonlinearity test, then noise titration can begin. It consists on adding to the time-series white noise increasing its standard deviation until nonlinearity is no longer detected. The level of noise, called noise limit (NL) denotes the relative intensity of nonlinearity, so NL=0 can be interpreted as the neutralization of nonlinear dynamics by the background noise in the data. The entire process is illustrated in Figure VI.2.

4 Nonlinear indexes

We designed nonlinear indexes to quantify the overall degree of nonlinearity when certain nonlinear methods need to range their parameters between certain values to characterize nonlinear dynamics. This is the case of mutual information and sample entropy, where the lag τ is set to several delays, and correlation dimension, where several values of the embedding dimension d are required to estimate the attractors. The particularity of these nonlinear indexes is that they are relative to the values of AMI, SpEn or D_2 computed in a set of surrogates from the original time-series, where any nonlinear content is removed. Appendix C describes the methodology involved in surrogate-data.

The three nonlinear indexes, NI, are named depending on the employed method: NI_{AMI} for mutual information, NI_{SpEn} for sample entropy and NI_{D2} for correlation dimension. The latter, for instance, is constructed in a signal X as follows:

- Estimate $D_2(X)$, the correlation dimension on the original signal across the chosen range of embeddings d.
- Obtain a set of i surrogates, S_i, from the original signal, with $i > 20$ to obtain reliable results.
- Estimate the correlation dimension on the surrogate set and obtain the median, $\widetilde{D}_2(S_i)$.
- Obtain the nonlinear index, $NI_{D2} = \text{mean}(\widetilde{D}_2(S_i) - D_2(X))$.

Equivalently, NI_{AMI} and NI_{SpEn} can be obtained varying τ.

Since randomized signals have higher chaos dimension, $\widetilde{D}_2(S_i) > D_2(X)$, hence $NI_{D2} > 0$ when nonlinearity is detected. Likewise, $NI_{SpEn} > 0$ because surrogates exhibit higher entropy than the original signal.

In Figure VI.3 we illustrate how nonlinear indexes are computed in an example, a regular breathing excerpt of one minute. Panels b) and c) contain two return maps for different time lags. Return maps constitute a geometrical tool to see the attractors of the dynamical system for a given lag. In d) the correlation dimension, D_2, is computed across the range of embeddings, both for the original signal and the set of surrogates. As it can be observed, the medians of the surrogates have higher values, meaning that complexity is higher when nonlinearity is removed. Similarly, in e) the sample entropy computed in surrogates leads to greater values.

5 Clinical application

In this section, we first apply nonlinear tools onto the filtered breathing signals to verify whether they provide more insights about the maturation of the respiratory control. Infants of the PHYSIDEV database and not immunized infants from VACCIN records were included to enlarge the age and weight bounds. Then, taking advantage of the here described methods, we get back to EEG burst signals to perform a quick analysis measuring their complexity.

5.1 Nonlinear analysis of temporal signals

Because many of the tools quantifying the degree of nonlinearity in rhythm generators are based in chaos theory, the selected data must be as clean as possible. Nevertheless, continuous data having enough quality in the NICU are scarce even during the steadiest recording periods, so the challenge of this analysis is to find indicators of maturity from short samples of regular breathing.

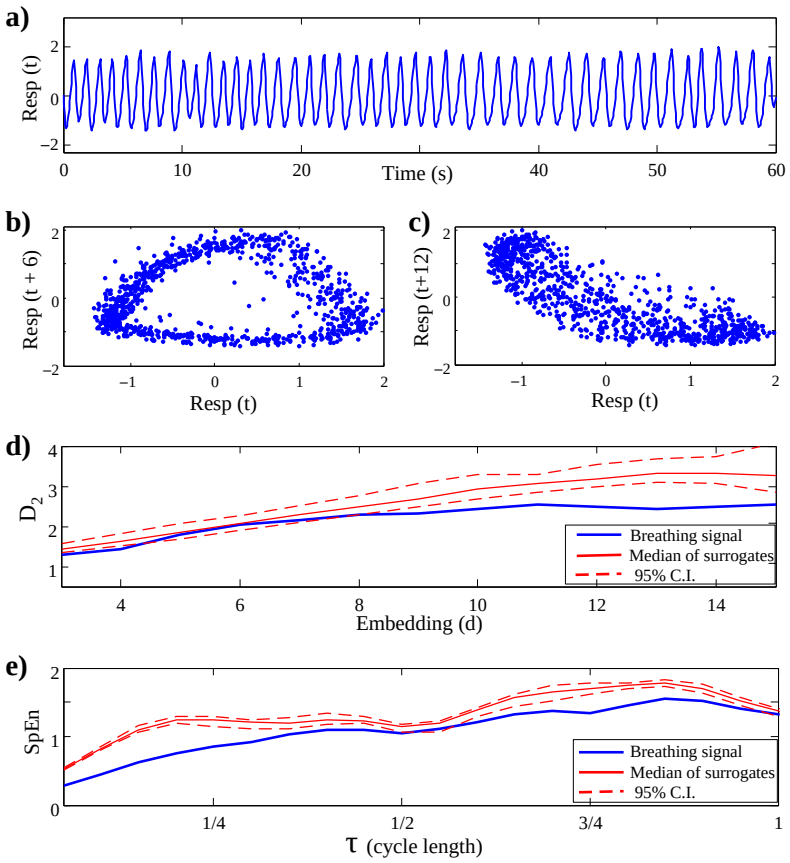

Figure VI.3 – Example of one-minute breathing signal (a). Return maps plotted from the original and time-delayed data sets: 6 points (b) and 12 points (c). Estimated D_2 with embeddings from 3 to 15 (d). SpEn of breathing signal (e) calculated with τ from 1 to 20 points (1-cycle length). The respective nonlinear indexes are obtained by the averaged difference of original and surrogate curves.

State of the art

About nonlinearity in breathing, it has been observed that adult human ventilation shows chaotic dynamics under certain conditions, but it remains still unclear whether the rhythm generated by respiratory control centers is governed by deterministic chaos [37]. One of the reasons contributing to this uncertainty is that experimental conditions, as the presence of noise, or an insufficient series length makes detection of chaos quite arduous. Thus, quantifying underlying nonlinear dynamics rather than chaos in physiological data seems more realistic and reliable.

Very few works are devoted to analyze nonlinear dynamics in breathing during infancy. In healthy full-term newborns, nonlinearity is present in ventilation, as demonstrated by Small et al. measuring abdominal movements by inductive plethysmography [41]. The same authors proposed later some modeling methods for these signals and stated that chaos was probably present during quiet sleep by examining the correlation dimension [42].

To our knowledge, no specific research have been done using preterm infants' breathing signals from the NICU until we presented a preliminary study of several nonlinear measures [43]. Exploiting only PHYSIDEV patients, we observed that nonlinearity slightly increased regarding maturative aspects such as postmenstrual

age and weight using noise titration and the nonlinear indexes described in Section 4. Including more infants, the present study will confirm if nonlinearity exists in our breathing signals and to which extent it evolves with maturation by evaluating the same measures of nonlinearity.

Data selection and parameter tuning

As introduced before, temporal signals need to be carefully selected because nonlinear methods are very sensitive to both observational and dynamic noise. Thus, breathing signals only during quiet sleep were included:

- Respiratory signals are the most regular and clean. The absence of body movements reduces the observational noise with respect other sleep states.
- Infant is deeply asleep. Reactivity to external stimuli is minimal and other sleep characteristics adding complexity (such as eye or limb movements) does not exist. Therefore, dynamical noise in these conditions should be the lowest.

An algorithm designed purposely selected automatically five fragments of ninety seconds along the entire recording (see Figure VI.4), ensuring that they had the smallest variances in the breathing rhythm ($\sigma(t_{tot})$) and in the amplitude of breaths ($\sigma(A)$). Since the hypnograms are not always available, this rule ensures that the infant is in quiet sleep. Then we down-sampled signals (originally sampled to 400 or 512 Hz) to 16 Hz, a frequency high enough to describe the shape of breaths but optimizing the length of signals to ensure computational efficiency.

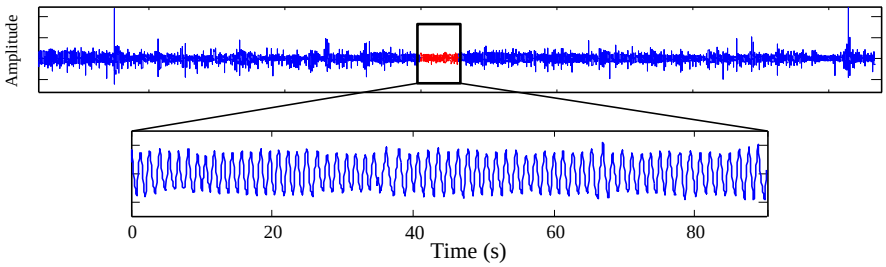

Figure VI.4 – Selection the 90-second most regular excerpt in an breathing signal.

After several tests, we found this selection criteria a good compromise between a convenient number of samples per excerpt (90 s × 16 Hz = 1500 samples), an acceptable regularity and an appropriate representativity (90 s × 5 excerpts = 450 seconds per patient). A brief description of the excerpts is depicted in Table VI.1. While longer excerpts, more reliable to calculate nonlinear measures, were more difficult to find in many patients, shorter fragments were potentially more numerous but could risk the reliability of the analysis.

	Infants	$\overline{t_{tot}}$	$\sigma(t_{tot})$	$\overline{t_i/t_e}$	$\overline{A}$	$\sigma(A)$
PHYSIDEV	27	1.05 ± 0.27	0.23 ± 0.14	0.61 ± 0.18	0.85 ± 0.63	0.16 ± 0.11
VACCIN	31	1.09 ± 0.23	0.14 ± 0.08	0.62 ± 0.19	24.6 ±16.1	3.59 ± 1.99

Table VI.1 – Characteristics of the steadiest excerpts of 90 seconds selected for nonlinear analysis.

To detect nonlinearity in each fragment we employed noise titration. To avoid false detections and over-parametrization, third order polynomials ($d = 3$) and a memory $\kappa = 5$, producing $\eta = 56$ terms, were used to

fit linear and nonlinear Volterra-Wiener models, following the recommendations given by Roulin et al. in an exhaustive analysis [44]. This procedure was computed using the Matlab code supplied by Barahona and Poon [40], which added noise in steps of 1% and repeated titration ten times to obtain reliable estimations.

The units are expressed as noise limit (NL), the percentage of noise necessary to neutralize nonlinearity. If nonlinearity was detected (NL>0), the other methods quantifying nonlinear properties were thereafter performed. To do so, the autocorrelation function (ACF) of the breathing signals were computed previously to find the time lag τ, a necessary delay to avoid temporally correlated points in periodic signals and capture nonlinear dynamics appropriately. As mentioned in previous chapter, l is chosen to be the first minimum of the ACF [27]. Another previous step was to find the embedding dimension m –necessary to unfold the attractor in the correlation integral– using the false nearest neighbor technique [26]. In summary, the nonlinear measures were computed as follows:

- Sample entropy (SpEn). Because of the quasi-periodicity of signals, to compute $SpEn(\tau)$ the time lag τ was set from 1 to l. The tolerance (r) was set to the 20% and the pattern size $m = 2$. The final value of the entropy was the mean of $SpEn(\tau)$.
- Correlation dimension (D_2). This value was the average correlation dimension using the lag l and embedding dimensions from m to $m + 5$.
- Average mutual information (AMI), expressed in bits. It was computed by introducing τ from 1 to l, then this function was divided by its value at $\tau = 1$ (the maximum) and averaged.
- Nonlinear indexes from entropy (NI_{SpEn}) and correlation dimension (NI_{D2}). Twenty iAAFT surrogates per excerpt were performed to obtain every index.

In Figure VI.5 we show an illustration of the computation of the above parameters.

Results and discussion

The computation of noise titration revealed that not all the selected excerpts in a patient contain nonlinearity. The mean number of excerpts having NL=0 was 1.28 ± 1.24, without significant differences between groups. Only in one patient nonlinear dynamics were absent in the five fragments. Hence, nonlinearity is in average present (or al least, detected) in the 75% of regular breathing periods.

For each infant, the mean of the nonlinear measures was obtained from the excerpts having NL>0 (see Figure VI.6). From the representation of the four variables, it can be observed that NL does not show a clear trend within groups regarding ages and weight, being its overall value about 10–20%, indicating a moderate level of nonlinearity.

However, SpEn and D_2 decreases with increasing maturative groups, meaning that the level of randomness or complexity decreases. The same phenomenon is observed in the evolution of AMI, as increasing values reveal a decrease of randomness of the underlying processes. It must be said that, although the mentioned trends are noticeable by inspecting the boxes, not all of these differences are statistically significant and interpretations should be done carefully.

According to our results, it can be elucidated that younger and low weight infants have higher levels of randomness governing the breathing oscillator, so complexity in respiration decreases with maturation during the preterm period. To reinforce these results, the nonlinear indexes were also computed (see Figure VI.7). However,

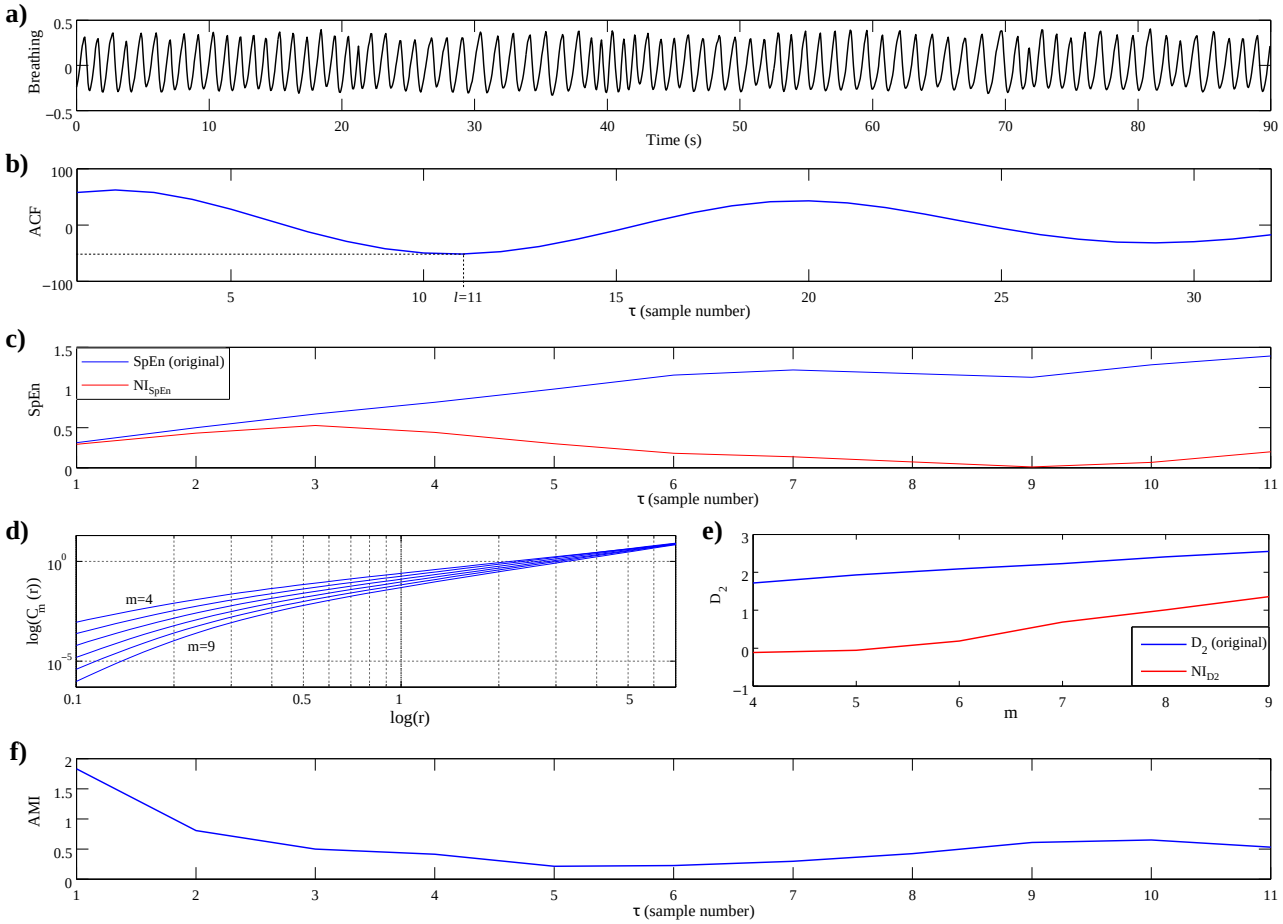

Figure VI.5 – Example of computing nonlinear measures in a respiratory signal. a) Regular 90-second excerpt. b) Autocorrelation function, where the first minimum is used as lag to compute the other measures ($l = 11$). c) Sample entropy and its nonlinear index across the delays 1 to l. SpEn and NI_{SpEn} were obtained by averaging the l values. d) Correlation integrals obtained by using $l = 11$ and embedding dimensions m from 4 to 9. e) Values of D_2 and its nonlinear index (NI_{D2}) across embeddings, averaged to obtain the final values. f) Average mutual information computed from delays 1 to l, subsequently normalized by its maximum value and averaged to obtain AMI.

we only consider valid NI_{SpEn} because it yields positive values, which confirms that nonlinearity in surrogates is lower. On the other hand, NI_{D2} produced negative values, in contradiction with the overall trends. The irregular behavior obtaining D_2 in surrogates was already reported by Small et al. [42], who suggested that the degradation of signals generated by surrogate-data perturbs the estimations of the correlation dimension, a highly sensitive technique. Consequently, we discard the use of NI_{D2} to measure nonlinearity.

