KU LEUVEN

ARENBERG DOCTORAL SCHOOL
Faculty of Engineering Science

Advanced models for monitoring stress and development trajectories in premature infants

Mario Lavanga

Supervisors:
Prof. dr. ir. S. Van Huffel
Prof. dr. E. Ortibus
Prof. dr. ir. A. Caicedo Dorado

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Engineering Science (PhD): Electrical Engineering

September 2020

Advanced models for monitoring stress and development trajectories in premature infants

Mario LAVANGA

Examination committee:
Prof. dr. ir. O. Van Der Biest, chair
Prof. dr. ir. S. Van Huffel, supervisor
Prof. dr. E. Ortibus, supervisor
Prof. dr. ir. A. Caicedo Dorado, supervisor
Prof. dr. ir. M. Verhelst
Prof. dr. G. Naulaers
Prof. dr. R. Slater
(University of Oxford)
Prof. dr. D. Marinazzo
(University of Ghent)

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Engineering Science (PhD): Electrical Engineering

September 2020

© 2020 KU Leuven – Faculty of Engineering Science
Uitgegeven in eigen beheer, Mario Lavanga, Kasteelpark Arenberg 10 box 2446, B-3001 Leuven (Belgium)

Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotokopie, microfilm, elektronisch of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever.

All rights reserved. No part of the publication may be reproduced in any form by print, photoprint, microfilm, electronic or any other means without written permission from the publisher.

Preface

Five years have passed since that humid and hot summer 2015. Those days were passing by with a mix of euphoria for a very close graduation and the anxiety of a thesis deadline. Nothing seems to change in our lives. I find myself almost in the same situation and things keep oscillating around a zero-mean like a sine. Time and space keep moving and bending without us realizing that we are in it. However, someone taught us that if we look reality from a different frame and scale, we could see the dynamics of it. If I may, I will guide you through a couple of thanks to rewind the last years. We will do it just in a different space-temporal scale.

I must start thanking my promoter Prof. Sabine Van Huffel, who had the will and the strength to listen to me after long conference session in that hot summer. There is a great need of gut to listen to a random Italian student, to assess his capabilities in less than 45 minutes and to accept him in your group after 14 days. Sabine, I could write how much I learned from you or how I am impressed about your work endurance or your grammar abilities. However, the most important trait that characterizes you as person is surely the trust that you put in people. I have been witnessing the different cultural, scientific and social backgrounds of the people that joined our group. Each of us had a fair shot to contribute to the research of our group. Each of us was welcome for a dinner at your place or a personal conversation about feelings of the present and ideas for the future. Each of us was accepted for the scientific contribution independently from our origin, gender, economical background, sexual or political orientation. This diversity that you built up year by year created one of the most bonded research group that you can find in the scientific community. I could not thank you more for your trust and the opportunity to feel myself at home far away from home.

The second place of this preface must go to one of the most committed researcher of all times and a truly public servant. My gratitude has to go to Prof. Alexander Caicedo. Like the tons and tons of emails that we sent each other, I will go

through my personal thoughts with a bullet points list. First, I have to get used to the fact that you have a title next your name. Second, I will never forget that the first day that I spent in my PhD was with you. Hopefully, it would be the same for the last. Third, I will never forget your initial teasing and our start was definitively rocky. Fourth, I will never forget the times that people could not understand my English and I asked you for help. Fifth, I will never forget those moments of loneliness that you could soothe with your patience, kindness and support. Sixth, I will never forget the day that you told me that I could be whoever I want to be, without strings attached or judgments from our peers. Seventh, I will never forget the attachment that I felt when you left for Colombia. Like any engineers, my emotions are hidden behind the gears of our mind, but I could only realize how close we were only when you left. Eighth, I will never forget the frustration that I shared with my research throughout my last year and I will never forget the way you guided me out of it. Ninth, I will never forget the support that your gave me during my mental lockdown of the last months. N-th, this list can go to the infinite, but the simple the better, right? However, I might still need more time to learn how to be concise.

I wanted to thank Prof. Gunnar Naulears, Prof. Els Ortibus and Prof. Katrien Jansen. Getting the time in your agenda to listen to an Italian engineer was hard. Getting that extra time to understand what he wanted to say was even harder. Getting even more extra time to wait the end of his never-ending presentation was a titanic effort. I cannot be sorrier for the precious time that I wasted. I cannot be merrier for the time that you granted me to get the best possible research outcome. Another special thanks also goes to Dr. Bieke Bollen and once again to Katrien. You are the people that listened to me the most from the clinical side. You did it during and beyond your working schedule. And yet, you did it for the passion to see what the electrophysiology can tell us about our mental health. You did it to help this clumsy and tall engineer to get out of his shell and communicate with the doctors. You taught me how to step up my analysis to be useful and understandable to the clinical community. There were ups and downs like any relationship. I am sorry for the downward trends and I am glad that we shared the upward trends. If we apply a repeated-measurements ANOVA test, the ups are significantly more that downs. I am sure that the findings of the work that we did together and the time that we spent together confidently support the statement: "I am happy to have worked with you".

A list of other people from the clinical world has to be thanked for their patience in collaborating, working or listening to who is writing this text: Dr. Anneleen Dereymaeker, Dr. Bart Boets, Katrien Lemmens, Prof. Lieven Lagae, Dr. Jessie De Ridder, Dr. Guy Bosmans. I should thank all the other collaborators of the Resilience study and the EPISTOP consortium for the time that they

spent in reading my manuscripts or providing useful feedback for my research. It should also be highlighted the generous contribution of the Fonds voor Wetenschappelijk Onderzoek – Vlaanderen with the strategic basic research grant which supported me in any step of my research.

After an extensive acknowledgment, it is finally the time for the fun part. Where could I start with Biomed? I should probably give a list of thanks and stop it there, but it wouldn't be me, right? Therefore, let me cluster you in an unsupervised manner. The links among clusters are estimated by imaginary coherence. I could not start without a cluster of Maricas Cabronas di Girona, comprised of Amalia AmaraVilla, John El Silvio Morales Tellez, CaroLinda MasHermosa Varon. Papito John, I simply do not deserve you. The amount of kindness and support that you provided are not matched by anybody in the world. I did not only find a smart colleague and a humble person, but I am proud to call you friend. The reason why you are so special and remain so special is the amount of people, love and *gente* that you have around you. I am sure that you know the etymology of those word by now. Never forget that your success will always be proven by the beautiful people around you. This also reminds me to say thanks to Gibran Tonatiu, Sandra, Andrea, Popochita, la Martha y la familia Tellez Morales. Gibranito, be the guardian of John while I am traveling around the world. I wanted to say that John is the smartest person on Earth, but this is merely false. However, the world already knows the smartest person in our lifetime. The galaxy awaits for her intelligence to bring the Force into balance. Carolina, I could keep writing a lot of sentences like this one. The truth is that you are one of a kind, a diamond in the rough ready to shine in her career. Your beauty does not rely only on your appearance, youth or intelligence, but it is woven in your listening skills that gave me so much relief in the last years. Steven, protect and support Carolina until she reaches the stars. Amalia, Mea Amor, la chica mas roja, a perfect character for a book about resilience, ecology, adventure and femininism. I have never spent one single day without being grateful that I met you. The moments that we spent talking about gender fluidity and the role of women in our society, the discussion about migration and the books about ecology, the dinners at your place and the one dinner at my place, the evenings with drinks and ice-cream will be the source of strength to go forward. In the sad and the good days, there was always you ready to hear about my stories, plans, dates, failures and frustrations. Our political differences and our will to make this world a better place make us a unique dyad and I cannot wait to have more discussions, more dinners and more evenings with you. No matter wherever it is. No matter who we are with. No matter who we are gonna be. I just hope that the place can be reached by bike. By the way, give my regards to Kranti, Jorge, Apu, Antoine, Mar (I should write a book for this girl too). Giselle and Alex, I did not forget about you and I hope to reach you in Cuba to spend time together

again. Esteban, you look amazing in sport clothes. If you can also take care of Amalia when I leave, it would be great.

The second cluster is the Turkish-Greek group made out of Christos and Cem. First, you are both Europeans for me. Second, I am happy that I met both of you. Christos, I will never forget your Greek Wedding. It was a good occasion to learn new skills. Most importantly, it reminded me of that great human being that you are. The passion and ideas that you bring in the political arena are always great food for thoughts. I cannot wait to debate with you again and to see how the future will unfold for you. Jenny, just take care of this big guy. Cem, I appreciated all the time that you spent to support and listen to me. I am sure that a great future is ahead of you. Let me add an Indian appendix to this cluster. Abhi, I will never stop being amazed by your soccer knowledge. Besides our common interests in the same media outlets, I am happy that we could spend time together at coffee corner of the cheerful Italian. Neetha, I am truly happy to be your colleague/friend and I will always be wherever you end up working. The dinners and discussions that we had together were special moments in our lives and I will never forget that you were always available for my many surprise parties. I just hope that I could visit you in Kerala for your anniversary and participate to the party afterwards.

The third cluster is the The Flanders Awaken group of Dries, Simon and Ofelie. Simon, although we have to take into the constraints of Carolina's presence, you are one of the smartest person that I have ever met. I am sure that we should have quarreled more about science. However, I cannot stop thinking that you are one of the people that supported me the most in my first years spent in Flanders. Beatriz, please take care of this big boy for me. Dries, *I was wondering* which questions or sentence I could write about you. The only one that I have in mind is: Coffee, free refill? I just hope to travel once again together with you and an eco-monster truck on the American motorways. Honestly, I want to write more, but I think that I will send you a very long vocal message with the rest of my preface and a beautiful presentation in Jean-Luc Doumont's style. Ofelie, you are probably a cluster on your own. Let me apologize for my voice and for my behavior. Nobody knows like you how hard the first months in Belgium were for me. Everybody knew the cheerful Mario, but you had the will to listen to the sad Mario. You brought me tea and candies when I felt far away from home. You showed me your city and host me at your place. We cooked together so many cakes in double the time that was required. We ran trails and half-marathons together. You sent me a beautiful picture by post in this difficult lockdown. We went a long way together and I have a small request. I hope to meet you once again in our future and laugh at ourselves looking back at our PhD.

The Gym Boys, Jonathan and Jasper, make the fourth cluster. Mr. Wouters,

I will never forget the *mariofy.py* program and your tongue-in-cheek humour. However, I always appreciate your help at any given time and the effort to listen to my vocal messages. Your (hidden) kindness was never unnoticed and I am sure that this is your effort to make the world 10% happier. Mr. Moeyersons, your open-arms policy is the true gem of your virtues. The *inburgering* is the toughest challenge for someone who moves to another country, but you made it the best time ever. Lene (or Nele?), be sure that this boy keeps helping the world in the amazing way he does. Try also to free up space in his agenda. I can also attach Thomas De Cooman to this cluster. I am sure that you had to break your stay-away-from-me policy to get to know me. Many people knows you for the jokes and funny comments, but I am very grateful for the time that you spent with me beyond your comfort zone. Nele (or Lene?), just take care of him. I am not so sure what's the Euclidean distance between this cluster and those two guys, but let me also thank Lieven and Adrian IM. *Lieveni, ego eram vultus in amicitia et thesaurum in auxilio tuo inveni.* Adrian, you had a rough time to bear my voice and my presence, but let me thank you again for the friendship that you granted me. Who is missing? Wait... Alexander B., you are an outlier and you did not expect a location in this preface for you, for sure. Like many others, you had to endure my gravitational-sound waves in your office. Unlike many others, you clearly state our friendship and my heart simply melted for that. I cannot wait to see your future research achievements and the amazing PhDs that you are going to guide.

The Vlaamse Meisjes is the fifth cluster with Griet, Dorien, Kaat, Margot and Laure. Kaat, thanks again for the swimming time and I apologize for being competitive. Margot, I just hope that your future parties always have a DJ for the entrance of the wine. Dorien, keep staying the true Austrian that you are. Griet, you can blackmail me at any time. And yet, I am so proud how our friendship turned out to be. We truly share a lot of interests, such as a vivid interest in Politics. I am sure more than ever that I found a new friend. Laure, just never give up your ideas. Your passion and your interest in science is what makes you the great person that you are. Since I do not like homogeneous groups, let me add Ying and Bori. Ying, you were an amazing neighbor and the residence was never the same without you. Bori, any English sentence with a double meaning can be a threat to my freedom and you understood that I *am out like a balcony.* However, you always found the time to the listen to my very very long vocal messages. Thanks once again. Just say to Adhi: *Beautiful there!* This cluster has also an internal module of Honor students, Elisabeth and Laura. Girls, you were for sure brave to choose me as supervisor. Let me say that I am proud to see your contribution to the scientific community.

The NeoGuard people is the seventh cluster with Amir, Tim and Nick. Amir, the adventures that we had together are legendary. The time that we spent

together in the office was less legendary for you, but my research would have never seen the light of the day without your μ-God help. Neda, take care of this guy. Tim and Nick, you were unlucky to meet me and I apologize if I pulled your leg. I am happy that we shared time together and let me say one thing: *Get the PhD done!*

The newcomers group is the eight cluster with Simon G., Jorian, Tim, Kenneth, Thomas S., Jonathan D., Stijn, Nithin, Miguel, while the tensor group is comprised of Nico, Rob and Martijn. I apologize for the loud voice and I am happy that we shared lunch together. Simon G. and Jonathan, you will always be newcomers for me. Nico, I am sure that a great future is ahead of you. I am happy that you could laugh and accept my extravagance. Rob, we had our differences, but I am sure that we are going to miss each other.

The Indo-Flemish crew is the cluster of Sheena, Harini, Natasha, Camille, Laura and Bobby. Sheena and Bobby, you are just my friends. I want to say out loud: friends! You gave me an unconditional trust without any scientific or working related interest. You love me for my strange way of life, doing sport or being loud. We discovered the world by running around it. I had with you one of the best trip of my life in Munich. I hope that we can do it again to remind us the great group of friends that we are. Harini and Natasha, I am truly grateful to all our gods for our friendship. You will never admit that, but we like each other and we cannot live without texting something ridiculous. I cannot wait to see how our next discussion turns out to be. Camille, you have one of the most amazing skin of all times and I hope that I always have the chance to hug you. Can I still add another person? Marleen, I was probably the craziest student that you have ever had. I am extremely grateful for you guidance. I am not only able to speak a better English thanks to you, but I can also deliver emotions. This skill is what brings bliss to our lives.

Il cluster Belgium Unposted-World of Lipsticks contiene Gaia, Federica, Manuela, Giorgia Carrà Carraramente Carrara, Agnese, Jessica. Ragazze, siete sicuramente le più belle ragazze del Belgio. Anzi, le più belle ragazze dell'Europa. Anzi, le più belle ragazze del mondo. Federica, sei molto meglio della persona disperata che credi di essere. La tua forza di volontà ti ha fatto conoscere una lingua ostica come l'olandese o ti ha portato in Vietnam con Gaia. Sappi che sono felice di essere tuo amico e sono sicuro che uscirai vittoriosa da tutte le faccende affaccendate in cui ti trovi. Gaia, sei davvero una persona unica e semplicemente non ti merito. Se ti guardi attorno, il mondo è pieno di uomini che ti potranno rendere felice. La tua bellezza, bontà e intelligenza sono perle rare e mi hanno permesso di sopravvivere nei momenti più difficili del lockdown. Farò sempre quel che posso per regalarti un sorriso (con un film Disney o un bento B) e ti aspetterò fino ai 40 anni, se questo ti rende felice. Sono convinto però che il futuro avrà molte sorprese in serbo per te.

Non dimentichiamoci il cluster del presidente! Parisi, Valentina, Themelis, Alessandra, Agnese, Melania, Giovanna, Giulio e la Prosciutto. Parisi, sei stato il primo ad ospitarmi in una casa accogliente nei primi mesi di quest'avventura e le cene (di tua mamma) sono state riferimento essenziale per la mia ricerca. Valentina, spero di poterti raggiungere ovunque facendo il pendolare su qualsiasi Freccia Rossa europeo. La nostra amicizia è cominciata sicuramente in maniera particolare. Spero solo che un giorno ci potremmo rivedere per una bella cena del Primo di Maggio (a base di coniglio?). Themelis, sei più bello e intelligente di quanto vuoi far credere. Alessandra, nulla potrà ripagare le cene nel più bell'attico di Leuven. Sappi che la tua bellezza, forza e intelligenza ti porteranno molto lontano, piú di quello che credi. Agnese, Giovanna, Melania, Giulio e Federica, le cene a casa vostra sono state fondamentali per riscaldare e colorare il grigiume delle giornate in Belgio. Notare che questo paragrafo ha una densità elevata del pattern "mangiare", in vero stile prof. Prosciutto (il cui riferimento è anche nella mia tesi adesso).

Non potevo non concludere con il cluster del Bettolino con Luca, Federica, Maura, Antonio, Alessio e la Ferri. Dovrei scrivere un milione di scuse per tutti i compleanni che mi sono dimenticato, ma il mio Italiano si è arruginito a tal punto che è meglio evitare (altro che *arena arenata*). Ragazzuoli, che ci piaccia o meno, il nostro legame è più eterno delle banche vaticane o dell'ASOCROMICHE (concedetemi la licenza poetica). Nella buona e nella cattiva sorte, mi avete sempre ascoltato e supportato anche se ho gli atteggiamenti da lupo solitario. Non vedo l'ora di tornare a fare un giro-pizza all'Oca Giuliva e non potrei essere più orgoglioso di chiamarvi amici.

E infine, il core-cluster o il gruppo più importante di tutti è la mia famiglia. Vito, non hai mai visto di buon occhio le mie scelte da girovago, ma le hai sempre rispettate. Apprezzo tutti i tuoi sforzi per accettarmi come sono e sono sicuro che dovunque il mio destino mi porterà sarai sempre al mio fianco. Elena, Leonardo, Massimo, Arianna, siete e rimanete i regali più grandi che la vita mi ha fatto. Per quanto io possa essere lontano, non smetterò mai di volervi bene. Mamma e Papà, quanto dolore vi ho provocato con le mie scelte. Spero che io possa essere una ragione di orgoglio e ogni successo che ho ottenuto e otterrò è soprattutto merito vostro. Isabella, l'unica vera donna della mia vita, la mia cantante preferita e la prossima popstar mondiale. Nel bene e nel male, sei e sei sempre stata la mia ancora di salvezza. Il nostro indissolubile legame ci supporta e ci rafforza per le nostre sfide. Ogni volta che ci separiamo è doloroso. Ma sappi che in ogni battaglia, in ogni sconfitta e in ogni vittoria tu rimani la stella polare e la colonna sonora che mi guidano nei prossimi passi.

Mario
Leuven, September 2020

Abstract

This thesis focuses on the design of various automatic signal processing algorithms to extract information from physiological signals of preterm infants. Overall, the aim was to improve the neurodevelopmental outcome of the neonate. More specifically, three main research objectives were carried out. The first objective was to describe the maturation of neonates during their stay in the neonatal intensive care unit. The second objective was to assess the stress and pain in premature infants and their impact on the development of neonates. The third objective was to predict developmental disabilities, such as autism.

The first part of this thesis presents an extensive overview of various developmental models to describe the maturation of premature infants. Three main strategies were proposed. The first strategy proposed an investigation of EEG connectivity networks. A variety of functional and effective connectivity methods were combined with an assessment of graph properties of the resulting EEG network. A set of topological and spectral indices were used to predict the age of the infants and to demonstrate that the functional and effective connectivity decreased with development, characterized by a shift from a small-world network to a more random network. The second strategy focused on the multifractal properties of the EEG signal, which were investigated to detect quiet sleep in premature infants on the one hand, and to describe developmental changes of the EEG on the other hand. The fractal paradigm showed a decrease of EEG regularity with increasing age. The last strategy reported the trends of heart-rate variability in premature infants to describe the development of the autonomic nervous system. Similar to the EEG analysis, the autonomic parameters showed a decrease of fractal regularity and an increase in short- and long-term variability of the tachogram. Most importantly, a full description of the autonomic maturation in premature has to take into account the roles of heart-drops, known as bradycardias, and new metrics to define the sympathovagal development.

The second part focuses on the quantification of perinatal stress in premature

infants. Firstly, the effect of stress load was investigated during bradycardia and hypoxic events. We established that stress-load and early-life experiences can enhance desaturations, EEG regularity and brain-heart connectivity during bradycardias. Secondly, we processed physiological background activity of premature infants without any pain elicitation. This unobtrusive approach showed a moderate association between physiological features and perinatal stress. More specifically, stress seemed to induce a more dysmature EEG, characterized by more discontinuity and slow-wave activity, a more synchronized EEG network, and a stronger brain-heart coupling. In addition, the effect of stress on preterm development during their hospital stay was examined and the discontinuous EEG seemed to persist throughout the infants' maturation.

The third part reports an application of early-life EEG analysis, in which we aimed to predict neurodevelopmental disabilities, such as autism spectrum disorder. Based on EEG dysmature traits in the first month of life, such as discontinuity and slow-wave persistence, we showed that different classification models were capable to predict the autism diagnosis at 24 months and other developmental abnormalities (such as a lower cognitive outcome). The quantitative analysis of EEG confirmed the central role of dysmaturity attributes (the lower EEG entropy, the higher EEG regularity and a greater brain network resilience) as the most predictive attributes of developmental abnormalities and disorders.

Beknopte samenvatting

Dit proefschrift richt zich op het ontwerp van verschillende automatische signaalverwerkingsalgoritmen om informatie te extraheren uit fysiologische signalen van premature baby's. Het voorname doel hierbij was het verbeteren van de neurologische ontwikkeling van de pasgeborene. Hiertoe werden drie onderzoeksdoelstellingen uitgevoerd. Het eerste doel was het beschrijven van de maturatie van de pasgeborenen tijdens hun verblijf op de neonatale intensive care unit. Het tweede doel was het beoordelen van stress en pijn bij premature baby's en het bepalen van de impact van beide op de ontwikkeling van pasgeborenen. De derde doelstelling was het voorspellen van ontwikkelingsstoornissen, zoals autisme.

Het eerste deel van dit proefschrift presenteert een uitgebreid overzicht van verschillende ontwikkelingsmodellen om de maturatie van premature baby's te beschrijven. Hierbij werden drie verschillende strategieën gevolgd. De eerste strategie bestudeerde EEG-connectiviteitsnetwerken. Een aantal functionele en effectieve connectiviteitsmethoden werden gecombineerd met de beoordeling van verschillende graafeigenschappen van het resulterende EEG-netwerk. Een verzameling van topologische en spectrale indices werd gebruikt om de leeftijd van de zuigelingen te voorspellen en om aan te tonen dat de functionele en effectieve connectiviteit afnam met ontwikkeling. Qua structuur werden de netwerken gekenmerkt door een verschuiving van een 'small-world' netwerk naar een meer willekeurig netwerk. De tweede strategie bestudeerde de multifractale eigenschappen van het EEG-signaal, enerzijds om stille slaap bij premature kinderen te detecteren en anderzijds om de ontwikkelingsveranderingen van het EEG te beschrijven. Het fractale paradigma toonde een afname van de EEG-regelmaat met toenemende leeftijd aan. De laatste strategie rapporteerde de hartslagvariatie bij premature baby's om de ontwikkeling van het autonome zenuwstelsel te beschrijven. Gelijkaardig aan de EEG-analyse, vertoonden de autonome parameters een afname van de fractale regelmaat en een toename van de variabiliteit van het tachogram, zowel op de korte als op de lange termijn. Belangrijk hierbij is dat een volledige beschrijving van de autonome maturatie

bij premature baby's rekening moet houden met de rol van trage hartslagen, bekend als bradycardieën, en met nieuwe meetgegevens om de sympathovagale ontwikkeling te definiëren.

Het tweede deel van dit proefschrift richt zich op het kwantificeren van perinatale stress bij premature baby's. Hiertoe werd allereerst het effect van stressbelasting onderzocht in periodes van bradycardie en hypoxie. Hier stelden we vast dat stressbelasting en vroege levenservaringen desaturaties, de regelmaat van het EEG en de connectiviteit tussen de hersenen en het hart in periodes van bradycardie kunnen vergroten. Vervolgens onderzochten we de fysiologische achtergrondactiviteit van premature baby's zonder pijn. Deze onopvallende benadering toonde een matig verband tussen fysiologische kenmerken en perinatale stress. Stress leek een meer dysmatuur EEG, gekenmerkt door meer discontinuïteit en langzame golfactiviteit, een meer gesynchroniseerd EEG-netwerk en een sterkere hersen-hartkoppeling te veroorzaken. Daarnaast werd het effect van stress op de vroegtijdige ontwikkeling van de neonaat tijdens het ziekenhuisverblijf onderzocht en leek de discontinue EEG aan te houden tijdens de maturatie van de neonaten.

Het derde deel van dit proefschrift rapporteert een toepassing van EEG-analyse op jonge leeftijd, met als doel het voorspellen van neurologische ontwikkelingsstoornissen, zoals autismespectrumstoornissen. Op basis van EEG-dysmaturiteitskenmerken in de eerste levensmaand, zoals discontinuïteit en trage-golf persistentie, toonden we aan dat verschillende classificatiemodellen zowel de autismediagnose na 24 maanden, als andere ontwikkelingsstoornissen (zoals een lagere cognitieve uitkomst) konden voorspellen. De kwantitatieve analyse van het EEG bevestigde de centrale rol van de dysmaturiteitskenmerken (een lagere entropie van het EEG, een hogere regelmaat van het EEG en een grotere veerkracht van het EEG-netwerk) als de meest voorspellende attributen van ontwikkelingsstoornissen.

List of Abbreviations

$\mathbf{5_{days}}$	Recording around 5 days from birth
$\mathbf{34_{weeks}}$	Recording at 34 PMA weeks
ACF	Autocorrelation function
AS	Active Sleep
AUC	Area under the ROC curve
BSID-III	Bayley Scales of Infant and Toddler Development
CA	Corrected Age
CCA	Canonical Correlation Analysis
CWT	Continuous wavelet transform
GA	Gestational Age
GC	Granger Causality
HF	High frequency band of the tachogram
HRV	Heart-rate variability
HSD	Hilbert-Schmidt dependence Criterion
IBI	Inter-bursts intervals
ICA	Independent Component Analysis
IC	Imaginary Coherence
IVH	Intraventricular hemorrhage
LASSO	Least absolute shrinkage and selection operator
LDA	Linear discriminant analysis
LF	Low frequency band of the tachogram
LME	Linear Mixed-Effects
LPS	Leuven Pain Score

LS-SVMs	Least-squares support
MAE	Mean Absolute Error
MFA	Multifractality
MSC	Mean Squared Coherence
MSE	Multiscale Entropy
MRI	Magnetic Resonance Imaging
NICU	Neonatal Intensive Care Unit
NQS	Non-Quiet Sleep
OLS	Ordinary least-squares
PLI	Phase Lag Index
PLV	Phase Locking Value
PMA	Post-menstrual Age
PSD	Power Spectral Density
PSG	Polysomnography
PVL	Periventricular leukomalacia
QS	Quiet Sleep
RMSE	Root Mean Squared Error
ROC	Receiving operating characteristic
SampEn	Sample Entropy
SAT	Spontaneous activity transients
SBP	Skin Breaking Procedures
SS	Singularity Spectrum
SVMs	Support Vector Machines
TE	Transfer Entropy
VLF	Very-Low frequency band of the tachogram

List of Symbols

δ	EEG band [0.5 - 4) Hz
θ	EEG band [4 - 8) Hz
α	EEG band [8 - 16) Hz
β	EEG band [16 - 32) Hz or [16 - 20) Hz
$\varphi(x)$	Feature map applied to x
ρ	Pearson's Correlation coefficient
τ	Scale Factor
$\tau(q)$	Scaling Exponent
μ_{RR}	Mean of the RR time series
σ_{RR}	Standard Deviation of the RR time series
$\mathbf{A}$	Adjacency Matrix
A_{ij}	Entries of the Adjacency Matrix
c_1	Main Hurst Exponent or maximum location of $D(h)$
c_2	Width of $D(h)$
c_3	Asymmetry of $D(h)$
$D(h)$	Singularity Spectrum
$\mathcal{F}$	Granger Causality F-score
H	Hurst Exponent
HF	Tachogram band (0.2-4] Hz
$k^2_{xy}(f)$	Mean Squared Coherence
$K(x, x_k)$	Kernel function
$I_{xy}(f)$	Imaginary Coherence
m	embedding dimension
$MSE(3)$	Multiscale Entropy at scale 3
$MSE(20)$	Multiscale Entropy at scale 20
n_{sup}	number of superfluous connections
LF	Tachogram band (0.08-0.2] Hz
$P_{xy}(f)$	Cross-Spectrum of signal x and y
$P_{xx}(f)$	Spectrum of signal x
R^2	Coefficient of Determination

r	Tolerance for sample entropy computation
RR_i	RR intervals
$S_{xx}(t,f)$	Time-frequency Auto-Spectrum
$S_{xy}(t,f)$	Time-frequency Cross-Spectrum
SS	Sum of squares of estimated $\hat{y}$
$x, s, a, b...$	vectors
$x_i{}_{i=1}^{N}, x_i \in \mathbb{R}$	Set of N points
VLF	Tachogram band (0.0033-0.08] Hz
$W(t,f)$	Wavelet Transform
$Z(a,q)$	Partition function

Contents

Abstract ix

Beknopte samenvatting xi

List of Abbreviations xiii

List of Symbols xv

Contents xvii

List of Figures xxiii

List of Tables xxxix

I Introduction 1

1 Introduction 3
- 1.1 Problem Statement . . . 3
- 1.2 The Resilience study and other collaborations . . . 5
- 1.3 Outline of the thesis . . . 7
- 1.4 Conclusion . . . 9

2 Physiological Introduction: Premature birth, functional monitoring and perinatal stress 11
- 2.1 Premature birth and development outcome . . . 11
- 2.2 EEG: functional monitoring of the brain . . . 15
  - 2.2.1 EEG recordings: Conventional EEG and aEEG . . . 15
  - 2.2.2 EEG Maturational Patterns . . . 16
  - 2.2.3 Abnormal EEG . . . 21
- 2.3 HRV: functional monitoring of the autonomic nervous system . 22

2.4 A definition of dysmature EEG and HRV 25
2.4.1 The dysmature EEG 27
2.4.2 The dysmature HRV 27
2.5 Pain, Stress and developmental outcome 28
2.6 Summary 30

3 Mathematical background **31**
3.1 Univariate Analysis 31
3.1.1 Spectral analysis 32
3.1.2 Entropy Measures 33
3.1.3 Multifractality 36
3.2 Multivariate analysis 41
3.2.1 Functional Connectivity 42
3.2.2 Effective Connectivity 50
3.2.3 Surrogates and phase analysis 58
3.2.4 Graph indices 59
3.3 Regression Models 65
3.3.1 Ordinary Least Squares regression 65
3.3.2 Linear mixed-effect regression 67
3.3.3 LASSO 69
3.4 Classification Models 70
3.4.1 Support vector machines 70
3.4.2 Least squares support vector machines 74
3.4.3 Linear discriminant analysis 76
3.4.4 Fixed-size LS-SVM 77
3.4.5 Subspace Linear discriminant analysis 79
3.4.6 Performance Metrics 79
3.5 Summary 81

II Maturation of the premature infant **83**

4 A brain-age model for preterm infants based on EEG connectivity **85**
4.1 Introduction 86
4.2 Methods 89
4.2.1 Data 89
4.2.2 EEG Data 90
4.2.3 Connectivity analysis 91
4.2.4 Algorithmic pipeline and statistical analysis 92
4.3 Results 97
4.3.1 Simulated data 97
4.3.2 EEG data: Effective connectivity indices 99
4.3.3 EEG data: Functional connectivity indices 99

4.3.4 Regression results . . . 103
4.4 Discussion . . . 105
4.4.1 Simulated dataset . . . 108
4.4.2 EEG data . . . 109
4.5 Summary . . . 112

5 A multifractal framework for quiet sleep detection and the estimation of age 113
5.1 Introduction . . . 114
5.1.1 Dataset . . . 115
5.1.2 Multifractal features and classical EEG features . . . 116
5.1.3 Classification approach . . . 117
5.1.4 Regression model . . . 118
5.2 Results . . . 118
5.3 Discussion . . . 122
5.4 Summary . . . 125

6 The maturation of the autonomic nervous system in premature infants 127
6.1 Introduction . . . 128
6.2 Methods . . . 130
6.2.1 Dataset . . . 130
6.2.2 Preprocessing . . . 130
6.2.3 Feature extraction . . . 132
6.2.4 Multifractal analysis . . . 134
6.2.5 Algorithmic pipeline and statistical analysis . . . 135
6.3 Results . . . 137
6.4 Discussion . . . 140
6.5 Summary . . . 147

III Perinatal Stress quantification 149

7 A bradycardia-based stress calculator for the neonatal intensive care unit: a multisystem approach 151
7.1 Introduction . . . 152
7.2 Material and methods . . . 154
7.2.1 Patient sample . . . 154
7.2.2 Data acquisition . . . 154
7.2.3 Bradycardia detection and data preprocessing . . . 155
7.2.4 Features extraction . . . 156
7.2.5 Bradycardia-based Classification . . . 160
7.3 Results . . . 162

7.4 Discussion . . . 166
7.5 Summary . . . 170

8 A perinatal stress calculator for the neonatal intensive care unit: an unobtrusive approach 171
8.1 Introduction . . . 172
8.2 Materials and methods . . . 173
8.2.1 Patient sample . . . 173
8.2.2 Data acquisition . . . 174
8.2.3 Preprocessing and sleep-stage analysis . . . 175
8.2.4 Univariate features extraction . . . 176
8.2.5 Multivariate features . . . 179
8.2.6 Sleep-based Classification . . . 181
8.3 Results . . . 184
8.4 Discussion . . . 185
8.5 Summary . . . 195

9 The description of the physiological maturation under early-life procedural pain 197
9.1 Introduction . . . 198
9.2 Methods . . . 200
9.2.1 Participants . . . 200
9.2.2 Data acquisition . . . 201
9.2.3 Electrophysiological data collection . . . 201
9.2.4 Functional measurement . . . 202
9.2.5 Statistical analysis . . . 204
9.3 Results . . . 206
9.3.1 Patients demographics . . . 206
9.3.2 Relationship between EEG and HRV functional variables and early skin breaking procedure . . . 206
9.4 Discussion . . . 210
9.5 Summary . . . 214

IV Autism spectrum disorder and development abnormalities 215

10 Early-life EEG quantitative analysis to predict development disorders in young infants 217
10.1 Introduction . . . 218
10.2 Material and methods . . . 219
10.2.1 Patient sample . . . 219
10.2.2 EEG recordings . . . 219

10.2.3 Neuropsychological assessment 220
10.2.4 Quantitative EEG analysis 220
10.2.5 ASD and development outcome prediction 225
10.3 Results 230
10.3.1 Automatic Classification of ASD patients 230
10.3.2 Automatic Classification of abnormal outcomes 231
10.3.3 Multiclass analysis of abnormal Bayley outcomes, ASD and normal patients 232
10.3.4 Regression analysis for the Bayley score 237
10.4 Discussion 238
10.5 Summary 241

V Conclusions 243

11 Conclusions and future directions 245
11.1 Conclusions 245
11.1.1 Preterm development: the assessment of the brain and autonomic maturation 246
11.1.2 Perinatal stress quantification 247
11.1.3 Later development outcome: the effect of stress and the importance of early-life EEG 248
11.2 Future directions 248
11.2.1 Graph theory 248
11.2.2 Network-Physiology age model 249
11.2.3 Stress classification 250
11.2.4 Stress Maturation 251

A Supplementary Tables of autonomic maturation trends 253

Bibliography 257

Curriculum Vitae 285

List of Publications 287

List of Figures

1.1 Thesis architecture chapter-by-chapter. The central parts focus on physiological or functional maturation of the healthy premature infants, the classification of stress and pain in infants and the relationship between functional trends and the development outcome of patients. 10

2.1 The figure show the multiple definition of age that are used in neonatology. The gestational age is the amount of time that the fetus stays in the womb before the delivery and the counting starts from the last menstrual period of the mother. The chronological age is the number of days after birth to the observation point, while the postmenstrual age is the sum of the former two. At last, the corrected age is computed as the number of days between the expected day of delivery and the observation point. 14

2.2 The figure reports the electrodes displacement on the human scalp according to the 10-20 international system. The grey circles and arrow show the electrodes that were used in the analyzed EEG recording of this thesis. Reprinted from [163]. 16

2.3 The figure reports the functional maturation trend of the premature infants in function of postmenstrual age. The charts are adapted from [261],[8],[197]. The top part reports the evolution of the sleep-wake cyclicity. The middle part reports the changes in EEG information (such as discontinuity and amplitude) and type of transients, while the last part shows the evolution of the different EEG graphoelements. The label AS stands for the active sleep trend, while QS is the quiet sleep and IS is the indeterminate sleep. 18

2.4 An example of discontinuous EEG, where by period of activities, known as bursts or spontaneous transient activities (SATs), and period of quiescence, known as interbursts intervals (IBIs), are respectively shown in blue and green. 19
2.5 Overview of the different graphoelements, which characterizes the neonatal EEG. The top left shows delta waves, commonly present in the first days of life and a common sign of dysmaturity, while the top right shows delta brushes, which are bursts in the delta band with superimposed faster information and emerge also as a sign of somatosensory cortex modulation. The bottom left shows theta temporal sawtooths, which are normally present in the infants below 30 weeks of gestation, while the bottom right shows frontal transients, which are good sign of maturation of the frontal area. 21
2.6 The figure shows the nervous control of the heart rate via the vagal and sympathetic innervation of the autonomic nervous system. Besides the conduction velocity, their activity acts as controller of the sino-atrial node anticipating or delaying the depolarization of the heart. The net effect is a variation of distance between the R-peaks (ventricle repolarization) or beat-to-beat change of R-peak location. The variations of R-peak-to-R-peak intervals and its inverse (known as instantaneous hear-rate) are known as heart-rate variability or tachogram. Reprinted from [186]. . . . 24
2.7 Functional trends of the respiratory and cardiovascular system. The charts are adapted by [58]. The upper part shows the evolution of the respiratory information, with a sharp decrease in apnea-frequency with aging and a u-shaped change of the respiratory rate. The lower part shows the cardiovascular trends, which are a decrease in heart-rate and increases of the low-frequency and high-frequency power. The label AS stands for the active sleep, while QS is quiet sleep. 26

3.1 The two steps to compute the Multiscale Entropy. The left panel shows the coarse-graining procedure, which is an averaging and downsampling at each scale (as shown by the blue and green dots). The right panel shows the computation of the $SampEn$, which is repeated at each scale. The pattern match is indicated by the blue lines and the color-coding of the red,green and blue samples. The figure is adapted from [65]. 35

3.2 Example of MSE curves computed for the following sets of signals: white Gaussian noise (golden diamonds), pink noise (pink circles), Brownian motion (green squares), an example of squared modulated (blue hexagons) noise and an example of EEG (silver stars). The white Gaussian noise have a decreasing entropy, the pink noise have a persistent entropy over all scales and the two regular signals have increasing entropy over the different scales. The neonatal EEG rapidly increases for the short-term scales and reach a plateau of entropy at higher scales similar with a comparable $Samp_{En}$ to the pink noise. 37
3.3 Example of singularity spectrum obtained with the multifractal framework (black line). The main multifractal parameters obtained from the singularity spectrum are the following: the cumulant c_1 or H represents the location of the $D(h)$ maximum, c_2 the width of the $D(h)$ distribution and c_3 is its asymmetry, which is intended as deviation from a concave parabola centered on c_1 (grey dashed lines). The extremes H_{max} and H_{min} respectively represent the maximal regularity and the minimal regularity. Another measure of the width of SS obtained as $\Delta H = H_{max} - H_{min}$. 38
3.4 Examples of two singularity spectra $D(h)$. The top panel shows a Brownian motion with one single Hurst exponent, while the bottom panel displays modulated noise obtained by the product between white noise and periodic Gaussian waves (green curve). The right panel shows the singularity spectra of the Brownian motion (purple circles) and of the modulated noise (green diamonds). The latter has a very wide $D(h)$ and higher global H, while Brownian motion has a narrow SS around its real Hurst exponent. 40
3.5 Examples of application of the coherence $C_{xy}(f)$ and the phase-locking vector PL_{xy} on two sines at the same frequency without any delay. The coherence will be a real number on the complex plane (the red bar in the top panel) as well as the PL_{xy} (the red bar in the bottom panel). 45
3.6 Examples of application of the coherence $C_{xy}(f)$ and the phase-locking vector PL_{xy} on two sines at the same frequency with a phase delay $\Delta\phi = \frac{\pi}{2}$. The coherence will be an *imaginary* number on the complex plane (the red bar in the top panel) as well as the PL_{xy} (the red bar in the bottom panel). 46

3.7 An explanation of the meaning of Granger Causality. The uncertainty of the prediction of the signal EEG_Y is given by the two funnels in grey. The future of Y_t has greater variance (and hence uncertainty) if we only consider the the past samples of Z_t. This estimation risk $\sigma^{(r)}$ will diverge the longer the time span that one wants to investigate into the future (as indicated by the bigger black dashed double arrow). On the contrary, the variance of estimation $\sigma^{(u)}$ is smaller in case we also consider the past samples X_{t-k} and the divergence of the estimation will be limited (as indicated by the smaller double blue arrow). 52
3.8 The typical processing pipeline that leads a multivariate analysis to a weighted graph. Given a starting set of signal $S_j = x_j$, one computes the $M \times M$ pairwise interactions $C_{x_i \leftrightarrow x_j}$ and derives the adjacency matrix $\mathbf{A}$. The latter can provide the information to draw the associated weighted graphs: each column represents a node of the matrix and the entries of the matrix represent both the existence and the weight of each link. The intensity of the coupling is normally highlighted with a different thickness of the edge. 60
3.9 Three examples of graphs and key organizational properties that can be computed with graph theory. The figure has been adapted from [43]. The two top figures show some key topological properties: the left graph shows the meaning of triangle motifs, while the right one reports an example of path length and a node with high clustering coefficient. The bottom graph shows a graph with different distinctive spectral properties: the two clear-cut communities will increase the algebraic connectivity of this graph, while the dominant hub on the right will lead to a high spectral radius and spectral gap. 65
3.10 Linear mixed effect model examples 69
3.11 The SVMs find the optimal hyperplane by maximizing the margin between the two classes. On the left, the perfect linearly-separable case is depicted and, on the right,the linearly-separable case with some tolerance expressed by the slack variable. 75
3.12 Example of application of a fixed-size LS-SVM. The toy-example shows that the classification between red squares and the blue circles could be solved by using all the data points in the two spirals. However, one can select fewer data points on two spirals (highlighted by the bigger markers) and would still be able to perform a classification. The M markers is normally selected with (3.96) to resemble the original dataset. 78

3.13 An example of a Receiver Operating Characteristic curve (green dash curve). The grey dot-dashed line is an example of a random classifier and the blue square curve represents an example of perfect classifier. 81

4.1 The figure displays the graph associated to the model 4.1. . . 90
4.2 The figure shows the results for the simulation dataset. The first three panels show how the graph indices behave for different level of coupling in the model 4.1. The last three panels investigate how the transfer entropy can estimate the network indices for different level of SNR. In particular, the figure compares the two cases when the data is filtered and when the raw data is used. . 98
4.3 The figure shows the results for EEG data. The first four panels show OLS regression between 4 main graph indices vs the age for GC in QS. The grey area is the confidence interval at 95%. On the top of the panel, the associated R^2 and $\sqrt{MSE}$ in PMA weeks on the test set. The last two panels show the trend of the clustering coefficient and the path length in three distinct age groups. The results are reported about GC in QS. 102
4.4 The figure shows the average connectivity graph (GC for three different age groups. The strength of the coupling among the electrodes is decoded by the color (the closer to the red color, the higher the coupling) and by the width of the arrow. The connectivity values have been normalized between 0 and 1 for the three groups together. The panels cleraly show the weakening of the coupling among EEG channels with maturation. The consequence is the increase of path length and the decrease of the clustering coefficient Figure 4.3. 103
4.5 The figure shows the linear mixed-effect regression between the clustering coefficient and the age (Panel a), and between path length and the age for the $ImCoh$ in θ bands (Panel b) during QS. Specifically, the black continuous line represents the estimated regression line, while the dashed lines are the confidence intervals. Each color in the plot represents a different patient. In order to show the specific trend for a single patient, we highlighted the result for a chosen infant in thicker red line. Above each panel, we reported the Pearson correlation coefficient and the associated p_{value}. 104

4.6 The figure shows the linear mixed-effect regression between different MSC density indices and the age. Panel a) displays the density in posterior hemisphere in the β band during only QS, while panel b) shows the density in anterior hemisphere in the θ band during QS. The black continuous line represents the estimated regression line, while the dashed lines are the confidence intervals. In this figure, we reported only the data cluster without highlighting each patient. Above each panel, we reported the Pearson correlation coefficient and the associated p_{value}. 104
4.7 Topographic distribution for $ImCoh(\theta)$ in QS: PMA $\leq 31\ Weeks$, PMA $\in (31-37)\ Weeks$, PMA $\geq 37\ Weeks$. 107

5.1 Example of EEG signal during QS and NQS epoch (prematurely born neonate at 42w PMA). 116
5.2 Median and 25-75 percentile for the c_2 values in QS (dashed line) and NQS (continuous line) epochs. On the left panel, the values from the dataset with the youngest neonates. On the right panel, the values from the dataset with the oldest neonates. 119
5.3 The left panel shows an example of a mean singularity spectrum in QS epochs (stars) and NQS epochs (diamonds) for one specific recording (prematurely born neonate at 42w PMA). The right panel shows the ROC curve for LS-SVM classifiers, using $\mathbf{X}_{c_2,\Delta H}$ as feature set, in the three different age groups. The continuous line represents the ROC curve for all recordings in the youngest group, the dotted line represents the ROC for all recordings in the middle age group, while the dashed line represents the ROC curve for all recordings in the oldest group. 119
5.4 Results for the linear-mixed effect regression that models the relationship between EEG multifractal features and the post-menstrual age. Specifically, the Hurst exponent for T_3, the c_2 and c_3 for the same channel and the difference ΔH for the channel Fp_1 (H_{exp,T_3},c_{2,T_3},c_{3,T_3} and ΔH_{Fp_1}) are displayed. The left column reports the result for the non-quiet sleep (NQS), while the right column reports the result for the quiet sleep (QS) epoch. The symbol ρ is the correlation coefficient of the regression and p_{value} represents the significance of the correlation. 123

6.1 Visual representation of the three windowing schemes applied in this investigation: postbradycardia scheme (PB), between-bradycardia scheme (BB) and within-bradycardia scheme (WB). The selected windows in each trace are indicated with grey dashed boxes, while the dot-dashed boxes show examples of annotated bradycardia. In case of BB windowing, a period T greater that 10 minutes is present between the end of the bradycardia and the first available window. 133

6.2 The block diagram shows the main steps of the age estimation. For each RR signal, artifact preprocessing is performed and associated resampling of the tachogram. The signal is split in different windows according to the scheme of Figure 6.1. For each of these epochs, temporal, spectral and fractal features, which undergo a grand-median process if there is more than one epoch per scheme. The three datasets are then used to estimate the age of the recording in a linear mixed effects (LME) regression. 138

6.3 The figure shows the linear-mixed effect regressions between the post-menstrual age and the following HRV features: the standard deviation of the tachogram σ_{RR}, the absolute and the relative power in the LF band $\left(\log_{10}(LF), \frac{LF}{LF+VLF}\right)$, the Hurst exponent $H_{exp,[j_1,j_2=5,12]}$ and the parameter C_2. The sampling frequencies for the fractal indices is $f_s = 8\ Hz$. The left column - magenta circles report the results for the bradycardia epochs, while the right column - indigo diamonds the results for the between-bradycardia epochs. ρ is the correlation coefficient with the associated significance p_{value}. 141

6.4 The figure shows results for the linear-mixed effect regression that models the relationship between $H_{exp,[j_1,j_2=5,12]}$ and $\frac{VLF}{LF}$ in the first row and between $H_{exp,[j_1,j_2=3,12]}$ and $\frac{LF}{HF}$ in the second row. The three columns respectively represents bradycardia epochs (magenta circle data points), post-bradycardia epochs (light-blue squares data points) and between-bradycardia epochs (indigo diamonds data points). ρ is the correlation coefficient of the regression and p_{value} represents the significance of the correlation. 143

7.1 Results of the bradycardia-based classification in three main datasets. The three colors represent different level of desaturation to consider the bradycardic event in the stress classification. The left panel displays the *area under the curve* in the three monitoring groups, while the right reports Cohen's kappa. . . . 164

7.2 Results of the bradycardia-based classification in three main datasets. The figure here reports the results based on the different feature groups. The left panel reports the area under the curve for desaturations greater than 3%, while the right panel report the results for desaturation greater than 10%. The three colors represent different feature groups: EEG stands for EEG features, HR-SpO_2 represent the cardiovascular features and B-H is related to the brain-heart connectivity. 165
7.3 The desaturation levels and the EEG regularity are more pronounced in case of stress. The left panel reports the SpO_2 during the bradycardic spell and the right panel shows the boxplot for the Hurst exponent of channel C_3 for the period before and after each bradycardia. The data are reported for the 5days group. All the events with a desaturation greater than 10% were included in this figure. The p-values in the boxplot are derived via the Kruskal-Wallis test. 165
7.4 The desaturation levels and the connectivity between delta oscillations and the heart-rate are more pronounced in case of stress. The left panel reports the SpO_2 during the bradycardic spell and the right panel shows the path length derived from the network with EEG channels and the HRV. It is important to remind that the lower the path length, the higher the connectivity. The data are reported for the 34w group. All the events with a desaturation greater than 10% were included in this figure. . . 166
7.5 The intensity of bradycardias and the parasympathetic activity are more pronounced in case of stress. The left panel reports the normalized HRV power in the HF band and the right panel shows the normalized power in boxplots before, during and after each bradycardic event. The data are reported for the PSG group. All the events with a desaturation greater than 10% were included in this figure. The p-values in the boxplot are derived via the Kruskal-Wallis test. 166

8.1 Performance for the stress classifier in function of the Leuven Pain Score (LPS) collected the day before the recording. The results are reported for the three monitoring group for the non-quiet sleep epoch: the recordings within 5 days from birth (5_{days}), the recordings at 34 *weeks* (34_{weeks}) and the polysomnographies at discharge (PSG). The left panel displays the area under the curve, while the right one reports the kappa score. The legend reports the threshold applied on the LPS to define the stress group ($LPS > 0$, $LPS > 1$, $LPS > 4$). 185

8.2 Performance for the stress classifier in function of the Leuven Pain Score (LPS) collected the day before the recording. The results are reported for the three monitoring group for the quiet sleep epoch: the recordings within 5 days from birth (5_{days}), the recordings at 34 *weeks* (34_{weeks}) and the polysomnographies at discharge (PSG). The left panel displays the area under the curve, while the right one reports the kappa score. The legend reports the threshold applied on the LPS to define the stress group ($LPS > 0$, $LPS > 1$, $LPS > 4$). 186
8.3 Performance for the stress classifier in function of the Leuven Pain Score (LPS) collected the day before the recording. The results are reported for the 4 post-menstrual age groups for the non-quiet sleep epoch: ≤ 32 *weeks*, $(32-34]$ *weeks*, $(34-36]$ *weeks*, > 36 *weeks*. The left panel displays the area under the curve, while the right one reports the kappa score. The legend reports the threshold applied on the LPS to define the stress group ($LPS > 0$, $LPS > 1$, $LPS > 4$). 186
8.4 Performance for the stress classifier in function of the Leuven Pain Score (LPS) collected the day before the recording. The results are reported for the 4 post-menstrual age groups for the quiet sleep epoch: ≤ 32 *weeks*, $(32-34]$ *weeks*, $(34-36]$ *weeks*, > 36 *weeks*. The left panel displays the area under the curve, while the right one reports the kappa score. The legend reports the threshold applied on the LPS to define the stress group ($LPS > 0$, $LPS > 1$, $LPS > 4$). 187
8.5 Classification performance for each set of features in discriminating stress with threshold $LPS > 0$. The results are reported for the 4 post-menstrual age groups for the quiet sleep epoch: ≤ 32 *weeks*, $(32-34]$ *weeks*, $(34-36]$ *weeks*, > 36 *weeks*. The left panel displays the area under the curve in non-quiet sleep, while the right one reports the AUC for quiet sleep. The legend reports the four sets of attributes : heart-rate variability features (HRV), EEG features (EEG), EEG connectivity features (CONN), brain-heart synchrony features (BH). 187
8.6 EEG complexity shows a marked decrease in case of stress during nQS for patient with *post-menstrual age* < 32 *weeks*. The left panel reports the topoplot of the complexity index on the scalp (area under MSE curve), while the right one reports the boxplot of the complexity index for each channel (p-value reported with Kruskal-Wallis test). The reduction of EEG complexity can be explained by the greater discontinuity in case of stress. 188

8.7 The synchrony in the θ band shows a marked increase in case of stress during nQS for patient with *post-menstrual age* < 32 *weeks*. The left panel displays the topoplot for the eccentricity on the scalp (computed with *phase lag index* in θ band), while the right one reports the boxplot of eccentricity for each channel (p-value reported with Kruskal-Wallis test). The eccentricity is a measure of distance, therefore the lower the eccentricity the higher the connectivity. 188
8.8 The rEEG showed a marked increase in case of stress during nQS for patients with *post-menstrual age* (PMA) between 32 and 36 weeks. The left panel reports the boxplot of rEEG lower margin for the different frequency bands in patients with PMA between 32 and 34 weeks, while the right one reports the boxplot of rEEG asymmetry for the different frequency bands with PMA between 34 and 36 weeks (p-value reported with Kruskal-Wallis test). . 189
8.9 The δ_2 power showed a marked increase in case of stress ($LPS > 0$) during QS for patients with post-menstrual age > 36 *weeks*. The left panel displays the topoplot of the δ_2 power on the scalp, while the right one reports the boxplot of δ_2 power for each channel (p-value reported with Kruskal-Wallis test). 189
8.10 The Brain-heart connectivity measured via wavelet coherence in the LF band during nQS for patients with a *post-menstrual age* between 34 and 36 weeks. Stress exposure increases coupled dynamics between brain and heart. However, most of the linked dynamics is shared between EEG channels C_3 and C_4. The left panel reports the graph of brain-heart network, while the right panel displays a boxplot for its eccentricity. 190
8.11 The synchrony in the δ_1 band shows a marked increase in case of stress during nQS for patients with *post-menstrual age* > 36 *weeks*. The left panel reports the topoplot of the eccentricity (computed with phase lag index in δ_1 band) on the scalp, while the right one reports the boxplot of the eccentricity for each channel (p-value reported with Kruskal-Wallis test). The eccentricity is a measure of distance, therefore the lower the eccentricity the higher the connectivity. 190

8.12 EEG complexity in nQS (panel a) for patients with post-menstrual age ≤ 32 *weeks*, the range EEG asymmetry in nQS (panel b) for patients in the range between 34 and 36 weeks and the δ_2 power during QS for patients with post-menstrual age > 36 *weeks* (panel c). Data are reported in three groups $LPS = 0$, $LPS = [1 - 4]$, $LPS = 4$. The p-values are obtained with a Kruskal-Wallis test and the multicomparison tests significance is expressed by $(^{**})$ for $p \leq 0.01$ and $(^{*})$ for $p \leq 0.05$. Similarly to the binary comparison, higher stress values are related to higher δ_2 power and *asymmetry* as well as a lower complexity. These results show how stress can increase the level of dysmaturity of EEG with a higher level of cumulated pain. 191

9.1 The chart of the skin breaking procedure is displayed as average over the different patients. The curve shows the SBP dynamics during the neonatal intensive care unit stay. The average chart is represented by the continuous thick line, while the upper bound and the lower bound represent the error margin. The curve has been defined as $\mu(t) \pm \frac{\sigma(t)}{\sqrt{N(t)}}$, where $\mu(t)$ is the mean of the patient distribution, $\sigma(t)$ is the standard deviation and $N(t)$ represents the number of patient. The variable t represents each postnatal day. The exponential decay highlights that most procedural pain is concentrated in the first days of life. 209

9.2 Complexity Index in non-quiet sleep. The panel a) reports how complexity index trends with development differ in case of high or low amount of skin breaking procedures. The magenta curve shows the trend with SBPs above 50 or HIGH SBPs, while the green curve shows the maturation curve for SBPs below 50 or LOW SBPs. Both the curve shows an increasing CI with increasing age, but the magenta curve has slower development (lower slope) and lower CI at full-term age. The blue/green triangles represent the patient with low SBPs ($N = 73$) and red/orange circles the HIGH SBPs ($N = 19$). The 3D plane confirms the effects on complexity index by stress. The panel b) represents the cloud of complexity index data points for each PM age and each SBPs value and the associated linear plane fitting. The highest value of complexity on the plain is obtained for the oldest age and the lowest SBPs, while the lowest complexity is reached for the youngest age and the highest SBPs. This aspect is also highlighted by the color scale, which goes from blue for the lowest value to yellow for the highest value. The yellow is reported for the highest complexity value on the top left of the plane, while the color blue is reported on the right bottom of the same plane. 210

9.3 The figure shows the association between rEEG asymmetry in the θ band (asymm) and early skin-breaking procedure (SBPs) (Left Panel A) and the association between HF oscillations of the tachogram P(HF) and SBPs (Right Panel B). Data are reported for the entire dataset during QUIET sleep (QS). The panels show how the asymm and P(HF) maturational trend differs in case of HIGH SBPs (magenta curve, SBP ≥ 50) compared to LOW SBPs (green curve, SBP < 50). SBPs seem to increase the level of asymm throughout the development as well as the power of HF oscillations. 211

9.4 rEEG Asymmetry and HF Oscillations: the most vulnerable patients. The figure shows the association between the rEEG asymmetry in the α band (asymm, Left Panel A) and early skin-breaking procedure (SBPs) and the association between HF oscillations of the tachogram P(HF) and SBPs (Right Panel B). Data are reported for the patients with gestational age below 29 weeks during QUIET sleep (QS). The panels show how the asymm and PHF maturational trend differs in case of HIGH SBPs (magenta curve, SBP $\geq$ 50) compared to LOW SBPs (green curve, SBP $<$ 50). SBPs seem to increase the level of asymm throughout the development. Furthermore, the HF oscillations have consistently higher magnitude with HIGH SBPs compared to the LOW SBPs group. 212

9.5 LF Oscillations in QUIET Sleep: vulnerable patients. The panel a) reports how with long-term HRV ($P(LF)$) trends with development differ in case of high or low amount of skin breaking procedures. The magenta curve shows the trend with SBPs above 50 or HIGH SBPs, while the green curve shows the maturation curve for SBPs below 50 or LOW SBPs. Data are reported for the patients with gestational age below 29 weeks ($N = 31$) during QUIET sleep (QS). Steeper increase in LF power are observed in HIGH SBPs group compared to the LOW SBPs group, which is confirmed by the 3D plane on the right panel b). P(LF) increases in case of higher post-menstrual age (PMA) and high SBPs (in logarithmic scale), while the LF power seems to have a lower value (in blue) when both SBPs and the age are low. This aspect is also highlighted by the color scale, which goes from blue for the lowest value to yellow for the highest value. 213

10.1 A schematic overview of the four strategies implemented in this analysis. The BSID-III stands for the Bayley Developmental Scale, ASD is autism spectrum disorder and THR is the threshold applied to define an abnormal Bayley outcome. After EEG preprocessing, a set of power, entropy, fractality, range EEG and connectivity features are derived to quantify the level of EEG dysmaturity. Those features are fed in different supervised discrimination strategies: 1) binary classification for ASD diagnosis with linear discriminant analysis (LDA), 2) binary classification with LDA for development abnormality diagnosis (ABN), 3) multiclass classification with LDA to determine if infants have ASD or if infants have only an abnormal BSID-III, 4) Regression analysis to extrapolate the Bayley score. The four supervised strategy blocks are fed with both features and the ground truth labels ASD, BSID-III (blue thin lines, BSID-III after threshold THR or the actual continuous variable for the regression analysis) and they have the prediction of ASD or abnormal outcome (dashed grey lines). 226

10.2 The figure shows two examples of full montage of EEG from the no ASD group (NO ASD, upper panel) and the autism spectrum disorder (ASD) group (lower panel). The data from the 8 channels are reported in a window of 150 sec and the associated entropy at scale 20 $MSE(20)$ and the Hurst exponent are displayed if and only if the threshold criteria discussed in the text are met. The ASD EEG has lower entropy and higher Hurst exponent, which leads to a more discontinuous channel with slower rhythms. On the contrary, the EEG of patients without ASD has faster frequencies and higher entropy at larger scale. Therefore, the ASD EEG can be considered a dysmature EEG. The sensitivity of the plotting is reported on the top left of the chart. 227

10.3 The figure shows the entropy at scale 20 $MSE(20)$ and the Hurst exponent in the two groups (NO ASD = No ASD, ASD = Autism Spectrum Disorder). Both describe the persistence of slow-waves and discontinuity. In the case of ASD, the EEG presents a dysmature or disorganized pattern, since the regularity of the signal is higher due to spikiness, lack of smoothness and general discontinuity. The enhanced and abnormal slow-wave information in the disorganized EEG lowers the entropy at lower frequencies ($MSE(20)$). P-values have been derived with Kruskal-Wallis test. The symbols * and ** respectively represent post-hoc comparison with $p \leq 0.05$ and $p \leq 0.01$. 234

10.4 Binary classification performance in function of the threshold applied to BSID-III ($Bayley_{THR}$) to define the positive class. The two panels report the linear discriminant analysis area under the ROC curve (AUC) and Kappa score for the classification of the normal vs abnormal development group for each group of features. The abnormal group contains at least one developmental problem (ASD or abnormal Bayley). The Bayley Threshold has been used to define the abnormal Bayley outcome for each of the three investigated scores (cognitive, motor and language outcome). The reported groups of features are power, entropy (MSE), range-EEG or NEURAL features (rEEG), connectivity (Conn), fractality (MFA) and all the features combined (All). Each AUC and Kappa score chart is reported in different colors and different symbols for each feature group. 235

10.5 The figure shows the entropy at scale 20 $MSE(20)$ and the Hurst exponent in the three groups (Normal = Normal Development, ABN-Bayley = only abnormal Bayley outcome , ASD = ASD patients). Both describe the presence of slow-waves persistence and discontinuity. The combination of ASD and abnormal development shows a more severe increase of regularity compared to other groups (especially, in the occipital area) and a more severe decrease of entropy compared to other groups (especially, in the frontal area).P-values have been derived with Kruskal-Wallis test. The symbols * and ** respectively represent post-hoc comparison with $p \leq 0.05$ and $p \leq 0.01$. 236

10.6 The figure shows the entropy at scale 20 $MSE(20)$ and the Hurst exponent in function of the Cognitive BSID-III score. The two panels reported the cluster of data points (circles with a color for each patient), the expected value of the considered feature for each Bayley score (pink thick line) and the 95% confidence intervals (pink dashed lines). The left panel clearly shows that the higher the Bayley score the higher the entropy at scale 20, while the right panel shows that the higher the cognitive score the lower the Hurst exponent. The figure confirms that both entropy and the regularity are features of brain dysmaturity, and are related to developmental outcome in ASD subjects. In addition, the increase of $MSE(20)$ underlines how the persistence of slow-waves is related to a worse developmental outcome. . . 237

10.7 The measure of network redundancy is shown for the three different study groups: normal, abnormal Bayley outcome (Bayley < 80) and ASD with an abnormal Bayley. As shown by the second and third panel, the number of superfluous connections (n_{sup}) is higher in case of development abnormalities. 238

List of Tables

3.1 An overview of all graph metrics that have been used in the study. 64

4.1 An overview of all graph metrics that have been used in the multivariate analysis of this chapter. 93
4.2 The main integration and spectral features in three discrete time points. The table shows the indices for both sleep states (QS = quiet sleep, NQS = non-quiet sleep) and they were computed on the Transfer Entropy connectivity graph. The results are reported as median(IQR), where IQR stands for *InterQuartile Range*. The symbol ρ stands for the Pearson correlation coefficient, while # represents a significant correlation with $p \leq 0.01$. The values 10^{-3} or 10^{-2} mean the reported results are multiplied by a factor 10^{-3} or 10^{-2} 100
4.3 The main integration and spectral features in three discrete time points. The table shows the indices for both sleep states (QS = quiet sleep, NQS = non-quiet sleep) and they were computed on the Granger Causality connectivity graph. The results are reported as median(IQR), where IQR stands for *InterQuartile Range*. The symbol ρ stands for the Pearson correlation coefficient, while # represents a significant correlation with $p \leq 0.01$. The values 10^{-3} or 10^{-2} mean the reported results are multiplied by a factor 10^{-3} or 10^{-2}. 101

4.4 The table displays the path length for the different FC methods during QS in the different age groups: PMA ≤ 31 $Weeks$, PMA $\in (31-37)$ $Weeks$, PMA ≥ 37 $Weeks$. Results are reported as median(IQR), where IQR stands for *interquartile range*. In the last column, the Pearson correlations of the path length with PMA is reported and the associated p_{value} is reported. ** stands for $p \leq 0.01$, * stands for $p \leq 0.05$, while $^{n.s}$ means $p = not\ significant$. The values $*10^3$ means that reported results are divided by $*10^3$. 105

4.5 The table displays the path length for the different FC methods during NQS in the different age groups: PMA ≤ 31 $Weeks$, PMA $\in (31-37)$ $Weeks$, PMA ≥ 37 $Weeks$. Results are reported as median(IQR), where IQR stands for *interquartile range*. In the last column, the Pearson correlations of the path length with PMA is reported and the associated p_{value} is reported. ** stands for $p \leq 0.01$, * stands for $p \leq 0.05$, while $^{n.s}$ means $p = not\ significant$. The values $*10^3$ means that reported results are divided by $*10^3$. 106

4.6 Multivariate regression model performances. The table shows the error on the test set (Error), the R^2 and the F-statistics ($\mathcal{F}$-stat) and the p-value obtained with the different connectivity methods in the different sleep states. The results are reported as median(IQR), where IQR stands for *InterQuartile range* over the 100 random splits of the dataset. The labels reported are TE = transfer entropy, GC = Granger causality, QS = quiet sleep, NQS = non-quiet sleep. 107

4.7 The performance of the different linear mixed-effect models for the investigated FC methods is displayed. The table reports the adjusted explained variance R^2_{adj} on the training set and the MAE on the test set as median(IQR), where IQR stands for *interquartile range*. The results are reported for both sleep states, QS and NQS. The last row reports the result with selection of the best features via LASSO. 108

5.1 The AUCs for the different LS-SVM classifiers. The rows represent the different input feature sets, while the columns represent the different age groups. 120

5.2 The performance of the different linear mixed-effect models for each multifractal features set is reported for both NQS (upper half) and QS (bottom half). The table reports the adjusted explained variance R^2_{ord} on the training set, the coefficient R^2_{test} on the test set and the MAE on the test set as median(IQR), where IQR stands for *interquartile range*. The first four models are multivariate regression model with the same feature for each channel. The last row reports the result of the model based with selected features via LASSO, reported in Table 5.3. 121

5.3 The selected features of the regression models in the last rows of the two blocks of Table 5.2. The features reported are : the Hurst exponent, the parameters c_2 and c_3 for the width and the asymmetry of $D(h)$ and the difference between minimal and maximal Hurst exponents (H_{exp}, c_2, c_3 and ΔH_{Fp_1}). 121

6.1 Demographics of the 25 patients: average duration of the tachogram in minutes (Duration_{Rec}), average duration of the annotated bradycardias in s (Duration_{WB}),average number of the annotated bradycardias (Number_{WB}), average RR amplitude during the bradycardia in ms (RR_{WB}), postmenstrual age in weeks (PMA) and gestational age in weeks (GA). The last three rows represent the number of recording for each age subgroup (below 32 PMA weeks, between 32 and 36 weeks and above 36 weeks). 131

6.2 Overview of all computed features. Besides the first and the second order moments, the table reports all the relative and absolute power features in the frequency domain, namely the power in VLF, LF and HF bands, the ratio between the VLF and LF bands and between LF and HF bands and the normalized LF band power with respect to VLF band and HF band. The spectral features are reported for each the PSD estimation approaches that were used in this investigation. The last section reports the computed multifractal features, i.e. Hurst exponent (H_{exp}) and the (C_2) for the investigated scale ranges: $[j_1, j_2] = [3, 12]$ and $[j_1, j_2] = [5, 12]$. 137

6.3 The main temporal, spectral and fractal features are reported for the three windowing schemes (PB, BB and WB). The results are reported as median(IQR). IQR stands for *interquartile range*. The fractal indices are reported for $f_s = 8\ Hz$. The symbol ρ stands for the Pearson correlation coefficient. The symbol ** represents a significant correlation with $p \leq 0.01$ and $^{n.s.}$ is used to indicate a non-significant correlation. 139

6.4 The fractal features are reported in three different age categories and for the investigated sampling frequencies $f_s = [6, 8, 12]$ Hz. The results are reported as median(IQR) for the between-bradycardia and bradycardia periods. IQR stands for *interquartile range*. The symbol ρ stands for the Pearson correlation coefficient. The symbol ** represents a significant correlation with $p \leq 0.01$, and * is used for a significant correlation with $p \leq 0.05$. $^{n.s.}$ is used to indicate a non-significant correlation. 142

6.5 Linear mixed-effect model performances. For each feature set and sampling frequency f_s, the table shows the median(IQR) over the 20 iterations for R^2_{train} on the train set as well as R^2_{test} and the MAE in weeks on test set. IQR here stands for the difference between 25% and 75% quantiles. 144

6.6 LASSO selected features for the linear mixed-effect model. For each feature set, the features that have been selected more than 40% of times have been reported. σ_{RR} and μ_{RR} stand for the standard deviation and the mean of HRV. $log_{10}(LF)$, $log_{10}(VLF)$, $\frac{VLF}{LF}$ stand for the absolute power in the LF and VLF bands and the ratio between the two. Results are reported for the Wigner-Ville distribution (SPWD) or the wavelet transform. $H_{exp,[j_2,j_2]}$ and $c_{2,[j_2,j_2]}$ stand for the Hurst exponent and the parameter c_2 in the range $[j_1, j_2]$. 145

7.1 Summary of patient data demographics at different time points: GA (gestational age), birth weight (in g), PMA (postmenstrual age) at EEG and ECG recording, LPS (Leuven Pain Score). Data is reported as $q_{50}[q_{25}$-$q_{75}]$, where q_{50} is the median and q_{25}-q_{75} is the IQR. 155

7.2 Overview of the univariate features derived from the physiological signal in the study. For each signal (HRV, SpO_2, EEG), the temporal, spectral and nonlinear attributes are reported. The total count for HRV is 26: 2 temporal features, 8 spectral features, 4 nonlinear features from the Poincaré plot (PP) and 12 PRSA features. The total count for SpO_2 is 18: 2 temporal features, 4 nonlinear features and 12 PRSA features. The total count for EEG is 64: 2 temporal features, 2 spectral features and 4 fractal features repeated for each channel.RR and $P(\delta)$ respectively represent the tachogram or HRV and the EEG power in the δ band. μ and σ stand for mean and standard deviation. LF and HF represent the high and low-frequency bands of HRV. H_{exp} and c_2 are the main Hurst exponent and the width of the singularity spectrum derived with the multifractality framework. C_{xy} and SD_{12} are the PP features. $Slope_{OV}(T)$ represents the overall PRSA slope from the start of its window, while $slope_{AP}(T)$ is the slope around each anchor point. Six different window lengths T were selected to define each anchor point: $[1, 5, 10, 20, 50, 100]$ s. 162

7.3 Overview of the multivariate features derived from the different monitoring groups and the possible interaction combinations among the different modalities (EEG-SpO_2, EEG-RR_i, EEG-EEG, EEG-SpO_2-RR_i). For 5_{days}, the interaction was derived specifically for the VLF band, while the interaction for other two groups was assessed in LF band. The set of attributes for each monitoring group is 84: 21 features for EEG-SpO_2, 21 features for EEG-RR_i, 19 for EEG-EEG, 23 features for EEG-SpO_2-RR_i.The clustering coefficient ($Clust_c$) and the eccentricity (Ecc) are node-dependent features, which explain the variation in numbers for each interaction group. The actual count rises to 168 since both mean and standard deviation are considered. n_{sup} is the number of superfluous connections.The label $Efficiency$ represents the global efficiency of the network. VLF and LF represent the very-low and low-frequency bands of HRV. 163

8.1 Summary of patient data set at different time points: GA (gestational age), birth weight (in g), PMA (postmenstrual age) at EEG and ECG recording. Data are median [interquartile range]. 175

9.1 Overview of all different SBP categories monitored in the study. The first column is related to all puncture procedures, the second column relate to ventilation procedures and the last one is related to other painful procedures. 203

9.2 Summary of patient data set at different time points: SBP (Skin Breaking Procedure at 5 days), GA (gestational age), birth weight (in g), PMA (postmenstrual age) at different recording dates (5days, 34w, PSG) and CRIB score. Data are split in HIGH SBPs (SBP ≥ 50) and LOW SBPs (SBP < 50). Data are median [IQR]. 205
9.3 Summary of most vulnerable patient data with GA ≤ 29 *weeks* : SBP (Skin Breaking Procedure at 5 days), GA (gestational age), birth weight (in g), PMA (postmenstrual age) at different recording dates (5days, 34w, PSG) and CRIB score. Data are split in HIGH SBPs (SBP ≥ 50) and LOW SBPs (SBP < 50). Data are median [IQR]. 205
9.4 Results of the linear mixed effect model for the non-quiet sleep (nQS) state in the ENTIRE dataset. Only the features that have a significant association with early skin-breaking procedure are reported. The values B represent the fixed-effect coefficients for the SBP, GA and PMA variables with the associated significance level: the symbol ** means $p_{value} < 0.01$ and the symbol * means $p_{value} < 0.05$. For each feature, the mean-absolute error (MAE), the coefficient of determination and the significance of the log-likelihood ratio (LLR) test are reported for each regression model. CI stands for complexity index and P(HF) stands for HF power in the tachogram. 206
9.5 Results of the linear mixed effect model for the QUIET sleep (QS) state in the ENTIRE dataset. Only the features that have a significant association with early skin-breaking procedure are reported. The values B represent the fixed-effect coefficients for the SBP, GA and PMA variables with the associated significance level: the symbol ** means $p_{value} < 0.01$ and the symbol * means $p_{value} < 0.05$. For each feature, the mean-absolute error (MAE), the coefficient of determination and the significance of the log-likelihood ratio (LLR) test are reported for each regression model. CI stands for complexity index, P(HF) stands for HF power in the tachogram and asymmetry is the asymmetry of rEEG in the different frequency bands. 207

9.6 Results of the linear mixed effect model for the non-quiet sleep (QS) state in the patients with a GA below 29 weeks. Only the features that have a significant association with early skin-breaking procedure are reported. The values B represent the fixed-effect coefficients for the SBP, GA and PMA variables with the associated significance level: the symbol ** means $p_{value} < 0.01$ and the symbol * means $p_{value} < 0.05$. For each feature, the mean-absolute error (MAE), the coefficient of determination and the significance of the log-likelihood ratio (LLR) test are reported for each regression model. VLF/(LF+VLF) stands for the relative power in the VLF band of HRV. 207

9.7 Results of the linear mixed effect model for the quiet sleep (QS) state in the patients with a GA below 29 weeks. Only the features that have a significant association with early skin-breaking procedure are reported. The values B represent the fixed-effect coefficients for the SBP, GA and PMA variables with the associated significance level: the symbol ** means $p_{value} < 0.01$ and the symbol * means $p_{value} < 0.05$. For each feature, the mean-absolute error (MAE), the coefficient of determination and the significance of the log-likelihood ratio (LLR) test are reported for each regression model. P(HF) and P(LF) stand for HF and LF power in the tachogram, while asymmetry is the asymmetry of rEEG in the different frequency bands. 208

10.1 Summary of patient data set: GA (gestational age or age at birth), PMA (postmenstrual age or chronological at time of the recording) and development quotients at 2 years of age based on Bayley Scales of Infant and Toddler Development III (BSID-III): cognitive, language and motor scores. The first column reports the demographics of the entire pool of EEG recordings, while the data are split for ASD (Autism Spectrum Disorder) and NO ASD outcomes at 2 years of age in other two columns. Statistics are reported as median [IQR]. The last row represents the number of tuberous sclerosis complex mutations caused by the gene TSC2 in the two groups. 221

10.2 Overview of the univariate and multivariate features derived from the physiological signal in the EEG analysis divided in 5 categories: power, multiscale entropy, multifractality, connectivity features derived by the mean squared coherence (MSC) and connectivity features derived by the imaginary coherence. For the univariate feature groups (first three rows), each attribute was derived for each channel (as indicated by the third column). For the spectral features and the connectivity features, each attribute was derived for the main EEG frequency bands (as indicated by the second column). The last column represents the number of features extracted for each category. CI stands for complexity index, $MSE(3)$ and $MSE(20)$ is $SampEn$ at scale 3 and at scale 20 and ΔH stands for the difference between the minimal and maximal Hurst exponents. N_{sup} is the number of superfluous connections, $Clust_{Coeff}$ is the average clustering coefficient, $Path_{length}$ is the path length and $\frac{PL}{GE}$ is the average between long-range and short-range connections. . 222

10.3 Overview of the NEURAL Features obtained with the toolbox described in [188], where the reader can find more details. Spectral and rEEG features were computed for the main EEG frequency bands (as reported by the second column). max IBI – maximum interburst interval. The last column indicates the number of features extracted for each category. 224

10.4 An overview of different classification and regression strategies used to detect development disorders abnormalities. In the case of simple binary classification, the positive class is represented by the ASD diagnosis (ASD vs NO ASD) and the performance metrics are the area under the curve, the Cohen's kappa and the misclassification error. In the second binary strategy, the positive class is represented by at least one development abnormality (ASD or abnormal Bayley Score) which is represented by one of the Bayley Scores under a certain threshold (ABN vs Normal). In case of multiclass classification, the first positive class is represented by patients with only an abnormal Bayley score, while the second positive class is the ASD patients (Normal vs ABN-Bayley vs ASD). Therefore, one obtains two AUCs: AUC_1 for all vs patients with only an abnormal Bayley score and AUC_2 for all vs patients with ASD. The last strategy is the regression between the EEG features and the BSID-III at two years. The common metrics are the coefficient of determination or explained variance (R^2) and the mean absolute error (MAE). 230

10.5 Results for the binary classification models for ASD diagnosis (ASD vs NO ASD). The model was trained with linear discriminant analysis reported in terms of misclassification error (percentage of misdiagnosis, E(%)), area under the receiver operating characteristic (ROC) curve (AUC) and the agreement between machine and clinical labels via Cohen's Kappa. 231
10.6 Results for the multiclass classification to discriminate patient with normal development, ASD and patients with only abnormal development (Normal vs ASD vs ABN Bayley). The model was trained with linear discriminant analysis. The performance is reported in terms of misclassification error (percentage of misdiagnosis, E (%)) and the agreement between machine and clinical labels via Cohen's Kappa. Since the model was trained according to the OneVsAll scheme, the AUCs are reported as the AUC of the class with only development abnormalities vs all (AUC_1) and the AUC of the combination of ASD vs all (AUC_2). The results are reported for entropy (MSE), fractality (MFA) and the combination of all features. The performance is reported for the Bayley score threshold range [73,75,80] to define the development delay. 232
10.7 Selected features for the binary classification of NO ASD vs ASD and the multiclass scenario when all features are combined together. The results for the multiclass design are reported for the Bayley score threshold range [73,75,80] to define the development delay. CI stands for Complexity Index, while n_{sup} is the number of superfluous connections of the Scalp-EEG network as explained in the main text. 233
10.8 Regression model to predict the BSID-III score with the most correlated features and the presence of TSC mutation. The features reported here represent the best model to explain the score in terms of mean absolute error and coefficient determination (R2). The first column reports the type of Bayley score, the second column the involved features, while the last ones report the regression performance (R2, MAE and the associated p-value). 233

A.1 The main temporal, spectral and fractal features are reported for the post-bradycardia periods. The results are reported as median(IQR). IQR stands for *interquartile range*. The fractal indices are reported for $f_s = 8\ Hz$. The symbol ρ stands for the Pearson correlation coefficient. The symbol ** represents a significant correlation with $p \leq 0.01$, and * is used for a significant correlation with $p \leq 0.05$. $^{n.s.}$ is used to indicate a non-significant correlation. 254

A.2 The main temporal, spectral and fractal features are reported for the between-bradycardias (BB) periods. The results are reported as median(IQR). IQR stands for *interquartile range*. The fractal indices are reported for $f_s = 8\ Hz$. The symbol ρ stands for the Pearson correlation coefficient. The symbol ** represents a significant correlation with $p \leq 0.01$, and * is used for a significant correlation with $p \leq 0.05$. $^{n.s.}$ is used to indicate a non-significant correlation. 255

A.3 The main temporal, spectral and fractal features are reported for the within-bradycardia (WB) periods. The results are reported as median(IQR). IQR stands for *interquartile range*. The temporal and spectral indices are reported for $f_s = 6\ Hz$ and the fractal index is reported for $f_s = 8\ Hz$. The symbol ρ stands for the Pearson correlation coefficient. The symbol ** represents a significant correlation with $p \leq 0.01$, and * is used for a significant correlation with $p \leq 0.05$. $^{n.s.}$ is used to indicate a non-significant correlation. 256

Part I

Introduction

Chapter 1

Introduction

1.1 Problem Statement

The worldwide premature populations account for 10% of all births [28]. Although the survival rates of those patients have steadily increased in the last 10 years, premature infants are prone to a variety of different neurological sequelae as well as lower cognitive outcomes compared to their matched age peers [32],[150],[164]. Therefore, prematurity remains a primary concern for healthcare authorities for both the associated health costs and the necessity to prevent and address the social and educational costs of a lower neurodevelopmental outcome population.

Brain insults, such as ventricular hemorrhage and leukomalacia, can be easily monitored with imaging tools, but chronic brain abnormalities and other subtle factors can be invisible to structural analysis [262]. In addition, tools such as MRI or ultrasound might be either expensive or burdensome for the patient. On the contrary, monitoring vital signs, such as heart-rate variability, or the EEG might not simply be cheaper alternatives, but they can probe and investigate the autonomic and central nervous system state under chronic abnormalities [123],[131], [184],[262]. Functional monitoring and the assessment of maturation of the brain and the autonomic nervous system can assist in the early detection of perinatal factors which can affect the development of the infant. Among these factors, pain and early-life exposure to stress might impact brain volumes and subcortical areas such that premature infants can undergo lower cognitive and behavioral outcomes [101],[215].

However, the interpretation of electrophysiological patterns, especially in the

EEG, can be extremely complex for clinicians and the visual detection of any pattern risks to be time-consuming. Therefore, the automatic detection of exposure to stress, pain or any early-life subtle factor might be helpful for the pain and stress management and preserve the future development of the infant. In general, clinicians look for automatic methods to track maturation in the NICU and possibly relate pain or early-life experience to the future neurodevelopment outcome in order to tune therapies and preventive strategies.

There is a variety of methods to automatically detect hemorrhage or leukomalacia, as well as seizures [197]. Furthermore, different algorithms have been developed to detect quiet sleep and sleep-wake cyclicity [9],[67]. In parallel, different algorithms to assess pain using physiological signals have also been developed. Under different types of noxious stimuli and intensities, several authors showed that pain can be detected from the EEG, fMRI, HRV or EMG [41],[102],[170]. However, these studies employed pain elicitation, focused on adults and they did not look into the effects of cumulated stress. As far as the premature infants' physiology is concerned, the cortical activity of the brain was investigated under heel lance condition [236], but there is no automatic method to retrieve pain or background stress from premature infants. In order to tune both analgesia strategies and preventive care for the development, clinicians need physiological biomarkers linked to stress and relate them to development. Pioneer studies examined the maturation of infants based on the quantitative analysis of EEG [190],[192]. However, they mainly concentrated either on good outcome population or severe pathologies and they focused on the decrease of EEG discontinuity. Fewer studies investigated different modalities or the interaction among those to assess the development of the infants. Furthermore, up to our knowledge, studies about stress detection by infant's electrophysiology and about infant's development under stress or pain load were not published yet.

The objective of this thesis is then threefold. The first aim is to investigate whether biomedical signal processing automatically track maturation of premature patients in the NICU. The second aim is to investigate methods to assess stress and procedural pain automatically using the physiological data and its effect on patient's maturation. The third aim is to investigate the relationship between abnormalities of EEG and the neurocognitive outcome of infants, such as autism or intellectual disability.

1.2 The Resilience study and other collaborations

The Resilience study is a multidisciplinary project that seeks a greater understanding of *perinatal stress* and its effects on infants during their stay at the *neonatal intensive care unit*. The concept of *perinatal stress* and the potential impact on development are a multifactorial concepts, that include physiological, psychological and interrelation aspects. The complexity of *perinatal stress* and the challenges to record and analyze data of preterm infant required a collaboration among the following KU Leuven research groups:

- Prof. Gunnar Naulaers, Dr. Anneleen Dereymaeker, Prof. Christine Vanhole. Department of Development and Regeneration, Woman and child. University Hospitals Leuven, Neonatology.
- Prof. Sabine Van Huffel, Mario Lavanga, Department of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, KU Leuven.
- Prof. Johan Verhaeghe, Department of Development and Regeneration, Woman and Child. University Hospitals Leuven, Obstetretics and Gynaecology.
- Prof. Katrien Jansen, Department of Development and Regeneration, Locomotoric and Neurological Disorders. University Hospitals Leuven, Child Neurology
- Prof. Els Ortibus, Department of Development and Regeneration, Locomotoric and Neurological Disorders. University Hospitals Leuven, Child Neurology.
- Prof. Stephan Claes, Department of Neurosciences, KU Leuven.
- Prof. Bart Boets, Department of Neurosciences, KU Leuven.
- Prof. Guy Bosmans, Department of Psychology and Educational Sciences, KU Leuven.

The project also relies on the expertise and the valuable contribution of Prof. Alexander Caicedo. Originally from the STADIUS group, he is currently professor at School of Engineering, Science and Technology, Universidad del Rosario, Bogotá, Colombia. The project had the following objectives: to collect physiological, psychological and endocrinological data related to perinatal stress from a cohort of preterm patients and their parents; to quantify the level of stress by automatic scoring derived from the electrophysiological data; to assess

its impact on the infant's development by a variety of assessments after the infant's discharge.

The study population consisted of 136 preterm infants, which were included in the study project between July 2016 and 2018 in the Neonatal Intensive Care Unit (NICU) of the University Hospitals Leuven, Belgium. Parents of preterm infants born before 34 weeks gestational age (GA) and/or with a birth weight less than 1500 g were informed within the first three days after birth. Exclusion criteria were parents age < 18 years, absence or limited knowledge of Dutch or English, medical (somatic or psychiatric) condition in the parent(s) that impeded participation, and the presence of a major congenital malformation or central nervous system pathology (grade 3 or grade 4 intraventricular hemorrhage or periventricular leukomalacia) at the time of consent. The research protocol was examined and approved by the Ethical Committee of University Hospitals Leuven, Belgium. The study was performed in accordance with the Guidelines for Good Clinical Practice (ICH/GCP) and the latest version of the Declaration of Helsinki. It was registered at Clinical Trials.gov (NCT02623400).

The current thesis focuses on the STADIUS's contribution to develop a model which automatically detects perinatal stress and quantifies the impact on development. The research of Mario Lavanga was generously supported by the Strategic Basic Research grant from Fonds voor Wetenschappelijk Onderzoek (FWO), Vlaanderen. The automatic analysis to detect stress and assess development was firstly developed on a dataset of 30 healthy preterm neonates, who were recruited for a larger study to assess brain development between 2012 and 2014 at the same neonatal intensive care unit (NICU), in the University Hospitals Leuven, Belgium, with informed parental consent. A total of 26 neonates presented normal neurodevelopment outcome at 2 years from the birth, while 4 subjects were declared normal at 9 months (Bayley Scales of Infant Development - II, mental and motor score > 85). This recruitment was conducted and coordinated by Dr. Anneleen Dereymaeker and Prof. Gunnar Naulaers. Furthermore, the potential applications of the automatic scoring of electrophysiological data in relationship with developmental outcome was also tested on a cohort of 61 patients of the EPISTOP project, which is a multicentre and prospective study (NCT02098759). The EPISTOP consortium aims to assess biomarkers of epileptogenesis in tuberous sclerosis complex and to analyze whether preventive treatment before seizure onset could improve epilepsy and neurodevelopmental outcome in TSC children at two years of age. The patients were enrolled from November 2013 to August 2016 at 10 different sites, after the approval of local ethical committees. Informed consent was obtained from caregivers in accordance with the Declaration of Helsinki.

1.3 Outline of the thesis

This thesis is divided into five parts. Part I deals with the physiological and mathematical background behind this research and the different algorithms which were employed for the extraction of physiological biomarkers to assess stress, development and their relationship with the future neurodevelopment outcome. Part II shows a variety of maturational models to estimate the age of the subject directly from the physiological data. Part III discusses how stress can be assessed from physiological measurements, either looking at specific events, behavorial states or interaction among different modalities. Part IV shows how the same features used to describe maturation models can be used to predict autism or developmental delays. Part V summarizes the results of all chapters and provide future directions for this research. The codes used in the studies of Chapters 4, 5, 6, 7 and 8 will be made publicly available on the following GitHub page `https://github.com/mlavanga`. A graphical representation of the thesis structure is reported in Figure 1.1. A more detailed explanation of the thesis is presented in the following section.

Part I: Introduction

Chapter 1 is the current chapter and is meant to introduce the problem statement and the research objectives of this thesis. The organization of the different chapters is also laid down.

Chapter 2 provides an overview of the epidemiology of prematurity and the associated neurological sequelae and the neurodevelopment outcome. It proceeds with the description of the different functional monitoring, such as EEG and HRV, which can be used for the assessment the neurophysiological state and maturation of the infants. The last part focuses on the description of the epidemiology pain and stress in the premature infants and their impact on the patients' development. The chapter concludes on the recent work to assess pain directly from the EEG reactivity on the cortex.

Chapter 3 describes the different automatic methods that can be used to extract biomarkers from EEG and HRV to automatically assess brain maturation, stress and pain exposure and neurodevelopment delays. The first part is dedicated to the univariate analysis of the physiological time series. The second part explains the multivariate analysis to assess the relationships among biomedical signal and the topology of the associated network. The last part of the

chapter delves into the different supervised learning strategies, both classification and regression models, to detect stress and describe the development of the patients and the associated development metrics.

Part II: Preterm Maturation

Chapter 4 presents an automated algorithm for the assessment of brain maturation in a preterm EEG dataset based on effective and functional connectivity. Different regression models based graph features are presented to predict the age of each recording. The capacity to properly describe different network topologies is also been tested via simulated data. Furthermore, the effect of volume conduction is also discussed in relationship to the different coupling methodologies.

Chapter 5 presents a fractal analysis of the neonatal EEG in relationship with sleep staging and brain maturation. First, a fixed-size LS-SVMs to discriminate sleep states by means of multifractality features is presented and its relationship with brain maturation is assessed. Second, a linear-mixed effects regression model based fractal attributes is used to estimate the age of the patient.

Chapter 6 presents an automated algorithm to assess the maturation of the preterm autonomic nervous system. A variety of temporal, spectral and multifractal features are extracted from the heart-rate variability (HRV) data of the same patient's cohort reported in the previous chapters. A linear-mixed effect regression based on HRV features is used to estimate the age of patient at the moment of the recording. Additionally, the impact of bradycardias on autonomic maturation is specifically taken into account.

Part III: Stress quantification

Chapter 7 presents a first classification model to discriminate stress or procedural pain in the preterm infants. This analysis focuses on the relationship between bradycardia events that a patient undergoes in their NICU stay and stress in the NICU. A supervised analysis based on EEG and HRV features during hear-rate drops determines if patients experienced pain the days before the recording. The dataset consists of 136 patients recorded in three different times during their NICU stay.

Chapter 8 presents a second classification model to discriminate stress or procedural pain in premature infants. This study focuses on a stress detection from physiological background activity. By means of EEG and HRV features, a subspace LDA model detects in different sleep states if the patient experienced pain in the day before the recording.

Chapter 9 investigates the impact of early-life pain or perinatal stress on the the development of the child in the same patient's cohort of the two previous chapters. Based on the features obtained in study reported in Chapter 8 and Chapters 4, 5 and 6, a regression model is designed to investigate if maturation trajectory of premature infants with high amount of cumulated painful procedures deviate from the trajectory of patients with a low amount.

Part IV: Developmental disabilities

Chapter 10 presents a quantitative analysis of early-life EEG for the prediction of autism spectrum disorder and neurodevelopment outcome in patients with tuberous sclerosis complex. Different linear-discriminant models are reported to show how EEG features can discriminate normal patients, patients with autism and patients with other development abnormalities. Regression models to estimate the development outcome of each patients are also discussed.

Part V: Conclusion

Chapter 11 gives a summary of the main findings reported in this manuscript and the associated implication for the clinical practice. The last part is dedicated to a discussion about the future research into both the area of stress and pain management and developmental delays detection in the NICU.

1.4 Conclusion

This chapter provided the problem statement of this thesis and its three main goals: automatic monitoring of functional maturation in infants, the stress detection in premature infants and its impact on patients' development. Moreover, the structure and the design of this thesis is presented in a chapter-by-chapter fashion.

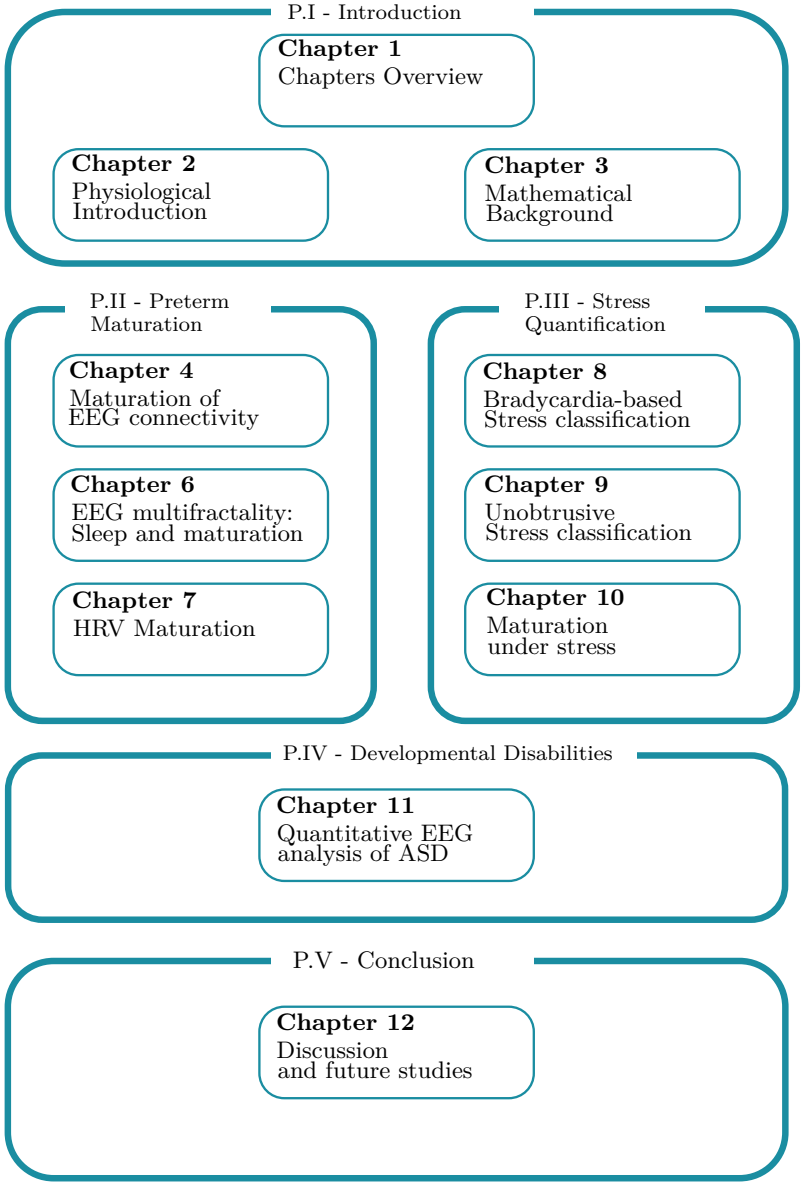


Figure 1.1: Thesis architecture chapter-by-chapter. The central parts focus on physiological or functional maturation of the healthy premature infants, the classification of stress and pain in infants and the relationship between functional trends and the development outcome of patients.

Chapter 2

Physiological Introduction: Premature birth, functional monitoring and perinatal stress

The present chapter introduces the necessary physiological background to fully understand the objectives of this thesis. First, an epidemiological overview of the premature population will be provided in terms of clinical and development outcomes. Second, an overview of the different physiological monitoring methods that are commonly used in the NICU is presented. At last, the current state of the art related to pain and stress and their relationship with infants' outcome will be discussed.

2.1 Premature birth and development outcome

Premature infants are neonates born before 37 weeks of gestation and they account for 10% of the worldwide newborn population [28]. Over the past decades, the survival rates of premature births have steadily increased, while the mortality and morbidity rates were on a sharp decline, especially in premature population below 30 weeks of gestation [32],[131],[150], [164]. This great success can be attributed to the latest improvements and technological achievements in delivery rooms and neonatal intensive care units (NICU) [32]. However, the

declined mortality of premature infants was paralleled by an increased incidence of neurological sequelae and an increased recognition of development disabilities in children at school-age [32],[150]. Premature infants are at risk of vast range of acute brain insults, such as intraventricular or periventricular hemorrhage (IVH/PVH) or periventricular leucomalacia (PVL), which consists in necrosis of white matter next to the lateral ventricles. Cerebral lesions like PVH and PVL account for 34.8% of abnormalities in infants at 30 weeks and 7% at 34 weeks, while cardiorespiratory abnormalities, such as respiratory distress syndrome and bronchopolmunar dysplasia, can impact 54.7% of infants at 30 weeks and 2.6% at 34 weeks [164]. Other perinatal postnatal complications can be maternal or nosocomial infections, inflammations or poor nutrition [131], which can also lead to PVH or PVL. Those events can occur in proximity of birth (perinatal) or later during their stay at the NICU (postnatal), and they might trigger severe neurodevelopmental deficits and cognitive dysfunctions, such as cerebral palsy, visual and hearing deficiencies, epilepsy or lower motor outcome [164],[262], among others. As shown by Marret et al. [164], the intensity of those impairments is directly proportional to the gestational age of the patients.

However, infants at 34 weeks, without any complications, can still have a lower cognitive development compared to their full-term peers at 2 age [164]. Although long-term developmental deficits remain a primary concern, with 5% to 10% of very preterm children having cerebral palsy (CP) and 40% with milder motor deficits, the incidence of cognitive deficits is still higher, with 30% to 60% of premature infants experiencing cognitive impairment and social or emotional difficulties [238]. Namely, prematurity can lead to subtle abnormalities such as attention disorders like ADHD [32]. The perinatal course might shed light on this higher risk of cognitive impairment: premature infants are not simply prone to postnatal complications and chronic lung disease, but they are also subjected to maternal separation and experience painful procedures [56],[238]. Multiple authors have discussed and investigated the reasons of those lower neurodevelopmental outcome. Severe complications, like IVH, PVH or PVL, can be responsible of different degrees of cerebral and white matter injuries [150],[164], while the respiratory distress syndrome (RDS) is the leading cause of preterm mortality [179] . It is interesting to notice that brain infarction or brain lesions can come as sequelae of multilevel and multisystem immaturities, such as immune system immaturities (in case of infections), cardiorespiratory immaturities (in case of RDS) or other congenital factors [4],[179]. In case of more subtle factors that can lead to cognitive dysfunction, premature birth can precipitate immature cell death in the brain. This higher rate of cell death can lead to volumetric losses in sensory-motor cortex and other subcortical areas, such as the hippocampus, which are ultimately responsible for cognitive and behavioral disorders [32]. Therefore, prematurity remains a primary concern for healthcare authorities. The focus is not only related to greatest health

cost that premature birth will impose in the coming decades, but there is need to prognosticate, prevent and address the social and educational costs from a population with a lower neurodevelopmental outcome in our societies.

The gold standard for the detection of cerebral lesions, PVH and PVL, and other acute brain abnormalities remain structural imaging tools, such as MRI and cranial ultrasound [184],[262]. The latter is considered a more accessible methodology, but it has poorer sensitivity in diffuse or subtle brain lesion [262]. The MRI is the finest method to obtain high-resolution images without any ionization. Magnetic Resonance images allow a precise estimation of a cerebral lesion location and extension. In addition, MRI can be used to detect white matter density and directionality thanks to the diffusion tensor imaging [169]. However, MRI is an expensive method, which is not always accessible, and it is very sensitive to movement artifacts [161]. MRI also requires patients' transportation and sedation and is specifically designed for structural imaging [169].

Chronic brain abnormalities are subtle changes in brain functioning and they might reveal structural lesions or differences only later in life [262]. As Mento et al. highlighted in a recent review, the structural imaging reveals "how the brain is", but the functional monitoring is required to tell "how the brain works" [169]. It should not also be excluded that functional monitoring methods can be used to monitor brain insults thanks for the availability and prognosis velocity [184],[262]. Therefore, there are several clinical studies that look into the functional techniques to monitor brain's and patient's homeostasis to detect both acute abnormalities or chronic stages that can lead to a lower neurodevelopment outcome [131],[150],[232],[184]. It is important to remember that studies that focus on developmental delays or cognitive and motor outcomes rely on different definitions of age. The assessment of maturation in the womb and the effects of prenatal factors are based on gestational age, which simply represents the weeks of gestation. In contrast, development inside the NICU is investigated based on the number of postnatal days after birth, which is the chronological age. A more common metric to investigate the development after birth is the combination of gestational age and chronological age, known as the postmenstrual age. If the focus is the comparison between developmental outcomes between preterm and full-term infants (born after 37 weeks), researchers rely on the corrected age, which is the number of days or weeks between the expected delivery day and the observation point. This metric allows to compare at expected similar level of maturation. A simple overview of the different age definition is reported in Figure 2.1.

The next paragraphs will introduce two of the most common functional monitoring tools in the neonatal intensive care unit. The first is the EEG, which monitors the electrical activity of the cortex. A longstanding clinical

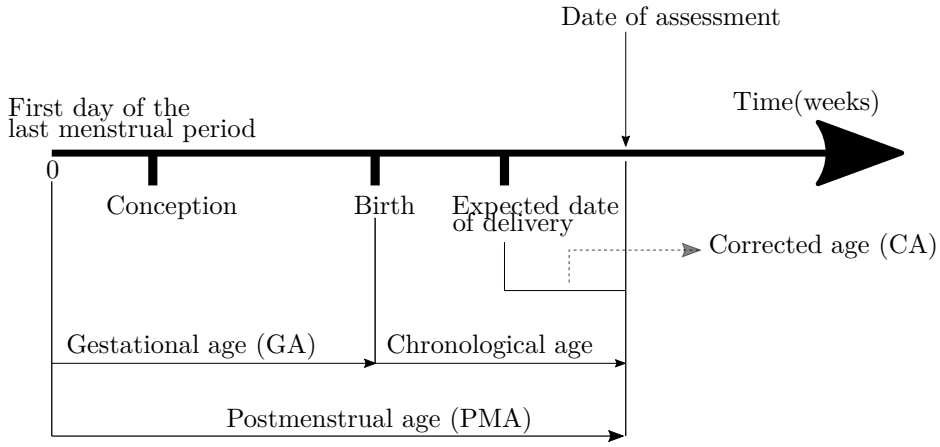


Figure 2.1: The figure show the multiple definition of age that are used in neonatology. The gestational age is the amount of time that the fetus stays in the womb before the delivery and the counting starts from the last menstrual period of the mother. The chronological age is the number of days after birth to the observation point, while the postmenstrual age is the sum of the former two. At last, the corrected age is computed as the number of days between the expected day of delivery and the observation point.

literature showed how the EEG can be used to assess both the impact of a variety of brain insults and the level of brain maturation. The second is the HRV, which monitors the influence of the autonomic nervous system on the heart. It is normally considered a probe of the status and development of the neurovegetative functions and it found specific application in detecting sepsis or acidosis. The objective of the next part of the chapter is to show how EEG and HRV can be used to monitor the development of the premature infants in the NICU in the clinical practice and their relationship with developmental outcome.

2.2 EEG: functional monitoring of the brain

The EEG provides a direct insight of the brain functioning. It consists of the recording of electrical activity along the scalp produced by the firing of neurons within the brain, reflecting correlated activity caused by the post-synaptic potentials of cortical neurons. It is currently used as diagnostic tool for different type of abnormalities, such as seizure or brain insults. Especially in the NICU, EEG can provide critical real-time information of cerebral dysfunction, especially in case of lesions or systemic diseases [196]. In addition, EEG can provide information of subjects' maturation. EEG does not only change with the development of the infant, but it has also been suggested that those changes relate to the anatomical development of the brain [33], [169].

2.2.1 EEG recordings: Conventional EEG and aEEG

The conventional set-up for EEG measurement includes Ag/AgCl electrodes, whose positioning typically follows the 10-20 international system. The 10-20 refers to distance between neighboring electrodes which equals 10% or 20% of the total distance from the nasion to the inion, which are respectively the suture that joins the nasal bone with the frontal bones and the occipital protuberance at the back of the skull. This system guarantees a uniform and consistent monitoring of different brain regions according to the different international standard. Each electrode is then labelled by its anatomical position (F for frontal, T for temporal, C for central, P for parietal and O for occipital) and a number, which indicates the left hemisphere if it is odd and the right hemisphere if it is even (Figure 2.2). Each biopotential information is a difference in voltage between two electrodes. In the case of EEG, the computational scheme to define those differences is defined as montage and it can be bipolar or monopolar. The former represents a proximity electrodes difference (e.g. Fp_1 - T_3), while the latter considered common ground, which is either an electrode on the aurical point (A_1) or an electrode at vertex of the scalp (C_z). The letter z is used to represent the middle-line of the cortex. In this thesis, all studies were performed with monopolar montage referred to C_z. Newborns have typically a reduced set of electrodes or EEG channels, namely Fp_1,Fp_2, C_3,C_4,T_3, T_4, O_1,O_2 and C_z (as also highlighted in Figure 2.2). This is dictated by the necessity to reduce skin injuries and avoid the discomfort for the patient, as well as by the limited area due to the small head size.

Due to the complex spatiotemporal information contained, the full-channel EEG is difficult to interpret with clinical eye. Medical doctors normally use the amplitude-integrated EEG or cerebral functional monitoring in clinical practice.

This methodology was introduced at the end of the 60s to quickly scan the brain function of the adults [165]. It is normally recorded in the NICU with two electrodes (usually C_3,C_4) and the signal is normally filtered in $[2-15]$ Hz band, compressed and semi-logarithmically transformed. As pointed by Pavlidis et al., the aEEG requires for very little training and it can be used not just for brain insult detection, but also sleep stage-monitoring and extraction of fairly simple functional features for the developing brain [196]. However, a full appreciation of the maturational trends requires the analysis of full-channel EEG, as explained in the next paragraphs.

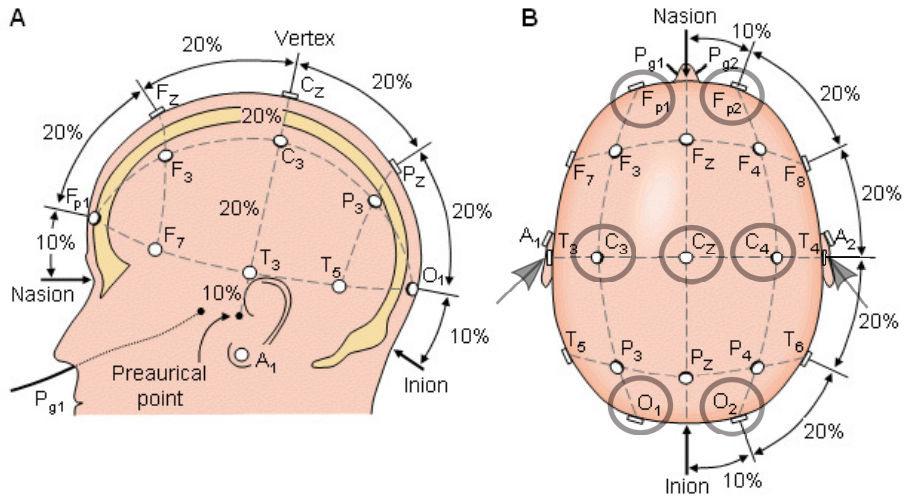


Figure 2.2: The figure reports the electrodes displacement on the human scalp according to the 10-20 international system. The grey circles and arrow show the electrodes that were used in the analyzed EEG recording of this thesis. Reprinted from [163].

2.2.2 EEG Maturational Patterns

The EEG evolves from the most immature infants at gestational age 23/24 weeks until full-term age with four major characteristics: (i) a decreasing discontinuity, (ii) the appearance of sleep-wake cyclicity, (iii) change in hemisphere synchrony and (iv) the appearance of transient waveforms [197]. The two latter trends can be detected only with conventional EEG, while the first two are easily assessed with aEEG. Those four trends are described in short in order to provide a

formal definition of dysmature EEG, i.e. an EEG which is not appropriate given the GA/PMA of the infant. A summary of those trends are reported in Figure 2.3 (adapted from [261],[8],[197]).

Discontinuity

The neonatal EEG baseline is characterized by period of activities, known as bursts or spontaneous transient activities (SATs), and period of quiescence, known as interburst intervals (IBIs), as shown by the example in Figure 2.4. There are different physiological reasons why the brain generates endogenous events in the EEG. The intermittent activity might be a prior form of sensory modulation or the GABAergic transmission might not be inhibitory in the early development [197]. The amplitude of the burst activity can be greater than 300 μV in 25 weeks and it can reduce up to 50 μV at full-term age (36-37 weeks). The burst is centered in low-delta band, which results in a slow-wave type of pattern. The peak frequency (which defines the main periodicity of the burst-pattern) is located in the band $[0.3 - 1]$ Hz at 25 weeks and it moves towards the upper-delta band $[2 - 4]$ Hz after full term age, as shown by the evolution of the dominant frequency in Figure 2.3. On average, the discontinuity is concentrated in the band $[0.5 - 2]$ Hz [8],[261]. The last normative value that clinicians look at to assess the discontinuity of EEG is the length of the IBIs, which goes from 60 seconds in extreme-early prematurity up to 10 seconds in the EEG of infants close to full-term age.

Sleep-wake cyclicity

Sleep-wake cyclicity stands for the organization in behavioral states of sleep. In neonates, experts distinguish wakefulness, active sleep (comparable to REM sleep) and quiet sleep (comparable to non-REM sleep) [8]. Behavioral states can be clinically detected in the EEG as early as 30 weeks GA. However, the states of younger infants can be detected if polysomnography is available, where the combination of EMG, EOG, vital signs and video can help doctors in the sleep assessment. Normally, the quiet sleep shows periods of discontinuous EEG tracing with absence of eye-movements or muscular activity, while active sleep is characterized by a more continuous tracing with more rapid muscular or eye activity [8],[197]. The sleep progression is such that infants start spending 90% of their time sleeping and this amount decreases to 70% in full-term infants [16]. During the development, the indeterminate sleep, which is a state that cannot be associated to AS or QS, decreases and the REM sleep becomes more structured [58], as highlighted by the sleep architecture trends in Figure 2.3. At term-age, the infants can have up to four states: wakefulness, active sleep 1, active sleep 2

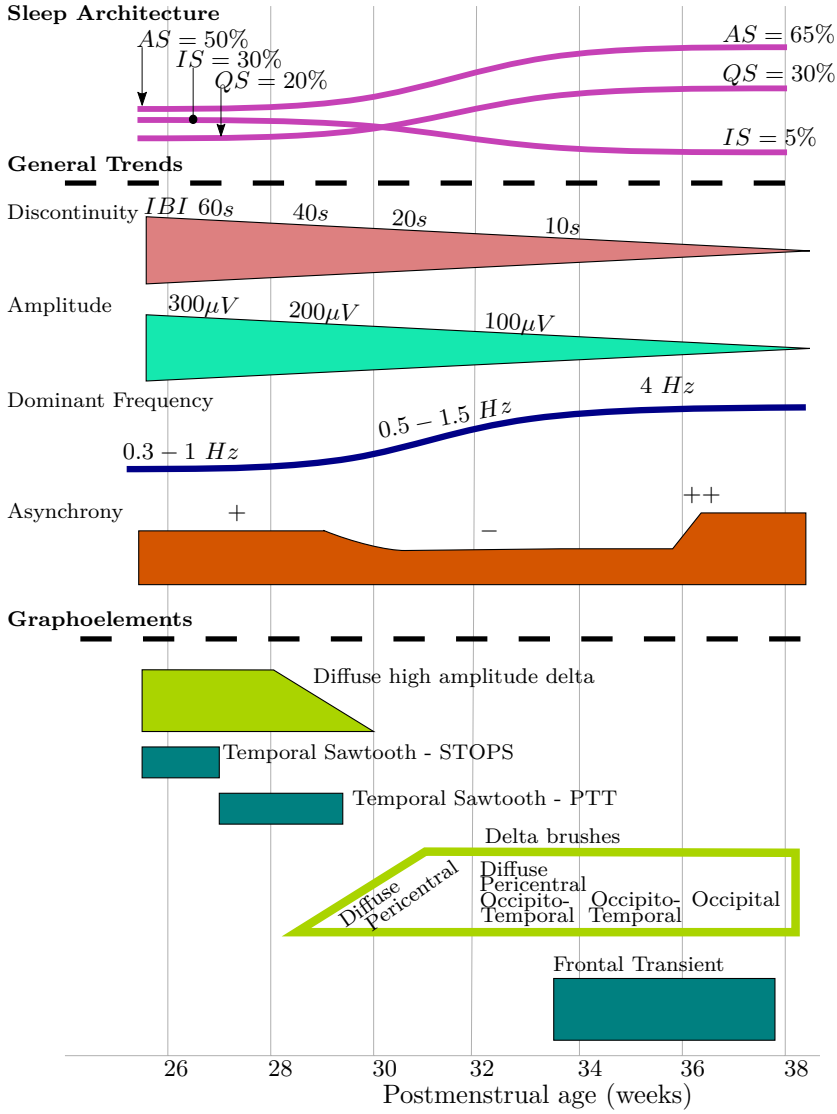


Figure 2.3: The figure reports the functional maturation trend of the premature infants in function of postmenstrual age. The charts are adapted from [261],[8],[197]. The top part reports the evolution of the sleep-wake cyclicity. The middle part reports the changes in EEG information (such as discontinuity and amplitude) and type of transients, while the last part shows the evolution of the different EEG graphoelements. The label AS stands for the active sleep trend, while QS is the quiet sleep and IS is the indeterminate sleep.

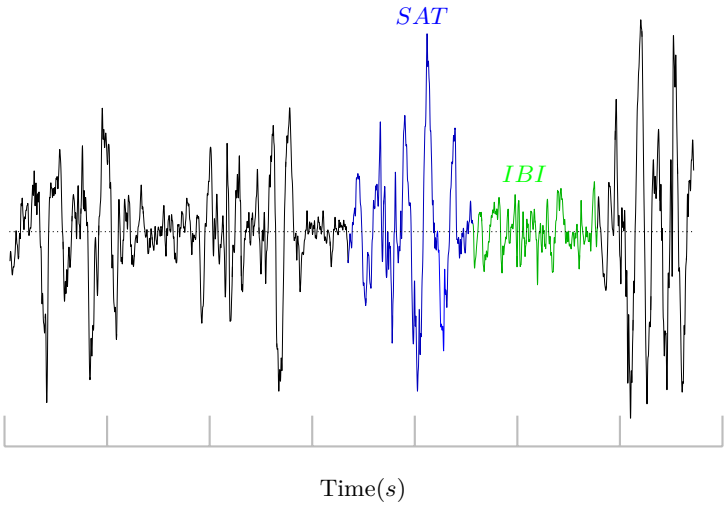


Figure 2.4: An example of discontinuous EEG, where by period of activities, known as bursts or spontaneous transient activities (SATs), and period of quiescence, known as interbursts intervals (IBIs), are respectively shown in blue and green.

and quiet sleep. The quiet sleep results in a semi-discontinuous tracing, called trace' alternant, and the active sleep is a multifrequency continuous tracing [8], [197].

Synchrony

Synchrony stands for the appearance of the same features in the two hemispheres simultaneously [197]. Synchrony of high-amplitude burst is present at GA < 30 weeks as consequence of the development of the corpus callosum and the predominance of the subcortical control [134]. Asynchrony between two hemispheres emerges at 30 weeks and it tends to disappear at term age [197]. Although asynchrony can be an important feature to assess the development of the corpus callosum, the full-channel EEG is required and an appropriate automatic assessment is necessary [197].

Transient waveforms

Diffused delta activity and delta brushes refer to high amplitude delta activity, mainly concentrated in the lower delta band $[0.3 - 1]$ Hz and mainly

present at very young GA. They dominate the burst activity in the first period of life. After 28 weeks, one of the most important traits of neonatal EEG emerges and is known as delta brush, which is a burst type of activity, centered in the delta-band $[0.5-2]\ Hz$ and superimposed with faster activity. They are present throughout the preterm development, but they peak between 32-34 weeks and they start disappearing after 38 weeks. While maturation progresses, delta activity sees a greater interference or superposition of faster frequencies (from theta to alpha/beta band), a shift from slow band $[0-1]\ Hz$ to the upper delta band $[2.5-3]\ Hz$ and decrease in amplitude (from 300 μV to 50 μV). Most importantly, the topography of the delta brushes changes with development. Although the spread of those graph-elements is pretty much diffused at early-stages, they tend to concentrate in the temporal-occipital region. This might be a fundamental step of the organization of sensory map on the cortex: delta brushes appear to be immature evoked potentials for infants younger than 35 weeks GA [236]. The location of those brushes becomes topography-specific with different type of stimuli: for example, visual stimuli concentrate delta brushes in the occipital area, while noxious stimuli elicit burst activity in the temporal area [78],[197],[255].

Temporal and occipital theta activity refer to theta band bursts and include "Sharp Theta on the occipitals of premature infants" (STOPS) and "Premature Temporal Theta" (PTT), also known as temporal sawtooth. Both those transients are centered in the $[4-7]\ Hz$ band and their amplitude increases with decreasing age. STOPS have a sinusoidal pattern, mostly located in the occipital region, and they disappear at 28 weeks GA, when PTT becomes more frequent in the temporal regions. Temporal sawtooths will fade around 34 weeks GA.

Frontal transients include anterior dysrhythmias and encoches frontales. They both can refer to a smooth transient in the frontal area. The dysrhythmias manifest as patterns of delta waves of medium amplitude (above 50 μV) in AS_1 from 36 weeks GA onwards. The encoches start as immature, smooth and asymmetrical transients from 33 weeks GA and they become a smooth, symmetric and diphasic patterns from 35 weeks GA. They are generally of high amplitude (200 μV) and they are considered as good sign of maturation.

An example of the different EEG graphoelements is given in Figure 2.5.

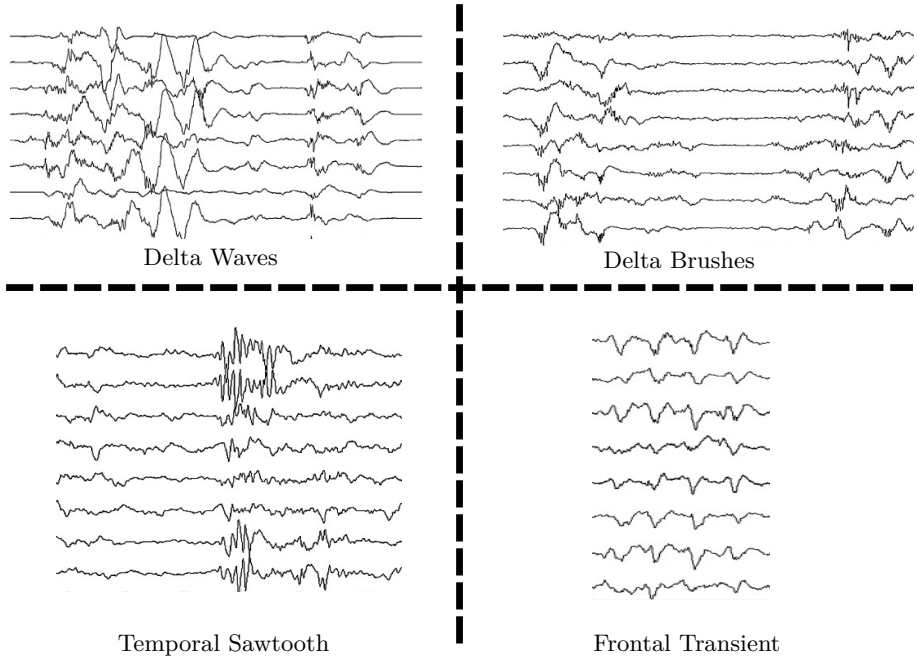


Figure 2.5: Overview of the different graphoelements, which characterizes the neonatal EEG. The top left shows delta waves, commonly present in the first days of life and a common sign of dysmaturity, while the top right shows delta brushes, which are bursts in the delta band with superimposed faster information and emerge also as a sign of somatosensory cortex modulation. The bottom left shows theta temporal sawtooths, which are normally present in the infants below 30 weeks of gestation, while the bottom right shows frontal transients, which are good sign of maturation of the frontal area.

2.2.3 Abnormal EEG

In short, a healthy preterm EEG should show a decreasing discontinuity, a decreasing and more asynchronous delta activity and different types of transients with the infant's development. Therefore, the monitoring of the EEG background continuity is an essential indicator to assess if the maturity of the infants is appropriate for a given GA. In general, various degrees of background depression and degrees of discontinuity are associated to acute brain insults and they are generally graded from I to IV. This discontinuity scale starts with prolonged IBIs and attenuated transients and proceed with different level of delta-burst attenuation up to low-voltage, or flat-line patterns. The most severe grades are characterized by solely delta waves or very slow-frequencies. Triggers of

abnormal EEG are brain insults like PVH or PVL, but acidosis, apneas and drugs can play a role in EEG suppression. Those patterns are acute abnormalities and they can be easily associated to specific insults. They can also be easily recognized with aEEG or CFM [197],[262].

However, EEG can show signs of abnormalities without findings of acute suppression. Those abnormalities are usually defined chronic, and consist mainly of dysmature and disorganized patterns, which means an inappropriate pattern for the patient's GA. In addition, dysmaturity implies the young-preterm patterns such as high-amplitude delta-waves, theta-bursts or other rhythms waves and absence of mature, slow patterns can last until term-age and beyond, generally leading to poor development outcome. In general, the persistence of discontinuity and slow-waves for a period of time, especially after 36 weeks GA, is considered a dysmature pattern. The clinical literature distinguishes the dysmature pattern from the disorganized EEG, which is a deformed activity compared to normal without any sign of acute insult [262]. Interestingly, the disorganized EEG might be associated with white-matter injuries, which are not initially visible with imaging tools, while a dysmature pattern does not always show association with structural lesion [262]. The dysmature EEG can lead to mental and cognitive impairment.

The clinical practice of EEG diagnosis reported by Watanabe et al. shows the power of full and continuous EEG monitoring in the NICU in order to timely intervene and provide right therapy [262]. High grading of EEG abnormalities (grade $>$ II) leads to more than 50% of cases to neurological sequelae, especially cerebral palsy, while chronic dysmature patterns were associated with retardation or borderline outcome in 68%, defined as a patient without clear sign of insults. Most importantly, mental retardation and cerebral palsy were respectively associated to a postnatal event in 71% of cases and to a perinatal event in 51% of cases.

2.3 HRV: functional monitoring of the autonomic nervous system

The heart-rate or tachogram represents the beat-to-beat variability of the cardiac frequency and its analysis is considered one of the most common indicators of the autonomic nervous system (ANS). The heart-rate represents the frequency of stimulation of the heart muscle and its autonomous activity is set by the sino-atrial node, which is a pacemaker group of cells. This activity is normally monitored with electrocardiogram (ECG). Although the pacemaker stimulation is autonomous, the ANS can anticipate or delay the heart depolarization, which

can ultimately change the distance among R-peaks. Specifically, the ANS can increase the heart-rate via the sympathetic innervation and decrease it via the parasympathetic innervation or vagus. The variations of the heart-rate are usually associated to those sympathetic and parasympathetic interplay. The short-term (or high-frequency, HF) variations are associated to the latter, while the long-term variations (or low-frequency, LF) are usually associated to the sympathetic stimulation. The HF is normally related to respiratory activity and carries information to the respiratory sinus arrhythmia, while the LF is normally split into a slower component (or very-low frequency, VLF) and faster component (usually regarded as LF or middle-frequency MF by certain authors). The latter is normally linked to the short-term control of blood pressure, or baroreflex sensitivity, while the VLF is considered of less-clear origin. A schematic representation how to derive the HRV from the ECG is reported in Figure 2.6.

Premature infants present an immature cardiovascular control and ANS. The first peculiarity is that the intracardiac regulation (based on Starling's law) shows a dominance of the frequency control, which means the cardiac output is more dependent of the heart-rate. The second peculiarity is the immature nervous control of the heart-rate which manifests with an extreme low vagal tone at rest, which is reflected by a very high heart-rate. The immaturity manifests also itself as a weaker blood-pressure control and chemoregulation. The baroreflex sensitivity is reduced in infants and the heart-rate reactions to hypercapnia are delayed or inhibited. However, the most common hallmark of preterm HRV reactivity to asphyxias are bradycardias, which are heart-rate drops anticipated by a period of a fast heart-rate with low-variability (known as tachycardias). Different studies of HRV data processing in both preterm infants and fetuses confirm that the weak parasympathetic tone is accompanied by a predominant LF tone [50],[58],[59],[123]. However, the infants's HRV is characterized by a strong-wave baseline, therefore the definition of the frequency band limits is essential. The preterm tachogram might have two dominant tones, LF and VLF, which might relate to different aspects of the sympathetic stimulation. This strong dominance of the long-term variations reduces the complexity or entropy of the HRV signal [231].

Similarly to the EEG, the heart-rate evolves while the brain and the autonomic nervous system mature. The most relevant changes in these features are the decrease of the cardiac frequency with infants' development and the increase of the heart-rate variability. Specifically, the vagal-tone increases with increasing GA, but the development of the LF and VLF is faster than the vagal increase [50],[58]. The increase of variability of the heart-rate seems to involve equally all the frequency bands. However, the development of the different tones is sleep-state dependent [58]: the vagal activity increases faster than any other tone

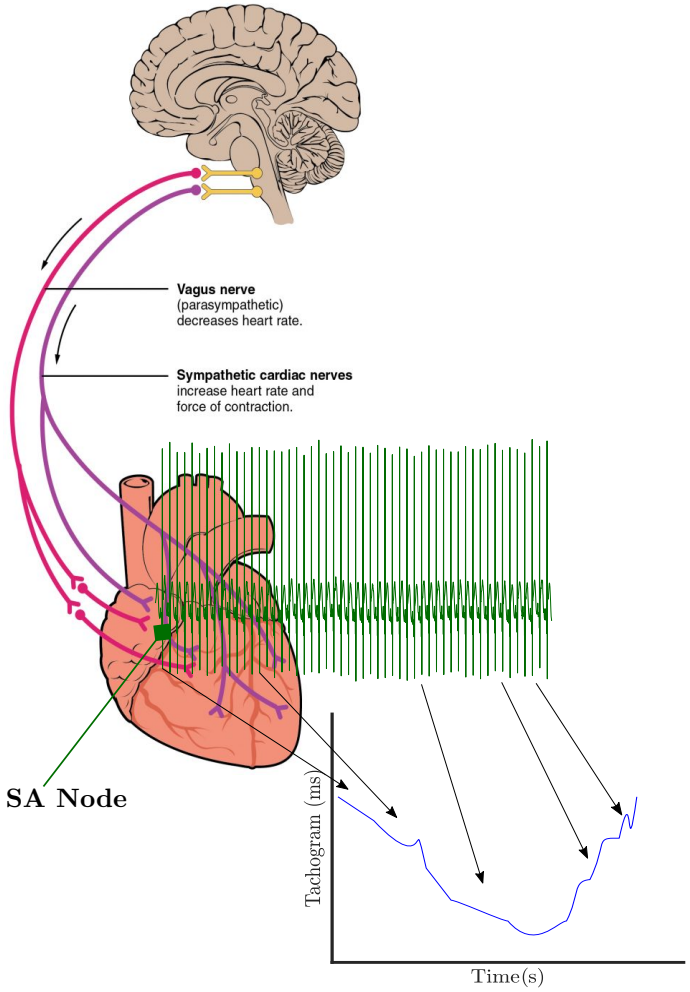


Figure 2.6: The figure shows the nervous control of the heart rate via the vagal and sympathetic innervation of the autonomic nervous system. Besides the conduction velocity, their activity acts as controller of the sino-atrial node anticipating or delaying the depolarization of the heart. The net effect is a variation of distance between the R-peaks (ventricle repolarization) or beat-to-beat change of R-peak location. The variations of R-peak-to-R-peak intervals and its inverse (known as instantaneous hear-rate) are known as heart-rate variability or tachogram. Reprinted from [186].

in the quiet sleep, while the rate of increase is higher for long-term variations in active-sleep. The ANS development also shows a decrease of frequency and severity of bradycardias, mainly driven by the reducing apnea load. The development of the cardiorespiratory control does not only show a decrease of hypoxic events, but it shows a first decrease of the respiratory frequency until term-age and then a further increase [58]. A schematic overview of the ANS development is reported together with EEG features in Figure 2.7.

The tachogram information has also been used to assess the acute insults to the brain. A low HRV has been predictive of respiratory distress syndrome, asphyxia, acidosis, sepsis and PVH [71],[138]. In the latter case, PVH can drastically reduce HRV, while infections and septic conditions can be predicted by low information entropy in the neonate HRV. More interestingly, the variability of the heart-rate seems predictive of the mental and cognitive outcome. Fetuses with slower and more variable heart rates tend to have significantly higher Mental and Psychomotor Development Indices (MDI, PDI) and had better speech development than those with faster HR and reduced HRV [123]. Similar results and associations were found between vagal tone and neonatal attentional orientation have been shown by [84], between HRV and MDI scores at the age of 1 year by [219], and between respiratory sinus arrhythmia and standardized cognitive test scores in middle childhood [76]. What is also pivotal of the heart-rate is that its variability seems primarily explained by maternal care and genetic factors. Therefore, any event that can affect the relationship between mother and child can have long-lasting impact on the autonomic control and the cognitive outcome of the infants [123].

2.4 A definition of dysmature EEG and HRV

The medical research community delved into the maturation of a variety of different traits of neonatal EEG and HRV and related them to possible developmental and clinical outcome. In order to provide a unique and reproducible framework, the biomedical literature tried to provide a definition of dysmature EEG and HRV. Considering a specific patient under monitoring, dysmaturity means an electrophysiological recording whose traits and aspects resemble the physiological data of a much younger patient. It could generally imply a possible adverse outcome. Therefore, two general definitions of dysmature EEG and HRV are provided and the concept of dysmaturity should be intended as the functional dysmaturity discussed below.

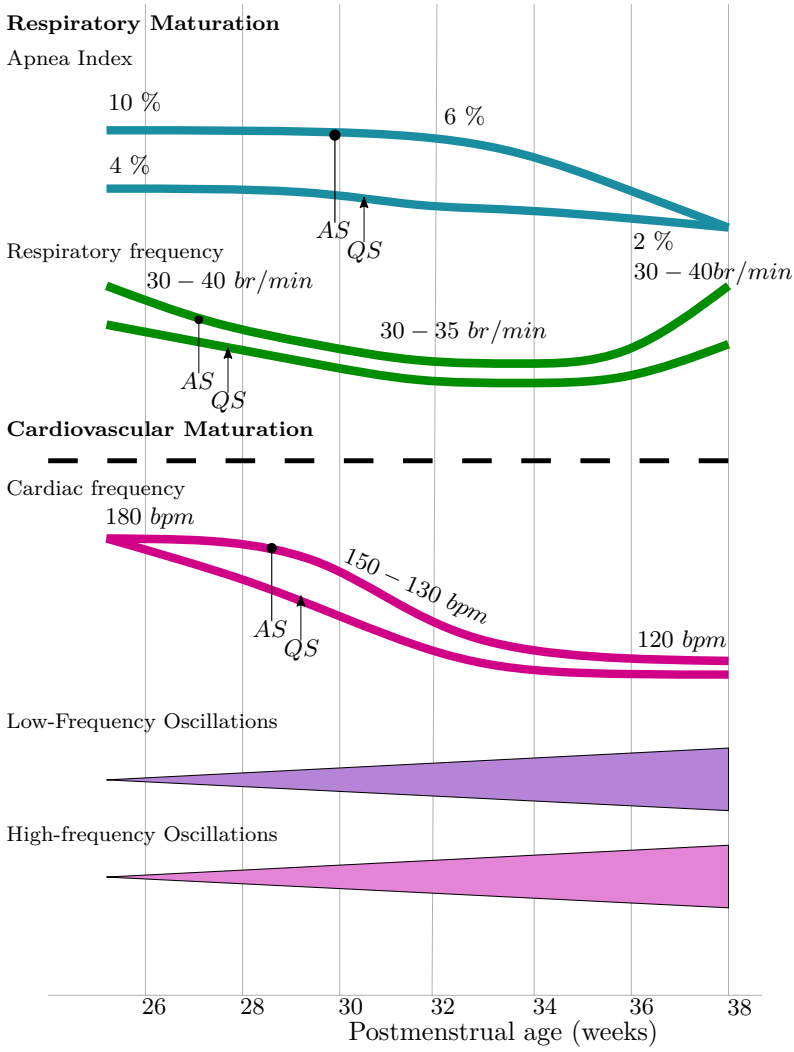


Figure 2.7: Functional trends of the respiratory and cardiovascular system. The charts are adapted by [58]. The upper part shows the evolution of the respiratory information, with a sharp decrease in apnea-frequency with aging and a u-shaped change of the respiratory rate. The lower part shows the cardiovascular trends, which are a decrease in heart-rate and increases of the low-frequency and high-frequency power. The label AS stands for the active sleep, while QS is quiet sleep.

2.4.1 The dysmature EEG

Pavlidis et al. [196] and Watanabe et al. [263] defined the main traits of a disorganized EEG in case of dismaturity or preterm infants with a poor clinical outcome as follows:

1. Discontinuity, intended as a background depression or prolonged interburst interval, which can reach a total absence of continuous pattern [150]
2. Persistence of slow-waves, intended as signal with mainly delta activity or increased presence of transients such as delta-brush
3. Asynchrony, intended as delayed occurrence or lagged asynchrony among bursts

In a normal maturation profile, all those features are expected to disappear, while their persistence might be symptom of an underlying abnormality. The engineering literature showed different methods to quantify the level of discontinuity and persistence of slow-waves. The former is usually assessed via the detections of spontanoeus transient activities (SATs). The distance between them is usually known as interburst intervals (IBIs) and different metrics were derived based on those two patterns, such as percentage of time spent with SATs, maximal and mean length of the IBIs or IBIs distribution, as shown by [194], [133], [191] and many others. In case of the persistence of the slow-waves, simple methodologies, such as the amplitude distribution of the EEG or the spectral power in the delta band [223],[130], were investigated to assess the presence of slow-wave transients in the EEG.

2.4.2 The dysmature HRV

Unlike the dysmature EEG, there is no clear-cut definition for dysmature heart-rate variability. However, multiple authors have investigated the differences between term and preterm population in terms of HRV and some similarities have been found between fetuses and preterm infants. Javorka et al. showed that the greatest consensus in the literature concerns the increase of variability of the heart-rate with maturation, together with a sharp decrease of the cardiac frequency [58],[123] . In addition, the bradycardias and apnea frequency should decrease [58]. Probably, the closest research to a definition of a dysmature HRV was the study by Hoyer et al. [111]. They found the neurovegetative maturation in fetuses will decrease the slow-wave or very-low frequency information in infants relatively to other frequency bands, while complexity of the signal

increases. Similarly, David et al. showed that fetal HRV has strong slow-wave baseline, while Clairambault et al. showed that the power of all frequency bands, especially the faster ones (low-frequency and high-frequency band), will increase over age [59], [50]. Therefore, a dysmature HRV has extreme slow-wave baseline, low variability and high regularity.

2.5 Pain, Stress and developmental outcome

Although many perinatal factors have been linked to an adverse neurodevelopmental outcome, much of the variations in premature infants' outcome has yet to be explained. During premature development, the brain might undergo subtle changes that can lead to attention disorders like ADHD. Besides the sequelae of postnatal complications, the incidence of cognitive, social and emotional impairment remains higher in the premature population. Multiple studies suggest that exposure to stressors in the NICU might impact the neurodevelopment outcome of premature infants [101],[238],[100],[258]. This exposure, known as Perinatal stress, is usually comprised of two factors: painful procedure and maternal separation. Pain-related stress has been associated to a lower cognitive and motor function at 8 and 18 months and higher internalizing behavior (linked to the depression) at 8 and 18 months [101].

Multiple theories are reported in the clinical literature about why and how stress and early procedural pain can impact the infant's development. Premature infants develop a central sensitization to pain and a lower pain threshold, as a consequence of peripheral sensitization [85]. This has been seen by the hyperinnervation of the periphery and the increase of receptive fields of the dorsal horns [218],[248]. At the central level, the subplate neurons are the first cells to be generated in the mammalian cortex and the first one to receive sensory inputs. Repeated procedural pain might lead to excitotoxicity and apoptosis of those neurons, while the premature cortex showed a larger neural activity compared to the term brain in response to noxious procedure [42], [236]. This overstimulation might reduce brain volume, white and gray matter connectivity, decrease cortical thickness and reduce the thalamic volume [32],[42],[73],[214]. Topology wise, stressors seem to target the volume of the frontal and parietal areas, as well as the functional and microstructure of temporal lobes [42],[238]. Interestingly, the altered temporal architecture might be a reflection of the reduced volume of the hippocampus, which has been observed in rat models as consequence of chronic stress and post-traumatic stress disorders. In addition, procedural pain might activate immune system modulators, which release inflammatory molecules, which already makes the glia cells vulnerable to immune signaling molecules, such as cytokines, and sensitize further the periphery to pain [73],[258].

However, perinatal stress does not only include pain-related information, but maternal separation might also play a role. Animal models using rats show that maternal separation might diminish frontal EEG activity [175], and parenting seems a protective factor to mitigate the stress impact on development and the insurgence of internalizing behavior [101].

The clinical literature clearly shows that stress can impact brain development, but it is important to remember that procedural pain or pain-related stress can depend on gestational age, first days of life, birth weight, respiratory support, infections and drug administration [101],[100],[238]. Although there are already multiple pain scales to assess behavioral and physiological states of the premature infants, the clinicians are faced with the difficult task of discriminating and appropriately managing pain in infants born prematurely. As Smith et al. pointed out, stress seems to be a pathway from severe illness to altered brain development rather than a marker of severe illness [238].

However, clinicians cannot yet fully disentangle pain, several illnesses and the effect on development. Unmanaged pain may have substantive effects on the developing brain and the stress-response system of premature neonates. In the current NICU scenario, the detection and assessment of the developmental effects of stress via structural imaging might be burdensome for the patients. Measures like MRI require further transportation and sedation [169] and, most importantly, they risk to be unhelpful. Alternatively, chronic brain abnormalities might be invisible to imaging and still lead to sequelae like mental retardation, as highlighted by [262]. The latter is especially associated to dysmature EEG patterns, which can be associated to mild or absent brain lesions. On top of that, premature infants show cortical reactivity under noxious stimuli. Continuous exposure to stress and procedural pain might induce pain sensitization and lower the pain threshold, which might be the biological link to affect the functional activity of the infant. Responses to the heel lance show dispersed neuronal bursts on the scalp up to 35 weeks gestational age, while full-evoked potentials are recorded from 36 weeks onwards [78],[236]. The immature burst that is recorded is a slow-wave transient with superimposed faster rhythms, which is defined as delta brush. Therefore, one might expect that the effect of procedural pain can be monitored with functional monitoring, especially EEG activity, which seems to be characterized by increased burst activity in case of pain. Interestingly, the emergence of delta brushes on the somatosensory cortex are also sign of the development of sensory modulation and they are expected to disappear after 35 weeks GA [197]. This EEG reactivity, especially to noxious stimuli, is the fundamental link among pain, background stress and HRV [125].

In general, the possibility to track infant's maturation in the NICU might be helpful to monitor the impact of a variety of perinatal and postnatal factors at cot-side. Although clinicians recognize the importance of electrophysiology, the

complexity of EEG channels and heart-rate variability dynamics might require a level of training, which is not always available. The capacity to automatically track maturation, detect pain, stress or any other abnormal condition, and assess their impact on the developmental trajectory might help to further tune preventive treatment or specific therapies to preserve the developmental outcome. If most of the subtle factors that can affect the development of infants are invisible to the clinical eye or imaging methods, physicians might require new methodologies to detect stress and pain levels from the abnormalities of physiological data. Clinician want to understand if early-life physiological activity can be compromised by stress, cumulated pain or other perinatal factors. Automatic tools based on digital signal processing might then be helpful to detect those abnormalities and predict the impact on the future development of the infants.

2.6 Summary

Although the survival rates of premature infants have steadily increased in the recent decades, the risk of lower cognitive outcomes compared to their age-matched peers is higher. Functional monitoring, such as the EEG and HRV, can assess the level of dysmaturity or development delay and they can be related to the developmental outcome of the patients. EEG shows a decrease of discontinuity and amplitude with infants' maturation, while HRV is characterized by an increase of heart-rate and the power of the main frequency bands. Both the EEG and HRV have been investigated under pain stimuli and perinatal factors such as pain and stress have been proven detrimental for the infant's development. However, automatic tools to detect and monitor stress based on physiological signals are required.

Chapter 3

Mathematical background

This chapter provides a mathematical overview of the different methods to process EEG and heart-rate variability. First, a description of the univariate or one-channel based methodologies to assess dysmaturity will be presented. Second, the connectivity methodologies among the different modalities will be listed. In addition, graph theory will be explained for the topological assessment of functional networks. Third, an overview of regression models and supervised learning methods will be provided.

3.1 Univariate Analysis

The previous chapter clearly defined the following traits of dysmature EEG: discontinuity, persistence of slow-waves and asynchrony. In parallel, a dysmature HRV was also described as a signal with slow-wave oscillations and higher regularity. The current section provides an overview of the features that can be extracted from one signal to assess discontinuity and persistence of slow-waves. Those features span a variety of properties of the EEG and HRV signals, from linear and spectral attributes to nonlinear and fractal parameters. This section starts with classical spectral approaches to continue with a description of entropy features and the multifractal framework. Examples using simulated data are also reported to show how and why this univariate analysis should be tailored to the neonatal EEG and HRV.

3.1.1 Spectral analysis

To assess the contribution of the slow-waves in the EEG or the sympathovagal balance in the HRV, the most common approach is to use the power spectral density and perform an analysis in different frequency bands. In case of EEG, the power is normally computed in the following frequency bands: $\delta_1 = (0.5 - 2]$ Hz, $\delta_2 = (2 - 4]$ Hz, $\theta = (4 - 8]$ Hz, $\alpha = (8 - 16]$ Hz and $\beta = (16 - 20]$ Hz. The division of the δ band between δ_1 and δ_2 is a common procedure in preterm infants due to the maturational changes in the slow-waves bands of the EEG [181], [261]. The HRV power is normally assessed in the following frequency bands: high-frequency $HF = (0.2 - 4]$ Hz, low-frequency $LF = (0.08 - 0.2]$ Hz and very-low frequency band $VLF = (0.0033 - 0.08]$ Hz [59]. In case of HRV, the relative power indices $\frac{VLF}{LF}$,$\frac{LF}{HF}$,$\frac{LF}{LF+VLF}$,$\frac{LF}{HF+LF}$ are commonly derived to express the interplay between the sympathetic and parasympathetic stimulation.

In general, both the EEG and the tachogram are not stationary time-series, which could require time-frequency (TF) analysis to describe the oscillation shifts in the signals. Therefore, the power spectral density (PSD) is usually estimated using Welch's periodogram in non-overlapping windows as well as with time-frequency approaches, such as the quadratic smoothed pseudo Wigner-Ville distribution (SPWD) [187] and the continuous wavelet transform [59]. Given a fixed window size, Welch's algorithm estimates multiple periodograms in overlapping subwindows and averages them. In contrast, the time-frequency approaches estimate the instantaneous autospectrum S_{xx} of the signal $x(t)$. In case of SPWD, S_{xx} is estimated as follows:

$$S_{xx}(t,f) = \int\int_{-\infty}^{+\infty} \Phi_{xx}(\tau,\nu) A_{xx}(\tau,\nu) e^{j2\pi(t\nu-\tau f)} d\tau d\nu, \tag{3.1}$$

where $A_{xx}(\tau,\nu)$ is the ambiguity function, which is defined as the Fourier Transform of the time-dependent auto-correlation of $x(t)$ as follows

$$A_{xx}(\tau,\nu) = \int_{-\infty}^{+\infty} x(t+\tau/2) x^*(t-\tau/2) e^{-j2\pi t\nu} dt, \tag{3.2}$$

where τ represents the lag-domain of the auto-correlation function, ν is the Doppler shift in the frequency domain and * is the complex conjugate operator. The smoothing of the time-frequency cross-coupling in Eq. 3.1 is done via the exponential kernel in the ambiguity domain defined as:

$$\Phi_{xx}(\tau,\nu) = \exp\left\{ -\pi\left[\left(\frac{\nu}{\nu_0}\right)^2 + \left(\frac{\tau}{\tau_0}\right)^2\right]^{2\lambda}\right\}, \tag{3.3}$$

where ν_0, τ_0, λ are kernel function parameters and they are set such that a desired TF resolution of $[\Delta t,\ \Delta f]$ is reached [268]. Therefore, τ and ν can be simply intended as degree of time and frequency filtering [187]. Similarly, $S_{xx}(t,f)$ can be computed as the scalogram of the continuous wavelet transform of the signal as follows:

$$W_{xx}(t,s) = \int_{-\infty}^{+\infty} x(\tau)\psi^*\left(\frac{t-\tau}{s}\right)d\tau, \tag{3.4}$$

$$S_{xx}(t,f) = |W_{xx}(t,f)|^2, \tag{3.5}$$

where ψ is the mother wavelet (normally, Analytic Morlet), while s stands for the scale of the wavelet transform and, in the general, $s \approx f^{-1}$. The lag τ represents the shift in the time domain of the wavelet base ψ.

Based on the given methodologies, the instantaneous power in the band of interest $\beta = [f_1, f_2]$ can be obtained as

$$P_\beta(t) = \int_{f1}^{f2} S_{xx}(t,f)df. \tag{3.6}$$

3.1.2 Entropy Measures

The persistence of slow-waves as well as the discontinuity and the predictability of a signal can be measured using entropy, which assesses the degree of irregularity in a signal. Lake and Moorman proposed the sample entropy ($SampEn$) as statistical estimator of entropy [138]. Generally, this estimator can be conceived as the conditional probability that two short templates or patterns will continue to match inside the signal when their size is increased within a certain tolerance.

Given a data record $x_1, x_2, ...x_N$, let $x_m(i)$ denote a m-long template like $x_i, x_{i+1}, ..., x_{i+m-1}$, with $m < N$ and a starting point i. A match is defined if and only if $x_m(i)$ is within a certain tolerance r of any other template $x_m(j)$, i.e. $|x_m(i) - x_m(j)| < r$. The tolerance r is defined as a fraction of the standard deviation of the signal (normally, 20%), while a common choice for m is 2 or 3. Let B_i denote the matches of dimension m with template $x_m(i)$ and let A_i denote the matches of dimension $m+1$ with template $x_{m+1}(i)$. The cumulative numbers of matches for a certain tolerance r are respectively defined

as $A^{m+1}(r) = \sum A_i$ and $B^m(r) = \sum B_i$. The *SampEn* is defined as the negative natural logarithm of the ratio between $A^{m+1}(r)$ and $B^m(r)$ as follows:

$$SampEn(m, r) = -\log \frac{A^{m+1}(r)}{B^m(r)} \tag{3.7}$$

Figure 3.1 shows how to compute the sample entropy on the right panel. The matches between a certain sample and other future sample is reported by the same color-coding (blue for top-part of the signal, red in the center and green for the lower margin) if inside the r tolerance tunnel. In order to increase the counting of A^{m+1} and B^m, the same pattern should repeat over time, which is indicated by the blue lines in the Figure: the combination of red, green and blue sample repeats over time and, therefore, increases the counting of matches.

The *SampEn* (and its variations, such as quadratic sample entropy or coefficient of sample entropy) was successfully used in a variety of applications with EEG and HRV. They span from atrial fibrillation detection, through the prediction of sepsis in premature infants to the impact assessment of Kangaroo Care on sleep quality and EEG maturation [138], [139], [128],[271].

However, the entropy information might differ across the different scales or time-frame of the signal. Especially, in neonatal HRV and EEG there is a clear distinction between slow-wave baseline and the information in faster rhythms [59],[198]. The complexity of those signals requires not only to understand how the information changes for different scales, but how the acceleration or deceleration of different rhythms might interact among each other [110]. Hoyer et al. found that fetal heart-rate is characterized by a decrease of short-term decelerations and increase of long-term accelerations with development [110]. Similarly, the neonatal EEG presents delta brushes with superimposed faster frequencies than the original theta rhythms, while the infant grows [198],[190].

Therefore, in order to take into account the irregularities at multiple scales, Costa el al. [55] proposed the multiscale sample entropy (MSE), which measures *SampEn* at different scales τ using a coarse-grained version of the signal of interest. Let y^τ denote the coarse grained version of the signal $x(t)$, which is computed as:

$$y_j^\tau = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i, 1 \leq j \leq \frac{N}{\tau} \tag{3.8}$$

For $\tau = 1$, $y^\tau = x(t)$, while y^τ represents a signal average in the window frame of length τ or a moving average of order τ when $\tau \geq 1$ [112]. After averaging, a

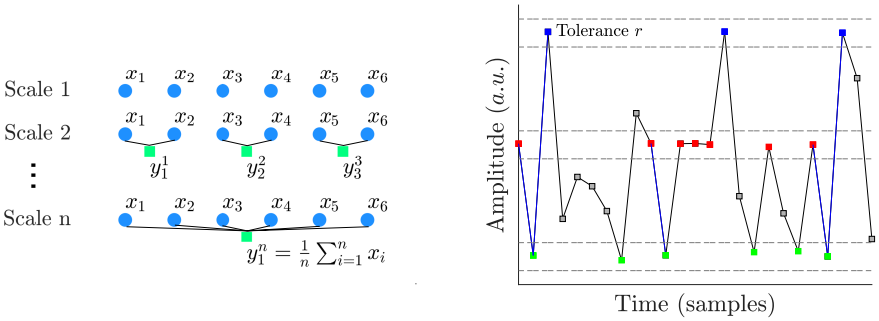


Figure 3.1: The two steps to compute the Multiscale Entropy. The left panel shows the coarse-graining procedure, which is an averaging and downsampling at each scale (as shown by the blue and green dots). The right panel shows the computation of the $SampEn$, which is repeated at each scale. The pattern match is indicated by the blue lines and the color-coding of the red,green and blue samples. The figure is adapted from [65].

downsampling of factor τ follows. The MSE is simply the $SampEn$ function of the different y^τ, as follows

$$MSE(m, r, \tau) = SampEn(m, r, y^\tau) \tag{3.9}$$

where τ normally spans from 1 to 20. The two steps of MSE are reported in Figure 3.1. The left panel shows the coarse graining procedure starting from the original signal down to a generic scale n: each level represents an average of the samples in a window of n samples (as shown by the blue and green dots). The right panel shows the computation of the $SampEn$, which is repeated for each scale.

The complexity of a signal is then computed as the area under the MSE curve, $CI = \sum_\tau MSE(\tau)$, which is a general measure of irregularity across scales. Different authors also investigated MSE at specific scales, such as $MSE(\tau = 3)$ and $MSE(\tau = 20)$, since they represent represent the information at small scales or high frequency and at longer scales or lower frequency, respectively [110].

The complexity of those rhythms might underlie specific physiological processes (like autonomic stimulation) or might be of greater development value (the

contribution of the slow-waves to EEG) [110],[198]. In order to illustrate which information the complexity index provides, we computed the $MSE(\tau)$ for the following set of signals: white Gaussian noise, pink noise, Brownian motion, an example of squared modulated noise and an example of EEG (Figure 3.2). The modulated noise has been generated as product between a white noise and a periodic signal of Gaussian waves in order to mimic the Burst-InterBursts pattern of the neonatal EEG (Figure 2.4). Examples of a Brownian motion and a modulated noise are reported in Figure 3.4. In general, a regular signal should have low entropy value for all scales, while stochastic signals may have different behaviors [55]. The white noise and uncorrelated signals have a decreasing MSE, which means a high entropy at short scales and a lower entropy at higher scales (golden diamonds curve). On the contrary, a pink-noise or $1/f$ signal have a persistent enropy at all scales since it contains a long-range correlations [55] (pink circles curve). Furthermore, regular signals like a fractal Brownian motion and modulate noise have far lower entropy than noise signals (green-squares curve and blue hexagons, respectively). This reduced entropy is due to regular structure of this signal. In addition, both signals have an increasing entropy to highlight the predominance of the long-range correlation over the short-term entropy. Especially, the modulated noise resembles the structure of the discontinuous neonatal EEG (especially in comparison Figure 2.4) and the squared version adds nonlinear phase interactions which enhances the long-range correlations (as discussed by [230] in case of a quadratic phase signal). The example of neonatal EEG locates in the middle of this set (silver-stars curve). The entropy will increase over the different scales, but the entropy in the short-term is general higher than a regular signal such as a Brownian motion. In the long-term scale, the entropy of pink-noise and EEG are similar to each other (Figure 3.2), which also implies that the physiological signal reach an entropy plateau. Eventually, the complexity index and $MSE(\tau)$ can describe the regularity or the "richness" of information, especially for a discontinuous pattern. Interestingly, the MSE does not only capture the information dynamic at different scales, but it also describes the infant's EEG as a process which mixes stochastic patterns with fractal and nonlinear behavior.