Stating that breathing complexity decreases with maturation seems incongruous with the brain development, which shows an increasingly complex EEG activity as the infant maturates (see Section 5.3). However, some works analyzing the early evolution of the autonomic nervous system already reported an irregular complexity behavior, as for instance, the study of Mrowka et al. [45], where the complexity of HRV in healthy term infants measured with the largest Lyapunov exponent increased in the fist five postnatal days, then decreased dramatically remaining low in the following two months and finally increased again progressively from the third month. Regrettably, in healthy preterm newborns, we did not find specific works assessing complexity measures

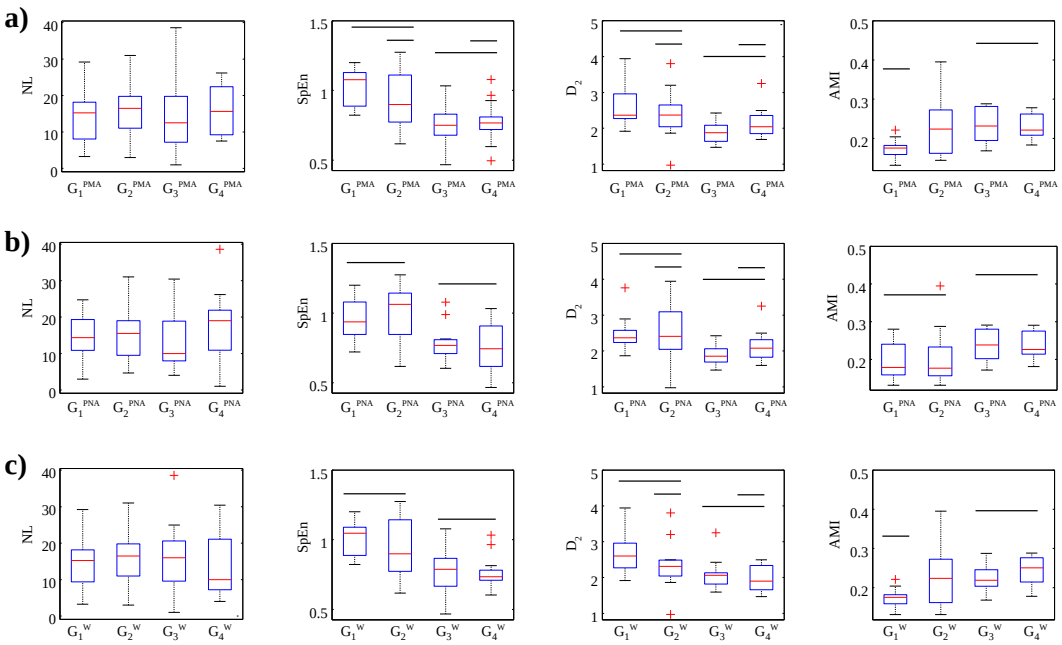

Figure VI.6 – Nonlinear measures computed in the most regular excerpts of 90 seconds. The represented values are the averages of the five excerpts per patient. Maturity groups according PMA (a), PNA (b) and weight (c). Horizontal lines of different heights indicate significant differences in a Mann-Whitney U test. Complementary data about these measures is available in Appendix D (means ± std).

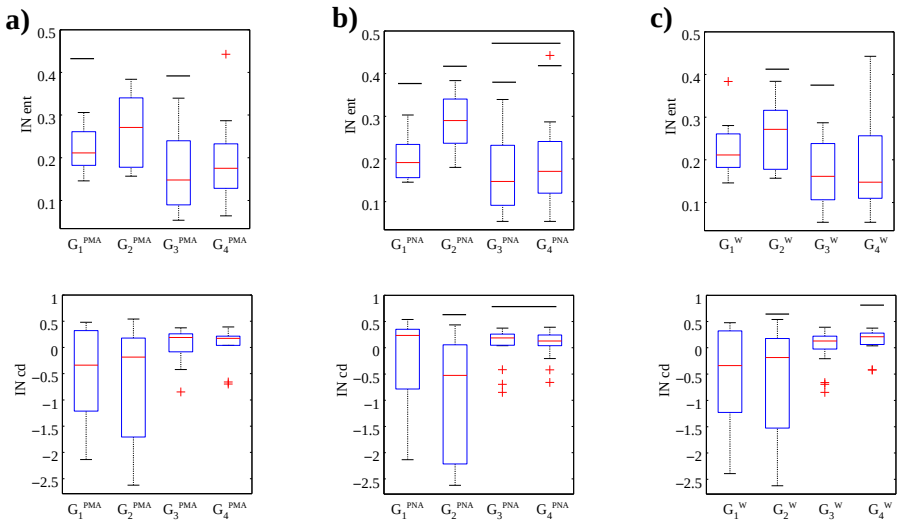

Figure VI.7 – Nonlinear indexes obtained by computing the difference of SpEn in original signals and a set of 20 surrogates. Values over zero denote a positive detection of nonlinearity, also confirmed by noise titration. In a) groups assorted by PMA, in b) by PNA and c) by weight. Complementary data about these measures is available in Appendix D (means ± std).

during maturation.

Concerning breathing activity, a similar valley-like evolution (increase-decrease-increase) during postnatal life was observed in piglets analyzing the vagal activity –responsible for both breathing rate and lung tidal volume–

with approximate entropy [46, 47]. However, our results only evidenced this kind of evolution in NI_{SpEn}.

It is clear then, that there is not a straight line development for the parameters describing the functional maturation of extrauterine life. Moreover, prematurity adds more difficulty to this scenario. During gestation, the maturity of the ANS follows a constant evolution evidenced by an increase of the complexity during the 20th to 36th week [48], but preterm birth interrupts this process. The incomplete lung function and the exaggerated baroreflex sensitivity alter cardiovascular responses [49], implying an irregular evolution of the complexity exhibited by breathing.

In the mentioned preliminary study [43], the weak increase of nonlinearity regarding age and weight was found analyzing younger patients from the PHYSIDEV database, only including one excerpt of 60 or 120 seconds per record. However this could be a serious limitation knowing the irregular behavior of immature breathing, so increasing trends might have been serendipitous. Adding more breathing excerpts and more infants, the present analysis revealed opposed trends. Thus, further investigation is needed to improve the conditions of the analysis and the availability of horizontal records to verify the maturity indexes within individuals.

We are also aware that the short length of data is a serious limiting condition to estimate accurately the nonlinear measures. Other works dealing with short respiratory data in animals obtained more conclusive results since they exploited the vagal nerve activity obtained invasively [7, 46]. This quality of data is not realistic in the NICU, so alternative manners of exploiting the respiratory activity should be found to study maturation.

5.2 Nonlinear analysis of variability signals

Unlike temporal signals, containing both information about the amplitude and frequency, RV data only represents the respiratory rhythm. But this apparent limitation constitutes also their main advantage, because changes in the temporal trace dues for instance, to postural or impedance changes, are irrelevant to compute the respiratory rhythm. In the following lines we describe the methodology employed to analyze these data and verify if they could be more useful than the temporal signals to quantify maturation.

State of the art

It is well known that the rhythms produced by the immature control centers in the brain are unstable, leading to disturbances in the cardio-respiratory regulation such as apnea-bradycardia, respiratory distress or periodic breathing. To assess the underlying dynamics of these observed rhythms, a number of mathematical tools have been proposed. Entropy, complexity and long-range dependence are some of the most commonly used.

The complexity in the heartbeat ratio of preterm infants is a well studied issue. It has been found to be decreased versus term infants [50] and to be progressively increasing with maturation [51]. Analyzed with the numerical noise titration, HRV from preterm newborns was also found useful to distinguish REM different sleep states [52].

However, the characterization of preterm's RV signals using nonlinear techniques remains less explored. Engoren et al. [53] studied the entropy in relatively short RV signals, concluding that while some specific techniques increased significantly with age and weight, SpEn and ApEn did not. Using longer series, Frey et al.[54] found differences analyzing power law distributions in preterm and term infants, having the latter significantly higher slopes in a log scale. In term infants, more authors explored nonlinearity, as for instance, Terrill et al. [55] who characterized inter-breath intervals with recurrence plots, or Cernelc et al. [56], who reported long-range dependent (LRD) properties in RV data, albeit any associations with maturative variables

were found significant. Furthermore, LRD in the respiratory rhythm might be related to sickness. As we observed in a previous study [57], patients with declared late onset sepsis presented a loss of fractal organization in long recordings of RV signals compared to healthy preterm infants.

Knowing that complexity in respiratory rhythms may be related to maturity and/or sickness, in the present section we will carry out a specific analysis for each purpose.

Selection of data and parameter tuning

From the same cohort of infants analyzed previously, we selected first the three most regular 10-minute excerpts. According to other works analyzing RV data, ten minutes is an appropriate window length (about 600 samples) for newborns [55, 56]. Moreover, we selected the five most regular 5-minute excerpts (approximately 300 samples) to compare our results with [53].

Finding indicators of maturity in variability signals differs from the strategy employed in temporal signals. RV signals only contain discrete information of time duration of cycles, so their properties are substantially different. With regard to their exploitation using nonlinear methods, we focus on two main constraints:

1. RV signals appear as random and aperiodic. Chaos-based methods relying in quasi-periodic properties of data (as correlation dimension) cannot be used. If lags need to be applied, the appropriate value for discrete, aperiodic data is 1 [55].

2. RV data are notably shorter for the same time of observation. This requires to analyze longer excerpts and employ nonlinear methods robust to short data, but not too long to avoid including different sleep states.

In view of the first constraint, we chose suitable methods: the numerical noise titration, the sample entropy (and its associated nonlinear index) and the Lempel-Ziv complexity.

To find NL, we first performed a resampling a of RV data to 2 Hz as done by Vandeput et al [52], then it was computed using the same degree and memory of Volterra-Wiener polynomials than in previous section ($d = 3, \kappa = 5$).

SpEn was applied directly in RV data with $\tau = 1$, $r = 0.2$ and $m = 2$ and NI_{SpEn} by performing twenty surrogates. The Lempel-Ziv complexity required previously the transformation of RV signals in binary coded sequences. To minimize dependences on quantification criteria, we simply coded as '1' an increase of the respiratory rhythm ($x_{i+1} > x_i$) and decreasing or unchanging rates as '0' ($x_{i+1} \leq x_i$), as done by Ferrario et al. [58] to find algorithmic complexity in HRV data.

Results and discussion

Tests applying nonlinear measures in RV data revealed some disconcerting results if they are compared to those obtained by the analyses of temporal signals. We expected that both signals could have similar trends regarding the same nonlinear measures, but it turned out that for the sample entropy, a priory robust, the evolutions with maturity were antagonistic. In RV data *SpEn* showed a significantly increasing tendency with age (see Figure VI.8) in five-minute excerpts and (less meaningfully) in ten-minute excerpts, but in temporal signals *SpEn* (see Figure VI.6) was clearly decreasing with maturity.

However, the evolutions of NI_{SpEn} from both RV and temporal data (see Figure VI.7b) showed results more in agreement. The noise titration evidenced higher noise limits for RV signals and a slight tendency to decrease

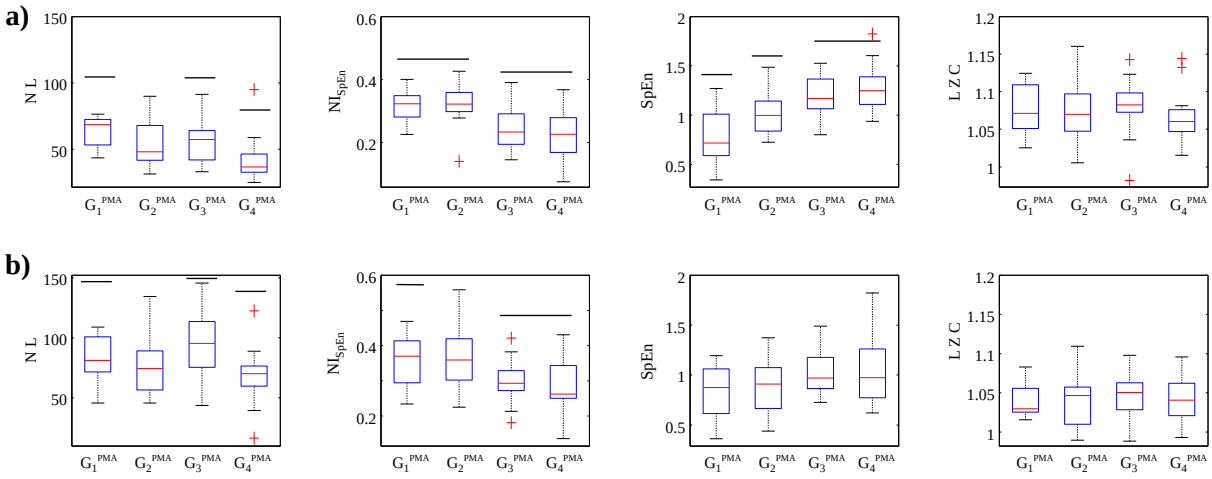

Figure VI.8 – Results from the nonlinear analysis of RV data. a) Five-minute excerpts b) Ten-minute excerpts.

with age in an irregular manner, whereas in temporal signals the overall NL was lower without clear trends. Similarly, the LZ complexity did not evidence differences by comparing age and weight.

It is worth pointing out that the analyses of five and ten-minute windows were almost reproducible, a fact that gives consistency to the measures because the signals analyzed in each test does not contain necessarily the same breathing periods. Moreover, the increasing measures of complexity by means of *SpEn* are in accordance to the observations of Engoren et al. [53] using other entropy measures. Therefore, it could be speculated that RV measures are more reliable than those obtained from temporal signals.

The reasons explaining the different evolutions in RV and temporal data could be diverse. As empathized before, the analyzed signals in the two tests are not the same, besides having different time durations. In such sensitive measures, probably the length of the analyzed signals may result insufficient to fully characterize nonlinearity or, eventually, the methodology to select a number of separated, short excerpts could be misleading.

5.3 Nonlinear analysis of EEG bursts

Although this chapter only deals with breathing signals, we show here a quick application to test if nonlinear tools could provide additional information for the assessment of the brain activity (see Section 5.2 in Chapter IV).

For instance, taking advantage of the binary coding of inter-burst interval signals, we analyzed the complexity of '1' and '0' sequences by means of the Lempel-Ziv method. We recall that the compared groups only concern VACCIN patients, thus the range of ages and weights are more reduced than in the other analyses performed in this chapter. As shown in Figure VI.9, some changes can be observed within different sleep states and maturation groups.

The most relevant finding is that in quiet sleep, the complexity of burst patterns increase significantly with maturation. In non-quiet sleep a small augmentation is observed, but not statistically significant. These results seems in concordance with other works assessing the complexity of brain activity during maturation [59, 60].

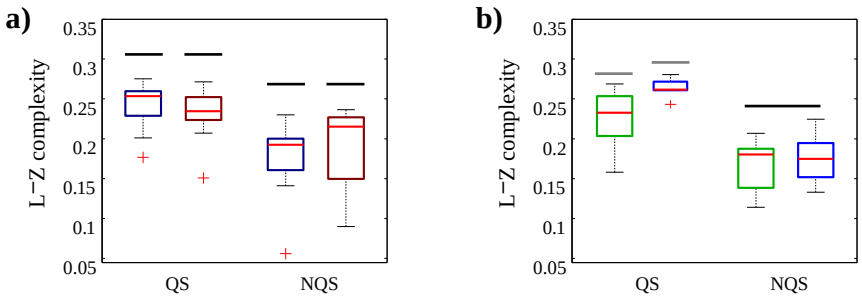

Figure VI.9 – Example of applying Lempel-Ziv complexity to burst/inter-burst coded sequences. a) Blue and red boxes are, respectively, sequences before and after vaccination. b) Green boxes correspond to G_1^{PMA} and blue boxes to G_2^{PMA}.

6 Conclusion

In this chapter, we employed nonlinear measures to study the underlying dynamics in breathing signals to find possible indicators of maturity. We therefore analyzed periods having the smallest perturbations (quiet sleep) to best capture the nonlinearity generated by the respiratory oscillators. We believe that short, clean excerpts would suffice to capture the underlying dynamics associated to these oscillators and hence robust methods such as noise titration and sample entropy have been employed.

The breathing rhythm, generated by the autonomic nervous system, has been found to contain nonlinearity, however interpreting the underlying dynamics is less obvious because during the firsts weeks of extra-uterine life the infants have many challenges, requiring a rapid adaptability to the external conditions. This might be the reason why higher complexity (and thus, better adaptability) is found in healthy newborns being 3 – 10 days old.

Nonetheless, nonlinear measures account for two main limitations. First, the existence of noise, at several levels, perturbing the results and restricting the length of analysis. The second one is the lack of horizontal databases, indispensable to corroborate the observations in different individuals. Other nonlinear methods (Lyapunov exponents or alternative entropy measures) should be also considered to help in the understanding of the breathing's complexity evolution. It is well known that, as a result of disrupted development following preterm birth, the cardio-respiratory control becomes unstable, but how this instability is reflected in nonlinear measures during maturity is much less documented.

Finally, the application concerning the measure of complexity in inter-burst interval signals showed an increase of complexity with maturation, opening a door to continue, in future works, the investigation of nonlinear methods applied to the brain activity.

Bibliography

[1] G. P. Williams. *Chaos theory tamed*. Joseph Henry Press, Washington, D.C. (1997).

[2] C. E. Shannon and W. Weaver. *A Mathematical Theory of Communication*. University of Illinois Press, Champaign, IL, USA (1963).

[3] J. Xu, Z.-r. Liu, Ren Liu, and Q.-f. Yang. *Information transmission in human cerebral cortex*. Physica D: Nonlinear Phenomena **106**(3), 363–374 August (1997).