3.1.3 Multifractality

A more discontinuous EEG signal or a more dysmature HRV is normally characterized by a lower entropy, since its pattern has higher predictability or memory-persistence. This means that the signal can be easily predicted by its past sample and, therefore, a more discontinuous signal is characterized by higher regularity. In particular, Figure 3.2 highlights that neonatal EEG is more than a stochastic process. Normally, regular signals are characterized by

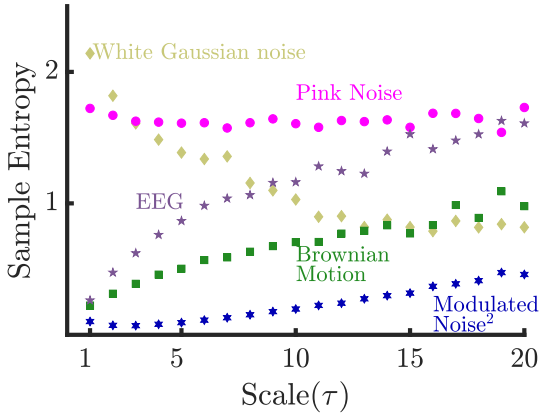


Figure 3.2: Example of MSE curves computed for the following sets of signals: white Gaussian noise (golden diamonds), pink noise (pink circles), Brownian motion (green squares), an example of squared modulated (blue hexagons) noise and an example of EEG (silver stars). The white Gaussian noise have a decreasing entropy, the pink noise have a persistent entropy over all scales and the two regular signals have increasing entropy over the different scales. The neonatal EEG rapidly increases for the short-term scales and reach a plateau of entropy at higher scales similar with a comparable $Samp_{En}$ to the pink noise.

a wide autocorrelation function and a power-law spectrum [71], whose rate of decay is controlled by the Hurst exponent as follows

$$S_{xx}(f) = |C| f^{-2(H-1)} \tag{3.10}$$

In case that time series is normally characterized by one exponent, this signal is also defined as self-similar (it repeats itself over time) or monofractal. However, regular or discontinuous signals might have multiple exponents h to control the degree of regularity over time [118]. Those signals are known as multifractals. In general, signals with a high-degree of regularity are known as fractals or scale-free signals. Small values of h represent sharp and transient regularity or singularity (known also as the *minimal regularity*), while large values represent smooth changes (known also as the *maximal regularity*) [151] [2]. In case of fractal signals, it is possible to quantify the embedding or fractal dimensions associated to each Hurst exponent. This function $D(h)$ is called the *singularity spectrum* (SS) and its determination is pivotal to assess the amount of singularities in the signal. Consequently, the monofractal signal has one embedding dimension,

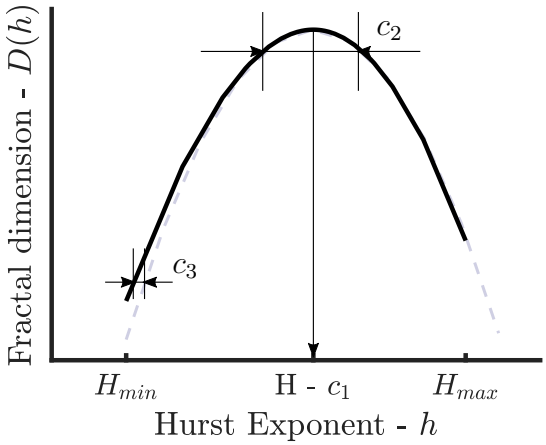


Figure 3.3: Example of singularity spectrum obtained with the multifractal framework (black line). The main multifractal parameters obtained from the singularity spectrum are the following: the cumulant c_1 or H represents the location of the $D(h)$ maximum, c_2 the width of the $D(h)$ distribution and c_3 is its asymmetry, which is intended as deviation from a concave parabola centered on c_1 (grey dashed lines). The extremes H_{max} and H_{min} respectively represent the maximal regularity and the minimal regularity. Another measure of the width of SS obtained as $\Delta H = H_{max} - H_{min}$.

while a multifractal signal has distribution of dimensions associated to each value of h.

A possible way to estimate the SS is the multifractal formalism based on the wavelet leaders. Let $\psi_0(t)$ be a mother wavelet with a positive number of vanishing moments. The discrete wavelet transform (DWT) is defined by the inner product $d_{j,k} = \int_{\mathbb{R}} x(t)\psi_{j,k}(t)dt$, which decomposes $x(t)$ into elementary time-frequency components by means of translation $2^j k$ and dilation or scale $a = 2^j$ of the mother wavelet [266], since $\{\psi_{j,k}(t) = 2^{-j}\psi_0(2^{-j}t - k), j \in \mathbb{R}, k \in \mathbb{R}\}$. Large scales describe smooth and low frequency oscillations, while small scales describe the sharp transitions in the signal. Wavelet leaders $L_f(j,k) \equiv L_\lambda$ represents the maximum wavelet coefficient in the narrow time neighborhood 3λ over all finer scales. Let $\lambda \equiv \lambda_{j,k} \equiv [k2^j, (k+1)2^j)$ be a dyadic interval, such that $d_\lambda = d_f(j,k)$ and let $3\lambda \equiv \lambda_{j,k-1} \cup \lambda_{j,k} \cup \lambda_{j,k+1}$ the union of three dyadic intervals, the wavelet leader is then defined as:

$$L_f(j,k) \equiv L_\lambda = \sup_{\lambda' \subset 3\lambda} \{|d_{\lambda'}|\}, \tag{3.11}$$

According to [266] and [211], a partition function $Z(a,q) = Z(2^j,q)$ can be estimated using the *wavelet leader* $L_f(j,k)$, as follows:

$$Z(a,q) = Z_L(2^j,q) = \frac{1}{n_k}\sum_{k=1}^{n_k} |L_f(j,k)|^q \sim 2^{j\tau(q)}, \tag{3.12}$$

where $\tau(q)$ is known as the scaling exponent and controls the rate of decay of the partition function. This formulation includes also the special case $q = 2$, which represents the scalogram or the power spectral density of the signal. In general, for certain values of q, the exponent $\tau(q)$ has a specific meaning: for positive q, $Z(a,q)$ reflects large fluctuations in the signal, while $Z(a,q)$ reflects short fluctuations for negative q. In general, for each q, the partition function exhibits a power-law decay characteristic, such as the power spectrum of the 1/f noise in (6.2). The scaling exponent associated to this decay can be obtained by computing the slope of Z versus the scales in a log-log diagram. The use of the log-transform clearly shows the advantage of defining the scale as $a = 2^j$. Formally, the function $\tau(q)$ is estimated as follows:

$$\tau(q) = \lim_{j\to 0} \inf \left(\frac{log_2(Z_L(2^j,q))}{j} \right), \tag{3.13}$$

Computation-wise, $\tau(q)$ is estimated in the log-log diagram from a certain scale $a_1 = 2^{j_1}$ to $a_2 = 2^{j_2}$, which are normally defined according to the oscillations that one wants to investigate.

In case of a monofractal signal, $\tau(q)$ is a linear function $\tau(q) = qH - 1$, where H is the global Hurst exponent. In case of a multifractal signal, $\tau(q)$ is a nonlinear function of the local exponents h as expressed by $\tau(q) = qh - D(h)$, where $D(h)$ is the singularity spectrum. Consequently, $D(h)$ can be written as $D(h) = qh - \tau(q)$ where $h = \frac{d\tau(q)}{dq}$. Since $\tau(q)$ can be decomposed in terms of cumulants or coefficient c_p using a Taylor expansion, $\tau(q) = \sum_{p=1}^{\infty} c_p \frac{q^p}{p!}$, the singularity spectrum can be also reformulated in terms of coefficients c_p. Figure 3.3 shows the meaning of first three cumulants: c_1 represents the location of the maximum, c_2 the width of the $D(h)$ distribution and c_3 is its asymmetry, which is intended as deviation from a concave parabola centered on c_1 (grey dashed line in the Figure 3.3). The parameter c_1 is normally considered as the main or global Hurst exponent of the signal and is then referred to as H_{exp} or H. In order to assess the smooth and sharp transients in the multifractal signal, it is standard practice to derive the maximal and minimal Hurst exponent and the difference between the two ΔH, which is another measure of dispersion of the fractals inside the signal similar to c_2 (Figure 3.3). The parameter H_{max} represents the maximal regularity or smooth transitions, while the parameter H_{min} represents the minimal regularity or sharp transitions.

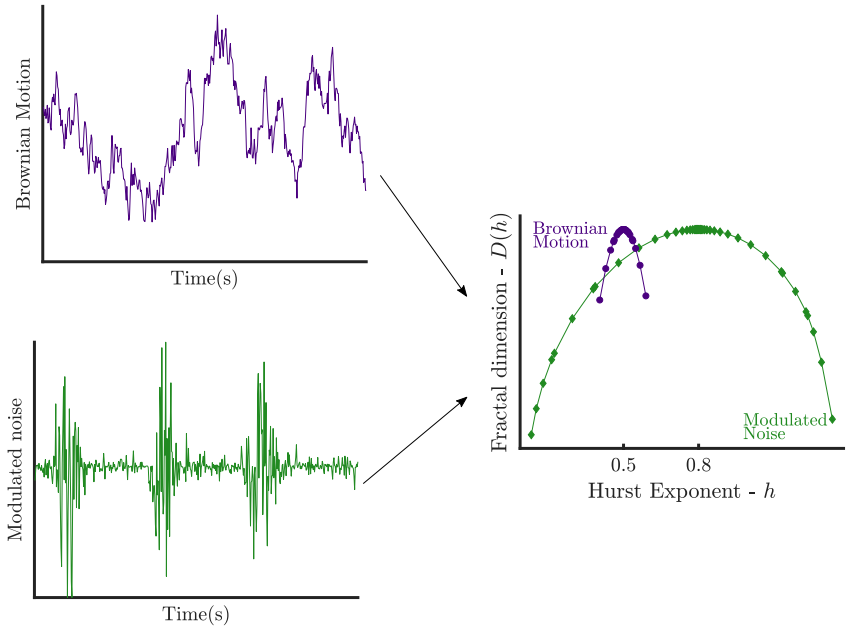


Figure 3.4: Examples of two singularity spectra $D(h)$. The top panel shows a Brownian motion with one single Hurst exponent, while the bottom panel displays modulated noise obtained by the product between white noise and periodic Gaussian waves (green curve). The right panel shows the singularity spectra of the Brownian motion (purple circles) and of the modulated noise (green diamonds). The latter has a very wide $D(h)$ and higher global H, while Brownian motion has a narrow SS around its real Hurst exponent.

In practice, both monofractal and multifractal signals can be investigated with SS. The approach with wavelet leaders always derives a $D(h)$ as indicated in Figure 3.4, which reports two examples for a Brownian Motion with one single Hurst exponent and a modulated noise. The former is a monofractal signal by design, while the latter is the product of white noise with a periodic signals of Gaussian waves to mimic the neonatal EEG. Brownian motion has a very narrow SS around its real Hurst exponent, while modulated noise has a very wide SS and a much higher parameter c_1. Based on these simulated data, one can draw some basic conclusions. First, the spectrum $D(h)$ and its triplet $c_{1,2,3}$ seems to detect a multifractal pattern for a discontinuous signal. Second, since the discontinuity is expected to change with infant's development, $D(h)$

seems uniquely placed to describe changes in EEG or HRV or discriminate continuous from discontinuous patterns, especially by combining H and the width parameter c_2. Third, the multifractal framework always generates a SS and the location of its maximum can be used to represent the overall fractality of the signal.

3.2 Multivariate analysis

The univariate time-series analysis is commonly used to study features of a single signal of the neurophysiological system under investigation. Nevertheless, Pereda et al. pointed out that an increasing number of experiments are carried out with simultaneous recording of several neurophysiological signals [200]. Additionally, Bartsch et al. highlighted the increased number of experiments where signals of different organs and systems are recorded in last decades [23]. The assessment of the interdependence among signals can give new insights into the functioning of the systems that generate them. As Pereda et al. described in [200], the univariate analysis alone cannot accomplish this investigation and a multivariate analysis is required.

Alongside the decrease in the persistence of slow-waves and of discontinuity, different studies show how neonatal EEG presents changes in the interactions among channels, known also as connectivity [197]. In particular, the asynchrony among bursts or SATs is expected to increase, especially in the lower frequency bands [251], [216],[261]. In contrast, the high-frequency information is expected to synchronize, either as bursts with alpha/beta rhythms or ERPs due to noxious or tactful stimuli [198],[237]. Numerous authors pointed out that this key feature is normally visually assessed by the temporal coincidence of SATs [251], [216], [197]. However, a refined analysis of the connectivity among EEGs requires both a full-montage EEG synchronous recordings and more sophisticated methodologies to study the coupling. As shown in Section 2.2.1, the conventional EEG channels are more difficult to interpret and clinicians rely more on quantitative EEG analysis, especially for synchrony which lack of an unanimous definition of the clinical community [216].

In order to investigate the interdependence among neurophysiological signals, we referred to the established framework of Brain Connectivity defined by Friston et al. [91]. They distinguished three levels of connectivity: structural, functional and effective connectivity. The structural connectivity refers to the anatomical architecture of white and grey fibers, which allow the actual transfer of information among brain regions via physical "wiring". The functional and the effective connectivity refer to statistical correlations or dependencies among

cortical activities of different brain regions. The difference between these two levels lies on the estimation of the directionality. The functional connectivity only estimates the coupling without specifying the direction of the information flow, while the effective connectivity is able to determine both the intensity of the coupling and the direction. The investigation of the functional level of the brain relied on the use of different methods, which investigate either the time or the frequency domain as well as the linear or nonlinear couplings.

In recent years, those methodologies and the functional connectivity framework were also applied beyond the world of neuroscience and neuroimaging. Although the medical community traditionally focused on each single physiological system, the human organism is an integrated network, where the different organ systems interact among each others [23]. The field of Network Physiology aims to investigate this horizontal integration among organ systems and a variety of examples were already reported in the scientific literature, such as the neurovascular coupling (EEG-NIRS), cardiorespiratory interactions and brain-heart interactions [80],[119],[45],[254],[107]. In the neonatology research, Pfurtscheller et al. showed that neonatal EEG bursts are synchronized with increases of heart-rate in the first weeks of life [204]. Interestingly, the same methodologies that Pereda et al. [200] and Friston et al. [91] illustrated for the investigation of neurophysiological brain networks, can be applied in the more extensive framework of Network Physiology, where signals may present different temporal and spatial properties.

Given the increasing availability and tractability of large datasets of different physiological systems and their connections, the application of multivariate analysis in biomedicine led to the rise of network science as a way to characterize network structure and functions [221]. The macroscopic behavior of brain networks or physiological networks is characterized by a specific regularity which emerges by their interacting elements [43]. Therefore, the multivariate analysis is presented in two steps. The first paragraphs describe how to estimate interactions among signals and how to assess their statistical validity. The multivariate analysis section subsequently focuses on graph theory and network properties to investigate the organizational architecture of those interacting systems.

3.2.1 Functional Connectivity

According to Friston et al. [91], functional connectivity is an observable phenomenon that can be quantified by measures of statistical correlation, without any explicit assumption on the data-generation model. For that purpose, different methods have been developed to test whether there is any dependency

among signals [200]. In this chapter, we are going to review some of the most common multivariate methods to test interaction in the time and frequency domain as well linear and nonlinear approaches.

Coherence

One of the most common methods to investigate the dependence in the frequency domain is the coherency or coherence. The definition is:

$$C_{xy}(f) = \frac{P_{xy}(f)}{(P_{xx}(f)P_{yy}(f))^{1/2}}, \tag{3.14}$$

which measures the strength of the linear relationship between two signals x and y at a specific frequency f and it is in general a complex number. $P_{xy}(f)$ is the cross-spectrum between channels x and y, while $P_{xx}(f)$ and $P_{yy}(f)$ are autospectra of each channel, respectively. Consequently, the coherence is a normalized cross-spectrum. The most common way to describe the correlation of two signals in the frequency domain is the squared magnitude of the coherence

$$k^2_{xy}(f) = |C_{xy}(f)|^2, \tag{3.15}$$

which is also known as the *magnitude squared coherence* (MSC). However, Nolte [182] pointed out that the true interactions can be measured via the imaginary part of the coherence ($ImCoh$)

$$ImCoh_{xy}(f) = \mathcal{I}(C_{xy}(f)), \tag{3.16}$$

which assesses only the lagged interactions. This distinction between lagged and instantaneous interaction is pivotal because the link among signals is not necessarily unique, especially in case of scalp EEG networks. Although an increasing number of channels might increase the spatial resolution in the functional description of the brain, the interactions among channels might be explained by the same underlying source which is projected on two different channels. However, couplings like $ImCoh$ emerge only if there is a phase shift among the time series. In this way, the lagged interaction should be robust to spurious connectivity, which can arise from any linear mixing of independent sources due to volume conduction [104]. According to Nolte [182], if a linear superposition of K sources $s_k(f)$ is assumed for any channel x

$$X(f) = \sum_{k=1}^{K} a_{xk} s_k(f), \tag{3.17}$$

the only distorted part of cross-spectrum $P_{xy}(f)$ will be the real part, as shown below:

$$P_{xy}(f) =< X(f), Y^*(f) >= \sum_{k=1}^{K} a_{xk} a_{yk'} < s_k(f), s^*_{k'}(f) >= \tag{3.18}$$

$$= \sum_{k=1}^{K} a_{xk} a_{yk} < s_k(f), s^*_k(f) >= \sum_{k=1}^{K} a_{xk} a_{yk} |s_k(f)|^2, \tag{3.19}$$

where $< X(f), Y^*(f) >$ is the dot product between two signals. Consequently, this instantaneous mixing is only reflected in the real part of coherency and might only affect the MSC. For these reasons, if independent sources are mixed at the sensors' level, the imaginary coherency will vanish to zero since it measures only the time-shifted couplings.

A possible interpretation of $C_{xy}(f)$ is normally a correlation or cosine-similarity in the frequency domain between signal x and y. Given two sine signals of amplitudes A_x and A_y and frequencies f_x and f_y respectively, the complex coherence in (3.14) can be re-written as:

$$C_{xy}(f) = \frac{A_x A_y e^{j(\phi_x - \phi_y)}}{\sqrt{|A_x|^2}\sqrt{|A_y|^2}} = \frac{z_x z_y^*}{\sqrt{|z_x|^2}\sqrt{|z_y|^2}} \tag{3.20}$$

where $A_x e^{j\phi_x} = X(f = f_x)$ and $A_y e^{j\phi_y} = Y(f = f_y)$, which represent the Fourier transform of the sine at their specific frequencies. Therefore, ϕ_y and ϕ_x will be the phase of the signal at frequencies f_x and f_y. The right part of (3.20) of the inner product between the complex numbers z_x, z_y highlights the interpretation as correlation or similarity of two signals at each frequency f. If we consider $f_x = f_y$ and zero delay between two sines, the $C_{xy}(f)$ will be a real number on the complex plane and it will be bounded between 0 and 1 based on the amplitudes of the sinusoid signals, as highlighted by the red bar in the top panel of Figure 3.5. The value 1 is reached when $A_x = A_y$. Consequently, instantaneous interaction among periodic signals will be reflected by the real part $C_{xy}(f)$ and in general by $k^2_{xy}(f)$. In contrast, if we consider $f_x = f_y$ and an additional phase delay $\Delta\phi = \frac{\pi}{2}$, the cosine similarity in (3.20) will be still bounded between 0 and 1, but it will be a complex number on the imaginary

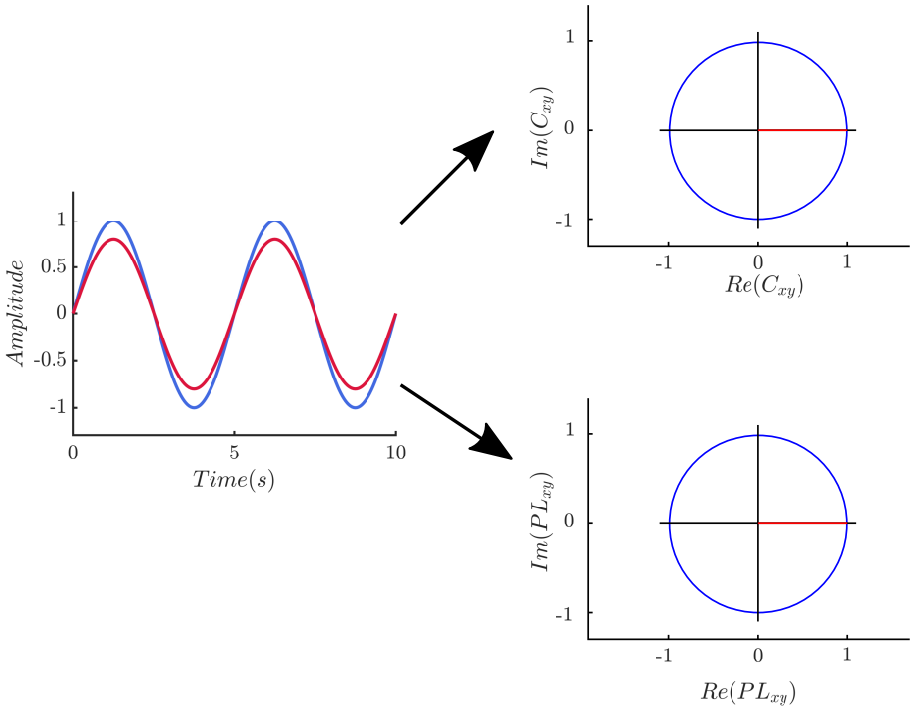


Figure 3.5: Examples of application of the coherence $C_{xy}(f)$ and the phase-locking vector PL_{xy} on two sines at the same frequency without any delay. The coherence will be a real number on the complex plane (the red bar in the top panel) as well as the PL_{xy} (the red bar in the bottom panel).

axis (see the red bar in the top panel of Figure 3.6). Those two figures clearly highlight how $ImCoh_{xy}(f)$ measures the only lagged interaction between the signals. Therefore, one might interpret the coherence as correlation measure at each frequency point. However, one may decide to focus on the lagged or instantaneous interaction based on the objectives of the analysis.

Coherence found numerous applications in the world of neurophysiology [200]. In case of the neonatal EEG, it is worth to report the studies in neurodevelopment by [95],[177],[168]. The normal processing pipeline consists of the computation of MSC and IC in the main EEG frequency bands, which are $\delta_1 = (0.5 - 2]$ Hz, $\delta_2 = (2 - 4]$ Hz, $\theta = (4 - 8]$ Hz, $\alpha = (8 - 16]$ Hz and $\beta = (16 - 20]$ Hz. After proper surrogate testing Section 3.2.3, one can derive the mean of the coherence function or the maximum value in these bands [96].

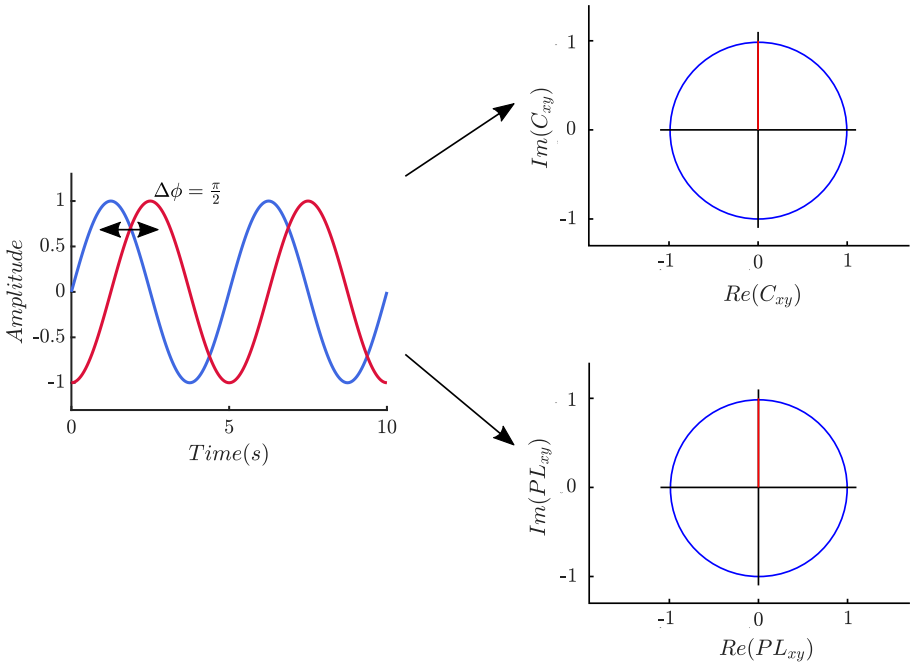


Figure 3.6: Examples of application of the coherence $C_{xy}(f)$ and the phase-locking vector PL_{xy} on two sines at the same frequency with a phase delay $\Delta\phi = \frac{\pi}{2}$. The coherence will be an *imaginary* number on the complex plane (the red bar in the top panel) as well as the PL_{xy} (the red bar in the bottom panel).

Phase synchrony

Alongside coherence and other linear measures, physiological systems may also synchronize even if signals are not correlated [200]. By synchronization, it is meant that two phase of two coupled oscillators might lock at any time t, which implies that:

$$|m\varphi_x(t) - n\varphi_y(t)| < \text{constant} \tag{3.21}$$

where $\varphi_x(t)$ and $\varphi_y(t)$ are the (unwrapped) phases of signals x and y in the system under investigation. The coefficients $m : n$ are integer multipliers and they are used to indicate synchronization of oscillators at different velocities.

The current framework will only discuss the theory related to $m = 1$ and $n = 1$. In order to investigate the degree of phase synchronization stated in (3.21), one might compute two synchrony indices which rely on the phase difference $\gamma_{xy}(t) = \varphi_x(t) - \varphi_y(t)$. The first one is the *phase locking value* PLV_{xy}, which is defined as:

$$PLV_{xy} = \frac{1}{N}\left|\sum_{t=0}^{N} e^{j\gamma_{xy}(t)}\right|, \tag{3.22}$$

where N is the length of the data record. The second is one is the *phase lag index* PLI_{xy}

$$PLI_{xy} = \frac{1}{N}\left|\sum_{t=0}^{N} \mathrm{sgn}(\gamma_{xy}(t))\right|, \tag{3.23}$$

where sgn is the sign function. Similarly to the coherence, one can consider the complex number

$$PL_{xy} = \frac{1}{N}\sum_{t=0}^{N} e^{j\gamma_{xy}(t)}, \tag{3.24}$$

to understand the meaning of the phase synchrony indices. For simplicity, we can also consider two sines at the same frequency, as reported in Figure 3.5 and Figure 3.6. By definition, the phases of two sines will be linearly increasing, but the phase difference will be constant. Therefore, the complex *phase-locking vector* will be $PL_{xy} = \frac{1}{N}\sum_{t=0}^{N} e^{j\gamma_{xy}(t)} = e^{j\Delta\phi_{xy}}$, where $\Delta\phi_{xy}$ is the phase-delay between two sines. If we consider an instantaneous interaction, PL_{xy} will be a real number on the complex plane (see the red bar in the bottom panel of Figure 3.5). If there is a lagged interaction with $\Delta\phi = \frac{\pi}{2}$, PL_{xy} will be an imaginary number on the complex plane (see the red bar in the bottom panel of Figure 3.5). One can clearly see that PLV_{xy} is just the absolute value of PL_{xy} that captures the instantaneous interactions, while the imaginary part of PL_{xy} and the distribution of $\gamma_{xy}(t)$ only focus on the lagged interactions.

Both PLV_{xy} and PLI_{xy} were involved in numerous applications that considered a variety of clinical domains (e.g. Parkinson, Alzheimer and so on) and different recording modalities (e.g. MEG and EEG) as well as simulated and real datasets [200],[240]. Stam et al. also highlighted the parallelism between phase-locking and coherence indices: PLV normally measures the instantaneous connectivity, while PLI focuses on the lagged interactions and is, therefore, resistant to

volume conduction. By definitions in (3.22), the phase locking will reflect both zero-lag differences and lagged interactions. Although it is insensitive to the amplitude since it is based on the phase, PLV_{xy} will be high in case of sources mix. On the contrary, PLI_{xy} investigates the asymmetry in distribution of $\gamma_{xy}(t)$: if a difference in phase persists, the index will be non-zero. An asymmetry in phase difference implies a consistent lag or time-delay between signals.

Phase synchrony requires phase extraction before its assessment. The phase of the signal is normally derived in narrower bands than the original full-band in order to obtain an oscillator at specific frequency. The normal processing pipeline in a band-pass filtering either by using a discrete wavelet transform or a continuous wavelet transform consists of the main EEG frequency bands δ_1, δ_2, θ, α and β. The Hilbert-transform is applied to derive the phase of the signal in each frequency band. Therefore, a coupling value can be computed in the different frequency bands for the different phase methods.

Mutual Information approaches

Since the multivariate analysis tries to establish whether there is any common information between signals, one can investigate this interdependence by means of information theoretical tools [200]. The *mutual information* between signals X and Y is defined as:

$$MI_{XY} = \sum p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \tag{3.25}$$

where the $p(x)$ and $p(y)$ are the probability distributions of the two signals, while $p(x,y)$ is the joint probability distribution of two signals. The advantage in this formulation is the generic definition of independence: two signals are totally independent when $p(x,y) = p(x)p(y)$. Although the information-theory approach is more generic than synchrony and coherence, the probability distribution estimation can be computationally expensive, e.g. when using a parametric-free approach [173].

In case of unknown density distribution, a cost-sound estimation is based on kernel estimation and reproducing kernel Hilbert spaces (RKHS). Given a joint distribution $p(x,y)$, we can define the cross-covariance operator as:

$$C_{xy} = \mathbb{E}([\phi(x) - \mu_y] \otimes [\psi(y) - \mu_x]), \tag{3.26}$$

where $\phi(x)$ and $\psi(y)$ are nonlinear feature maps of x and y (mapping from domain $\mathcal{X}, \mathcal{Y}$ to the RKHS domain $\mathcal{F}, \mathcal{G}$). Furthermore, μ_x and μ_y are the expected value of the two feature maps: $\mu_x := \mathbb{E}(\phi(x))$ and $\mu_x := \mathbb{E}(\psi(y)))$, while $\otimes$ is the tensor product [244],[98]. The cross-covariance operator is a generalization of the cross-covariance matrix between two random vectors or signals.

Based on the (3.26), one can define *Hilbert - Schmidt Dependence* (HSD) as follows

$$HSD(p(x,y), \mathcal{F}, \mathcal{G}) = ||C_{xy}||^2_{HS}, \tag{3.27}$$

which indicates that two signals are independent if HSD is null, given a certain RKHS. The powerful advantage to estimate MI via HSD is the empirical estimate of the RKHS dependence [98], which is computed as follows:

$$HSD = (1-N)^{-1}\mathbf{tr}(KHLH), \tag{3.28}$$

where $K, H, L \in \mathbb{R}^N$, $\mathbf{tr}$ is the trace of the matrix $KHLH$ and N is the length of the signal. K is the associated kernel to $\phi(x)$ and L is the associated to $\psi(x)$, i.e. $K = k(x_i, x_j) =< \phi(x_i)\phi(x_j) >$ and $L = l(x_i, x_j) =< \psi(x_i)\psi(x_j) >$. The matrix H is defined as $\mathbf{H} = \mathbf{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^T$. In a nutshell, one can estimate HSD in the moment a kernel function is defined and substitute the MI with HSD.

The mutual information found also numerous applications in MEG and EEG in variety of clinical domains, such as synchrony and coherence. In the context of this thesis, we decided to compute HSD in narrower bands than the original full-band similarly to the synchrony approaches. Therefore, we applied a discrete wavelet transform or a continuous wavelet transform to main EEG frequency bands δ_1, δ_2, θ, α and β and compute MI in different frequency bands.

Time-frequency coherence

The methods considered so far implied stationarity or stable dynamics. However, signals such as the neonatal EEG can vary their spectral properties over time [251]. The approach of coherency can be extended to a time-frequency distribution approach thanks to the continuous wavelet transform (CWT), already described in section 3.1.1. The time-frequency coherency is computed as the continuous wavelet coherence:

$$C_{x_i \leftrightarrow x_j}(t,f) = \frac{s_{x_i \leftrightarrow x_j}(t,f)}{\sqrt{s_{x_i}(t,f)s_{x_j}(t,f)}}, \tag{3.29}$$

where $s_{x_i \leftrightarrow x_j}(t,f)$ is the cross-scalogram between the signal x_i and x_j , $s_{x_i}(t,f)$ and $s_{x_j}(t,f)$ are the autoscalogram of the signals [207]. If $s_{x_i}(t,f)$ is normally computed as the squared magnitude of the wavelet transform, i.e. $s_{x_i}(t,f) = |W_{ii}(t,f)|^2$, $s_{x_i \leftrightarrow x_j}(t,f)$ is the cross-product of the wavelet-transform x_i and x_j, i.e. $s_{x_i \leftrightarrow x_j}(t,f) = W^*_{ii}(t,f)W_{jj}(t,f)$. The wavelet transforms are computed according to (3.4) and the mother wavelet is analytic Morlet, which is a complex function, for both the cross-scalogram and the autoscalograms. The clear advantage of this type of wavelet is clear in a multivariate analysis, because one may decide to compute both a magnitude squared wavelet coherence and an imaginary wavelet coherence as follows:

$$k^2_{x_i \leftrightarrow x_j}(t,f) = |C_{x_i \leftrightarrow x_j}(t,f)|^2, \tag{3.30}$$

$$ImCoh_{x_i \leftrightarrow x_j}(t,f) = \mathcal{I}(C_{x_i \leftrightarrow x_j}(t,f)), \tag{3.31}$$

The time-variant coherence has been instrumental in the description of the interaction between heart-rate variability and EEG, whose interdependence can variate during epilepsy [207] or neonatal burst activity [230]. The equations (3.30) and (3.31) clearly show that wavelets investigate the dynamics of the coupling over time. Similarly to the synchrony approach, the signal is band-pass filtered in the reported frequency bands for EEG or in VLF, LF and HF in case of HRV. As shown by another study [225], the match between temporal scales of HRV and EEG is normally reached in two steps. First, the EEG oscillations are extracted with wavelet-based approaches. Second, the Hilbert-transform amplitude is derived and the signal is downsampled to sampling frequency of the HRV.

3.2.2 Effective Connectivity

Unlike functional connectivity, effective connectivity tries to infer the direction of the interdependence or coupling and, therefore, relies on a specific model how data were caused in the observed system [91]. In general, a common strategy to estimate the direction and the coupling in the multivariate analysis is the computation of different models and the comparison among those. The main concept is the following: if the computed models are statistically different, one can infer the directions or a first approximation of the investigated connectivity. However, directionality methods have specific underlying hypotheses which is always fundamental to report to provide a full overview of the estimated interaction [91]. Consequently, the scientific literature

about effective connectivity presents different methods to assess directional coupling among time series [35].

In the context of this thesis, the focus of most of the methods is on the causal investigation of *G-causality*. Based on the importance of temporal ordering of events, both Wiener and Granger defined causal relationship as follows: if two signals X and Y are recorded simultaneously and signal Y is better predicted if both the past samples of two signals are incorporated in the prediction, signal X is said to cause Y [14],[200]. Granger specifically developed regression models to explain the future Y in case of linear forecasting modeling. Therefore, it is commonly said that X Granger-causes to Y, in order to highlight the specific domain and hypotheses of this definition [14]. Figure 3.7 shows an example of this type of causal-relationship in the context of neurophysiology: the future sample of the EEG signal Y_t can be predicted by the past samples Y_{t-k} and X_{t-k}.

However, Figure 3.7 also shows that multiple recordings are often collected in biomedicine and the applicability of the Granger-cause definition should go beyond the bivariate case X and Y. A common framework to estimate directional interdependencies with multiple channels is the conditional multivariate Granger causality (cMVGC), which is based on Vector AutoRegressive (VAR) modeling and often referred to as *G-connectivity* or *G-causality* [18],[14], [224], [173]. This type of modeling aims to estimate the active direct link between two channels in the time domain given the existence of other signals in the experiment of interest. Specifically, the cMVGC can be estimated by fitting AutoRegressive (AR) models or via information dynamics, which are respectively described in [18] and [173].

Granger Causality

The first method employed to assess the *G-causality* among EEG channels was Granger Causality (GC) with a VAR modelling framework, following the formulation by [20]. A pth order multivariate autoregressive model VAR(p) takes the form:

$$\mathbf{U}_t = \sum_{k=1}^{p} \mathbf{A}_k \cdot \mathbf{U}_{t-k} + \epsilon(t), \tag{3.32}$$

where $\mathbf{U}_t$ is a collection of process U_i, while $\mathbf{A}_k$ and $\epsilon(t)$ are respectively the regression coefficient matrices and the stochastic process residuals. In the conditional scenario [20], where we want to know the pairwise influence of process $\mathbf{X}_t$ over $\mathbf{Y}_t$ considering the presence of the other U_i variables, $\mathbf{U}_t$ can be rewritten as:

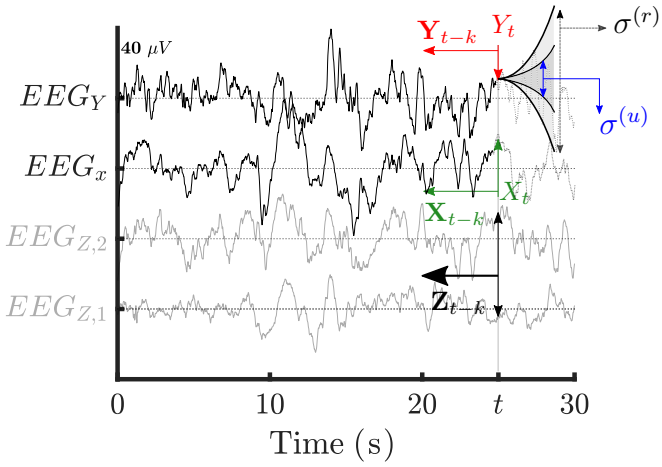


Figure 3.7: An explanation of the meaning of Granger Causality. The uncertainty of the prediction of the signal EEG_Y is given by the two funnels in grey. The future of Y_t has greater variance (and hence uncertainty) if we only consider the the past samples of Z_t. This estimation risk $\sigma^{(r)}$ will diverge the longer the time span that one wants to investigate into the future (as indicated by the bigger black dashed double arrow). On the contrary, the variance of estimation $\sigma^{(u)}$ is smaller in case we also consider the past samples X_{t-k} and the divergence of the estimation will be limited (as indicated by the smaller double blue arrow).

$$\mathbf{U}_t = \begin{bmatrix} \mathbf{X}_t \\ \mathbf{Y}_t \\ \mathbf{Z}_t \end{bmatrix}, \tag{3.33}$$

where $\mathbf{X}_t$ and $\mathbf{Y}_t$ are the two time series of interest, while $\mathbf{Z}_t$ represents the third set of variables involved in the analysis. In our study, $\mathbf{X}_t$ and $\mathbf{Y}_t$ can be two EEG channels, e.g. EEG_Y and EEG_x in Figure 3.7. Consequently, $\mathbf{Z}_t$ will be a vector time series which contain the remaining EEG channels ($EEG_{Z,1}$ and $EEG_{Z,2}$ in Figure 3.7). Based on the $VAR(p)$ model (3.32) and the split defined in (3.33), we may actually explain the future of $\mathbf{Y}_t$ based on

the following full and reduced regressions

$$\mathbf{Y}_t = \sum_{k=1}^{p} \mathbf{A}_{yy} \cdot \mathbf{Y}_{t-k} + \sum_{k=1}^{p} \mathbf{A}_{yx} \cdot \mathbf{X}_{t-k} + \\ + \sum_{k=1}^{p} \mathbf{A}_{yz} \cdot \mathbf{Z}_{t-k} + \epsilon_y(t), \tag{3.34}$$

$$\mathbf{Y}_t = \sum_{k=1}^{p} \tilde{\mathbf{A}}_{yy} \cdot \mathbf{Y}_{t-k} + \sum_{k=1}^{p} \tilde{\mathbf{A}}_{yz} \cdot \mathbf{Z}_{t-k} + \tilde{\epsilon}_y(t). \tag{3.35}$$

In both cases we consider the conditioning variable $\mathbf{Z_t}$, although only the first model considers explicitly the influence of $\mathbf{X_t}$. The difference is also highlighted by the two regression coefficient matrices $\mathbf{A}_{yy}$ and $\tilde{\mathbf{A}}_{yy}$. In order to test whether the coefficient matrices $\mathbf{A}_{yx}$ are significantly different from zero, the Granger Causality is defined as the log-likelihood ratio:

$$\mathcal{F}_{X \to Y|\mathbf{Z}} = \log \frac{|\tilde{\Sigma}_{yy}|}{|\Sigma_{yy}|} = \log \frac{\sigma^{(r)}}{\sigma^{(u)}}, \tag{3.36}$$

where Σ_{yy} and $\tilde{\Sigma}_{yy}$ are the covariance matrices of the residuals $\epsilon_y(t)$ and $\tilde{\epsilon}_y(t)$. As reported in Figure 3.7, if we consider only target EEG signal Y_t, $|\Sigma_{yy}|$ and $|\tilde{\Sigma}_{yy}|$ coincide with the variances $\sigma^{(r)}$ and $\sigma^{(u)}$ of the (monodimensional) residuals $\epsilon_y(t)$ and $\tilde{\epsilon}_y(t)$. The letter r stands for the *reduced* model in (3.35), while u is for the full model (3.34). Based on this multivariate framework, the G-causality basically quantifies the reduction of the prediction error if the variable X is added to explanatory variables of Y, one of which is the variable $\mathbf{Z}$ [18]. An explanation of the meaning of $\mathcal{F}_{X \to Y|\mathbf{Z}}$ is reported in Figure 3.7. The future of the signal EEG_Y starts at the time-point t and the variance or the uncertainty is given by the two funnels in grey. The future of Y_t has greater variance (and hence uncertainty) if we only consider the past samples of Z_t. This estimation risk $\sigma^{(r)}$ will diverge with time. On the contrary, the variance of estimation $\sigma^{(u)}$ is smaller in case we also consider the past samples X_{t-k} and the divergence of the estimation will be limited. Based on Figure 3.7, one can understand the concept of comparison of different models to infer causality.

Friston et al. highlighted the necessity to compare the models to test the statistical significance of the model [91]. Although the value of GC can be used as amount of the information transfer, the clear-cut advantage of this parametric modeling is one of its statistical properties. According to [173] and

[18], the GC follows an asymptotic F-distribution, like R^2 statistics. Therefore, one can implement an F-test or log-likelihood ratio test to assess the statistical validity of this coupling.

Unlike the classical Granger causality test limited to the bivariate case, $\mathcal{F}_{X\to Y|\mathbf{Z}}$ is defined as *conditional multivariate Granger causality*. A special case of the conditional scenario is the pairwise conditional G-causality, where the pairwise effective couplings among all pair of variable U_i and U_j contained in $\mathbf{U}_t$ are measured. In other words, when all the possible combinations of X and Y are tested and the dimension of Y and X strictly limited to one signal. In principle, the cMVGC can consider Y and X are multidimensional as well. However, the methods of this thesis will be limited to pairwise conditional causalities.

Since we take in account the spurious effect due to the presence of other variables (i.e. the coupling between U_i and U_j conditioned by the presence of the other variables), the cMVGC framework defines a set $M \times M$ pairwise conditional causalities as follows:

$$\mathcal{G}_{ij}(\mathbf{U}) = \mathcal{F}_{U_i\to U_j|\mathbf{U}_{[\mathbf{ij}]}}, \tag{3.37}$$

where M is the number of processes. Therefore, all the pairwise coupling estimates are contained in one single matrix. Granger-causality was investigated in multiple applications, especially involving EEG or even subcortical activity [200]. In case of neonatal EEG, GC gives the specific opportunity to investigate the aforementioned asynchrony and assess whether the cortical connectivity changes with the neurodevelopment. Similarly to the imaginary coherence and PLI, the GC is uniquely placed to investigate the lagged-interactions, which have a specific role in the definition of asynchrony among neonatal bursts. Unlike other brain-connectivity approaches, effective connectivity were far less-applied in the neonatology research and therefore new insights might be obtained.

Transfer Entropy

The second method applied to assess the *G-causality* among EEG channels is Transfer Entropy, which is defined according to information dynamics framework by Schreiber et al. [227] as follows:

$$T_{X\to Y} = \sum p(Y_t, \mathbf{X}_t^m, \mathbf{Y}_t^n) \log \frac{p(Y_t|\mathbf{X}_t^m, \mathbf{Y}_t^n)}{p(Y_t|\mathbf{Y}_t^n)}, \tag{3.38}$$

where $p(Y_t|\mathbf{X}_t^m, \mathbf{Y}_t^n)$ is the conditional probability that Y at times t is explained by past values of X and Y with respective memory order m and n. $p(Y_t|\mathbf{Y}_t^n)$ is the conditional probability that Y at times t is only explained by past values of Y. $p(Y_t, \mathbf{X}_t^m, \mathbf{Y}_t^n)$ is the joint probability distribution among the three variables

Y_t, $\mathbf{X}_t^m$, $\mathbf{Y}_t^n$. The second term is also defined as Kullback-Leibler entropy or divergence and it communicates a very similar idea to the one reported in Figure 3.7: the signal X causes signal Y if the probability to have the future sample of Y_t considering the past information of X and Y is higher than the probability considering only the past samples of Y_t. Similarly to GC, TE also relies on a log-likelihood ratio. Therefore, Transfer entropy inherently implies directionality (i.e. effective connectivity) [227], [256].

The definition in (3.38) has two main problems. First, it is a bivariate definition. Second, it requires a method to estimate the conditional probabilities. In order to estimate the information dynamics coupling in a multivariate context, Montalto [173] proposed to estimate TE as the difference of two following conditional entropies (CE):

$$T_{X\to Y|\mathbf{Z}} = H(Y_t|\mathbf{Y}_t^n, \mathbf{Z}_t^p) - H(Y_t|\mathbf{Y}_t^n, \mathbf{X}_t^m, \mathbf{Z}_t^p), \tag{3.39}$$

where $\mathbf{Z}_t^p$ is introduced. Since we cannot discriminate the exclusive relationship between two processes X and Y [97], [18], it is possible to investigate the contribution of time series X in the evolution of time series Y with respect to all the other agents involved in the analysis. Consequently, we can actually define $\mathbf{Z}_t^p$ as a vector variable that does contain neither Y nor X or their past values, and the $T_{X\to Y|\mathbf{Z}}$ illustrates the transfer of information from X to Y taking into account other time series involved. One may immediately notice the similarities of notation between $\mathcal{F}_{X\to Y|\mathbf{Z}}$ and $T_{X\to Y|\mathbf{Z}}$.

If we then assume that X, Y, $\mathbf{Z}$ have a joint gaussian distribution, the two CE in (3.39) can be expressed as linear regression of past values of the vector variables involved in the multivariate system as follows

$$\begin{aligned} Y_t &= A^u V^u + \xi(t) \\ Y_t &= A^r V^r + \xi(t) \end{aligned} \tag{3.40}$$

The first equation explains Y_t with a regression on the vector $V^u = [V_x, V_y, V_z]$, where V_x,V_y,V_z approximates respectively $\mathbf{X}_t^m$, $\mathbf{Y}_\mathbf{t}^\mathbf{n}$, $\mathbf{Z}_t^p$ with a vector of size p as follows:

$$\begin{aligned} V_x &= [X_{t-1}, X_{t-2}, \dots, X_{t-p}] \\ V_y &= [Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}] \\ V_z &= [\mathbf{Z}_{t-1}, \mathbf{Z}_{t-2}, \dots, \mathbf{Z}_{t-p}] \end{aligned}$$

The second equation explains Y_t with a regression on the vector $V^r = [V_y, V_z]$, which only contains V_y,V_z. Equations (3.40) are usually referred to as full and restricted regressions. The reader can easily recognize the same structure of

Granger causality equations (3.34) and (3.35), which is asymptotically equivalent to Transfer entropy [227]. In light of the previous results, $T_{X\to Y|\mathbf{Z}}$ can be rewritten as

$$T_{X\to Y|\mathbf{Z}} = \frac{1}{2}\log\frac{\sigma^{(r)}}{\sigma^{(u)}} \tag{3.41}$$

where $\sigma^{(r)}$ and $\sigma^{(u)}$ are the variances of the white noise residuals in (3.40). Except for a scalar factor, (3.41) is the same expression of (3.36). Therefore, the interpretation of (3.41) is the same one for (3.36) and reported in Figure 3.7: the combination of that past samples of X and Y reduces the uncertainty of prediction compared to the a model that considers only the past samples of Y, as highlighted by the two grey funnels ($\sigma^{(u)} < \sigma^{(r)}$).

Similarly to cMVGC, we can define a set $M \times M$ pairwise conditional causalities as follows:

$$\mathcal{G}_{ij}(\mathbf{U}) = T_{U_i\to U_j|\mathbf{U}_{[\mathbf{ij}]}}, \tag{3.42}$$

where M is the number of processes. Once again, all the pairwise coupling estimates are contained in one single $M \times M$ matrix. Similarly to Granger-causality, TE was investigated in multiple neurophysiological applications [200]. In the context of this thesis, we only looked to a specific case of TE, which resembles the GC definition. The reason to apply TE in our research was to add enough redundancy to investigation of the lagged asynchrony and changes the cortical connectivity throughout the neurodevelopment.

Directionality Index

A more generalized approach than bivariate Granger Causality is the estimation of directionality in case a dependent dynamics exists between two systems after phase reconstruction [220]. Similar to the Granger-Causality, this methodology estimates if the phase dynamics or the future phase of a signal is explained by the past phase of another signal. Therefore, this method is based on the phases of two signals or require a phase extraction of those methods, like the synchrony approaches in Section 3.2.1. The key difference is estimation of a directionality index (DI) to assess the direction and the intensity of the coupling.

Given two self-sustained oscillators with a weak coupling, their phase dynamics $\dot{\phi}_{1,2}$ can be expressed as:

$$\dot{\phi}_{1,2} = \omega_{1,2} + \epsilon_{1,2} f_{1,2}(\phi_{2,1}\phi_{1,2}) + \xi_{1,2}(t), \tag{3.43}$$

where $\omega_{1,2}$ are the natural pulse frequencies of the two systems. The functions $f_{1,2}$ describe how the interaction between the phases $\phi_{1,2}$ drives the two systems

and $\epsilon_{1,2}$ are parameters to describe the intensity of the coupling (usually $\epsilon_{1,2} \ll \omega_{1,2}$). $\xi_{1,2}(t)$ are noise perturbations. The coupling $\epsilon_{1,2}$ $f_{1,2}$ $(\phi_{2,1}, \phi_{1,2})$ can be inferred with a least squares fitting procedure, where the digital phase increments $\phi_{1,2}$ are related with a probe function $F_{1,2}^{m,n}$, which is a finite Fourier series of order m,n as follows:

$$\Delta_{1,2} = \phi_{1,2}(t+\tau) - \phi_{1,2}(t) = \omega_{1,2}(t)\tau + F_{1,2}^{m,n}(\phi_{1,2}(t), \phi_{2,1}(t)) + \eta_{1,2}(t), \quad (3.44)$$

$$\epsilon_{1,2} f_{1,2} \approx F_{1,2}^{m,n} = \sum_{m,n} A_{m,n} e^{jm\phi_1 + jn\phi_2}, \quad (3.45)$$

where $A_{m,n}$ are the coefficients of the Fourier series in (3.45) and τ is a fixed interval to sample the phase dynamics. From the smooth function $F_{1,2}^{m,n}$, the cross-dependencies between phase dynamics can be estimated as:

$$q_{1,2}^2 = \int \frac{\partial F_{1,2}}{\partial \phi_{2,1}}^2 d\phi_2 d\phi_2 \quad (3.46)$$

Eventually, the directionality index $d_{1\to 2}$ is introduced as

$$d_{1\to 2} = \frac{q_1 - q_2}{q_1 + q_2} \quad (3.47)$$

which is value bounded between [-1,1]. Consequently, this index gives an integrated measure of how strongly a system drives (a value close to 1) or how sensitive it is to be driven (close to -1).

In layman terms, the DI estimates whether the surface in (3.45) has a specific gradient along one of two phases $\phi_{1,2}$. In case that gradient is non-zero, one of the two quantities $q_{1,2}$ will not only be non-zero, but it might happen that $q_1 \gg q_2$ (or $q_2 \gg q_1$), which implies a specific direction from one system to other. In case that the gradient is non-zero, the dynamics of two phases is explained by the coupling contribution in (3.43). In contrast, if the surface is flat and, therefore, the gradient in (3.45) is null, we have that $q_1 \approx q_2$ and the estimation of $\epsilon_{1,2} f_{1,2} \approx 0$, which implies that there is no coupling. In conclusion, the two oscillators have independent dynamics, which means $\dot{\phi}_{1,2} = \omega_{1,2} + \xi_{1,2}(t)$.

The advantage of phase dynamics is the estimation of both linear and nonlinear coupling. The usage of phases and a generic coupling harmonic function between them extend the applicability beyond VAR models. However, the disadvantages have to be stressed before its application. The method is bivariate by design, which strongly limits its application in brain connectivity or Network Physiology. The hypothesis of weakly coupled oscillators has to be investigated and a proper estimation of the phases is required before the assessment of DI. Therefore,

Rosenblum et al. [220] suggests to apply this methodology only in known cases of the existence of oscillatory processes. Among those phenomena, it is worthwhile to mention the influence of the respiration over the heart-rate (known as respiratory sinus arrhythmia) and the interaction between neonatal bursts and the heart-rate increases, as shown by [230].

3.2.3 Surrogates and phase analysis

Section 3.2.2 touched upon the concept of statistical validity and statistical comparison among models to infer the direction of interdependence among signals. The estimated dependencies derived with the different functional connectivity methods (but also with the directionality index) might be completely spurious. In order to test if the signals are really dependent, one can refer to data surrogates [200].

Surrogates were originally introduced to determine whether a certain property of a signal significantly differs from N_{surr} time-series constructed by either shuffling the original data or by annihilating specific traits of the considered signals [140]. Specifically, they were vastly applied to determine if processes have certain nonlinear properties. If this property or statistic of the original time series deviated from the surrogates distribution of the same property, the data record is unlikely to be originated by a linear model. This statistical test was subsequently extended to multivariate analysis [140], [200]. The concept is similar to the univariate case: a set of N_{surr} surrogates is first constructed with the objective to destroy any actual dependency among those artificial time-series. The estimated coupling is actually statistically significant if the estimated coupling differs from the distribution between signals and the surrogates. The statistical test is normally built on the null-hypothesis that the signals are independent and the surrogate test rejects the null-hypothesis.

Since the described functional connectivity methods rely on lagged interactions (coherence) or synchrony of phase (PLV and PLI), the generation of surrogate for this type of multivariate analysis is based on the original record such that only the linear properties are preserved, while the phase of the signal is disrupted. Specifically, the data are shuffled via the *adjusted-amplitude Fourier transform algorithm* (AAFT) in each epoch, which rescales the signals with a Gaussian process and randomizes the phase to destroy any lagged interaction or synchrony among the surrogate signals [228]. Consequently, the AAFT surrogates data testing consists in the re-estimation of the coupling with the surrogates and test whether the original lagged interaction is actually different from the distribution of lagged interactions of the surrogates. In case of the directionaliy index (Section 3.2.2), the coupling specifically concentrates on the phase dynamics and

surrogates via the Cycle-Phase permutation testing, which randomly permutes cycles of a signal phase to generate surrogates [140]. In both cases, whether the null hypothesis is rejected and the value obtained with the experimental data is different, the coupling is defined as significant and it is normally preserved in set of $M \times M$ dependencies. Otherwise, the coupling is set to zero.

Surrogation might be useful to assess the robustness of dependencies in the world of neurophysiology, where large datasets are collected and the risk of obtaining a coupling by chance is present. However, the surrogation might also be computationally expensive since it requires re-estimation of the interdependence among surrogates time series. In order to reduce the computational costs, Gonzalez et al. suggested to reduce the number of surrogates to the specific level of statistical significance that is required [95]. For each coupling between two generic signals X and Y, the couplings between X and 19 surrogates of Y are estimated. The coupling is considered significant if and only if it is higher than the surrogate-based 19 interaction intensities. The number 19 guarantees an $\alpha = 0.05$ significance level since $N_{surr} = \frac{1}{\alpha} - 1 = 19$. Another approach is to know the distribution of the estimated coupling and avoid surrogates, which is only possible in case the data-generation model is known. This has already been introduced for Granger Causality and Transfer Entropy in Section 3.2.2, which are known to follow an F-distribution and have the advantage to be automatically tested for statistical significance.

3.2.4 Graph indices

The list of reported multivariate approaches has the common characteristic to derive a set of $M \times M$ couplings. Even if the surrogate testing is supposed to preserve only the real coupling, the general pipeline of any connectivity method is reported in Figure 3.8. Given a set of signal S_j or x_j, one estimates the generic dependencies $\mathcal{C}_{x_i \leftrightarrow x_j}$ which can be reorganized in the following coupling matrix $M \times M$:

$$\mathbf{A} = \mathcal{A}_{ij} = \mathcal{C}_{x_i \leftrightarrow x_j}, \tag{3.48}$$

where M is the number of processes and $i, j = 1, .., M$. The matrix $\mathbf{A} = A_{ij}$ is defined as an adjacency matrix A and reports the intensity of interaction among all possible pairwise interactions, as shown in the right-top panel of Figure 3.8. The intensity can be the original coupling or can be rescaled between 0 and 1. In general, $\mathbf{A}$ describes a weighted graph [43],[20], as highlighted by the central bottom of Figure 3.8. A graph is a diagram of points or nodes, whose relations are defined by the connecting lines among them. The adjacency matrix

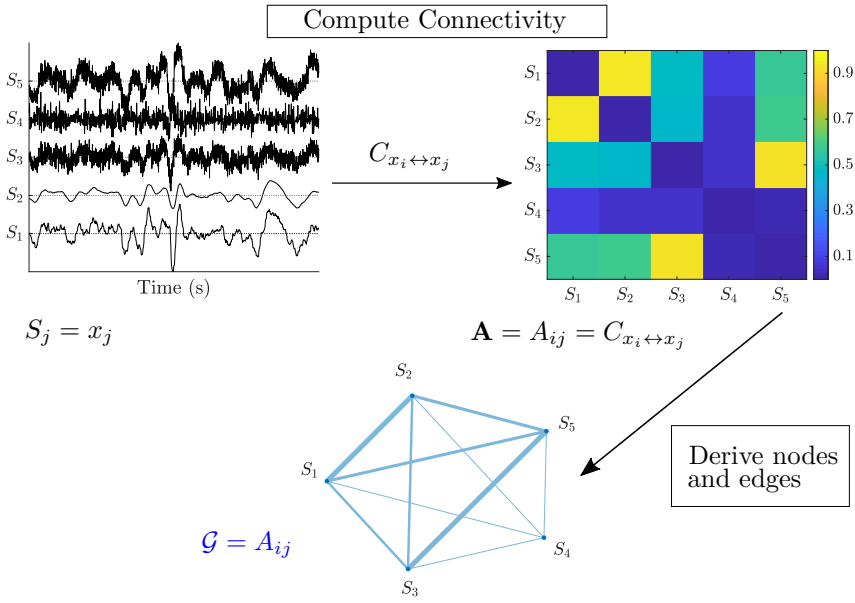


Figure 3.8: The typical processing pipeline that leads a multivariate analysis to a weighted graph. Given a starting set of signal $S_j = x_j$, one computes the $M \times M$ pairwise interactions $C_{x_i \leftrightarrow x_j}$ and derives the adjacency matrix $\mathbf{A}$. The latter can provide the information to draw the associated weighted graphs: each column represents a node of the matrix and the entries of the matrix represent both the existence and the weight of each link. The intensity of the coupling is normally highlighted with a different thickness of the edge.

will allow to represent this graph (as also indicated by the name adjacency). Each column represents a node and a signal in case of multivariate analysis, while the entries represent the link or edges among the nodes. Therefore, if each node represents a signal (as also shown in the graph of Figure 3.8), the existence of dependency among those nodes is represented by the drawn blue link. The intensity of the coupling is represented by the weight associated to the edge, which is normally represented by a different thickness of the graph. As described in details by [43], one might consider to apply a specific threshold to the coupling in $\mathbf{A}$ to eliminate spurious connections and binarize the network. It means that the intensity of the information is lost for the sake of representing the most important connections. As far as the application of graph is concerned, the full information of adjacency matrix will be considered and the couplings

will be pruned if and only if they do not pass surrogate testing.

Although any connectivity matrix generates an adjaceny matrix, functional connectivity and effective connectivity matrices have different properties. The FC methods generate a symmetric matrix $\mathbf{A}$, which results in an undirected graphs (as highlighted by $x_i \leftrightarrow x_j$ and by the absence of any arrow in any graph of Figure 3.8). On the contrary, effective connectivity methods generate an asymmetric matrix as follows,

$$\mathbf{A} = \mathcal{A}_{ij} = \mathcal{C}_{x_i \to x_j}, \tag{3.49}$$

which results in a directed graph, as highlighted by $x_i \to x_j$. It normally modifies the final graph with an arrow to highlight the direction of the coupling.

Eventually, each multivariate analysis will start from a set of signals, compute the pairwise interactions, derive the matrix $\mathbf{A}$ after surrogate testing and represent the associated weighted graphs. The purpose to the associated network is twofold. On one hand, the number of couplings that a multivariate approach generates in general $M^2 - M$ and a dedicated investigation to each of them risks to be unfeasible. Graph theory can significantly reduce the amount of information generated by the multivariate analysis [43],[38]. On the other, the complexity properties of the network can be investigated. The random interactions among different signals might emerge in a regular architecture whose key organizational principles might be shared among different physiological systems and they can be modified by neurodevelopment or different physiological states [43].

Consequently, those organizational principles can be assessed by a list of topological and spectral indices to describe the architecture of the estimated network. A visual representation of some of those topological and spectral indices is reported in three graph examples reported in Figure 3.9. The most common metrics to assess the level of network's integration are the path length, the global efficiency, the diameter, the clustering coefficient and the eccentricity.

- The **path length** is the minimum number of edges or the shortest path that must be traversed to go from one node to another, as shown in the blue path reported in Figure 3.9. The shortest route to move from the node in the extreme bottom left to extreme top right node is a diagonal walk throughout the graph. Any other walk will risk to have a longer route. The average of all path lengths or characteristic path length is the mean of the nodes' shortest paths and can be considered a measure of network integration capacity [26]. The shorter the blue paths of all

nodes, the higher the integration in the network, because few steps will be needed to move from any node to any other one.

- The **global efficiency** is the average of the inverse path length. It is another measure of integration and it can be also interpreted in a energy perspective: the higher the efficiency, the lower the energy dissipated to move inside the network. However, Peters et al. [202] highlighted that the global efficiency is primarily driven by shorter paths (stronger connections) while characteristic path length is primarily driven by longer paths (weaker connections).

- The **clustering coefficient** is defined as the average of all weighted triangles around a node and mirrors the graph coupling density [81]. A visual interpretation of this index is reported in Figure 3.9. A triangle is here intended as the smallest non-trivial motif that a node can form with two neighboring nodes, as represented in the top-left panel graph. The clustering coefficient quantifies the number of connections that exist between the nearest neighbours of a node as a proportion of the maximum number of possible connections [43]. It is possible to derive as the average of all clustering coefficients as overall measure of clustering or level of interconnection of all nodes in the network as further measure of integration.

- The **eccentricity** of a node represents the maximum distance from one node to any other node in the graph and the the **diameter** of graph is then defined as the maximum eccentricity in the graph.

- One can compute also the **graph density** as the sum of all significant couplings of the adjacency matrix [21], which represents the energy of the entire network [43].

Alongside those network metrics, the spectral indices such as the spectral radius, the spectral gap and the algebraic connectivity were considered.

- The **spectral radius** is the maximal absolute eigenvalue of the adjacency matrix [30]. Also known as "Page Rank", it represents the dominance degree of a node in the network: the higher the value, the higher the centrality of the dominant node in the network. In other words, the dominant node behaves as the center of a hub [30]. An example of graph with a high spectral radius is reported in Figure 3.9, where the blue node has a dominant role in the provincial hub of Figure 3.9.

- Another way to look into the change of dominance is the **spectral gap**, which is the difference between the first and the second absolute eigenvalues

of the adjacency matrix, [77]. If the difference decreases, the node with highest eigenvalue (the spectral radius) is less dominant with respect to the other nodes in the network. Therefore, the blue node in Figure 3.9 will determine a graph with a high spectral radius. However, Estrada [77] argued that the spectral gap behaves as a measure of clustering in the undirected graph: in case of lower values of spectral gap, the graph can present small clusters in the network.

- The last spectral measure is the **algebraic connectivity**, which is the second smallest eigenvalue of the Laplacian matrix L and it illustrates how easily a graph can be divided into clusters or communities [38]. For example, the graph reported in Figure 3.9 has high algebraic connectivity because of the two distinct communities. In case of an undirected graph, the Laplacian matrix can be obtained as follows

$$L = D - A \tag{3.50}$$

where A is the adjacency matrix obtained by the effective connectivity tools described above and D is the degree matrix of the associated graph [49]. Since both A and D are symmetric, L will be also symmetric and the eigenvalues will be real. In case of a directed graph, the Laplacian matrix L_{dir} will not be symmetric. Therefore L_{sym} can be obtained as follows

$$\begin{aligned} L_{dir} &= D - A \\ L_{sym} &= \frac{1}{2}(L_{dir} + L_{dir}^T) = D - \frac{A + A^T}{2} \end{aligned} \tag{3.51}$$

where A is the transposed adjacency matrix.

Alongside the integration and spectral indices, the amount of superfluous connections of the adjacency matrix can also be investigated. Any graph can risk to be overly connected or under-connected. Even after surrogate testing, some connections can emerge as significant, but they result from physiological redundant connectivity [202]. Therefore, we can test the resilience, which is the capacity of the network to keep the global connectivity high even if some connections are removed. Therefore, suppose we order all the connections in descending order based on the intensity of the coupling. Suppose also the set of original weights of A is defined as w_{ij}^0. The **number n_{sup} of superfluous connections** is derived as the number that maximizes the following quantity

$$\begin{aligned} &\max_{n} H(w_{ij}(n)) + E(w_{ij}(n)) = \\ &= -\sum_{ij} w_{ij}(n) \log(w_{ij}(n)) + \sum_{ij} (w_{ij}(n) - w_{ij}^{0})^2 \end{aligned} \tag{3.52}$$

where $H(w_{ij}(n))$ is the entropy of the matrix A when n weights were removed. The values $w_{ij}(n)$ represent the remaining non-zero weights, while $E(w_{ij}(n))$ is the squared error between the new matrix A and the original matrix. The number n_{sup} represents the number of removed connections that maintain the global connectivity high without significant deviation from the original matrix. The goal is to remove as many connections as possible, which results in an increasing error and decreasing connectivity entropy. In an extreme scenario, a redundant network keeps the global entropy high even if there are few non-zero coupling in the matrix; and, in general, they will have a higher n_{sup}. An overview of all graph indices is reported in Table 3.1.

Table 3.1: An overview of all graph metrics that have been used in the study.

Overview of integration indices	
Path length	Mean of the nodes' shortest paths
Global Efficiency	Average of the inverse path lengths.
Clustering coefficient	Mean of the nodes' triangles intensity around each node
Eccentricity	Maximum distance from one node to any other node in the graph
Diameter	Maximum graph eccentricity
Causal density	Sum of all significant couplings in A
Overview of spectral indices	
Spectral radius	$\lambda_1(A)$: first eigenvalue of adjacency matrix A
Spectral gap	$\lambda_1(A) - \lambda_2(A)$: the difference between the first two eigenvalues of matrix A
Algebraic connectivity	$\lambda_{M-1}(L)$: the second smallest eigenvalue of the Laplacian matrix L or L_{sym}
Overview of resilience indices	
Superfluous connections	$n_{sup}(A)$: number of superfluous connections of adjacency matrix A

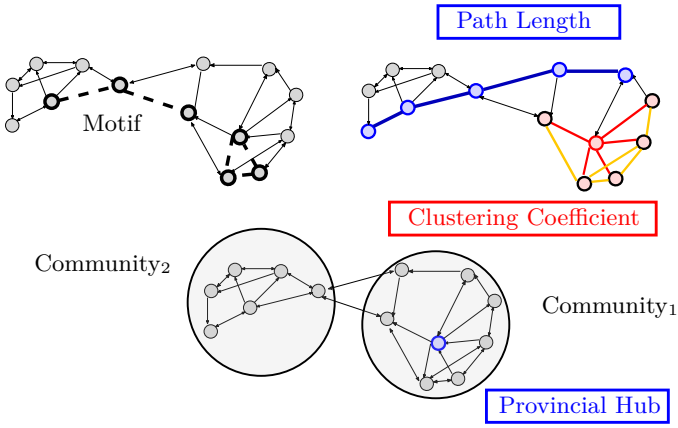


Figure 3.9: Three examples of graphs and key organizational properties that can be computed with graph theory. The figure has been adapted from [43]. The two top figures show some key topological properties: the left graph shows the meaning of triangle motifs, while the right one reports an example of path length and a node with high clustering coefficient. The bottom graph shows a graph with different distinctive spectral properties: the two clear-cut communities will increase the algebraic connectivity of this graph, while the dominant hub on the right will lead to a high spectral radius and spectral gap.

3.3 Regression Models

3.3.1 Ordinary Least Squares regression

The ordinary least squares linear regression or fixed-effect regression is the most common model to examine the relationship between an output or dependent variable and one or more input or independent variables [217] [226]. The most generic case refers to the multiple regression where the dependent variable $\mathbf{y}$ and its variance can be explained by a set of dependent variables $\mathbf{x}_p$:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ...\beta_p x_{ip} + \epsilon_{i1} \tag{3.53}$$

where the index i indicates the number of instances in the dataset, while β_p represents the coefficient of the regression model. The model is normally derived

with a pseudoinverse of the dataset $\mathbf{X}$, which collects all instances x_i and variables $\mathbf{x}_p$, as follows:

$$\boldsymbol{\beta} = (\mathbf{X^T X})^{-1}(\mathbf{X^T y}) = \mathbf{X}^{+}\mathbf{y} \tag{3.54}$$

where $\boldsymbol{\beta}$ is the vector of the coefficient β_p and $\mathbf{X^T}$ is the transpose of $\mathbf{X}$. The matrix $\mathbf{X}^{+}$ is the Moore-Penrose pseudoinverse matrix. Equation 3.54 is derived as the convex solution of the sum of squared errors ϵ_i minimization. The error is defined as follows:

$$\epsilon_i = y_i - \hat{y}_i \tag{3.55}$$

where $\hat{y}_i$ represents the estimated value for the instance i of the dataset, while the sum of squared errors SS_{res} is computed as:

$$SS_{res} = \sum \epsilon_i^2 = \sum (y_i - \hat{y}_i)^2 \tag{3.56}$$

The most common performance metric for the linear regression is the coefficient of determination R^2, which is basically the fraction of the total variation explained by the linear regression model:

$$R^2 = \frac{\text{explained variation}}{\text{total variation}} = \frac{SS_{reg}}{SS_{tot}} = 1 - \frac{SS_{res}}{SS_{tot}} \tag{3.57}$$

where SS_{reg} and SS_{tot} are respectively defined as the sum of squares of the estimated variables $\hat{y}_i$ and the total sum of squares of the output variable y_i as follows:

$$SS_{reg} = \sum (\hat{y}_i - \bar{y})^2 \tag{3.58}$$

$$SS_{tot} = \sum (y_i - \bar{y})^2 \tag{3.59}$$

The variable $\bar{y}$ represents the mean of all observations of the output variable. Other common performance metrics are the *root mean squared error* (RMSE) and the *mean absolute error* (MAE), which are defined as follows:

$$RMSE = \frac{1}{N}\sqrt{SS_{res}} = \frac{1}{N}\sqrt{\sum (y_i - \hat{y}_i)^2} \tag{3.60}$$

$$MAE = \frac{1}{N}\sqrt{\sum |y_i - \hat{y}_i|} \tag{3.61}$$

where N is the number of instances of the dataset. If the RMSE is simply the square root of the SS_{reg}, MAE represents the average absolute deviation from the estimated curve. The absolute value is normally used to reduce the influence of outliers on the performance metrics.

3.3.2 Linear mixed-effect regression

Equation 3.54 underlines some fundamental hypothesis. The first assumption is the independence of the input variable or absence of collinearity. Although this might seriously hinder the pseudo-inversion of $\mathbf{X}$, the collinearity is normally solved with a proper selection of the input variables to avoid interdependence. A greater concern of biomedical datasets is the randomness of the different instances. Although the different values should be drawn randomly from the underlying data distribution, this is not generally the case if the data comes from patients recorded at different time points. This might seriously affect the hypothesis of equal variance of ϵ_i for each instance (homoscedasticity), since there are some hidden autocorrelations in the same patients' data. Therefore, a random-effect has to be added to the fixed effect of β_p coefficients. The model reported in the equation 3.53 is corrected as follows:

$$\begin{aligned} y_i = \beta_0 + \beta_1 x_{i1}^{(m)} + \beta_2 x_{i2}^{(m)} + ... + \beta_p x_{ip}^{(m)} + \\ b_0^{(m)} + b_1^{(m)} x_{i1}^{(m)} + b_2^{(m)} x_{i2}^{(m)} + ... + b_p^{(m)} x_{ip}^{(m)} + \epsilon_i^{(m)}. \end{aligned} \tag{3.62}$$

The coefficients β_p still represent OLS coefficients which do not vary and hence, they are fixed, while the coefficients $b_p^{(m)}$ are random variables themselves. The variability of the coefficients is normally introduced by a grouping variable whose M levels are indicated by the variable (m) $(m = 1..., M)$. The grouping variable considers data points that are not fully independent in one single group or level (e.g. data of the same patient). The distribution of random effects is:

$$b^{(m)} \sim \mathcal{N}(0, \Sigma_b), \tag{3.63}$$

which means that $b^{(m)}$ follows a normal variate and Σ_b is the covariance matrix of the random coefficients, i.e.

$$\Sigma_b = \begin{bmatrix} \sigma_1 & \rho_{1,2} & \cdots \\ \vdots & \ddots & \\ \rho_{1,p} & & \sigma_p \end{bmatrix}$$

.

One can further simplify the expressions in equations 3.62 and 3.63 in one single normal variate $\gamma_p \sim \mathcal{N}(\beta, \Sigma_b)$, where the expected value of the coefficients is the fixed effect β_p with some variations added by the Σ_b of the random effects. Hence, the model is called linear-mixed effect model [217].

In order to estimate both the vector of fixed coefficients $\boldsymbol{\beta}$ and the covariance matrix Σ_b, one may introduce the random effect variable $z_{ip}^{(m)}$ to substitute the group-level variable $x_{ip}^{(m)}$. The z variable expands and nests the x variable for all possible M-levels up to the dimensionality $\mathbb{R}^{qM\times 1}$. The dimension q represents the number of coefficients which will have a random effect upon the fixed-effect and, in general, $q \leq p$. Consequently, one may rewrite (3.62) as follows:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\epsilon} \tag{3.64}$$

where $\mathbf{y} = \{y_i\} \in \mathbb{R}^{N\times 1}$, $\mathbf{X} = \{x_{ip}\} \in \mathbb{R}^{N\times p}$, $\mathbf{Z} = \{z_{ip}^{(m)}\} \in \mathbb{R}^{N\times qM}$ and $\boldsymbol{\epsilon} = \{\epsilon_i^{(m)}\} \in \mathbb{R}^{N\times 1}$. The vector of $\mathbf{u}$ is the random complement to $\boldsymbol{\beta}$, which implies $\mathbf{u} = \{b_{ip}^{(m)}\} \in \mathbb{R}^{qM\times 1}$.

In layman terms, one expects that the variations Σ_{b_p} induced by each group should be small enough not to hinder the overall model designed by the fixed effect and, therefore, the term $\mathbf{Z}\mathbf{u}$ negligible. A visual explanation is reported in Figure 3.10. Each group is indicated by different colors. The randomness of the different groups on the intercept β_0 and the slope β_1 shift and change the orientation of the models in the Cartesian plane. However, the variations are not big enough to deviate the overall trend in the left panel. On the contrary, the right panel of the figure shows the variations introduced by the different groups can distort and annihilate the expected trend.

Similarly to OLS regression, the linear-mixed effect models use R^2, RMSE and MAE as performance metrics. However, the R^2 is corrected to take into account the effect of the randomization due to the grouping variable, which is also defined as *conditional* R^2, [178]. It is normally expressed as follows:

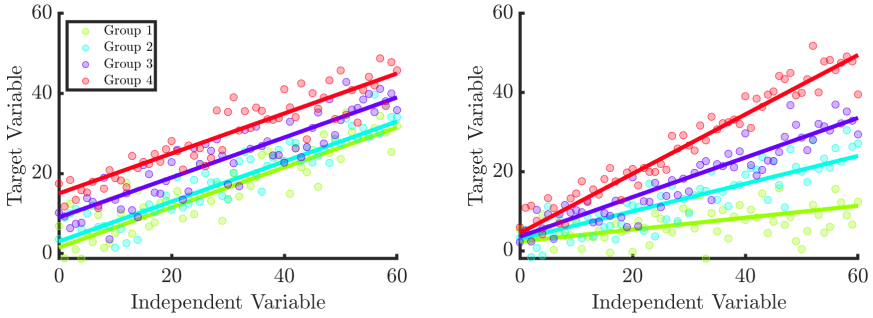


Figure 3.10: Linear mixed effect model examples

$$R^2 = \frac{\text{explained variation}}{\text{total variation}} = \frac{SS_{reg}}{SS_{tot}} = \frac{SS_{reg,fixed} + SS_{reg,group}}{SS_{reg,fixed} + SS_{reg,group} + SS_{res}}, \tag{3.65}$$

where the difference from (3.57) consists in the split of contribution of the variance of the estimated $\hat{y}_i$ given by the fixed effects $SS_{reg,fixed}$ and the variance of the grouping variable due to the normal distribution of the random coefficients b_p ($SS_{reg,group}$). The latter is actually computed as the determinant of Σ_b [178].

3.3.3 LASSO

The OLS regression in (3.54) can be rewritten as the following optimization problem:

$$\min_{\boldsymbol{\beta}} ||\mathbf{X}\boldsymbol{\beta} - \mathbf{y}||_2^2, \tag{3.66}$$

which highlights that the vector $\boldsymbol{\beta}$ is obtained as minimization of the sum of the squared errors ϵ_i. However, one can introduce a penalty term to satisfy certain conditions, such as model simplicity. The regression should not only be able to estimate the target variable, but it should consider as few variables as possible. This implies that some coefficients in $\boldsymbol{\beta}$ are set to zero. A very common way to select some of the variables in the regression model is adding a ℓ_1-norm penalty to (3.67) as follows:

$$\min_{\boldsymbol{\beta}} ||\mathbf{X}\boldsymbol{\beta} - \mathbf{y}||_2^2 + \lambda||\boldsymbol{\beta}||_1. \tag{3.67}$$

This approach is known as Least-Absolute Shrinkage and Selection Operator (LASSO) and the hyperparameter λ makes a trade-off between approximation error and sparsity [247]. Although it is defined for regression, LASSO is one of the most common features selection method. In a supervised learning strategy, LASSO is commonly used to reduce the feature subset before obtaining any regression or classification model.

3.4 Classification Models

Classification as a supervised learning strategy tries to distinguish two or more classes in a group of data points. The classification models aim to find the best possible decision boundary or function that does not only separate the considered data (training dataset), but it can generalize to new data points (test set). Several ways to find a decision boundary exist, such as neural networks, support vector machines, linear discriminant analysis or random forest.

Although neural networks (NN) are well-known to approximate any possible boundary (they are known as *universal approximators*), the weights of the neural networks are normally computed with non-convex approaches. Therefore, the solution is not unique and their convergence might be slow. In addition, their high number of weights and layers might increase the risk of overfitting [243]. Similarly, random forest (RF) might approximate any decision boundary, but they are not exempted from the risk of overfitting. Thanks to the *bagging* process, multiple decision trees can be trained to approximate the decision function in an optimal way. However, all features have to be considered in order to select the final random forest model [108]. In this section, different approaches are discussed to address some of the shortcomings of NN and RF. These methodologies are support vector machines, least-squares support vector machines, linear discriminant analysis (LDA) and subspace LDA.

3.4.1 Support vector machines

Support vector machines are a supervised learning method originally proposed for a 2-class problem and later extended to multiclass problems [252],[54],[152].

The main idea of SVMs is based on the construction of an optimal separating hyperplane between a positive class cluster (C^+) and a negative class cluster

(C^-) with maximal margin. Consider a given training set $\{x_k, y_k\}_{k=1}^N$ with input data $x_k \in \mathbb{R}^d$ and the output data $y_k \in \mathbb{R}$ with class labels $y_k \in \{-1, +1\}$ and the linear classifier

$$y(x) = \text{sign}(w^T x + b), \tag{3.68}$$

where w and b represents the parameters of the hyperplane for the class separation. The two groups C^+ and C^- can be separated if:

$$\begin{cases} w^T x_k + b \geq +1 & \text{if } y_k = 1 \\ w^T x_k + b \leq -1 & \text{if } y_k = -1, \end{cases} \tag{3.69}$$

which can be summarized in one single inequality as follows:

$$y_k[w^T x_k + b] \geq 1, \;\; k = 1, ...N \tag{3.70}$$

In the context of the optimization theory, the support vector machines are defined such that the weights of the hyperplane maximize the so-called margin, which is the distance between the hyperplanes defined in equation 3.69, as also shown in Figure 3.11. Therefore, the weights are defined by the solution of the following problem:

$$\begin{aligned} &\min_{w,b} J_p(w) = \frac{1}{2} w^T w \\ &\text{such that } y_k[w^T x_k + b] \geq 1, \qquad\qquad k = 1, ...N \end{aligned} \tag{3.71}$$

Applying (3.71), we defined the Langragian as the equation:

$$\mathcal{L}(w, b, \alpha) = \frac{1}{2} w^T w - \sum_{k=1}^{N} \alpha_k (y_k[w^T x_k + b] - 1), \tag{3.72}$$

where the coefficients α_k are known as Langrangian multipliers and, in general, $\alpha_k \geq 0$ for $k = 1, ...N$. The solution of (3.72) is characterized by the saddle point of the Lagrangian:

$$\max_{\alpha} \min_{w,b} \mathcal{L}(w, b, \alpha). \tag{3.73}$$

However, most of the problems in real-life problems consists in non-separable cases since data inputs might be missing or data might be noisy or unreliable. Therefore, one has to tolerate some data points misclassification. This is normally done by adding a slack or error variable to (3.70) as follows:

$$y_k[w^T x_k + b] \geq 1 - \xi_k, \;\; k = 1, ...N \tag{3.74}$$

where $\xi_k \geq 1$ allows violation of the inequalities (3.70) and, in general, $\xi_k \geq 0$. Therefore, one might reformulate (3.71) as follows:

$$\begin{aligned} \min_{w,b} J_p(w, \xi) &= \frac{1}{2} w^T w + c \sum_{k=1}^{N} \xi_k, \qquad (3.75) \\ \text{such that } y_k[w^T x_k + b] &\geq 1 - \xi_k, \qquad k = 1, ...N \\ \xi_k &\geq 0, \qquad k = 1, ...N \end{aligned}$$

where c is a positive real constant. Consequently, one might consider the following Lagrangian

$$\mathcal{L}(w, b, \xi; \alpha, \nu) = J_p(w, \xi) - \sum_{k=1}^{N} \alpha_k (y_k[w^T x_k + b] - 1 + \xi_k) - \sum_{k=1}^{N} \nu_k \xi_k, \tag{3.76}$$

where $\nu_k \geq 0$ for $k = 1, ...N$ is a second set of Lagrange multipliers added for the additional slack variables ξ_k. The new solution is then given by the saddle point of the new Lagrangian:

$$\max_{\alpha,\nu} \min_{w,b,\xi} \mathcal{L}(w, b, \xi; \alpha, \nu) \tag{3.77}$$

Applying the Karushn-Kuhn-Tucker conditions, one obtains:

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial w} = 0 \rightarrow w = \sum_{k=1}^{N} \alpha_k y_k x_k \\ \frac{\partial \mathcal{L}}{\partial b} = 0 \rightarrow \sum_{k=1}^{N} \alpha_k y_k = 0 \\ \frac{\partial \mathcal{L}}{\partial \xi_k} = 0 \rightarrow 0 \leq \alpha_k \leq c \end{cases}, \tag{3.78}$$

which gives the dual quadratic programming problem after replacing (3.78) in (3.75):

$$\min_{w,b} J_D(\alpha) = -\frac{1}{2}\sum_{k=1}^{N} y_k y_l x_k^T x_l \alpha_k \alpha_l + \sum_{k=1}^{N} \alpha_k, \tag{3.79}$$

$$\text{such that } \sum_{k=1}^{N} \alpha_k y_k = 0,$$

$$0 \leq \alpha_k \leq c,\ k = 1, ...N$$

for $k, l = 1, ...N$. In contrast with (3.79), the formulation (3.75) is defined as the primal problem. Hence, one might reformulate (3.68) as follows:

$$y(x) = \text{sign}(\alpha_k y_k x_k^T x + b). \tag{3.80}$$

The advantage of (3.79) is twofold. The focus shifts from the weight search of (3.75) to the search of α_k. Based on the conditions reported in (3.78) and (3.79), the weights are defined by the instances x_k. On top of that, since some of the α_k are zeros due to the sparsity condition $\sum_{k=1}^{N} \alpha_k y_k = 0$, only specific data points x_k contribute to define the hyperplane in (3.75) and they are called *support vectors*.

It is clear from (3.80) and from (3.79) that the optimal hyperplane is defined by the correlation distances $x_k^T x_l$ among data points. A nonlinear separation between C^+ and C^- can be defined thanks to a generic function $\varphi(x)$, which maps the data points in feature space where they result linearly separable. The classifier in (3.68) is reformulated as:

$$y(x) = \text{sign}(w^T \varphi(x) + b). \tag{3.81}$$

The dual problem of (3.79) can be rewritten as:

$$\min_{w,b} J_D(\alpha) = -\frac{1}{2}\sum_{k=1}^{N} y_k y_l K(x_k, x_l)\alpha_k\alpha_l + \sum_{k=1}^{N}\alpha_k, \tag{3.82}$$

$$\text{such that } \sum_{k=1}^{N}\alpha_k y_k = 0$$

$$0 \leq \alpha_k \leq c,\ k = 1, ...N,$$

where $K(x_k, x_l) = \varphi(x_k)^T\varphi(x_l)$ is a kernel function the expresses a distance in high-dimensional feature space. The advantage of the so-called *Kernel trick* is to avoid the definition of feature map $\varphi(.)$, since the final classifier will result in

$$y(x) = \text{sign}(\alpha_k y_k K(x_k, x) + b), \tag{3.83}$$

In the linear case, one might express the kernel function as correlation distance $K(x_k, x_l) = x_k^T x_l$ ($\varphi(x_k) = x_k$), but other distance functions exist, such as:

$$K(x_k, x_l) = \exp\left(-\frac{||x_k - x_l||^2}{\sigma^2}\right) \tag{3.84}$$

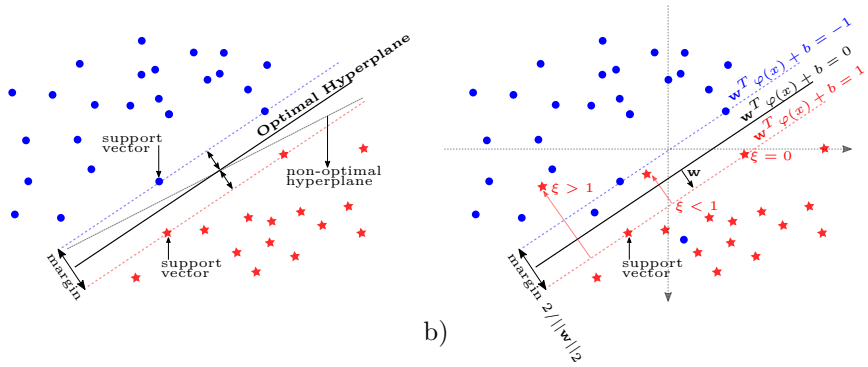
or

$$K(x_k, x_l) = (x_k^T x_l + \tau)^p, \tag{3.85}$$

where σ,τ and p are parameters of the kernel function and they are normally defined hyperparameters of the SVMs classifier together with the regularization variable (3.75) and (3.82). All hyperparameters are normally computed with a nonconvex solution of an additional optimization problem (in contrast to the convex solutions of the α_k). Standard procedure is to compute these hyperparameters using crossvalidation, e.g. $10 - fold$ crossvalidation.

3.4.2 Least squares support vector machines

A further simplification of SVMs was proposed by Suykens et al. [242]. Starting from the formulation in (3.75), one might convert the inequality condition in an equality condition:



a) b)

Figure 3.11: The SVMs find the optimal hyperplane by maximizing the margin between the two classes. On the left, the perfect linearly-separable case is depicted and, on the right,the linearly-separable case with some tolerance expressed by the slack variable.

$$\min_{w,b} J_p(w,e) = \frac{1}{2}w^T w + \gamma\frac{1}{2}\sum_{k=1}^{N} e_k^2, \tag{3.86}$$

$$\text{such that } y_k[w^T x_k + b] = 1 - e_k, \qquad k = 1,...N \tag{3.87}$$

The variable e_k behaves as like the slack variable ξ_k. However, the second term of $J_p(w,e)$ aims to minimize the squared errors, unlike (3.75).

$$\mathcal{L}(w,b,e;\alpha) = J_p(w,e) - \sum_{k=1}^{N}\alpha_k(y_k[w^T x_k + b] - 1 + e_k). \tag{3.88}$$

By solving the conditions of optimality reported in (3.78), one obtains a system of equations as formulation of the dual problem

$$\left[\begin{array}{c|c} 0 & y^T \\ \hline y & \Omega + \gamma^{-1}I \end{array}\right]\left[\begin{array}{c} b \\ \hline \alpha \end{array}\right] = \left[\begin{array}{c} 0 \\ \hline 1_v \end{array}\right], \tag{3.89}$$

where $\alpha = [\ \alpha_1\ \ \alpha_2\ \ ...\ \ \alpha_N\ ]^T$ are the Langrange multipliers, $y = [\ y_1\ \ y_2\ \ ...\ \ y_N\ ]^T$, $1_v = [\ 1\ \ 1\ \ ...\ \ 1\ ]^T$ and $\Omega = \{\Omega_{kl}\} = K(x_k, x_l)$ for $k,l = 1,...N$. The advantage of this formulation is that the α solution is

obtained by solving the linear regression in (3.89). This simplification comes to the price of a lack of sparsity, since the Lagrange multipliers are proportional to the error, i.e. $\alpha_k = \gamma e_k$. Consequently, all data points contribute to definition of the hyperplane, which means that all instances are *support vectors*. Based on the minimization of the squared errors in (3.86), this classifier has been named *least-squares support vector machines* (LSSVMs).

3.4.3 Linear discriminant analysis

Similarly to LSSVMs, the linear discriminant analysis (LDA) aims to minimize the squared error of misclassification, while maximizing the margin between the two groups C^+ C^- [243]. Based on the following data projections:

$$f(x) = w^T x + b, \tag{3.90}$$

the aim is to find that set of weights w that maximize the *Rayleigh quotient*, defined as follows:

$$\max_{w,b} J_{FD}(w,b) = \frac{w^T \Sigma_{\mathcal{B}} w}{w^T \Sigma_{\mathcal{W}} w}, \tag{3.91}$$

In layman terms, $J_{FD}(w,b)$ represents the ratio between the *between class* variance and *within class* variance. The matrices $\Sigma_{\mathcal{B}}$ and $\Sigma_{\mathcal{W}}$ are covariance matrices defined as follows:

$$\Sigma_{\mathcal{B}} = [\mu^{(+)} - \mu^{(-)}][\mu^{(+)} - \mu^{(-)}]^T, \tag{3.92}$$

$$\begin{aligned} \Sigma_{\mathcal{W}} &= \Sigma_{\mathcal{W}_1} + \Sigma_{\mathcal{W}_2} = \\ &= \mathbb{E}([x - \mu^{(+)}][x - \mu^{(+)}]^T) + \mathbb{E}([x - \mu^{(-)}][x - \mu^{(-)}]^T) \end{aligned} \tag{3.93}$$

where $\mu^{(+)}$ and $\mu^{(-)}$ represent respectively the mean of C^+ and C^-. By taking the $\partial J_{FD}(w)/\partial w$, the optimality condition leads to the generalized eigenvalue problem:

$$\Sigma_{\mathcal{W}} w = \Sigma_{\mathcal{W}} \frac{w^T \Sigma_{\mathcal{W}} w}{w^T \Sigma_{\mathcal{B}} w}, \tag{3.94}$$

and, by using (3.92), one obtains that:

$$w_{FD} = \Sigma_{\mathcal{W}}^{-1}[\mu^{(+)} - \mu^{(-)}]. \tag{3.95}$$

The formulation in (3.95) shows the optimal projection hyperplane is to maximize distance between the means of positive and the negative class and, therefore, the margin. However, the variability of the data of each group should be taken into account in order to minimize the error terms in the classification process. The interested reader can find more detail about the links between LDA and LS-SVMs in [243].

3.4.4 Fixed-size LS-SVM

In case of extremely large datasets, the computation of the LS-SVMs kernel on the training set can become computationally expensive. Therefore, a low-rank approximation of the kernel matrix in (3.89) can be performed by the Nyström method with a selection of M columns. In case of LS-SVMs, this approximation consists in a selection of support vector machines, which ultimately coincides with a subsampling of the training set to a size M. Since the selection of M is fixed or decided a-priori, this dataset reduction approach is known as *fixed-size LS-SVMs*, as discussed in [60]. The selection of the data points or support vectors can be random or an active selection scheme can be applied. Namely, one can maximize the quadratic Renyi entropy (QRE) in a iterative fashion. As first step, a training set is subsampled from the original dataset with a reduced size M ($M \ll N$), where N is the number of instances in the original dataset. For each iteration, one data point of the subset can be replaced with a data point coming from the original dataset and the QRE can be computed as:

$$H_R = -\log \int p^2(x)dx, \tag{3.96}$$

where the $p(x)$ is the probability density function (PDF) of the the considered subset. The argument of the logarithm is normally estimated as

$$\int p^2(x)dx = \frac{1}{N^2} 1_N^T \Omega 1_N^T, \tag{3.97}$$

where the Ω represents the kernel matrix of the fixed size training set. Therefore, the inherent advantage is that the Ω based on the subset has dimension $M \times M$, which is a much easier matrix to invert in the solution of the dual problem

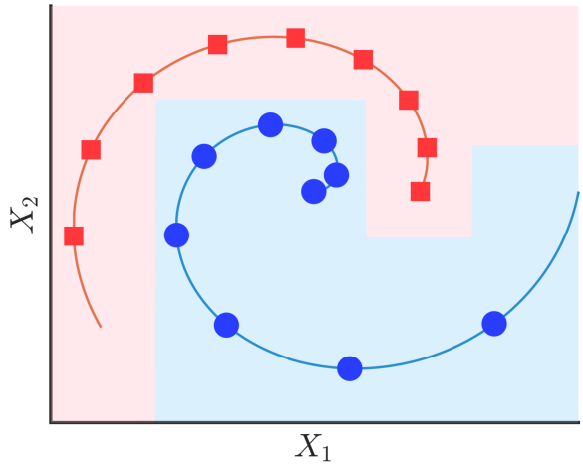


Figure 3.12: Example of application of a fixed-size LS-SVM. The toy-example shows that the classification between red squares and the blue circles could be solved by using all the data points in the two spirals. However, one can select fewer data points on two spirals (highlighted by the bigger markers) and would still be able to perform a classification. The M markers is normally selected with (3.96) to resemble the original dataset.

(3.89). The QRE is a common metric to describe the variety of selected data points in the subset. The entropy maximization leads to a selection of data points that is as representative as possible of the original dataset. The iterative scheme is normally characterized by a steady increase of QRE until reaches a plateau, when the procedure is normally stopped. This procedure is normally meant to define the LS-SVMs hyperplane with those data points that fully characterize the manifold without large computational costs. More details can be found in [60],[243]. An illustrative examples of the fixed-size LS-SVMs is reported in Figure 3.12. The toy-example shows that the classification between red squares and the blue circles could be solved by using all the data points in the two spirals (which have N data points). However, one can select a subset of M data points (3.96) to resemble the original dataset and would be still be able to perform a classification by means of LS-SVMs. The selected data points are represented by bigger markers in Figure 3.12.

3.4.5 Subspace Linear discriminant analysis

Similarly to the fixed-size LSSVMs, one might think to reduce the feature size together with training size of the dataset. In this context, the *bagging* process can be first used to randomly subsample the training set. It is normally done to train multiple classifiers with each subsample before obtaining the final classification function. If the bagging process is performed with a random subsampling of the features to find the best feature subsets to separate the data, one might refer to a *subspace process* [108]. The clear advantage is to span a greater number of features and allow the model to tune for the best subset, while different classifiers are trained similarly to the *random forest*. Since the classifier chosen is LDA for the subspace approach, we will call it subspace linear discriminant analysis in the context of this thesis.

3.4.6 Performance Metrics

In order to evaluate the performance of the developed classifiers, some metrics should be defined and a precise set of performance indices was taken as reference for studies reported in this thesis.

Confusion Matrix

The first step in evaluation of the classifier is the construction of the confusion matrix, which is comprised of True Positives (TP), True Negatives (TN), False Positives (FP) and False Negative (FN):

- TP and TN are data points correctly predicted as positive and negative by the model, as shown by the diagonal in the matrix below
- FP and FN are data points *incorrectly* predicted as positive and negative by the model, as shown by the antidiagonal in the matrix below

		Predicted	
		True	False
Actual	True	True Positive	False Negative
	False	False Positive	True Negative

Accuracy Indices

Based on the confusion matrix, one may derive the following indices:

- Accuracy is the percentage of the correctly classified data points out of the entire dataset:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FN + FP}. \tag{3.98}$$

  The complementary measure is the misclassification error, which is simply computed as

$$\mathcal{E} = 1 - \text{Accuracy}. \tag{3.99}$$

- Sensitivity is the fraction of positive samples classified by the model:

$$\text{Sensitivity} = \frac{TP}{TP + FN}. \tag{3.100}$$

- Specificity is the fraction of negative samples classified by the model:

$$\text{Specificity} = \frac{TN}{TN + FP}. \tag{3.101}$$

Cohen's kappa

Originally developed as measure of agreement among different raters, Cohen's kappa (k) is employed to measure the agreement between the predicted labels and the ground truth. This statistical measure is considered a more reliable and more robust measure, especially in case of unbalanced datasets [167]. This metric is normally derived as:

$$k = \frac{p_o - p_e}{1 - p_e}, \tag{3.102}$$

where p_o is the observed agreement, which is the Accuracy in (3.98), and p_e is the expected agreement. The latter is normally computed as the sum of the probabilities of occurrence of the positive class and the negative class in the dataset, which means:

$$p_e = \frac{TN * FN}{N}\frac{TN * FP}{N} + \frac{TP * FP}{N}\frac{TP * FN}{N}. \tag{3.103}$$

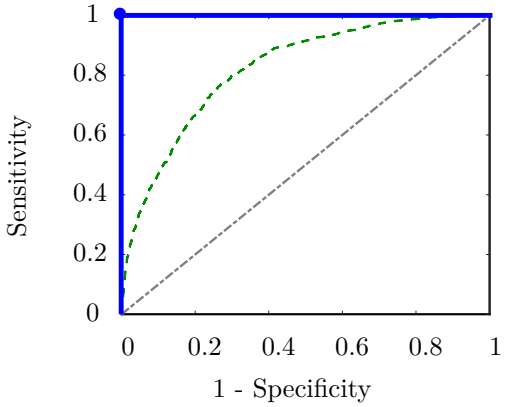


Figure 3.13: An example of a Receiver Operating Characteristic curve (green dash curve). The grey dot-dashed line is an example of a random classifier and the blue square curve represents an example of perfect classifier.

The k metric is a value bounded between -1 and 1, where $k = 1$ represents a perfect agreement and $k = 0$ is considered agreement by chance. The value -1 is an agreement worse than the agreement expected by chance. Normally, Cohen's kappa in the range $0 < k \leq 0.3$ is a weak agreement, while the interval $0.3 < k \leq 0.6$ is a moderate association and $k > 0.6$ is considered a strong agreement [167].

Receiver operating characteristic

The classification metrics strictly depends on the decision threshold present in each classifier (e.g. the bias b in the SVMs). Therefore, one might think to move this threshold to assess how the performance indices vary accordingly. The Receiver operating characteristic (ROC) curve plots the Sensitivity in function of the false positive rate (1 - Specificity), as reported by the green dashed curve in Figure 3.13. Normally, the area under the ROC curve (AUC) is considered a further measure of classification performance, since an AUC closer to 0.5 indicates a random classifier (grey dot-dash line in Figure 3.13) and an AUC closer to 1 indicates a perfect classifier (black square in Figure 3.13).

3.5 Summary

The current chapter discussed in depth an extensive background of all the different methods and extracted features to describe the dysmature EEG and

HRV and assess their evolution or association with perinatal stress. Spectral features were presented, taking into account the nonstationary nature of the EEG and HRV signal. In order to assess the persistence of slow-waves and discontinuity of EEG, entropy measure and multifractal features were reported. Furthermore, a detailed view of the multivariate analysis was described both for the assessment of signals interaction and for the description of the graph topology of the time-series network. In addition, regression and classification approaches were reported together with the associated performance metrics.

Part II

Maturation of the premature infant

Chapter 4

A brain-age model for preterm infants based on EEG connectivity

This chapter has been published as the two following articles:

1. *Lavanga M., De Wel O., Caicedo A., Jansen K., Dereymaeker A., Naulaers G., Van Huffel S. (2018). "A brain-age model for preterm infants based on functional connectivity." Physiological Measurement, 39(4), 044006;*

2. *Lavanga M., De Wel O., Caicedo A., Jansen K., Dereymaeker A., Naulaers G., Van Huffel S. (2017). "Monitoring Effective Connectivity in the Preterm Brain: A Graph Approach to Study Maturation." Complexity, 2017(ID 9078541), 1–13.*

Lavanga M. has developed the methodology, conducted the experiments and has written the manuscript. Compared to the publications, minor textual and notational changes have been implemented for better integration in this thesis.

The neonatal EEG is expected to develop with infant's maturation. Specific morphological traits such as discontinuity and slow-wave patterns are expected to visually disappear. However, other features like asynchrony are also expected to change with increasing age, but the variations in EEG dependencies are difficult to monitor even for an expert clinical eye. This chapter introduces different

algorithms to monitor brain development via EEG connectivity measures. As first step, a variety of methods evaluates the functional and effective connectivity in a full-scalp EEG dataset of infants with normal development outcome. Subsequently, a set of network measures is derived to describe the topology of the EEG network at different ages of the infants. At last, those features are used to develop a Brain-Age model, which is a regression model based on physiological biomarkers to describe the development of the infant.

4.1 Introduction

Premature birth normally leads to an early exposure to the extrauterine environment, which is linked with the disruption of the normal brain development. Due to the possible impact of prematurity on neurodevelopmental outcome, a need for maturation charts to detect neural disorders in advance has been discussed in the literature. In [89], Franke et al. used fMRI to describe cerebral maturation, while other studies investigate the changes in the neonatal EEG to describe infants' development [191],[192].

The brain can be seen as a complex network of interacting regions and hierarchical communications, which are constrained by the anatomy, but not limited to it [94]. The neuronal clusters can actually work together and communicate to perform a joint task beyond their structural locations. The clinical literature [141] distinguishes this type of connectivity from the anatomical one, which is often called structural. The consequence of this functional infrastructure is the generation of complex electrophysiological patterns, which are temporally correlated, by distant cerebral areas [260]. Those spatiotemporal patterns are dynamic and they change according to the individual development trajectory [213].

The last trimester of gestation is a period of brain development, which includes both anatomical rewiring [26] and electrophysiological modifications [8]. Different authors illustrated that the cortical regions undergo differentiation, folding and gyrification with the premature infant's development, while the subcortical areas experience synaptogenesis and myelination as well as neural pruning to establish thalamo-cortical connections or long distance cortical connections [245], [72]. Based on MRI scans of preterm babies, Dubois et al. [72] showed that the white matter volume and the inner cortical surface increases with gestational age. Furthermore, the same authors [72], [113] demonstrated that fractional anisotropy (FA) of the brain fiber bundles increases with post-birth maturation, although the different connection pathways seem to develop in an asynchronous way. According to Batalle et al. [26], FA is a measure of

anatomical directivity that describes the connectivity strength among brain regions together with the density and the percentage of connecting streamlines. Besides thalamo-cortical connections and changes in anatomical connectivity, the development of the infant brain during the perinatal period is characterized by an increased relevance of cortico-cortical connections [134]. Based on this anatomical findings, Huppi et al. [113] argued that diffusion tensor imaging parameters, such as anisotropy, can be structural markers of network functional organization [26]. Specifically, the slow wave burst-modulation pattern is due to the cortical interactions, while subcortical patterns generate high-frequency oscillations [270]. The latter type of connections are dominant up to 32 weeks PMA, when the former starts prevailing [134]. Consequently, the anatomical development drives a change in neuroelectrical waveforms and the temporal correlations among brain regions is also expected to change since the functional connectivity is related to the structural topology [94].

However, less is known about the correlations among neuroelectrical waveforms and their maturation [241]. As discussed in details in Section 3.2 and in [90], three types of brain connectivity are defined. The first one consists of the description of the anatomical wiring, known as structural connectivity, and usually described by imaging techniques. The second type is the functional connectivity (FC), which delves into the statistical relationships among the neural activities of different brain regions and provides insights into how the brain manages its neuronal resources to engage any type of tasks. The third one is called effective connectivity (EC) and is a special case of FC since it delves into the directionality of those functional dependency and therefore the causality of the coupling. Functional connectivity in developmental science received increasing attention in the recent years [94]. Moreover, different tools to describe functional connectivity became available in the last two decades, such as graph theory (Section 3.2.4). Brain networks can be represented as a graph, where the nodes are brain regions and the edges are the connection strengths. Moreover, functional connectivity methods generate undirected graphs since the statistical dependencies are represented without any directionality. On the contrary, the effective connectivity methods are associated to directed graphs, since they highlights which area or node causes the other (Section 3.2.4). In different studies of the preterm infant brain [26], [185], the neural activity has been measured using fMRI. Although this neuroimaging technique can investigate the subcortical structures, it is an expensive method and it is not suitable to measure effective connectivity due to fMRI low temporal resolution [229]. In contrast, EEG is a suitable measurement for effective and functional connectivity and has been employed in recent papers to study the brain connectivity maturation [229], [146].

The neonatal EEG connectivity studies have three main drawbacks: the

limited investigated maturation period (especially for EC), the absence of graph theory metrics to investigate the connectivity changes and the investigation of connectivity at sensors' level. Studies about the long term development of effective connectivity based on EEG in neonates were rarely performed [229]. In particular, there is a lack of research, that investigates the effective and functional connectivity from birth until full-term age in premature infants. Secondly, in case of perinatal development, the common approach is to describe how the coupling of a selected pair of channels changes over age ([95],[168], [177]). However, the combination of channels to study their interactions increases exponentially with the number of channels. Unlike the application with fMRI [26], graph indices, which describe the topological distribution of the network, were seldom applied in premature EEG connectivity [185]. The third drawback is the investigation of the connectivity at the sensors' level, which can generate spurious couplings due to volume conduction [104]. Some authors suggested to retrieve sources to investigate the coupling. And yet, many doubts can be raised on this approach for two reasons. The first one is the requirement of a proper head-model for neonates to estimate the sources, which can turn into an expensive and unneeded process, as already mentioned in previous studies [168]. The second reason is that the sources of spontaneous EEG are close to the surface of the scalp [99], therefore to the EEG recording sites, which questions the necessity to retrieve the sources to avoid the computation of spurious connectivity. Based on these considerations, FC can be described via different pairwise EEG connectivity methods, such as Nolte's imaginary part of the coherence ($ImCoh$), which studies the lagged interactions among signals [182]. Haufe [104] reports that $ImCoh$ highlights the strong and true interactions and disregards the weak interactions due to the volume conduction (see Section 3.2.1). The $ImCoh$ is not new to perinatal development studies, such as [95]. However, previous research has not dealt with this method via growth charts as [177], [168] or [132] proposed.

The main aim of the two studies reported in this chapter was to provide a systematic comparison of different EEG-based functional and effective connectivity approaches in order to track the maturation of the infant brain, exploiting different network metrics (as in [26]). In particular, effective connectivity was investigated by means of transfer entropy [227] and Granger Causality [97]. Concerning functional connectivity, Imaginary Coherence ($ImCoh$), Mean squared Coherence (MSC), Phase locking value and the *Hilbert - Schmidt dependence* (HSD) [244] were employed in the investigation of the healthy preterm cerebral networks. On the obtained directed and undirected graphs, different networks metrics were computed to track their evolution from 27 to 42 post-menstrual age (PMA) weeks. This network analysis was complemented by a regression to predict the age of the patient. The ultimate goal was to provide a brain-age prediction model based on graph topological

indices, similarly to [89], [191].

4.2 Methods

4.2.1 Data

Different datasets were used. A dataset comprised of EEG signals was the main part of the maturation study. In addition, a simulated dataset was also employed in order to show how effective connectivity analysis performs in a controlled case. In particular, the main objective of the simulation was to illustrate the meaning of the most common graph indices used in the literature in relation with the employed time series models.

Simulated data

The first experiment of this investigation consisted of a simulation based on a linear Gaussian regression model (derived from [173]), expressed by the following equations

$$\begin{aligned}
x_{1,t} &= 0.9x_{1,t-1} + 0.9\beta x_{2,t-1} + 0.7\beta x_{3,t-1} + \xi_{1,t} \\
x_{2,t} &= -0.9\beta x_{1,t-1} + 0.7\beta x_{3,t-1} + \xi_{2,t} \\
x_{3,t} &= 0.8\beta x_{1,t-1} + 0.7\beta x_{2,t-1} + \xi_{3,t} \\
x_{4,t} &= -0.25\beta x_{1,t-1} - 0.6x_{4,t-1} + \xi_{4,t} \\
x_{5,t} &= 0.25\beta x_{1,t-1} + 0.9\beta x_{4,t-1} + \xi_{5,t} \\
x_{6,t} &= 0.9\beta x_{4,t-1} + 0.9x_{5,t-1} - 0.7\beta x_{6,t-1} + \xi_{6,t}
\end{aligned} \tag{4.1}$$

The parameter β is a scalar value that varies between 0 and 1 and it influences the strength of the coupling among the $x_{i,t}$ variables. The $\xi_{i,t}$ variables represent white noise with unit variance. The associated graph is reported in Fig.4.1. In order to give the reader a clear insight of the graph theory measures, the objective of the simulation was twofold: firstly, we investigated the influence of the strength of the coupling, provided by the parameter β, and show the effect of the weakening of causality among the variables. Secondly, we investigated the influence of the noise using different levels of signal to noise ratio (SNR), which were obtained varying the variance of white noise in (4.1). In the latter

objective, time series were simulated using the AR model (4.1) and the coupling was estimated with the effective connectivity methods discussed in the following paragraphs. We also studied the impact of filtering on the graph based metrics by computing the scores from filtered signals, using a sampling frequency which was set at 200 Hz and a band-pass filter between 1 and 80 Hz was applied on the simulated data. The reason to investigate the impact of filtering is explained in Section 4.2.3. The band-pass frequencies were obtained from [87].

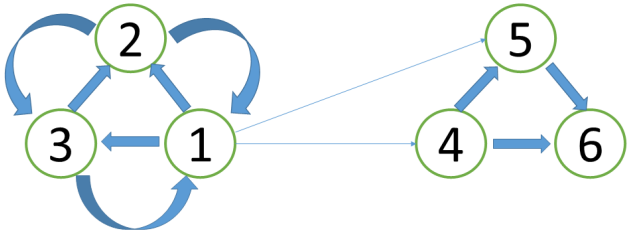


Figure 4.1: The figure displays the graph associated to the model 4.1.

4.2.2 EEG Data

The second dataset comprised of 30 preterm neonates, who were recruited for a larger study to assess brain development [144],[132] between 2012 and 2014 at the same neonatal intensive care unit (NICU), in the University Hospitals Leuven, Belgium, with informed parental consent (Section 1.2). A group of 26 neonates presents normal neurodevelopment outcome at 2 years from the birth, while 4 subjects are declared normal at 9 months (Bayley Scales of Infant Development - II, mental and motor score > 85). EEG measurements for each subject took place at least once during their stay at the unit and lasted at least 2 hours. The data were recorded with 9 electrodes (F_1,F_2,C_3,C_4,T_3,T_4,O_1,O_2 and the reference C_z) according to the 10-20 international system monopolar set-up. The sampling frequency for the data collection was 250 Hz (BRAIN RT, OSG equipment, Mechelen, Belgium). A total number of 103 recordings was obtained from the patients, whose *post-menstrual age* (PMA) ranges from 27 weeks to 42 weeks. The two sleep stages, namely quiet sleep (QS) and non-quiet sleep (NQS, all other states), were annotated by two independent readers. Each recording has at least one QS-NQS cycle. The average (standard deviation) duration for the EEG recording, for the QS and NQS are respectively 263.69(129.39) min, 73.58(33.01) min and 190.11(106.90) min, while the amount of multichannel EEG, QS and NQS data are respectively 452 h, 326 h, 126 h. Sleep states QS and NQS are respectively 29.46(9.83)% and 70.54(9.83)% of the trace.

The FC connectivity analysis focused on the entire EEG dataset, while the EC analysis has originally been performed in a subset of patients. Specifically, the directionality among signals was investigated on 25 preterm infants, who presented good neurodevelopment outcome at 2 years from the birth. This dataset resulted in 88 recordings ranging from postmenstrual age of 27 weeks to 42 weeks.

4.2.3 Connectivity analysis

Functional connectivity methods

In order to assess the functional connectivity (FC), the most common linear and nonlinear approaches to derive bidirectional dependencies, as reported in details in section 3.2.1. The first two measures were *magnitude squared coherence* (MSC) and imaginary part of the coherence ($ImCoh$). Similarly, the nonlinear interrelationships and synchrony was assessed via the *phase locking value* (PLV) and the *Hilbert - Schimdt dependence* (HSD). In this multivariate analysis, the HSD analysis was performed by means of the ITE toolbox, Matlab (Mathworks, Natick, Ma, USA) [244].

Effective connectivity methods

As anticipated in section 3.2.2, the literature about effective connectivity presents different methods to assess coupling among time series, such as *directed transfer function* (DTF), *partial directed coherence* or *Granger causality test* (GCT) [35]. These approaches generate a type of coupling known as *G-causality* [14]. In case of multiple time series, the conditional multivariate Granger causality (cMVGC) is normally estimated instead of the bivariate GCT [18]. According to [14] and [224], both the DTF and cMVGC are based on the same multivariate or Vector AutoRegressive (VAR) modeling. However, they can show different aspects in the causality analysis. While the DTF estimates the *reachability* from one channel to another (defined also as *G-influentiability* by [14]), both PDC and cMVGC look into the direct active link between two channels, which is defined as *G-connectivity* [14]. Those two methods investigate the same connectivity model in two different domains, which are the frequency (PDC) and time (cMVGC) domain. In this chapter, the results by cMVGC in its AutoRegressive (AR) and information dynamics implementations, as described in section 3.2.2, are reported.

Graph theory indices

As illustrated in section 3.2.4, connectivity method generates an adjacency matrix $M \times M$, defined as follows

$$\mathcal{A}_{ij} = \mathcal{C}_{x_i \leftrightarrow x_j}, \tag{4.2}$$

where $\mathcal{C}_{x_i \leftrightarrow x_j}$ is the general coupling intensity and M is the number of processes. Functional connectivity methods generate a symmetric adjacency matrix A, which describes a weighted undirected graph [43]. In contrast, effective connectivity methods generate an *asymmetric* adjacency matrix A, which describes a weighted directed graph. Those graphs are normally treated as complete or full without applying a specific threshold (unlike binary networks). Since the maximal number of couplings is $M^2 - M$ and, therefore, increases exponentially, graph theory measures can be used to summarize brain connectivity [234]. Although there is no minimal theoretical number of graph nodes, 8 EEG channels can be considered quite limited for a graph analysis. However, there are studies in the neural processing literature where graph theory was applied on a limited number of time series in order to give a concise view of a high-density (or complete) network [234], [83], in particular if the sources-level is concerned.

A set of topological and spectral indices can be computed for an EC-directed and FC-undirected networks (see Section 3.2.4). Among the topological indices, the average characteristic path length, the clustering coefficient, the diameter and the causal density were considered. Among the the spectral indices, the spectral radius, the spectral gap and the algebraic connectivity were computed. A quick overview of the considered indices is newly reported in Table 3.1.

4.2.4 Algorithmic pipeline and statistical analysis

Effective connectivity processing

According to different authors [19], [87], filtering could add spurious connectivity in the effective connectivity analysis or make the estimation of the underlying Granger causality VAR model unstable. Although a theoretical invariance of causality estimation has been demonstrated under filtering, the G-causality works in practice if, and only if, the data are stationary. The use of filtering as mean to reach stationarity with filtering has already been investigated by [19]. However, three main reasons can undermine this approach. The first reason is the increase of the estimated VAR model order due to the fitting of *filtered*

Table 4.1: An overview of all graph metrics that have been used in the multivariate analysis of this chapter.

Overview of graph indices	
Path length	Mean of the nodes' shortest paths
Clustering coefficient	Mean of the nodes' triangles intensity around each node
Diameter	Maximum graph eccentricity
Causal density	Sum of all significant couplings in A
Spectral radius	$\lambda_1(A)$: first eigenvalue of adjacency matrix A
Spectral gap	$\lambda_1(A) - \lambda_2(A)$: the difference between the first two eigenvalues of matrix A
Algebraic connectivity	$\lambda_{M-1}(L_{sym})$ or $\lambda_{M-1}(L)$: the second smallest eigenvalue of the Laplacian matrix L_{sym} or L

process of theoretical *infinite* order with a numerical *finite* order. This could lead to a poor, and not robust, parameter estimation or even unstable VAR model, which is the second reason to avoid band-pass filtering. The last reason is the ill-conditioning of the Toeplitz matrix of the autocorrelation sequence $\Gamma_k = cov(\mathbf{U}_t, \mathbf{U}_{t-k})$, which is necessary for VAR model estimation. All the related theoretical details are explained in [19]. The practical downsides of the band-pass filtering are the dramatic increase of the VAR model order compared to the unprocessed data or the increase of false positive detection of connectivity links with both the GC [19] or the PDC [87]. In particular, the amount of false detections increases with narrowing of the filtering frequency band or the increase of the filter order. Interestingly, there is no distinction between FIR or IIR filtering for both the authors. However, in both studies [19] and [87], it is pointed out notch filtering and differentation (or high pass filtering at 1 Hz) as methods to keep the VAR model order low and reduce the false detection, even compared to unprocessed data. Those approaches help to achieve stationarity or keep the VAR model order low for nearly-nonstationary processes like EEG. In addition, the presence of trends or seasonality can add unit-roots to the time series (poles on the unit circle or outside it in the complex plane), which violates the hypothesis of covariance stationarity. In the presence of unit-roots, the impulse response of the VAR model would be oscillatory or diverge to infinity. Consequently, it is suggested to eliminate trends and seasonality by first order differentiation or differentiation at various lags (notch-filtering). Theoretical and numerical details of the different type of filtering are reported in [19] and [18]. Given the stated literature results, on the one hand, we decided to investigate the impact of filtering on the simulated data and, on the other hand, we applied

only notch filtering at 50 Hz and 100 Hz and differentiation on the EEG data in order to reduce nonstationarities in the time series. However, the EEG can be affected by muscle artifacts, which can spread out among the different electrodes and bias the connectivity analysis. To mitigate this effect, we applied canonical correlation analysis (CCA) to remove the artefacts caused by the change in the EMG signals [269]. To investigate the effect of the CCA on the analysis, we compared the output of the connectivity analysis in two scenarios: the first one did not consider the usage of CCA; the second one applies CCA and reconstructed the EEG removing the 3 sources with lowest autocorrelation. In the first case, the authors segmented the EEG in 5 s windows and computed the effective coupling with both listed methods. In the second scenario, CCA was applied on 5 s EEG segments. However, before performing any connectivity analysis, we recombined the segments in 30s intervals in order to increase the rank of the reconstructed EEG matrix. This step was required since both GC and TE request data that do not present collinearity (i.e. the time series matrix cannot be rank-deficient, [18]). The window length was suggested as a further step to keep time series stationarity [18]. For each recording, we computed the adjacency connectivity matrix for EEG segments and we averaged over the QS epochs and NQS epochs. This averaging step did not consider coupling values that were not statistically significant. According to [173] and [18], the GC and TE follow an asymptotic F-distribution, like R^2 statistics. Therefore, an F-test was implemented to test the significance of coupling among EEG channels and all the couplings with p-value $p < 0.05$ were set to zero (see Section 3.2.3).

Functional connectivity processing

In case of functional connectivity, each EEG signal was resampled to 128 Hz and low-pass filtered at 32 Hz. Each FC method was computed in 1 min non-overlapping windows. MSC and $ImCoh$ were computed with Welch's method using a 1 sec subwindow and the overlap was 70%. Those 2 parameters were mainly chosen for a consistent and unbiased coherency estimation [39]. The MSC was then extracted as the average of $|C_{xy}(f)|^2$ in the following frequency bands: δ_1 [0-2] Hz, δ_2 [2-4] Hz, θ [4-8] Hz, α [8-16] Hz, β [16-32] Hz. The $ImCoh$ was computed as the maximum of $\mathcal{I}(C_{xy}(f))$ in the same frequency bands. This choice is a consequence of the fact that the $ImCoh$ is expected to be centered on one single frequency, as reported in [182]. In case of the HSD and PLV, the signal was split in the specified frequency bands via wavelet decomposition. The chosen basis was Daubechies 4, which is known to be the best wavelet for epileptic EEG, but also for neonatal EEG since it is characterized by transient waveforms. The coefficients of each scale were used to compute the phase synchrony and the information entropy. In both types of segmentation, the mean was subtracted and the window was discarded wherever

the maximum value was above 3000 μV. It is interesting to notice how shorter windows for HSD and PLV could have affected the significance of the coupling estimation. In case of PLV, enough basic periods are required to assess the oscillator interactions and the lowest frequency oscillations are just reported by the empirical threshold defined by [31]. MSC, PLV and MI generate symmetric adjacency matrices, while the latter one is antisymmetric in case of $ImCoh$. In order to apply the undirected graph theory, we simply decided to ignore the sign of the entries and take $\mathcal{A}_{ij} = |A_{ImCoh_{ij}}|$. Consequently, each method generates a 8×8 matrix for each method and each window. In order to guarantee the reliability of each connectivity method, we applied a surrogate data test (see Section 3.2.3). Specifically, the data were shuffled via the adjusted-amplitude Fourier transform algorithm (AAFT) in each epoch, which rescales the signals with a Gaussian process and randomizes the phase [228]. Subsequently, we did not only compute the coupling between two generic channels X and Y, but also the coupling between X and 19 surrogates of Y in each epoch. As suggested by Pereda et al. [200], the coupling was significant if and only if it was higher than the surrogate-based 19 interaction intensities. The number 19 guarantees an $\alpha = 0.05$ significance level since $N_{surr} = \frac{1}{\alpha} - 1 = 19$. Each coupling value which did not pass the surrogate test was set to zero. The AAFT surrogates were computed with the MATS toolbox [137]. Since the label of sleep stage was known, the different FC adjacency matrices were averaged throughout the QS and NQS periods, such that two adjacency matrices $\mathbf{A}$ were obtained per recording.

Regression analysis: effective connectivity

After the F-test, the EC pipeline obtained a total of $352 = 88 * 2 * 2$ coupling directed graphs, where 88 is the total number of recordings, 2 is the employed causality methods, 2 is the considered number of sleep states. On the average matrices we computed the network indices described above, which results in a tensor $88 \times 7 \times 4$, where 7 is the number of graph features and 4 is the number of combinations considering the number of connectivity methods and sleep states involved (TE in QS, TE in NQS, GC in QS, GC in NQS). For each feature, we evaluated the maturation trend in three distinctive age groups (≤ 31, $\in (31-37)$, ≥ 37 PMA weeks) as median(IQR), where IQR is *InterQuartile Range*. Besides, we computed the Pearson correlation coefficient between the variable age and each single feature and its statistical significance were computed. Those results were meant to give a general overview of the feature maturation trend for each connectivity method, for each sleep state, with or without CCA preprocessing. In addition, we computed an ordinary least squares (OLS) regression for each single feature in case of GC during QS and the associated confidence interval at 95%, in order to give a visual representation of any network index prediction

power. In particular, we split randomly the single graph index dataset 100 times in 70% training set and 30% testing set and we computed the prediction error on the test set as *root mean squared error*, $\sqrt{MSE}$, as shown by [199]. Furthermore, each slice of the tensor (a matrix 88×7) was used to predict the patient's age at the moment of the recording with a multivariate linear regression. Similarly to the single feature approach, we randomly split the dataset 100 times, as in [199] and we assessed $\sqrt{MSE}$ in the test set. At each iteration, the R^2 index and the F-test statistics were computed.

Regression analysis: functional connectivity

After surrogation, the next step in the FC pipeline was the computation of the topological and spectral indices, whose output was arranged as a $103 \times 12 \times 5$ tensor for each method and sleep state. The number of recordings was 103. We obtained 12 network indices and we used 5 frequency bands. Due to the large number of features, we decided to report only the path length as median(IQR), where IQR is *inter-quartile range*, in three PMA groups ($\leq 31\ Weeks$, $\in (31-37)$ $Weeks$, $\geq 37\ Weeks$) to give a general overview of the predictive power for each method. Furthermore, the Pearson correlation coefficient between the path length and PMA on the entire dataset has been reported. The topological features from each FC method were used to develop a linear mixed-effect (LME) model to predict the PMA in each sleep state. In the random-effects modeling, the subject or patient ID label was considered as a grouping variable. Since the topology features can be characterized by collinearity due to the similar definitions of graph indices, we decided to apply a least absolute shrinkage operator (LASSO) to select the features for the regression model (Section 3.3.3). In particular, a LASSO algorithm was run 100 times on each dataset and only features selected more than 40% of times were included in the final model. The LME model with LASSO-filtered features was then estimated on 100 random splits of the dataset. In each split, the model was tuned on 70% of data, the training set, and tested on the remaining 30%, the test set. The model performance has been assessed via the *mean absolute error* (MAE) on the test set and the *adjusted coefficient of determination* R^2_{adj} on the training set. The last step was to develop a model independently from the type of FC method or frequency band. Specifically, a subset of features with the highest predictive power to track PMA have been selected in this way: each attribute was used to predict age in a univariate LME fashion and ranked based on MAE. The 10 features with the lowest MAE were used to develop a final age regression model. In this phase, the FC coupling intensities together with the network indices were also used as candidate features. Eventually, the final model with 10 features was tested on 100 random splits of the dataset and evaluated using R^2_{adj} and MAE.

4.3 Results

In the following paragraphs, the results obtained in both the simulated and the real dataset are discussed. In particular, we will show how the network indices behave in both examples, the simulated data and the EEG maturation dataset. The last part summarizes the predictive power of those graph metrics to infer the *post-menstrual age* of the subject.

4.3.1 Simulated data

Figure 4.2, Figure 4.2.b and Figure 4.2.c display the integration graph indices for the AR model defined in (4.1), whose network is depicted in Figure 4.1. In the original model ($\beta = 1$), the graph presents two distinct clusters, which are loosely connected by the edges between node 1 and nodes 4 and 5. This two-hubs structure is reflected by the first two panels in Figure 4.2. When the intra-cluster connectivity is high ($\beta = 1$), the clustering coefficient reaches its highest level, while the path length is at its lowest level. Those results are in line with a rich-club or small world network [43]. However, when the coefficient β starts decreasing, the clustering coefficient proportionally decreases, while the path length increases. The spectral radius decreases with vanishing values of β, as the clustering coefficient does. Figure 4.2.d, Figure 4.2.e and Figure 4.2.f display the ratio between the network indices estimated from the simulated time series via Transfer Entropy and the original measures (in particular, in the first panel, $\frac{CC_{TE}}{CC_{orig}}$, in the second one, $\frac{length_{orig}}{length_{TE}}$, in the third one, $\frac{\lambda_{TE}}{\lambda_{orig}}$). Different noise levels have been used. The results are quite straightforward for the clustering coefficient and the spectral radius: the higher the signal-to-noise ratio, the higher the values of two indices are and the closer they are to the original values (Figure 4.2.d and Figure 4.2.f). In the case of the path length, the absolute value is decreasing with higher SNR. However, the estimated value becomes similar to the original one for very low noise variance (Figure 4.2.e). Figure 4.2.d, Figure 4.2.e and Figure 4.2.f show also the effect on band-pass filtering on the graph indices estimation. The most remarkable results are related to the path length and the clustering coefficient. The estimated clustering coefficient tend to underestimate the original one, while the estimated path length tends to be persistently higher than the original one. On the contrary, the ratio for the spectral radius in case of filtering behaves similarly to the one obtained using the raw data.

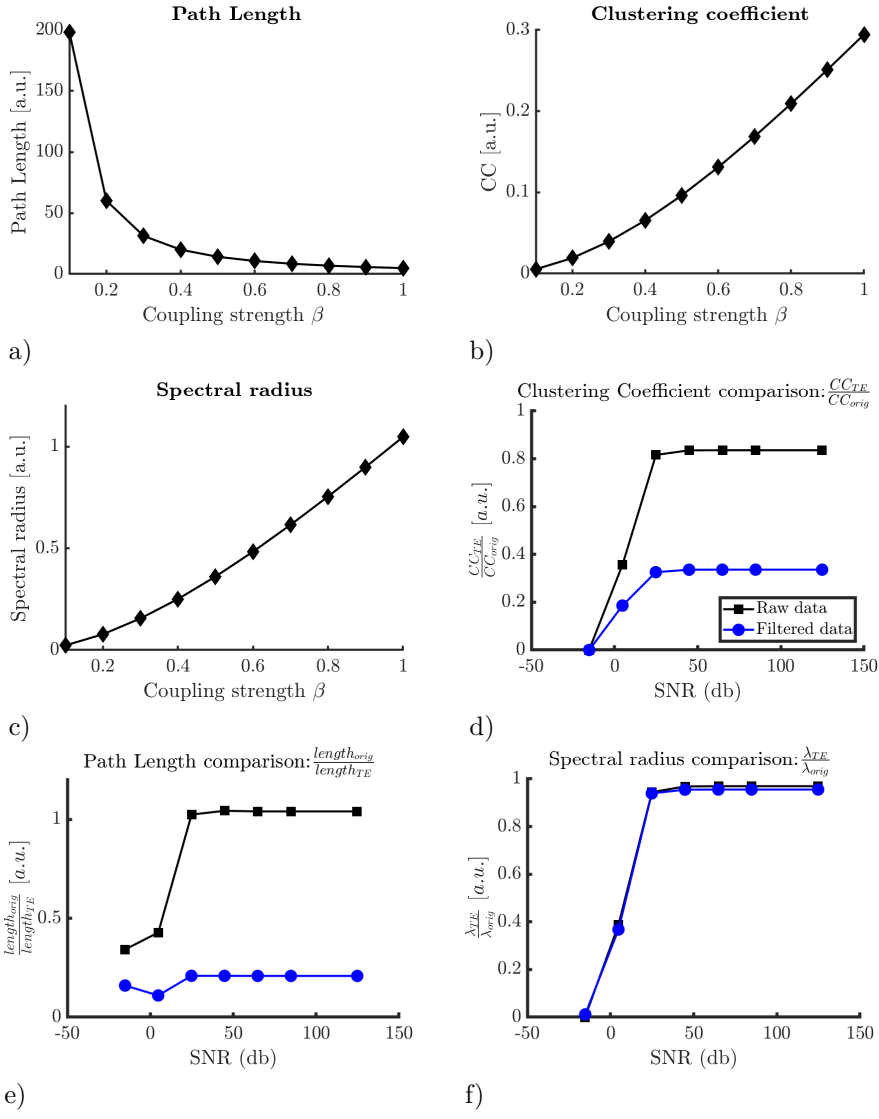


Figure 4.2: The figure shows the results for the simulation dataset. The first three panels show how the graph indices behave for different level of coupling in the model 4.1. The last three panels investigate how the transfer entropy can estimate the network indices for different level of SNR. In particular, the figure compares the two cases when the data is filtered and when the raw data is used.

4.3.2 EEG data: Effective connectivity indices

Figure 4.3 shows graph metrics for the adjacency matrices obtained using the EEG measurements and Granger causality in QS (GC in QS). Figure 4.3.a, Figure 4.3.b, Figure 4.3.c and Figure 4.3.d show the scatter plots with the fitted OLS regression model, while Figure 4.3.e and Figure 4.3.f display the clustering coefficient and path length dynamics in three distinct age groups. If Figure 4.3 gives a visual representation of the trends for the graph features, Table 4.2 and Table 4.3 provide a complete overview for all coupling methods (TE,GC) and all sleep states. Each single feature has a significant trend with age, although the Pearson correlation coefficient $\rho(\%)$ increases when CCA is used as a pre-processing step. Specifically, the trend for the clustering coefficient, the spectral radius and the spectral gap is negative, while the path length is increasing with age. This result is persistent in each method and each sleep state. The connectivity weakening for GC in QS is also reported in Figure 4.4, which shows the average connectivity graph for three distinct age groups. The three panels show how the coupling among time series decreases by the reduction in arrows width and the color shift from red to blue.

4.3.3 EEG data: Functional connectivity indices

Table 4.4 and 4.5 give a general overview of different functional connectivity methods to describe the connectivity over age by means of the path length. The different age groups in the two tables have respectively the following number of recordings: 24, 62, 17. The first remarkable aspect is a stronger correlation with PMA for $ImCoh$ compared to other FC methods, in particular for the frequency bands δ_2, θ and α, together with $k^2_{xy}(\theta, \beta)$ and $HSD\ (\theta, \alpha)$. Secondly, PLV, $HSD\ (\theta, \alpha)$ and $k^2_{xy}(\delta_1, \delta_2, \theta, \alpha)$ show a negative correlation between path length and PMA, while $k^2_{xy}(\beta)$ and $ImCoh$ present a positive trend. Since the path length is a measure of distance, a decrease of this topological index relates with an increase of connectivity, while its increase is associated to a lesser degree of interaction. For the $ImCoh(\theta)$, LME-regression chart with the associated confidence intervals and data cloud of the path length and the clustering coefficient have been reported in Figure 4.5. Each color in the plot represents a different patient. The topographic distribution of the coupling value for $ImCoh(\theta)$ is reported in Figure 4.7. Figure 4.6 shows different density indices computed via MSC in β and θ bands. In particular, panel a shows posterior density for $k^2_{xy}(\beta)$ during QS, while panel c shows the anterior density for $k^2_{xy}(\theta)$ for QS. The anterior density $k^2_{xy}(\theta)$ increases with postnatal maturation, while we found a negative trend of $k^2_{xy}(\beta)$ with postnatal maturation.

Table 4.2: The main integration and spectral features in three discrete time points. The table shows the indices for both sleep states (QS = quiet sleep, NQS = non-quiet sleep) and they were computed on the Transfer Entropy connectivity graph. The results are reported as median(IQR), where IQR stands for *InterQuartile Range*. The symbol ρ stands for the Pearson correlation coefficient, while # represents a significant correlation with $p \leq 0.01$. The values 10^{-3} or 10^{-2} mean the reported results are multiplied by a factor 10^{-3} or 10^{-2}

Network indices - Transfer Entropy in three age groups				
Median(IQR) - PMA weeks	≤ 31	$\in (31-37)$	≥ 37	$\rho(\%)$
Clustering coefficient				
QS	.025(.008)	.021(.005)	.017(.002)	-53 #
NQS	.025(.006)	.020(.006)	.018(.002)	-49 #
Path length				
QS	3.73(.30)	3.89(.22)	4.07(.10)	59 #
NQS	3.71(.20)	3.91(.25)	4.04(.14)	54 #
Spectral radius				
QS	.18(.07)	.16(.05)	.12(.01)	-49 #
NQS	.18(.05)	.15(.04)	.13(.02)	-48 #
Spectral gap				
QS	.15(.09)	.12(.04)	.10(.02)	-51 #
NQS	.15(.06)	.12(.04)	.11(.02)	-48 #
Network indices - Transfer Entropy - CCA				
Median(IQR)	≤ 31	$\in (31-37)$	≥ 37	$\rho(\%)$
Clustering coefficient				
QS(10^{-3})	9.82(6.1)	6.21(3.3)	4.98(0.9)	-57 #
NQS(10^{-3})	9.13(3.5)	6.29(2.0)	5.90(1.0)	-49 #
Path length				
QS	4.70(.57)	5.11(.47)	5.32(.18)	64 #
NQS	4.73(.40)	5.09(.27)	5.16(.18)	52 #
Spectral radius				
QS(10^{-2})	8.20(4.7)	4.68(3.1)	3.63(.6)	-55 #
NQS(10^{-2})	6.94(2.8)	4.74(1.6)	4.30(.7)	-48 #
Spectral gap				
QS(10^{-2})	5.84(3.1)	3.93(2.0)	3.30(.6)	-52 #
NQS(10^{-2})	5.84(3.1)	3.93(2.0)	3.30(.6)	-49 #

Table 4.3: The main integration and spectral features in three discrete time points. The table shows the indices for both sleep states (QS = quiet sleep, NQS = non-quiet sleep) and they were computed on the Granger Causality connectivity graph. The results are reported as median(IQR), where IQR stands for *InterQuartile Range*. The symbol ρ stands for the Pearson correlation coefficient, while # represents a significant correlation with $p \leq 0.01$. The values 10^{-3} or 10^{-2} mean the reported results are multiplied by a factor 10^{-3} or 10^{-2}.

Network indices - Granger Causality				
Median(IQR)- PMA weeks	≤ 31	$\in (31-37)$	≥ 37	$\rho(\%)$
Clustering coefficient				
QS	.024(.006)	.019(.006)	.015(.002)	-56 #
NQS	.024(.007)	.019(.005)	.016(.002)	-51 #
Path length				
QS	3.77(.20)	3.95(.25)	4.17(.13)	61 #
NQS	3.75(.33)	3.95(.27)	4.10(.14)	56 #
Spectral radius				
QS	.18(.06)	.15(.04)	.10(.02)	-54 #
NQS	.18(.05)	.14(.03)	.10(.01)	-51 #
Spectral gap				
QS	.14(.06)	.13(.04)	.09(.02)	-58 #
NQS	.15(.05)	.11(.03)	.11(.01)	-51 #
Network indices - Granger Causality - CCA				
Median(IQR)- PMA weeks	≤ 31	$\in (31-37)$	≥ 37	$\rho(\%)$
Clustering coefficient				
QS(10^{-3})	12.87(4.8)	9.34(3.2)	7.39(1.2)	-68 #
NQS(10^{-3})	12.29(3.4)	9.08(2.1)	8.51(1.2)	-61 #
Path length				
QS	4.36(.37)	4.69(.33)	4.92(.16)	73 #
NQS	4.41(.28)	4.71(.21)	4.77(.14)	63 #
Spectral radius				
QS(10^{-2})	9.52(3.5)	6.79(2.6)	5.28(.8)	-68 #
NQS(10^{-2})	8.98(2.5)	6.50(1.5)	6.05(.9)	-61 #
Spectral gap				
QS(10^{-2})	7.28(2.4)	5.87(2.3)	4.91(1.1)	-63 #
NQS(10^{-2})	6.91(3.0)	6.07(1.0)	5.35(1.4)	-56 #

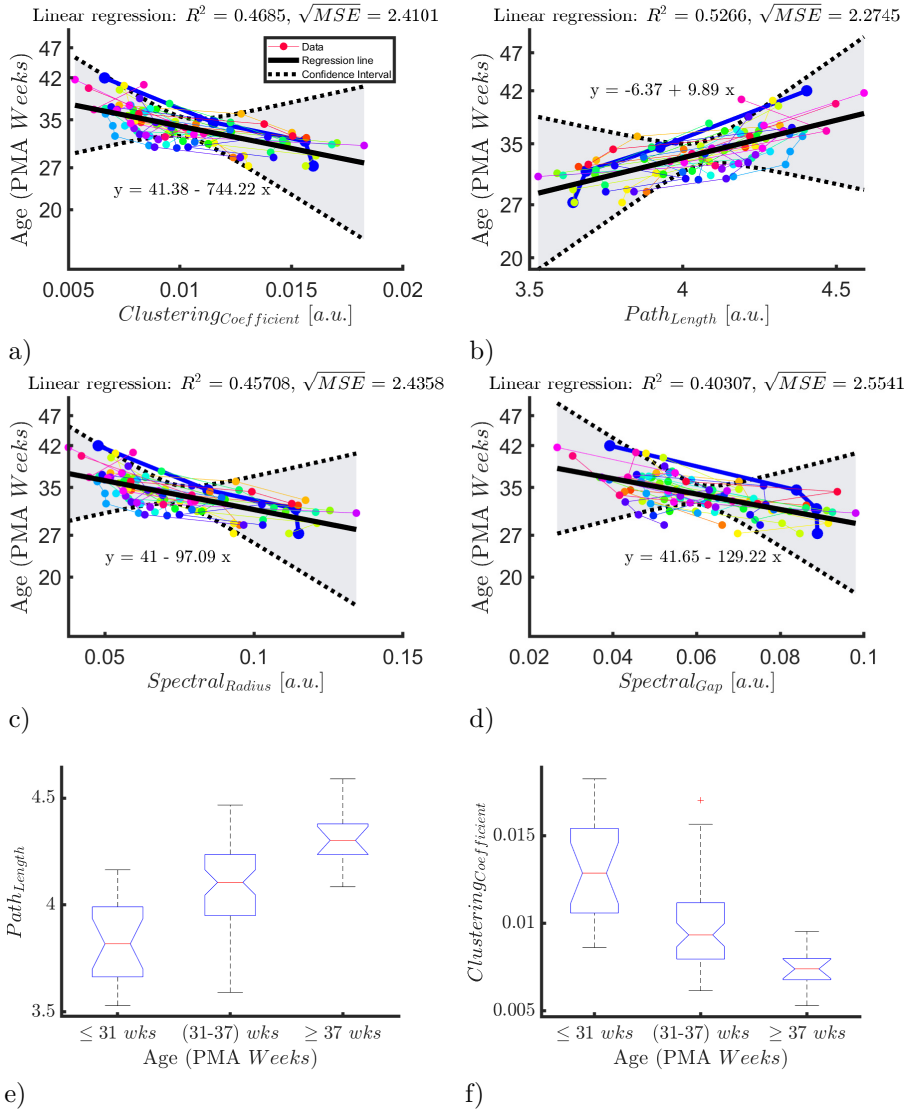


Figure 4.3: The figure shows the results for EEG data. The first four panels show OLS regression between 4 main graph indices vs the age for GC in QS. The grey area is the confidence interval at 95%. On the top of the panel, the associated R^2 and $\sqrt{MSE}$ in PMA weeks on the test set. The last two panels show the trend of the clustering coefficient and the path length in three distinct age groups. The results are reported about GC in QS.

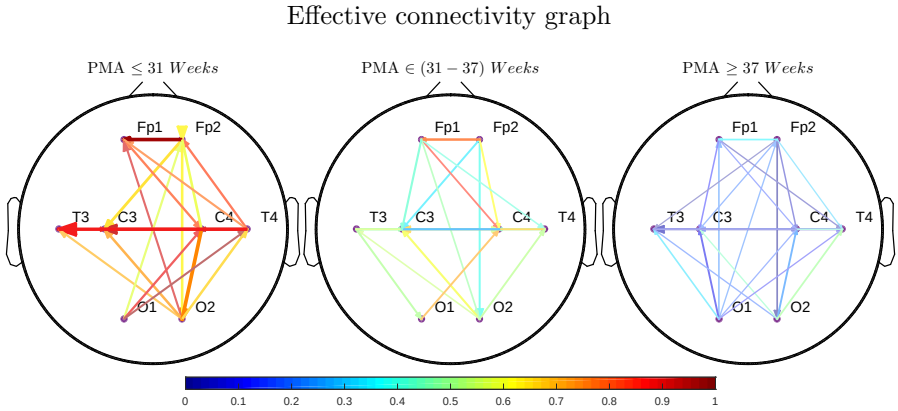
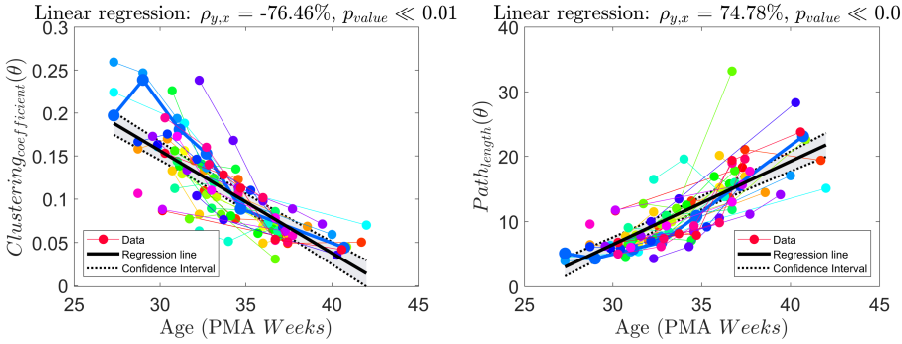


Figure 4.4: The figure shows the average connectivity graph (GC for three different age groups. The strength of the coupling among the electrodes is decoded by the color (the closer to the red color, the higher the coupling) and by the width of the arrow. The connectivity values have been normalized between 0 and 1 for the three groups together. The panels cleraly show the weakening of the coupling among EEG channels with maturation. The consequence is the increase of path length and the decrease of the clustering coefficient Figure 4.3.

4.3.4 Regression results

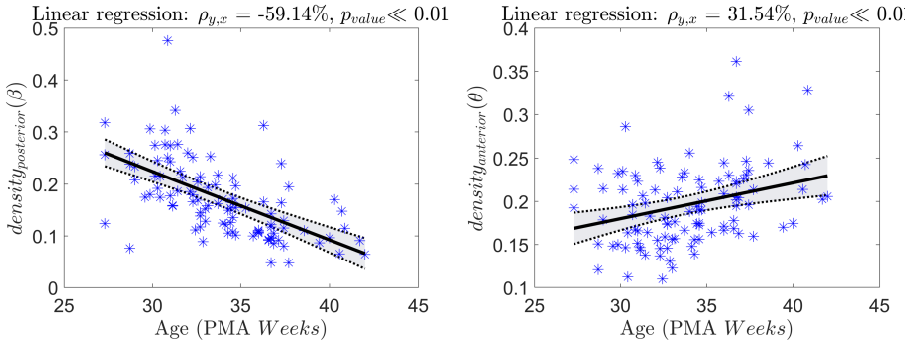
Table 4.6 reports the results for age prediction by means OLS regression and effective connectivity features. All the models can predict the age of the infant recording with a $\sqrt{MSE}$ between 2 to 3 weeks and the CCA models always outperform the model without CCA as a preprocessing step. Furthermore, the explained variance (R^2) is higher with the models that include CCA. It is also interesting to notice that best prediction results are obtained with GC during QS ($\sqrt{MSE}_{simple} = 2.52$ PMA weeks, $\sqrt{MSE}_{CCA} = 2.10$ PMA weeks). Table 4.6 does not report the results for one single model estimation, but the median and InterQuartile Range (IQR) of the evaluation parameters for 100 bootstrap iterations. In each single iteration, the model proved to be significant as reported by the p-value column in the Table 4.6 ($p < 0.01$).

Table 4.7 reports the results for the prediction via a linear mixed-effects model and FC network features. $ImCoh$ clearly outperforms any other FC methods in PMA prediction both in terms of MAE and R^2_{adj}. Considering the frequency bands, the θ band consistently outperforms the others in PMA prediction for



a) b)

Figure 4.5: The figure shows the linear mixed-effect regression between the clustering coefficient and the age (Panel a), and between path length and the age for the $ImCoh$ in θ bands (Panel b) during QS. Specifically, the black continuous line represents the estimated regression line, while the dashed lines are the confidence intervals. Each color in the plot represents a different patient. In order to show the specific trend for a single patient, we highlighted the result for a chosen infant in thicker red line. Above each panel, we reported the Pearson correlation coefficient and the associated p_{value}.



a) b)

Figure 4.6: The figure shows the linear mixed-effect regression between different MSC density indices and the age. Panel a) displays the density in posterior hemisphere in the β band during only QS, while panel b) shows the density in anterior hemisphere in the θ band during QS. The black continuous line represents the estimated regression line, while the dashed lines are the confidence intervals. In this figure, we reported only the data cluster without highlighting each patient. Above each panel, we reported the Pearson correlation coefficient and the associated p_{value}.

Table 4.4: The table displays the path length for the different FC methods during QS in the different age groups: PMA $\leq 31\ Weeks$, PMA $\in (31-37)$ $Weeks$, PMA $\geq 37\ Weeks$. Results are reported as median(IQR), where IQR stands for *interquartile range*. In the last column, the Pearson correlations of the path length with PMA is reported and the associated p_{value} is reported. ** stands for $p \leq 0.01$, * stands for $p \leq 0.05$, while $^{n.s}$ means $p = not\ significant$. The values $*10^3$ means that reported results are divided by $*10^3$.

Path Length - QS				
PMA	$\leq 31\ Weeks$	$\in (31-37)\ Weeks$	$\geq 37\ Weeks$	$\rho\ \%$
$ImCoh(\delta_1)$	14.54(5.32)	20.24(8.12)	24.81(16.75)	$49.53\%^{**}$
$ImCoh(\delta_2)$	8.97(3.63)	13.42(5.84)	20.17(5.01)	$75.48\%^{**}$
$ImCoh(\theta)$	6.11(1.85)	9.72(5.4)	17.71(6.19)	$74.78\%^{**}$
$ImCoh(\alpha)$	5.53(1.72)	7.2(3.21)	11.01(3.53)	$74.49\%^{**}$
$ImCoh(\beta)$	5.78(1.09)	6.03(1.31)	6.39(1.82)	$17.63\%^{n.s.}$
$k^2_{xy}(\delta_1)$	6.18(1.54)	6.18(2.01)	5.53(2.09)	$-9.61\%^{n.s.}$
$k^2_{xy}(\delta_2)$	6.17(2.36)	6.19(2.1)	5.52(1.77)	$-18.89\%^{n.s.}$
$k^2_{xy}(\theta)$	6.77(2.1)	6.67(1.92)	4.92(1.31)	$-44.85\%^{**}$
$k^2_{xy}(\alpha)$	7.52(2.15)	8.04(2.83)	7.08(2.7)	$-21.03\%^{*}$
$k^2_{xy}(\beta)$	6.45(2.67)	8.45(4.01)	10.1(6.25)	$48.49\%^{**}$
$PLV(\delta_1)$	3.86(1.78)	3.48(1.13)	3.66(1.35)	$-11.41\%^{n.s.}$
$PLV(\delta_2)$	3.02(0.68)	3.22(0.92)	3.09(0.67)	$-9.91\%^{n.s.}$
$PLV(\theta)$	2.99(0.74)	3.17(0.8)	2.95(0.52)	$-9.51\%^{n.s.}$
$PLV(\alpha)$	3.37(1.3)	3.54(0.91)	3.31(0.51)	$-15.94\%^{n.s.}$
$PLV(\beta)$	3.71(1.59)	4.13(1.05)	4.57(1.15)	$19.97\%^{*}$
$HSD(\delta_1)$	154.57(65.05)	132.93(47.37)	130.55(68.01)	$-20.45\%^{*}$
$HSD(\delta_2)$	167.54(63.25)	155.04(55.88)	135.17(36.72)	$-37.45\%^{**}$
$HSD(\theta)$	216.53(96.42)	172.87(52.74)	135.18(32.48)	$-56.62\%^{**}$
$HSD(\alpha)$	237.61(78)	205.71(65.18)	182.09(31.04)	$-46.08\%^{**}$
$HSD(\beta)$	236.16(100.17)	268.16(79.4)	274.02(98.75)	$14\%^{n.s.}$

the $ImCoh$, while MSC sees the β band as the most informative one, PLV the δ_1 and δ_2 band and HSD the θ and α band. These findings tend to be consistent in both sleep states. Table 4.7 also displays the result with the best selected features, which mainly come from $ImCoh$ in θ and α bands. Consequently, the prediction power is similarly to that of the $ImCoh$.

4.4 Discussion

In this chapter, we investigated recent approaches to describe EEG-based functional and effective connectivity to track and assess brain maturation in

Table 4.5: The table displays the path length for the different FC methods during NQS in the different age groups: PMA ≤ 31 $Weeks$, PMA $\in (31-37)$ $Weeks$, PMA ≥ 37 $Weeks$. Results are reported as median(IQR), where IQR stands for *interquartile range*. In the last column, the Pearson correlations of the path length with PMA is reported and the associated p_{value} is reported. ** stands for $p \leq 0.01$, * stands for $p \leq 0.05$, while $^{n.s}$ means $p = not\ significant$. The values $*10^3$ means that reported results are divided by $*10^3$.