[4] J. Jeong, J. C. Gore, and B. S. Peterson. *Mutual information analysis of the EEG in patients with alzheimer's disease*. Clinical Neurophysiology **112**(5), 827–835 May (2001).

[5] D. Abásolo, J. Escudero, R. Hornero, C. Gómez, and P. Espino. *Approximate entropy and auto mutual information analysis of the electroencephalogram in alzheimer's disease patients*. Medical & biological engineering & computing **46**(10), 1019–1028 October (2008).

[6] B. Julitta, M. Vallverdu, U. Melia, N. Tupaika, M. Jospin, E. Jensen, M. Struys, H. Vereecke, and P. Caminal. Auto-mutual information function of the EEG as a measure of depth of anesthesia. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society,EMBC*, pages 2574–2577 (2011).

[7] R. R. Dhingra, F. J. Jacono, M. Fishman, K. A. Loparo, I. A. Rybak, and T. E. Dick. *Vagal-dependent nonlinear variability in the respiratory pattern of anesthetized, spontaneously breathing rats*. Journal of Applied Physiology (Bethesda, Md.: 1985) **111**(1), 272–284 July (2011).

[8] R. A. Katz, M. S. Lawee, and A. K. Newman. Method and apparatus for diagnosing sleep breathing disorders, June (2003). U.S. Classification: 600/513; 600/529 International Classification: A61B/50472.

[9] J. S. Richman and J. R. Moorman. *Physiological time-series analysis using approximate entropy and sample entropy*. American Journal of Physiology - Heart and Circulatory Physiology **278**(6), H2039 –H2049 June (2000).

[10] H. V. Huikuri, J. S. Perkiömäki, R. Maestri, and G. D. Pinna. *Clinical impact of evaluation of cardiovascular control by novel methods of heart rate dynamics*. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **367**(1892), 1223–1238 April (2009).

[11] D. E. Lake, J. S. Richman, M. P. Griffin, and J. R. Moorman. *Sample entropy analysis of neonatal heart rate variability*. American Journal of Physiology. Regulatory, Integrative and Comparative Physiology **283**(3), R789–797 September (2002).

[12] A. Beuchée, G. Carrault, J. Y. Bansard, E. Boutaric, P. Bétrémieux, and P. Pladys. *Uncorrelated randomness of the heart rate is associated with sepsis in sick premature infants*. Neonatology **96**(2), 109–114 (2009).

[13] S. M. Pincus, I. M. Gladstone, and R. A. Ehrenkranz. *A regularity statistic for medical data analysis*. Journal of Clinical Monitoring **7**(4), 335–345 October (1991). PMID: 1744678.

[14] F. Kaffashi, R. Foglyano, C. G. Wilson, and K. A. Loparo. *The effect of time delay on approximate & sample entropy calculations*. Physica D: Nonlinear Phenomena **237**(23), 3069–3074 December (2008).

[15] G. J. Chaitin. *Algorithmic Information Theory*. Cambridge University Press December (2004).

[16] A. Lempel and J. Ziv. *On the complexity of finite sequences*. IEEE Transactions on Information Theory **22**(1), 75–81 (1976).

[17] T. A. A. Watanabe, C. J. Cellucci, E. Kohegyi, T. R. Bashore, R. C. Josiassen, N. N. Greenbaun, and P. E. Rapp. *The algorithmic complexity of multichannel EEGs is sensitive to changes in behavior.* Psychophysiology **40**(1), 77–97 January (2003).

[18] X. S. Zhang and R. J. Roy. *Derived fuzzy knowledge model for estimating the depth of anesthesia.* IEEE transactions on bio-medical engineering **48**(3), 312–323 March (2001).

[19] N. Radhakrishnan and B. N. Gangadhar. *Estimating regularity in epileptic seizure time-series data. a complexity-measure approach.* IEEE engineering in medicine and biology magazine: the quarterly magazine of the Engineering in Medicine & Biology Society **17**(3), 89–94 June (1998).

[20] T. Cover and T. M. . *Elements of information theory.* Wiley-Interscience, Hoboken, N.J. (2006).

[21] T. Schreiber and H. Kantz. *Noise in chaotic data: Diagnosis and treatment.* Chaos **5**, 133–142 (1995).

[22] O. Rössler. *An equation for continuous chaos.* Physics Letters A **57**(5), 397–398 July (1976).

[23] E. Ott. *Chaos in Dynamical Systems.* Cambridge University Press August (2002).

[24] F. Takens. Detecting strange attractors in turbulence. In D. Rand and L.-S. Young, editors, *Dynamical Systems and Turbulence, Warwick 1980*, number 898 in Lecture Notes in Mathematics, pages 366–381. Springer Berlin Heidelberg January (1981).

[25] W. Liebert, K. Pawelzik, and H. G. Schuster. *Optimal embeddings of chaotic attractors from topological considerations.* EPL (Europhysics Letters) **14**(6), 521 March (1991).

[26] M. B. Kennel, R. Brown, and H. D. I. Abarbanel. *Determining embedding dimension for phase-space reconstruction using a geometrical construction.* Physical Review A **45**(6), 3403–3411 March (1992).

[27] H. S. Kim, R. Eykholt, and J. D. Salas. *Nonlinear dynamics, delay times, and embedding windows.* Phys. D **127**(1-2), 48–60 March (1999).

[28] P. Grassberger and I. Procaccia. *Characterization of strange attractors.* Physical Review Letters **50**(5), 346–349 January (1983).

[29] J. Theiler. *Estimating fractal dimension.* Journal of the Optical Society of America A **7**(6), 1055–1073 June (1990).

[30] C. Diks. *Estimating invariants of noisy attractors.* Physical Review E **53**(5), R4263–R4266 May (1996).

[31] D. Yu, M. Small, R. G. Harrison, and C. Diks. *Efficient implementation of the gaussian kernel algorithm in estimating invariants and noise level from noisy time series data.* Physical Review E **61**(4), 3750–3756 April (2000).

[32] C.-S. Poon and M. Barahona. *Titration of chaos with added noise.* Proceedings of the National Academy of Sciences **98**(13), 7107–7112 June (2001).

[33] A. Serletis, A. Shahmoradi, and D. Serletis. *Effect of noise on estimation of lyapunov exponents from a time series.* Chaos, Solitons & Fractals **32**(2), 883–887 April (2007).

[34] U. S. Freitas, C. Letellier, and L. A. Aguirre. *Failure in distinguishing colored noise from chaos using the noise titration technique.* Physical Review E **79**(3), 035201 March (2009).

[35] U. Freitas, E. Roulin, J.-F. Muir, and C. Letellier. *Identifying chaos from heart rate: The right task?* Chaos: An Interdisciplinary Journal of Nonlinear Science **19**(2), 028505 (2009).

[36] M. Wysocki, M. N. Fiamma, C. Straus, C. S. Poon, and T. Similowski. Chaotic dynamics of ventilatory flow in humans. In *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, pages 759–762. IEEE January (2006).

[37] M.-N. Fiamma, C. Straus, S. Thibault, M. Wysocki, P. Baconnier, and T. Similowski. *Effects of hypercapnia and hypocapnia on ventilatory variability and the chaotic dynamics of ventilatory flow in humans.* American Journal of Physiology - Regulatory, Integrative and Comparative Physiology **292**(5), R1985 –R1993 May (2007).

[38] L. Mangin, G. Lesèche, A. Duprey, and C. Clerici. *Ventilatory chaos is impaired in carotid atherosclerosis.* PLoS ONE **6**(1), e16297 January (2011).

[39] S. Vandeput, B. Verheyden, A. Aubert, and S. Van Huffel. Nonlinear heart rate variability in a healthy population: Influence of age. In *Computers in Cardiology, 2008*, pages 53–56 (2008).

[40] M. Barahona and C.-S. Poon. *Detection of nonlinear dynamics in short, noisy time series.* Nature **381**(6579), 215–217 May (1996).

[41] M. Small and K. Judd. *Comparisons of new nonlinear modeling techniques with applications to infant respiration.* (1998).

[42] M. Small, K. Judd, M. Lowe, and S. Stick. *Is breathing in infants chaotic? dimension estimates for respiratory patterns during quiet sleep.* Journal of Applied Physiology **86**(1), 359–376 January (1999).

[43] X. Navarro, F. Porée, A. Beuchée, and G. Carrault. *Assessing nonlinear properties in breathing signals from preterm infants.* International Journal of Bioelectromagnetism **15**(1), 102–108 (2013).

[44] E. Roulin, U. S. Freitas, and C. Letellier. *Working conditions for safe detection of nonlinearity and noise titration.* Physical Review E **83**(4), 046225 April (2011).

[45] R. Mrowka, A. Patzak, E. Schubert, and P. B. Persson. *Linear and non-linear properties of heart rate in postnatal maturation.* Cardiovascular Research **31**(3), 447–454 March (1996).

[46] M. Akay, T. Lipping, K. Moodie, and P. Hoopes. *Effects of hypoxia on the complexity of respiratory patterns during maturation.* Early Human Development **70**(1-2), 55–71 December (2002).

[47] M. Akay. *Influence of the vagus nerve on respiratory patterns during early maturation.* IEEE Transactions on Biomedical Engineering **52**(11), 1863–1868 (2005).

[48] P. Van Leeuwen, S. Lange, H. Bettermann, D. Grönemeyer, and W. Hatzmann. *Fetal heart rate variability and complexity in the course of pregnancy.* Early Human Development **54**(3), 259–269 April (1999).

[49] N. B. Witcombe, S. R. Yiallourou, S. A. Sands, A. M. Walker, and R. S. C. Horne. *Preterm birth alters the maturation of baroreflex sensitivity in sleeping infants.* Pediatrics **129**(1), e89–96 January (2012).

[50] F. A. Selig, E. R. Tonolli, E. V. C. M. d. Silva, and M. F. d. Godoy. *Heart rate variability in preterm and term neonates.* Arquivos Brasileiros de Cardiologia **96**(6), 443–449 June (2011).

[51] L. A. Lipsitz, S. M. Pincus, R. J. Morin, S. Tong, L. P. Eberle, and P. M. Gootman. *Preliminary evidence for the evolution in complexity of heart rate dynamics during autonomic maturation in neonatal swine.* Journal of the Autonomic Nervous System **65**(1), 1–9 July (1997).

[52] S. Vandeput, G. Naulaers, H. Daniels, and S. Van Huffel. *Heart rate variability during REM and non-REM sleep in preterm neonates with and without abnormal cardiorespiratory events.* Early Human Development **85**(10), 665–671 October (2009).

[53] M. Engoren, S. E. Courtney, and R. H. Habib. *Effect of weight and age on respiratory complexity in premature neonates.* Journal of Applied Physiology **106**(3), 766–773 March (2009).

[54] U. Frey, M. Silverman, A. L. Barabási, and B. Suki. *Irregularities and power law distributions in the breathing pattern in preterm and term infants.* Journal of Applied Physiology (Bethesda, Md.: 1985) **85**(3), 789–797 September (1998).

[55] P. I. Terrill, S. J. Wilson, S. Suresh, and D. M. Cooper. *Characterising infant inter-breath interval patterns during active and quiet sleep using recurrence plot analysis.* Annual International Conference of the IEEE Engineering in Medicine and Biology **2009**, 6284–6287 (2009).

[56] M. Cernelc, B. Suki, B. Reinmann, G. L. Hall, and U. Frey. *Correlation properties of tidal volume and end-tidal o2 and CO2 concentrations in healthy infants.* Journal of Applied Physiology (Bethesda, Md.: 1985) **92**(5), 1817–1827 May (2002).

[57] X. Navarro, F. Porée, A. Beuchée, and G. Carrault. Respiration signal as a promising diagnostic tool for late onset sepsis in premature newborns. In *Computing in Cardiology, 2010*, pages 947–950. IEEE September (2010).

[58] M. Ferrario, M. G. Signorini, and S. Cerutti. *Complexity analysis of 24 hours heart rate variability time series.* Annual International Conference of the Engineering in Medicine and Biology Society **2**, 3956–3959 September (2004).

[59] E. Biagioni, M. Frisone, S. Laroche, B. Kapetanakis, D. Ricci, M. Adeyi-Obe, H. Lewis, N. Kennea, G. Cioni, F. Cowan, M. Rutherford, D. Azzopardi, and E. Mercuri. *Maturation of cerebral electrical activity and development of cortical folding in young very preterm infants.* Clinical Neurophysiology **118**(1), 53–59 January (2007).

[60] D. Zhang, H. Ding, Y. Liu, C. Zhou, H. Ding, and D. Ye. *Neurodevelopment in newborns: a sample entropy analysis of electroencephalogram.* Physiological measurement **30**(5), 491–504 May (2009).

Chapter VII

Measuring long-range dependence

In last chapter, we employed nonlinear methods based on complexity measures and chaos estimations in short and clean signals, attempting to quantify the underlying dynamics that characterize respiratory center's oscillators. However, these measures are limited by their high sensitivity to noise, making difficult their application on long data, indispensable to study infection processes.

Indeed, abnormalities provoked by sickness require long-term data to observe unbalanced homeostatic responses in changing conditions such as sleep transitions or environmental changes. In this context, the study of long-range dependence (LRD) as an inherent property of integrative control systems [1], has become an issue of interest in clinical investigations since a loss of fractal organization is often related to disease and aging [2].

Therefore, this chapter focuses on the LRD property in breathing variability signals from preterm infants to study its potential to detect neonatal sepsis. After giving the background of long-range dependent processes, we perform a thorough analysis on artificial and surrogate data to find the best method to estimate the Hurst exponent, a parameter characterizing this property. Then, we apply the best methods on real data to verify their utility to detect infection.

1 Long-range dependence and the Hurst exponent

Long-range dependence or long-range memory is a property naturally present in many physical phenomena. It is characterized by self-similar (or fractal) behavior, meaning that similar statistical properties are preserved at different scale levels, which are related by a constant, the Hurst exponent [3]. By contrast, short-range dependence (SRD) exhibits statistical similarity only at short scales.

LRD was firstly reported in hydrology [4] and subsequently used in geophysics, econometrics, network traffic and biology and, as well as its diverse application fields, its formalism is discussed from a variety of points of view. However, the approach based on the second-order properties of stochastic processes is probably the most commonly accepted in the literature.

1.1 Definition of the Hurst exponent

Let X_i be a stationary stochastic process with $i = 1, 2, ..$ the discrete sequence of observations. Let μ, $\gamma(k)$, $\rho(k)$ $(k \geq 0)$ denote the mean, the auto-covariance and the autocorrelation, and $\sigma^2 = \gamma(0)$ the variance of the process. X_i is then considered self-similar if it has an autocorrelation of the form:

$$\rho(k) \sim k^{-\beta} L(i), \quad \text{as } k \to \infty, \tag{VII.1}$$

where $\beta \in [0,1]$, and L is a slowly varying function at infinity, i.e.:

$$\lim_{i\to\infty} \frac{L(ix)}{L(i)} = 1 \quad \text{for any } x > 0.$$

Given m a positive scalar representing a scale greater than 1 ($X_k^{(1)} = X$), a new data-series, the mean aggregated stochastic process, is obtained for every scale:

$$X_k^{(m)} = \frac{1}{m} \sum_{i=km}^{(k+1)m-1} X_i, \tag{VII.2}$$

with an autocorrelation denoted $\rho^{(m)}(k)$ at each scale m.

The process X is called second-order (exactly) self-similar with a characteristic Hurst exponent ($H = 1-\beta/2$) if the variance and autocorrelation are the same at each scale, so for all $m = 1, 2, 3, ..,$

$$\text{var}[X^{(m)}] = \sigma^2 m^{2H-2}, \tag{VII.3}$$

$$\rho^{(m)}(k) = \rho(k) = \frac{1}{2}[(k+1)^{2H} + (k-1)^{2H}] - k^{2H}. \tag{VII.4}$$

In finite processes (with k large enough), X is defined asymptotically self-similar if $\rho^{(m)}(k) \to \rho(k)$ as $m \to \infty$. Self-similar processes are then scale-invariant, i.e. autocorrelations at low scale levels (aggregate processes $X_k^{(m)}$) are similar at higher scales (X).

The long-range dependence property is given by the fact that autocorrelation is non-summable, $\sum_k \rho(k) = \infty$, and described by a slow, positive decaying function. The Hurst exponent ranges from 0 to 1 but LRD, which describes natural phenomena, implies $H > 0.5$. A value equal to 0.5 means that the process is completely uncorrelated or random (white noise) and values under 0.5 describe processes with negative correlations or with an anti-persistent behavior.

1.2 *H* estimation techniques

A considerable number of methods to estimate H, providing different approaches to quantify the self-similar behavior, have been proposed and analyzed in the literature. In this work, we consider the detrended fluctuation analysis (DFA), the discrete wavelet transform-based (DWT) –already employed in respiratory signals– and the least squares based on standard deviation (LSSD), a method used in hydrology yet unexplored in biomedical data.

Detrended fluctuation analysis

DFA, introduced by Peng et al. [5] to estimate long-range dependence in non-stationary signals, has already been used to quantify the fractal content in RV series from adults [6] and infants [7]. The data-series X_n of length N are first integrated and then divided into blocks of equal size m. A least squares line, representing the local trend, is fitted to the data in each block. The y coordinate of the fitted line, $y_m(k)$, is then subtracted to the integrated series $y(k)$ to remove the trend in each box. Next, $F(m)$, the root-mean-square fluctuation of this

integrated and detrended time series is calculated:

$$F(m) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}[y(k) - y_m(k)]^2}. \tag{VII.5}$$

This function is repeated successively over all time scales (box sizes) to characterize its relationship with the box size m. A power law, given by $F(m) \sim m^{\alpha}$ as $m \to \infty$, indicates the presence of LRD. α is the scaling exponent, a generalization of the Hurst exponent, and it is obtained by finding the slope of the line relating $F(m)$ to m in a log plot.