Path Length - NQS				
PMA	≤ 31 $Weeks$	$\in (31-37)$ $Weeks$	≥ 37 $Weeks$	ρ %
$ImCoh(\delta_1)$	18.79(4.49)	21(6.11)	20.42(7.43)	$-2.8\%^{n.s.}$
$ImCoh(\delta_2)$	10.88(2.53)	14.73(3.87)	15.14(3.54)	$40.89\%^{**}$
$ImCoh(\theta)$	7.55(2.01)	12.44(5.69)	14.83(3.42)	$59.12\%^{**}$
$ImCoh(\alpha)$	6.68(1.31)	9.7(2.79)	10.72(1.39)	$68.83\%^{**}$
$ImCoh(\beta)$	6.5(1.04)	6.6(0.98)	6.51(0.8)	$1.95\%^{n.s.}$
$k^2_{xy}(\delta_1)$	5.99(1.43)	5.82(1.7)	4.85(1.85)	$-28.8\%^{**}$
$k^2_{xy}(\delta_2)$	6.18(2.49)	5.45(1.88)	5.68(1.36)	$-18.86\%^{n.s.}$
$k^2_{xy}(\theta)$	8.05(2.64)	6.19(2.29)	5.21(1.04)	$-45.49\%^{**}$
$k^2_{xy}(\alpha)$	10.29(3.62)	9.12(3.4)	8.22(4.64)	$-24.77\%^{*}$
$k^2_{xy}(\beta)$	7.45(2.22)	11.59(6.1)	14.44(6.98)	$56.21\%^{**}$
$PLV(\delta_1)$	3.8(1.3)	4.01(1.22)	3.84(1.4)	$-1.98\%^{n.s.}$
$PLV(\delta_2)$	3.3(0.7)	3(0.79)	3.23(0.66)	$-21.03\%^{*}$
$PLV(\theta)$	3.25(0.91)	3.09(0.7)	2.97(0.56)	$-24.43\%^{*}$
$PLV(\alpha)$	3.82(0.82)	3.74(0.99)	3.65(0.57)	$-20.84\%^{*}$
$PLV(\beta)$	4.17(1.09)	4.55(1.18)	4.99(1.38)	$32.21\%^{**}$
$HSD(\delta_1)$	140.41(51.88)	137.28(38.59)	127.26(43.46)	$-15.78\%^{n.s.}$
$HSD(\delta_2)$	165.8(44.65)	133.66(44.25)	133.97(41.04)	$-34.96\%^{**}$
$HSD(\theta)$	195.53(76.71)	152.3(57.33)	131.62(32.15)	$-53.99\%^{**}$
$HSD(\alpha)$	268.11(46.97)	220.82(69.59)	193.22(37.79)	$-44.76\%^{**}$
$HSD(\beta)$	266.93(77.01)	292.73(104.85)	355.53(170.72)	$36.15\%^{**}$

premature infants. Firstly, different methodologies were employed to describe the connectivity development; secondly, different graph features were used to define a regression models to extrapolate age, known as brain-age models [89].

The current chapter illustrates the application of two well-known methods to estimate directional coupling between processes, like GC and TE, as well as a range of linear and nonlinear methods to investigate temporal correlations among time series ($ImCoh$,MSC,PLV,HSD). Based on the obtained connectivity matrices, integration and spectral network indices were estimated for both directed and undirected graphs. The EC graph features were firstly applied on a simulated dataset to investigate how the network measures behaved in a controlled case. Subsequently, the FC and EC graph features were used

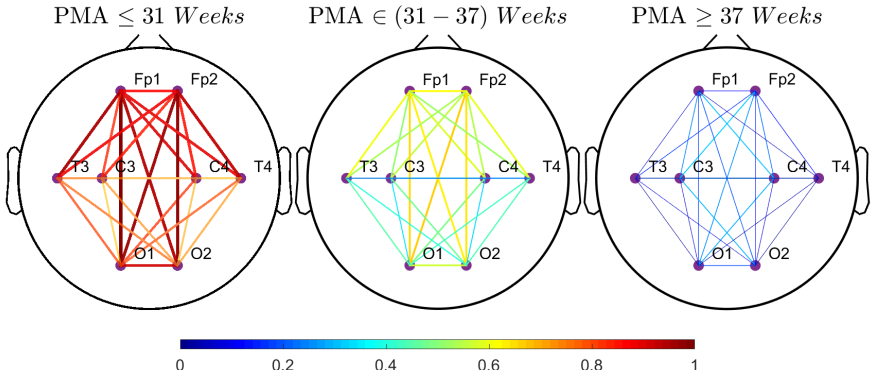


Figure 4.7: Topographic distribution for $ImCoh(\theta)$ in QS: PMA $\leq$ 31 *Weeks*, PMA $\in$ (31 – 37) *Weeks*, PMA $\geq$ 37 *Weeks*.

Table 4.6: Multivariate regression model performances. The table shows the error on the test set (Error), the R^2 and the F-statistics ($\mathcal{F}$-stat) and the p-value obtained with the different connectivity methods in the different sleep states. The results are reported as median(IQR), where IQR stands for *InterQuartile range* over the 100 random splits of the dataset. The labels reported are TE = transfer entropy, GC = Granger causality, QS = quiet sleep, NQS = non-quiet sleep.

Multivariate regression performances				
Median(IQR)	Error(weeks)	R^2	$\mathcal{F}$-stat	P-value
Simple filtering				
TE - QS	2.54(0.41)	0.57(0.07)	11.64(3.46)	$p < 0.01 * 100$
TE - NQS	2.88(0.39)	0.40(0.07)	6.23(1.72)	$p < 0.01 * 100$
GC - QS	**2.52(0.37)**	**0.52(0.07)**	**10.01(2.64)**	$p < 0.01 * 100$
GC - NQS	2.79(0.53)	0.44(0.07)	7.20(2.09)	$p < 0.01 * 100$
CCA				
TE - QS	2.23(0.29)	0.63(0.06)	12.90(3.14)	$p < 0.01 * 100$
TE - NQS	2.54(0.51)	0.57(0.06)	10.36(2.66)	$p < 0.01 * 100$
GC - QS	**2.10(0.38)**	**0.67(0.05)**	**15.96(3.64)**	$p < 0.01 * 100$
GC - NQS	2.35(0.42)	0.63(0.04)	13.38(2.61)	$p < 0.01 * 100$

Table 4.7: The performance of the different linear mixed-effect models for the investigated FC methods is displayed. The table reports the adjusted explained variance R^2_{adj} on the training set and the MAE on the test set as median(IQR), where IQR stands for *interquartile range*. The results are reported for both sleep states, QS and NQS. The last row reports the result with selection of the best features via LASSO.

FC methods	QS		NQS	
	R^2_{Adj} [.]	MAE ($Weeks_{PMA}$)	R^2_{Adj} [.]	MAE ($Weeks_{PMA}$)
$ImCoh(\delta_1)$	0.37(0.1)	2.06(0.38)	0.09(0.06)	2.62(0.45)
ImCoh(δ_2)	**0.72(0.09)**	**1.61(0.27)**	0.26(0.07)	2.31(0.34)
ImCoh(θ)	**0.81(0.08)**	**1.54(0.22)**	**0.62(0.06)**	**1.59(0.31)**
ImCoh(α)	**0.66(0.14)**	**1.7(0.3)**	**0.61(0.08)**	**1.54(0.27)**
$ImCoh(\beta)$	0.49(0.09)	2.08(0.29)	0.39(0.08)	2.21(0.42)
$k^2_{xy}(\delta_1)$	0.08(0.06)	2.64(0.36)	0.36(0.08)	2.07(0.39)
$k^2_{xy}(\delta_2)$	0.2(0.07)	2.46(0.4)	0.26(0.09)	2.47(0.45)
$\mathbf{k^2_{xy}}(\theta)$	**0.38(0.08)**	**2.03(0.25)**	0.39(0.1)	2.05(0.38)
$k^2_{xy}(\alpha)$	0.38(0.09)	2.14(0.39)	0.28(0.09)	2.35(0.34)
$\mathbf{k^2_{xy}}(\beta)$	**0.41(0.12)**	**2.16(0.4)**	**0.47(0.08)**	**1.94(0.36)**
PLV(δ_1)	0.18(0.08)	2.47(0.36)	**0.44(0.09)**	**1.96(0.36)**
PLV(δ_2)	**0.27(0.07)**	**2.33(0.33)**	**0.32(0.11)**	**2.32(0.39)**
$PLV(\theta)$	0.24(0.07)	2.31(0.33)	0.23(0.09)	2.44(0.42)
$PLV(\alpha)$	0.29(0.12)	2.26(0.36)	0.26(0.08)	2.4(0.35)
$PLV(\beta)$	0.15(0.07)	2.43(0.33)	0.34(0.14)	2.25(0.33)
$HSD(\delta_1)$	0.17(0.09)	2.66(0.37)	0.08(0.06)	2.67(0.36)
$HSD(\delta_2)$	0.54(0.1)	1.97(0.35)	0.26(0.07)	2.27(0.41)
HSD(θ)	**0.59(0.07)**	**1.82(0.33)**	**0.51(0.05)**	**1.87(0.33)**
HSD(α)	**0.54(0.07)**	**1.92(0.34)**	**0.45(0.06)**	**1.96(0.34)**
$HSD(\beta)$	0.28(0.08)	2.27(0.4)	0.28(0.09)	2.4(0.34)
Best features	**0.8(0.11)**	**1.51(0.31)**	**0.65(0.07)**	**1.64(0.29)**

to predict the age of the patient on a wide maturation period (from birth to full-term age) in an EEG dataset of health preterm infants.

4.4.1 Simulated dataset

According to [43], graphs with high clustering coefficient and low path length behave like a rich-club network, while graphs with low clustering coefficient and high path length denote a random network, where the number of edges for each node is normally distributed. Figure 4.1 portraits two club networks, where the nodes are connected to each other with a short distance. This leads to a high clustering coefficient and low path length when the coupling coefficient is equal to 1. However, when β decreases, the intra-cluster connectivity weakens and

the graph becomes more similar to a random network. This type of graph is characterized by low clustering coefficient and high path length: indeed, the nodes are less connected among each other in a more homogeneous network. This result is also supported by the direct proportionality between spectral radius and coupling coefficient. Another interesting point is related to the filtering. Figure 4.2 illustrates clearly how the clustering coefficient and the path length estimation can be highly affected by the filtering. Those results are in line with the analysis by [19], which demonstrated that careless filtering can add spurious connectivity in the time courses. In our simulation, the effect of filtering weakens the intracluster connectivity (adding inter-cluster connectivity). The net effect is a decrease in clustering coefficient and an increase in path length. Therefore, filtering must be carefully applied when methods like GC and TE are applied. In the context of this research, the notch filter and differentiation were the only preprocessing steps before computing EEG effective connectivity.

4.4.2 EEG data

In the literature, a number of studies can be found to have assessed the brain maturation in children and adolescents by graph theory [241],[94],[213] and a few papers focused on preterm brain maturation by network metrics [185], [26]. The findings obtained in the reported studies confirm that the both effective and functional connectivity change with the development of the premature brain. Although Schumacher et al. [229] used a different method, they also concluded that there is a change in effective connectivity mainly driven by postnatal maturation (Figure 4.3). Regarding functional connectivity, Grieve and Meijer have also shown an increase of the $k^2_{xy}(\theta)$ in the frontal area (Figure 4.6 and Table 4.4 and 4.5) and a decrease of $k^2_{xy}(\beta)$ in the occipital area (at least, during gestational maturation) [99], [168]. Similarly, Gonzalez has shown a decrease in $ImCoh(\delta_2)$ [95].

In this analysis, we also observed a change in graph parameters that suggest that the EEG-scalp network moved from a rich-club network to a more random network, especially if the effective connectivity and the $ImCoh$ are considered (Figure 4.3 and 4.5, Table 4.3, 4.2, 4.4 and 4.5). The integration and spectral indices decreased with age, while the path length increased. It might reflect a segregation of nodes due to a higher graph distance as well as less intense triangle patterns around the nodes themselves. Hub-networks have high clustering coefficients since they have central club nodes, which are surrounded by triangle patterns. On the contrary, a random network present nodes, which are connected to any other node in the network with a weak coupling. The net effect is a low clustering coefficient and high path length. The results of the multivariate regression models further highlight the shift from a rich-club network at younger

age to a more random one at full-term age. In case of effective connectivity, the lowest $\sqrt{MSE}$ on the test set is around 2.11 weeks (Table 4.6), which is comparable to other studies [146], [190]. The negative trends for clustering coefficients and positive trends for the path length features are consistent for each effective connectivity method and each sleep state. Furthermore, both models with and without CCA found the same trends, but source filtering increased the prediction power of the model. It is possible that the EMG artifacts disturbed the connectivity analysis and biased the prediction model in the first considered scenario. In case of functional connectivity, the lowest MAEs on the test set was 1.51 *weeks* for QS and 1.64 *weeks* for NQS with combination of different features (Table 4.7). Nonetheless, the best regression performance is normally obtained with $ImCoh$, which also presented a positive trend for the path length and a negative trend for the clustering coefficient.

Additionally, this emergence of a normal-distributed network is also confirmed by the decrease of the spectral gap, spectral radius and the algebraic connectivity. The latter two indices emphasize how easily the graph can be clustered and a negative trend would suggest the absence of modules or groups in the graph. On the contrary, a negative trend of the former index should suggest an increase of modularity, as shown by [77]. However, since the spectral gap is also inversely proportional to the path length, its reduction just shows that the spectral radius is less dominant with respect to the other eigenvalues [30] and it becomes another measure of modularity like the other two spectral indices.

At first sight, the obtained results seem to be in contradiction with maturation trends that can be found in children or adolescents, where a shift from random to a rich-club network has been discovered. However, it should be taken in account that, on one side, only 9 electrodes on the scalp were used, due to the small size of the preterm brain and, on the other hand, there is a fast development of the brain during this monitoring period with different trajectories for the different cerebral regions [134]. This composite evolution is mainly driven by the different disappearance timing of cortical subplate in the various brain areas [134]. In particular, two main changes took place. The first one is the faster development of the thalamo-cortical connections compared to the cortico-cortical ones [26]. Especially, the thalamo-cortical connections are dominant in the first part of the preterm development (up to 31 PMA $Weeks$), in order to develop a substrate for sensory-cortical activation. The cortico-cortical connections prevail in the second part of the development (after 31 PMA $Weeks$), in order to shape the folding of the different cortical areas [134], which causes the vanishing discontinuity in the neonatal EEG [134], [270]. This could induce either a burst de-synchronization or instantaneous synchronization, which both will diminish the $ImCoh$ and might explain why this feature specifically outperform the other FC methods (Table 4.7). The second important change is the negative

correlation between age and short-range cortico-cortical connections, as shown by [26] via fMRI. Their study results showed that long distance connections develop faster than the short ones. Furthermore, the development is characterized by a strengthening of the former connections and weakening of the short-range couplings [94]. Consequently, the EEG electrodes/nodes (which measure short-range connectivity) tend to separate each other, with an increase of the path length and a reduction of the clustering coefficient. This hypothesis is also supported by the decrease in fronto-frontal and occipito-occipital functional coupling measured by fMRI [26].

Therefore, the development of thalamo-cortical and long cortico-cortical connections and the weakening and short-range wiring may trigger a separation and segregation of the different brain regions. This segregation hypothesis was also been debated in [241]. The author reports the different studies which investigated a decrease of FC in general brain development. A possible explanation of this connection reorganization is a shift from a more diffuse cortical organization to a focalized one, probably driven by the cortical folding. Similarly, Meijer also described a segregation on the cortex due to the development of specialized tasks associated to the different brain areas, which also finds support in the decrease of the $ImCoh$ and in the hypothesis that sources are on the cerebral cortex of the infant [168]. This result can be further confirmed by the decrease of connectivity in $k^2_{xy}(\beta)$ (Figure 4.6). This interpretation can be contradicted by positive trends of MSC, PLV and HSD in the other bands. A possible explanation is that $k^2_{xy} = |c_{xy}|^2$, which contains the real part of the coherence, and the $ImCoh$ have different information. Grieve argued that the lower $k^2_{xy}(\theta)$ in preterm babies is a sign of higher reactivity to sensory inputs during sleep [99]. An increase in MSC and HSD in that band could mean a decrease in reactivity (higher instantaneous interaction), but a lesser degree of lagged interaction with a lower $ImCoh$. This opposite trend between the two indices was also reported by Gonzalez [95], who argued that stimuli and intervention in the postnatal maturation could influence those trends. Whether this hypothesis is true, it should be tested with a proper study. However, Myers [177] has already shown that the skin-to-skin contact (which is a tactile stimulus) drives a sharper decrease rate in the MSC with postnatal maturation. This approach could be extended to $ImCoh$ in future research.

It is important to point out that this segregation can be emphasized by the fact that there are a few electrodes on infant scalp. However, a study with high-density EEG on preterm infants [185] found an increased modularity on the scalp EEG network and a reduced clustering coefficient in the postcentral network, while the clustering coefficient increases in the precentral network. This result could confirm the segregation or the more local integration of the brain network due to the pruning of short-range connections, as also shown by

[82] in the comparison between adults and children.

4.5 Summary

This chapter provides new functional and effective EEG-based brain connectivity methods for post-menstrual age estimation in healthy premature neonates. The main novelty lies on the extensive usage of graph features to describe EEG interactions, as well as, the comparison of multiple regression models to quantify the predictive power of the different connectivity methods. In particular, the clustering coefficient decreased with maturation, while the path length increased. This perspective suggests that the EEG graph shifted from a small-world network to a random network. This apparent nodes' segregation can be a consequence of the thalamo-cortical connections development and the strengthening of the long-range cortical connections. Those hypothesis are confirmed by the regression results. In case of effective connectivity, the lowest age prediction MSE was 2.11 PMA weeks and was obtained with GC in QS. In case of FC, $ImCoh$ outperforms the other methods to predict the age of the patient and the best regression performance parameters were were $MAE = 1.51$ weeks and $R^2_{adj} = 0.80$. This result could be explained by a more diffused functional architecture to perform different tasks on the cortex.

Chapter 5

A multifractal framework for quiet sleep detection and the estimation of age

This chapter has partially been published as ***Lavanga M.****, De Wel O., Caicedo A., Heremans E., Jansen K., Dereymaeker A., Naulaers G., Van Huffel S. (2017). " Automatic quiet sleep detection based on multifractality in preterm neonates: Effects of maturation." Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), EMBS, (pp. 2010–2013). Lavanga M. has developed the methodology, conducted the experiments and has written the manuscript. Compared to the publication, the text has been expanded to report the assessment of brain maturation by means of the multifractal framework.*

So far, two age-prediction models have been reported for premature infants based on the EEGs interdependence, normally referred to as Brain-Age models. This chapter presents the first model based on a single-channel or univariate analysis of the EEG, which can ideally predict the maturation of the infants with a set of reduced electrodes. The methodology presented here uses the multifractality framework, which investigates the changes in fractality of the different scales of the EEG. In particular, the variations in regularity and multifractal EEG properties in both development and in different sleep stages will be reported. The first part of the chapter is dedicated to the outline of a quiet sleep classifier based on multifractality and fixed-size LS-SVM. The second part is the description of a regression model to predict the age of the infants based on the multifractal

features. The study reported in this chapter led to a publication for the part concerning the sleep staging [144], while the maturation model is based on unpublished results.

5.1 Introduction

Assessment of brain maturation and sleep-wake ciclicity by means of EEG analysis has become standard clinical practice [198], [17]. However, the vast amount of EEG data that are normally collected in neonatal intensive care units (NICUs) has made manual sleep scoring or visual assessment of brain dysmaturity extremely time consuming. Due to advances in biomedical signal processing, automatic neonatal EEG monitoring became a viable option. Different algorithms to track sleep stages by means of linear and nonlinear EEG features have been published [208]. These methods tend to rely on the fact that quiet or NREM sleep (QS) is characterized by a discontinuous tracing, while the active or REM sleep exhibits a more continuous tracing. Similarly, it is possible to quantify the level of infant's development by tracking the EEG discontinuity or other electrophysiological changes. Based on different EEG-derived features, different authors designed regression models to automatically estimate the age of the infant [191],[192],[149],[65].

Among nonlinear features, fractal analysis received an increased interest to discriminate sleep stages or describe the neurodevelopment of an infant, because it can describe the morphology and the complexity of the EEG signals [3]. In particular, Accardo et al. showed that the fractal dimension decreases during the slow-wave sleep of full-term newborns, while it reaches its maximum value when the subject is awake [3]. O'Toole et al. showed that the fractal dimension of the full-channel EEG is expected to decrease with increasing age [192] and Hartley et al. investigated long-term range correlations of the EEG activity in the first days of life in premature infants [103]. In parallel, other studies showed that entropy of the EEG (even at different scales) increases with the development of infant [271], [65], which nicely complements the decrease of the regularity shown by fractal measures (see Section 3.1.2 and Section 3.1.3). It is important to stress that the measurement of the fractal dimension investigates only self-similarity or the global scaling behavior (*long-term persistance*), which assumes that the signal is monofractal. However, Popivanov et al. [211] showed that the EEG presents different local scaling behavior, known as multifractality. The fact that the fractal dimension is changing through the different sleep stages highlights the higher complexity of EEG signals compared to a self-similar process like the Brownian motion (Section 3.1.2). If we consider that the neonatal EEG in the premature infant is in constant development, the intertwined nature of sleep and

development can not only affect how the EEG differs among sleep stages, but it can also show a structure of EEG development which is not trivial and might go beyond the simple decrease of discontinuity. If sleep is taken into account, the duality between the two common behavioral states (quiet and active sleep) evolves during development (see Figure 2.3). Below 31 weeks postmenstrual age (31 w PMA) the different sleep stages can not be distinguished, while at 32 to 36 weeks there is a maximal separation between the 2 sleep states. After 37 weeks, different sleep stages start flourishing in the EEG signal with the first seeds of *Tracé Alternant*, which is fully present at full-term age (see section 2.2.2). If the development is considered, the evolution of EEG is characterized by emergence of different rhythms, which might consist in burst overlapped with higher frequency information (brushes with α frequency or temporal sawtooths), or activity localized in specific area (such as the encoches frontales).

Therefore, one might expect the fractal properties of the signal to differ among sleep stages, to change with age and to evolve differently according to the sleep state. The univariate analysis in this chapter exploits multifractality of neonatal EEG for automatic quiet sleep detection in preterm babies and the assessment of the neurodevelopment of the infant. It also describes the impact of maturation on the performance of this sleep stage classifier and the different development trends of the EEG according to sleep state.

5.1.1 Dataset

Twenty-five preterm neonates, with normal developmental outcome at 2 years, were recruited at the neonatal intensive care unit (NICU) in the University Hospitals Leuven. The patients have a PMA ranging from 27 to 42 weeks and they represent a subset of the dataset in Chapter 4 and Section 1.2. For each subject, eight EEG signals were recorded according to the restricted 10-20 international system (F_1,F_2,C_3,C_4,T_3,T_4,O_1,O_2) with the electrode C_z as reference. The measurements were performed at least twice during their stay at the unit (at different PMA) and lasted at least 2 hours, producing a dataset of 88 recorded EEGs. The monopolar EEG signals were recorded at a sampling frequency of 250 or 500 Hz. Each channel was filtered using a band-pass filter with a band-pass between 1-20 Hz and subsequently they were downsampled to 128 Hz for uniformity in the analysis. Clinicians manually detected the quiet sleep (QS) and non-quiet sleep epochs (NQS) in the polygraphs, which is required to develop a supervised automated hypnogram. This study mainly investigated QS detection, because the awake states are difficult to discriminate from the active sleep periods at very young age. For this reason, only QS/NQS stages will be referred to in the remaining sections. Figure 5.1 shows a representative

segment for QS and NQS epochs from a recording of a prematurely born neonate at 42w PMA.

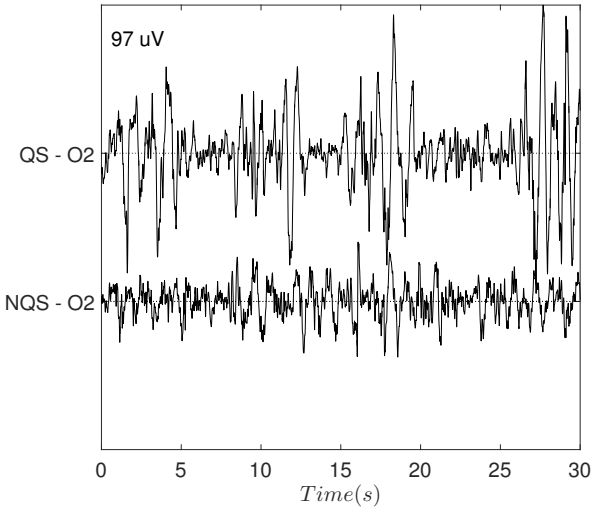


Figure 5.1: Example of EEG signal during QS and NQS epoch (prematurely born neonate at 42w PMA).

5.1.2 Multifractal features and classical EEG features

A thorough description of the multifractal formalism has been reported in section 3.1.3. The main hypothesis of this study are that the multifractality features c_1, c_2, c_3 and Δh are able to discriminate QS epochs from NQS periods in neonatal EEG and they can describe the evolution of EEG discontinuity and track infants' maturation.

To evaluate whether the first hypothesis is correct, we compared their classification power with the features traditionally used in the literature for this task, as described in [208]. In particular, the power in the main frequency bands, the spectral edge frequency, as well as the spectral moment and the spectral entropy were selected for this subset. Furthermore, Shannon amplitude entropy and fractal dimension (according to Katz's algorithm) were also computed. For each recording, all the mentioned features were computed for 30 s non-overlapping windows in each channel. Each window was labeled as QS or NQS according to the clinicians' labelling. The multifractal features were computed

using the Wavelet p-Leader and Bootstrap based MultiFractal analysis (PLBMF) MATLAB toolbox, described in [266]. This toolbox can be downloaded from `https://www.irit.fr/~Herwig.Wendt/software.html`. The feature c_1 is also normally referred to as the main Hurst exponent of the signal H_{exp}.

In order to test the second hypothesis, we developed regression models to estimate the age of the patients based on the same multifractal features derived to discriminate sleep stages. Since the age-model aims to predict the postmenstrual age associated to each recording, the median of each feature was derived for the windows associated to quiet sleep and non-quiet sleep in order to obtain only two data point for each recording (grand-averaging or grand-median process).

5.1.3 Classification approach

A feature matrix $\mathbf{X} \in \mathbb{R}^{N \times d}$, where $N = 102209$ and $d = 112$, was produced for the complete dataset. The dimensions of d include 4 multifractal measurements, 9 spectral features and 1 monofractal dimension for each one of the 8 EEG channels ($d = 14 * 8 = 112$). In addition, each row was associated to a classification vector $\mathbf{Y} \in \mathbb{R}^{N \times 1}$ with 1 for QS and -1 for NQS. In order to develop a supervised model, least-squares SVMs (LS-SVM) were chosen as a suitable classification method to discriminate QS epochs, as described in [242] and Section 3.4.2. To study the maturational effect on QS detection, the dataset $\mathbf{X}$ was first split into three groups according to the infant's PMA. The first group contains all data points that belong to recordings from neonates up to 31 weeks PMA ($N = 11541$). The second one contains all data points from neonates between 31 and 37 weeks PMA ($N = 57053$). The remaining dataset contains all recordings of neonates beyond 37 weeks PMA ($N = 33615$). Despite the splitting, the number of data points is extremely large to train and tune the classification model. In order to reduce the size of the dataset for the training of the classifiers, we used a modified approach to a fixed-size LS-SVM, where a training set with a reduced size M ($M \ll N$) is selected from the available data. In fixed-sized LS-SVM these training points are selected with an iterative process that maximizes the quadratic Renyi entropy. Further details can be found in Section 3.4.4. However, according to Varon et al. [253], it is possible to choose a reduced size dataset as centroids of M clusters in the original data cloud, which are defined with the k-medoids method. In this way, a dataset with a high entropy can be provided as initial guess for this iterative process [253]. Due to the fact that maximizing the Renyi entropy might lead to the selection of outliers, we reduced the dataset with k-medoids clustering without applying the iterative entropy process, in contrast to [253]. QS classifiers were implemented for the different age groups, as well as for different sets of input features. Specifically, $\mathbf{X}_{katz} \in \mathbb{R}^{N \times 8}$, $\mathbf{X}_{c1,c2,c3} \in \mathbb{R}^{N \times 24}$,

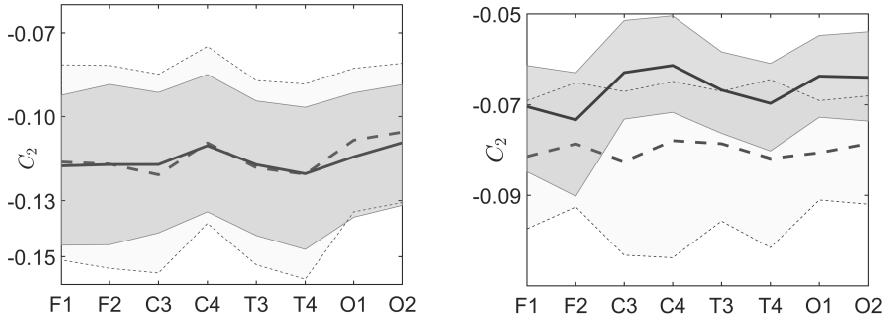
$\mathbf{X}_{\Delta h} \in \mathbb{R}^{N\times 8}$, $\mathbf{X}_{spectral} \in \mathbb{R}^{N\times 72}$, $\mathbf{X}_{c2,\Delta h} \in \mathbb{R}^{N\times 16}$, $\mathbf{X}_{c_2,c_1,c_3,\Delta h,spec} \in \mathbb{R}^{N\times 104}$, $\mathbf{X}_{c_2,c_1,c_3,spec} \in \mathbb{R}^{N\times 96}$, $\mathbf{X}_{all} \in \mathbb{R}^{N\times 112}$ represent the different set of input features used to train the classifiers. The training set was reduced to size $M = 1500$, using the k-medoids method explained before, and the model was tested on the remaining data points for each age group ($N_{test,\leq 31w} = 10041$, $N_{test,\in(31-37)w} = 55553$, $N_{test,\geq 37w} = 32115$). The LS-SVM model was tuned with a 10-fold crossvalidation on the training set. The classifiers' performance was measured as *area under the curve* (AUC) of the ROC curve.

5.1.4 Regression model

After the grand-median process, the feature matrix was shrunk to the $\mathbf{X}_{mat} \in \mathbb{R}^{P\times d}$, where $P = 88$ and $d = 32$. The dimension P represents the total number of recordings in the dataset, while d is the total number of multifractal features. Each set c_1,c_2,c_3 and Δh is normally multiplied by the number of channels ($4 * 8 = 32$). The predictive power of the different sets of features were tested with a linear mixed-effect model,as reported in section 3.3.2. For each type of multifractal feature, a multivariate regression model with a random split of the data in 70%/30% of train/test set based on the patient label and for 10 iterations. The random effect was represented by the patient ID as grouping variable, while the fixed effects were represented by the same feature for each channel (8 indipendent variables). In order to test the global predictive power of all features, the total set of 32 features were considered in the regression modeling. The model was tuned on 10 random split of the data in 70%/30% of train/test set. For each iteration, a feature selection was applied via the least absolute shrinkage and selection operator (LASSO, see Section 3.3.3). The performances were then assessed as *mean absolute error* MAE on the test set, as well as explained variance R^2 both on train and test set (R^2_{train}, R^2_{test}). See section 3.3.1 for more details. The results were reported as median(IQR).

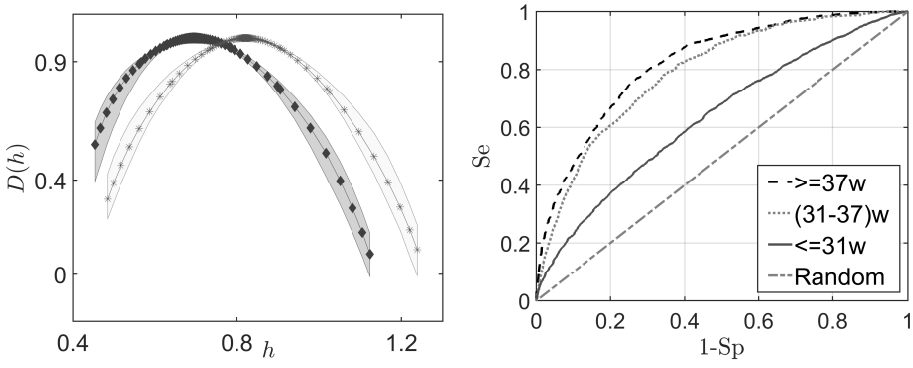
5.2 Results

Figure 5.2 depicts the median and the 25 - 75 percentiles for c_2 in QS epochs (dashed line) and NQS ones (continuous line) for the group with the youngest neonates (left panel) and the group with the oldest neonates (right panel). The x-axis in the figures represent each of the EEG channels. The right panel shows that the parameter c_2 discriminates the sleep epochs for the oldest patients (Age $\geq$ 37w PMA) in the channels $[C_3, C_4, O_1, O_2]$, while the left one does not show any discriminative power from c_2 (Age $\leq$ 31w PMA). In Figure 5.3.a an example of SS for NQS (diamonds) and QS (stars) segments in a recording (prematurely



a) b)

Figure 5.2: Median and 25-75 percentile for the c_2 values in QS (dashed line) and NQS (continuous line) epochs. On the left panel, the values from the dataset with the youngest neonates. On the right panel, the values from the dataset with the oldest neonates.



a) b)

Figure 5.3: The left panel shows an example of a mean singularity spectrum in QS epochs (stars) and NQS epochs (diamonds) for one specific recording (prematurely born neonate at 42w PMA). The right panel shows the ROC curve for LS-SVM classifiers, using $\mathbf{X}_{c_2,\Delta H}$ as feature set, in the three different age groups. The continuous line represents the ROC curve for all recordings in the youngest group, the dotted line represents the ROC for all recordings in the middle age group, while the dashed line represents the ROC curve for all recordings in the oldest group.

born neonate at 42w PMA) are shown. The latter is wider than the former, leading to the hypothesis that QS has a greater number of singularities, i.e. the QS epochs are more multifractal than the NQS ones. Table 5.1 shows the AUC for different LS-SVM classifiers for the different feature sets in the different age groups. Although the spectral features always outperform the fractality measures, it can be seen that the performance of the latter increases with age, especially for the $\mathbf{X}_{c1,c2,c3}$ feature set. This maturation effect is also shown in Figure 5.3.b, which reports the ROC curves for different age groups using $\mathbf{X}_{c_2,\Delta h}$ as feature set. At the youngest age, the classifier exhibits the lowest performance. Beyond 31w PMA, the AUC dramatically increases reaching its maximum at full-term age. It is interesting to notice that all performance results were obtained with a test set which was at least ten times bigger than the training set (1500 vs 10041, 1500 vs 55553, 1500 vs 32115). In addition, the EEG data were not manually preselected to compute the features.

Figure 5.4 and Table 5.2 show the change of the multifractal parameters according to the maturation of the infant. Figure 5.4 respectively reports the results for NQS and QS on the left column and the right column, while the Table 5.2 respectively shows the results on the upper half and the bottom half. The figure displays the current set of feature: the Hurst exponent for T_3, the c_2 and c_3 for the same channel and the difference ΔH for the channel Fp_1 (H_{exp,T_3},c_{2,T_3},c_{3,T_3} and ΔH_{Fp_1}). All the fractal features show a significant correlation with age: both H_{exp,T_3} and ΔH_{Fp_1} decrease with age, while the parameters c_{2,T_3},c_{3,T_3} increase with maturation. However, the Pearson correlation ρ_{xy} shows stronger correlation during NQS (left column) than QS. Especially, the parameters c_{2,T_3},c_{3,T_3} have respectively ρ_{xy} equal to 83% and

Table 5.1: The AUCs for the different LS-SVM classifiers. The rows represent the different input feature sets, while the columns represent the different age groups.

AUC for different classifiers			
Age (PMA w)	≤ 31	$\in (31-37)$	≥ 37
$\mathbf{X}_{Katz}$	.63	.69	.70
$\mathbf{X}_{c_1,c_2,c_3}$	.70	.82	.88
$\mathbf{X}_{\Delta h}$	.67	.78	.79
$\mathbf{X}_{spectral}$	.83	.93	.91
$\mathbf{X}_{c_2,\Delta h}$	.63	.79	.82
$\mathbf{X}_{c_2,c_1,c_3,\Delta h,spec}$	.83	.94	.93
$\mathbf{X}_{c_2,c_1,c_3,spec}$	.84	.93	.93
$\mathbf{X}_{all}$	.82	.94	.92

Table 5.2: The performance of the different linear mixed-effect models for each multifractal features set is reported for both NQS (upper half) and QS (bottom half). The table reports the adjusted explained variance R^2_{ord} on the training set, the coefficient R^2_{test} on the test set and the MAE on the test set as median(IQR), where IQR stands for *interquartile range*. The first four models are multivariate regression model with the same feature for each channel. The last row reports the result of the model based with selected features via LASSO, reported in Table 5.3.

NQS			
	R^2_{Ord}	MAE	R^2_{test}
H_{exp}	0.52(0.07)	1.6(0.22)	0.53(0.15)
c_2	0.72(0.09)	1.41(0.28)	0.68(0.09)
c_3	0.64(0.05)	1.93(0.18)	0.57(0.09)
ΔH	0.75(0.08)	1.52(0.28)	0.68(0.12)
All	0.83(0.08)	1.38(0.23)	0.69(0.18)
QS			
H_{exp}	0.52(0.09)	1.72(0.31)	0.51(0.11)
c_2	0.23(0.14)	2.61(0.48)	0.19(0.18)
c_3	0.63(0.1)	2.3(0.4)	0.39(0.16)
ΔH	0.61(0.06)	1.99(0.61)	0.39(0.26)
All	0.78(0.07)	1.96(0.49)	0.42(0.25)

Table 5.3: The selected features of the regression models in the last rows of the two blocks of Table 5.2. The features reported are : the Hurst exponent, the parameters c_2 and c_3 for the width and the asymmetry of $D(h)$ and the difference between minimal and maximal Hurst exponents (H_{exp}, c_2, c_3 and ΔH_{Fp_1}).

	Features
All - NQS	H_{exp,O_1},c_{2,T_3}, c_{3,T_3}, ΔH_{Fp_1},ΔH_{O_1}
All - QS	H_{exp,T_4},c_{3,T_3},ΔH_{Fp_1},ΔH_{O_1}

72% in NQS and equal to 42% and 67% in QS. This difference in correlation is mirrored in the regression results reported in Table 5.2. The NQS state have better performance in terms of R^2 and MAE, but for H_{exp}. Indeed, the results in terms of the main Hurst exponent are comparable between two states, while c_2 shows a significant plunge in performance ($MAE_{NQS} = 1.41\ wks$ vs $MAE_{QS} = 2.61\ wks$). Overall, the best performance in age prediction is reached when all features are pooled together, but QS consistently underperform compared to NQS ($MAE_{NQS} = 1.38\ wks$ vs $MAE_{QS} = 1.96\ wks$). The features selected at least 50% of the iterations by LASSO are reported in Table 5.3 for both behavioral state.

5.3 Discussion

Multifractality seems able to discriminate the QS from the NQS. Although the most known spectral features outperform the multifractal parameters, it should be taken into account that the number of spectral features ($d = 72$) is higher than the number of multifractal moments ($d = 24$) or the Hurst differences ($d = 8$). However, if the multifractal features are combined with the spectral ones, there is an increase in AUC. Interestingly enough, a marked influence of the age was observed in the performance of the classifiers using the multifractal features, especially when using the input feature set $\mathbf{X}_{c_2,\Delta h}$. The mentioned input set contains features that are related to the width of the SS, which represents the degree of multifractality of the signal (the wider the SS, the more multifractal the signal is). This result indicates that the separation between NQS epochs and the QS epochs based on multifractality in neonatal EEG is more apparent in the oldest group. A possible explanation is that the discontinuous trace is present both in QS and NQS epochs in the youngest premature neonates, while the duality of the two states is well defined above 31 weeks PMA [8]. At full-term age, the trace-alternant replaces QS. These results are in agreement with the clear pattern of fractal dimension changes throughout sleep stages in full-term neonates reported in [3]. Furthermore, Piryatinska [208] showed that the features that better discriminate the sleep stages in preterm neonates are different from the ones in the full-term ones. Unlike Accardo [3] and Piryantiska [208], we observed a wider number of Hurst exponents or singularities for QS, which could indicate that the EEG signals are more complex or, at least, characterized by more fractalities or local singularities during QS. However, in the youngest group recordings, the NQS has approximately the same amount of Hurst exponents, as during QS, and they can be regarded as equally complex.

The maturation charts in Figure 5.4 also confirm that there are more local

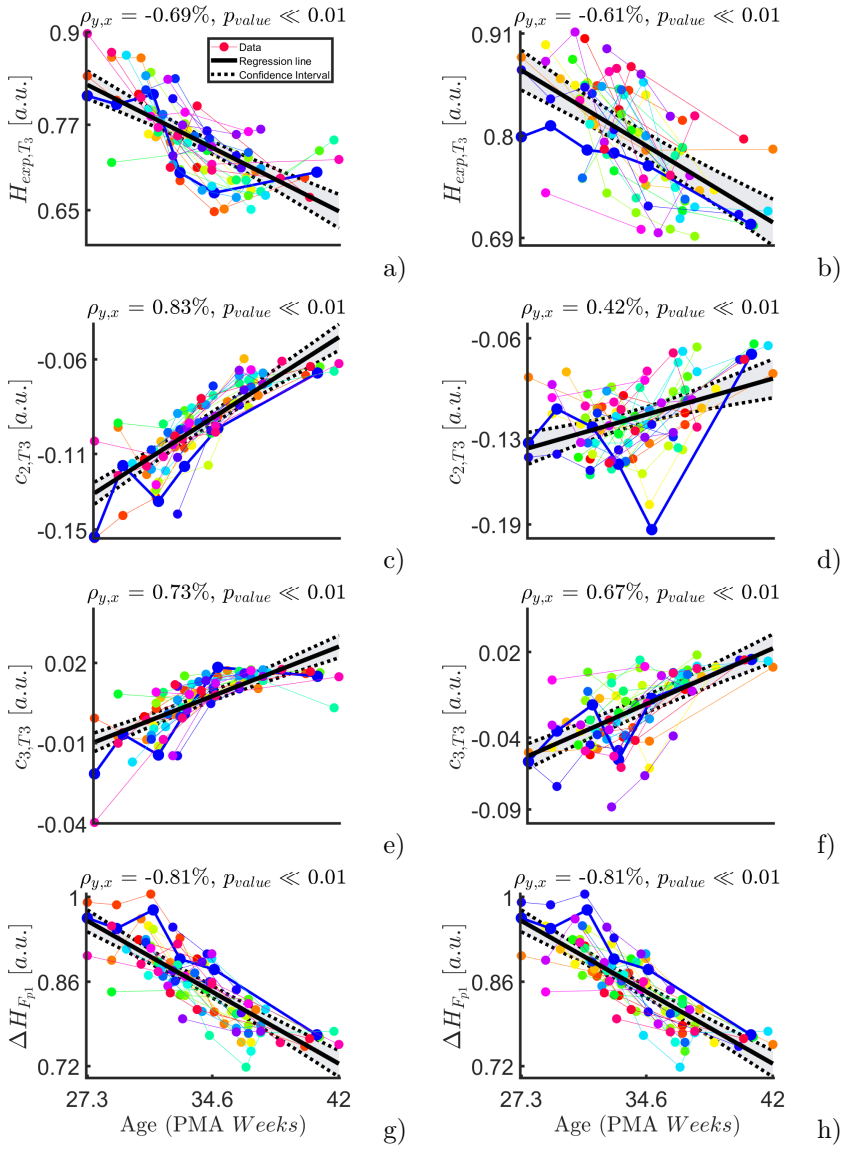


Figure 5.4: Results for the linear-mixed effect regression that models the relationship between EEG multifractal features and the post-menstrual age. Specifically, the Hurst exponent for T_3, the c_2 and c_3 for the same channel and the difference ΔH for the channel Fp_1 (H_{exp,T_3},c_{2,T_3},c_{3,T_3} and ΔH_{Fp_1}) are displayed. The left column reports the result for the non-quiet sleep (NQS), while the right column reports the result for the quiet sleep (QS) epoch. The symbol ρ is the correlation coefficient of the regression and p_{value} represents the significance of the correlation.

singularities during QS, especially throughout the development. The parameter c_2 has very similar median value in both states below 30 weeks PMA, but the reduction of the singularity spectrum width is faster during NQS than QS. The net effect is not simply an ability to discriminate sleep states from 30 weeks onward, but the prediction of brain maturation has higher performance during NQS ($MAE_{NQS} = 1.38\ wks$ vs $MAE_{QS} = 1.96\ wks$, see Table 5.2). Similarly to c_2, the other two c_p parameters do not only show slower changes with increasing age (correlation $\rho_{NQS} = -69\%$ vs $\rho_{QS} = -61\%$ in case H_{exp} and $\rho_{NQS} = 73\%$ vs $\rho_{QS} = 67\%$ in case of c_3), but the y-axis confirms the amplitude of all the three c_p parameters of the singularity remains higher in QS throughout the development. Those results are complementary to the studies of [65] and [271], that shows a general increase of entropy and the complexity in the neonatal EEG with development and a sustained higher value of complexity during QS with maturation. As reported in Section 3.1.2 and Section 3.1.3, a dysmature EEG is a process which have both stochastic and fractal properties. Based on our findings and the findings of De Wel et al. [65], an increase of complexity and randomness seems to be matched with a decrease in regularity. Monofractal signals like the Brownian motion or the white Gaussian noise have one single H_{exp} with an ideal null width [2]. In real-life examples, a signal can get closer to one single and global Hurst exponent with a narrow width of the $D(h)$. In particular, if the H_{exp} becomes approximately 0.5, the signal resembles a white Gaussian noise [121]. Therefore, the neonatal EEG becomes less multifractal and more similar to a white noise with development. However, this is more evident in NQS, which is the first state to become a continuous tracing [8]. This decrease in multifractality and regularity is also reflected by Table 5.3 and Table 5.2. The dominant group of features is represented by the width of the singularity spectrum (either as c_2 or ΔH) in NQS and the ΔH group of features outperform the main H_{exp} in both states. However, the best regression performance is achieved in combination of the two sets. The regression performance is comparable to or outperform other age model reported in literature [191],[65],[145],[144],[192].

In summary, multifractal features can be used to detect sleep states and predict brain maturation. In case of automatic sleep staging, the development of the neonatal EEG has a clear impact on the performances of the classifiers for QS detection in premature neonates, especially when multifractal features are used. This might be caused by the morphological changes in the EEG activity that evolves from a burst-like waveform to a more continuous trace during NQS epochs. This is also reflected by the quantitative assessment of brain maturation of the EEG.

5.4 Summary

In this chapter, the multifractal formalism, as defined by [266], was presented to analyze neonatal EEG in order to detect QS epochs from NQS epochs in premature infants as well as to assess brain maturation. Although the multifractal features did not outperform the classical spectral features, they could explain the effect of maturation on the classification performances. This investigation suggested that the changes in performance can be attributed to the changes in multifractal behavior of the neonatal EEG. In particular, the width of the singularity spectrum and the amount of singularities reduced in NQS compared to QS epochs. However, this difference was less clear in premature infants close to birth and explains why the detection of QS is possible only form 32 weeks PMA onward by means of EEG [8]. In support to that, the multifractal parameters could estimate the age of the patient in both sleep states, but the best performance is achieved during NQS. The signal became less regular and multifractal with the development, but this decrease was highly pronounced in NQS. The regression performance was comparable to other maturity indices reported in the literature ($R^2_{NQS} = 0.83$, $MAE_{NQS} = 1.38$ wks,$R^2_{NQS} = 0.78$, $MAE_{QS} = 1.96$ wks). These results support the concept that EEG generation mechanisms are highly non-linear and the description of both sleep staging and maturation is not a trivial problem.

Chapter 6

The maturation of the autonomic nervous system in premature infants

This chapter has been submitted for publication as Lavanga M., Heremans E., Moeyersons J., Bollen B., Jansen K., Ortibus E., Van Huffel S., Naulears G., Caicedo, A. (2020). "Maturation of the autonomic nervous system in premature infants: estimating development based on heart-rate variability analysis". Lavanga M. has co-developed the methodology, conducted the experiments and has written the manuscript. Compared to the submitted manuscript, minor textual and notational changes have been implemented for better integration in this thesis.

This chapter introduces the first model to describe the development of premature infants' autonomic nervous system (ANS) based on a quantitative analysis of the heart-rate variability (HRV). Additionally, the specific role of heart-rate drops, known as bradycardias, was examined in relationship to common clinical indices. The HRV data was obtained from the ECG recordings of the same 25 preterm infants with normal developmental outcome at 2 years reported in the previous chapters. The total number of recordings was 74. After a bradycardia-based segmentation, a variety of temporal, spectral and fractal indices were investigated to describe the changes of HRV with the development and estimate the age of the patient. Three main novelties can be reported. First, the obtained maturation models based on HRV have comparable performance to other development models in the previous chapters. Second, the selected features for age prediction show a

predominance of power and fractal features in the very-low and low frequency bands in explaining the infants' sympatho-vagal development from 27 PMA weeks until 40 PMA weeks. Third, bradycardias might disrupt the relationship between common temporal indices of the tachogram and the age of the infant and the interpretation of sympathovagal indices. This approach might reveal better insights in the investigation of postnatal maturation. The results of this chapter are reported in [147].

6.1 Introduction

Since the preterm neonatal population is at higher risk for development disorders that can lead to adverse outcome [13], the investigation of maturation via multiple physiological biomarkers is part of the clinical practice to prevent lower cognitive, motor or language outcomes later on in life [132], [89]. A common probe to inspect the development of the neurovegetative functions or *Autonomic Nervous System* (ANS) is the heart-rate fluctuation, simply known as *Heart-rate variability* (HRV).

The guidelines of the adults HRV task force clearly specify the association between the different frequency tones of the tachograms and the stimulation of the ANS branches [46]. The contribution of the sympathetic branch is represented by the low-frequency band (LF, $[0.04-0.15]$ Hz) of the HRV, while the high-frequency band (HF, $[0.15-0.5]$ Hz) reflects the parasympathetic branch. The sympathovagal balance can be expressed by the power ratio of the two frequency bands $\left(\frac{LF}{HF}\right)$, while the very-low frequency band (VLF, $[0-0.04]$ Hz) is usually associated to thermal and hormonal regulation. On the contrary, the fetal and preterm HRV frequency bands are still subject of an intensive discussion in the literature. The early exposure to the ex-utero environment induces an aberrant sympathetic response and delays autonomic maturation [123],[239]. The association between the common HRV frequency bands and the sympathovagal regulation is far less documented in infants and fetuses [71]. Other factors are known to play a role in the definition of the oscillations of the heart-rate, such as intermittent breathing cycle with high respiratory frequency and the actual delay in maturation of the autonomic nervous system. Therefore, Doret et al. [71], David et al. [59] and Hoyer et al. [111] suggested that new ways to investigate the sympathovagal balance should be examined. Since the fetal heart-rate is characterized by a strong slow-wave baseline, David et al. redefine the frequency bands for fetuses as follows: $VLF = [0.02-0.08]$ Hz, $LF = [0.08-0.2]$ Hz, $HF = [0.2-3]$ Hz. While adults normally presents an HRV spectrum with two clear peaks at HF and LF [46], infants and fetuses have a $1/f$ spectrum up to 0.1 Hz [129]. Consequently, the full description of the

preterm ANS has to consider all the possible frequency bands (VLF,LF,HF) [50],[160],[58],[166]. This could explain why the $\frac{LF}{HF}$ ratio can give contradictory results: Krueger et al. [136] did not find any specific change in this ratio in a longitudinal study with preterm patients, while Longin et al. [160] found a decrease in $\frac{LF}{HF}$ from preterm to term age. The rapid development and the unclear definition of the sympathovagal frequency bands might not give a simple interpretation of $\frac{LF}{HF}$ as it is for adults. Surprisingly, infants show bigger changes in the absolute power of the three main bands VLF, LF, and HF than relative power [160]. Hoyer et al. [111] argued that predominant principles of autonomic development are not only an increase in heart-rate variability, but also increasing complexity and pattern formation. Consequently, HRV indices can be chosen to reflect these principles in order to describe the sympathovagal balance maturation. Pattern formation can be described by tachogram skewness and the new ratio $\frac{VLF}{LF}$, while the increasing complexity is characterized by an increasing HRV entropy. It should also be stressed that the computation of power ratios, such as $\frac{LF}{HF}$, requires stationarity, which can be questioned in case of infants heart-rate time series. Therefore, Abry et al. [2] and Doret et al. [71] proposed fractal analysis as an alternative method to investigate the sympathovagal balance in fetal heart-rate. It focuses on quantities such as oscillations or increments at different scales to tackle the absence of stationarity and determines a specific relations between the fractal exponents (such as the Hurst Exponent) and the $\frac{LF}{HF}$ ratios. However, those methods were never applied on premature infants.

One example of non-stationarity is the presence of bradycardias. These are normally heart-rate drops below 70% of the heart-rate average, which last at least for 4 s and may be associated with apneas. These drops can alter oxygen saturation and blood flow putting organs at risk of damage [195]. In general, apneas and bradycardias can be uncorrelated, but apneic spells that occur with bradycardias are most likely to affect brain homeostasis. In addition, apneas and bradycardias are the probable consequence of the immature respiratory system, and the apneas' occurrence tends to decrease with increasing age [12] [212]. The bradycardia can then be considered a consequence of heart-rate disregulation, which can disrupt the state-space and the probability density function of the tachogram [92]. Any proper model that tries to describe the development of the infants' ANS has to include not simply the slow variation of the basal heart-rate, but the sudden drops of the tachogram. Those nonstationary events possibly affect the most common HRV temporal or spectral features used in clinical practice and can bias conclusions of the medical community. For example, bradycardias can forcefully increase the variability of the tachogram or its regularity.

In order to address the shortcomings by the studies outlined above, a new

framework to describe autonomic maturation in healthy preterm babies is provided in this chapter. This research can be divided in two main strands. First, both spectral analysis and multifractal analysis are employed to investigate the neurovegetative development of the sympathovagal balance and its complexity and track maturation. Second, the impact of bradycardias on ANS maturation is investigated. This chapter tries to provide a complete overview of autonomic maturation in premature neonates, including non-stationary events such as bradycardias. The final clinical objective is to provide novel maturation charts for the premature autonomic nervous system in the first weeks of life and correct the effect of heart-rate events on common clinical HRV indices. Those normative charts might be used as reference to investigate early-life and ex-utero factors that can deviate from a normal premature development and define suitable therapies in the neonatal intensive care unit.

6.2 Methods

6.2.1 Dataset

The dataset consists of electrocardiograms (ECG) of 25 preterm infants, that were recorded at the Neonatal Intensive Care Unit of the University Hospital of Leuven. It was collected in a multimodal setting for another research study related to brain development and a sleep-stage analysis (see [67], [132] and Section 1.2). Inclusion criteria were: a normal neurodevelopmental outcome at 9 and 24 months corrected age (Bayley Scales of Infant Development-II, mental and motor score > 85), no severe brain lesions, assessed by ultrasound, and not taking any sedative or antiepileptic drugs during the EEG registration. The sampling ECG frequency was 250 or 500 Hz and the average length of the recording was 4h 44 min. An overview of the dataset is reported in Table 6.1.

6.2.2 Preprocessing

The HRV represents the instantaneous fluctuations of heart rate and is usually expressed by the tachogram which visualizes the variations of the time interval between two consecutive R-peaks (RR intervals, RR_i). In order to compute a RR_i time series, the R peaks of the ECG have been detected via the Matlab toolbox by Moeyersons et al. [172], which is based on enveloping procedure. This graphical user-interface also allows for correction and deletion in case of erroneous R-peaks. In case of a single missing R-peak, the value was replaced by using the formula:

Table 6.1: Demographics of the 25 patients: average duration of the tachogram in minutes (Duration_{Rec}), average duration of the annotated bradycardias in s (Duration_{WB}),average number of the annotated bradycardias (Number_{WB}), average RR amplitude during the bradycardia in ms (RR_{WB}), postmenstrual age in weeks (PMA) and gestational age in weeks (GA). The last three rows represent the number of recording for each age subgroup (below 32 PMA weeks, between 32 and 36 weeks and above 36 weeks).

Number of patients = 25	
Duration_{Rec} (min)	208.435 ± 115.657
Duration_{WB} (s)	18.881 ± 8.332
Number_{WB}	7 ± 12
RR_{WB}	631.405 ± 78.677
PMA (wks)	33.689 ± 3.049
GA (wks)	28.315 ± 2.318
$PMA \leq 32\ wks$	22
$PMA \in (32 - 36]\ wks$	35
$PMA > 36\ wks$	17

$$\hat{R}_t = \frac{R_{t-1} + R_{t+1}}{2}, \tag{6.1}$$

where $\hat{R}_t$ is the estimated position of the missed R-peak, while R_{t-1} and R_{t+1} are the location of the previous and following R-peak. In case of two or more missing R-peaks due to ECG flat lines or muscle artifacts which made the QRS detection impossible, the contaminated parts of the signal were discarded. In case that less 20 minutes of noise-free signal remained, the signal was discarded. All included recordings had at least 50 minutes of available data for a total of 74 recordings.

Besides the preprocessing of artifacts and before the feature extraction, we also dealt with the sudden drops of heart-rate, known as bradycardias. Although those phenomena are completely natural in the developing infant, they can suddenly increase the frequency content of the RR_i series. Therefore, traditional linear spectral and temporal analysis might not be suitable since the instantaneous variance and mean of the heart-rate can vary over time, as explained in detail by [92]. According to the same study, the heart-rate activity that precedes sudden drops might differ from the drops itself and other bradycardia-free periods. Consequently, bradycardias have been detected in

the current studies before any further processing. Based on the definitions of apnea of prematurity and bradycardias by [195], the bradycardia spells were detected as sudden RR_i increases above $1.5*\overline{RR_i}$ that persist for more than 4 s, where $\overline{RR_i}$ is the median tachogram of the entire recording. Subsequently, three different windowing strategies were applied:

1. Post-bradycardia (PB) windowing: the immediate 10 minutes after the bradycardic event were considered a candidate for features' extraction. It is important to remember this window did not include the bradycardia itself.

2. Between-bradycardias (BB) windowing: all non-overlapping 10 minute windows contained between bradycardic events were considered as candidate epochs for features' extraction. The first viable window was at least 10 minutes away from the bradycardic event in order to guarantee that the signal was stabilized.

3. Within-bradycardia (WB) windowing: a 10 minute window was considered from the moment that the signal exceeds the threshold $1.5*\overline{RR_i}$ in the tachogram, i.e. the supposed start of the bradycardic event. This windowing should involve both the information related to the heart-rate drop and the recovery period.

A visual description of the windowing scheme is reported in Figure 6.1. The grey dashed boxes highlight the three type of windows (WB,BB,PB) that can be determined in a single trace, while the dot-dash box shows typical bradycardia events.The average duration and amplitude of a bradycardia event is reported in Table 6.1, which also shows the average number of bradycardias in the entire dataset. Some of the recordings did not have any heart-rate drop according to the reported definition. Therefore, the windowing scheme based on bradycardias was not applicable. In this specific case, the design choice was a segmentation in non-overlapping 10 minutes windows and assign the results of feature extraction to post-bradycardia windowing scheme (see Figure 6.2).

6.2.3 Feature extraction

In each of the windows defined according to the PB,BB and WB schemes, a set of temporal, spectral and fractal features were derived to describe the autonomic nervous system of the premature infants and its relationship with development. An overview of the different attributes is reported in Table 6.2.

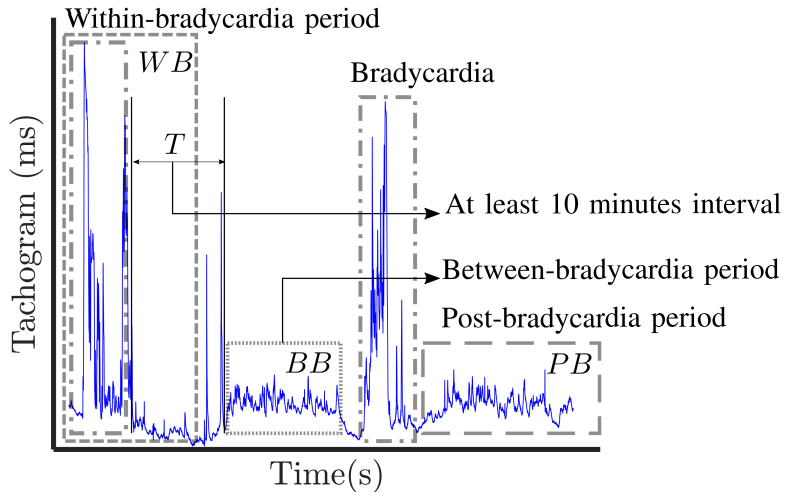


Figure 6.1: Visual representation of the three windowing schemes applied in this investigation: postbradycardia scheme (PB), between-bradycardia scheme (BB) and within-bradycardia scheme (WB). The selected windows in each trace are indicated with grey dashed boxes, while the dot-dashed boxes show examples of annotated bradycardia. In case of BB windowing, a period T greater that 10 minutes is present between the end of the bradycardia and the first available window.

Temporal Indices

Based on the most common guidelines related to HRV processing [123] and [46], the first and the second order moments of the RR_i, i.e. the mean of the tachogram (μ_{RR}) and the standard deviation (σ_{RR}), were computed in order to assess the level of the variability.

Spectral analysis

The sympathovagal activity is normally assessed by the computation of the spectral power in the different HRV frequency bands [46]. Unlike adults, the premature infants have a higher mean heart-rate with very slow oscillation around it [59],[50]. Therefore, the frequency bands of the premature patients were defined as follows: $VLF = [0, 0.08]\ Hz$, the low-frequency $LF = [0.08, 0.2]\ Hz$ and high frequency $HF = [0.2, 3.0]\ Hz$. Additionally, the RR time series of the premature infant can be nonstationary due to a series of events, like

bradycardias or other heart-rate disregulation (see section 2.3). Therefore, the power spectral density was computed with time-frequency (TF) methodologies reported in section 3.1.1: the Welch's periodogram, the quadratic smoothed pseudo Wigner-Ville distribution (SPWD) [187] and the continuous wavelet transform (CWT) [59].

In particular, one can compute the absolute power in three bands VLF, LF and HF as the reported integral in (3.6). Besides, the ratios $\frac{VLF}{LF}$ and $\frac{LF}{HF}$ were also computed alongside two indices to represent the normalized LF power: $\frac{LF}{VLF+LF}$ and $\frac{LF}{HF+LF}$. In case of Welch's algorithm, there is no dependency from the time variable t. On the contrary, CWT and SPWD generates a time series for each selected frequency band, as highlighted by (3.1),(3.4) and (3.6). In order to obtain one single value for each window, the median of this time-series was taken into account. The set of spectral features derived for each methodology is reported in the central block of Table 6.2.

6.2.4 Multifractal analysis

Since spectral analysis requires stationarity of data and the very definition of the tachogram series frequency bands have been questioned, the HRV was also analyzed according to the fractal or multifractal paradigm. As shown in [71], the infant's tachogram is a fractal or scale free signal, which presents a power-law decay spectrum as follows:

$$S_{RR}(f) = |C| f^{-2(H_{exp}-1)} \tag{6.2}$$

where H_{exp} is known as the Hurst exponent and controls the decay of the power function. H is also a representative parameter for fractal time series and there can be more than one exponent for each signal. A signal with one single exponent is commonly known as monofractal, while a signal with multiple exponents h is known as multifractal [118]. Small values of h represent sharp and transient regularity or singularity, while large values represent smooth changes [151].

An efficient method to determine the amount of exponents or singularities h is the multifractal formalism reported in Section 3.1.3. Thanks to the partition function $Z(a, q)$, the scaling exponents $\tau(q)$ (SE) associated to this decay can be obtained by computing the slope of Z versus the scales in a log-log diagram from a certain scale $a_1 = 2^{j_1}$ to a certain scale $a_2 = 2^{j_2}$. The log-transform clearly shows the advantage to define scales as power quantities.

The multifractality parameters (H_{exp}, C_2) were computed in the entire non-overlapping window according to three schemes discussed in Section 6.2.2. Similar to Chapter 5, the multifractal features were derived using the Wavelet p-Leader and Bootstrap based MultiFractal analysis (PLBMF) MATLAB toolbox, described in [266]. This toolbox can be downloaded from `https://www.irit.fr/~Herwig.Wendt/software.html`. A fundamental design choice is the scale range $[2^{j_1}, 2^{j_2}]$ from which the exponent $\tau(q)$ is estimated from (3.12) [2], [266]. In case of HRV, the exponents $[j_1, j_2]$ are normally set equal to $[3, 12]$. Given the fact that the scale can be written as $a = 2^j = (f_s/2)/f$ with f_s as sampling frequency of the signal, it follows that the range $[j_1, j_2] = [3, 12]$ approximately represents the frequency band $\approx [0.375, 0.001]$ Hz with $f_s = 6$ Hz. In case that $[j_1, j_2] = [5, 12]$, the chosen scale range approximately represents the frequency band $\approx [0.094, 0.001]$ Hz. It is clear the first range considers part of the HF band, while the latter solely focuses on the combination of LF and VLF. Since the chosen scale range might influence the multifractal attributes, both ranges were tested to investigate which frequency bands mostly reflects the sympathovagal balance. In fact, the main Hurst exponent H_{exp} or C_1 parameter is able to influence the ratio $\frac{LF}{HF}$. Based on (6.2), one can rewrite the spectral ratios as:

$$\frac{LF}{HF} = \frac{\int_{f_m}^{f_I} S_{RR}(f)df}{\int_{f_I}^{f_M} S_{RR}(f)df} = \frac{(f_I^{2-2H_{exp}} - f_m^{2-2H_{exp}})}{(f_M^{2-2H_{exp}} - f_I^{2-2H_{exp}})} \tag{6.3}$$

where $[f_m, f_I]$,$[f_I, f_M]$ represent the frequency bands of LF and HF. Taking into account that the Hurst exponent and the $\frac{LF}{HF}$ are related and taking also into account that the chosen $[j_1, j_2]$ decides which frequency bands the multifractal parameters are related to, the scales investigation of the fractal properties can shed a light which bands mainly reflect the sympathovagal activity. The set of fractal features derived for each methodology is reported in last block of Table 6.2.

6.2.5 Algorithmic pipeline and statistical analysis

The processing pipeline to estimate the age of the infant is reported in Figure 6.2. For each HR time series, the signal was split according to the PB,BB and WB windowing scheme reported in Figure 6.1 and all the features reported in Table 6.2 were extracted. Besides artifact removal, a fundamental preprocessing step is the resampling of the tachogram. The behavior of nonlinear features can depend on the sample rate, as also shown by the definition of the scales and their range for the multifractal parameters (Section 6.2.4). Based on the findings by

[37], the following sampling frequencies were tested for fractal indices: $[6, 8, 12]$ Hz. In contrast, the sampling frequencies for the spectral and temporal indices was set to 6 Hz in order to include the higher respiratory frequency of premature infants [46],[123].

As described in Section 6.2.3, the tuning and designed parameters for the spectral and fractal analysis were chosen in accordance to the absence of stationarity and the persistent slow-wave baseline of the premature HRV signal. The necessity to investigate long-range fluctuations and a recovery period after events such as bradycardias justify the segmentation in 10 minutes. Normally, time-frequency approaches use windows longer than 600 s to describe evolution in HRV spectrum [268],[187] and the fractal indices also requires windows of this size to fully investigate changes in regularity [2],[71]. Additionally, the BB and WB schemes can generate a set of windows and therefore an array of features based on the number of bradycardias present in each recording (on average, 7 bradycardias per recording, as reported by Table 6.1). In order to obtain one representative value for each recording in each windowing scheme, the median of this array of attributes over the different windows was computed for each recording, as highlighted by the grand-median block in Figure 6.2.

After the features' extraction process and the grand-median step, three datasets were then obtained according to the three different windowing schemes. The number of features extracted for each dataset was then 27 in total: 21 for the spectral attributes, 2 for the temporal ones, 4 for fractal indices, as shown in Table 6.2.

In order to investigate the ANS maturation, the HRV features were used to estimate the PMA of the patient, as shown by the last block of the diagram in Figure 6.2. Since the PMA is known for each recording, a linear mixed effects (LME) regression model was developed for each dataset with PMA as response variable [149]. The actual regression consisted of two steps. First, the features were selected via the least absolute shrinkage and selection operator (LASSO) due to the high number of features, after that the absolute power features were log-transformed (see Section 3.3.3). Specifically, the LASSO was repeated for 20 iterations on the entire dataset and the features which were selected in more than 40% of the total number of iterations (8 iterations out of 20) were included in the regression model [149]. Second, a linear mixed-effect regression model was built with the selected subsets with multiple random splits of the data. In particular, the dataset was split into 70% training set and 30% test set for 20 iterations and the model was developed on the train set and tested for test set for each iteration. The performance was then assessed as *mean absolute error* MAE on the test set, as well as explained variance R^2, both on train and test set (R^2_{train}, R^2_{test}). The results were reported as median(IQR) (where IQR stands for *InterQuartile Range*) over the 20 iterations. A linear-mixed

effect model requires the definition of a grouping variable which introduces the random effect and the patient ID was taken as grouping variable since a set of one or more recordings belong to a patient (as discussed in a previous study [149]) . Furthermore, the LME regression with the LASSO procedure was not simply examined for the entire subset of features, but also for the three subset feature groups: temporal, spectral and fractal attributes. In case of temporal features regression, the LASSO step was not performed.

On top of that, the trend for the ANS features throughout the patients' development was also reported as median(IQR) in three age groups (PMA $\leq$ 31 weeks, PMA $\in (31-36]$ weeks, PMA $>$ 36 weeks) as well as Pearson correlation coefficient with PMA.

Table 6.2: Overview of all computed features. Besides the first and the second order moments, the table reports all the relative and absolute power features in the frequency domain, namely the power in VLF, LF and HF bands, the ratio between the VLF and LF bands and between LF and HF bands and the normalized LF band power with respect to VLF band and HF band. The spectral features are reported for each the PSD estimation approaches that were used in this investigation. The last section reports the computed multifractal features, i.e. Hurst exponent (H_{exp}) and the (C_2) for the investigated scale ranges: $[j_1, j_2] = [3, 12]$ and $[j_1, j_2] = [5, 12]$.

Temporal features	
Statistical moments	μ_{RR}, σ_{RR}
Spectral features	
Welch	$P(VLF)$,$P(LF)$,$P(HF)$, $\frac{VLF}{LF}$,$\frac{LF}{HF}$,$\frac{LF}{LF+HF}$,$\frac{LF}{LF+VLF}$
SPWVD	$P(VLF)$,$P(LF)$,$P(HF)$, $\frac{VLF}{LF}$,$\frac{LF}{HF}$,$\frac{LF}{LF+HF}$,$\frac{LF}{LF+VLF}$
Wavelet	$P(VLF)$,$P(LF)$,$P(HF)$, $\frac{VLF}{LF}$,$\frac{LF}{HF}$,$\frac{LF}{LF+HF}$,$\frac{LF}{LF+VLF}$
Fractal features	
Multifractality	$H_{exp,[j_1,j_2=\mathbf{5},12]}$, $C_{2,[j_1,j_2=\mathbf{5},12]}$, $H_{exp,[j_1,j_2=3,12]}$,$C_{2,[j_1,j_2=3,12]}$

6.3 Results

The overview of the dataset is reported in Table 6.1, which shows certain traits of the annotated bradycardias. The mean length is around 18 s. On average,

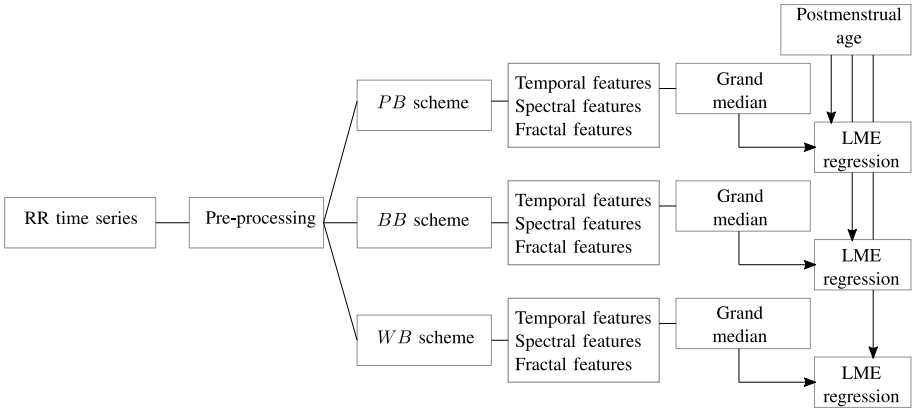


Figure 6.2: The block diagram shows the main steps of the age estimation. For each RR signal, artifact preprocessing is performed and associated resampling of the tachogram. The signal is split in different windows according to the scheme of Figure 6.1. For each of these epochs, temporal, spectral and fractal features, which undergo a grand-median process if there is more than one epoch per scheme. The three datasets are then used to estimate the age of the recording in a linear mixed effects (LME) regression.

there are 7 bradycardias per recording and the mean RR_i during bradycardias is approximately 631 ms. The overview shows that the infants have an average PMA of 34 weeks and 35 recordings are collected in the range (32-36] weeks. A total of 22 recordings is collected in the first days of life, while a set of 17 recording has been included from the weeks close to discharge.

Figure 6.3 and Table 6.3 report the trends in the three different windowing schemes for the following features: the mean μ_{RR} and standard deviation of the HRV σ_{RR}, the absolute power in the LF band (and its logarithmic transform), the relative $\frac{LF}{LF+VLF}$ power, the Hurst Exponent in the range $[j_1, j_2] = [5, 12]$ $H_{exp_{[j_1,j_2=5,12]}}$ and the width of the singularity spectrum in the same range $C_{2_{[j_1,j_2=5,12]}}$. The overview of all features for all different windowing schemes (PB,BB and WB) are reported in the supplementary Tables A.1, A.2 and A.3. Figure 6.3 reports the results for the windowing scheme for the within-bradycardia epochs on the left column, while the results for the between-bradycardia epochs are shown on the right column. The power in the LF band and the relative LF power ($\frac{LF}{LF+VLF}$) increase with increasing PMA in both scenarios: in particular, Pearson correlations ρ_{xy} are respectively 69% (72% with logarithm transform) and 64% for bradycardia epochs, while ρ_{xy}

Table 6.3: The main temporal, spectral and fractal features are reported for the three windowing schemes (PB, BB and WB). The results are reported as median(IQR). IQR stands for *interquartile range*. The fractal indices are reported for $f_s = 8\ Hz$. The symbol ρ stands for the Pearson correlation coefficient. The symbol ** represents a significant correlation with $p \leq 0.01$ and $^{n.s.}$ is used to indicate a non-significant correlation.

Median(IQR) - PMA weeks	≤ 32	$(32-36]$	> 36	$\rho(\%)$
POST-bradycardia (PB) group				
μ_{RR}	374.65(366.38-391.36)	377.07(364.33-393.69)	387.2(374.98-416.44)	0.39^{**}
σ_{RR}	16.71(12.02-22.05)	25.5(21.65-31.1)	28.47(24.03-32.08)	0.49^{**}
$P(LF)_{Wavelet}$	2.14(0.96-3.68)	4.17(2.16-8.9)	17.74(4.94-25.51)	0.69^{**}
$\frac{LF}{LF+VLF}_{Wavelet}$	5.38(3.45-8.7)	4.9(3.67-12.17)	14.06(11.77-18.09)	0.57^{**}
$H_{exp,[j_1,j_2=5,12]}$	0.61(0.52-0.7)	0.55(0.45-0.59)	0.5(0.44-0.56)	-0.47^{**}
$C_{2,[j_1,j_2=5,12]}$	-0.2(-0.26 - -0.17)	-0.19(-0.21 - -0.13)	-0.14(-0.15 - -0.11)	0.45^{**}
BETWEEN-bradycardia (BB) group				
μ_{RR}	370.51(359.96-388.36)	377.42(363.11-389.25)	394.93(370.01-427.45)	0.47^{**}
σ_{RR}	13.89(10.97-18.49)	19.81(15.72-23.82)	29.1(21.99-30.66)	0.64^{**}
$P(LF)_{Wavelet}$	1.3(0.86-3.38)	4.23(1.93-6.28)	11.34(8.4-15)	0.71^{**}
$\frac{LF}{LF+VLF}_{Wavelet}$	6.93(4.89-8.53)	7.9(4.63-11.05)	12.57(8.75-13.95)	0.48^{**}
$H_{exp,[j_1,j_2=5,12]}$	0.6(0.52-0.68)	0.54(0.5-0.59)	0.48(0.45-0.52)	-0.5^{**}
$C_{2,[j_1,j_2=5,12]}$	-0.19(-0.23 - -0.14)	-0.17(-0.2 - -0.14)	-0.09(-0.12 - -0.08)	0.43^{**}
WITHIN-BRADYCARDIA (WB) group				
μ_{RR}	384.9(369.62-398.91)	384.16(369.2-397.51)	389.12(377.65-425.8)	0.37^{**}
σ_{RR}	38.31(32.22-44.62)	40.61(32.5-49.81)	35.89(28.43-40.74)	$-0.04^{n.s.}$
$P(LF)_{Wavelet}$	2.3(1.12-4.03)	4.68(2.8-9.86)	18.66(5.35-26.46)	0.69^{**}
$\frac{LF}{LF+VLF}_{Wavelet}$	2.57(1.13-3.97)	3.34(2.19-8.97)	10.96(6.69-16.78)	0.64^{**}
$H_{exp,[j_1,j_2=5,12]}$	0.61(0.49-0.71)	0.55(0.43-0.62)	0.49(0.43-0.52)	-0.45^{**}
$C_{2,[j_1,j_2=5,12]}$	-0.26(-0.3 - -0.21)	-0.21(-0.24 - -0.17)	-0.13(-0.18 - -0.11)	0.54^{**}

are respectively 71% (71% with logarithm transform) and 48% for between-bradycardia windows. Concerning the post-bradycardia period, the Table 6.3 shows a Pearson correlation of 69% for the power in LF band and 57% for $\frac{LF}{LF+VLF}$. Results are here reported for the wavelet approach, but the other spectral methodologies exhibit similar trends (see Tables A.1, A.2 and A.3). In addition, the Hurst exponent (derived as the c_1 of the singularity spectrum) decreases with development (ρ_{xy} are -45% in the bradycardias scenario, -47% in post-bradycardia scenario and -50% in the between-bradycardia scenario), while the width of the singularity spectrum (c_2 parameter) increases with increasing PMA (ρ_{xy} are 54% in the bradycardia scenario, 45% in the post-bradycardia scenario and 43% in the between-bradycardia scenario). The greatest contrast was found with the variability of the heart-rate, σ_{RR}. While the standard deviation increases with infants' maturation in the between-bradycardia epochs, the σ_{RR} does not increase with age within the bradycardic event. Moreover, it is higher in the bradycardia epochs than in the between-bradycardia scenario ($\rho_{bradycardia} = -4\%$ vs $\rho_{between} = 64\%$ with $p_v = 0.77$ vs $p_v \leq 0.01$). The multifractal parameters are reported for $f_s = 8\ Hz$ in Table 6.3. A full overview

of the effect of the sampling frequency on the fractal indices is reported in Table 6.4.

Table 6.5 shows the regression results for the linear mixed-effect models, while Table 6.6 reports the features selected by LASSO. Those two tables report the results for the three different windowing schemes (PB,BB,WB) in three different blocks, while the rows report the results for the different feature groups (temporal, spectral and fractal attributes) and sampling frequencies. The different columns respectively report the explained variance in the training set (R^2_{train}), the mean absolute error (MAE) and the explained variance in the test set (R^2_{test}). The best performance is reached for the combination of all features at $f_s = 12\ Hz$ in the PB scheme ($R^2_{train} = 0.75$, $MAE = 1.83\ weeks$, $R^2_{test} = 0.57$) as well as between bradycardias ($R^2_{train} = 0.68$, $MAE = 1.56\ weeks$, $R^2_{test} = 0.59$). During the bradycardia event (WB), the best performance is achieved with the spectral features ($R^2_{train} = 0.73$, $MAE = 1.9\ weeks$, $R^2_{test} = 0.62$). Table 6.6 shows that the selected features are the absolute spectral power in LF and VLF together with C_2 parameter in the range $[j_1, j_2] = [5, 12]$ for the first two schemes. For the WB scheme, the selected feature is simply the power in LF band.

Figure 6.4 shows that the relationship of (6.3) between the H_{exp} and the ratios $\frac{VLF}{LF}$ and $\frac{LF}{HF}$. The first row shows the relationship between $\frac{VLF}{LF}$ and $H_{exp,[j_1,j_2=5,12]}$ in the three windowing schemes: WB in magenta circles, PB in light-blue squares and BB in indigo diamonds. The Pearson correlation coefficients are respectively 21%, 49% and 43%. The second row shows the relationship between $\frac{LF}{HF}$ and $H_{exp,[j_1,j_2=3,12]}$ in the same three schemes. The Pearson correlation coefficients are respectively 18%, 20% and 36%.

6.4 Discussion

This chapter provides an overview of autonomic nervous system maturation in preterm infants and aims to estimate the postmenstrual age of the infants based on the HRV. Since the neonatal tachogram is signal characterized by lack of stationarity and strong slow-wave baseline [111],[2],[71], the reported study investigated the maturation of sympathetic and parasympathetic branches with the combination of temporal, spectral and fractal indices. Three main novel findings can be reported. First, Table 6.5 shows that the maturation of infants can be assessed with different spectral and fractal HRV indices with comparable performances to other maturation models for fetal and preterm development by [111], [65], [144], [149]. Second, Figure 6.4 reports that the spectral ratio $\frac{VLF}{LF}$ and the Hurst exponent in the range $[j_1, j_2] = [5, 12]$ are more correlated

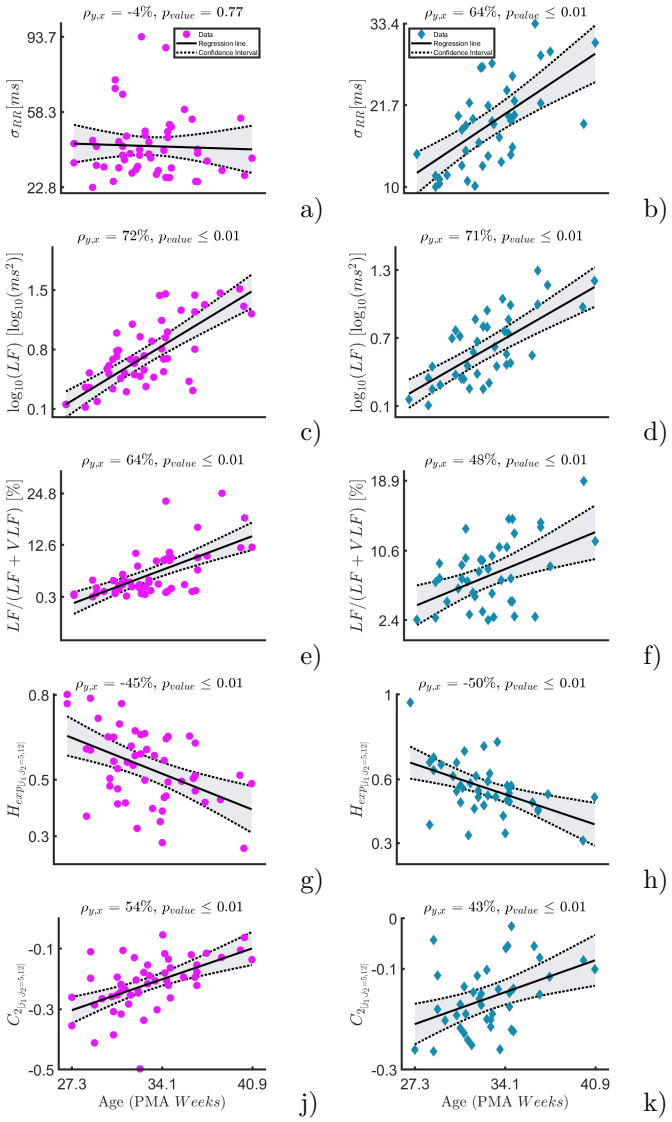


Figure 6.3: The figure shows the linear-mixed effect regressions between the post-menstrual age and the following HRV features: the standard deviation of the tachogram σ_{RR}, the absolute and the relative power in the LF band $\left(\log_{10}(LF), \frac{LF}{LF+VLF}\right)$, the Hurst exponent $H_{exp,[j_1,j_2=5,12]}$ and the parameter C_2. The sampling frequencies for the fractal indices is $f_s = 8\ Hz$. The left column - magenta circles report the results for the bradycardia epochs, while the right column - indigo diamonds the results for the between-bradycardia epochs. ρ is the correlation coefficient with the associated significance p_{value}.

Table 6.4: The fractal features are reported in three different age categories and for the investigated sampling frequencies $f_s = [6, 8, 12]\ Hz$. The results are reported as median(IQR) for the between-bradycardia and bradycardia periods. IQR stands for *interquartile range*. The symbol ρ stands for the Pearson correlation coefficient. The symbol ** represents a significant correlation with $p \leq 0.01$, and * is used for a significant correlation with $p \leq 0.05$. $^{n.s.}$ is used to indicate a non-significant correlation.

Median(IQR) - PMA weeks	≤ 32	$(32-36]$	> 36	$\rho(\%)$
Fractal features in the PB group, $f_s = 6\ Hz$				
$H_{exp,[j_1,j_2=5,12]}$	0.53(0.46-0.67)	0.49(0.41-0.58)	0.45(0.36-0.5)	-0.45**
$C_{2,[j_1,j_2=5,12]}$	-0.2(-0.25 - -0.18)	-0.19(-0.22 - -0.15)	-0.13(-0.17 - -0.1)	0.57**
$H_{exp,[j_1,j_2=3,12]}$	0.65(0.56-0.71)	0.6(0.55-0.68)	0.56(0.53-0.6)	-0.4**
$C_{2,[j_1,j_2=3,12]}$	-0.17(-0.18 - -0.13)	-0.13(-0.16 - -0.11)	-0.09(-0.12 - -0.08)	0.41**
Fractal features in the PB group, $f_s = 8\ Hz$				
$H_{exp,[j_1,j_2=5,12]}$	0.61(0.52-0.7)	0.55(0.45-0.59)	0.5(0.44-0.56)	-0.47**
$C_{2,[j_1,j_2=5,12]}$	-0.2(-0.26 - -0.17)	-0.19(-0.21 - -0.13)	-0.14(-0.15 - -0.11)	0.45**
$H_{exp,[j_1,j_2=3,12]}$	0.67(0.6-0.71)	0.66(0.59-0.69)	0.62(0.58-0.65)	-0.33*
$C_{2,[j_1,j_2=3,12]}$	-0.14(-0.16 - -0.1)	-0.11(-0.14 - -0.08)	-0.09(-0.11 - -0.09)	0.2$^{n.s.}$
Fractal features in the PB group, $f_s = 12\ Hz$				
$H_{exp,[j_1,j_2=5,12]}$	0.62(0.52-0.7)	0.56(0.49-0.64)	0.53(0.44-0.54)	-0.45**
$C_{2,[j_1,j_2=5,12]}$	-0.2(-0.23 - -0.16)	-0.17(-0.19 - -0.15)	-0.11(-0.13 - -0.1)	0.57**
$H_{exp,[j_1,j_2=3,12]}$	0.67(0.61-0.73)	0.64(0.61-0.71)	0.62(0.6-0.63)	-0.33*
$C_{2,[j_1,j_2=3,12]}$	-0.13(-0.15 - -0.11)	-0.11(-0.13 - -0.09)	-0.09(-0.11 - -0.08)	0.22$^{n.s.}$

Median(IQR) - PMA weeks	≤ 32	$(32-36]$	> 36	$\rho(\%)$
Fractal features in the BB group, $f_s = 6\ Hz$				
$H_{exp,[j_1,j_2=5,12]}$	0.55(0.45-0.65)	0.52(0.43-0.55)	0.45(0.4-0.48)	-0.43**
$C_{2,[j_1,j_2=5,12]}$	-0.19(-0.24 - -0.16)	-0.17(-0.2 - -0.13)	-0.11(-0.12 - -0.09)	0.52**
$H_{exp,[j_1,j_2=3,12]}$	0.65(0.56-0.69)	0.61(0.57-0.67)	0.55(0.52-0.59)	-0.39*
$C_{2,[j_1,j_2=3,12]}$	-0.15(-0.17 - -0.12)	-0.13(-0.15 - -0.1)	-0.08(-0.1 - -0.06)	0.39*
Fractal features in the BB group, $f_s = 8\ Hz$				
$H_{exp,[j_1,j_2=5,12]}$	0.6(0.52-0.68)	0.54(0.5-0.59)	0.48(0.45-0.52)	-0.5**
$C_{2,[j_1,j_2=5,12]}$	-0.19(-0.23 - -0.14)	-0.17(-0.2 - -0.14)	-0.09(-0.12 - -0.08)	0.43**
$H_{exp,[j_1,j_2=3,12]}$	0.68(0.61-0.73)	0.65(0.6-0.67)	0.6(0.55-0.62)	-0.36*
$C_{2,[j_1,j_2=3,12]}$	-0.12(-0.15 - -0.1)	-0.12(-0.14 - -0.1)	-0.08(-0.09 - -0.05)	0.23$^{n.s.}$
Fractal features in the BB group, $f_s = 12\ Hz$				
$H_{exp,[j_1,j_2=5,12]}$	0.62(0.52-0.68)	0.57(0.52-0.63)	0.52(0.48-0.53)	-0.43**
$C_{2,[j_1,j_2=5,12]}$	-0.18(-0.23 - -0.16)	-0.15(-0.19 - -0.12)	-0.09(-0.11 - -0.08)	0.53**
$H_{exp,[j_1,j_2=3,12]}$	0.68(0.6-0.71)	0.64(0.6-0.69)	0.59(0.54-0.64)	-0.31$^{n.s.}$
$C_{2,[j_1,j_2=3,12]}$	-0.12(-0.14 - -0.11)	-0.11(-0.12 - -0.1)	-0.08(-0.1 - -0.06)	0.26$^{n.s.}$

Median(IQR) - PMA weeks	≤ 32	$(32-36]$	> 36	$\rho(\%)$
Fractal features in the WB period, $f_s = 6\ Hz$				
$H_{exp,[j_1,j_2=5,12]}$	0.52(0.42-0.7)	0.48(0.39-0.52)	0.43(0.4-0.46)	-0.36**
$C_{2,[j_1,j_2=5,12]}$	-0.23(-0.29 - -0.2)	-0.21(-0.24 - -0.17)	-0.14(-0.18 - -0.12)	0.55**
$H_{exp,[j_1,j_2=3,12]}$	0.62(0.58-0.67)	0.6(0.53-0.63)	0.57(0.5-0.6)	-0.31*
$C_{2,[j_1,j_2=3,12]}$	-0.19(-0.23 - -0.15)	-0.16(-0.19 - -0.14)	-0.1(-0.12 - -0.09)	0.48**
Fractal features in the WB period, $f_s = 8\ Hz$				
$H_{exp,[j_1,j_2=5,12]}$	0.61(0.49-0.71)	0.55(0.43-0.62)	0.49(0.43-0.52)	-0.45**
$C_{2,[j_1,j_2=5,12]}$	-0.26(-0.3 - -0.21)	-0.21(-0.24 - -0.17)	-0.13(-0.18 - -0.11)	0.54**
$H_{exp,[j_1,j_2=3,12]}$	0.66(0.62-0.71)	0.64(0.58-0.68)	0.61(0.58-0.62)	-0.36**
$C_{2,[j_1,j_2=3,12]}$	-0.15(-0.2 - -0.12)	-0.14(-0.17 - -0.11)	-0.11(-0.12 - -0.09)	0.31*
Fractal features in the WB period, $f_s = 12\ Hz$				
$H_{exp,[j_1,j_2=5,12]}$	0.58(0.5-0.68)	0.56(0.5-0.6)	0.5(0.49-0.53)	-0.36**
$C_{2,[j_1,j_2=5,12]}$	-0.22(-0.28 - -0.19)	-0.19(-0.22 - -0.17)	-0.12(-0.14 - -0.1)	0.58**
$H_{exp,[j_1,j_2=3,12]}$	0.64(0.61-0.68)	0.65(0.61-0.68)	0.63(0.6-0.65)	-0.18$^{n.s.}$
$C_{2,[j_1,j_2=3,12]}$	-0.16(-0.18 - -0.11)	-0.13(-0.16 - -0.11)	-0.11(-0.11 - -0.09)	0.32*

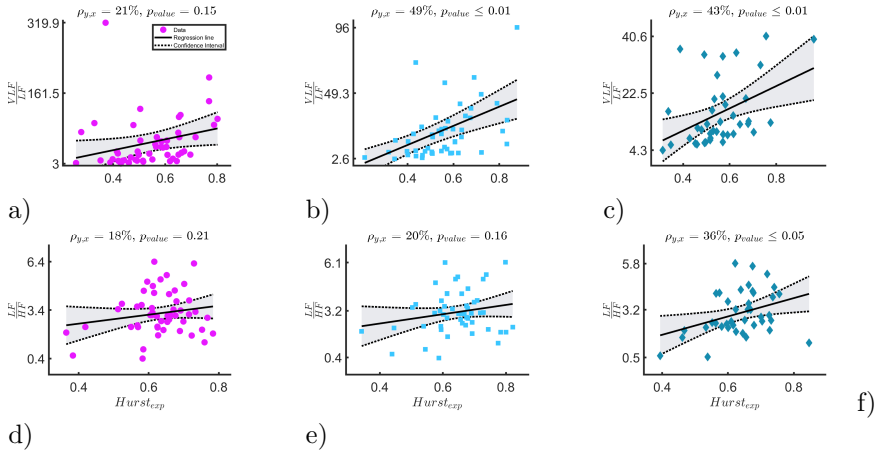


Figure 6.4: The figure shows results for the linear-mixed effect regression that models the relationship between $H_{exp,[j_1,j_2=5,12]}$ and $\frac{VLF}{LF}$ in the first row and between $H_{exp,[j_1,j_2=3,12]}$ and $\frac{LF}{HF}$ in the second row. The three columns respectively represents bradycardia epochs (magenta circle data points), post-bradycardia epochs (light-blue squares data points) and between-bradycardia epochs (indigo diamonds data points). ρ is the correlation coefficient of the regression and p_{value} represents the significance of the correlation.

than the $\frac{LF}{HF}$ and the Hurst exponent $[j_1, j_2] = [3, 12]$. This might indicate that neonates do not have sympathovagal balance that rely on the typical interplay between LF and HF [2],[71]. Third, the bradycardias can impact HRV maturational features, especially the most common temporal indices that are used in clinical practice, as highlighted in italic by the regression with only temporal features in the WB block of Table 6.5 and the correlation coefficients in Table 6.3. Additionally, the relationship between spectral ratios and H_{exp} is strongly diminished in the WB scheme, as stressed by Figure 6.4.

The different age models that were derived in this chapter can outperform or can be compared to the other developmental models reported in the literature [111], [149]. Specifically, Table 6.5 highlights the capacity of spectral features to outperform all other features in the PMA estimation in all three windowing schemes ($R^2_{train} = [0.74, 0.59, 0.72]$, $R^2_{test} = [0.57, 0.69, 0.62]$, $MAE = [1.83, 1.56, 1.9]$ $Weeks$). Furthermore, the LF power is consistently selected by LASSO for all the different f_s and with any type of windowing scheme. These results are not simply in line with decrease of $\frac{VLF}{LF}$ by [111], but they are also supported by other clinical findings. Namely, an increase of the short-term variability of the tachogram was found during first days of life

Table 6.5: Linear mixed-effect model performances. For each feature set and sampling frequency f_s, the table shows the median(IQR) over the 20 iterations for R^2_{train} on the train set as well as R^2_{test} and the MAE in weeks on test set. IQR here stands for the difference between 25% and 75% quantiles.

Post-bradycardia (PB) epochs			
Feature type $\|_{fs}$	R^2_{train}	$MAE(weeks)$	R^2_{test}
All features $\|_{6Hz}$	0.73(0.11)	1.88(0.39)	0.57(0.29)
All features $\|_{8Hz}$	0.75(0.1)	1.85(0.3)	0.56(0.23)
All features $\|_{12Hz}$	0.75(0.09)	**1.83(0.41)**	0.57(0.22)
Temporal features	0.44(0.28)	2(0.56)	0.35(0.19)
Spectral features	0.74(0.12)	2.01(0.42)	0.5(0.11)
Fractal features $\|_{6Hz}$	0.33(0.17)	2.21(0.42)	0.33(0.34)
Fractal features $\|_{8Hz}$	0.35(0.13)	2.31(0.54)	0.21(0.22)
Fractal features $\|_{12Hz}$	0.26(0.1)	2.18(0.46)	0.43(0.28)
Between-bradycardia (BB) epochs			
Feature type $\|_{fs}$	R^2_{train}	$MAE(weeks)$	R^2_{test}
All features $\|_{6Hz}$	0.55(0.12)	1.82(0.28)	0.57(0.24)
All features $\|_{8Hz}$	0.6(0.11)	1.81(0.38)	0.55(0.19)
All features $\|_{12Hz}$	0.68(0.11)	**1.56(0.39)**	0.59(0.16)
Temporal features	0.6(0.33)	2.06(0.38)	0.44(0.24)
Spectral features	0.59(0.19)	1.93(0.54)	0.59(0.15)
Fractal features $\|_{6Hz}$	0.3(0.22)	2.57(0.43)	0.15(0.19)
Fractal features $\|_{8Hz}$	0.22(0.26)	2(0.25)	0.24(0.36)
Fractal features $\|_{12Hz}$	0.34(0.28)	2.16(0.53)	0.18(0.31)
Within-Bradycardia (WB) epochs			
Feature type $\|_{fs}$	R^2_{train}	$MAE(weeks)$	R^2_{test}
All features $\|_{6Hz}$	0.73(0.1)	1.97(0.42)	0.58(0.25)
All features $\|_{8Hz}$	0.7(0.15)	1.91(0.21)	0.5(0.25)
All features $\|_{12Hz}$	0.72(0.15)	1.95(0.33)	0.57(0.24)
Temporal features	*0.14(0.1)*	*2.79(0.35)*	*0.13(0.13)*
Spectral features	0.73(0.17)	**1.9(0.21)**	0.62(0.21)
Fractal features $\|_{6Hz}$	0.33(0.07)	2.16(0.4)	0.23(0.18)
Fractal features $\|_{8Hz}$	0.36(0.13)	2.03(0.56)	0.43(0.22)
Fractal features $\|_{12Hz}$	0.4(0.16)	2.13(0.56)	0.29(0.28)

[62] and the absolute LF power can discriminate preterm and full term infants with 84% accuracy [176]. However, the highest performances in the PB and BB schemes are achieved when the fractal and spectral features are combined, as highlighted in bold in Table 6.5 and suggested by the concomitant increase of entropy and short-term variability of HRV found by [62]. Interestingly, the highest performances are also achieved when the between bradycardias epochs are considered ($MAE = 1.56\ Weeks$), which further reveals a bias effect of bradycardias in the description of autonomic maturation. In line with [111],[160],[50],[58], we found that the tachogram mean μ_{RR} and its standard deviation σ_{RR} increase with maturation together with the absolute power in all investigated frequency bands. If the relative power is considered, both $\frac{LF}{VLF+LF}$ and $\frac{LF}{HF+LF}$ increase with age (Table 6.3 and supplementary Tables A.1, A.2

Table 6.6: LASSO selected features for the linear mixed-effect model. For each feature set, the features that have been selected more than 40% of times have been reported. σ_{RR} and μ_{RR} stand for the standard deviation and the mean of HRV. $log_{10}(LF)$, $log_{10}(VLF)$,$\frac{VLF}{LF}$ stand for the absolute power in the LF and VLF bands and the ratio between the two. Results are reported for the Wigner-Ville distribution (SPWD) or the wavelet transform. $H_{exp,[j_2,j_2]}$ and $c_{2,[j_2,j_2]}$ stand for the Hurst exponent and the parameter c_2 in the range $[j_1, j_2]$.

Post-bradycardia (PB) epochs			
Feature type $\vert_{fs}$			
All $\vert_{6Hz}$	$\log_{10}(LF)_{SPWVD}$	$\log_{10}(LF)_{Wavelet}$	$C_{2,[j_1,j_2=5,12]}$
All $\vert_{8Hz}$	$\log_{10}(LF)_{Wavelet}$		
All $\vert_{12Hz}$	$\log_{10}(LF)_{SPWVD}$	$C_{2,[j_1,j_2=\mathbf{5},12]}$	
Spectral	$\log_{10}(LF)_{SPWVD}$	$\log_{10}(LF)_{Wavelet}$	
Fractal $\vert_{6Hz}$	$C_{2,[j_1,j_2=\mathbf{5},12]}$		
Fractal $\vert_{8Hz}$	$H_{exp,[j_1,j_2=\mathbf{5},12]}$	$C_{2,[j_1,j_2=\mathbf{5},12]}$	
Fractal $\vert_{12Hz}$	$C_{2,[j_1,j_2=\mathbf{5},12]}$		
Between-bradycardia (BB) epochs			
Feature type $\vert_{fs}$			
All $\vert_{6Hz}$	μ_{RR}	$\log_{10}(LF)_{SPWVD}$	
All $\vert_{8Hz}$	$\log_{10}(LF)_{SPWVD}$		
All $\vert_{12Hz}$	$\log_{10}(VLF)_{Wavelet}$	$\log_{10}(LF)_{SPWVD}$	$C_{2,[j_1,j_2=\mathbf{5},12]}$
Spectral	$\log_{10}(LF)_{SPWVD}$		
Fractal $\vert_{6Hz}$	$H_{exp,[j_1,j_2=\mathbf{5},12]}$	$C_{2,[j_1,j_2=\mathbf{5},12]}$	$H_{exp,[j_1,j_2=3,12]}$
	$C_{2,[j_1,j_2=3,12]}$		
Fractal $\vert_{8Hz}$	$H_{exp,[j_1,j_2=\mathbf{5},12]}$		
Fractal $\vert_{12Hz}$	$C_{2,[j_1,j_2=\mathbf{5},12]}$		
Within-Bradycardia (WB) epochs			
Feature type $\vert_{fs}$			
All $\vert_{6Hz}$	$\log_{10}(LF)_{Wavelet}$	$C_{2,[j_1,j_2=\mathbf{5},12]}$	$C_{2,[j_1,j_2=3,12]}$
All $\vert_{8Hz}$	$\log_{10}(LF)_{Wavelet}$	$C_{2,[j_1,j_2=\mathbf{5},12]}$	
All $\vert_{12Hz}$	$\log_{10}(LF)_{Wavelet}$	$C_{2,[j_1,j_2=\mathbf{5},12]}$	
Spectral	$\log_{10}(LF)_{Wavelet}$		
Fractal $\vert_{6Hz}$	$C_{2,[j_1,j_2=\mathbf{5},12]}$	$C_{2,[j_1,j_2=3,12]}$	
Fractal $\vert_{8Hz}$	$H_{exp,[j_1,j_2=\mathbf{5},12]}$	$C_{2,[j_1,j_2=\mathbf{5},12]}$	
Fractal $\vert_{12Hz}$	$H_{exp,[j_1,j_2=\mathbf{5},12]}$	$C_{2,[j_1,j_2=\mathbf{5},12]}$	

and A.3), which means the spectral ratios $\frac{VLF}{LF}$ and $\frac{LF}{HF}$ had opposite directions. While the latter increases with age, we also found that $\frac{VLF}{LF}$ decreases as shown by [111]. It should be stressed that the relative power features are less correlated with age (see suppplementary Tables A.1, A.2 and A.3) and they are never selected in the LASSO procedure (Table 6.6). These results confirm the findings by [160] and [58], who found also greater increase in the absolute power during maturation, and the findings by [176], which also displays the superiority of the absolute power in the classification between preterm and full-term neonates compared to the relative indices. The lack of stationarity and the $1/f$ spectrum behavior make difficult to describe the autonomic maturation without all the frequency bands in place. Furthermore, the regularity measure by the fractal indices with the H_{exp} and c_2 decreases with development and they play a role in the age estimation models (Table 6.6).

These findings seem to suggest that the ratio $\frac{LF}{HF}$ might not be the best suitable index for the sympathovagal balance and the common HRV frequency bands are suitable for infants. As anticipated, David et al. noticed that the fetal heart-rate has such enhanced slow-wave baseline, which increases the power in the VLF band, such that both Hoyer et al. and David et al. used the ratio $\frac{VLF}{LF}$ as a possible index to describe the sympatho-vagal interplay [59], [111]. This approach seems confirmed by the results in Figure 6.4. As discussed by Abry [2] and Doret [71], the spectral ratio is linked to H_{exp} via (6.3). The panels suggest that the ANS modulation and its fractal regularity lie in the lower-frequency bands, since the H_{exp} is more correlated with $\frac{VLF}{LF}$. The Pearson correlation coefficients ρ_{xy} between $\frac{VLF}{LF}$ and H are respectively 21%,49%,43% according to different windowing schemes compared to ρ_{xy} between $\frac{LF}{HF}$ and H_{exp} (18%, 20% and 36%). It is important to notice that the H_{exp} matches the spectral ratio if its estimation range $[j_1, j_2]$ matches the frequency bands with most of the exponential decay in the PSD. In this analysis, $[j_1, j_2 = 5, 12]$ covers specifically the lower frequency bands and its importance is confirmed by the features selected by LASSO (Table 6.6). In line with Doret [71], the current findings clearly suggest a redefinition of $\frac{LF}{HF}$ with an extension of frequency bands from the most common adults' range, e.g. $LF = [0.02 - 0.15]\ Hz$ and $HF = [0.15 - 1.3]\ Hz$. They also highlights the greater prominence of the slower oscillations in the description of premature ANS.

However, the results also highlight the disruptive role of bradycardias in maturation analysis. As anticipated, the best regression results are achieved in the between bradycardia epochs (Table 6.5) and the relationship between the spectral ratios and the H_{exp} is disrupted with WB windowing (panels with magenta circle 6.4). Most importantly, the relationship between the temporal features and maturation is lost, as highlighted by the poor R^2 (Table 6.5 and panels with magenta circles in Figure 6.3). In addition, Gee et al. [92] observed that the LF power, the variance and the regularity of the heart-rate increase before bradycardias. The results in Figure 6.3 and Tables A.2, A.3 support this increase in variance and regularity (as can be easily noticed by the y-axis of σ_{RR} or any other features of the left column in Figure 6.3). This finding clearly implies that exclusion of bradycardias is fundamental whenever using the standard deviation and the mean of the tachogram to assess the maturation of ANS. Figure 6.4 also shows that bradycardias annihilate the expected relationship between the spectral ratios and the Hurst exponent. This is a further proof that bradycardias disrupt the vagal tone [212], which can distort the PSD power-law in (6.2). However, Table 6.5 shows that the autonomic age models within bradycardia can maintain comparable performance to the other two windowing schemes thanks to the spectral features. In particular, Table 6.6 confirms that the most selected power attribute is the power in $LF = [0.08 - 0.2]\ Hz$, as further proof of the central role of this range in the bradycardic event and ANS

maturation [92], [59], [111].

It is also important to highlight some limitations. Bolea et al. showed a dependency for nonlinear metrics (such as Sample Entropy) with the resampling frequency of the tachogram [37]. Based on their results, they concluded that a resampling frequency correction of nonlinear parameters is needed in cardiovascular applications in order to detect meaningful results (such as experiments with body position changes) [37]. The employed fractal indices clearly show a increasing trend with an increasing sampling frequency (Section 6.2.4 and Table 6.4). However, a correction for the sampling frequency was not implemented. The multifractal properties were investigated with different sampling frequencies and our analysis seems to show mild differences in term of regression results and correlation with age (Table 6.4). One may also object the exclusion of proper sleep-staging in the current analysis, as normally done by [58]. However, the specific focus on the bradycardia effect strongly limit the number of the windows available. On top of that, bradycardias are events normally associated to active sleep [212] and most of the annotated bradycardias in this investigation were found during states that were not associated to quiet sleep. Similarly, one may also find the number of patients limited, but it was caused by the difficulties in the follow-up. All the included patients had normal developmental outcome at 2 years and the development assessment process is normally characterized by large drop-outs. Concerning the methodology, the different spectral methods (Welch, Wavelet and Wigner-Ville) show very similar spectral trends, but LASSO more frequently tends to select time-frequency distribution features (Wavelet and Wigner-Ville, Table A.2, A.3). Although there are studies that claim the superiority of the quadratic time-frequency methods [268],[187], the current findings show the wavelet approach would suffice for the spectral analysis.

In a nutshell, the HRV analysis might a useful tool for development monitoring, but two important factors have to be taken into account. First, the neonatal HRV is characterized by a very-low frequency tone which requires a redefinition of the different frequency bands to the autonomic stimulation. Second, bradycardia have a disruptive role in the assessment of maturation.

6.5 Summary

In the present chapter, the maturation of the preterm autonomic nervous system was investigated by means of temporal, spectral and fractal features of HRV. Three main findings can be reported. First, infants' maturation can be described by means of multifractal and spectral analysis, which show an increasing trend

of LF power as well as decreasing trend of fractal regularity with increasing post-menstrual age. The best obtained regression performances ($R^2 = 0.68$ and $MAE = 1.56$ $weeks$) are obtained as combination of fractal and spectral features and are comparable to other developmental models reported by different authors [111], [144], [149], [65]. Second, this predominance of LF and VLF bands as well as the lower scales for the multifractal indices suggest that the sympathovagal balance of neonates might not simply be related to the ratio $\frac{LF}{HF}$, but the entire HRV band and the regularity of the tachogram should be included to have better understanding of the ANS maturation. Third, bradycardias might forcefully increase the variance of the heart-rate and disrupt the relationship between autonomic indices and age. The PMA estimation models based on novel HRV indices provides a more comprehensive understanding of postnatal autonomic maturation. They can be also considered an alternative automated maturity index to other electrophysiological data analysis for the NICU. This research might be a first step to design personalized therapies or preventive care to preserve infants' development.

Part III

Perinatal Stress quantification

Chapter 7

A bradycardia-based stress calculator for the neonatal intensive care unit: a multisystem approach

This chapter has been published as Lavanga M., Bollen B., Jansen K., Ortibus E., Van Huffel S., Naulaers G., Caicedo A. (2020). "A bradycardia-based stress calculator for the neonatal intensive care unit: a multisystem approach". Frontiers in Physiology, 11, 74. Lavanga M. has developed the methodology, conducted the experiments and has written the manuscript. Compared to the publication, minor textual and notational changes have been implemented for better integration in this thesis.

Early life stress in the neonatal intensive care unit (NICU) can predispose premature infants to adverse health outcomes and neurodevelopment delays. Hands-on-care and procedural pain might induce apneas, hypoxic events and sleep-wake disturbances, which can ultimately impact maturation. However, a data-driven method based on physiological fingerprints to quantify early-life stress does not exist. This chapter provides an automatic stress detector by investigating the relationship between bradycardias, hypoxic events and perinatal stress in NICU patients. A binary classifier was designed with physiological data from 136 patients and the stress burden assessed by the Leuven Pain Scale. The results of this study were published in [142].

7.1 Introduction

Premature infants are at risk of maladaptive outcomes and neurodevelopment delays. Patients who spend their early life in the neonatal intensive care unit (NICU) can undergo profound alterations of sleep-patterns as well as exposure to painful procedures and noxious stimuli [16],[100]. Grunau et al. have shown how stress exposure can induce a cascade of physiological consequences, behavioral and hormonal responses [100]. In addition, Brummelte et al. highlighted how procedural pain can affect structural connectivity of the subcortical areas during neurodevelopment [42].

In particular, routine day-care has been reported to affect sleep quality inside the NICU [16]. Levy et al. showed that prolonged contact in NICU can have multiple consequences [153]. 57% of the sleeping infants experience awakening because of hands-on care. Handling is usually followed by respiratory events, such as hypoapneas and apneas, or desaturations. Surprisingly, clinical handling is more likely to initiate oxygen desaturation and bradycardias. Monitoring of respiratory and hypoxic events is pivotal since experience of long bradycardias and apneic spell in very-low weight infants are known to impact the development of the patients [109],[122], [205]. In particular, Janvier et al. showed that a higher apnea burden (total amount of apnea days in the ward) is associated to a worsening of the cognitive and motor outcome [122]. Prolonged oxygen desaturations associated with bradycardias are known to have greater negative effect on cerebral oxygenation [205] and the persistence of their effect can be even prolonged at 5-6 months corrected age, with worse SpO_2 and heart-rate drops compared to full-term infants at equivalent age [109]. Furthermore, bradycardias were under scrutiny in different studies as a sign of autonomic nervous system development. Gee et al. showed how the heart-rate variance and entropy dramatically change before any heart-rate drop [92]. This could be the consequence of a dysfunction of vagal stimulation, which induces the bradycardia, according to the polyvagal theory by Porges and colleagues [212]. Those events are usually preceded by low-heart rate variability as a sign of fetal distress [212].

Although a possible link exists between stress burden and cardiorespiratory events, an automated method to quantify stress exposure in the NICU based on physiological signal activity, especially during oxygen desaturations or bradycardias, has not been described yet. However, the literature provides an overview on how physiological signals can be used to investigate pain and apneic spells in adults. Multiple authors described machine-learning models to classify pain-patterns using different modalities, such as EEG or EMG [170],[102]. In parallel, other authors described an algorithm to detect apnea events based on SpO_2 analysis [68]. In addition, some authors already investigated a possible link

between modalities that describe brain activity and modalities that describe cardiovascular activity, in the case of apneic spells or desaturation events. Specifically, a recent study proposed a model to explain how pre-frontal cortex dysfunctions in adults and children can be caused by obstructive sleep apneas due to disruption of sleep and chemical homeostasis [29]. Pitson et al. showed that SpO_2 dips due to apneas are related to the patients' daily sleepiness, which can affect the emotional and behavorial state. Interestingly, those desaturation events seem to significantly correlate to other physiological events, such as EEG and heart-rate arousals [209].

This inherent coordination of different physiological systems in case of apneas, as highlighted by Pitson et al. [209], or the necessity to rely on different modalities to classify biopotential information, as shown by several authors [102],[170], strongly suggest a horizontal interaction among organs, which might be altered in case of stress or hypoxia, and might require different tools to approach the alteration of the physiological state of the patients [120]. This synchronization among different organs or signal modality is known as Network Physiology and was specifically applied to show the alteration between brain activity and parasympathetic tone of the HRV [127] and the synchrony between the neonatal EEG bursts and the heart-rate accelerations of the infants [204]. However, one might investigate network physiology in the infants and relate that to a specific physiological condition. As highlighted by Bashan et al. [25], physiological systems under neural regulation exhibit a high degree of complexity with non-stationary, intermittent, scale-invariant and nonlinear behaviour and change in time under different physiological states and pathological conditions. One can not only simply derive the integration among the different physiological systems, but might also try to summarize the topological properties of the physiological network and investigate their evolution over time [22],[23]. The clinical literature also suggested that the overall activity of the individual physiology cannot simply be summarized as the sum of the individual organs' physiology, but it requires an investigation of the interaction among the different sub-systems, especially in the intensive care setting [174].

Since the clinical literature already showed a unique relationship between handling of infants and apneas or hypoxic events [153], the aim of this study is the development of a classification model to relate hypoxias to patient's stress exposure. A binary classifier was developed to classify whether a bradycardic event belonged to a patient with stress or without stress burden. Due to the interdisciplinary nature of hypoxic events and stress exposure, the study aimed not only to derive the features from different modalities, but assess the network physiology of the patients and its relationship with stress load and bradycardias. Stress is here intended as the accumulation of procedural pain, as described in Section 2.5 and in [100]. The ultimate goal of this analysis is to demonstrate

if stress is associated to cardiorespiratory abnormalities or other physiological abnormalities. Bebee et al. [29] showed that hypoxias and sleep disruptions might impact the prefrontal cortex and therefore the cognitive function of the patient. The specific detection of hypoxic events related to stress might lead to a possible explanation how the accumulation of pain influences the development outcome.

7.2 Material and methods

7.2.1 Patient sample

Data from preterm infants were collected as part of the Resilience Study, which has been carried out in the Neonatal Intensive Care Unit (NICU) of the University Hospitals Leuven, Belgium. A total of 136 preterm infants born before 34 weeks gestational age (GA) and/or with a birth weight less than 1500 g was included in the cohort from July 2016 to July 2018. The exclusion criteria and information related to the research protocol or the inform consent are reported in Section 1.2.

7.2.2 Data acquisition

During the NICU stay, pain levels were daily recorded with a multidimensional scale for premature infants known as the Leuven Pain Scale (LPS). This scale varies in the range [0,14] and is obtained as the sum derived by seven categories (such as crying, grimace or heart-rate) [7], [6]. LPS scores were routinely daily recorded by bed-side nurses, every hour for the intensive care patients and every three hours for the intermediate care.

Based on the association between stress and pain, perinatal stress has been defined as the presence of non-zero LPS in the patient record the day before the recording, i.e. $LPS > 0$, which means experience of any pain the day before the recording.

According to the clinical protocol, EEG, ECG and SpO_2 measurements were recorded for at least 3 hours in three monitoring groups: the first measurement took place around 5 days after birth (5days), while the second and the third recording were respectively planned around 34 weeks postmenstrual age (PMA) (34w) and in the week before discharge home. The last recording usually consisted of a 24 hours polysomnography, therefore the last group was labeled as PSG. Only one of the first two recordings was performed for infants born at

Table 7.1: Summary of patient data demographics at different time points: GA (gestational age), birth weight (in g), PMA (postmenstrual age) at EEG and ECG recording, LPS (Leuven Pain Score). Data is reported as $q_{50}[q_{25}\text{-}q_{75}]$, where q_{50} is the median and q_{25}-q_{75} is the IQR.

	5days (n=118)	34w (n=67)	PSG (n=117)
GA($weeks$)	31.14 [28.86-32.43]	28.86 [26.86-30.71]	30.29 [27.29-31.71]
Birth Weight (g)	1475 [1120-1725]	1140 [900-1480]	1225 [950-1540]
PMA($weeks$)	32.14 [30-33.43]	34.14 [33.86-34.29]	38.43 [37.29-39.57]
LPS	1 [0-3]	0 [0-2]	0 [0-2]

33-34 weeks. In the course of their NICU stay, some infants were transferred to level II units in hospitals closer to home. Therefore, not all infants have multiple recordings and some LPS measures are missing. A total of 245 recordings had corresponding pain scores available and were analyzed. A total of 39 patients had three recordings with associated pain score. A set of 38 patients had two recordings and the remaining 52 had 1 recording ($39 * 3 + 38 * 2 + 52 = 245$). Table 7.1 summarizes the clinical characteristics of patients at each measuring point. EEG set-up included nine monopolar electrodes (F_{p1}, F_{p2} , C_3, C_4 , C_z, T_3, T_4, O_1, O_2) and the EEG signals were referenced to the electrode C_z. The sampling frequencies for EEG, ECG and SpO_2 were 256, 500 and 1 Hz, respectively. They were monitored with the OSG system (OSG BVBA, Brussel). The R-peaks of the ECG were detected via the R-DECO toolbox [172] and the tachogram or HRV signal was derived as subsequent R-peak to R-peak intervals (RR_i).

7.2.3 Bradycardia detection and data preprocessing

Multiple studies have shown that hands-on-care and clinical handling can disrupt the sleep cycle and induce oxygen desaturations and apneic spell [153],[16]. The most threatening desaturations for the brain physiology and the development of the infant are usually events concurrent with bradycardia, i.e. a sudden drop in heart-rate, [205], [109]. Since Levy et al. showed that bradycardias, apneas, hypoapneas and hypoxic events are linked to stress exposure [153] and Porges et al. related bradycardias to fetal distress [212], the definition of apnea prematurity was followed to detect cardiorespiratory events or desaturations in the physiological signal [195]. Clinically relevant apneas were characterized by RR elongation above $1.5 * \overline{RR_i}$ for at least 4 s, where $\overline{RR_i}$ is the average of the entire tachogram, with a variation of SpO_2 greater than 10% with respect to the baseline [122]. Consequently, hypoxic events were detected as events with

concomitant variations of HRV and oxygen saturation, defined by increases above $1.5*\overline{RR_i}$ for more than 4 s and SpO_2 desaturations exceeding the following different thresholds: 3%, 5% and 10%. The saturation drops from the baseline were detected according to [68] and the different thresholds were used to test whether stress exposure induces more pronounced hypoxic events. Normally, apneas are defined as breathing cessation for more than 20 secs. However, both Barbeau et al. [16] and Levy et al. [153] have shown that events due to NICU handling are not necessary full apneic spells, but mostly physiological events like hypopneas and desaturations which last shortly and do not reach the level of clinical alarm. Gee et al. [92] and Porges et al. [212] outlined the solely and specific importance of bradycardias as a sign of dismaturity and distress of the premature infant. In addition, the respiration signal in our study was frequently distorted by artifacts and usually derived from the ECG for the younger patients. Therefore, the event detection specifically targeted bradycardias, instead of looking at the general breathing cessations. For each of those events, a window of 3 minutes before and after each bradycardia peak was the starting interval to develop a stress classifier. Specifically, a bradycardia peak is the moment of maximal heart-rate drop or RR intervals elongation. For each epoch, the EEG signal was filtered between [0.5-20] Hz and possible EOG artifacts were filtered using independent component analysis.

7.2.4 Features extraction

Multiple features were extracted from the EEG, HRV, SpO_2 from each bradycardic spell to assess its relationship with stress. They were computed at least in two moments: the period before the bradycardic event, i.e. from the start of the window until the RR_i exceeds $1.5 * \overline{RR_i}$ threshold, and the period after the bradycardic event, which goes from the moment RR_i comes back to stationarity until the end of the window. According to the different methodologies, features were also computed during the bradycardia or during the entire hypoxic spell. The computation within the bradycardia was not always possible since indices like fractality require a larger number of samples that were not available. Furthermore, the epoch durations were variable depending on the length and the intensity of the bradycardic event. An overview of the different features are reported in Table 7.2 and Table 7.3.

Cardiovascular analysis: HRV and SpO_2 features

The tachogram's reactivity was investigated with classical temporal and spectral indices. Specifically, the power spectral density (PSD) of the tachogram was computed with the continuous wavelet transform using analytical Morlet as

mother wavelet (Section 3.1.1). The absolute powers in the high-frequency (HF) and low-frequency (LF) range were derived as sum of the PSD bins in the following frequency bands: $HF = (0.2 - 4]\ Hz$ and $LF = (0.08 - 0.2]\ Hz$ [59]. The indices $\frac{LF}{HF}$ and $\frac{HF}{LF+HF}$ were used to assess the contribution of the multiple autonomic branches [111]. Since the wavelet-approach derives the time-frequency distribution of a signal, both the mean and the standard deviation of those indices, together with the temporal mean and standard deviation of the HRV, were derived as features in the epochs before, during and after each bradycardic spell. Additionally, the heart-beat dynamics were assessed via the Poincaré Plot (PP) analysis. The PP are two-dimensional scatter-plots where $RR(t)$ is plotted versus the lagged sample $RR(t + \tau)$. This graphical representation is a simplification of Taken's theorem to represent the phase space in order to assess the nonlinear behaviour of the signal. The lag τ was estimated as the first zero of the autocorrelation function of the signal and the PP can then be described by the matrix $\mathbf{X} = [RR(t), RR(t + \tau)]$, where $RR(t)$ is a vectorial representation of the HRV time series of dimension $\mathbb{R}^{(N-\tau)\times 1}$, where N represents the length of the signal. Most commonly, the standard deviations SD_1 and SD_2 of the minor and major axis of the cloud defined by X are computed to represent the short and long-term RR variability. In this study, the information in the PP was quantified via SD_2 and SD_1 as the first two singular values of X and via the centroids C_x and C_y of the same matrix as the column-wise mean of matrix X. The PP was represented and investigated using the entire bradycardic window.

Similarly to HRV, temporal features, such as mean and standard deviation, as well as the PP features were derived from SpO_2. Concerning the epochs for SpO_2 features computation, the epoch before and after SpO_2 dips were considered, i.e. the epoch that starts from the beginning of the window until SpO_2 exceeds the considered threshold and the epoch that starts from the moment that SpO_2 goes back to stationarity until the end of the window.

Desaturation events and bradycardic spells never occur alone, especially when driven by hands-on-care. The periodicity of both SpO_2 dips and heart-rate can be characterized by Phase Rectified Signal Averaging (PRSA), which searches for all time points where the signal goes downward (or upward) in the 6 minute segments [27]. Fragments of 120 s duration were extracted around each time point, known as anchor point, and they were subsequently aligned and averaged. From this average curve, the overall slope and the slope before and after each anchor point were derived to describe the rate of increase or decrease, such as a desaturation trend or bradycardia increase [27]. However, the computed average rate is sensitive to the definition of the anchor points, which ultimately represent an increase or decrease for a specific time window of length T according to the properties of the signal. Therefore, multiple parameters T were investigated in

the range $[1, 5, 10, 20, 50, 100]$ s to define the best set of PRSA features .

Neurological analysis: EEG features and multivariate attributes

Pitson et al. showed how EEG arousals are related to SpO_2 dips in respiratory events due to obstructive sleep apneas [209]. Those arousals were defined as an increase in the main carrier frequency of EEG in windows of 10s or more. Furthermore, different authors showed swings in burst activity as a consequence of HR variations in premature infants [204], [230]. The increase in discontinuity and burst-like type of activity are known biomarkers for brain dismaturity or pain elicitation [198], [78]. Therefore, multiple features were computed from the EEG to describe the level of discontinuity in terms of slow-wave persistence, regularity and lack of smoothness [198]. In addition, the concurrent variations of heart-rate, SpO_2 and EEG were investigated to assess whether they are related to the stress load or not.

EEG time-frequency analysis

The cortical activity was analyzed both in the time and frequency domain. The EEG power in the band $\delta = (0.5 - 4]\ Hz$ was computed via the continuous wavelet transform, using the analytical Morlet as mother wavelet (Section 3.1.1). The reason to focus on the δ band is two-fold. On the one hand, the δ band represents the sensitive band to pain stimuli and contains the dominant frequency of the neonatal EEG, which is the frequency with the highest power [78], [261], [1]. On the other hand, this frequency band represents subcortical areas, such as the thalamus, which are involved in stress management and autonomic control of the nervous system [203]. Similarly to the cardiovascular variables, the mean and the standard deviation for the EEG and the power in the δ band in each channel was derived for the three epochs around the bradycardic peak.

Multifractality

As anticipated in section 3.1.3, a more discontinuous EEG signal is characterized by higher regularity or self-similarity. Signals with such property are defined as fractals or scale-free signals. These time series have long-exponentially decaying autocorrelation functions (ACF) or a power-law spectrum, whose rates of decay can be defined by the Hurst exponent (H), which assess the level of similarity [71]. However, complex and discontinuous signals can vary in fractal properties over time, i.e. the Hurst exponent and therefore the rate of ACF decay can differ [121]. Wendt et al. proposed an efficient way to estimate the different fractal

properties based on wavelet leaders [71]. Their method estimates the spectrum of singularities $D(h)$ (SS), which measures the different Hurst Exponents in the signal and the associated fractal dimension [266].

The most interesting attributes of the singularity spectrum are the location of the maximum and its width which are usually defined as c_1, c_2 [121]. According to Jaffard [121], c_1 is usually considered the main Hurst exponent (H_{exp}) of the multifractal signal, while c_2 is a variational index to represent the amount of fractals inside the signal. Wendt reported further details of the methodology and of the WLBFM toolbox implemented in MATLAB to estimate the fractal parameters [266]. The parameters c_1, c_2 were estimated for each EEG channel and the associated δ oscillations.

Multivariate analysis: Brain-heart interactions

The interaction among the cortical activity and the cardiovascular variables can be estimated with the time-frequency coherence between the δ oscillations derived with CWT, the HRV and the SpO_2 [207]. In order to match the temporal scale, all signals were resampled at 8 Hz. The continuous wavelet coherence $C_{x_i \leftrightarrow x_j}(t, f)$ was then computed between the signal x_i and x_j, as described in section 3.2.1. Specifically, the coupling was derived as the maximum absolute imaginary part of $C_{x_i \leftrightarrow x_j}$ in the band of interest and the statistical validity of each coupling was then tested with amplitude adjusted Fourier transform (AAFT) surrogates (Section 3.2.1, [149]). The signal x_i can be the delta oscillation of an EEG channel, HRV or the SpO_2. The wavelet transform was computed with analytic Morlet as mother wavelet and the coherence was investigated in the very-low-frequency band $VLF = (0.033 - 0.08]\ Hz$ in the 5 days group and the low-frequency band $LF = [0.08 - 0.2]\ Hz$ in the 34w group and the PSG group. As discussed in previous studies [148], [111], this shift in frequency band is due to undergoing maturation of the autonomic nervous system.

However, due to the large number of channels and exponential number of associations, the pairwise coupling risks to produce collinear features for stress discrimination. Therefore, $C_{x_i \leftrightarrow x_j}$ were interpreted as entries of adjacency matrix $A = A_{ij}$ of a weighted undirected graph and a set of topological indices, such as the path length, the clustering coefficient, the eccentricity, and a network resilience metric, the number of superfluous connections n_{sup}, were computed, as discussed in section 3.2.4 and [38],[43].

In this study, graph theory was applied as follows: EEG delta oscillations (8 channels), HRV and SpO_2 were involved in the analysis as nodes setting the number M of processes to 10. Since the interaction estimation is based on

wavelet coherence, the adjacency matrix was computed for each time sample and therefore it was possible to derive the charts of the different topological indices. The average and the standard deviation of each topological feature was computed before, during and after each bradycardic spell. In order to test the contribution of a specific modality or signal to the stress classification, graph theory indices were not only computed for the entire set of processes, but we used also partitions of the adjacency matrix A_{ij}. Specifically, we considered connections only related to EEG channels (EEG-EEG), the connections between EEG channels and SpO_2 (EEG-SpO_2), the connections between EEG channels and RR_i (EEG-RR_i) and the entire set of connections (EEG-SpO_2-RR_i), as reported in Table 7.3. For each of those partitions, the described list of topological indices was computed.

7.2.5 Bradycardia-based Classification

A customized software tool was developed with MATLAB libraries to detect whether each bradycardic event belonged to a patient with or without stress burden. In summary, the following groups of features were derived for each hypoxic event:

- Temporal and periodicity features: 14 features in total for HRV, 14 features for SpO_2 and 16 features for the EEG
- Spectral features for both HRV and EEG: 8 features for HRV and 16 features for EEG
- Nonlinear features: 4 features for HRV, 4 features for SpO_2 and 32 features for EEG
- Brain-heart connectivity topological indices: 168 features in total for HRV, EEG and SpO_2

A complete overview is reported in Table 7.2 and Table 7.3. Given the fact that features were derived for three epochs (before, during and after each bradycardia), the total number of extracted features was 748.

The power-features were log-transformed. The study investigated whether there was any association between those features and the bradycardic spell in a patient with a stress exposure in the NICU. As mentioned earlier, the presence of stress was defined as experience of pain the day before the recording ($LPS > 0$). However, Gruss et al. have shown that more intense pain can be discriminated in an easier way [102]. On top of that, there is no clear consensus on the level of desaturation that can be considered threating for premature infants [122], [153],

[210]. Therefore, different levels of hypoxia were tested in the classification, i.e. $SpO_2 > 3\%$, $SpO_2 > 5\%$, $SpO_2 > 10\%$.

The objective of the classification was to discriminate whether a bradycardic event belonged to a patient with or without stress. After testing different classification algorithms such as support vector machines (SVMs) and linear discriminant analysis (LDA), a classifier based on subspace ensemble with LDA has been designed to separate bradycardic spell in two groups [108]. Subspace LDA is an ensemble method like random forest, where the bagging process (random subsampling of the training set) is performed together with a random subsampling of the features to find the best feature subsets to separate the data (see Section 3.4.5 and [108]). The clear advantage is to span a greater number of features and allow the model to tune for the best subset. The choice of the subspace LDA was based on generalization and simplicity. Due to the great variety of features and the initial filter for features selection based on linear separation (see below), the comparison among different classification methods prioritized the models obtained with subspace LDA. The model was tested according to a leave-one-patient-out (LOPO) scheme for each monitoring group (5days, 34w, PSG), which meant that all bradycardic event belonging to one patient were put in the test set. The following set of performance indices were derived each monitoring group: the area under the curve (AUC) and Cohen's kappa between machine learning labels and the clinical labels (Section 3.4.6). It is important to remind the only one set of indices was obtained for each classifier since they were obtained by pooling all test sets of the different patients together. The classification hyperparameters of the subspace LDA, which are the γ parameter of the discriminant function and the number of ensemble cycles, were tuned via a 10-fold cross-validation in the training set. Like any random-forest family or ensemble method, the subspace LDA has one learner (normally a decision tree, but in this case linear discriminant function) and a number of cycles to tune the classifier with the bagging process [108]. The number of cycles represents also the number of "trees" in the "forest". In a subspace LDA method, the optimization focuses on the γ parameter of the discriminant functions of the different trees and the number of cycles necessary to reach the minimal classification error on the validation set. The comparison metrics to select classification methods were the AUC and the kappa-scores, which were the highest with subspace LDA.

Given that the number of features should be below one tenth of the training dataset, the subspace of features has been restricted to 1/10 (one-tenth) of the data as a generic rule of thumb to avoid overfitting [170]. However, before tuning of the model, features were further reduced before the subspace ensemble algorithm was applied. The considered attributes had intra-feature correlation below 90% and the highest F-scores. The F-score is a simple measure to assess

Table 7.2: Overview of the univariate features derived from the physiological signal in the study. For each signal (HRV, SpO_2, EEG), the temporal, spectral and nonlinear attributes are reported. The total count for HRV is 26: 2 temporal features, 8 spectral features, 4 nonlinear features from the Poincaré plot (PP) and 12 PRSA features. The total count for SpO_2 is 18: 2 temporal features, 4 nonlinear features and 12 PRSA features. The total count for EEG is 64: 2 temporal features, 2 spectral features and 4 fractal features repeated for each channel.RR and $P(\delta)$ respectively represent the tachogram or HRV and the EEG power in the δ band. μ and σ stand for mean and standard deviation. LF and HF represent the high and low-frequency bands of HRV. H_{exp} and c_2 are the main Hurst exponent and the width of the singularity spectrum derived with the multifractality framework. C_{xy} and SD_{12} are the PP features. $Slope_{OV}(T)$ represents the overall PRSA slope from the start of its window, while $slope_{AP}(T)$ is the slope around each anchor point. Six different window lengths T were selected to define each anchor point: $[1, 5, 10, 20, 50, 100]$ s.

	Temporal	Spectral	Nonlinear
HRV	μ_{RR}, σ_{RR}, $Slope_{OV}(T)$, $Slope_{AP}(T)$	μ_{HF}, σ_{HF}, μ_{LF}, σ_{LF} $\mu_{HF_{nu}}$, $\sigma_{HF_{nu}}$, $\mu_{\frac{LF}{HF}}$, $\sigma_{\frac{LF}{HF}}$	C_x, C_y SD_1, SD_2
SpO_2	μ_{SpO_2}, σ_{SpO_2}, $Slope_{OV}(T)$, $Slope_{AP}(T)$		C_x, C_y SD_1, SD_2
EEG	μ_{EEG}, σ_{EEG},	$\mu_{P(\delta)}$, $\sigma_{P(\delta)}$	$H_{exp,EEG}$, $H_{exp,P(\delta)}$ $c_{2,EEG}$, $c_{2,P(\delta)}$

the discrimination between the positive and the negative class. It is computed as the ratio between the separation between positive and negative class (intra-class variability) and the separation within each class (inter-class variability). The details of the procedure are reported here [47]. In addition, the features were corrected by the baseline effect of age in case subject's PMA was a covariate of the feature of interest (i.e. significant Pearson correlation or $p < 0.05$).

7.3 Results

The results for the bradycardia-based stress classification are reported for the three monitoring groups in Figure 7.1. The AUC and kappa scores are reported in function of the desaturation threshold used to define which events should have been included in the classifier. Each color represent a threshold: blue for desaturations higher than 3%, yellow for desaturations higher than 5% and red

Table 7.3: Overview of the multivariate features derived from the different monitoring groups and the possible interaction combinations among the different modalities (EEG-SpO_2, EEG-RR_i, EEG-EEG, EEG-SpO_2-RR_i). For 5_{days}, the interaction was derived specifically for the VLF band, while the interaction for other two groups was assessed in LF band. The set of attributes for each monitoring group is 84: 21 features for EEG-SpO_2, 21 features for EEG-RR_i, 19 for EEG-EEG, 23 features for EEG-SpO_2-RR_i.The clustering coefficient ($Clust_c$) and the eccentricity (Ecc) are node-dependent features, which explain the variation in numbers for each interaction group. The actual count rises to 168 since both mean and standard deviation are considered. n_{sup} is the number of superfluous connections.The label $Efficiency$ represents the global efficiency of the network. VLF and LF represent the very-low and low-frequency bands of HRV.

	EEG-SpO_2	EEG-RR_i	EEG-EEG	EEG-SpO_2-RR_i
5_{days}	$Path_{length}(VLF)$, $Efficiency(VLF)$, $Clust_{c,node}(VLF)$, $Ecc_{node}(VLF)$, $n_{sup}(VLF)$	$Path_{length}(VLF)$, $Efficiency(VLF)$, $Clust_{c,node}(VLF)$, $Ecc_{node}(VLF)$, $n_{sup}(VLF)$	$Path_{length}(VLF)$, $Efficiency(VLF)$, $Clust_{c,node}(VLF)$, $Ecc_{node}(VLF)$, $n_{sup}(VLF)$	$Path_{length}(VLF)$, $Efficiency(VLF)$, $Clust_{c,node}(VLF)$, $Ecc_{node}(VLF)$, $n_{sup}(VLF)$
34_w, PSG	$Path_{length}(LF)$, $Efficiency(LF)$, $Clust_{c,node}(LF)$, $Ecc_{node}(LF)$, $n_{sup}(LF)$	$Path_{length}(LF)$, $Efficiency(LF)$, $Clust_{c,node}(LF)$, $Ecc_{node}(LF)$, $n_{sup}(LF)$	$Path_{length}(LF)$, $Efficiency(LF)$, $Clust_{c,node}(LF)$, $Ecc_{node}(VLF)$, $n_{sup}(LF)$	$Path_{length}(LF)$, $Efficiency(LF)$, $Clust_{c,node}(LF)$, $Ecc_{node}(LF)$, $n_{sup}(LF)$

for desaturations higher than 10%. The results suggest a moderate association between the bradycardia features and the clinical labels: the AUC lies in the range [0.80-0.96] and the kappa score lies in the range [0.41-0.80]. The SpO_2 threshold for the desaturation seems to have a mild effect on classification: only the PSG group reports an increasing Kappa score for higher threshold.

The effect of the threshold is also reported in Figure 7.2, where the classification results are shown based on the different feature groups. The left panel shows the AUC for a 3% desaturation threshold, while the right panel shows results for the 10% threshold. The feature groups are respectively indicated with the labels EEG, HR – SpO_2 and BH for the EEG features, the cardiorespiratory features and the brain-heart features. In the 5 days group and 34w group, either the brain-heart features or the EEG features outperform the HR-SpO_2 group. In addition, the desaturation threshold seems to increase the AUC for the brain-related attributes. On the contrary, the performance seems to be comparable for all different groups at PSG and the effect of the threshold is

equally beneficial for the three groups.

In order to give an idea of the selected features or the most discriminative information for stress classification, Figure 7.3, 7.4, 7.5 reported either the behavior of the selected time-series or the boxplots of the most-discriminative features in epochs before, during and after each bradycardia for the three different monitoring groups. Figure 7.3 reports the desaturation charts for the 5days group with $LPS > 0$ (in blue) and $LPS = 0$ (in green) on the left panel, while the Hurst regularity is reported in the period before and after each bradycardia for a 10% threshold on SpO_2. The Hurst exponent shows a higher regularity in case of stress and the SpO_2 charts show higher desaturation in case of stress. Figure 7.4 reports the desaturation charts and the path length among EEG channels and HRV in the LF band for the 34w group with a 10% threshold on SpO_2. Results reveal a higher desaturation in case of stress as well as a stronger association between the tachogram and the delta-oscillations of the EEG. It is important to remember that the lower the path length, the higher the connectivity. Figure 7.5 reports the normalized power in the HF band both as time-series and as boxplots for the PSG group with a 10% threshold on SpO_2. The figure does not only suggest a higher and more intense bradycardic spell, but also a more variable bradycardia.

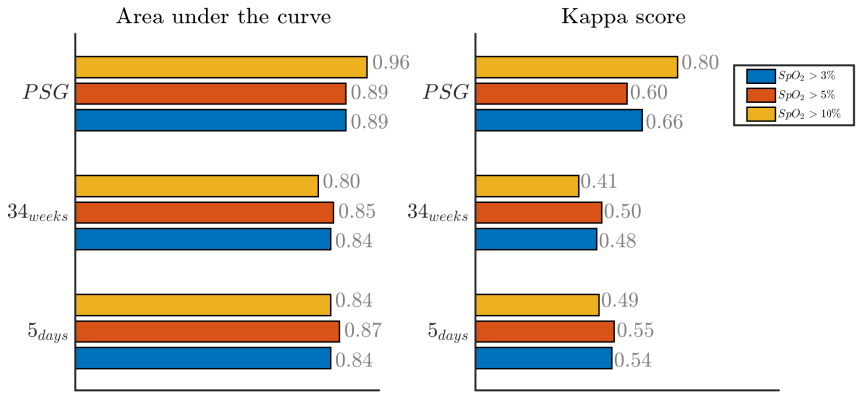


Figure 7.1: Results of the bradycardia-based classification in three main datasets. The three colors represent different level of desaturation to consider the bradycardic event in the stress classification. The left panel displays the *area under the curve* in the three monitoring groups, while the right reports Cohen's kappa.

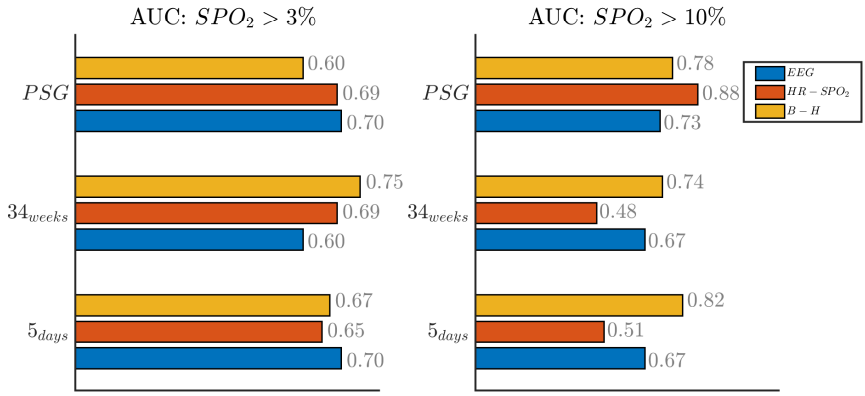


Figure 7.2: Results of the bradycardia-based classification in three main datasets. The figure here reports the results based on the different feature groups. The left panel reports the area under the curve for desaturations greater than 3%, while the right panel report the results for desaturation greater than 10%. The three colors represent different feature groups: EEG stands for EEG features, HR-SpO_2 represent the cardiovascular features and B-H is related to the brain-heart connectivity.

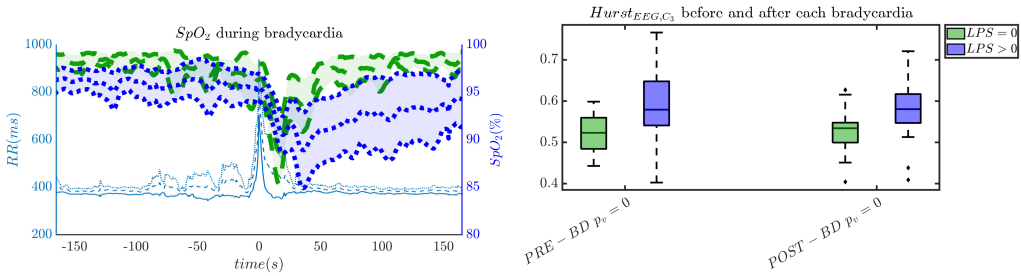


Figure 7.3: The desaturation levels and the EEG regularity are more pronounced in case of stress. The left panel reports the SpO_2 during the bradycardic spell and the right panel shows the boxplot for the Hurst exponent of channel C_3 for the period before and after each bradycardia. The data are reported for the 5days group. All the events with a desaturation greater than 10% were included in this figure. The p-values in the boxplot are derived via the Kruskal-Wallis test.

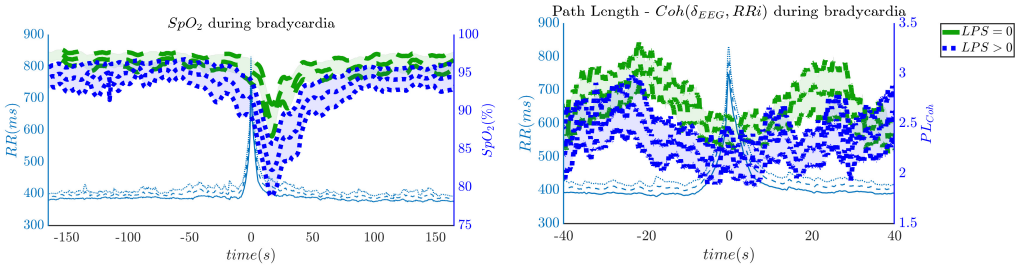


Figure 7.4: The desaturation levels and the connectivity between delta oscillations and the heart-rate are more pronounced in case of stress. The left panel reports the SpO_2 during the bradycardic spell and the right panel shows the path length derived from the network with EEG channels and the HRV. It is important to remind that the lower the path length, the higher the connectivity. The data are reported for the 34w group. All the events with a desaturation greater than 10% were included in this figure.

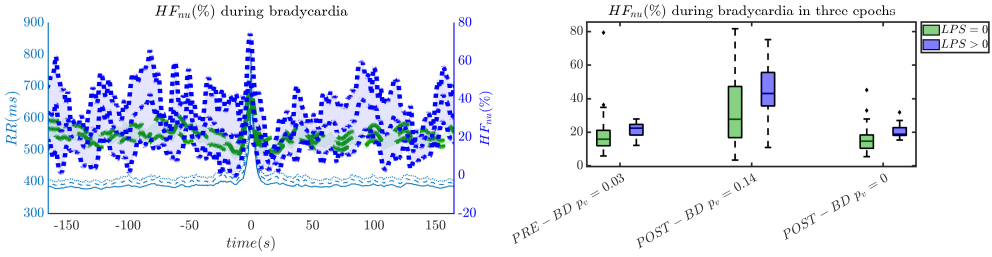


Figure 7.5: The intensity of bradycardias and the parasympathetic activity are more pronounced in case of stress. The left panel reports the normalized HRV power in the HF band and the right panel shows the normalized power in boxplots before, during and after each bradycardic event. The data are reported for the PSG group. All the events with a desaturation greater than 10% were included in this figure. The p-values in the boxplot are derived via the Kruskal-Wallis test.

7.4 Discussion

In this chapter, the relationship between bradycardic spells and stress burden in premature infants was examined and it suggests that stress load can enhance the desaturation and the bradycardic effects. Two novel findings can be reported.

First, this research supports the feasibility of the automatic stress classification based on the physiological reactivity in bradycardias. Levy et al. showed how routine contact in the NICU could induce respiratory events, such as apneas and hypoapneas, and long oxygen desaturations [153]. This result was confirmed by the classification performance reported in Figure 7.1 and 7.2 and the desaturation charts displayed in Figure 7.4 and 7.5. The definition of routine handling by Levy et al. follows the notion of stress exposure or procedural pain by Grunau et al. [100], who defines perinatal stress as accumulation of pain and noxious stimuli. The experienced hands-on-care and pain might trigger a completely different physiological reactivity which could induce a greater desaturation or respiratory burden, as also reported by Levy et al. [153]. Interestingly, the results showed a moderate association between the features and the classification outcome (with kappa score between 0.3 and 0.6 for the most of the groups). Although similar studies that perform classification of pain stimuli based on physiological information show strong association between features and the outcome [41],[102],[170], it is important to remind that the study protocol does not elicit any pain in the patient. And yet, it shows that babies experiencing pain the day before the measurement react differently to stress as shown by the stress calculator but also by looking at individual parameters like the desaturation chart, Hurst exponent of the EEG and the HRV in the LF and HF bands.

Second, hypoxic events can impact brain homeostasis. Sleep fragmentation and sleepiness might result from either hands-on-care (especially in infants, [153]) or from desaturation events (especially in apneic patients, [209]). Sleep fragmentation is able to impact the daily behavior of both adult and NICU patients and is commonly considered a category of pain scoring [100]. Interestingly, Pitson et al. did not only show that the sleepiness and desaturation loads are related in apneic patients, but SpO_2 appears to be related to heart-rate and EEG arousals, intended as increases in frequency [209]. These EEG arousals can be seen in the increase of EEG regularity (Figure 7.3), while the relationships among SpO_2 dips, heart-rate and EEG arousals might support the higher connectivity between EEG and HRV in the 34 weeks group (Figure 7.4). In adults, those physiological fingerprints might be the sign of an altered cardiovascular control [127] or disrupted emotional regulation by the prefrontal cortex [29]. Based on these results, one might speculate a possible impact on the brain development and the autonomic regulation development of those infants. However, the exact mechanisms responsible for those events remain still unclear even in adults and further research is still required.

The increase of EEG regularity and desaturation is normally a feature of the first two monitoring groups (Figure 7.3 and 7.4), while the PSG group is characterized by a greater vagal activity in case of stress exposure (Figure 7.5). Furthermore,

Figure 7.1 and Figure 7.2 show better classification performance for the PSG data. One might speculate that the effect of stress on the patients' physiology might be easier to discriminate due to a lower apnea - bradycardia burden with increasing age and the overall maturation of the ANS [58]. The autonomic development can also explain the increase in performance of cardiovascular features ($HR - SPO_2$) at PSG, while the dominant features are EEG and BH in the first two recording groups (Figure 7.2, Second Panel). It seems that stress initiates a desaturation effect and regular EEG patterns in the first days of life, while the stress-related HRV patterns only arise at full-term age with the maturation of ANS. It is possible that regular EEG patterns are especially present at younger age because of enhanced hypoxia by hands-on-care [153] or a more dysmature EEG. Hypercapnia and reduced cerebral blood flow are common factors to enhance discontinuity of the cerebral activity [264],[267]. However, the discontinuous EEG might also be triggered by the cumulated pain of the NICU, which increases neonatal burst activity [236]. In general, dysmature EEG patterns are especially present at younger age and any EEG disruption might be the consequence of subtle effects that can impact the later-life outcome [263]. This relationship between regularity and dysmaturity might further support the hypothesis of an effect on brain development due to enhanced desaturation and exposure to stress.

Similarly to Lin et al. [155], the interaction between the EEG delta waves showed a strong positive correlation, which increases during the bradycardia spells and under stress exposure (Figure 7.4). This stronger positive interaction between the slow rhythm of the EEG and the HRV is normally concomitant with a vanishing negative modulation when a sleep state shifts from deep sleep to wake [23],[155]. This sudden increase in connectivity might indeed be caused due to an underdeveloped parasympathetic control, and the hypoxia event might be considered as a sudden shift towards an awake state. Apneas and other respiratory events are known to lead to sleep fragmentation [153] and therefore this increase in connectivity might be a consequence of this sleep disruption. Bartsch et al. [23] showed that awake and REM states exhibit stronger physiological connectivity than deep sleep. Especially, the brain-heart interaction increases during REM and awake [158]. It is possible that the combination of bradycardia and stress exposure might lead the subject to a condition closer to an awake state, with an overall increase of network connectivity.

However, this study has limitations, which were already considered in the clinical studies by Levy et al. [153] and Janvier et al. [122]. Bradycardias and apneas are physiological events, whose frequency and severity vary throughout the hospitalization [122]. Therefore, there could not be enough events to classify stress levels for the late preterm, since there are fewer bradycardias and apneas

at full-term age. Moreover, the definition of stress or hands-on-care might also influence the design of the classification. Although Levy et al. pointed out that the clinical handling initiates apneas or hypoapneas, technical contact was also likely to induce desaturations [153]. This study relies on a specific pain scale (LPS), but future research could involve different multidimensional pain scales to have a more in-depth view of the preterm physiology under stress [125]. The definition of bradycardias or the physiological events of interest might also impact the current analysis. Levy et al. pointed out the different consequences of clinical handling, which does not only include apneas, but also sleep fragmentations, hypoapneas and general desaturation events [153]. Gee et al. had a more generic approach, which include all possible bradycardias in his prediction analysis [92]. Specifically, Gee et al. considered any heart-rate drops for more than 1.2 sec as bradycardic event [92], while Paolillo focused only on bradycardias that last for more than 4 sec and were concurrent to desaturation events [195]. Based on the fact that the most dangerous de-oxygenation happens with bradycardias [205], the pursued strategy of this investigation focused on events that looked both to desaturations and bradycardias, but it might be possible to reconsider the entire analysis to have only bradycardias. However, the long-term studies on stress aim to quantify the impact on the development of early-life experiences in the NICU and the specific effect of hypoxia was proven detrimental for the development outcome of preterm patients [122]. This study might also be complemented by a longitudinal analysis, using repeated measurement ANOVA or a balanced linear mixed-effect model. However, it presents an event-based dataset, where the number of bradycardias vary for each patient and recording time. The number of bradycardias normally reduces with the development of the infant [58] and the uneven distribution of those events risk to make any within-subject analysis invalid and unrevealing. This is also the reason why the three datasets were kept separated instead of merging them in one single cluster of data. The unbalance amount of bradycardias would have favored age as a predominant factor in the stress detection, since most of the bradycardias are concentrated in the first days of life [58]. Therefore, a future study should be designed to monitor bradycardic spell in a longitudinal sense in order to assess whether stress has a persistent effect over the different recordings.

Future steps of this analysis might include a further proof of the development delays in case of apnea load and stress. The multiple attributes derived in this study might be included in a regression model to assess the differences in Bayley scores or other clinical scales [122]. Furthermore, the same methodology can be applied to assess the effect of parents-infant interaction with scales, such as the emotional availability scale [272]. It is important to highlight that stress can be the cause of the physiological instability, but it can also be that the most fragile infants tend to overreact with hypoxic abnormalities. Therefore, future

studies should also focus on causality to establish if stress is the source of the physiological instability. The final goal of the perinatal stress quantification should not simply be a modelling based on physiology, but it should help ultimately to understand the impact of stress on development. Therefore, it is important to show how stress and physiological abnormalities derived from stress might impact the development in terms of Bayley score [101], [70].

In a nutshell, stress seems to induce more intense desaturations, apneic and bradycardic events and cortical activation, which can be the trigger of neurodevelopment impairment. Janvier et al. have shown how apnea burden can impact the patients' development in terms of cognitive and motor outcome [122]. Pichler et al. highlighted how long bradycardias can induce severe cerebral deoxygenation in premature infants [205] and Horne et al. stressed that the cumulated effect of apneas has a long-term negative impact on the cerebral oxygenation of the patients at 5-6 months corrected age [109]. Therefore, an exacerbation of respiratory or hypoxic events due to patient handling or procedural pain can ultimately affect the development of the preterm infants. This analysis shows how perinatal stress is associated to cardiorespiratory abnormalities and other physiological instabilities. These findings might provide a possible link between stress and the maturation of premature infants and explain how stress-related physiology can influence the development outcome of those infants.