Discrete wavelet transform-based

Abry et Weitch [8] proposed a semi-parametric joint estimator of H based on the DWT, probed to be robust with non-stationarities even when signals contains SRD. It takes advantage of the scaling properties of the wavelet basis, which captures optimally the scaling self-similar nature of LRD processes.

Briefly, the method performs first a wavelet decomposition of a given discrete time series X_n, providing $d_x(j,k)$, the wavelet coefficients or details. Next, at each fixed octave j the details are squared and then averaged across k to produce an estimate of the variance of the wavelet coefficients, called μ_j. A plot of $\log_2(\mu_j)$ against j is done to identify the range of octaves where scaling occurs. Finally, H is computed by performing a weighted linear regression over those scales. The algorithm employed in the present work utilizes Daubechies wavelets to perform DWT and identifies automatically the octaves with scaling [9].

Least squares based on standard deviation

Classical statistics assume that a sample is constituted by independent random variables, but this is not consistent with finite processes exhibiting long range dependence. In particular, the variances of the sample mean decay slower than those from the sample size, hence important bias can be introduced if the properties of self-similarity are not taken into account [10]. It has been shown in [11] that a robust estimator of variance $\widehat{S}^2$ with known H is given by:

$$\widehat{S}^2 = \frac{n-1}{n-n^{2H-1}}S^2 = \frac{1}{n-n^{2H-1}}\sum_{i=1}^{n}(X_i - \overline{X})^2, \tag{VII.6}$$

with $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Note that when $H = 0.5$ (absence of long-term memory), the above expression becomes the classic variance estimator.

Koutsoyiannis proposed a method to estimate both H and standard deviation, σ, using modified statistics for long-range dependent processes [12]. Combining expressions (VII.1) and (VII.5), assuming that $\mathrm{E}(\widehat{S}) = \sigma$, the standard deviation of $X^{(m)}$ at each scale m could be approximated by $\widehat{S}^{(m)} \approx C_m(H)m^{H+1}\sigma$, where:

$$C_m(H) = \sqrt{\frac{n/m - (n/m)^{2H-1}}{n/m - 1/2}}. \tag{VII.7}$$

Then, the next step was to minimize the fitting error e^2 to estimate both H and σ:

$$e^2(H,\sigma) = \sum_{m=1}^{m'} \frac{[\ln \sigma^{(m)} - \ln S^{(m)}]^2}{m^p}, \tag{VII.8}$$

where m' is the maximum scale level so that $m' = n/10$ and p is a weight factor. After a Monte-Carlo study, it was found that $p = 2$ provided the best H estimation. A complete analytical and iterative procedure is detailed in the original paper [12] and a performance analysis in [13].

2 Performance analysis of *H* estimators

The Hurst exponent can be estimated by a large number of time or frequency domain methods, but their performances can differ substantially depending on the way they are defined or the context where they are applied. Given that we did not find any recommendations in the literature to compute H reliably on respiratory signals from preterm infants, this section is devoted to evaluate a few estimation techniques. The different analyses presented here led to the publication cited in [14].

2.1 State of the art

The diverse areas in which the Hurst exponent can be utilized has provided the literature with numerous comparative analyses. Basically, they prescribe H in simulated data, and then it is estimated by different methods to evaluate the error. The works of Taqqu et al. [15], and more recently Rea et al. [16], presented an extensive review of this kind of empirical study, analyzing several estimators with different lengths of time-series generated by fractional Gaussian noise (fGn) [17] and fractional autoregressive-integrated moving average (FARIMA) models [18]. Other works focusing on non-Gaussian conditions [19, 20] tested the estimators with infinite variance FARIMA series and concluded that robustness decreases when short-range dependence structures are present. The effect of non-stationarities on the estimated value of H ($\widehat{H}$) has also aroused numerous studies in view of the fact that real data often contains local trends and shifts in the mean, which are known to produce a false detection of LRD and bias [19]. In such conditions, the detrended fluctuation analysis (DFA) [5] and a wavelet-based (DWT) estimator [8] result in good performances [21, 22].

However, the models to generate simulated time-series in the mentioned works might not be realistic for some experimental data, and more specifically, variability signals from preterm infants. In a preliminary study [23], we analyzed five estimators through a basic model which approximated apneic patterns adding spikes ad hoc to fGn, finding out that DFA, DWT and LSSD showed the best performances. This section not only analyzes accurately the structure of respiratory variability series, but also formalizes more rigorous models to generate artificial signals, with the purpose of evaluating the three aforementioned methods under realistic conditions.

2.2 Data selection

We selected twenty-four breathing signals from different patients of the SEPSIS database in order to characterize the three respiratory patterns: regular, erratic and periodic. The cohort was aged between 29 and 33 weeks GA and three to ten days of postnatal age. After processing raw signals by the filtering framework and the

cycle detector described in previous chapter, 20-minute excerpts were selected by clinicians according to their predominant breathing pattern so that they were classified in three groups of eight signals.

Once variability series were obtained, we computed the mean duration of cycles, $\overline{t_{tot}}$, the standard deviation, $std(t_{tot})$, the number of apneas per hour, N_{Ap}, and the main duration of apneas, D_{Ap}, for each RV signal. Regarding the three groups of infants, these parameters showed statistically significant differences in a Mann-Whitney U-test, whereas the gestational age (GA) and the weight did not (see Table VII.1).

	Regular	Erratic	Periodic
GA (weeks)	31.9 ± 1.5	31.2 ± 0.9	31.1 ± 0.5
Weight (g)	1338 ± 353	983 ± 249	1061 ± 149
$\overline{t_{tot}}$ (s)	0.93 ± 0.17*	1.29 ± 0.22§	1.39 ± 0.18*§
$std(t_{tot})$	0.29 ± 0.08*	0.93 ± 0.25*§	1.58 ± 0.56*§
N_{Ap} (num/h)	10.6 ± 7.70*	68.2 ± 19.5*§	122.9 ± 29.7*§
D_{Ap} (s)	2.26 ± 1.42*	5.58 ± 1.31*†	7.20 ± 1.65*†
μ_L	-0.12 ± 0.17†‡	0.15 ± 0.22†	0.02 ± 0.14‡
σ_L	0.25 ± 0.06*	0.46 ± 0.02*§	0.58 ± 0.08*§

*, § paired groups with $p < 0.01$
†, ‡ paired groups with $p < 0.05$

Table VII.1 – Description and some characteristics of the selected signals of each pattern. μ_L and σ_L correspond to the mean and standard deviation of the lognormal law, respectively.

With regard to the marginal laws, the presence of apneas (spikes), variable with the individual breathing pattern, and the fact that there are no negative times, determines right skewed distributions in RV signals, which were best fitted with lognormal laws [24]. In our case, the fitted lognormal parameters between the three patterns were also significantly different. Later on, the mean and standard deviation of the distribution parameters will be employed to generate simulated signals.

2.3 Evaluation methodology

Surrogate-data tests

Surrogate-data tests are utilized to statistically infer the nature of the process generating the observed data [25]. The realization of a surrogate set, artificial signals preserving certain statistical properties of the original data, determines the formulation of the null hypothesis, H_0, which is to be tested with a statistic capable to discern original and surrogates properties. If the statistics from the surrogate set and from the empirical data are significantly different, H_0 can be rejected.

In this work, we designed two surrogate tests, each posing different null hypotheses and modifying differently the original data:

- Random-shuffled surrogates (RSS) test: Temporal correlations are destroyed (theoretically, $H = 0.5$) and spectrum is whitened in surrogate-data. The addressed null hypothesis is that LRD in empirical data is generated by uncorrelated noise. H can be used directly as discriminant statistic thanks to its ability to detect LRD.

- Small-shuffled surrogates (SSS) test: Surrogate-data have their samples shuffled in a small scale, consequently

local correlations (SRD) are lost and long-term structures preserved (H is theoretically unchanged). The null hypothesis addressed with SSS is that empirical data do not contain short term dependent processes. The auto-mutual information computed in short lags to capture short-term dynamics can be used as discriminant statistic [26].

Although the main goal of RSS and SSS tests is to demonstrate the existence of LRD and SRD, the set of artificial data generated by both methods will be used for a second purpose: to evaluate the performance of estimators measuring the bias from the theoretical values of H.

Generation of simulated signals

The utilization of surrogate-data to evaluate the estimators has the advantage of reproducing realistic conditions, but at the expenses of imposing approximated values of H on artificial signals. Therefore, to conduct more complete performance studies, simulated self-similar signals with an exact prescription of H are necessary.

In order to generate processes exhibiting LRD, several methods can be utilized. One of the most employed is the fractional Gaussian noise, a stochastic model able to synthesize exact self-similar realizations [17]. The FARIMA model is another alternative if both short-range and long-range dependence need to be introduced but, for simplicity, and knowing that the original series does not contain important SRD processes (see section 2.4), we utilized the fGn model.

Fractional Gaussian model is originated from the family of fractional Brownian motions (fBm) $B_H = \{B_H(t), t \in \mathbb{R}\}$, that are Gaussian self-similar stochastic processes with stationary increments indexed by the Hurst parameter H. So, a process $X = \{X_k, k \in \mathbb{Z}\}$, with an autocorrelation function $\rho(k)$ as equation (VII.1), is fractional Gaussian noise if $X_k = B_H(k+1) - B_H(k)$.

We employed the *circulant embedding* technique [27] to synthesize stationary, normally-distributed realizations of fGn with $H \in \{0.5, 0.55, ...0.9\}$. However empirical data are described by lognormal distributions, hence fGn need to be transformed to follow the same marginal law to test the estimators under similar conditions.

A Gaussian discrete series, $X_G(n)$ can be transformed into a log-normally distributed series $X_L(n)$ using the relation [28]:

$$X_L(n) = e^{\mu_G + \sigma_G X_G(n)}, \tag{VII.9}$$

where μ_G and σ_G, the mean and standard deviation of the normal distribution, are given by:

$$\mu_G = \ln \frac{\mu_L}{\sqrt{1+cv^2}} \quad ; \quad \sigma_G = \sqrt{\ln(1+cv^2)}, \tag{VII.10}$$

μ_L is the mean of the lognormal distribution and $cv = \sigma_G/\mu_G$ its coefficient of variation, a measure of the relative dispersion of the series. The transformation of normally distributed series to lognormal processes can be applied to fGn, obtaining fractional Lognormal noise (fLn) with approximately the same characteristic Hurst exponent [29]. Values of cv in close proximity to zero produce a highly peaked distribution with low variance, whereas greater values increase the tail. A few examples of fLn with several values of cv and H can be seen in Figure VII.1.

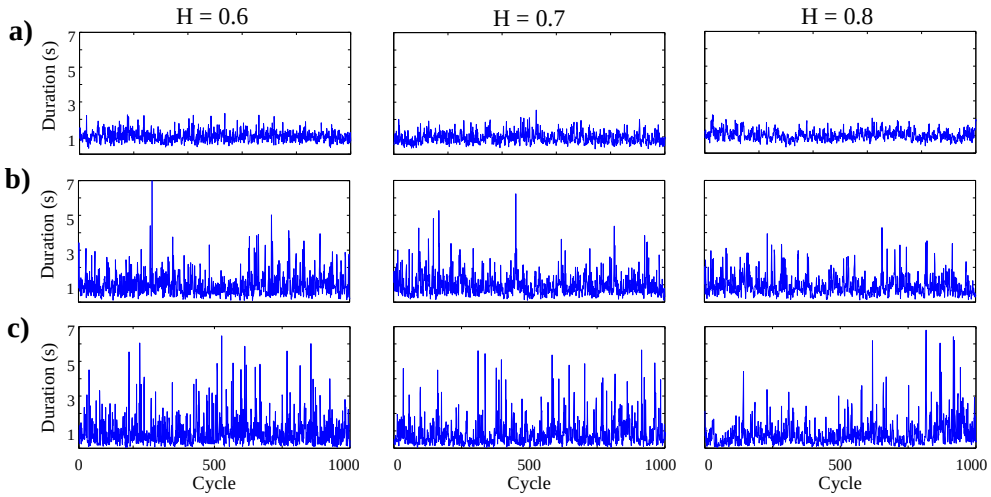

Figure VII.1 – Examples of fLn realizations with several coefficients of variation ($cv = \sigma_G/\mu_G$) at $H = 0.6, 0.7$ and 0.8. a) $cv = 0.3$ b) $cv = 0.7$ c) $cv = 1$.

Artificial signals with Hurst adjusted surrogates

As introduced before, the RSS and SSS eliminate long- and short-range dependence, respectively. But in some cases it can be useful to modify the LRD in a signal, producing surrogate-data with a prescribed value of H. This is the reason why the Hurst-adjusted surrogates (HAS) are introduced. Briefly, the main intention of this technique is to replicate the degree of correlation (governed with H) in an ordinary realization of fGn to an empirical record X. This procedure is done as follows:

1. Generate Y a realization of fGn with the desired H and the same standard deviation of X and form iy, the vector of indexes of Y.
2. Obtain Y' by sorting in ascending order Y, and get iy', a vector containing the original positions of the fGn samples.
3. Form a new series X' by sorting X, also in ascending order.
4. Reorder X' using the indexes of iy'.

These signals equally distributed and parametrized with H will be referred subsequently as Hurst-Adjusted Surrogates (HAS).

2.4 Long and short-range dependence in real RV signals

Given that LRD is detected on the real RV signals, it should be verified if 1) LRD is not caused by uncorrelated processes in RV (apneas and local trends), and if 2) RV signals do not contain strong short-term dependent dynamics, known to perturb the estimation of the Hurst exponent [19].

Hence, long- and short-range dependence need to be examined by a RSS and a SSS test. In next lines, the following notation concerning H will be utilized:

- H_r and $\widehat{H_r}$: The real value of H of an empirical signal and its estimation, respectively.

- H_s and $\widehat{H_s}$: The theoretical value of H imposed by a surrogate technique and its estimated value.

In both tests, a set of 100 surrogates was obtained from each original signal. Then, $\widehat{H_s}$ and $\widehat{H_r}$ were compared by a Mann-Whitney U-test to find possible statistically significant differences.

Results of random-shuffled surrogates tests

The RSS provided data-series with $H_s = 0.5$, meaning that LRD is eliminated by unsorting randomly the empirical data. Statistically significant differences were found between $\widehat{H_s}$ and $\widehat{H_r}$ using the three estimators (p_{DFA}, p_{DWT} and $p_{LSSD} < 0.0001$) in regular and erratic group. However, these differences were not found in the periodic group ($p_{DFA} = 0.585; p_{DWT} = 0.919; p_{LSSD} = 0.176$), presumably due to the nature of this breathing pattern, which contains less LRD and hence closer values of H_r to 0.5.

Therefore, the null hypothesis, that long-range dependence is caused by an uncorrelated process, can be rejected by regular and erratic groups, but not by the periodic group.

Results of small-shuffled surrogates tests

The SSS produced a set of signals without SRD containing approximately the same LRD ($H_s \approx H_r$). Comparing $\widehat{H_s}$ and $\widehat{H_r}$ it could be stated that SRD does not affect the estimation of LRD because no statistically significant differences were found, but the use of the statistic given by AMI, as introduced previously, is a more rigorous criterion. AMI was computed with lags from 1 to 10 to quantify the dependence between close cycles in each original signal and in a set of 100 SSS. The analysis of the 24 RV signals (see Figure VII.2) revealed that there are statistically significant differences regarding the AMI of original and surrogate populations in the closest cycles: breaths one to three in regular patterns, one and two in erratic and only the first breath ahead in periodic breathing. Therefore, the null hypothesis, that there is no short range dependence in RV signals, can be rejected only in the mentioned lags.

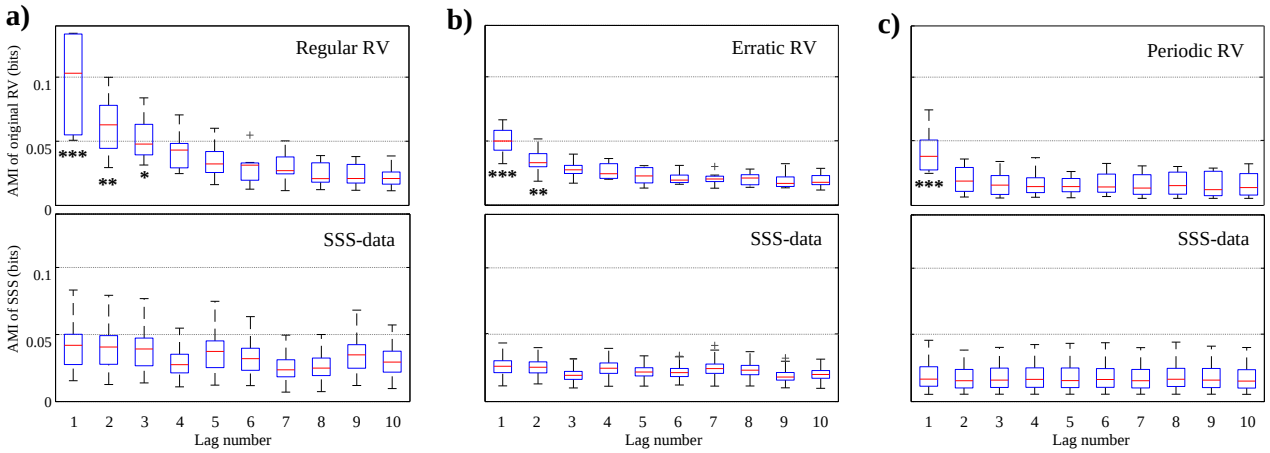

Figure VII.2 – Average Mutual Information (AMI) computed for the ten first lags in original RV series (upper panels) and small-shuffled surrogates (lower panels). There were significantly significant differences in the first three breaths for regular breathing (a), in the first two breaths for erratic (b) and in the first breath for erratic patterns (c) after performing a Mann-Whitney U-test ($***, p < 0.001; **, p < 0.005; *, p < 0.05$).

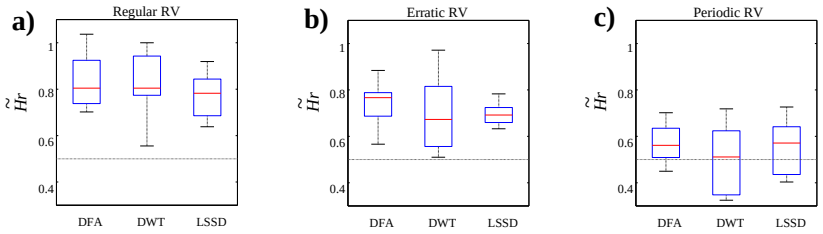

Figure VII.3 – Illustration of the discrepancy within the estimations of the Hurst exponent ($\widehat{H_r}$). Regular RV (a) showed in general the highest LRD, followed by erratic (b) and periodic (c) patterns. In all cases, $\widehat{H_r}$ in periodic RV presented statistically significant differences ($p < 0.05$) with regard to regular and erratic breathing.

2.5 Robustness of estimators

As can be observed in Figure VII.3, the meaningful differences regarding $\widehat{H_r}$ poses a dilemma to choose the most appropriate estimator in each pattern. The following tests elucidate and compare the robustness of DFA, DWT and LSSD employing artificial data generated by surrogate techniques and simulated data generated by the fLn model.

Results in surrogate-data

The knowledge of H_s in surrogate-data allowed to carry out a first evaluation of the estimators with signals identically distributed than the real ones. The robustness was measured by the error of estimations, ϵ_{sur}, defined as:

$$\epsilon_{sur} = H_s - \widehat{H_s}. \tag{VII.11}$$

We used the set of 100 RSS ($H_s = 0.5$) and SSS ($H_s \approx H_r$) generated in previous tests to compute ϵ_{sur} on the 3 groups of RV (see first and second columns in Figure VII.4). The sample size of surrogate-data, as real RV, is comprised approximately between 500 and 1000 samples due to the variations on breathing rates and apneas. In view of the boxplots dispersion and bias from the middle lines ($\epsilon_{sur} = 0$), DWT is the least accurate and least precise technique, whereas LSSD exhibits very small bias and the least scattered interquartile ranges (IQRs).

The HAS method was employed to obtain surrogates with different levels of LRD, setting H_s from 0.6 to 0.9 in steps of 0.05. The evaluation of the robustness in this case was measured by the mean surrogate error, $\overline{\epsilon_{sur}}$, computed as the mean of the different ϵ_{sur} obtained by 100 realizations of HAS for each value of H_s (see third column in Figure VII.4). The boxplots show, in general, more dispersed and positive-biased IQRs, revealing that the presence of a stronger LRD perturbs the estimations. Of note, only LSSD has its IQRs below the simulation step ($\Delta H_t = 0.05$) in all patterns, despite having the greatest bias.

Tests with fLn

Prior to evaluating the estimation errors, we first verified if the apneas generated by the fLn model are realistic. One hundred fLn series of 1024 samples were realized for cv from 0.2 to 1 in steps of 0.05, then the mean duration (D_{Ap}) and number of apneas per hour (N_{Ap}) were obtained. Both parameters are plotted in Figure VII.5 as a function of cv, together with the corresponding points from empirical RV series. As it can be observed, fLn signals can reproduce the frequency of real apneas, with the 87.5% of empirical N_{Ap} into the 95% confidence

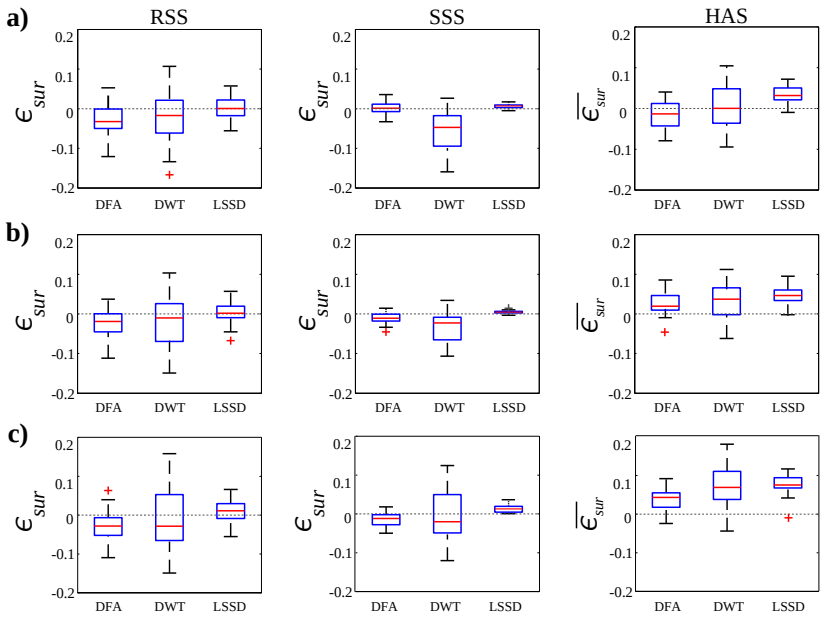

Figure VII.4 – First and second columns: Estimation error from surrogates (ϵ_{sur}) in RSS ($H_s = 0.5$) and SSS ($H_s \approx H_r$). Third column: Averaged estimation error ($\overline{\epsilon_{sur}}$) for HAS prescribing H_s $(0.6, ...0.9)$. Results for regular (a), erratic (b) and periodic (c) RV patterns.

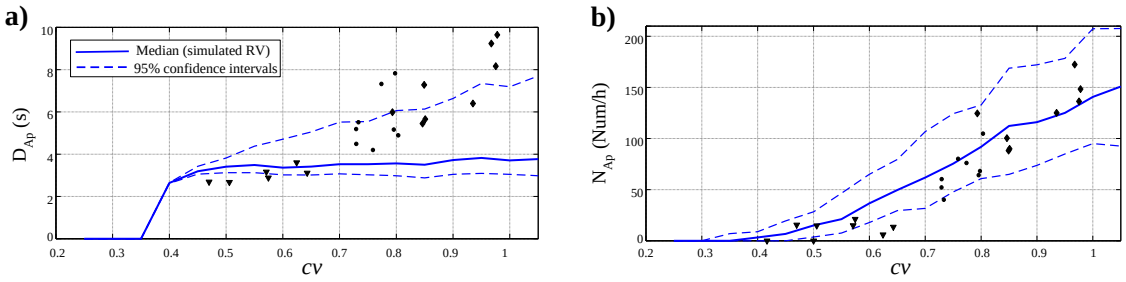

Figure VII.5 – Comparison of the mean duration (D_{Ap}) and the number of apneas per hour (N_{Ap}) obtained by simulated RV (grey lines) and from real signals (regular RV in triangles, erratic in circles and periodic in diamonds). These plots were obtained by simulated RV at $H = 0.6$, but similar curves were obtained with values between 0.55 and 0.9.

intervals (CI) of simulated values. However, the duration of apneas shows some limitations based on the fact that the 47.5% of empirical D_{Ap} (specially long apneas) are not included into the 95% CI obtained by the simulated fLn.

Simulations were performed by generating 100 fLn signals using the average parameters σ_L and μ_L for the three breathing patterns (see Table 1), and H was prescribed from 0.5 to 0.9 in steps of 0.05. To evaluate the robustness of DWT, DFA and LSSD, we utilize the estimation error from simulated values, ϵ_{sim}, the homologous of expression (VII.11):

$$\epsilon_{sim} = H_r - \widehat{H_r}, \tag{VII.12}$$

where H_r is in this case the prescribed value of H by fLn and $\widehat{H_r}$ its estimation. In Figure VII.6, ϵ_{sim} is plotted as a function of H_r for 1024-sample data, comparing the curves obtained for each simulated pattern and estimator.

It can be noticed that, in general, LSSD is the most precise technique, showing the narrowest 95% CI.

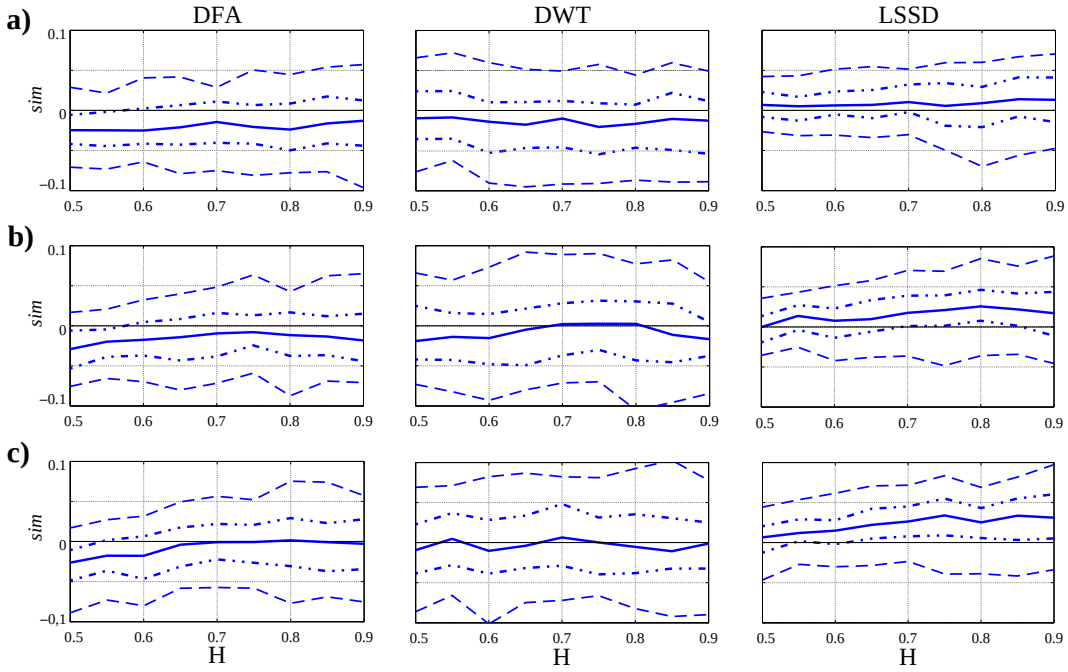

Figure VII.6 – Performance of estimators in 100 realizations of fLn. The median of the simulation error (ϵ_{sim}) in solid lines and 75% and 95% confidence intervals in dashed-dotted (-.-) and discontinuous (- -) lines, respectively. Curves for regular breathing in (a), erratic in (b) and periodic in (c).

However, its accuracy diminishes slightly as the simulated pattern becomes more spiky and the strength of LRD increases. On the other hand, the DWT-based method has the worst precision yet enhances the accuracy with irregular and periodic patterns, and DFA shows the best performance in periodic-like patterns. Observing the sense of bias, DFA and DWT are prone to underestimate the real value of H whereas LSSD behaves conversely.

Simulations using 512-sample signals were also undergone, and the confidence intervals of the estimation errors and bias (not shown here) exceeded 20 to 40% the values obtained in simulations for 1024 samples. Therefore, it can be stated that robustness estimating H decreases as the length of data diminishes, as reported previously [16].

3 Clinical application

This section studies the long range dependence on real breathing variability signals to examine its potential to detect the presence of septicemic infection. Therefore, the Hurst exponent will be estimated in infected and healthy order to compare the differences between them.

3.1 Data selection and parameter tuning

We employed the SEPSIS selection from the PHYSIDEV database (described in Table V.3 from chapter V) which contains 16 infected (*Sepsis*) and 16 healthy (*No-sepsis*) infants having each cohort similar ages and weights to minimize possible differences due to the maturational status.

To estimate the Hurst exponent, we employed DFA and LSSD, the most performing algorithms. According

to Peng et al., an appropriate range of the box size to capture slow dynamics in heartbeat data in adults is $m = 2$ to $m = 64$ [6]. However, breathing dynamics are slower so we computed the function $F(m)$ from 2 to 256 samples (see illustration in Figure VII.7).

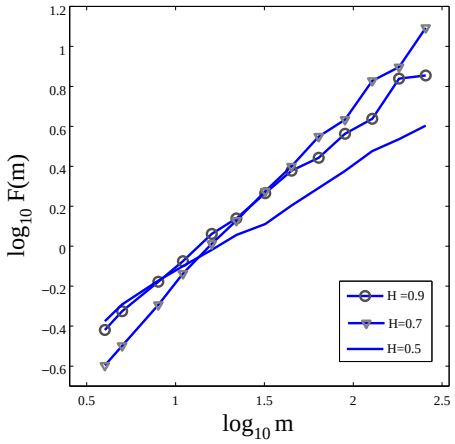

Figure VII.7 – Example of estimating the Hurst exponent by DFA. $F(m)$ is plotted against several box sizes (m) on a log-log scale for three synthetic series with prescribed H at 0.5, 0.7 and 0.9 (generated by fractional Gaussian noise).

On the other hand, since LSSD did not require any parameter setting as it is a non-parametric. To perform the tests, we employed several windows sizes (the entire signal and 2048, 1024 and 512 samples) with 10% overlapping.

3.2 Results and discussion

The Hurst exponent was first estimated on the whole datasets and then applying the different windows sizes. The final values for the latter were obtained by averaging H in the windows contained in each dataset (noted as $\overline{H}$ in Figure VII.8).

We also computed the estimations in the surrogates for every case, a supplementary test of reliability to see if by removing LRD there are differences between *Sepsis*/*No-sepsis* surrogates. If so, it could be thought that the differences observed in real signals have the origin in the nature of data, i.e. spikes or other elements perturbing the estimations. By comparing the pairs of surrogates we did not find significant differences, so it can be considered that estimations on real data correspond to true LRD processes carried by the variability signals.

Observing Figure VII.8, it can be stated that infected patients (red boxes) have slightly lower H values compared to healthy patients, regardless of the analyzed length or the estimation method. Differences were only statistically significant between the whole records, although some pairs revealed p-values close to the significance level in a Mann-Whitney U test. For instance, for 1024-sample tests we obtained $p = 0.095$ and $p = 0.074$ for LSSD and DFA estimations, respectively.

It must be said that data lengths are not homogeneous for the entire datasets although the recording time is almost identical (the mean and standard deviation of samples are 5638 ± 2437 for *Sepsis* and 6865 ± 1908 for *No-sepsis*, corresponding approximately to 2.5 hours). This is caused by the longer duration, in average, of the apneas in *Sepsis* patients (see Table V.4) so the patterns are prone to be more spiky in this group. This should not be a problem for the estimators, robust to irregular patterns, but as mentioned before, our models to tests

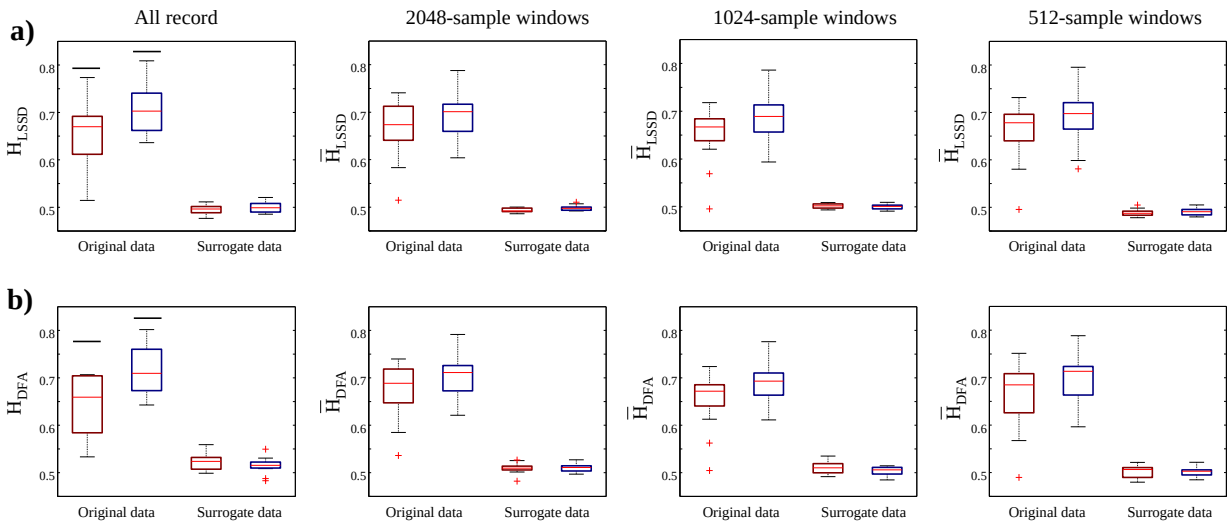

Figure VII.8 – Hurst exponent in infected (red boxes) and healthy (blue boxes) infants, computed on the entire records (first columns) and in sliding windows of 2048, 1024 and 512 samples (second to fourth columns). H was estimated using the LSSD method in a) and using DFA in b).

the performance of estimators cannot generate severe apneas.

Despite this limitation, the results are encouraging since the averaged H through different windows size, and by two different estimation approaches, the infected patients show, in general, a decreased long-range dependence.