7.5 Summary

In this chapter, we investigated the relationship between stress experience and bradycardias in preterm infants by means of physiological data. Two main findings have been observed. Larger desaturation levels are associated to stress experience. Larger brain-heart synchrony and EEG regularity are observed during hypoxic events linked to procedural pain. The results show that an automatic stress discrimination in premature infants can be implemented assessing the information of the bradycardic spells. In addition, a possible link between stress and neurodevelopment can be envisaged. The enhanced autonomic and hypoxic events we found in stressed infants might impact their frontal cortex activity, which could ultimately affect their developmental outcome. Future research might be required to test this hypothesis.

Chapter 8

A perinatal stress calculator for the neonatal intensive care unit: an unobtrusive approach

This chapter has been published as Lavanga M., Smets L., Bollen B., Jansen, K., Ortibus E., Van Huffel S., Naulaers G., Caicedo A. (2020). "A perinatal stress calculator for the neonatal intensive care unit: an unobtrusive approach". Physiological Measurement, 1–26. Lavanga M. has co-developed the methodology, conducted the experiments and has written the manuscript. Compared to the publication, minor textual and notational changes have been implemented for better integration in this thesis.

As anticipated in chapter 7, an early experience of pain and stress can affect the neurodevelopment of the preterm infant, but an automated method to quantify the procedural pain or perinatal stress in premature patients does not exist. In this study, an unobtrusive method based on the analysis of physiological signals to detect stress is presented. Without any pain elicitation protocols, features were extracted from the EEG and heart-rate variability in both quiet and non-quiet sleep segments to develop a stress classifier. The reported results suggest pain and stress processing in preterm neonates might cause a dysmature EEG. Therefore, the automatic monitoring of EEG and HRV can be helpful to assess stress load in premature patients via supervised classification strategies.

8.1 Introduction

Early experience of pain and stress in premature infants has been under greater scrutiny by the clinical investigators due to the long-term effects on development [100]. Concerns about the impact of infant pain on neurodevelopment have been raised in the 80s [201]. During their stay in the neonatal intensive care unit (NICU), infants undergo different procedures and care which can generate a cascade of behavioral, physiological and hormonal responses [100]. This accumulation of painful and/or stressful procedures can lead to higher stress, which is defined as perinatal stress. Although procedural pain is not necessarily associated to background stress, Ranger et al. [215] pointed out that pain and stress cannot be discriminated in clinical practice and Jones et al. [126] showed that there is an increase in pain reactivity in case of high exposure of background stress. Due to the possible developmental implications and the worse outcomes of premature infants, clinicians are interested to assess the effect of early-life stress during the patient's stay in the NICU and later on in their life. It is important to remember that stress is a broader concept than pain only. However, the pain-related stress is normally investigated since mainly validated pain scales are used as measurements. Therefore, stress or pain experience will be both used as synonyms of *perinatal stress*.

The growing interest in perinatal stress has been shown in a variety of recent studies. Brummelte et al. found that perinatal stress can affect the fractal anisotropy of the subcortical white matter during infant's neurodevelopment [42]. Cong et al. showed how kangaroo care (KC) can alleviate the effect of post-stimulus experience at the autonomic level after heel prick procedure [52]. Similarly, studies that investigated EEG recordings during pain exposure relied on heel lance procedures and inoculation to assess brain synchronization in nociceptive processing, as reported in [237], [235]. Specifically, the authors investigated differences between tactile and noxious stimuli in evoked related potentials (ERPs) on the one hand and the involvement of autonomic circuitry in pain processing on the other hand. Furthermore, handling of patients induces also sleep disturbances that can generate long bradycardias and apneas [16].

However, there is no automated method to classify exposure to stress in premature infants based on physiological signal background. What is more common is pain-patterns classification based on biopotentials information in the adult population. Misra et al. [170] developed a SVM model to classify low-pain and high-pain using EEG recordings in adults. In their study, they have shown that θ and γ power increase in the prefrontal region and β power in the sensorimotor region contralateral to the stimulus, which can predict the pain patterns with an accuracy of 89.58%. Vijayakumar et al. [257] used a wavelet-transform of EEG to derive power in multiple frequency bands and discriminated

multiple pain stimuli in adults by means of random forest classifiers. Other modalities have also been investigated to automatically classify pain patterns in the adult population. Gruss et al. [102] investigated patterns in EMG under thermal stimulation. A noteworthy aspect was not just the automatic approach to discriminate between baseline and pain patterns, but also the reported accuracy in function of the pain intensity. Another example comes from Brown et al. who used fMRI data to develop a SVM model to determine the absence and presence of pain during heat stimuli [41].

Besides machine learning approaches, the physiological reaction to pain has been investigated by means of the autonomic response or EEG connectivity patterns in adults. Loggia et al. [159] showed that the magnitude of the heart-rate and skin-conductance (SC) sweeps increase with increasing thermal pain stimulation. Functional cortical connectivity can also vary according to stress or pain disorders, as reported by De Tommaso et al. [63] and Imperatori et al. [115]. Migraine patients tend to have a greater connectivity at resting state than controls in laser-induced pain experiments [63]. In addition, post-traumatic stress disorder seems to lead to a higher power in the θ band and higher connectivity in the α band [115].

The objective of this study is to develop a classification model to detect the presence of stress exposure in premature infants. A binary model to discriminate between stress and low-stress patterns was derived by multiple features extracted for background physiological activity. As mentioned earlier, perinatal stress is here intended as pain-related stress. Since there is no clear definition of stress in the NICU and there is no consensus on the level under which the effect can be considered negligible or low, different levels of accumulation of procedural pain have been investigated to define stress. The final goal is to demonstrate if stress is associated to any specific abnormalities in the physiological signals. The detection of functional disruptions is of paramount importance to understand if stress can impact the future development of infants via physiological instabilities.

8.2 Materials and methods

8.2.1 Patient sample

Data from preterm infants were collected as part of the Resilience Study, which has been carried out in the Neonatal Intensive Care Unit (NICU) of the University Hospitals Leuven, Belgium. A total of 136 preterm infants born before 34 weeks gestational age (GA) and/or with a birth weight less than 1500 g was included in the cohort from July 2016 to July 2018. The exclusion

criteria and information related to the research protocol or the inform consent are reported in Section 1.2.

8.2.2 Data acquisition

Patients' pain levels were daily recorded at the cot-side. Pain scores were assessed using the Leuven Pain Scale (LPS), a validated multidimensional pain scale for preterm infants [7], [6]. The LPS assigns scores (0,1 and 2) to seven categories (facial expression, crying, irritability, drowsiness, muscle tension, comfort, and heart rate). Thus, LPS scores vary between 0 and 14, with higher scores corresponding to higher pain intensity. A score of 4 is regarded as the critical threshold for distinguishing a comfortable condition from an uncomfortable one (not pain-free and/or acute pain). The inter-rater reliability was 0.88 [7],[6]. Pain assessment was done routinely by bed-side nurses that were familiar with the LPS and noted it in the patient's electronic medical file. As part of the clinical routine, LPS was scored hourly in preterm infants receiving intensive care, and every three hours in infants receiving intermediate care.

Following the definition of stress exposure, stress was regarded as the experienced pain in the day before the experimental recording, that is the presence of non-zero LPS in the patient record. As already mentioned, a common threshold for uncomfortable conditions for LPS is 4. However, in line with [102], multiple thresholds for experience of pain in the day before the recording were tested: $LPS > 0$, $LPS > 1$, $LPS > 4$ to define a patient under stress conditions. The main assumption is that a greater threshold should lead to a better stress classification performance.

EEG and ECG measurements were simultaneously measured at three different time points for at least 3 hours. Importantly, infants had to be clinically stable at the time of recording. The first recording took place around 5 days after birth (5days). The second recording was planned around 34 weeks postmenstrual age (PMA) (34w). For infants born at 33-34 weeks GA, only one of the two recordings from birth was performed. A last recording consisted of a 24-hour polysomnography (PSG) that was conducted in the week before discharge home. The parents only consented for the polysomnography and opted out for additional EEG measurements. In the course of their NICU stay, some infants were transferred to level II units in hospitals closer to home. Therefore, not all infants have multiple recordings and some LPS measures are missing. However, we strived to readmit infants to our hospital for the PSG measurement. A total of 245 recordings had corresponding pain scores available and were analyzed. Table 8.1 summarizes the clinical characteristics of patients at each measuring

Table 8.1: Summary of patient data set at different time points: GA (gestational age), birth weight (in g), PMA (postmenstrual age) at EEG and ECG recording. Data are median [interquartile range].

	5days (n=118)	34w (n=67)	PSG (n=117)
GA($weeks$)	31.14 [28.86-32.43]	28.86 [26.86-30.71]	30.29 [27.29-31.71]
Birth Weight (g)	1475 [1120-1725]	1140 [900-1480]	1225 [950-1540]
PMA($weeks$)	32.14 [30-33.43]	34.14 [33.86-34.29]	38.43 [37.29-39.57]
LPS	1 [0-3]	0 [0-2]	0 [0-2]

point. EEG data were collected according to the 10-20 system using nine monopolar electrodes (F_{p1}, F_{p2} , C_3, C_4 , C_z, T_3, T_4, O_1, O_2) and monitored with the OSG system (OSG BVBA, Brussel). Each EEG signal was referenced to C_z, which was then excluded from further analysis, leaving a total amount of 8 channels. ECG was then used to derive the tachogram or HRV signal as subsequent R-peak to R-peak intervals (RRi). The R-DECO toolbox [172] was used for R-peaks detection.

8.2.3 Preprocessing and sleep-stage analysis

EEG was band-pass filtered between 0.5 and 20 Hz and independent component analysis was used to remove EOG artifacts. To determine whether the channels were contaminated by movement-related or non-cortical artifacts, multiple criteria were applied on a window-basis in the feature extraction step. Criteria were as follows: standard deviation below 50 μV, absolute difference sample-to-sample below 50 μV and absolute amplitude below 200 μV [116]. If more than 4 channels had an artifact in a window, that window was excluded.

In order to compensate the effect of premature ventricular contractions, the RR intervals were corrected as discussed in [172]. In addition, the R-peak locations were also used to derive the modulation signal $m(t)$ via the integral pulse frequency modulation (IPFM) model, which also accounts for premature contractions [15].

In order to assess whether sleep-wake ciclicity can influence stress classification, sleep states were automatically derived from the EEG recordings. Sleep is characterized by the development of behavorial states with the maturation of the infant and can be divided in quiet-sleep, active sleep and awake [8]. Generally, the EEG signal is relatively more discontinuous during quiet sleep, while the other two states present a more continuous tracing as well as a higher variability of the cardiorespiratory pattern and more body movements [66],[64].

Most of the data-driven algorithms focus on the detection of QS, while the other two states are normally merged in one single state defined as non-quiet sleep. The probability of Quiet Sleep (QS) was derived from the EEG signal with 2 data-driven approaches: a Convolutional Neural Network (CNN) model by [10] and the CLASS algorithm reported in [206]. These two methodologies generate a profile for QS occurrence from the full-channel EEG activity, which was used to derive two 20 minutes window, one associated to QS and the other to non-quiet sleep information (nQS). Each of the window was respectively located around the maximum or the minimum of the probability profile for QS and nQS, considering epoch 10 minutes before and after each prominence.

8.2.4 Univariate features extraction

Multiple features were extracted on a single channel basis. These attributes span from the classical spectral analysis to more refined nonlinear approaches and are meant to describe the presence of dysmature EEG and HRV. Following the definition by Pavlidis et al. [196], a dysmature EEG is generally characterized by discontinuity, persistence of slow-waves and general lack of smoothness, while a HRV of a dysmature neurovegetative system is characterized by lower complexity and a prominent slow-wave baseline. Different studies showed that pain elicitation can generate burst-type activity in premature infants [255] and the patients that stay longer inside the NICU are the candidates to have a greater negative impact on the neurodevelopment [100]. Therefore, the relationship among dysmature EEG, dysmature HRV and stress can be investigated. In the next section, a general overview of the algorithm to derive those features will be presented.

Power features

The power spectral density of the tachogram was computed with both the Welch-periodogram and the continuous wavelet transform (see Section 3.1.1). For the periodogram, 70% was chosen as overlap between windows, while analytical Morlet was picked as the mother wavelet for the wavelet transform [39]. In case of EEG, the power was only computed using the Welch Periodogram for the following frequency bands: $\delta_1 = (0.5 - 2]\ Hz$, $\delta_2 = (2 - 4]\ Hz$, $\theta = (4 - 8]\ Hz$, $\alpha = (8 - 16]\ Hz$ and $\beta = (16 - 20]\ Hz$. The power-features were computed in non-overlapping windows of 30 sec and 4 sec subwindows and it was grand-averaged along QS or nQS channel wise [265], [96].

In case of HRV, the absolute powers of high-frequency (HF), low-frequency (LF) and very-low frequency band (VLF) were computed as sum of the PSD bins

in the following frequency bands: $HF = (0.2 - 4]\ Hz$, $LF = (0.08 - 0.2]\ Hz$, $VLF = (0.0033 - 0.08]\ Hz$ [59], which slightly differs from the frequency bands definition reported in [46]. The relative power indices were also derived as $\frac{VLF}{LF}$,$\frac{LF}{HF}$,$\frac{LF}{LF+VLF}$,$\frac{LF}{HF+LF}$. Both, the modulated signal $m(t)$ and the resampled HRV, were used for this analysis. The spectral density was computed in QS and nQS epochs using the entire 20 minutes [71],[2]. The Welch periodogram subwindow was set to 5 minutes. In case of the wavelet transform, the time-frequency values were then averaged in both sleep epochs.

Multifractality

As discussed in Section 2.4.1 and in Section 2.4.2, a dysmature EEG is a more discontinuous signal and both dysmature EEG and dysmature HRV are characterized by slow-wave persistence. This type of pattern is normally characterized by a lower entropy and higher regularity, due to its higher predictability or memory-persistence. As shown in Chapter 5 and Chapter 7, regular or discontinuous signals might have multiple exponents to control the degree of regularity over time (multifractal signals). An efficient way to estimate the Hurst exponent H and the associated spectrum of singularities $D(h)$ is based on the wavelet transform (Section 3.1.3). The most important parameters to describe $D(h)$ are the c_1,c_2,c_3 parameters, which respectively represent the location of the maximum or the main Hurst exponent, the width and the asymmetry of the singularity spectrum. These multifractal parameters were computed with the WLBFM toolbox in MATLAB [266].

In case of EEG, the triplet c_1,c_2,c_3, the maximal and minimal Hurst exponents and their difference ΔH were estimated for each EEG channel in non-overlapping windows of 150 s and the results were averaged along the 20 minutes window channel wise for both QS and nQS. The window 150 s was picked as epoch to assess the long-term correlation of EEG, according to the findings reported by [65],[211],[69].

In case of HRV, the triplet c_1,c_2,c_3, the maximal and minimal Hurst exponents and ΔH were derived for the entire 20 minutes window for both QS and nQS epochs, in order to assess the slower oscillations which drives the HRV power-law spectrum [2],[71].

Multiscale Entropy

As discussed in Section 2.4.1 and Section 3.1.2, the discontinuity and slow-wave persistence of the EEG can be estimated by means of entropy features.

Multiscale entropy and $SampEn$ were computed to predict stress levels. In particular, $SampEn(m, r)$ and its variations quadratic sample entropy and coefficient of sample entropy $QSE(m)$ and $COSE(m)$ were computed with m spanning in the range $m = 2, 3, 6$, following the findings by Lake et al. [138] and Li et al. [154], while the MSE curve is derived for the scales $\tau = 1, ..20$, as originally discussed by [55]. The complexity index was then derived via the area under the MSE curve, $CI = \sum_{\tau} MSE(\tau)$. In addition, the MSE at scale 3 and 20 ($MSE(\tau = 3)$ and $MSE(\tau = 20)$) were also considered [111]. It is important to remember that $SampEn$ and the MSE at scale 3 represent the information at small scales or high frequency, while MSE at scale 20 represents the information at longer scales or lower frequency. CI is a general measure of irregularity across scales.

In case of EEG, those features were computed in 150 sec non-overlapping windows and grand-averaged along the time course of each sleep state, while the entire 20 minutes window was used for entropy features of the HRV. As already mentioned, the length of the epoch was derived according to the entropy analysis [65] and the multifractal analysis by [211],[69].

Neural features

The dysmature EEG features can also be quantified by means of the NEURAL toolbox [188], available on GitHub. The approach proposed by the authors yields an exhaustive range of features that can be obtained from amplitude, burst and spectral information, connectivity analysis and the range EEG (rEEG). As mentioned by O'Toole [188], the range EEG (rEEG) was proposed as an alternative to amplitude-integrated EEG (aEEG), since a clear definition of the amplitude integrated EEG is missing and different EEG machines implement different algorithms for its estimation. The range EEG represents a filtered, compressed and rescaled version of the EEG: it is normally band-pass filtered between [2-15] Hz, compressed and rectified in the time domain with a windowing procedure and rescaled in a linear-log amplitude scale. More details are reported in [188]. Among all the attributes, it is worth to mention that rEEG was used to quantify the level of dysmaturity by several morphology indices, such as its lower and upper margin, its standard deviation and the rEEG asymmetry. These two margins represent respectively the 5th and 95th percentiles of the rEEG, while the rEEG asymmetry expresses the difference in distance from the median to these 2 margins. More details are reported in [188]. These features were derived only for the EEG and they were obtained using windows of 2 s with 50% overlap for spectral features and 2 s windows without overlap for amplitude features. All the indices were then grand-averaged for each sleep state.

8.2.5 Multivariate features

Features were also extracted in a multivariate fashion to describe how signals interact in case of stress. Specifically, EEG cortical connectivity and brain-heart interaction were investigated. The former was estimated by means of functional connectivity methodologies [96], while the latter was derived via the wavelet coherence [207] and the phase dependent dynamics [220].

EEG connectivity

Functional cortical connectivity was analysed by means of lagged and instantaneous connectivity as well as synchronization of oscillatory processes from the EEG channels [96], [182], [145]. The greatest advantage of lagged connectivity is the resistance to volume conduction and low spatial resolution [240]. Therefore, the magnitude squared coherence $k^2_{xy}(f)$, the imaginary coherence, $I_{xy}(f)$, phase locking value PLV_{xy} and phase lag index PLI_{xy}, as described in Section 3.2.1. The statistical validity of each coupling was then tested with amplitude adjusted Fourier transform surrogates (see Section 3.2.3). In case of coherency, the surrogates approach generates a threshold $T(f)$ in function of the frequency, which was used to zeroed the frequency bins in both $k^2_{xy}(f)$ and $I_{xy}(f)$ that didn't pass the surrogate testing. In case of the magnitude squared coherence, the $k^2_{xy}(f)$ was averaged in the bands δ_1, δ_2, θ, α and β. In case of the Imaginary Coherence, the maximal amplitude of $I(C_{xy}(f))$ was then considered in the same frequency bands [182]. The EEG connectivity coherency was estimated using a Welch-periodogram approach for non-overlapping 30 s window. The subwindow was set to 4 secs and overlap to 70%, according to the findings by [39],[96]. Similarly, the Phase Locking Value and Phase Lag Index were computed in non-overlapping 30 s window, based on the previous analysis on premature infants EEG connectivity [96], [145].

Brain-heart interaction: Wavelet Coherence

The interaction between brain and heart can be estimated as the correlation in the frequency domain between the envelope of δ power and the heart-rate variability. The method by Piper et al. consists of the wavelet coherence between the δ oscillations derived with the continuous wavelet transform and the HRV (see Section 3.2.1, [207]). In order to match the temporal scale, both signals were resampled at 8 Hz. In order to assess the discriminatory power of the brain-heart, the coherence was investigated in the following frequency bands: VLF, LF and the combined band of VLF and LF band $VLFLF = (0.033 - 0.2]\ Hz$. The main reason is the frequency shift that HRV undergoes

with maturation of the neurovegetative system from VLF to LF band [111]. In addition, stress information is commonly associated to the low-frequency band of HRV. Therefore, the brain-heart interaction has been tested with LF oscillations as heart time-course. Similarly to the scalp connectivity, the coupling value was derived as mean of the squared time-frequency coherence in the considered band or the absolute maximum of the imaginary part of the time-frequency coherence. The continuous wavelet coherence was derived for the entire 20 minutes in both QS and nQS epochs and the investigation was limited to channels C_3 , C_4 and HRV, following the studies in adults by Faes et al. [80] and in premature infants by Pfurtscheller et al. [204]. In order to assess the statistical significance, we used amplitude-adjusted Fourier transform surrogates similarly to EEG functional connectivity (see Section 3.2.3).

Brain-heart interaction: Phase dependent dynamics

The dependent dynamics between two systems can be investigated via phase reconstruction, as shown by Rosenblum et al. [220] and Section 3.2.2. Their methodology estimates the phase dynamics of interacting oscillators and provides a directionality index, which describes the direction and the intensity of the coupling. The directionality index is a value bounded between [-1,1], which gives an integrated measure of how strongly a system drives (a value close to 1) or how sensitive it is to be driven (close to -1). The main limit of this methodology is the estimation of the phases, since different methods exists with different underlying hypotheses. The phase of a signal can be estimated via the Hilbert Transform in case of an oscillatory process (e.g. a band-passed EEG signal in the delta band). In case of a point-process like the HRV, the phase of the signal has to be derived taking into account this pulse-type of information. Therefore, the modulatory signal $m(t)$ was considered in the reconstruction of the phase dynamics between brain and heart. In addition, since both EEG and $m(t)$ are wide frequency band signals, the phases of the signals were derived as phases of the main carrier frequency from those signals. The main carrier frequency is defined as the frequency with highest power in the signal. In case of the neonatal EEG, the burst activity around 1 Hz is considered the main carrier frequency of the cortical activity [261], while the pacemaker of the modulatory activity is centered on 0.1 Hz, which is between VLF and LF bands [203]. A method to derive the phase of the carrier frequency is the ridge wavelet-transform, which is a wavelet transform that finds the carrier frequency by looking at the maximum of the spectrum in each time-point [114]. The directionality index was derived only for the channel C_3 and $m(t)$ and the entire 20 minutes window was used to derive the phases of the two signals and compute the interaction. The significance for the directionality index was tested with Cycle-Phase permutation testing, which randomly permutes cycles of a signal

phase to generate surrogates (see Section 3.2.3 and [140]). 19 surrogates were derived to reach a level of statistical significance $\alpha = 0.05$. However, the results of the directionality index strictly depends on the order of finite Fourier series to estimate the phase dependencies [220]. Therefore, the order was tested in the range [1:10]. The ridge-wavelet transform was estimated with MODA toolbox [51], while the directionality index was derived with the DAMOCO toolbox [135] (both implemented in MATLAB).

Graph theory

In multiple points of this thesis, the concept of graph was introduced (see Section 3.2.4, Chapter 4 and Chapter 7). Specifically, a multivariate methodology generates an adjacency matrix $A_{ij} = C_{x_i \leftrightarrow x_j}$, where $C_{x_i \leftrightarrow x_j}$ is coupling among nodes, and a set of topological and resilience indices can be computed to describe the physiological network. Namely, the path length, the clustering coefficient and the eccentricity are the most common indices to assess the network's level of integration. The number n_{sup} is the metric used to assess resilience and represents the number of removed connections that maintain the global connectivity high without significant deviation from the original matrix. A full overview of the all graph features is reported in Table 3.1.

In this study, graphs were applied in two scenarios. The first one was EEG scalp connectivity, where the number M of processes was set to the number of EEG channels ($M = 8$). Specifically, the adjacency matrix was derived every 30 s. The topological indices were computed for each window and averaged along the entire 20 minutes of QS and nQS. The second case relates to brain-heart connectivity computed with wavelet coherence. In this scenario, M shrunk to 3 because only two EEG channels (C_3 and C_4) and HRV were considered. Since the coupling method is based on a time-frequency approach, the adjacency matrix was computed for each time sample of the 20 minutes window and the derived topological indices were then grand-averaged for each sleep state.

8.2.6 Sleep-based Classification

A customized software tool was developed using MATLAB built-in machine learning libraries to classify patients into stress or absence of stress conditions. To summarize, the following groups of features were derived for each patient and for both QS and nQS:

- Power features for both HRV and EEG: 7 features for each HRV time-series and 5 features for each EEG channel

- Entropy features: 12 features for each EEG channel and HRV time-series
- Regularity features: 6 features for each EEG channel and HRV time-series
- NEURAL features for EEG: 30 for all EEG channels
- EEG – connectivity topological indices: 76 features for all EEG channels
- Brain-heart connectivity topological indices: 18 features for all EEG channels and HRV combined and 10 directionality indices.

Considering the number of channels, the different frequency bands and the different type of methodologies, 800 features in total were extracted.

The study investigated if there was any association between these features of the physiological background and stress experience in the NICU, which was defined as experience of pain the day before the recording. In order to test the different level of pain in the stress modeling, the following thresholds were considered $LPS > 0$, $LPS > 1$, $LPS > 4$ in the definition of the presence of stress (similarly to [102]). Multiple classification methods were tested such as SVMs and linear discriminant analysis (LDA), but subspace ensemble with LDA has been found superior in separating the two classes. This method was described in Section 3.4.5. The choice of the final method to design the classification model is mainly related to the generalization. Although methods such as SVMs were proven to be superior in classification accuracy in other studies, the extensive amount of features and the investigation via F-score analysis (see below) prioritized a linear boundary and the automatic selection of features via subspace LDA.

Given the extensive number of features, they were included in the subspace ensemble algorithm if intra-feature correlation was below 90% and they had the highest ratio between intra and inter variance ratio [47]. Specifically, the subspace of features has been restricted to 5, since the datapoints in the different datasets can go as low as 50 (1/10 of the training set as a rule of thumb to avoid overfitting) [47].

In addition, the subject's PMA is an important feature covariate given the maturational changes in both EEG and HRV features. Therefore, if there was a significant Pearson correlation between the selected feature and PMA ($p < 0.05$), the linear correction on the generic feature F $F_{corr} = F - b_1 * age - b_0$ was applied, where F_{corr} was the baseline value deprived of age effects. However, this correction works at best if and only if there is a linear dependence between age and the considered feature. The association between the selected features and the stress experience was also tested via a Kruskal-Wallis test, both for

the binary variable $LPS > 0$ and the three groups $LPS = 0$, $LPS = [1 - 4]$, $LPS = 4$.

The classification accuracy was estimated with a leave-one-patient-out (LOPO) scheme. The final accuracy of all test sets was obtained by pooling the class estimates of all patients together. Therefore, only one set of performance indices was obtained for each classifier. Specifically, each model was assessed via the misclassification error (E(%)), the area under the curve (AUC) and the Cohen's kappa between machine learning labels on the pooled test-set obtained with the LOPO scheme. The comparison among the different classification models (SVMs, Subspace LDA, etc.) was based on the AUCs and the Kappa-scores and the same testing procedure provided lower values for the models that are not reported. The classification hyperparameters of the subspace LDA, which are the γ parameter of the discriminant function and the number of ensemble cycles, were tuned via a 10-fold cross-validation in the training set. Like any random-forest family or ensemble method, the subspace LDA has one learner (normally a decision tree, but in this case linear discriminant function) and a number of cycles to tune the classifier with the bagging process [108]. The number of cycles represents also the number of "trees" in your "forest". In a subspace LDA method, the optimization focuses on the γ parameter of the discriminant functions of the different trees and the number of cycles.

The stress classification was tested in the three datasets based on 3 monitoring groups ($5days$, $34w$, PSG) and both sleep states. The stress classification was also tested on 4 different regroupings of the recordings based on PMA: ≤ 32 $weeks$, $(32 - 34]$ $weeks$, $(34 - 36]$ $weeks$, > 36 $weeks$. The reason to test the classification performance on those different groups is twofold. On the one hand, this classification might still be influenced by age in the monitoring groups (especially in the first group, where the frequency of hands-on care and the heterogeneity of PMA and gestational age are the highest). On the other hand, if a classifier has to be implemented at the cot-side, clinicians tend to be interested in a classifier that considers the age of the patient instead of the monitoring group. This age division follows clinical and physiological definition. The 32 weeks is the threshold to define early prematurity, while 36 weeks is the end of full gestation. The period that goes from 32 to 34 weeks is characterized by synaptogenesis from the subcortical areas towards cortical areas, while the phase after 34 weeks experiences the development of the white matter and the reduction in size of the subcortical plate [134].

8.3 Results

The results for the stress classification for the monitoring groups are reported in Figure 8.1 and 8.2, respectively for QS and nQS. The left panel of both Figures reports the AUC, while the right one shows the kappa score. The AUC, intended as measure for accuracy, is above 0.7 for both QS and nQS. Specifically, the AUC of two sleep states respectively lies in the range [0.73-0.97] and [0.70-0.90]. However, the description of the classifier would be incomplete without reporting the agreement between the ground truth and the predicted labels, expressed by Cohen's kappa. The kappa scores span the range [0.24-0.68] for QS and [0.18-0.68] for nQS, which suggests a moderate association between features and the outcome variable, but also an increased AUC due to an unbalanced design of the dataset. In fact, the most noticeable aspect is that kappa does not necessarily increase with increasing pain threshold, but it can also decrease (as shown in the PSG group), which can be due to the presence of an unbalanced dataset.

The results for age groups are reported in Figure 8.3 and 8.4. Similarly to the monitoring groups, AUC can respectively span in the range [0.78-0.93] for nQS and [0.77-0.96] for QS and the kappa score span the range [0.18-0.59] for nQS and [0.20-0.63] for QS. As mentioned in the monitoring groups, the AUC can increase beyond 0.7, however Cohen's kappa shows a moderate association between features and outcome variable or the possibility of an unbalanced dataset. Moreover, kappa score can either increase or decrease in function of the pain score and it especially decreases in groups above 36 PMA weeks in nQS. No substantial differences were noticed between the two sleep groups.

Figure 8.5 shows how the different groups of features contribute to the classification of stress. Results are here only reported for $LPS > 0$. In both sleep states, two main aspects emerge. HRV features consistently underperform, especially in nQS, with AUC never above 0.62. In addition, there is a consistent predominance of EEG and Scalp connectivity features (CONN) for stress discrimination in the young and old age group (respectively ≤ 32 weeks and > 36 $weeks$), while brain-heart interaction features can outperform the EEG or connectivity features in the period from 32 weeks to 36 weeks. This is also confirmed by feature's topoplots and boxplots, displayed in Figures from 6 to 10. Young patients show a marked decrease in complexity and phase-lag-index eccentricity in θ band in case of stress, as reported in figure 8.6 and 8.7. It is important to remember that the lower the eccentricity, the higher the integration in the network (and therefore the connectivity). This persistence of slow-wave activity, higher regularity and connectivity are also present in term babies: Figures 8.9 and 8.11 report a higher power in δ_2 band and connectivity in δ_1 band for $PMA > 36$ $weeks$. In the period between 32 and 36 weeks,

EEG is characterized by a higher rEEG lower margin as well as a higher rEEG asymmetry, as shown in Figure 8.8. On top of that, a higher brain-heart synchrony is present in case of stress, as shown by the lower eccentricity in Figure 8.10. Figure 8.12 specifically reports the results related to the selected univariate EEG features in a three group fashion, i.e. $LPS = 0$, $LPS = [1 - 4]$, $LPS > 4$, and for almost the entire age range investigated in this study. Those boxplots confirm that the reported differences for $LPS > 0$ persist even if the trend is investigated in three categories. In particular, the EEG shows a marked decrease in complexity with increasing level of stress for $PMA \leq 32$ $weeks$ as well as an increased power in δ_2 band for $PMA > 36$ $weeks$ and an increased range EEG asymmetry band for PMA = $PMA = (32 - 34]$ $weeks$.

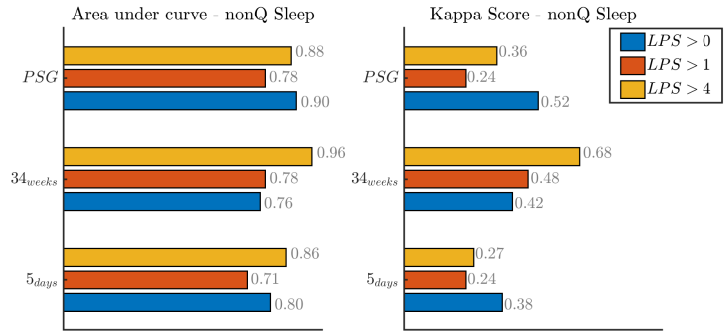


Figure 8.1: Performance for the stress classifier in function of the Leuven Pain Score (LPS) collected the day before the recording. The results are reported for the three monitoring group for the non-quiet sleep epoch: the recordings within 5 days from birth (5_{days}), the recordings at 34 $weeks$ (34_{weeks}) and the polysomnographies at discharge (PSG). The left panel displays the area under the curve, while the right one reports the kappa score. The legend reports the threshold applied on the LPS to define the stress group ($LPS > 0$, $LPS > 1$, $LPS > 4$).

8.4 Discussion

An analysis of brain activity and heart-rate was performed in order to extract features to develop a stress classifier for premature infants at cot-side. Physiological signals, such as EEG and HRV, were collected together with the pain-scale data. The main novelty of study lies on the absence of intrusive methods or pain elicitation protocols to develop the stress classifier. Three main

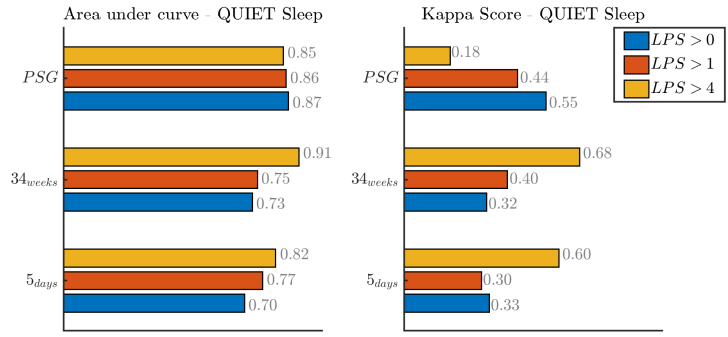


Figure 8.2: Performance for the stress classifier in function of the Leuven Pain Score (LPS) collected the day before the recording. The results are reported for the three monitoring group for the quiet sleep epoch: the recordings within 5 days from birth (5_{days}), the recordings at 34 $weeks$ (34_{weeks}) and the polysomnographies at discharge (PSG). The left panel displays the area under the curve, while the right one reports the kappa score. The legend reports the threshold applied on the LPS to define the stress group ($LPS > 0$, $LPS > 1$, $LPS > 4$).

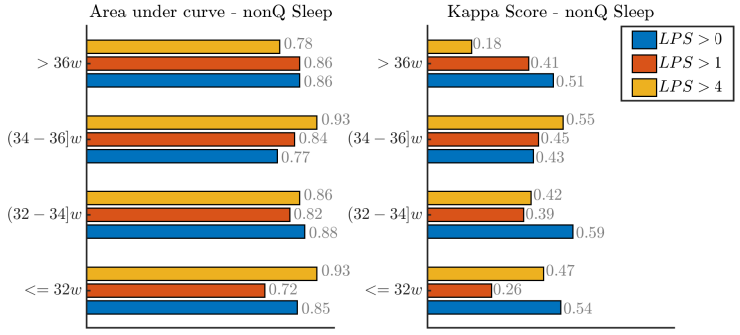


Figure 8.3: Performance for the stress classifier in function of the Leuven Pain Score (LPS) collected the day before the recording. The results are reported for the 4 post-menstrual age groups for the non-quiet sleep epoch: ≤ 32 $weeks$, $(32-34]$ $weeks$, $(34-36]$ $weeks$, > 36 $weeks$. The left panel displays the area under the curve, while the right one reports the kappa score. The legend reports the threshold applied on the LPS to define the stress group ($LPS > 0$, $LPS > 1$, $LPS > 4$).

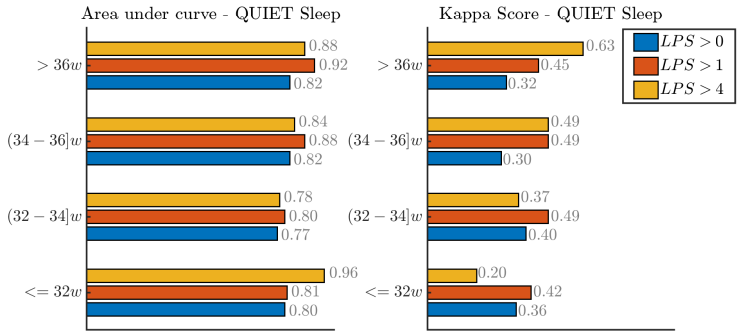


Figure 8.4: Performance for the stress classifier in function of the Leuven Pain Score (LPS) collected the day before the recording. The results are reported for the 4 post-menstrual age groups for the quiet sleep epoch: ≤ 32 *weeks*, $(32-34]$ *weeks*, $(34-36]$ *weeks*, > 36 *weeks*. The left panel displays the area under the curve, while the right one reports the kappa score. The legend reports the threshold applied on the LPS to define the stress group ($LPS > 0$, $LPS > 1$, $LPS > 4$).

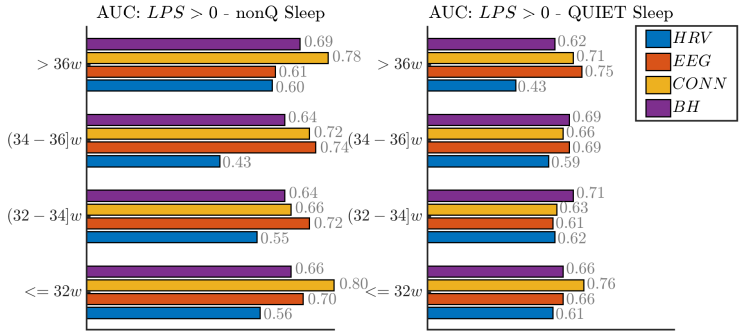


Figure 8.5: Classification performance for each set of features in discriminating stress with threshold $LPS > 0$. The results are reported for the 4 post-menstrual age groups for the quiet sleep epoch: ≤ 32 *weeks*, $(32-34]$ *weeks*, $(34-36]$ *weeks*, > 36 *weeks*. The left panel displays the area under the curve in non-quiet sleep, while the right one reports the AUC for quiet sleep. The legend reports the four sets of attributes : heart-rate variability features (HRV), EEG features (EEG), EEG connectivity features (CONN), brain-heart synchrony features (BH).

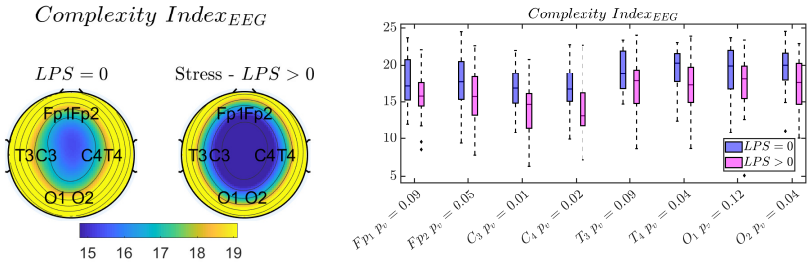


Figure 8.6: EEG complexity shows a marked decrease in case of stress during nQS for patient with *post-menstrual age* < 32 *weeks*. The left panel reports the topoplot of the complexity index on the scalp (area under MSE curve), while the right one reports the boxplot of the complexity index for each channel (p-value reported with Kruskal-Wallis test). The reduction of EEG complexity can be explained by the greater discontinuity in case of stress.

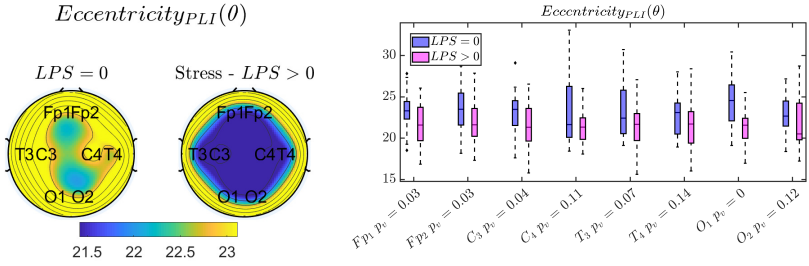


Figure 8.7: The synchrony in the θ band shows a marked increase in case of stress during nQS for patient with *post-menstrual age* < 32 *weeks*. The left panel displays the topoplot for the eccentricity on the scalp (computed with *phase lag index* in θ band), while the right one reports the boxplot of eccentricity for each channel (p-value reported with Kruskal-Wallis test). The eccentricity is a measure of distance, therefore the lower the eccentricity the higher the connectivity.

findings can be reported. First, discontinuous EEG is related with a higher stress load, especially at young age. Second, a higher scalp and brain-heart connectivity are biomarkers of stress experience. Third, the stress classification results are comparable to the pain-pattern classification literature.

The results in Figure 8.6 and 8.9 shows a lower complexity of EEG for premature infants (PMA ≤ 32 *weeks*) and higher δ_2 power in age-matched premature

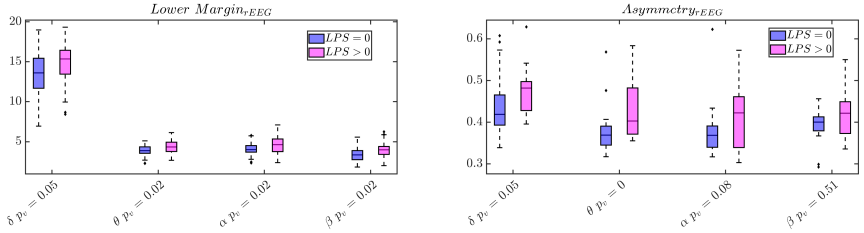


Figure 8.8: The rEEG showed a marked increase in case of stress during nQS for patients with *post-menstrual age* (PMA) between 32 and 36 weeks. The left panel reports the boxplot of rEEG lower margin for the different frequency bands in patients with PMA between 32 and 34 weeks, while the right one reports the boxplot of rEEG asymmetry for the different frequency bands with PMA between 34 and 36 weeks (p-value reported with Kruskal-Wallis test).

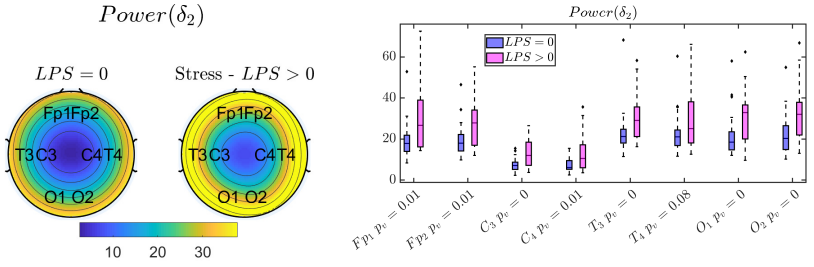


Figure 8.9: The δ_2 power showed a marked increase in case of stress ($LPS > 0$) during QS for patients with post-menstrual age > 36 *weeks*. The left panel displays the topoplot of the δ_2 power on the scalp, while the right one reports the boxplot of δ_2 power for each channel (p-value reported with Kruskal-Wallis test).

patients (PMA > 36 *weeks*). In addition, the subjects show a higher rEEG margin and asymmetry in the period from 32 to 36 weeks, as displayed in figure 8.8. These findings suggest the perinatal stress might increase EEG discontinuity or dysmaturity, which refers to a type of EEG signal characterized by persistence of slow-wave (intended as higher power in δ band and signal regularity), discontinuity (intended as burst pattern which increases regularity) and presence of transient waveforms [196]. The reasons to have a dysmature EEG might be multiple. Figure 8.5 shows predominance of connectivity and EEG features for the younger patients, who have EEG complexity as one of the most discriminative feature (Figure 8.6). This might simply be due to the fact that younger patients are the one to experience higher stress and therefore a

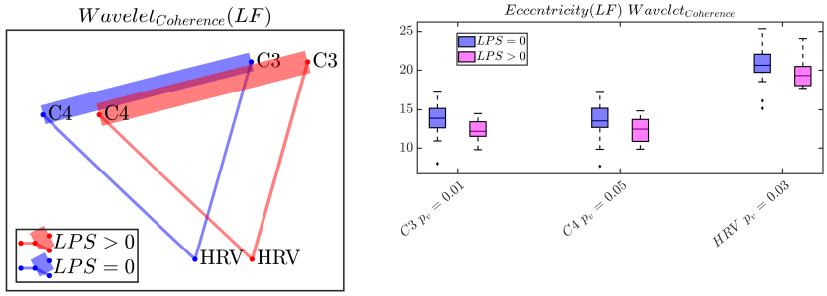


Figure 8.10: The Brain-heart connectivity measured via wavelet coherence in the LF band during nQS for patients with a *post-menstrual age* between 34 and 36 weeks. Stress exposure increases coupled dynamics between brain and heart. However, most of the linked dynamics is shared between EEG channels C_3 and C_4. The left panel reports the graph of brain-heart network, while the right panel displays a boxplot for its eccentricity.

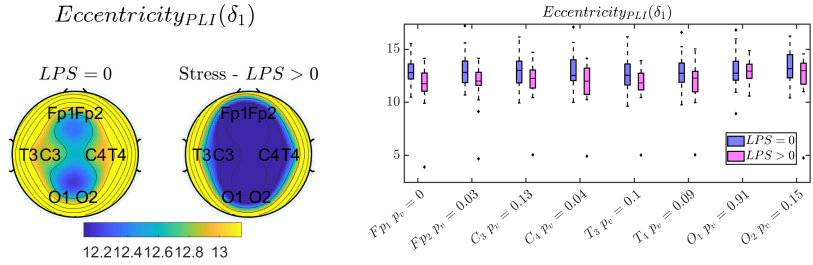


Figure 8.11: The synchrony in the δ_1 band shows a marked increase in case of stress during nQS for patients with *post-menstrual age* $>$ 36 *weeks*. The left panel reports the topoplot of the eccentricity (computed with phase lag index in δ_1 band) on the scalp, while the right one reports the boxplot of the eccentricity for each channel (p-value reported with Kruskal-Wallis test). The eccentricity is a measure of distance, therefore the lower the eccentricity the higher the connectivity.

discontinuous cortical activity might be one of the most discriminative features. However, according to Fabrizi [79], the full development of pain circuitry is completed by 35 weeks PMA, while infants at 28 weeks tend to have immature evoked potentials following painful stimuli. Under pain elicitation, Fabrizi et al. observed delta-brushes and a 10-fold increase in the δ_2 band in premature infants [79], which could explain results in Figures 8.6, 8.8, 8.9. Consequently,

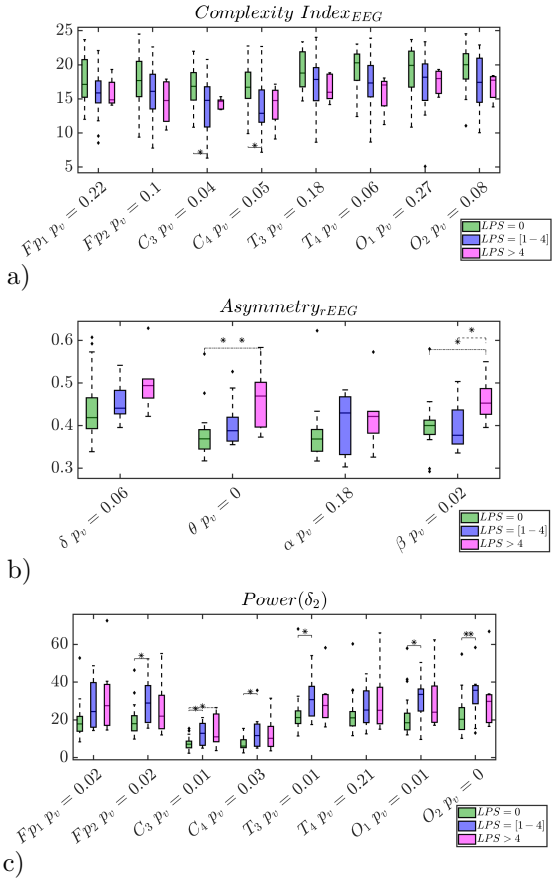


Figure 8.12: EEG complexity in nQS (panel a) for patients with post-menstrual age ≤ 32 *weeks*, the range EEG asymmetry in nQS (panel b) for patients in the range between 34 and 36 weeks and the δ_2 power during QS for patients with post-menstrual age > 36 *weeks* (panel c). Data are reported in three groups $LPS = 0$, $LPS = [1-4]$, $LPS = 4$. The p-values are obtained with a Kruskal-Wallis test and the multicomparison tests significance is expressed by $(^{**})$ for $p \leq 0.01$ and $(^{*})$ for $p \leq 0.05$. Similarly to the binary comparison, higher stress values are related to higher δ_2 power and *asymmetry* as well as a lower complexity. These results show how stress can increase the level of dysmaturity of EEG with a higher level of cumulated pain.

the higher discontinuity might be driven by a higher stress exposure, which is also a hypothesis supported by Figure 8.12. Although all reported features

in this chapter were tested in a three-group fashion, this figure specifically reports the results related to the selected univariate EEG features, which best illustrate the relationship between stress and EEG dysmaturity. If data are grouped in three stress categories ($LPS = 0$, $LPS = [1 - 4]$, $LPS > 4$), EEG complexity, δ_2 power and asymmetry show a trend related with stress levels, as also shown by the Kruskal-Wallis test and the multicomparison tests with the sign (*) for $p \leq 0.05$ and the sign (**) for $p \leq 0.01$. In other words, those boxplot trends might further suggest that an "increase of dysmaturity" can also be due to early-life pain experience. Although conclusions cannot be drawn without a further study with a multiclass framework, the majority of significant multiple comparison tests reveal differences between $LPS = 0$ and $LPS = [1 - 4]$, which might suggest that signs of discomfort are enough to detect differences in physiological data. This might also be linked to the asymmetric distribution reported in Table 8.1 and it suggests that stress can be discriminated with $LPS > 0$. However, the lack of significant comparison tests between $LPS = [1 - 4]$ and $LPS > 4$ can be simply due to a lower number of intense pain patients ($LPS > 4$), which is also highlighted by the asymmetric distribution of Table 8.1. In addition, the infants cortical responses are known to be related to pain intensity and animal models have shown that power in θ band can also increase in case of noxious stimuli, which is further confirmed in the lower margin from rEEG in the θ band (Figure 8.8) and higher connectivity in θ band (Figure 8.7).

Concerning the greater cortical and brain-heart synchrony in case of stress load, the more intense EEG interactions can be the result of a delayed development since the functional cortical connectivity as measured by EEG decreases with maturation [145], [177]. After 32 weeks PMA, neural fibers growth is dominated by interhemispheric connections to form the core of the adult white-matter. At the same time, those fibers grow into the cortical plate together with thalamo-cortical connections, while the subcortical plate diminishes in size [177]. Those events, together with cortex gyrification, lead to the establishment of cortical specialization and they are accompanied by profound changes in EEG, which should be less discontinuous [197] and less functionally connected [177]. The persistent low-frequency rhythms and connectivity can be a sign of delayed development or emergence of altered cognitive processing [130]. Those findings seem to support the fact that stress does not only increase the power, but also the connectivity, in the low-frequency bands (as reported by Figures 8.7, 8.11, and 8.10).

Interestingly, the autonomic activity seems less discriminative in detecting stress exposure, while Figure 8.5 shows that the brain-heart interaction have better performance in stress classification. Stress exposure might increase the synchrony between the cardiovascular activity and the cortical activity, as

reported in Figure 8.10. In addition, Pfurtscheller et al. showed that burst activity might induce a HR response, which is greater in case of closer burst-activity [204]. One might speculate the stress might modulate cortical activity, which can modulate the cardiovascular system. Similar discussion were reported for stroke [48] and epilepsy [225], but further testing is required to prove the presence of a synchronized physiological network in case of stress.

Based on the results provided from Figure 8.1 to 8.4, substantial differences between sleep states are not noticed, as also previously reported by Grunau et al., since EEG responses to noxious stimuli are not sleep dependent [100]. If these results are compared with the literature about pain classification [170], [102], the reported stress classifiers reached similar performances in terms of accuracy. However, several aspects should be highlighted. The pain classification literature usually does not report kappa scores, while this study suggests a moderate association between the predicted labels and the true labels (the Cohen's kappa mostly lies in the range [0.4-0.6]). Gruss et al. [102] reported Cramers's V to support the accuracy rates of thermal pain classification. Although their results suggest a strong association between their features and the level of pain (the variable V mostly lies in the range [0.6-0.8]), Gruss et al. focused on evoked pain, while this study focuses on perinatal stress detection based on physiological background activity. In addition, unlike Gruss [102], an increasing pain threshold does not always lead to an increase of classification performance. In particular, a decreased Kappa is observed for a higher pain threshold in old patients (> 36 *weeks*) and higher kappa for higher pain threshold in younger patients (≤ 36 *weeks*), which can indicate either a milder experience of pain or a more difficult discrimination of stress at older age. This result is line with findings by [73], which indicates a predominant effect of early procedural pain in brain development. Therefore, PMA becomes an essential factor in stress discrimination. This is not only due to the reduced amount of experienced pain later on in the NICU stay, but also the interplay between pain and development which can shift the set of features necessary to discriminate stress. In other words, the subdivision and correction by infant's age was also a necessary step due to different discriminative features for different age groups. As shown by the boxplots in Figure 8.6, 8.7, 8.8, 8.9, 8.10, 8.11, the features able to discriminate stress are mainly EEG and connectivity in the first weeks of life, while the most discriminative features in the last weeks of the NICU are connectivity or brain-heart synchrony related. Consequently, age correction is not only required for clinical purposes, but it is fundamental since the physiological response to perinatal stress differs with patient's development [73]. Furthermore, the employed features are related to the description of the functional maturation of the patients [143],[144],[149]. For sake of completeness, the author of this thesis tried to investigate a full-dataset model (without age subdivision). The classification consistently failed (AUC performances around 0.5) or found in the

PMA the best possible predictor of the stress value. This might be due to the shift in features (the features in first weeks of life cannot discriminate stress in the last weeks in the NICU) and a fundamental importance of age in describing the stress reaction. Therefore, future studies might concentrate on the role of PMA or GA and how the physiological information can further contribute in stress discrimination. However, the effect of postmenstrual age and sleep-wake cyclicity were thoroughly minimized. The analysis based on QS/nQS, the age correction and the division in age groups according to Pavlidis et al. [197] limited the influence of development on the stress classifier. It should also be mentioned that the stress definition was based on the pain score of the day before and represents the cumulated effect of the previous day. Therefore, the classification can be barely influenced by the sleep staging of the recording under investigation and the effect of cumulative pain or stress is expected to influence the physiology of the infant, as also shown in case of pain stimulation [237].

In addition, one may argue that the current analysis is an association or correlation type of analysis, which cannot disentangle the directionality of the stress discrimination. The reported study was meant to design a classifier and detect stress load, but two possible explanations can be found. It might be possible that stress can be the cause of the physiological instability, but it can also be that the most fragile infants tend to overreact to painful stimuli. Therefore, a causal type of analysis is then required to establish if stress can cause the physiological instability. The next chapter will partially address the link between stress and the state of fragility of patient. However, future studies should further deepen the investigation and demonstrate that stress induces physiological instability, which can ultimately impact the development outcome. The final goal of the stress analysis is not only a stress quantification, but it should help ultimately to understand the impact of stress on development. In order to pursue this goal, future investigations should explain if stress causes a reduction in development outcome and if this reduction can be explained via abnormalities in the physiological signals [101],[70].

In a nutshell, the current study suggests that preterm infants that show greater stress have a more dysmature or discontinuous EEG, as reported by the lower complexity (Figure 8.6), the lower rEEG asymmetry (Figure 8.8) and higher power in δ_2 band (Figures 8.9 and 8.12). Although the decrease in EEG complexity and the better performance of the stress classifier in the youngest patients (Figures 8.1, 8.2, 8.3, 8.4) might point out that patients under stress are most vulnerable at lower gestational age, the stress results seem not only to depend on the age of the patient, but they can vary based on the threshold used to define stress. In this perspective, the increase in δ_2 power might be related to the increase of the cortical activity due to accumulation of pain (Figure 8.9).

Furthermore, stress exposure seems to also induce a further synchronization among the different modalities, with an increase in EEG connectivity in θ and δ_1 band (Figures 8.7 and 8.11) and a higher brain-heart synchrony (Figure 8.10). Connectivity seems also to fundamentally contribute to the stress classification, as highlighted in Figure 8.5. Up to our knowledge, this is the first study that tries to automatically classify stress in preterm infants in an unobtrusive way, which can justify the moderate association between features and stress. Our results show that the relative contribution of the different features and modalities of the stress classification model change with the infants' age. Therefore, pain scores as well postmenstrual age are essential factors to take into account for stress discrimination in the NICU.

8.5 Summary

In this chapter, the first automatic discrimination of background stress in premature infants based on physiological data is presented. The main novelty of study lies on the absence of intrusive methods or pain elicitation protocols to investigate physiological data under stress conditions. This approach could provide insights in the relationship between neurodevelopment and stress impact in the neonatal intensive care unit. Alongside a well-established literature about EEG responses to noxious stimuli and procedural pain, the findings in this chapter show a possible relationship between dysmature EEG and stress. In addition, a more synchronized cortical activity as well as a tighter communication between the slow-wave cortical activity and the cardiovascular activity have been found in case of stress, while the autonomic activity does not show a clear link with perinatal stress. The current findings suggest that an automatic tool to investigate disorganized EEG can be used not only for brain development, but can also be helpful to assess the level of stress in infants at the cot-side.

Chapter 9

The description of the physiological maturation under early-life procedural pain

This chapter has been submitted for publication as Lavanga M., Bollen B., A. Dereymaeker, Jansen K., Ortibus E., Van Huffel S., Naulaers G., Caicedo A. "The effect of early procedural pain in preterm infants on the maturation of EEG and heart rate variability". Lavanga M. has developed the methodology, conducted the experiments and has written the manuscript. Compared to the submitted manuscript, minor textual and notational changes have been implemented for better integration in this thesis.

Premature infants present a higher incidence of cognitive, social and behavioral problems, even there no important complications during their stay in the neonatal intensive care unit. Several authors suggested that stress and procedural pain could play role in the altered development of cognition, behavior or motor patterns. Early-life experience of pain might impact the cerebral development and maturation resulting in significant lower cognitive scores in later life. However, it remains very difficult to assess the impact of procedural pain. This chapter focused on the description of maturation of EEG and heart-rate variability associated to different pain-related stress levels. Based on the physiological signal modeling, the research reports how perinatal stress or early procedural pain can affect the level of maturity of the physiological signals. The investigation

considered both the entire set of available data and a subset of patients with gestational age below 29 weeks.

9.1 Introduction

Premature infants remain at risk of developing neurological sequelae and motor dysfunction like cerebral palsy. Several studies have shown that clinical outcomes, such as cerebral palsy and visual deficiency are mostly linked to severe clinical conditions, such as hemorrhage or leukomalacia [164],[183],[262]. The incidence of cognitive and behavioral deficits is far higher than in full-term infants and often not related to specific cerebral lesions, with 30% to 60% of the preterm population experiencing cognitive, social or behavioral problems or lower cognitive outcome compared to their peers [157],[156],[164]. It is important to note that the infants might not show any sign of specific abnormality during their neonatal intensive care unit stay. This means that also functional abnormalities due to pre-, peri- and postnatal triggers may play a role in these neurodevelopmental deficiencies [131].

Several authors suggest that stress and repeated procedural pain could play a role in the altered development of cognition, behavior and motor patterns of premature infants [100], [238], [259]. Early-life painful experiences seem to impact the brain structure and function on both the cortical and subcortical level [73],[214],[238]. Smith et al. showed that the frontal and parietal cortex volumes are decreased in case of procedural pain [238]. Additionally, the same author showed that the functions and microstructure of the temporal regions are disrupted. Brummelte et al. found that white matter connectivity (expressed as fractional anisotropy) is reduced in case of high numbers of skin breaking procedures (SBPs), which count events like heel lance, intramuscular injection, chest tube insertion and central line insertion [42]. Similarly, Duerden et al. showed that the thalamic volume and thalamocortical projections are affected by pain, especially if early skin-breaking procedures and low gestational age (GA) are considered [73]. In most of the studies, convincing results like the reduction in parietal and frontal thickness were obtained by Ranger et al. [214], even after correction for clinical confounders (such as gestational age, severity of illness).

The reasons why brain region volumes change and reduce in case of stress exposure might be multiple. The possible trigger is the sensitization to pain processing and the lowering pain threshold, which is known to happen in premature infants [86],[101]. Sensitization is here intended as a process to shape nociceptive circuitry. The forced exposure of noxious stimuli in the ex-utero

environment might affect the subplate neurons of the cortex up to conditions of excitotoxicity and apoptosis [258]. This increased rate of neuronal death will emerge as smaller brain volumes or decreased cortical thickness [32]. The experience of pain might also compromise the thalamocortical connections which are also fundamental for the pain development [73]. Noxious stimuli can also affect the glia cells and make them vulnerable to inflammatory responses, which can be triggered by pain stimuli [258]. Smith et al. [238] also argued that exposure to stress can affect the hippocampal volume, which ultimately can affect the temporal lobe connectivity and microstructure.

Several studies suggest that the net effect of this accumulated pain in the NICU might play a role in the lower cognitive and behavioral outcome. Grunau et al. [101] did not only show that the cognitive outcome at 8 months and 18 months corrected age are lower in case of high number of SBPs, but also that infants with a higher exposure to procedural pain show more internalizing behavior [101]. This adverse impact on outcome seems plausible given the fact that pain-related stress seems to reduce volumes of multiple brain regions. A recent review showing that prematurity leads to a lower cognitive outcome, hypothesized that procedural pain might contribute to this dysfunction [131]. It is worth to note that the diagnostic tools that support the link between prematurity and cognitive delay were not just imaging techniques but also functional monitoring, such as EEG and HRV [123],[131],[150].

An important aspect of the stress analysis in the NICU is the tight relationship between pain and noxious stimuli. Ranger et al. pointed out that pain and stress cannot be discriminated in clinical practice [215]. Grunau et al. showed that cortisol, which is a common marker for stress, appears to be lower in case of stress [100]. Although premature infants have an immature hypothalamic-pituitary-adrenal (HPA) axis, cortisol appears to rise at 8 and 18 months corrected age, which might indicate a delayed effect of stress exposure. Therefore, clinicians rely mainly on pain scales in clinical practice [100].

The assessment of the developmental effect of stress via structural imaging is invasive for preterm babies. Measures like MRI require transportation and sedation [169] and, mainly give an impression regarding structural (anatomical) abnormalities. As discussed earlier, pain sensitization is the biological mechanism underlying the relation between pain exposure and adverse outcome and it can be measured in the cortical reactivity of preterm infants. Responses to the heel lance show dispersed neuronal bursts on the scalp up to 35 weeks gestational age, while full-evoked potentials are recorded from 36 weeks onwards [78],[237]. This diffused response that is recorded before 35 weeks is a slow-wave burst with superimposed faster rhythms, which is defined as delta brush. The emergence of delta brushes are also a sign of the somatosensory cortex development and they are expected to disappear after 35 weeks GA [198].

Specifically, Jones et al. showed that the EEG reactivity to noxious stimuli (either as delta brushes or evoked potientials) can be a fundamental link between pain and background stress [126]. Therefore one might expect that the effect of procedural pain can be monitored with functional monitoring, especially the EEG activity and reactivity. Additionally, Jones et al. [126] also highlighted a relationship between heart-rate variability of the infants and the noxious EEG evoked-potentials. HRV is normally monitored to evaluate stress response in adults. Moreover, Cong et al. showed a specific autonomic reactivity to heel lance in premature infants, while monitoring HRV [53].

Different research groups showed that both quantitative EEG and HRV features can be used to follow preterm maturation [65],[110],[143],[149],[192]. Normally, the rate of EEG discontinuity should decrease with increasing postmenstrual age, while the variability of the heart-rate with the development [65],[143]. O'Toole et al. looked at the mathematical quantification of amplitude integrated EEG considering both gestational and post-menstrual age and their findings showed that the lower-margin and the aEEG width between upper and lower margins increase as expected in clinical practice [192]. Curzi-Dascalova et al. showed that heart-rate variability in all frequency bands increase with increasing age [58].

In this chapter, the impact of procedural pain on brain development in a functional perspective is described. Based on the available clinical and quantitative results of EEG and HRV maturation, the aim was the assessment of those functional patterns according to the amount of procedural pain experienced by NICU patients. Specifically, functional growth charts were obtained to understand whether the accumulated pain can induce dysmature patterns in EEG and heart-rate variability and suggest an impact on the developmental outcome of the patient.

9.2 Methods

9.2.1 Participants

As in the two previous chapters, the data investigated in this analysis comes from the Resilience study, reported in Section 1.2. However, a subset of of (n = 92) preterm infants (<34 weeks GA) was considered in this investigation due to the limited available information on early-life procedural pain as *skin-breaking procedures*.

9.2.2 Data acquisition

Medical and nursing electronic charts were reviewed from birth till discharge by a physician and by two GCP qualified neonatal research nurses. Variables that were obtained include, but are not limited to, gestational age (GA) at birth, birth weight (BW), illness severity at birth (e.g. CRIB score [246]) and the number of skin-breaking procedures.

Neonatal procedural pain exposure was defined as the sum of all skin-breaking procedures (SBPs). To define these SBP's we drew on the work of Brummelte et al [42] and Grunau et al [101], as well as the work of Newnham et al. in the development of Neonatal Infants Stress Scale (NISS) [180]. We included medical procedures that yield a score of 4 and 5 on the NISS such as chest tube insertion, central line insertion, heel lance, intramuscular injection, eye examination, and other added painful procedures such as wound care. An overview of all procedures is reported in Table 9.1. SBPs were counted for every category separately for each day. If multiple attempts were carried out, each attempt was counted as a SBP. We calculated the sum of SBPs in the first five days of life, because most of SBPs was concentrated in first days of life (Figure 9.1).

9.2.3 Electrophysiological data collection

Physiological data were measured at three different time points in each patient. The recording protocol comprised a first measurement within five days from birth (5days), another session at 34 weeks post-menstrual age (PMA) (34w) and a last measurement before discharge to home. Measurements lasted for at least 3 hours as a rule, infants underwent a 24-hours polysomnography before leaving the NICU, therefore the last session was labeled as PSG. For some patients the recording schedule was adapted. For infants born at 33-34 weeks GA, only one recording (representing the first two coinciding recordings from birth) was performed since their PMA was around 34w when they were 5 days old. Some infants were transferred to level II units in hospitals closer to home. Therefore, not all infants have multiple recordings and, importantly, some SBP measures could not be tracked. However, we strived for the readmission of infants to our hospital for the PSG measurement before they went home. In total, 92 patients with 222 recordings and information about the number of SBPs were analyzed. Similarly to Brummelte et al. and Grunau et al., the data were split into high skin-breaking procedures (HIGH SBPs) and low-skin breaking procedures (LOW SBPs), using a 50 SBPs threshold at 5 days from birth [42],[73]. More details are reported in the statistical analysis section.

Physiological data included EEG and ECG. EEG data were collected according to the 10-20 system using nine monopolar electrodes (F_{p1}, F_{p2}, C_3, C_4, C_z, T_3, T_4, O_1, O_2) and monitored with the OSG system (OSG BVBA, Brussel). Each EEG signal was referenced to C_z, which was then excluded from further analysis, leaving a total amount of 8 channels. ECG was used to derive the tachogram or HRV signal as subsequent R-peak to R-peak intervals (RRi). The R-DECO toolbox [172] was used for R-peak detection.

Before any functional analysis, both EEG and HRV were preprocessed to remove artifacts and define sleep states. In order to remove movement-related and non-cortical information, the EEG signals were band-pass-filtered between 0.5 and 20 Hz and independent component analysis to filter EOG was then applied. The narrow filtering band is based on the quantitative analysis of discontinuous EEGs by De Wel et al. [65] and of the amplitude integrated EEGs by O'Toole et al. [191]. EEG data were then used to define two processing epochs of 20 minutes each related to quiet sleep slate and a combination of awake and active sleep, which is called nonQuiet sleep (nQS). The behavioral analysis was performed with the algorithms described in [206],[9], the probabilistic output of which was summed to look for the two 20 minute epochs. QS was derived as the a window around the maximum of a quiet-sleep probability profile, while nQS as window around the minimum. Before the extraction of any functional feature, the last step of the preprocessing for both sleep epochs was a segmentation in non-overlapping sliding windows and a threshold filtering according to the following criteria: standard deviation above 50 μV, absolute difference sample-to-sample above 50 μV and absolute amplitude above 200 μV [117]. The size of each window was different for each extracted variable and is specified in following section. It is important to highlight that the window was excluded from the processing if more than 4 channels exceeds one of the criteria thresholds. Since the RR intervals of the tachogram can be affected by premature ventricular contractions (PVC), the RR were corrected as discussed in [172] and the modulation signal $m(t)$ was derived via the integral pulse frequency modulation (IPFM) model to investigate the autonomic nervous system stimulation [15].

9.2.4 Functional measurement

Different features were computed to track the functional maturation of neonates in the NICU and assess the level of dysmaturity in the electrophysiological signals, based on the studies reported by [65],[111],[192]. A definition of dysmature EEG and and dysmature HRV are reported in Section 2.4.1 and 2.4.2, while the features to quantify dysmaturity are thoroughly described in Chapter 3. The subset of variables used in this investigation are reported below.

Table 9.1: Overview of all different SBP categories monitored in the study. The first column is related to all puncture procedures, the second column relate to ventilation procedures and the last one is related to other painful procedures.

SBP categories: Punctures	SBP categories: Ventilation	SBP categories: Other
Venous or arterial blood sample	Placing CPAP tube	Placing thorax drain
Placing peripheral venous catheter	endotracheal aspiration	Care thorax drain
Placing deep venous catheter	Aspiration mouth and nose	Placement gastro-intestinal tube
Placing umbilical arterial and/or venous catheter		Wound care
Lumbal Punction		Surgery
Intubation		Eye examination

HRV Power features The different tones of the HRV were assessed as the absolute power in the high-frequency (HF), low-frequency (LF) and very-low frequency (VLF) band. The band limits were the following ones: $HF = (0.2-4]$ Hz, $LF = (0.08-0.2]$ Hz, $VLF = (0.0033-0.08]$ Hz [59]. The following relative power indices were also derived as VLF/LF, LF/HF, $VLF/(VLF + LF)$, $LF/(LF + HF)$. The power spectral density was computed via the continuous wavelet transform, with analytical Morlet as mother wavelet, as shown in section 3.1.1. The CWT was estimated in both QS and nQS using the entire 20 minutes and power values were averaged over both epochs.

Multiscale Entropy The first index to assess the dysmaturity of EEG was the complexity index computed via the multiscale entropy [40],[65],[55]. As discussed in section 3.1.2, this feature is then derived via the area under the MSE curve, $CI = \sum MSE(\tau)$, which is a general measure of irregularity across scales. It was computed in 150 sec non-overlapping windows and grand-averaged along the time course of each sleep state [65],[69],[211]. Multiscale entropy was computed for each channel and each of these complexity indices was considered in the final procedural pain analysis.

Amplitude-integrated EEG features The second index to assess the dysmaturity of EEG was the range EEG (rEEG) asymmetry [189]. The rEEG represents an estimation of the amplitude integrated EEG (aEEG), since a clear definition of aEEG is missing [188]. The rEEG asymmetry quantifies the level of dysmaturity by the difference in distance from the rEEG median to lower and upper margin of rEEG itself. These two margins represent respectively the 5th and 95th percentiles of the rEEG [189]. Following the guidelines in [189], the

asymmetry was derived for the most important EEG frequency bands, which are $\delta = (0.5 - 4]\ Hz$, $\delta = (0.5 - 4]$, $\alpha = (8 - 16]\ Hz$ and $\beta = (16 - 20]\ Hz$. The rEEG was derived for 2 s windows without overlap and all indices were then grand-averaged for each sleep state.

9.2.5 Statistical analysis

Statistical analysis was performed with MatLab (Mathworks, Inc). A linear-mixed effect model was performed to evaluate the association between functional measurements (HRV power, EEG complexity and asymmetry) and early procedural pain. The latter was defined as the cumulated number of skin-breaking procedures (SBPs) in the first 5 days of life. Subsequently, SBP scores were binarized (LOW SBPs versus HIGH SBPs) with a threshold of 50 SBPs in the first 5 days of life to detect patients who experienced a high number of skin-breaking procedures, and thus a high level of early procedural pain. This threshold was chosen based on the earlier research on the impact of early procedural pain in preterm infants [42],[73]. These authors used a similar threshold for binarizing SBP data. Brummelte et al [42] and Duerden et al. [73] considered procedural pain as the cumulated SBP up to their first MRI scan (normally around PMA 32 weeks or 21 postnatal days). Our first recording was planned to be as close as possible to postnatal day 5. However, sometimes it had to be carried out a few days earlier or later because of the preterm infant's unstable medical condition, or because of EEG equipment or technician availability. In order to correct for the covariate effect of gestational age and post-menstrual age, the final regression model did not only include the binary SBP, but also the two age variables. Consequently, each functional variable was explained by three independent variables (SBP, GA, PMA). Additionally, the regression models were computed for both behavorial sleep states (quiet sleep, QS, and non-quiet sleep, nQS).

The significance of the association between early SBPs and functional measurements was tested in two ways. In order to correct for the multiple comparisons (multiple features coming from multiple channels), the p_{value} of the SBP coefficient was corrected via the false-discovery rate with $\alpha = 0.05$. In addition, the significant contribution of SBP in explaining the maturational trajectory of the functional variables was assessed via the log-likelihood ratio (LLR) test [21]. This ratio test specifically assesses whether adding "early SBPs" in the prediction model reduces the variance of the residuals of a model which contains only the age variables (PMA and GA). These prediction models were tested for both the entire set of patients (n=92) and for preterm infants with a GA $\leq 29\ weeks$ (n=31). We specifically focus on this more coherent group of extremely preterm babies because they are most vulnerable to require intensive

Table 9.2: Summary of patient data set at different time points: SBP (Skin Breaking Procedure at 5 days), GA (gestational age), birth weight (in g), PMA (postmenstrual age) at different recording dates (5days, 34w, PSG) and CRIB score. Data are split in HIGH SBPs (SBP ≥ 50) and LOW SBPs (SBP < 50). Data are median [IQR].

	HIGH SBPs (n=73)	LOW SBPs (n=19)	p_{value}
SBP	60 [56.25-69.5]	31 [24-37]	≤ 0.01
GA($weeks$)	26.57 [25.29-29.5]	32 [29.5-32.57]	≤ 0.01
$\text{PMA}_{5days}(weeks)$	28.57 [26.43-31.57]	32.86 [31-33.57]	≤ 0.01
$\text{PMA}_{34w}(weeks)$	34.29 [33.71-34.29]	34.07 [33.86-34.29]	0.32
$\text{PMA}_{PSG}(weeks)$	39.93 [38.71-42]	38.43 [37.36-39.21]	≤ 0.01
Birth Weight (g)	950 [820-1137.5]	1540 [1230-1800]	≤ 0.01
CRIB Score	2 [1-4.75]	0 [0-1]	≤ 0.01

Table 9.3: Summary of most vulnerable patient data with GA ≤ 29 $weeks$: SBP (Skin Breaking Procedure at 5 days), GA (gestational age), birth weight (in g), PMA (postmenstrual age) at different recording dates (5days, 34w, PSG) and CRIB score. Data are split in HIGH SBPs (SBP ≥ 50) and LOW SBPs (SBP < 50). Data are median [IQR].

	HIGH SBPs (n=14)	LOW SBPs (n=17)	p_{value}
SBP	65.5 [58-73]	40 [32-44]	≤ 0.01
GA($weeks$)	26.14 [25.14-26.86]	28.43 [26.71-28.71]	≤ 0.01
$\text{PMA}_{5days}(weeks)$	26.57 [26.21-28.64]	29.43 [27.54-29.68]	0.05
$\text{PMA}_{34w}(weeks)$	34.29 [34-34.29]	34 [33.86-34.14]	0.08
$\text{PMA}_{PSG}(weeks)$	39.93 [39.29-42]	38.64 [38.14-40]	0.08
Birth Weight (g)	890 [800-1050]	1090 [905.5-1203.25]	0.1
CRIB Score	2.5 [1-6]	2 [1-4.5]	0.64

care and thus more SPBs. Furthermore, the explained variance of coefficient of determination or explained variance (R^2) and the mean absolute error (MAE) were reported for each regression model.

Table 9.4: Results of the linear mixed effect model for the non-quiet sleep (nQS) state in the ENTIRE dataset. Only the features that have a significant association with early skin-breaking procedure are reported. The values B represent the fixed-effect coefficients for the SBP, GA and PMA variables with the associated significance level: the symbol ** means $p_{value} < 0.01$ and the symbol * means $p_{value} < 0.05$. For each feature, the mean-absolute error (MAE), the coefficient of determination and the significance of the log-likelihood ratio (LLR) test are reported for each regression model. CI stands for complexity index and P(HF) stands for HF power in the tachogram.