4 Conclusion

The application of the surrogate-data tests allowed first to evaluate the properties of RV signals. The random-shuffled surrogates found significant long-range dependence, as reported by earlier works [7, 24], and the small-shuffled surrogates proved the existence of short-term memory regarding the three first consecutive breaths in regular breathing, as it had been found in adults [30]. Not surprisingly, the SRD is weaker as the randomness and apneas increase, i.e. typically immature respiration, where it decreased to two and one consecutive breaths in erratic and periodic patterns, respectively. As randomness is associated with uncorrelated and irregular processes (weak LRD), this result is also consistent with the lowest values of $\widehat{H_r}$ characterizing periodic patterns. On the contrary, regular patterns show the highest $\widehat{H_r}$, suggesting a more predictable long-term behavior within breaths.

The Surrogate-techniques served as well to test the performance of the estimators using artificial data generated from experimental signals. Interestingly, the HAS method provides identically distributed signals than real RV approximating H, permitting to assess the estimation errors in a realistic scenario including spikes. Evaluating the results of this test, it can be stated that H tends to be overestimated in situations in which LRD becomes stronger in artificial signals, particularly if the pattern analyzed contains more spikes, as erratic and periodic breathing. A possible explanation could be the inefficacy of HAS to replicate the exact correlations of fGn on the basis of apneic patterns, but there is still an uncertainty if the presence of spikes itself introduces the before-mentioned bias.

This is the reason why our stochastic model to simulate RV, initially Gaussian, was modified to generate

apneas with approximately the same probability as the real signals. In this way, H can be prescribed reliably by the fLn while apneas can be regulated by the distribution parameters of empirical data. Nevertheless, the fLn approach has some limitations. Severe apneas, which in some patterns exceed frequently 10 seconds, cannot be reproduced within this model because the tail of the fitted distributions restricts greater values. Other marginal laws could be more suitable for highly spiked patterns, such as generalized extreme distributions, but they will be the object of further research.

However, it is noteworthy that the behavior of the estimators with simulated fLn is similar to the trends observed in the surrogates analysis: LSSD appears to be the most precise but least accurate, DWT the most imprecise and DFA seems to be a compromise between precision and accuracy. In spite of the limitations of the fLn model, it can be stated that the use of DWT on RV signals, regardless of the pattern, should be avoided due to the unacceptable bias in estimations (the 75% CI exceed the simulation step). LSSD could be a good choice in case-control studies where two populations need to be characterized by H in regular-erratic patterns, because its good precision gives the best discriminant power despite overestimating the real value of H. On the other hand, the robustness of DFA facing spiky signals makes it a more reliable choice in periodic breathing or mixed patterns.

In relation to the length of the breathing signals, it has to be considered that long and clean records are difficult to obtain at NICU, so the number of samples in RV series could be in some cases insufficient to compute H reliably. Therefore, the optimal length of analysis should be a trade-off between a realistic sample size and a tolerable level of bias in estimators, but keeping in mind that the interest of the sample size is not only technical. Indeed, the choice of very long RV signals (several hours) can result in capturing circadian rhythms such as corporal temperature, feeding or other behavioral elements [31], which might add uncertainty to the measures.

But from a methodological point of view, and considering the application of H to detect infection, the length of the analysis should include at least 512 to 1024 samples, equivalent to time intervals from 10 to 25 minutes. For longer data, the estimators are expected to provide better performances, however, estimations should not exceed 3 hours of recordings (the typical periodicity of food administration).

Finally, the clinical application of the Hurst exponent is encouraging, as this parameter alone might be an indicator of infection. The *Sepsis* population showed in effect lower H values, proving that the fractal organization of the breathing rhythm is degraded (lower H values) as a consequence of the abnormalities induced by the septicemic infection. This relationship between the increased randomness and sepsis is corroborated by previous observations on heart rate variability signals, in which infected infants evidenced abnormal cardiac rhythms, leading to significantly lower fractal measures estimated by the DFA [32].

Bibliography

[1] J. B. Bassingthwaighte, L. S. Liebovitch, and B. J. West. *Fractal Physiology*. American Physiological Society / Oxford October (1994).

[2] A. L. Goldberger. *Fractal dynamics in physiology: Alterations with disease and aging*. Proceedings of the National Academy of Sciences **99**(90001), 2466–2472 February (2002).

[3] B. B. Mandelbrot. *The Fractal Geometry of Nature*. W.H. Freeman (1983).

[4] H. E. Hurst. *Long-term storage capacity of reservoirs*. American Society of Civil Engineers (1950).

[5] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger. *Mosaic organization of DNA nucleotides*. Physical Review E **49**(2), 1685–1689 February (1994).

[6] C. K. Peng, J. E. Mietus, Y. Liu, C. Lee, J. M. Hausdorff, H. E. Stanley, A. L. Goldberger, and L. A. Lipsitz. *Quantifying fractal dynamics of human respiration: age and gender effects*. Annals of Biomedical Engineering **30**(5), 683–692 May (2002). PMID: 12108842.

[7] M. Cernelc, B. Suki, B. Reinmann, G. L. Hall, and U. Frey. *Correlation properties of tidal volume and end-tidal o2 and CO2 concentrations in healthy infants*. Journal of Applied Physiology (Bethesda, Md.: 1985) **92**(5), 1817–1827 May (2002).

[8] P. Abry and D. Veitch. *Wavelet analysis of long-range-dependent traffic*. Information Theory, IEEE Transactions on **44**(1), 2 –15 January (1998).

[9] D. Veitch, P. Abry, and T. M. S. *On the automatic selection of the onset of scaling*. Fractals **11**(2), 377–390 (2003).

[10] H. Kunsch. *Contrasts under long-range correlations*. The Annals of Statistics **21**(2), 943–964 June (1993). Mathematical Reviews number (MathSciNet): MR1232527.

[11] J. Beran. *Statistics for Long-Memory Processes*. Chapman & Hall October (1994).

[12] D. Koutsoyiannis. *Climate change, the hurst phenomenon, and hydrological statistics*. Hydrological Sciences Journal **48**(1), 3–24 (2003).

[13] H. Tyralis and D. Koutsoyiannis. *Simultaneous estimation of the parameters of the Hurst-Kolmogorov stochastic process*. Stochastic Environmental Research and Risk Assessment **25**(1), 21–33 (2011).

[14] X. Navarro, F. Porée, A. Beuchée, and G. Carrault. *Performance analysis of hurst exponent estimators using surrogate-data and fractional lognormal noise models: Application to breathing signals from preterm infants*. Digital Signal Processing **23**(5), 1610–1619 September (2013).

[15] M. S. Taqqu, V. Teverovsky, and W. Willinger. *Estimators for long-range dependence: An empirical study*. Fractals **3**, 785–798 (1995).

[16] W. Rea, L. Oxley, M. Reale, and J. Brown. *Estimators for long range dependence: An empirical study*. arXiv:0901.0762 January (2009).

[17] B. B. Mandelbrot and J. R. Wallis. *Computer experiments with fractional gaussian noises: Part 1, averages and variances*. Water Resources Research **5**(1), 228 (1969).

[18] G. Box, G. Jenkins, and G. Reinsel. *Time Series Analysis: Forecasting & Control (3rd Edition)*. Prentice Hall February (1994).

[19] M. S. Taqqu and V. Teverovsky. On estimating the intensity of long-range dependence in finite and infinite variance time series. In *A Practical Guide To Heavy Tails: Statistical Techniques and Applications*, pages 177–217 (1996).

[20] J. Barunik and L. Kristoufek. *On hurst exponent estimation under heavy-tailed distributions*. Physica A: Statistical Mechanics and its Applications **389**(18), 3844–3855 September (2010).

[21] F. Esposti, M. Ferrario, and M. G. Signorini. *A blind method for the estimation of the hurst exponent in time series: Theory and application*. Chaos: An Interdisciplinary Journal of Nonlinear Science **18**(3), 033126–033126–8 August (2008).

[22] S. Stoev, M. S. Taqqu, C. Park, and J. S. Marron. *On the wavelet spectrum diagnostic for hurst parameter estimation in the analysis of internet traffic*. Comput. Netw. **48**(3), 423–445 June (2005).

[23] X. Navarro, A. Beuchee, F. Poree, and G. Carrault. Performance analysis of hurst's exponent estimators in higly immature breathing patterns of preterm infants. In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 701–704. IEEE May (2011).

[24] U. Frey, M. Silverman, A. L. Barabási, and B. Suki. *Irregularities and power law distributions in the breathing pattern in preterm and term infants*. Journal of Applied Physiology (Bethesda, Md.: 1985) **85**(3), 789–797 September (1998).

[25] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. Doyne Farmer. *Testing for nonlinearity in time series: the method of surrogate data*. Physica D: Nonlinear Phenomena **58**(1–4), 77–94 September (1992).

[26] T. Nakamura, H. Horio, S. Miyashita, Y. Chiba, and S. Sato. *Identification of development and autonomic nerve activity from heart rate variability in preterm infants*. Biosystems **79**(1–3), 117–124 January (2005).

[27] C. R. Dietrich and G. N. Newsam. *Fast and exact simulation of stationary gaussian processes through circulant embedding of the covariance matrix*. SIAM J. Sci. Comput. **18**(4), 1088–1107 July (1997).

[28] J. A. C. Aitchison, J. & Brown. *The Lognormal Distribution*. Cambridge University Press (1969).

[29] B. Mandelbrot. *Possible refinement of the lognormal hypothesis concerning the distribution of energy dissipation in intermittent turbulence*. **12**, 333–351 (1972).

[30] M. J. Tobin, K. L. Yang, A. Jubran, and R. F. Lodato. *Interrelationship of breath components in neighboring breaths of normal eupneic subjects*. American journal of respiratory and critical care medicine **152**(6 Pt 1), 1967–1976 December (1995).

[31] M. Mirmiran and J. H. Kok. *Circadian rhythms in early human development*. Early human development **26**(2), 121–128 September (1991).

[32] A. Beuchée, G. Carrault, J. Y. Bansard, E. Boutaric, P. Bétrémieux, and P. Pladys. *Uncorrelated randomness of the heart rate is associated with sepsis in sick premature infants*. Neonatology **96**(2), 109–114 (2009).

Conclusion and perspectives

The objective of this work was to analyze the cerebral and respiratory activity recorded in Neonatal Intensive Care Units by means of signal processing methods and examine their clinical utility.

Concerning the brain activity, it was analyzed from the EEG recorded in the NICU. Since these signals are recorded by a reduced number of electrodes without the noise protection of standard systems, they are strongly contaminated by several noise sources. Thus, a first part of this work was focused in minimizing the artifacts superimposed on the EEG, taking into account the particularities of these signals in the premature infant: the presence of nonstationarities alternating transitory low frequency voltages (EEG bursts) and the quiescence periods.

An original approach, associating techniques of signal decomposition and noise cancellation, was therefore proposed and studied in-depth in Chapter III. We showed, using both realistic simulations and real signals, the superiority of the combination of the complete ensemble empirical mode decomposition and adaptive filtering to minimize the distortion of EEG. Indeed, when the signals are contaminated with cardiac and low frequency noise, this association increases the rate of noise suppression up to the 50% compared to the use of a classical bandpass filtering plus adaptive filters.

Once the signals were cleaned, we proposed in Chapter IV a new detector of EEG activity based on a logistic regression, permitting the classification of the EEG in burst and inter-burst intervals. A comparative study with a classical thresholding detector, used classically in neonatology, showed the superiority of our approach, presenting performances in terms of sensitivity of 96.1% (against 91.4%) and 95.9% in specificity (against 93.6%). The detectors employed as a reference the gold standard defined by three clinicians, which served, in turn, to examine inter-rater agreements and compare manual and automatic detections. These tests showed that the differences between raters criteria are in the same level than their discrepancies with automatic decisions, thus our algorithm could substitute the human eye satisfactorily.

From a clinical point of view, we were interested in the maturation of the premature baby ex-utero, more particularly, in the potential of the analysis of the EEG and inter-burst intervals to provide indexes quantifying this maturity in monitoring systems. The tests realized in our database stated that the frequency bands of EEG, the duration and distribution of IBIs are apparently related to the age and weight of the infants. These findings are reinforced by the Lempel-Ziv algorithm, whose measures revealed that the complexity in IBIs increased significantly with age, showing the potential of the nonlinear approach to study the perterm's maturity.

We were also interested in the impact of vaccination on the cerebral activity of the premature newborn. This analysis complemented a previous essay of the CHU of Rennes concerning the cardio-respiratory effects in post-immunized infants, in which a significant increase of bradycardia and desaturation was found. The analysis of the EEG, IBIs and hypnograms revealed that the immunization very slightly disturbs the sleep, however, we

observed a redistribution of energies in the spectrum of the EEG. In particular, we stated a relative reduction in the lower delta (0.5 - 1.5 Hz) band and an increase in the lower theta (5.1 – 8 Hz) band. Therefore, the eventual interrelations of the cardio-respiratory system and central nervous system during immunization should be tackled in future work.

Concerning the respiratory activity, the third part of this thesis studied abdominal signals as time-series but also as variability (RV) data, in a similar way as heart rate variability analysis. A respiratory cycle detector was thus developed in Chapter V. However, a prior denoising step was necessary to reject the artifacts of the frequent movements produced in the NICU and filter properly the breathing trace. Once again, this detector was based on a logistic regression classifier with sensitivities and a specificities around 86%, obtained by training its parameters from the annotations of CHU experts. Thanks to this preprocessing, an algorithm designed to identify breathing cycles reached a predictive positive value and a sensitivity of 97.7%.

Applied to the quantification of maturity, we found that the number of apneas and their durations, the length of respiratory cycles and the relative energy in low frequency bands –from a spectral analysis of RV signals– vary with the age and the weight of the baby. Regarding the detection of sepsis, a linear analysis of RV signals provided a first insight of infected patterns, as we observed that the standard deviation of the respiratory cycles and the duration apneas lengthened within infected newborns.

In Chapter VI we investigated the application of nonlinear methods such as correlation dimension, the sample entropy and the numerical noise titration for the study maturation. In a first instance, we checked the existence of nonlinearity in regular and short breathing periods. This property, present in the majority of signals, evolves with maturity in an irregular manner, tending to decrease. The computation of these methods, however, had probably meaningful pitfalls due to the difficulty to quantify nonlinearity in short and noisy signals. On the other hand, the analysis or RV data led to more consistent results, i.e. an increase of complexity with age and weight, as stated in counterpart analyses of heart rate variability. Nevertheless, the significance is still weak and the more relevant measure (sample entropy), should be complemented, in forthcoming tests, with other nonlinear measures not explored here.

The last chapter assessed the long-range dependence in RV signals. After demonstrating the existence of this property in our data, we conducted a thorough analysis to find the best estimator to compute the Hurst exponent (H). The detrended fluctuation analysis (DFA) and the least squares based on standard deviation (LSSD) algorithms were the more robust as they provided the lowest errors, evaluated with artificial -yet realistic- signals with prescribed H. They were generated by means of a stochastic model, the fractional lognormal noise, and a new surrogate-data technique, the H-adjusted surrogates. Both approaches permitted to synthesize RV-like signals that mimic the immature respiratory patterns: Regular, periodic and erratic.

The application of the Hurst exponent as a potential indicator of neonatal sepsis was thereafter explored. Indeed, the infected population of newborns showed a significant reduction in their value of H compared to the healthy population, a fact that evidenced the degradation of the fractal structures in breathing during septicemic conditions. Importantly, this observation is in accordance with similar studies on the heart rate variability that found more randomness in the cardiac rhythms of sick preterm infants.

This work also implemented software platforms already exploitable by clinicians in research. The first concerns the annotation of EEG bursts and the second the annotation of artifacts in breathing signals and the validation of respiratory cycles. These applications represent a fundamental starting point to conduct a large

scale study on the disagreement between several expert's annotations. PHRC-INTEM and CARESS-PREMI projects will enable us to constitute unified databases on several hospital centers and evaluate the differences between them.

The future works related to this thesis are multiple. With regard to the EEG analysis, it is important to constitute a horizontal database making possible the validation of the here-presented results. The latter can only be regarded as preliminaries insofar as the studied database is reduced. The before-mentioned projects will, without a doubt, contribute to this concern. From there on, additional analyses, such as the identification of EEG graphoelements taking advantage of new and effective classification methods, could be carried out to enhance the study of maturity.

With respect to respiration, it would be also appropriate to constitute a horizontal database if maturity wants to be assessed individually and more accurately. From a methodological point of view, the artifact detector for breathing signals should be compared with more sophisticated classification methods, such as the support vector machines or neural networks, to verify if the current performances could be improved.

Finally, concerning the various parameters from the linear and nonlinear tools, the design of a multivariate statistical analysis would probably ameliorate the performances in the detection of infection. To this purpose, it is indeed essential to exploit our former results founded on cardiac variability. This perspective is an explicit part of the clinical research protocol CARESS-PREMI, conceived to validate cardiac and respiratory variability based tools for the detection of sepsis. This protocol, coordinated by our clinical partners, implies three university hospitals (Rennes, Lille and Angers) and represents without a doubt a very positive framework for the clinical application, in the NICU's context, of the solutions described in this dissertation.

List of publications

International journals

1. **X. Navarro**, F. Porée, A. Beuchée, and G. Carrault. Performance analysis of Hurst exponent estimators using surrogate-data and fractional lognormal noise models: Application to breathing signals from preterm infants. *Digital Signal Processing* 23(5), 1610–1619 September (2013).
2. T. Mialet-Marty, A. Beuchée, W. Ben Jmaa, N. N'Guyen, **X. Navarro**, F. Porée, A. M. Nuyt, and P. Pladys. Possible predictors of cardiorespiratory events after immunization in preterm neonates. *Neonatology* 104(2), 151–155 July (2013).
3. **X. Navarro**, F. Porée, A. Beuchée, and G. Carrault. Assessing nonlinear properties in breathing signals from preterm infants. *International Journal of Bioelectromagnetism* 15(1), 102–108 January (2013).
4. **X. Navarro**, F. Porée, A. Beuchée and G. Carrault. Removal of ECG artifacts in preterm EEG using a combination of improved empirical mode decomposition and adaptive filtering. *Transactions on Biomedical Engineering* (Submitted).

International conferences

1. **X. Navarro**, F. Porée A. Beuchée, and G. Carrault. Assessing nonlinear properties in breathing signals from preterm infants. In *7th International Workshop on Biosignal Interpretation, (BSI)*, IEEE July (2012).
2. **X. Navarro**, F. Porée, and G. Carrault. ECG removal in preterm EEG combining empirical mode decomposition and adaptive filtering. In *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 661–664. IEEE, March (2012).
3. **X. Navarro**, A. Beuchée, F. Porée, and G. Carrault. Performance analysis of Hurst's exponent estimators in higly immature breathing patterns of preterm infants. In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 701–704. IEEE May (2011).
4. **X. Navarro**, F. Porée, A. Beuchée, and G. Carrault. Respiration signal as a promising diagnostic tool for late onset sepsis in premature newborns. In *Computing in Cardiology 2010*, pages 947–950. IEEE September (2010).

Appendix A
Blind Source Separation

1 Description

Blind Source Separation (BSS) is a major area of research in signal and image processing aimed at recovering source signals from their mixtures without detailed knowledge of the mixing process [1]. It is hypothesized that the original data is composed by m different unknown sources $\mathbf{s_t} = [s_1(t), s_2(t), ..., s_m(t)]$, subject to certain assumptions. Then, an unknown linear model $A_{n\times m}$ is considered to generate the n observed signals $\mathbf{x_t} = [x_1(t), x_2(t), ..., x_n(t)]$:

$$\mathbf{x_t} = A\mathbf{s_t}, \tag{A.1}$$

meaning that the observations at time t are instantaneous linear combinations of the sources. The estimation of the sources, $\mathbf{y_t}$, consists in identifying the inverse matrix of A, or separation matrix (W), according to the following equation:

$$\mathbf{y_t} \approx \mathbf{s_t} = W\mathbf{x_t} \tag{A.2}$$

In last years, many different algorithms have been proposed to solve the BSS problem. The difference between them relies on the adoption of different additional hypotheses on the sources. The **independent component analysis** (ICA) [2], probably the most popular BSS approach, assumes mutual statistical independence of non-Gaussian sources to find y_t by maximizing this property on the estimated components. A number of algorithms have been proposed to such purpose, being the most common:

- JADE (Joint Approximation Diagonalization of Eigen matrices) [3]: The sources are assumed mutually independent, stationary, without more than one being Gaussian. The algorithm minimizes cross-cumulants (fourth order moments) between the components of the whitened signal and the maximization of the autocumulants.
- SOBI (Second Order Blind Identification) [4]: Sources need to be temporarily correlated to perform a separation based in the second-order statistics (covariance matrices).
- TFBSS (Time-Frequency BSS): Proposes a joint diagonalization of a set of spatial time-frequency distribution matrices when the sources are non-stationary.
- FastICA [5]: A fast implementation of ICA by a fixed-point algorithm, taking advantage of third order statistics to perform separation.

The **canonical correlation analysis** (CCA) constitutes another approach to solve the BSS problem. Originally proposed to measure the linear relationship between two multidimensional random variables [6], it has recently applied to solve Equation A.1, forcing the sources $\mathbf{s_t}$ to be mutually uncorrelated and maximally autocorrelated [7]. More precisely, the solution consists on finding the linear combination of the observed set of vectors $\mathbf{x_t}$ that maximizes the correlation with a temporally delayed versions of itself.

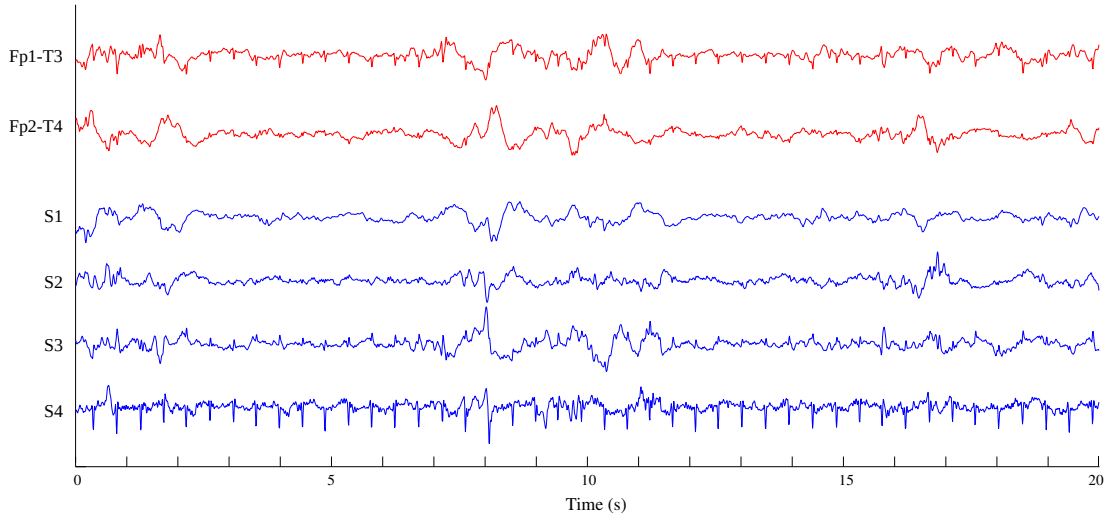

Figure A.1 – Example of application of BSS (FastICA algorithm) on a mixing model composed by four observed signals (in red, represented as Fp1-T3 and Fp2-T4) and four sources (S1 to S4, in blue).

2 Denoising EEG by BSS

The ICA approach is probably the most widely used to solve the BSS problem in electroencephalography. Nevertheless, there is no consensus for the best BSS solution to remove ECG and other noncephalic artifacts [8], being this choice mainly determined by the number of EEG channels, the availability of other signals (EOG, EMG, ECG) and their statistical properties.

To denoise our signals, we employed FastICA and CCA because, in principle, the four EEG channels have the requirements needed for both methods: There are at maximum one Gaussian signal, EEG signals are not correlated with artefacts and their autocorrelations have different structures.

The first test was done employing the four electrodes as the four observed inputs (n =4) to the mixing model. Since the ECG constitutes a source of artifacts, the number of electroencephalographic sources can be hypothesized to be equal or less than three. Following the reasoning of other works, it can be considered that the bilateral set Fp1-T3 and Fp2-T4 are originated by two focal sources, one at each side of the brain [9]. The remaining source should contain eventually other noise sources. Figure A.1 shows an example of applying FastICA with these assumptions. As it can be observed, S4 could be considered the ECG source because it is mainly composed by QRS complexes, but non cardiac waves (presumably other artifacts or low frequency EEG content) are also included. The inefficacy of the separation is also evidenced by the presence of peaks synchronized with heartbeats in source S3.

The application of CCA yielded better results (see Figure A.2) isolating the ECG noise in one source (S4) but unfortunately, a residual activity containing frequencies of interest, were also part of this component. Hence, removing S4 to reconstruct the denoised EEG would result in loosing true brain activity.

Observing qualitatively several examples from different patients and sleep states, we stated similar results than the two showed here. We also performed additional tests modifying the mixing model, supplying five signals (the four EEGs plus the ECG channel) and modifying the number of estimated sources to m=5, 4 and 3, but none of the combinations provided better results.

As mentioned before, to apply BSS some constraints must be accomplished. In our case, the four electrodes (Fp1, Fp2, T3, T4) and one ECG lead supplies five signals ($n = 5$) to the BSS mixing model. Supposing that the only source of

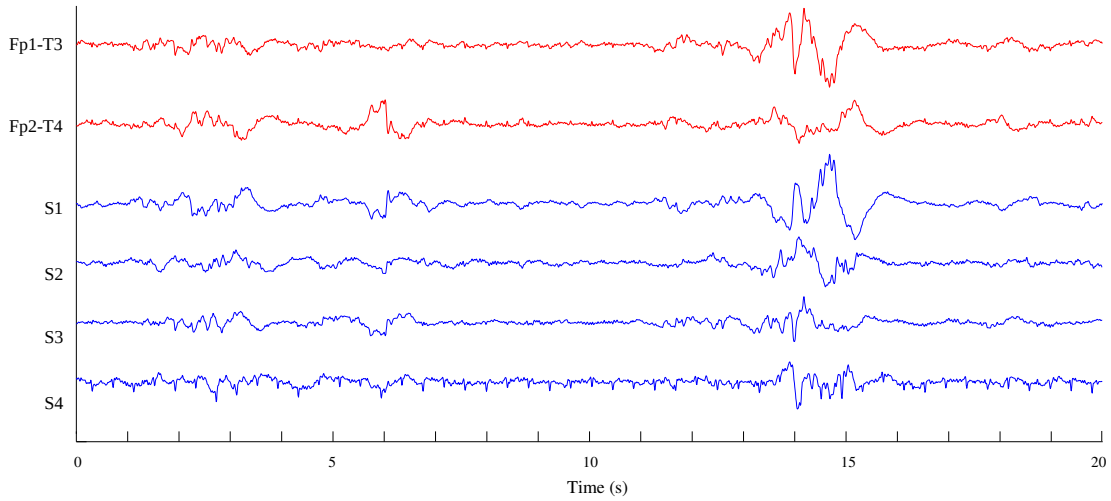

Figure A.2 – Example of application of BSS-CCA. In red, the EEG signals (expressed as Fp1-T3 and Fp2-T4) and in blue the four estimated sources.

artifacts in the EEG is ECG (EMG and low frequency noise were previously filtered), the number of EEG sources (m) can be hypothesized to be equal or less than four. Nevertheless, tests performed using several existing BSS methods on real data from NICU reported in general a very poor separation of the ECG and cross-talk was observed between the EEG information carried by the different channels. This can be caused by the following reasons:

1. Insufficient number of EEG channels. It is known that the best separations in BSS are achieved with an important number of EEG leads, grater to the number of sources. In our case, the number of sources probably exceeds the number of available input signals, leading to an under-determined mixture and consequently to the an ill-posed inverse problem. We speculated with the existence of two EEG sources, an ECG source and another source containing artifacts. But the last component potentially includes left and right ocular movements and unfiltered muscular activity and other artifacts. Since the EOG and EMG are not available, the remaining sources cannot be easily identified.
2. Irregular presence of ECG noise. In some channels ECG noise is highly variable or inexistent whereas in other the signal-to-noise ratio is low. This makes still more difficult the performance of BSS.

In conclusion, the lack of a BSS solution performing consistently on all signals declined the use of this approach in our particular case. However, it has to be mentioned that the literature also proposes algorithms capable of separating independent sources from a reduced number of channels. For instance, Davies and James [10] introduced a methodology for the extraction of multisource brain activity from a single channel, formalizing later the concept as SCICA, single channel ICA [11]. Other BSS algorithms were extended to more complex source mixture models, as for instance the solution of Bell and Sejnowsky [12] for convolutive mixtures of sources (i.e. the observations are FIR-filtered mixtures of the sources) based in the maximization of the joint entropy. In a similar manner, Devuyst et al [13] proposed a QRS complex suppressor with a modification of convolutive ICA using two channels, the noisy EEG and the ECG reference.

The above methods will remain unexplored in the present work, where EEG decomposition and adaptive filtering has been preferred, but further work should include them in a more complete comparative analysis.

Bibliography

[1] A. Cichocki and S.-i. Amari. *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*. John Wiley & Sons Ltd April (2002).

[2] P. Comon. *Independent component analysis, a new concept?* Signal Process. **36**(3), 287âĂŞ314 April (1994).

[3] J.-F. Cardoso and A. Souloumiac. *Blind beamforming for non-gaussian signals*. Radar and Signal Processing, IEE Proceedings F **140**(6), 362–370 (1993).

[4] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines. *A blind source separation technique using second-order statistics*. IEEE Transactions on Signal Processing **45**(2), 434–444 (1997).

[5] A. Hyvärinen and E. Oja. *Independent component analysis: algorithms and applications*. Neural Networks **13**(4âĂŞ5), 411–430 June (2000).

[6] H. Hotelling. *Relations between two sets of variates*. Biometrika **28**(3/4), 321–377 December (1936).

[7] O. Friman, M. Borga, P. Lundberg, and H. Knutsson. *Exploratory fMRI analysis by autocorrelation maximization*. NeuroImage **16**(2), 454–464 June (2002).

[8] M. Klemm, J. Haueisen, and G. Ivanova. *Independent component analysis: comparison of algorithms for the investigation of surface electrical brain activity*. Medical & Biological Engineering & Computing **47**(4), 413–423 April (2009).

[9] F. Porée, A. Kachenoura, H. Gauvrit, C. Morvan, G. Carrault, and L. Senhadji. *Blind source separation for ambulatory sleep recording*. Ieee Transactions on Information Technology in Biomedicine **10**(2), 293–301 April (2006).

[10] C. J. James and D. Lowe. *Extracting multisource brain activity from a single electromagnetic channel*. Artificial Intelligence in Medicine **28**(1), 89–104 May (2003).

[11] M. Davies and C. James. *Source separation using single channel ICA*. Signal Processing **87**(8), 1819–1832 August (2007).

[12] A. J. Bell and T. J. Sejnowski. *An information-maximization approach to blind separation and blind deconvolution*. NEURAL COMPUTATION **7**, 1129—1159 (1995).

[13] S. Devuyst, T. Dutoit, P. Stenuit, M. Kerkhofs, and E. Stanus. *Cancelling ECG artifacts in EEG using a modified independent component analysis approach*. EURASIP Journal on Advances in Signal Processing **2008**, 1–14 (2008).

Appendix B

Tables from Part II

This appendix provides complementary information about the analysis of sleep, EEG discontinuity parameters and energy bands presented in Chaper IV.

The tables show means ± standard deviations of several parameters, confronting the pair of groups G_1 (the most immature) and G_2 (the least immature). Depending on the maturative criteria, groups have superindexes PMA, PNA and W if they are divided according postmenstrual age, postnatal age and weight, respectively. Asterisks (*) denote significant differences at a statistical level $p<0.05$ in a Mann-Whitney U test.

	G_1^{PMA}	G_2^{PMA}	G_1^{PNA}	G_2^{PNA}	G_1^{W}	G_2^{W}
Transitions / h	10.4 ± 4.18	9.04 ± 5.12	10.0 ± 4.76	9.43 ± 4.63	10.3 ± 4.99	9.03 ± 4.22
% of time in QS	7.34 ± 6.61	6.22 ± 8.07	8.80 ± 6.74	4.66 ± 7.38*	5.34 ± 6.02	8.57 ± 8.40
% of time in NQS	87.3 ± 9.39	87.1 ± 9.97	85.7 ± 9.40	88.8 ± 9.67	89.5 ± 8.36	84.4 ± 10.3
% of time awake	5.38 ± 4.94	6.65 ± 6.90	5.52 ± 5.45	6.50 ± 6.51	5.16 ± 4.47	7.00 ± 7.33

Table B.1 – Description of sleep parameters, confronting age and weight groups (before vaccination).

	G_1^{PMA}	G_2^{PMA}	G_1^{PNA}	G_2^{PNA}	G_1^{W}	G_2^{W}
0.5–1.5 Hz	78.2 ± 3.36	**76.0 ± 4.39***	76.1 ± 3.16	78.3 ± 4.54	77.7 ± 4.62	76.4 ± 3.05
1.6–3.0 Hz	12.9 ± 1.90	13.8 ± 2.36	14.0 ± 1.64	12.6 ± 2.42	12.8 ± 2.37	14.0 ± 1.71
3.1–5.0 Hz	4.73 ± 0.71	**5.49 ± 1.14***	5.32 ± 0.82	4.87 ± 1.15	5.05 ± 1.22	5.16 ± 0.71
5.1–8.0 Hz	2.36 ± 0.48	2.82 ± 0.63*	2.63 ± 0.50	2.53 ± 0.70	2.57 ± 0.74	2.60 ± 0.37
8.1–14 Hz	1.84 ± 0.75	1.85 ± 0.62	1.91 ± 0.54	1.77 ± 0.82	1.87 ± 0.81	1.81 ± 0.51

Table B.2 – Relative energy in several EEG bands versus the age and weight groups in non-quiet sleep (before vaccination).

EEG Band	G_1^{PMA}	G_2^{PMA}	G_1^{PNA}	G_2^{PNA}	G_1^{W}	G_2^{W}
0.5–1.5 Hz	77.0 ± 2.45	74.7 ± 3.04*	75.8 ± 3.37	76.0 ± 2.45	76.2 ± 3.26	75.6 ± 2.67
1.6–3.0 Hz	13.2 ± 1.27	14.4 ± 1.63*	14.0 ± 1.79	13.6 ± 1.24	13.4 ± 1.62	14.2 ± 1.41
3.1–5.0 Hz	5.06 ± 0.45	5.90 ± 0.92*	5.49 ± 0.90	5.43 ± 0.76	5.42 ± 0.88	5.52 ± 0.79
5.1–8.0 Hz	2.54 ± 0.42	3.01 ± 0.46*	2.71 ± 0.51	2.83 ± 0.49	2.78 ± 0.58	2.75 ± 0.41
8.1–14 Hz	2.08 ± 1.01	1.93 ± 0.58	1.98 ± 0.55	2.04 ± 1.10	2.16 ± 1.01	1.83 ± 0.54

Table B.3 – Relative energy in several EEG bands versus age and weight groups in quiet sleep (before vaccination).

	G_1^{PMA}	G_2^{PMA}	G_1^{PNA}	G_2^{PNA}	G_1^{W}	G_2^{W}
Bursts per minute	5.89 ± 1.36	6.57 ± 0.59	6.65 ± 0.73	5.61 ± 1.33*	5.81 ± 1.33	6.57 ± 0.80
Mean duration of bursts (s)	5.01 ± 2.87	4.43 ± 1.42	4.37 ± 0.87	5.25 ± 3.39	4.98 ± 3.02	4.56 ± 1.45
Max duration of IBI (s)	17.8 ± 2.77	13.5 ± 6.16	17.2 ± 3.69	14.7 ± 5.74	17.2 ± 3.88	14.7 ± 5.59
Mean duration of IBI (s)	5.86 ± 1.69	4.58 ± 1.93	4.75 ± 1.13	6.03 ± 2.32	6.05 ± 1.78	4.56 ± 1.68
% of Bursts	0.45 ± 0.11	0.49 ± 0.17	0.48 ± 0.09	0.45 ± 0.18	0.44 ± 0.12	0.50 ± 0.15
L-Z complexity	0.23 ± 0.03	0.26 ± 0.01*	0.24 ± 0.02	0.24 ± 0.04	0.23 ± 0.04	0.25 ± 0.02

Table B.4 – Discontinuity parameters describing burst activity versus age and weight groups in quiet sleep (before vaccination).

Appendix C

The method of surrogate-data

Surrogate-data or surrogates are artificial signals generated from empirical data, but preserving certain statistical properties originally present. Several techniques are available to modify signals both in time and frequency domain, being the most relevant introduced in this appendix.

Surrogate-data tests are utilized to statistically infer the nature of the process generating the observed data [1]. The realization of a surrogate set determines the formulation of the null hypothesis, H_0, which is to be tested with a statistic capable to discern original and surrogates properties. If the statistics from the surrogate set and from the empirical data are significantly different, H_0 can be rejected.

1 Temporal domain surrogates

The **Random-Shuffled Surrogates** (RSS) [2] are generated by randomly exchanging the order of data in a time series. Any temporal correlations are thus destroyed and the spectrum is whitened. The resulting series are essentially random observations (Hurst exponent close to 0.5) drawn from the same probability distribution as the data. In surrogate-data tests, the addressed null hypothesis is that Long Range Dependence in empirical data is generated by uncorrelated noise.

In **Small-Shuffled Surrogates** [3](SSS) samples are shuffled in a small scale, so local correlations (Short Range Dependence) are lost and long-term structures (trends) are preserved. Thus, H is theoretically unchanged and the null hypothesis addressed with SSS should be that empirical data does not contain short-term dependent processes. Being $x(n)$ the original data and $i(n)$ the index of $x(n)$, $g(n)$ the Gaussian random series and $s(n)$ the surrogate-data, the generation of SSS data is done in four steps:

1. Obtain $i'(n) = i(n) + Ag(n)$ where A is an amplitude.

2. Sort $i'(n)$ by the rank order, i.e. the sequence order in which the values of different relative magnitude occur.

3. Form $\widehat{i}(n)$ with the index of $i'(n)$

4. Reorder the original data with the perturbed index to obtain surrogate-data, $s(n) = x(\widehat{i}(n))$.

2 Frequency domain surrogates

The basic idea of this kind of surrogates is to manipulate the discrete Fourier transform (DFT) of the data and obtain the surrogates by performing the inverse Fourier transform.

A firstly proposed algorithm by Theiler et al. [1] generates surrogates by shuffling the phases of the DFT but maintaining the amplitude of the complex conjugate pairs. The power spectrum is then preserved but the nonlinear determinism is eliminated. However, this procedure has the limitation of spurious low-frequency effects when the inverse transform is performed, so alternative algorithms were proposed by the same authors.

The **amplitude adjusted Fourier transformed** (AAFT) surrogates algorithm [1] generates surrogate data by first rescaling the values in the original time series so they are gaussian. Then the Fourier Transform is used to make surrogate time series which have the same Fourier spectrum as the rescaled data, and finally, the gaussian surrogate is then rescaled back to have the amplitude distribution as the original time series.

The **iterated amplitude adjusted Fourier transformed** (iAAFT) surrogates has been claimed to be superior to AAFT surrogates [4] in retaining the power spectrum. Let $x(n)$ be the original data, $s(n)$ be a sorted list of $x(n)$, and $X_k^2 = |\sum_{n=0}^{N-1} x_n e^{i2\pi kn/N}|^2$ be the squared amplitudes of the FT of $x(n)$, the algorithm consists in the following iteration scheme:

1. Do a random shuffle of the data, $x(n)^i$
2. Perform the FT of $x(n)^i$ replacing the squared amplitudes, $X_k^{2,i}$ by X_k^2 and do the inverse transform.
3. Rank order the resulting series in order to assume exactly the values taken by $x(n)$.
4. Since the spectrum of the resulting $x(n)^{(i+1)}$ will be modified, repeat from step 2.

The accuracy is checked at each iteration step by measuring the discrepancy of the original spectrum and the surrogate until a given accuracy is reached.

Bibliography

[1] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. Doyne Farmer. *Testing for nonlinearity in time series: the method of surrogate data*. Physica D: Nonlinear Phenomena **58**(1-4), 77–94 September (1992).

[2] J. A. Scheinkman and B. LeBaron. *Nonlinear dynamics and stock returns*. The Journal of Business **62**(3), 311–37 (1989).

[3] T. Nakamura, H. Horio, S. Miyashita, Y. Chiba, and S. Sato. *Identification of development and autonomic nerve activity from heart rate variability in preterm infants*. Biosystems **79**(1–3), 117–124 January (2005).

[4] T. Schreiber and A. Schmitz. *Improved surrogate data for nonlinearity tests*. Physical Review Letters **77**(4), 635–638 July (1996).

Appendix D

Tables from Part III

This appendix provides complementary information about the analyses carried out in Chapter V and Chapter VI. The tables listed below depict the variables within the four degrees of maturation (G_1 to G_4) according postmenstrual age, postnatal age and weight. The signals were selected from 58 patients of both the VACCIN and PHISIDEV databases.

Differences between groups have been analyzed with a Mann-Whitney U test, so that if statistical significance exists between a pair of groups, a super-index, described in Table D.1, is added next to the statistical in question.

	G_1	G_2	G_3	G_4
G_1	–	a	b	c
G_2	a	–	d	e
G_3	b	d	–	f
G_4	c	e	f	–

Table D.1 – Super-indexes denoting statistically significant differences between groups ($p < 0.05$).

	G_1^{PMA}	G_2^{PMA}	G_3^{PMA}	G_4^{PMA}
$\overline{t_{tot}}$	1.12 ± 0.21	1.24 ± 0.30	1.25 ± 0.28	1.30 ± 0.21
σt_{tot}	0.94 ± 0.25	1.02 ± 0.36	0.79 ± 0.22	0.79 ± 0.20
$\overline{ti/te}$	1.17 ± 0.21	**1.35 ± 0.21**^{a}	1.39 ± 0.44	1.30 ± 0.39
N_{ap}	85.7 ± 28.7	80.8 ± 29.3	**62.8 ± 22.7**^{b}	**57.1 ± 18.0**^{c,e}
D_{ap}	5.42 ± 1.02	5.79 ± 1.39	5.20 ± 0.87	5.26 ± 0.89
Std_{ap}	2.33 ± 0.58	2.18 ± 0.87	**1.27 ± 0.23**^{b,d}	**1.22 ± 0.37**^{c,e}
VLF	63.2 ± 18.9	57.1 ± 25.5	47.4 ± 19.5	**48.2 ± 15.6**^{c}
LF	26.4 ± 12.9	27.9 ± 14.5	37.6 ± 16.4	35.6 ± 12.4
HF	4.24 ± 5.97	6.13 ± 8.64	**5.51 ± 4.10**^{b}	**5.72 ± 3.65**^{c,e}

Table D.2 – Mean statistics computed in the whole recordings. Results from PMA groups.

	G_1^{PNA}	G_2^{PNA}	G_3^{PNA}	G_4^{PNA}
$\overline{t_{tot}}$	1.17 ± 0.23	1.18 ± 0.30	1.29 ± 0.28	1.26 ± 0.20
σt_{tot}	1.01 ± 0.26	0.94 ± 0.36	0.82 ± 0.24	**0.76 ± 0.17**^{c}
$\overline{ti/te}$	1.30 ± 0.24	1.21 ± 0.21	1.40 ± 0.44	1.29 ± 0.37
N_{ap}	88.1 ± 28.9	78.3 ± 29.1	**63.4 ± 24.4**^{b}	**56.5 ± 15.3**^{c,e}
D_{ap}	5.67 ± 1.06	5.52 ± 1.38	5.29 ± 0.95	5.16 ± 0.78
Std_{ap}	2.38 ± 0.58	2.12 ± 0.86	**1.21 ± 0.30**^{b,d}	**1.29 ± 0.30**^{c,e}
VLF	63.5 ± 18.5	56.8 ± 25.8	**46.7 ± 19.3**^{b}	**49.0 ± 15.8**^{c}
LF	26.3 ± 12.0	28.0 ± 15.3	37.1 ± 17.5	36.1 ± 10.8
HF	3.87 ± 4.87	6.53 ± 9.26	**5.73 ± 3.97**^{b}	**5.48 ± 3.80**^{c}

Table D.3 – Mean statistics computed in the whole recordings. Results from PNA groups.

	G_1^{W}	G_2^{W}	G_3^{W}	G_4^{W}
$\overline{t_{tot}}$	1.12 ± 0.22	1.25 ± 0.29	1.22 ± 0.23	**1.35 ± 0.25**^{c}
σt_{tot}	0.90 ± 0.26	1.05 ± 0.34	**0.77 ± 0.17**^{d}	0.81 ± 0.25
$\overline{ti/te}$	1.16 ± 0.21	**1.36 ± 0.19**^{a}	1.26 ± 0.30	1.44 ± 0.50
N_{ap}	81.1 ± 23.8	85.7 ± 33.7	**64.7 ± 20.9**^{b}	**54.5 ± 19.1**^{c,e}
D_{ap}	5.34 ± 1.07	5.87 ± 1.31	5.10 ± 0.87	5.39 ± 0.86
Std_{ap}	2.29 ± 0.61	2.22 ± 0.85	**1.31 ± 0.32**^{b,d}	**1.17 ± 0.27**^{c,e}
VLF	65.5 ± 15.4	54.6 ± 27.1	**52.3 ± 13.9**^{b}	**42.4 ± 20.1**^{c}
LF	27.1 ± 13.1	27.1 ± 14.4	33.2 ± 12.6	**40.7 ± 15.9**^{c,e}
HF	2.66 ± 1.09	7.84 ± 9.96	**5.14 ± 2.02**^{b}	**6.18 ± 5.30**^{c}

Table D.4 – Mean statistics computed in the whole recordings. Results from weight groups.

	G_1^{PMA}	G_2^{PMA}	G_3^{PMA}	G_4^{PMA}
NL (%)	14.5 ± 7.90	16.2 ± 7.44	14.7 ± 10.1	16.4 ± 6.49
$SpEn$	1.02 ± 0.13	0.94 ± 0.21	0.76 ± 0.15^{b,d}	0.78 ± 0.15^{c}
D_2	2.69 ± 0.65	2.38 ± 0.70	1.88 ± 0.29^{b,d}	2.14 ± 0.40^{c}
AMI	0.17 ± 0.03	0.23 ± 0.07	0.24 ± 0.04^{b}	0.23 ± 0.03^{c}
NI_{SpEn}	0.22 ± 0.05	0.27 ± 0.09	0.16 ± 0.09^{d}	0.19 ± 0.09
IN_{D2}	-0.50 ± 1.01	-0.70 ± 1.19	0.04 ± 0.34	0.06 ± 0.33

Table D.5 – Nonlinear measures averaged in the five most five regular 90-second excerpts. Results from PMA groups.

	G_1^{PNA}	G_2^{PNA}	G_3^{PNA}	G_4^{PNA}
NL (%)	14.6 ± 7.26	16.0 ± 8.04	13.8 ± 7.87	17.4 ± 9.00
$SpEn$	0.95 ± 0.16	1.00 ± 0.20	0.78 ± 0.12^{b,d}	0.75 ± 0.17^{c,e}
D_2	2.49 ± 0.50	2.57 ± 0.84	1.89 ± 0.28^{b,d}	2.12 ± 0.41^{c}
AMI	0.20 ± 0.05	0.20 ± 0.08	0.23 ± 0.04^{b,d}	0.23 ± 0.04^{c,e}
NI_{SpEn}	0.20 ± 0.05	0.29 ± 0.07^{a}	0.17 ± 0.09^{d}	0.19 ± 0.10^{e}
IN_{D2}	-0.28 ± 0.94	-0.94 ± 1.16	0.04 ± 0.38^{d}	0.06 ± 0.29^{e}

Table D.6 – Nonlinear measures averaged in the five most five regular 90-second excerpts. Results from PNA groups.

	G_1^W	G_2^W	G_3^W	G_4^W
NL (%)	14.6 ± 7.76	16.0 ± 7.61	16.6 ± 8.45	14.3 ± 8.74
$SpEn$	1.01 ± 0.13	0.95 ± 0.22	0.77 ± 0.17^{b,d}	0.77 ± 0.12^{c,e}
D_2	2.73 ± 0.64	2.35 ± 0.69	2.06 ± 0.37^{b,d}	1.94 ± 0.36^{c}
AMI	0.17 ± 0.03	0.23 ± 0.08	0.23 ± 0.04^{b}	0.24 ± 0.04^{c}
NI_{SpEn}	0.23 ± 0.07	0.26 ± 0.08	0.17 ± 0.07^{d}	0.19 ± 0.12
IN_{D2}	-0.53 ± 1.06	-0.67 ± 1.15	-0.01 ± 0.37	0.13 ± 0.26^{e}

Table D.7 – Nonlinear measures averaged in the five most five regular 90-second excerpts. Results from weight groups.

Ces travaux portent sur le traitement et l'analyse des signaux issus des unités de soins intensifs néonatales pour l'étude de la maturité et de l'infection généralisée (sepsis) chez le nouveau-né prématuré.
Dans une première partie, des électroencephalogrammes (EEG) enregistrés avec un nombre réduit d'électrodes sont analysés. Dans le contexte de l'USIN, ces signaux sont particulièrement bruités et de fait des méthodes de décomposition du signal et d'annulation optimale du bruit, adaptées aux particularités des EEG immatures, ont été spécifiquement développées. Les meilleurs résultats ont été obtenus en associant une version améliorée de la décomposition modale empirique et un filtrage adaptatif. L'analyse quantitative de l'EEG et des bouffées d'ondes delta sont ensuite étudiées. Un classifieur basé sur la régression logistique est proposé. Quelques applications cliniques sont finalement examinées. En particulier, nous avons constaté que certains paramètres (l'énergie des bandes EEG et les distributions des bouffées) pourraient Ãłtre utiles pour apprécier la maturation et pour étudier les effets de la vaccination.
Dans une deuxième partie, les signaux de respiration sont étudiés en profondeur. Un algorithme de rejection d'artéfacts, basé sur la régression logistique, et un détecteur de cycles respiratoires sont alors mis au point pour produire des séries temporelles de variabilité respiratoire. Ensuite, la dimension de corrélation, l'entropie d'échantillonnage et la titration de bruit sont exploitées pour évaluer la maturité des prématurés. Bien que la non-linéarité ait été appréciée dans la respiration, ses évolutions n'ont pas montré de réelle concordance avec la maturation. La dépendance à long terme est ensuite examinée pour détecter l'infection. Ceci nous amené à conduire une étude comparative rigoureuse de trois estimateurs de l'exposant de Hurst (H) sur les données de variabilité respiratoire. Pour cela, un simulateur de données synthétiques et des surrogates réalistes à la variabilité respiratoire sont proposés. Appliquées aux données réelles, on montre que l'exposant de Hurst et plusieurs descripteurs linéaires (durées de cycle respiratoire et nombre d'apnées) de la variabilité respiratoire sont des puissants indices de détection d'infection.

Mots clés : Nouveau-né prématuré, unité de soins intensifs néonatales, apnée, maturation, sepsis, EEG, respiration, détection automatique, chaos, fractale, titration du bruit, entropie, exposant de Hurst.

This work processed and analyzed signals from the neonatal intensive care units (NICUs) to study the maturity and the generalized infection (sepsis) in the premature infant.
In the first part, electroencephalograms (EEG) from a reduced montage has been denoised using signal decomposition techniques and noise cancellation optimally adapted to the particularities of immature EEGs. Bests results were obtained by applying a new association of an improved version of empirical mode decomposition and adaptive filtering. Then, from the quantitative analysis of the EEG and the delta bursts, detected by a logistic regression classifier, some clinical applications have been tested. In particular, we found that a few parameters could be useful to quantify maturation and study the effects of vaccination.
The second part of this work, processed breathing signals from the NICU to obtain both temporal and respiratory variability (RV) series. An algorithm to reject the artifacts based on logistic regression and a cycle detector ensured clean data, analyzed thereafter by means of several linear and nonlinear methods for clinical purposes. Correlation dimension, sample entropy and noise titration were exploited to assess maturity. Although nonlinearity was present in most of the breathing signals, there were not clear changes according to the infants maturation. The long-range dependence was applied to detect infection, so a performance analysis of three Hurst exponent (H) estimators applied on RV data was carried out, proposing new formalisations to generate realistic synthetic and surrogate data. It was demonstrated that computing H and basic linear descriptors (related to the breathing cycle durations and apnea) in RV data constitutes an interesting tool to detect infection.

Keywords : Preterm newborn, neonatal intensive care unit, apnea, maturity, sepsis, EEG, respiration, automatic detection, chaos, fractal, noise titration, entropy, Hurst exponent.