Feature	B(SBP)	B(GA)	B(PMA)	MAE	R^2	$p_{value}(LLR)$
CI(F_{p2})	-1.5848*	0.2409*	0.6689**	8.0395	0.4755	0.0328
$P(HF)$	0.0029*	0.0001	0.0005**	< 0.0001	0.2079	0.0097
CI(T_4)	-1.3332*	0.2661**	0.5414**	7.1098	0.4207	0.0382

9.3 Results

9.3.1 Patients demographics

Table 9.2 summarizes the clinical characteristics of patients based on SBP conditions, while a summary of the most vulnerable (GA ≤ 29 $weeks$) patient demographics are reported in Table 9.3. Figure 9.1 shows the chart of the skin breaking procedure as average over the different patients. The curve shows the SBP dynamics during the neonatal intensive care unit stay. The average chart is represented by the continuous thick line, while the upper bound and the lower bound represent the error margin. The curve has been defined as $\mu(t) \pm \frac{\sigma(t)}{\sqrt{N(t)}}$, where $\mu(t)$ is the mean of the patient distribution, $\sigma(t)$ is the standard deviation and $N(t)$ represents the number of patient. The variable t represents each postnatal day. The exponential decay clearly highlights that the most of the procedural pain is concentrated in the first days of life and supports the idea to focus the investigation on the association between early skin breaking procedure and functional development.

9.3.2 Relationship between EEG and HRV functional variables and early skin breaking procedure

Early SBPs were found to be significantly associated with a reduction in the EEG complexity index, during QS as well as nQS. Results are summarized in

Table 9.5: Results of the linear mixed effect model for the QUIET sleep (QS) state in the ENTIRE dataset. Only the features that have a significant association with early skin-breaking procedure are reported. The values B represent the fixed-effect coefficients for the SBP, GA and PMA variables with the associated significance level: the symbol ** means $p_{value} < 0.01$ and the symbol * means $p_{value} < 0.05$. For each feature, the mean-absolute error (MAE), the coefficient of determination and the significance of the log-likelihood ratio (LLR) test are reported for each regression model. CI stands for complexity index, P(HF) stands for HF power in the tachogram and asymmetry is the asymmetry of rEEG in the different frequency bands.

Feature	B(SBP)	B(GA)	B(PMA)	MAE	R^2	$p_{value}(LLR)$
Asymmetry(θ)	0.0515**	0.0025**	-0.0233**	0.0047	0.6188	0.0025
CI(C_3)	-2.2227**	-0.1822	1.0768**	9.7388	0.6148	0.0040
P(HF)	0.0035**	0.0001	0.0003**	< 0.0001	0.1292	< 0.0001
Asymmetry(δ)	0.0478**	0.0042	-0.0210**	0.0050	0.5496	0.0054
Asymmetry(α)	0.0558**	0.0058*	-0.0218**	0.0058	0.5595	0.0099
CI(F_{p2})	-1.7589**	-0.2112*	1.0495**	7.9296	0.6483	0.0147
CI(O_1)	-2.1371*	-0.1758	1.0406**	8.8279	0.6309	0.0067
CI(O_2)	-1.7045*	-0.1428	1.0313**	7.4683	0.6740	0.0131
CI(F_{p1})	-1.6282*	-0.2565*	1.0582**	8.0964	0.6506	0.0315
CI(T_3)	-1.7707*	-0.0989	1.0272**	6.1865	0.7286	0.0120
Asymmetry(β)	0.0468*	0.0050	-0.0196**	0.0053	0.5331	0.0281

Table 9.6: Results of the linear mixed effect model for the non-quiet sleep (QS) state in the patients with a GA below 29 weeks. Only the features that have a significant association with early skin-breaking procedure are reported. The values B represent the fixed-effect coefficients for the SBP, GA and PMA variables with the associated significance level: the symbol ** means $p_{value} < 0.01$ and the symbol * means $p_{value} < 0.05$. For each feature, the mean-absolute error (MAE), the coefficient of determination and the significance of the log-likelihood ratio (LLR) test are reported for each regression model. VLF/(LF+VLF) stands for the relative power in the VLF band of HRV.

Feature	B(SBP)	B(GA)	B(PMA)	MAE	R^2	$p_{value}(LLR)$
VLF/(LF+VLF)	-0.0498*	-0.0171*	-0.0098**	0.0061	0.2923	0.0645

Table 9.7: Results of the linear mixed effect model for the quiet sleep (QS) state in the patients with a GA below 29 weeks. Only the features that have a significant association with early skin-breaking procedure are reported. The values B represent the fixed-effect coefficients for the SBP, GA and PMA variables with the associated significance level: the symbol ** means $p_{value} < 0.01$ and the symbol * means $p_{value} < 0.05$. For each feature, the mean-absolute error (MAE), the coefficient of determination and the significance of the log-likelihood ratio (LLR) test are reported for each regression model. P(HF) and P(LF) stand for HF and LF power in the tachogram, while asymmetry is the asymmetry of rEEG in the different frequency bands.

Feature	B(SBP)	B(GA)	B(PMA)	MAE	R^2	$p_{value}(LLR)$
P(HF)	0.0047**	0.0015**	0.0006**	< 0.0001	0.2682	0.0069
Asymmetry(α)	0.0639*	0.0056	-0.0208**	0.0051	0.7015	0.0391
P(LF)	0.0090*	0.0019	0.0014**	< 0.0001	0.3056	0.0216
Asymmetry(θ)	0.0502*	0.0048	-0.0236**	0.0050	0.7383	0.0768

Table 9.4 for nQS and Table 9.5 for QS. Early SBPs were negatively associated with EEG complexity as shown by the negative B coefficients for SBP in both tables ($p < 0.05$) for both the FDR test (after ranking correction) and the likelihood ratio test. The two tables only show the features that passed the FDR test. Specifically, the functional variables are reported for the different sleep states, for the specific channel in case of CI and for the specific band in case of EEG Asymmetry and HRV power.

Figure 9.2 shows the effect of SBPs with the development for channel T_4 in nQS: the recordings associated with SBP ≥ 50 (magenta curve) have a lower complexity with increasing postmenstrual age compared to the LOW SBPs group (green curve). This is also illustrated in the 3D representation, which shows the complexity index of T_4 in function of PMA and SBPs as a continuous variable. The early SBPs is reported as log-transformed continuous variable for the sake of representation. The complexity plane results tilted for both the PMA axis and the SBP axis. This means that the highest complexity (in yellow) is obtained for low SBPs and high PMA, while the lowest complexity is obtained for low PMA and high SBPs.

Table 9.4 and 9.5 report the power in the HF band (for both sleep states) and the asymmetry in theta, alpha and beta bands for QS. Similarly to the complexity index, the charts of HF power and the asymmetry of the θ band are reported in Figure 9.3 for high number of SBPs and low number of SBPs. The asymmetry remains higher in case of a higher exposure to pain with increasing

PMA, while the HF oscillations of HRV have a higher power in case of higher SBPs.

The results for the patients with GA below 29 weeks are reported in both Table 9.6 and Table 9.6. Similarly to the entire dataset, the HF power and the asymmetry in the α and θ band in QS have a significant association with the SBP variable and with a positive coefficient both in terms of FDR test and log-likelihood test ($p < 0.05$). Figure 9.4 shows that the asymmetry of the α band and the power in the HF band stay higher in case of a high number of SBPs throughout the development. Additionally, the LF power also shows a positive association with SBP (FDR and LLR tests with p-value < 0.05), as shown in Figure 5. Both the 3D plotting and the growth charts show that a higher stress load increases the long-term (and short-term) heart-rate variability, whose increase is normally expected with infant's development. In periods of nQS, the only feature that shows a significant association with SBPs is $\frac{VLF}{LF+VLF}$ index by FDR test. However, the LLR test shows only a tendency of association between the feature and the stress variable.

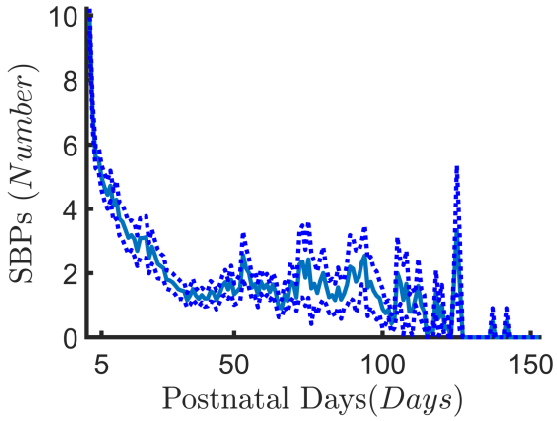
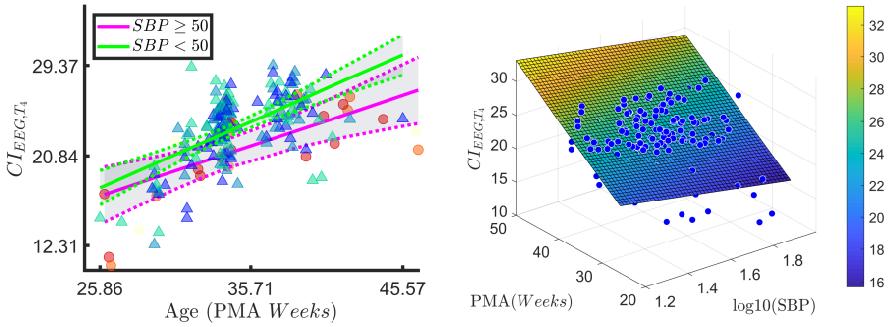


Figure 9.1: The chart of the skin breaking procedure is displayed as average over the different patients. The curve shows the SBP dynamics during the neonatal intensive care unit stay. The average chart is represented by the continuous thick line, while the upper bound and the lower bound represent the error margin. The curve has been defined as $\mu(t) \pm \frac{\sigma(t)}{\sqrt{N(t)}}$, where $\mu(t)$ is the mean of the patient distribution, $\sigma(t)$ is the standard deviation and $N(t)$ represents the number of patient. The variable t represents each postnatal day. The exponential decay highlights that most procedural pain is concentrated in the first days of life.

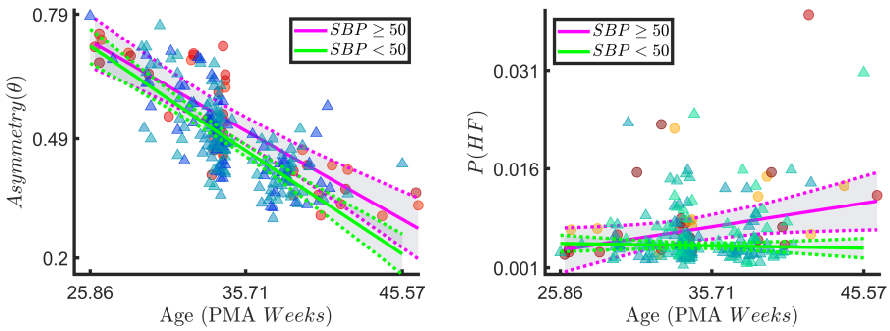


a) b)

Figure 9.2: Complexity Index in non-quiet sleep. The panel a) reports how complexity index trends with development differ in case of high or low amount of skin breaking procedures. The magenta curve shows the trend with SBPs above 50 or HIGH SBPs, while the green curve shows the maturation curve for SBPs below 50 or LOW SBPs. Both the curve shows an increasing CI with increasing age, but the magenta curve has slower development (lower slope) and lower CI at full-term age. The blue/green triangles represent the patient with low SBPs ($N = 73$) and red/orange circles the HIGH SBPs ($N = 19$). The 3D plane confirms the effects on complexity index by stress. The panel b) represents the cloud of complexity index data points for each PM age and each SBPs value and the associated linear plane fitting. The highest value of complexity on the plain is obtained for the oldest age and the lowest SBPs, while the lowest complexity is reached for the youngest age and the highest SBPs. This aspect is also highlighted by the color scale, which goes from blue for the lowest value to yellow for the highest value. The yellow is reported for the highest complexity value on the top left of the plane, while the color blue is reported on the right bottom of the same plane.

9.4 Discussion

Our results showed that a high number of early skin breaking procedures are significantly associated to a dysmature EEG and more variable heart rate. The dysmaturity pattern is here defined as signal with a lower complexity index of the MSE and a higher rEEG asymmetry, which have been proven to respectively increase and decrease with maturation [65],[192]. The association was present after correcting for GA, and PMA at the time of recording. Furthermore, similar results were obtained for both the entire dataset and the most vulnerable patients (GA ≤ 29 $weeks$). Interestingly, the results tend to be similar in QS state (Table 9.5 and Table 9.7), while nQS results were less informative in the



a) b)

Figure 9.3: The figure shows the association between rEEG asymmetry in the θ band (asymm) and early skin-breaking procedure (SBPs) (Left Panel A) and the association between HF oscillations of the tachogram P(HF) and SBPs (Right Panel B). Data are reported for the entire dataset during QUIET sleep (QS). The panels show how the asymm and P(HF) maturational trend differs in case of HIGH SBPs (magenta curve, SBP $\geq$ 50) compared to LOW SBPs (green curve, SBP $<$ 50). SBPs seem to increase the level of asymm throughout the development as well as the power of HF oscillations.

vulnerable group (Table 9.6).

Early pain might act on the infant's development through different mechanisms, mainly sensitization to pain of the central nervous system and lowering sensory threshold [86]. Early pain and injuries might developmentally regulate nociceptive pathways, such as hyperinnervation of the periphery and increasing receptive fields of the dorsal horn of neurons [218], [248]. This peripheral sensitization might lead to the central sensitization, which can affect the central nociceptive pathways and thalamus. Consequently, the thalamocortical projections (whose development peaks at 33 week GA) will be abnormally distributed since their topography and structural organization is activity-dependent [73],[134]. During the development of thalamocortical projections, the axons of the thalamus will contend the activity on the cortex. If early pain disrupts the thalamo-cortical connections, the activity on the cortex will also be different [73].

The net effect of this central sensitization might not only be a specific cortical response as shown by [237], but it might lead to a greater discontinuous tracing on the cortex as shown in Figure 9.2, 9.3 and 9.4. The cumulated EEG response to early pain as observed in [78],[126],[237],[255] might increase the discontinuity of the EEG as accumulated and dispersed delta brush reaction

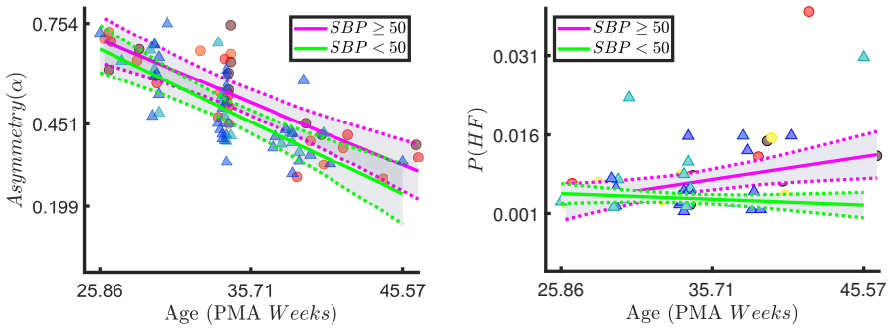
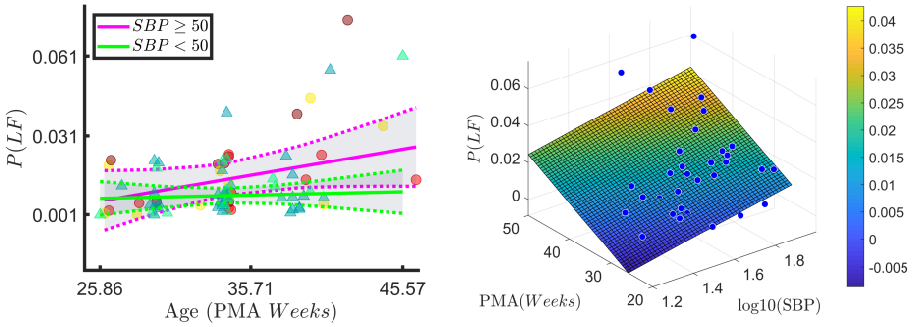


Figure 9.4: rEEG Asymmetry and HF Oscillations: the most vulnerable patients. The figure shows the association between the rEEG asymmetry in the α band (asymm, Left Panel A) and early skin-breaking procedure (SBPs) and the association between HF oscillations of the tachogram P(HF) and SBPs (Right Panel B). Data are reported for the patients with gestational age below 29 weeks during QUIET sleep (QS). The panels show how the asymm and PHF maturational trend differs in case of HIGH SBPs (magenta curve, SBP $\geq$ 50) compared to LOW SBPs (green curve, SBP $<$ 50). SBPs seem to increase the level of asymm throughout the development. Furthermore, the HF oscillations have consistently higher magnitude with HIGH SBPs compared to the LOW SBPs group.

and might ultimately contribute to a more dysmature pattern. Therefore, the greater rEEG asymmetry and the lower EEG complexity are the result of a more dysmature tracing. Clearly, other subtle factors or causes of this persistence EEG dysmaturity cannot be excluded. However, Doesburg et al. [70] also found that a different brain rhythmicity was associated with cumulated pain, which ultimately entailed a lower visual-perceptual abilities, and Grunau et al. showed that pain-related stress is associated with a lower cognitive outcome [101]. In parallel, several authors showed that a dysmature EEG in premature infants is associated with worse cognitive outcome [131],[150],[262]. Furthermore, early-life pain and the resulting disruption of thalamocortical pathways are also associated to a lower cognitive outcome [73]. The results in Table 9.4 and 9.5 and Figure 9.2, 9.3 and 9.4 might provide the functional link which can discriminate infants with a worse developmental outcome: a higher SBP load might induce a more dysmature EEG (potentially, for good part of the NICU stay) and induce a lower development outcome. Currently, follow-up developmental testing is taking place so we will be able to link the current results with developmental outcome measures in the future.



a) b)

Figure 9.5: LF Oscillations in QUIET Sleep: vulnerable patients. The panel a) reports how with long-term HRV ($P(LF)$) trends with development differ in case of high or low amount of skin breaking procedures. The magenta curve shows the trend with SBPs above 50 or HIGH SBPs, while the green curve shows the maturation curve for SBPs below 50 or LOW SBPs. Data are reported for the patients with gestational age below 29 weeks ($N = 31$) during QUIET sleep (QS). Steeper increase in LF power are observed in HIGH SBPs group compared to the LOW SBPs group, which is confirmed by the 3D plane on the right panel b). P(LF) increases in case of higher post-menstrual age (PMA) and high SBPs (in logarithmic scale), while the LF power seems to have a lower value (in blue) when both SBPs and the age are low. This aspect is also highlighted by the color scale, which goes from blue for the lowest value to yellow for the highest value.

Unlike EEG, the results related to HRV might be somewhat counterintuitive. Normally, a lower heart-rate variability in premature infants is associated to a lower mental outcome [123] and the maturation of ANS is characterized by an increase in short and long-term heart-rate variability, which means an increase in HF and LF oscillations [58]. Based on the cortical findings, one would expect that a higher SBP would decrease the HRV power in different bands. However, both HF and LF oscillations seem to increase with increasing age and increasing SBP. The increase in short and long-term oscillations of the tachogram might be the consequence of a faster ANS development to modulate the noxious response. This might be a further consequence of central pain sensitization and the alteration of the stress response at ANS control level and the HPA axis [73],[100],[258]. However, this hypothesis requires further testing in future studies.

A notable difference regards the sleep states: QS substantially presents more significant associations between stress and functional features than nQS.

Similarly to HRV, definitive conclusions cannot be drawn, but it is interesting to notice that the greater number of associations are revealed in the calmest behavioral state with the most discontinuous EEG tracing [8]. This can be a further proof that early-life stress can induce a more discontinuous EEG, with a long-lasting effect to infant's physiology that goes beyond a simple sleep interruption or deprivation. However, QS is the latest state to mature and is typically characterized by a discontinuous tracing, which is one of the main trait of dysmaturity [8]. Therefore, the association between stress and dismaturity might be easier to investigate in this stage than in nQS, which is normally mix of active sleep and awake with a more continuous tracing.

Another crucial point regards the most vulnerable patients, who have GA ≤ 29 *weeks*. Duerden et al. found a stronger association between SBPs and the thalamic volume [73] in case of vulnerability. One might expect that the association between SBPs and EEG dysmaturity is mainly driven by the young GA of the patients with high SBPs. However, the analysis was focused on the subgroup of the vulnerable patients with GA ≤ 29 *weeks* and the link between SBPs and EEG dismaturity remained significant. Therefore, the current results show similar EEG dysmaturity and HRV behavior for both the entire dataset and the vulnerable patient group in case of high SBPs. Interestingly, this similarity is present only for the QS state, while the nQS show different results. The similarity in the calmest state of the sleep cycle might further support the long-lasting effect of stress on sleep architecture and physiology.

9.5 Summary

In this chapter, we quantified the association between SBPs and EEG and HRV functional trends in preterm infants. We found a more discontinuous EEG and a larger HRV in infants that were exposed to high levels of early procedural pain. The larger burst activity might be due to a higher cortical response due to more frequent pain stimulation. The dysmaturity patterns can further prove that central pain sensitization in premature babies might affect their future development and cognitive and behavioural outcome. Pain and stress assessment via functional monitoring might help clinicians to optimize their pain treatment at cot-side. Since the early-life experience in the first days of life is crucial for future development, the use of EEG and HRV might complement clinicians' guideline about analgesia in order to ameliorate the neurodevelopmental outcome of the patients.

Part IV

Autism spectrum disorder and development abnormalities

Chapter 10

Early-life EEG quantitative analysis to predict development disorders in young infants

This chapter has been submitted for publication as Lavanga M., De Ridder J. et al. (2020). "Results of quantitative EEG analysis are associated with autism spectrum disorder and development abnormalities in infants with TSC". Lavanga M. has developed the methodology, conducted the experiments and has written the manuscript. Compared to the submitted manuscript, minor textual and notational changes have been implemented for better integration in this thesis.

Quantitative EEG analysis showed similar traits for the autism spectrum disorder (ASD) EEG and the premature neonates EEG with a poor developmental outcome. However, the quantitative EEG analysis for ASD patients focused only in infants older than 6 months. This chapter provides an investigation of early-life EEG dysmature background in infants with tuberous sclerosis complex (TSC) and its capacity to predict autism spectrum disorder and other developmental problems. EEG data were collected from TSC patients younger than 4 months and ASD risk and neurodevelopmental outcome were assessed at the age of 2 years. The EEGs at first visit were analyzed by means of Multiscale Entropy (MSE), multifractality (MFA) and EEG network features to predict both ASD and the developmental abnormalities using linear discriminant analysis. Quantitative

EEG analysis shows that a dysmature EEG, i.e. a signal with higher fractal regularity and lower entropy, is associated with autism spectrum disorder or abnormal Bayley outcome at 2 years of age. A quantitative analysis of the early-life EEG might provide a unique understanding of the brain functional processing for cognitive disorders or development disabilities, such as ASD, in TSC.

10.1 Introduction

Quantitative and automatic EEG analysis has been proven useful for early detection of autism spectrum disorder (ASD) [34],[202]. An extensive review recently showed which EEG features can be computed to possibly detect ASD, both at rest or during specific tasks [202]. Namely, multiscale entropy (MSE), signal fractality and EEG connectivity have been the subject of multiple studies to investigate ASD risk. By computing the MSE, Bosl et al. obtained a lower EEG entropy in infants at high risk of ASD and a detection accuracy of 70% at age 6 months [40]. Peters et al. [202] found EEG networks with higher resilience in case of ASD, where resilience was defined as a capacity to resist a failure in the network and it is measured as the reduction of global efficiency, after the removal of a node or edge [202],[221].

However, current EEG studies on ASD do not focus on patients below 6 months or close to birth [202],[40],[34], although such early-stage analysis could be beneficial for patients. Especially, patients who are diagnosed with genetic disorders such as Tuberous Sclerosis Complex (TSC) may present a variety of different outcomes. Tuberous Sclerosis Complex (TSC) is a multisystem and autosomal dominant and a number of studies have shown that the majority of infants with TSC develop epilepsy and a variety of neurodevelopment abnormalities including intellectual disability and autism spectrum disorder (ASD) [57]. However, studies with a specific focus on quantitative analysis of early-life EEG and TSC–associated Pervasive Developmental Disorders (PDD), especially ASD, were not published.

Research of EEG background abnormalities in premature infants and neonates showed that a dysmature EEG was associated with a poor cognitive outcome [33],[150],[184],[232],[233]. Pavlidis et al. [197] and Watanabe et al. [262] defined the main traits of a disorganized or dysmature EEG in preterm infants with a poor clinical or developmental outcome as discontinuity, persistence of slow-waves, asynchrony or lagged asynchrony among bursts (see Section 2.4.1).

Multiple studies quantified the presence of dysmature EEG in premature infants as a type of background characterized by a lower entropy, higher regularity and

higher EEG connectivity [65],[144],[145],[192],[271]. The state of the art clearly suggests that both the autistic EEG and preterm dysmature EEG presents the same attributes in term of entropy, regularity and connectivity, but a complete overview of all those features was never investigated in young infants at risk of developmental disabilities.

This chapter presents the assessment of the relationship between quantitative early-EEG characteristics, neurodevelopment and ASD in a cohort of infants with TSC. Based on the infants' data in the first months of life, we analyzed electrophysiological data to automatically predict ASD diagnosis and abnormal neurodevelopment. Additionally, we investigated the relationship between early-life EEG and cognitive, motor and language Bayley score.

10.2 Material and methods

10.2.1 Patient sample

This chapter presents an EEG analysis as part of the EPISTOP project, which is a multicentre and prospective study to assess biomarkers of epileptogenesis in tuberous sclerosis complex (NCT02098759) and to analyze whether preventive treatment before seizure onset could improve epilepsy and neurodevelopmental outcome in TSC children at two years of age. Patients with a definite diagnosis of TSC and age $\leq$ four months were enrolled from November 2013 to August 2016 at 10 different sites, after the approval of local ethical committees. Informed consent was obtained from caregivers in accordance with the Declaration of Helsinki.

The final number of patients involved in this EEG investigation was 61. The gestational age (age at birth) of the included patients had a range from 29.71 weeks until 42 weeks. The median GA was 38.14 weeks, as shown in Table 10.1.

10.2.2 EEG recordings

Patients had serial video EEG recordings, but the data processing focused only on the first-EEG, recorded at four weeks after patients' enrollment. The EEG was recorded according to the 10-20 system for at least one hour. The minimum sample rate was 128 Hz and the electrode impedance was below 5 kOhm. Normally, a set of 19 electrodes was used, but a reduced array with nine electrodes was allowed in infants below 3 months corrected age ($F_{p,1}$, $F_{p,2}$, C_3, C_4, T_4, T_3, O_1, O_2, C_z). The quantitative analysis focused only on this set of

electrodes and Cz was used as reference channel (which was then excluded from further processing). The median EEG recording time was 73.42 min (around 1 h 13 min) and the median age at the moment of the recording was 42.57 weeks (around 1 month and 1 week from birth) for the entire pool of recordings. The age at moment of the recording had a range from 35.86 weeks to 57 weeks.

10.2.3 Neuropsychological assessment

In order to assess the relationship between neurodevelopmental outcome and EEG abnormalities, we used the following scales to define developmental outcomes. First, the Autism Diagnostic Observation Scale 2 (ADOS-2) score or Diagnostic and Statistical Manual of Mental Disorders-5 (DSM-5) were used as clinical criteria to evaluate the presence of ASD at 24 months [171]. Second, the Bayley Scales of Infant and Toddler Development III (BSID-III) were used to evaluate the cognitive developmental quotient (DQ), language DQ and motor DQ at two years of age. The median Cognitive DQ was 75, the median language DQ was 68 and the median motor DQ was 73 for the entire dataset.

An overview of the patients' demographics is reported in Table 10.1 for all the EEG recordings and two groups: patient who had ASD diagnosis (ASD) and those who did not have an ASD diagnosis (NO ASD). The table reports the median gestational age (GA or age at birth), the median age at the time of the recording in weeks for two groups and the median BSID-III scores at 2 years. We also reported the total number of ASD diagnoses and TSC mutations caused by the gene TSC2 in the patients' cohort. Figure 10.2 shows two examples of EEG traces from an ASD patient and from a patient without ASD.

10.2.4 Quantitative EEG analysis

The quantitative EEG analysis was based on the extraction of features that are associated to the most important traits of dysmature EEG, which are discontinuity, persistence of slow waves and asynchrony [197]. After EEG preprocessing, the computed features were combined in different classification and regression models to determine the development abnormalities of the patient, as displayed in Figure 10.1. A summary of all attributes is reported in Table 10.2 and Table 10.3. The quantitative EEG analysis was performed with available MATLAB subroutines or toolboxes which are available online.

Table 10.1: Summary of patient data set: GA (gestational age or age at birth), PMA (postmenstrual age or chronological at time of the recording) and development quotients at 2 years of age based on Bayley Scales of Infant and Toddler Development III (BSID-III): cognitive, language and motor scores. The first column reports the demographics of the entire pool of EEG recordings, while the data are split for ASD (Autism Spectrum Disorder) and NO ASD outcomes at 2 years of age in other two columns. Statistics are reported as median [IQR]. The last row represents the number of tuberous sclerosis complex mutations caused by the gene TSC2 in the two groups.

	EEG (n=61)	ASD (n=19)	NON-ASD (n=42)
GA($weeks$)	38.14 [37-40]	38.43 [37-40]	38 [37-40]
PMA($weeks$)	42.57 [40.36-45.68]]	41.57 [39.11-43.43]	43.14 [40.86-46.43]
EEG(min)	73.42 [56.95-93.92]	82.08 [60.47-109.1]	71.04 [55-88.5]
BSID-III Cognitive	75 [65-90.25]	65 [55-80]	82.5 [70-95]
BSID-III Language	68 [61.25-78.25]	62 [50.75-66.5]	71 [65-94]
BSID-III Motor	73 [67-85]	67 [55.75-72.25]	77.5 [70-92]
TSC2	44/61	14/19	30/42

EEG preprocessing

EEG data were band-pass filtered between $[0.5-32]\ Hz$ (FIR filter, with limited ripple and 40 db attenuation) and resampled at 64 Hz. Each channel was split in non-overlapping windows of different length based on the computed feature and each window was kept in the processing if the following criteria were met: standard deviation below 50 μV, absolute difference sample-to-sample below 50 μV and absolute amplitude below 200 μV [117].

Power analysis

The power spectral density (PSD) was estimated in 30 s window with the non-parametric Welch approach (see Section 3.1.1). The length of the subwindow was 4s with 70% overlap. The power was computed in the following frequency bands $\delta_1 = (0.5-2]\ Hz$, $\delta_1 = (2-4]\ Hz$, $\theta = (4-8]\ Hz$, $\alpha = (8-16]\ Hz$, $\beta = (16-32]\ Hz$ and log-transformed. The power features were grand-averaged for each recording along the window dimension. The split of the delta band in δ_1 and δ_2 is explained by the shift in the spectral edge frequency and the power distribution of the EEG during maturation [261].

Table 10.2: Overview of the univariate and multivariate features derived from the physiological signal in the EEG analysis divided in 5 categories: power, multiscale entropy, multifractality, connectivity features derived by the mean squared coherence (MSC) and connectivity features derived by the imaginary coherence. For the univariate feature groups (first three rows), each attribute was derived for each channel (as indicated by the third column). For the spectral features and the connectivity features, each attribute was derived for the main EEG frequency bands (as indicated by the second column). The last column represents the number of features extracted for each category. CI stands for complexity index, $MSE(3)$ and $MSE(20)$ is $SampEn$ at scale 3 and at scale 20 and ΔH stands for the difference between the minimal and maximal Hurst exponents. N_{sup} is the number of superfluous connections, $Clust_{Coeff}$ is the average clustering coefficient, $Path_{length}$ is the path length and $\frac{PL}{GE}$ is the average between long-range and short-range connections.

Features	Frequency bands	Channels	Properties	Number
Power	δ_1, δ_2, θ, α, β	$F_{p,1}$, $F_{p,2}$, C_3, C_4, T_3, T_4, O_1, O_2	Band Power	40
Multiscale Entropy		$F_{p,1}$, $F_{p,2}$,C_3, C_4, T_3, T_4, O_1, O_2	CI, $MSE(3)$, $MSE(20)$, $SampEn(2)$, $SampEn(3)$	40
Multifractality		$F_{p,1}$, $F_{p,2}$, C_3,C_4, T_3, T_4, O_1, O_2	$Hurst_{Exp}$, c_2, c_3, ΔH	32
MSC - connectivity	δ_1, δ_2, θ, α, β		$Path_{length}$, $\frac{PL}{GE}$, $Clust_{Coeff}$, N_{sup}	20
ImCoh - connectivity	δ_1, δ_2, θ, α, β		$Path_{length}$, $\frac{PL}{GE}$ $Clust_{Coeff}$, N_{sup}	20

Entropy features

Section 2.4.1 and 3.1.2 together with Chapter 8 thoroughly described the relationship between the EEG dysmaturity and entropy, either measures as Sample Entropy ($Samp_{En}$) or Multiscale Entropy (MSE).

Following the pioneering studies of [65],[271], the main features derived in this analysis to describe the dysmaturity patterns were:

- $Samp_{En}$ with embedding dimension $m = 2$ and $m = 3$
- The MSE at scale 3 and 20, i.e. $MSE(\tau = 3)$ and $MSE(\tau = 20)$)
- The complexity index $CI = \sum_\tau MSE(\tau)$, which is the area under the MSE curve

The SampEn and $MSE(\tau = 3)$ represent the information at small scales or high frequency, while $MSE(\tau = 20)$ represents the information at longer scales or lower frequency [111]. The latter is fundamental to show the effect of the persistence of slow-waves, while CI is a general measure of overall irregularity or entropy. All features were estimated in non-overlapping 150s windows and grand-averaged for each recording.

Fractality features

Section 2.4.1 and 3.1.3 together with Chapter 5, 7 and 8 gave an ample overview of the relationship between the discontinuity and the slow-wave persistence of the dysmature EEG and the multifractal parameters.

As shown in the previous sections of this book, the distribution of Hurst exponents, known as singularity spectrum (SS), was computed via wavelet transform [121] and can be used to describe the dysmature EEG with the following set of parameters:

- The main Hurst exponent ($Hurst_{Exp}$) at the location of the SS maximum, which represents the main regularity in the signal
- The parameters c_2 and c_3, which respectively represent the width and the asymmetry of SS and measure the "amount of fractals" inside the time series.
- The difference between maximal and minimal Hurst exponents ΔH of the SS, which resembles c_2 and assesses the number of singularities in a signal. The higher the number of singularities, the higher the discontinuity.

These features were computed in non-overlapping 150s windows and by means of the WLBFM toolbox [266]. Subsequently, they were grand-averaged for each recording.

NEURAL features

The dysmaturity of EEG was also quantified by means of the NEURAL toolbox (available on GitHub) [188], which estimates an exhaustive range of features that can be obtained by amplitude, burst and spectral information, connectivity analysis and the range EEG (rEEG). The latter represents an estimation of the amplitude integrated EEG (aEEG) [188]. It is important to mention that rEEG was used to quantify the level of dysmaturity, by means of its lower and upper

Table 10.3: Overview of the NEURAL Features obtained with the toolbox described in [188], where the reader can find more details. Spectral and rEEG features were computed for the main EEG frequency bands (as reported by the second column). max IBI – maximum interburst interval. The last column indicates the number of features extracted for each category.

NEURAL Features	Frequency bands	Properties	Number
Amplitude		Mean, standard deviation skewness, kurtosis envelope mean, envelope standard deviation envelope standard deviation	25
Spectral	δ, θ, α, β	Band power, relative power entropy, edge frequency spectral flatness, spectral difference (Diff)	21
Range EEG (rEEG)	δ, θ, α, β	Mean, median, lower margin, upper margin, width, standard deviation coefficient of variation, asymmetry	32
Burst		Burst percentage, median and max IBI, Number of bursts	4

margin and the rEEG asymmetry, defined as proximity of the median towards one of the two margins, based on the studies of [106]. Besides the rEEG features, the NEURAL toolbox computes statistics of the EEG amplitudes, the power in the different frequency bands and the spectral difference, the burst percentage and the time interval between bursts. The complete list of features is reported in [188] and in Table 10.3. NEURAL features were computed in 2 secs for spectral features with 50% overlap, 2 secs for amplitude features without overlap and 4 secs for connectivity features with 75% overlap. As already mentioned, the NEURAL attributes were grand-averaged for each recording.

EEG connectivity

The Functional cortical connectivity was analysed by means of the Imaginary Coherence (IC) and Magnitude Squared Coherence (MSC), as reported in Section 3.2.1 and in Chapters 4 and 8.

The statistical validity of each coupling was then tested by means of 19 surrogates derived with amplitude adjusted Fourier transform (AAFT) surrogates (see Section 3.2.3 and [95]). The actual MSC coupling was then represented by the average of $k^2_{xy}(f)$ in the following frequency bands δ_1, δ_2, θ, α and β. In case of IC, the maximal amplitude of $|I(C_{xy}(f))|$ was then considered in the same frequency bands [182]. The EEG connectivity coherence was estimated using

a Welch-periodogram approach for non-overlapping 30 s windows, with 70% overlap and 4s subwindows [39],[95].

Graph Theory

Functional connectivity methods such as IC and MSC generate EEG undirected networks: the nodes represent the EEG channels and the weighted edges represent the intensity of the coupling between channels (see [43] and Section 3.2.4). Graphs can be characterized by different topological measures, as reported in [221] and Table 3.1. As short recap, the most common measure of integration is the path length (PL), which is the average length of the shortest paths among nodes in the network. The global efficiency (GE) is the average of all inverse path lengths. The PL is normally driven by weaker connections (or longer distances) in the graph, while the GE is driven by the strongest connections (or shortest distances) [202]. Therefore, one can also measure the long-range over short-range connections ratio as the ratio between PL and GE. Another common index is the average clustering coefficient, which represents the likelihood of two nodes to be connected to a third one. An overview of the computed topological indices for this analysis is reported in Table 10.2. The topological indices were computed by means of the Brain Connectivity toolbox [221].

Recently, Peters et al. investigated the resilience of EEG networks, which is defined as the ability to resist random or target attacks (removal of edges or nodes) and keep the global efficiency close to the original one [202]. As previously discussed, the resilience has been derived as the number of superfluous connections, which are the connections that can be removed without significant changes in the EEG network. The details of this methodology are reported in Section 3.2.4.

10.2.5 ASD and development outcome prediction

The pipeline followed for the quantitative EEG analysis was as follows: first the EEG channels were preprocessed in order to remove artifacts and to improve the quality of the signals. Second, EEG features were extracted in order to quantify dysmaturity. Finally, the dysmaturity features were related to the neurodevelopmental outcome of the infant. In order to meet the four stated objectives and investigate the relationship between EEG abnormalities and developmental abnormalities in TSC infants (at two years of age), four different strategies were pursued, as reported in Table 10.4 and Figure 10.1.

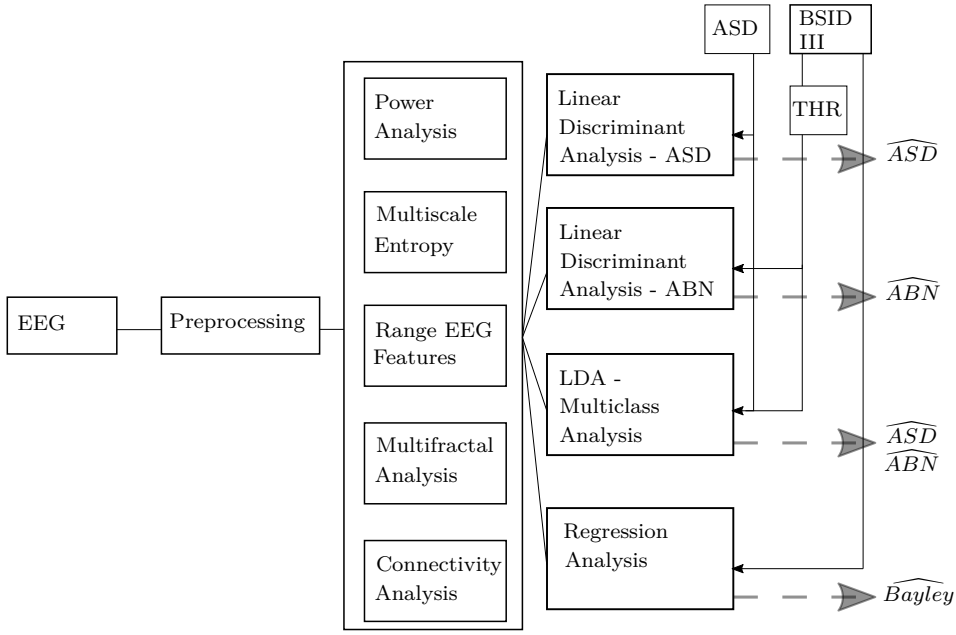


Figure 10.1: A schematic overview of the four strategies implemented in this analysis. The BSID-III stands for the Bayley Developmental Scale, ASD is autism spectrum disorder and THR is the threshold applied to define an abnormal Bayley outcome. After EEG preprocessing, a set of power, entropy, fractality, range EEG and connectivity features are derived to quantify the level of EEG dysmaturity. Those features are fed in different supervised discrimination strategies: 1) binary classification for ASD diagnosis with linear discriminant analysis (LDA), 2) binary classification with LDA for development abnormality diagnosis (ABN), 3) multiclass classification with LDA to determine if infants have ASD or if infants have only an abnormal BSID-III, 4) Regression analysis to extrapolate the Bayley score. The four supervised strategy blocks are fed with both features and the ground truth labels ASD, BSID-III (blue thin lines, BSID-III after threshold THR or the actual continuous variable for the regression analysis) and they have the prediction of ASD or abnormal outcome (dashed grey lines).

Automatic classification of ASD patients

The first strategy aimed to design a classifier to predict an ASD outcome at 2 years with EEG dysmaturity features. A set of 5 features was selected for a

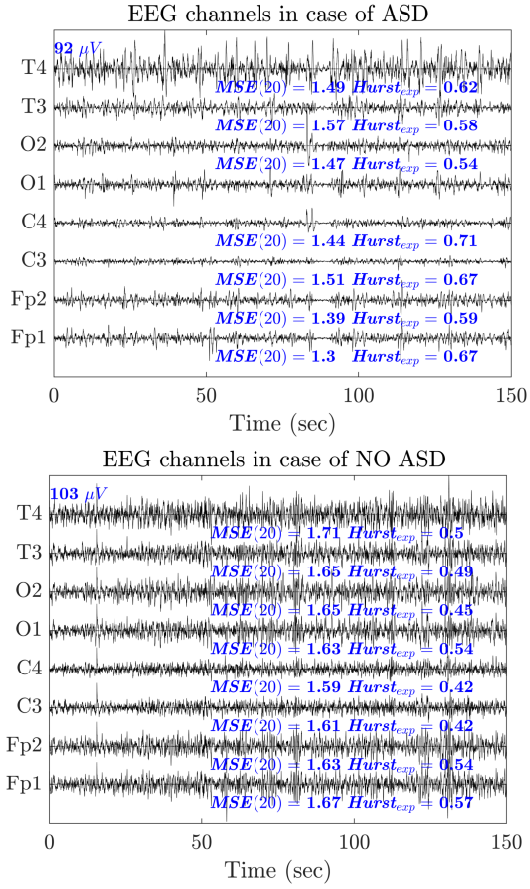


Figure 10.2: The figure shows two examples of full montage of EEG from the no ASD group (NO ASD, upper panel) and the autism spectrum disorder (ASD) group (lower panel). The data from the 8 channels are reported in a window of 150 sec and the associated entropy at scale 20 $MSE(20)$ and the Hurst exponent are displayed if and only if the threshold criteria discussed in the text are met. The ASD EEG has lower entropy and higher Hurst exponent, which leads to a more discontinuous channel with slower rhythms. On the contrary, the EEG of patients without ASD has faster frequencies and higher entropy at larger scale. Therefore, the ASD EEG can be considered a dysmature EEG. The sensitivity of the plotting is reported on the top left of the chart.

binary classification scheme, whose target variable was represented by autism diagnosis (ASD vs NO ASD). The feature selection was based on the highest ratio between higher inter-group variance and lower intra-group variance [47] for each feature group (power, fractality, entropy, NEURAL, connectivity) and for the entire pool of attributes. We selected 5 features in order to guarantee that the number of predictors is below 10% of the total number of patients, which is a common rule of thumb to avoid overfitting [88]. A classifier was implemented with a Linear discriminant analysis (LDA) algorithm with 3-fold testing: the algorithm was iterated 3 times and the classifier was first trained on 2/3 of data as training set and tested on the remaining 1/3. The LDA regularization hyperparameter γ was tuned using 10-fold cross-validation (9/10 of the training set is used for training set and 1/10 is used as validation set). The classification results were reported as classification error, area under the receiver operating characteristic (ROC) curve (AUC, defined as measure of classification accuracy) and Cohen's kappa score (Table 10.4). Cohen's Kappa is a statistical measure of the agreement between machine labels and true labels and it corrects for the agreement expected by chance (Section 3.4.6). The Cohen's kappa is usually a more suitable index for imbalanced datasets. Additionally, the statistical differences among features represented by autism diagnosis (ASD vs No-ASD) were tested with a Kruskal-Wallis test for each separate channel.

Automatic classification of abnormal outcomes

The second strategy aimed to predict an abnormal developmental outcome at 2 years. Therefore, it consisted of a binary classification of patients based on EEG dysmaturity features with a slight modification of the target variable compared to the first strategy. The positive class or abnormal outcome was represented by either the ASD diagnosis, an abnormal Bayley score or both. An abnormal Bayley score was here defined as one or more of the three BSID-III scores at 2 years below a certain threshold. Consequently, this strategy tries to include any possible abnormality (abnormal vs normal), whether it is a developmental disorder like ASD, or a global or specific developmental impairment (e.g. low motor DQ, low cognitive DQ). The definition of abnormal development outcome with Bayley score is linked to the threshold. The clinical cut-off is usually 70 or 80 since these are associated with a lower mental and cognitive outcome [124]. However, the literature does not show a clear consensus on which threshold has to be applied to define developmental delay [75],[124]. Therefore, multiple thresholds were investigated to define the positive class in the range from 55 to 80 with steps of width 3, i.e. [55:3:80]. The threshold 55 represents the minimum Bayley score in the dataset, while 80 is one of the highest thresholds in the literature to define developmental delay [75],[124]. The classifier was designed based on the same LDA algorithm used for ASD detection. The

classification results were reported as AUC and Cohen's kappa score in function of the threshold.

Multiclass analysis of abnormal Bayley outcomes, ASD and normal patients

The third strategy is a multiclass approach, where the target variable splits the patients in normal or total absence of any abnormal outcome (normal), patients with just an abnormal Bayley (ABN-Bayley) and patients with ASD diagnosis (ASD). It is important to notice that ASD is always associated with at least an abnormal Bayley language score (see Table 10.1). Therefore, this strategy has a target variable with three classes in order to assess if EEG dysmaturity features can discriminate abnormalities linked with ASD and Bayley abnormalities. Similarly to the two previous strategies, a classifier was designed via an LDA algorithm and with 3-fold testing (2/3 of data as training set and 1/3 test set). The LDA regularization hyperparameter γ was tuned tuned using 10-fold cross-validation (9/10 of the training set is used for training set and 1/10 is used as validation set). The multiclass classifier was trained according to a one-versus-all scheme. Results were reported as Cohen's kappa score and two AUCs. The AUC_1 was computed as area under the ROC curve for the classification Abnormal Bayley Score vs All and AUC_2 was computed as area under the ROC curve for the ASD vs All. Similarly to the second strategy, different thresholds were tested for the Bayley score. The chosen range was [73,75,80], which represent the best and the worst case scenarios of the second strategy. An overview of the different classification performance metrics is reported in Table 10.4. The statistical differences among features represented by three classes (Normal vs ABN-Bailey vs ASD) were tested with a Kruskal-Wallis test for each separate channel with an associated multiple comparison test.

Regression analysis for the Bayley score

The final strategy used a robust linear regression model to explain the three main Bayley scores. This aimed to investigate how the EEG dysmaturity features explain abnormalities of each score. For each of the three Bayley scales, the five most correlated features were used to develop the robust model to predict the clinical score, together with the presence of TSC_1 mutation. The regression results are reported in terms of explained variance (R_2), mean absolute error (MAE) and the selected features for the regression.

Table 10.4: An overview of different classification and regression strategies used to detect development disorders abnormalities. In the case of simple binary classification, the positive class is represented by the ASD diagnosis (ASD vs NO ASD) and the performance metrics are the area under the curve, the Cohen's kappa and the misclassification error. In the second binary strategy, the positive class is represented by at least one development abnormality (ASD or abnormal Bayley Score) which is represented by one of the Bayley Scores under a certain threshold (ABN vs Normal). In case of multiclass classification, the first positive class is represented by patients with only an abnormal Bayley score, while the second positive class is the ASD patients (Normal vs ABN-Bayley vs ASD). Therefore, one obtains two AUCs: AUC_1 for all vs patients with only an abnormal Bayley score and AUC_2 for all vs patients with ASD. The last strategy is the regression between the EEG features and the BSID-III at two years. The common metrics are the coefficient of determination or explained variance (R^2) and the mean absolute error (MAE).

Assessment strategies	Target variable	Performance metrics
Binary classification for ASD	ASD vs NO ASD	Error, AUC, Kappa Score
Binary Classification for abnormal development	Normal vs abnormal development	AUC, Kappa Score
Multiclass classification	Normal vs abnormal Bayley outcome vs ASD	AUC_1 (ALL vs only abnormal Bayley outcome), AUC_2 (ALL vs ASD), Kappa Score, Error
Burst	BSID-III variable	R^2, MAE

10.3 Results

The results of the quantitative EEG analysis were reported according to the different strategies explained in the section Methods.

10.3.1 Automatic Classification of ASD patients

Power, multifractal, entropy and connectivity features were computed to investigate if ASD was associated with a dysmature EEG and a linear discriminant analysis was used to automatically classify ASD patients. Figure 10.2 visually suggests that dysmature EEG predicted ASD symptoms at two years. This visual relationship is supported by the classification results described in the first strategy. Table 10.5 shows that the different LDA classifiers for the different feature subgroups have the AUCs in the range of 0.66 – 0.79 and the Cohen's kappa in the range of 0.26 - 0.48. The most important

biomarkers in terms of AUC and kappa were MFA features ($AUC(MFA) = 0.74$, $K(MFA) = 0.48$) and MSE features ($AUC(MSE) = 0.79$, $K(MSE) = 0.26$).

Figure 10.3 shows the entropy at higher scale ($MSE(20)$) and regularity ($Hurst_{Exp}$) for the different channels. The Kruskal-Wallis test shows a significant lower entropy at higher scales for the following channels ($F_{p,1}$ – $p < 0.01$, Fp_2 – $p = 0.01$, C_4 – $p = 0.04$, T_3 – $p = 0.02$, T_4 – $p < 0.01$) and higher regularity for the following channels (O_1 – $p = 0.01$, O_2 – $p = 0.05$, T_4 – $p = 0.04$) in patients with an ASD diagnosis by 2 years of age.

Table 10.7 shows that the fractal and entropy features are specifically selected in the ASD classification when the entire pool of features is considered. In fact, all attributes are either Hurst exponents or MSE features, like the complexity index and the entropy at scale 20.

10.3.2 Automatic Classification of abnormal outcomes

The same features were used in another LDA model to automatically classify abnormal development outcome according to different Bayley score thresholds. Figure 10.4 reports the results of the LDA classification of the abnormal outcomes based on dysmature EEG. The AUCs and kappa scores are reported for the tested range of Bayley thresholds to define abnormalities in the development. The second strategy confirms that MSE (red-circles curve) and MFA (green-diamonds curve) outperform the other biomarkers in the prediction of abnormal development outcome. For most of the thresholds, MSE and MFA report the highest AUC and Kappa. The peak-performances are reached at a threshold of 75: MSE reaches AUC up to 0.92 and Kappa up to 0.63, while MFA reaches

Table 10.5: Results for the binary classification models for ASD diagnosis (ASD vs NO ASD). The model was trained with linear discriminant analysis reported in terms of misclassification error (percentage of misdiagnosis, E(%)), area under the receiver operating characteristic (ROC) curve (AUC) and the agreement between machine and clinical labels via Cohen's Kappa.

Age (PMA w)	Error(%)	AUC	Kappa
Power	32.59	0.66	0.27
MSE	33.01	0.79	0.26
rEEG	30.26	0.69	0.27
Conn	37.50	0.61	0.16
MFA	22.12	0.74	0.48
All	26.67	0.74	0.35

AUC up to 0.88 and Kappa score up to 0.45. Beyond this 75 threshold, a drop in classification performance of each feature group has been observed except for the combination of all features.

10.3.3 Multiclass analysis of abnormal Bayley outcomes, ASD and normal patients

Eventually, the same pool of features was used to automatically discriminate normal patients, ASD patients and patients with only abnormal Bayley outcome. The results associated to the multiclass strategy are reported in Table 10.6. The findings of the third strategy are reported for the Bayley thresholds 73,75,80, which represent the peak and the drop of performance in the second strategy (Figure 10.4). In case of the multiclass approach, there is a steady increase of all parameters (AUC_1, AUC_2, Kappa) with increasing threshold, but the values

Table 10.6: Results for the multiclass classification to discriminate patient with normal development, ASD and patients with only abnormal development (Normal vs ASD vs ABN Bayley). The model was trained with linear discriminant analysis. The performance is reported in terms of misclassification error (percentage of misdiagnosis, E (%)) and the agreement between machine and clinical labels via Cohen's Kappa. Since the model was trained according to the OneVsAll scheme, the AUCs are reported as the AUC of the class with only development abnormalities vs all (AUC_1) and the AUC of the combination of ASD vs all (AUC_2). The results are reported for entropy (MSE), fractality (MFA) and the combination of all features. The performance is reported for the Bayley score threshold range [73,75,80] to define the development delay.

Bayley	Error(%)	AUC_1	AUC_2	Kappa
		MSE		
73	36.67	0.60	0.72	0.28
75	43.33	0.55	0.67	0.13
80	30.00	0.65	0.75	0.36
		MFA		
73	33.33	0.66	0.66	0.36
75	33.33	0.65	0.73	0.36
80	30.00	0.71	0.71	0.41
		All		
73	35.90	0.69	0.76	0.33
75	35.90	0.71	0.76	0.32
80	33.33	0.73	0.74	0.35

Table 10.7: Selected features for the binary classification of NO ASD vs ASD and the multiclass scenario when all features are combined together. The results for the multiclass design are reported for the Bayley score threshold range [73,75,80] to define the development delay. CI stands for Complexity Index, while n_{sup} is the number of superfluous connections of the Scalp-EEG network as explained in the main text.

Bayley	Features
	Binary
	$Hurst_{exp}(O_1)$, $CI(T_4)$, $CI(O_1)$, $MSE(20, T_4)$, $MSE(20, F_{p,1})$
	Multiclass
73	$Hurst_{exp}(O_1)$, $CI(T_4)$, $CI(O_1)$, $n_{sup}(\delta_2)$, $Path_{Length}(\delta_2)$
75	$Hurst_{exp}(O_1)$, $CI(T_4)$, $CI(O_1)$, $n_{sup}(\delta_2)$, $Path_{Length}(\delta_2)$
80	$Hurst_{exp}(O_1)$, $n_{sup}(\delta_2)$, $CI(T_4)$, $CI(O_1)$, $Path_{Length}(\delta_2)$

Table 10.8: Regression model to predict the BSID-III score with the most correlated features and the presence of TSC mutation. The features reported here represent the best model to explain the score in terms of mean absolute error and coefficient determination (R2). The first column reports the type of Bayley score, the second column the involved features, while the last ones report the regression performance (R2, MAE and the associated p-value).

Bayley	Features	R^2	MAE	P_{value}
Motor	$MSE(20, T_4)$,$\frac{PL}{GE}(\delta_1)$,$MSE(20, O_1)$, $MSE(20, Fp1)$,$Path_{Length}(\delta_1)$, TSC	0.3048	7.8502	0.0068
Language	$\Delta_H(C_3)$,$c_3(T_3)$, $c_3(C_3)$ $Hurst_{Exp}(C_4)$, $Hurst_{Exp}(O_1)$, TSC	0.3711	7.5014	0.0008
Cognitive	$Diff(\delta)$, $SD(\delta)$,$MSE(20, Fp2)$, $Width(\beta)$, $Width(\delta)$, TSC	0.4564	7.8529	0.0002

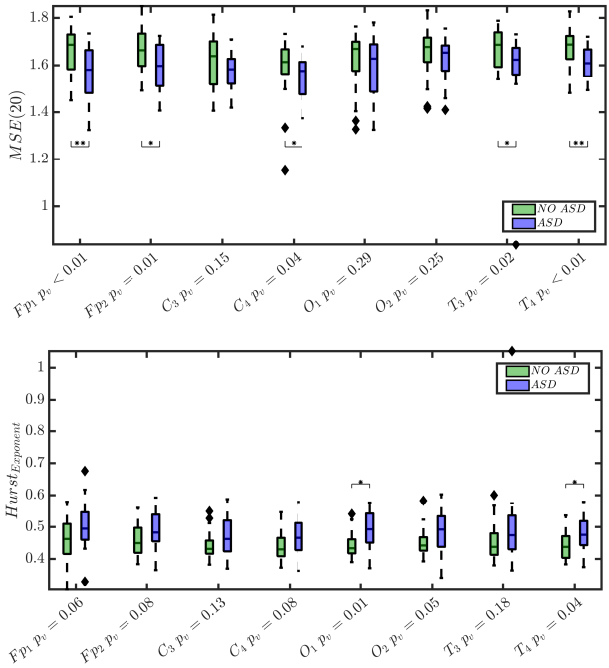


Figure 10.3: The figure shows the entropy at scale 20 $MSE(20)$ and the Hurst exponent in the two groups (NO ASD = No ASD, ASD = Autism Spectrum Disorder). Both describe the persistence of slow-waves and discontinuity. In the case of ASD, the EEG presents a dysmature or disorganized pattern, since the regularity of the signal is higher due to spikiness, lack of smoothness and general discontinuity. The enhanced and abnormal slow-wave information in the disorganized EEG lowers the entropy at lower frequencies ($MSE(20)$). P-values have been derived with Kruskal-Wallis test. The symbols * and ** respectively represent post-hoc comparison with $p \leq 0.05$ and $p \leq 0.01$.

of AUC_2 consistently score higher than those of AUC_1. For example, AUC_2 for MSE, MFA and all features together are 0.75, 0.71, 0.74 respectively, while AUC_1 are 0.65, 0.71, 0.73 respectively. This means that the ASD diagnosis score seems easier to discriminate compared to a simple abnormality in BSID-III.

Figure 10.5 depicts the boxplots for MSE(20) and Hurst-Exp in the three different groups defined for the multiclass problem (Normal vs ABN Bailey vs ASD). The Kruskal-Wallis test shows a significant difference for lower entropy at higher scales in the following channels ($F_{p,1}$ - $p < 0.01$,$F_{p,2}$ - $p = 0.01$, $T3$ - $p = 0.02$,T_4 - $p = 0.02$) and higher regularity for the following channels (O_1 -

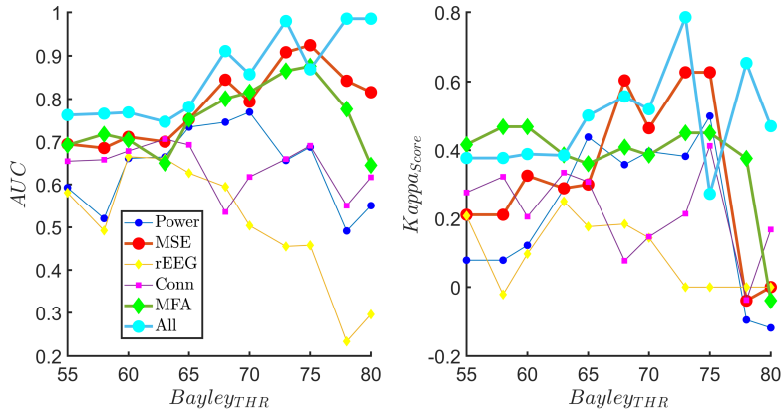


Figure 10.4: Binary classification performance in function of the threshold applied to BSID-III ($Bayley_{THR}$) to define the positive class. The two panels report the linear discriminant analysis area under the ROC curve (AUC) and Kappa score for the classification of the normal vs abnormal development group for each group of features. The abnormal group contains at least one developmental problem (ASD or abnormal Bayley). The Bayley Threshold has been used to define the abnormal Bayley outcome for each of the three investigated scores (cognitive, motor and language outcome). The reported groups of features are power, entropy (MSE), range-EEG or NEURAL features (rEEG), connectivity (Conn), fractality (MFA) and all the features combined (All). Each AUC and Kappa score chart is reported in different colors and different symbols for each feature group.

$p = 0.03$, O_2 - $p = 0.05$). The multicomparison test shows a tendency associated to the severity of abnormal development at 2 years for the $MSE(20)$ in $F_{p,1}$, as indicated by the asterisks associated to $p \leq 0.05$. The remaining channels with a significant p-value show a significant difference between normal patients and ASD. The entropy at scale 20 is lower in case of an ASD patient with abnormal Bayley, while the Hurst regularity increases in case of ASD.

Table 10.7 shows the selected features for the binary classification when the entire pool of features is considered. They mostly belong to the fractal and entropy family with the addition of graph and network resilience features in the δ_2 band. Figure 10.7 shows the boxplots of the number of superfluous connections (or network resilience), which suggest a higher network resilience is found in case of abnormal developmental outcome, as confirmed by the Kruskal-Wallis test for δ_2 ($p = 0.02$), θ ($p = 0.04$) and alpha ($p = 0.03$). The boxplots

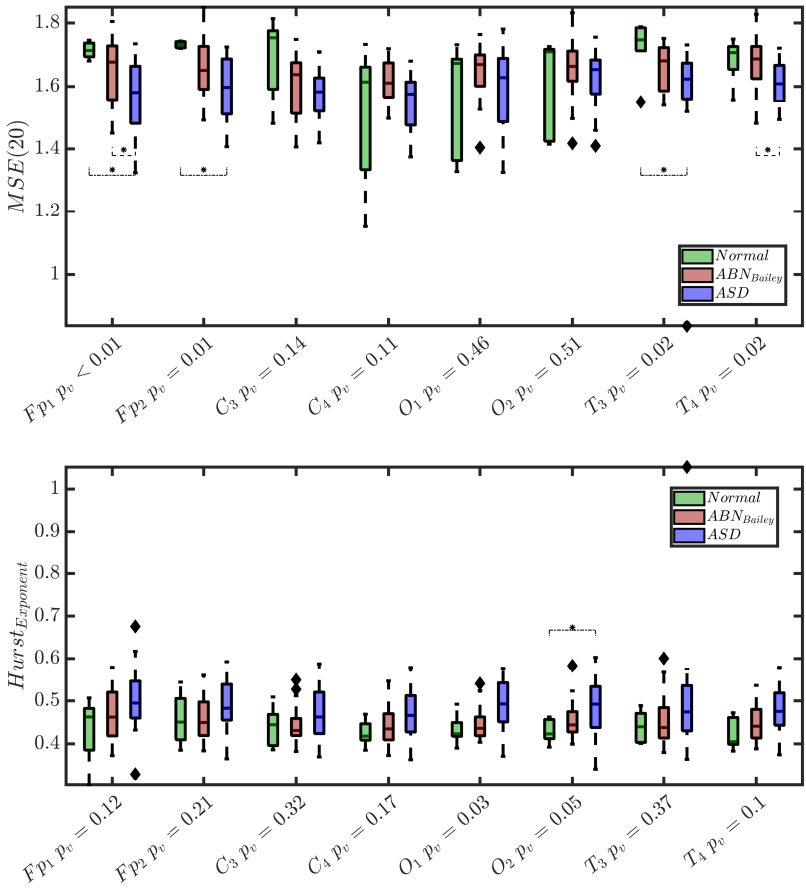


Figure 10.5: The figure shows the entropy at scale 20 $MSE(20)$ and the Hurst exponent in the three groups (Normal = Normal Development, ABN-Bayley = only abnormal Bayley outcome , ASD = ASD patients). Both describe the presence of slow-waves persistence and discontinuity. The combination of ASD and abnormal development shows a more severe increase of regularity compared to other groups (especially, in the occipital area) and a more severe decrease of entropy compared to other groups (especially, in the frontal area).P-values have been derived with Kruskal-Wallis test. The symbols * and ** respectively represent post-hoc comparison with $p \leq 0.05$ and $p \leq 0.01$.

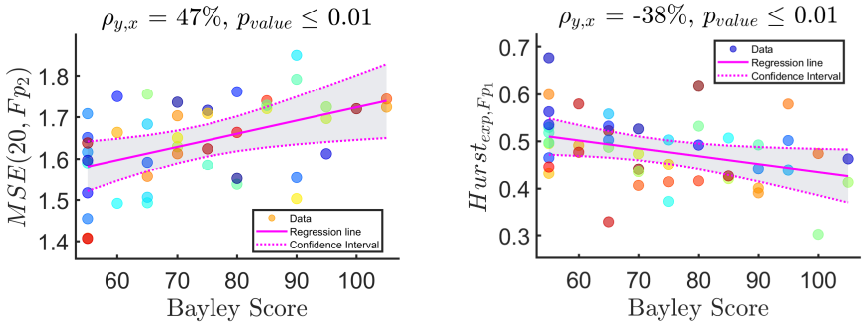


Figure 10.6: The figure shows the entropy at scale 20 $MSE(20)$ and the Hurst exponent in function of the Cognitive BSID-III score. The two panels reported the cluster of data points (circles with a color for each patient), the expected value of the considered feature for each Bayley score (pink thick line) and the 95% confidence intervals (pink dashed lines). The left panel clearly shows that the higher the Bayley score the higher the entropy at scale 20, while the right panel shows that the higher the cognitive score the lower the Hurst exponent. The figure confirms that both entropy and the regularity are features of brain dysmaturity, and are related to developmental outcome in ASD subjects. In addition, the increase of $MSE(20)$ underlines how the persistence of slow-waves is related to a worse developmental outcome.

were reported by assuming an abnormal development with Bayley score below 80.

10.3.4 Regression analysis for the Bayley score

The last step of this analysis was the regression analysis to automatically compute the different Bayley scores with quantitative EEG features. The relationship between the different Bayley outcome and the EEG dysmature features is reported in Table 10.8. The explained variance or coefficient R^2 for the Motor, Language and Cognitive score is respectively 0.30, 0.38, 0.46, while the mean absolute error is 7.85, 7.5 and 7.85. For each of the models, the five most correlated features with different Bayley scores which were used in the final model are also reported in Table 10.8, alongside the TSC mutation variable. Similarly to the classification strategies, the entropy at scale 20 ($MSE(20, Fp1)$) and the Hurst exponent contribute to define the Bayley score at 2 years. Figure 10.6 shows the relationship between $MSE(20)/Hurst_{Exp}$ and the cognitive Bayley score.

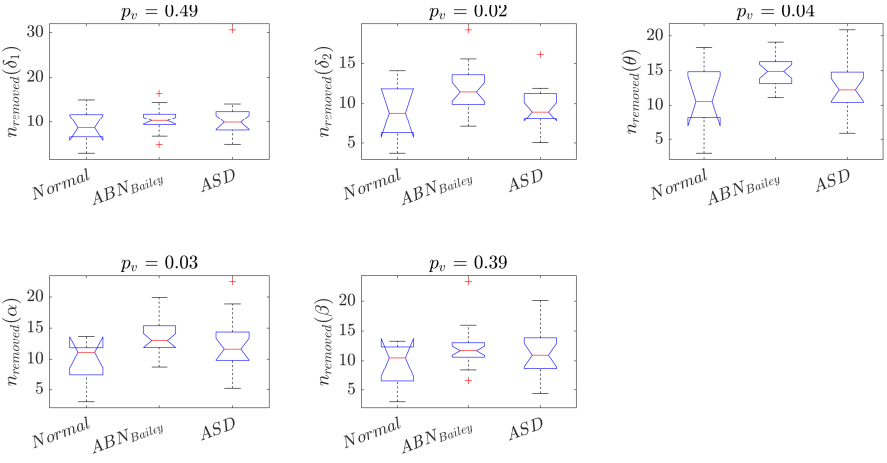


Figure 10.7: The measure of network redundancy is shown for the three different study groups: normal, abnormal Bayley outcome (Bayley < 80) and ASD with an abnormal Bayley. As shown by the second and third panel, the number of superfluous connections (n_{sup}) is higher in case of development abnormalities.

10.4 Discussion

The quantitative analysis of EEG background confirmed that an early-life abnormal or dysmature EEG background is associated with an abnormal development outcome in TSC infants. The examined biomarkers and classification strategies yield a high prognostic value for future developmental abnormalities in children with TSC. The current findings do not only support the earlier clinical EEG investigation of the early-life recordings by [61], which shows that background abnormalities are a significant predictor of ASD, but confirms the similarities between EEG dismaturity and the autistic EEG traits, which are normally investigated only in the later childhood.

The results of the models of the first and second strategy show that a dysmature EEG in the first months of life is associated with ASD traits at 24 months and with neurodevelopmental abnormalities. Table 10.5 underlines that dysmature EEG features can discriminate ASD patients, as also displayed in Figure 10.3. Those results are comparable to the ASD siblings reported by [40] at 6 months. Although the current results might underperform compared to more common studies related to classification of ASD [34],[74],[202], the literature has normally focused on children older than 2 years, while this investigation has specifically focused to young infants below 4 months. Furthermore, Figure 10.4 and

Table 10.6 do not only stress that these same features can detect abnormal development as shown by a wide range of studies in neonates [184],[262], but they can discriminate ASD patients from patients with abnormal Bayley scores. As further proof of the relationship between abnormal Bayley and EEG dysmaturity, we showed that characteristic quantitative EEG features were significantly able to explain the variance of the cognitive, language and motor BSID-III score (Table 10.8 and Figure 10.6). Interestingly, the reported analysis also inferred the best cut-off for the Bayley to define problems with the development based on data-driven methods. The best BSID-III threshold to discriminate healthy patients from controls seems to be 75 for the binary case and 80 for the multiclass case, in line with other clinical studies [75],[124].

Our findings suggest that the most informative features to discriminate ASD in TSC is MSE and MFA, as shown by [5],[40] and by the performance in Table 10.5 and 10.6, the selected features in Table 10.7 and the red and green charts in Figure 10.4. Those results do not only confirm known results of the risk of autism [34], but prove that ASD-prone patients and subjects with an abnormal Bayley score have a dysmature early-life EEG, as defined by [197] and quantified by [65],[95], [144],[191]. This type of signal is characterized by a lower complexity and higher discontinuity at very young age (< 4 months) (see Figure 10.3) and it is associated with adverse outcome [197]. Figure 10.5 might also suggest that EEG entropy at higher scale decreases even further in case of ASD compared to just a lower developmental outcome (especially in channel $F_{p,1}$). In addition, the selected features for both the classification and the regression models (Table 10.7 and 10.8) emphasize the role of lower frequencies for predicting developmental delay, which further support the persistence of EEG slow-waves in case of dysmaturity or poor developmental outcome. The fact that the lower complexity and higher regularity play a role in the detection of the poor outcome is confirmed by the regression results: a lower entropy and higher regularity in neonatal EEG lead to a worse developmental outcome in later life (Table 10.8). Interestingly, a higher EEG network resilience was also found in case of developmental abnormality (Figure 10.7), which confirms the results reported by Peters [202].

Based on the current results, the quantitative EEG analysis can provide insights on the neurobiology of the altered brain maturation in patients with TSC. The higher network resilience found in case of development abnormalities, as reported by Figure 7 and Peters et al., might be a consequence of local over-connectivity or reduced functional specialization in ASD patients (Peters et al., 2013). The dysmature neonatal EEG background is associated with a poor cognitive outcome [33],[150],[184],[232] and Table 10.8 and Figure 10.6 confirms that a lower Bayley score is associated with an early-life abnormal EEG in young TSC infants. Specifically, a background EEG with a lower entropy,

higher regularity and higher width of the range EEG is predictive of lower cognitive, motor and language outcomes. More importantly, TSC patients with abnormal development can be discriminated based on automatic biomarkers of the early-life EEG. The current results further claim that a dysmature EEG pattern might reflect cortical injury or a maturation delay in the development of neurons [105], [150],[232] or simply, the EEG of TSC patients with ASD is associated to a dysmature EEG, as it is the case for premature infants with a poor developmental outcome. Unlike other studies related to ASD [40],[202], the current research includes infants with age below 4 months and an automatic analysis of early-life EEG might be a pivotal tool to prevent and preserve the future development of the infant, as also shown in a recent review [131].

However, it is also important to point out some limitations. The number of ASD patients is limited and future validation studies should include a bigger sample to confirm the current results. Furthermore, the quantitative analysis pursued a different nested strategy to show that a dysmature EEG background is specifically predictive of ASD diagnosis and developmental abnormalities and the obtained kappa score shows a moderate association between features and outcomes. However, the moderate level might be also caused by the poor communication skills of ASD patients, which entails a lower Bayley language score (as seen by the values well below 70 in Table 10.1). Therefore, the ASD class might be difficult to discriminate from a generic abnormal Bayley score. A more interesting problem could be the prediction of intellectual disability in ASD patients (normally defined by cognitive scores < 70) compared to ASD patients with any cognitive problem. However, a bigger sample size of ASD patients would be required for this type of analysis. Moreover, an abnormal Bayley outcome might include patients with just a lower developmental motor outcome (known also as developmental dissociation) or patients with both intellectual disability and developmental dissociation. Future studies should also include a proper investigation of how the different aspects and traits of the dysmature EEG contribute to the definition of different abnormal Bayley outcomes.

In a nutshell, the current analysis delved into the association of a dysmature EEG background in TSC neonates and infants, the later diagnosis of ASD and its interplay with the developmental outcome at 2 years. Both entropy and fractal features can indeed discriminate ASD at early-age, but it can be simply the consequence of a worsening developmental outcome. And yet, MSE and MFA have been proven able to discriminate ASD from general development abnormalities in a multiclass scenario. Besides, multiple studies have shown that dysmature EEG predicts a worse developmental outcome, but the current studies show which of the early-EEG quantitative features correlate with the Bayley score at 2 years.

10.5 Summary

The investigation of the quantitative EEG analysis in TSC infants showed that discontinuity and dysmaturity of early-life EEG is associated with ASD diagnosis and neurodevelopmental abnormalities at the age of 2 years. Following the definition of Pavlidis and other pioneering studies on the autism spectrum disorder EEG, a dysmature EEG is characterized by low entropy and high regularity as a consequence of the discontinuity and the persistence of slow-waves. Those features are also the ones that are most predictive in detecting ASD or other abnormal neurodevelopmental outcomes, such as a lower cognitive, language or motor DQ's. In addition, a larger network resilience has also been found in case of development abnormalities [202]. Similarly, to other clinical studies, the definition of abnormal development was based on a Bayley Score threshold. Surprisingly, the value that led to the best binary and multiclass classification was in a range between [75-80], which are near the most common thresholds used in the clinical practice. Eventually, the quantitative EEG analysis at early-stage might provide a greater understanding of the brain functional processing for cognitive disorders or development disabilities such as autism spectrum disorder. The quantitative EEG analysis at the cot-side might lead to a further tuning of interventions or preventive care in case of deviation from the normal developmental trajectory.

Part V

Conclusions

Chapter 11

Conclusions and future directions

11.1 Conclusions

In the first chapter, we introduced the three main objectives of this thesis:

1. The description of infant's development inside the neonatal intensive care unit by means of biomedical signal processing
2. The assessment of stress in NICU patients by means of physiological biomarkers
3. The investigation of the stress impact on the development of the infant and the relationship between early-life EEG and the developmental outcome of the infants

In this thesis, different algorithms were presented to describe the maturation and automatically predict the age of infants in relationship to the first objective. They focused on the functional or the effective connectivity of the scalp EEG, the fractality of the cortical signal or the HRV and other temporal and spectral features related to the autonomic activity. The same developmental features were applied to design a stress classifier either in bradycardia epochs or sleep states in order to address the second objective. The last objective was tackled with the analysis of similar models used for the first goal, but the development was investigated under early-life pain conditions. Similarly, the multivariate

and univariate features that describe the discontinuity and slow-wave patterns of neonatal EEG were employed in the prediction of future neurodevelopment delays, such as autism diagnosis and lower cognitive or motor outcome.

Following the structure of this book, the conclusions for each part will be drawn separately and presented in the next paragraphs.

11.1.1 Preterm development: the assessment of the brain and autonomic maturation

The contribution of this work can be divided in two main fields, Brain-age models and autonomic development models. The first one relates to the prediction of age based on EEG data of the patient, while the latter refers to the maturation aspects of the heart-rate variability of the infant. In the case of EEG, the maturation models were based on both univariate and multivariate analysis.

The quantitative assessment of brain maturation

Chapter 4 investigated the relationship between the scalp EEG connectivity and the maturation of the infant. The regression models were based either on functional connectivity or effective connectivity methods applied on EEG datasets of preterm patients with normal developmental outcome. In both studies, the EEG network topology was normally summarized by means of graph theory. Both methods have shown that the connectivity can estimate the age of the recording with performance that are comparable with other published algorithms. Most importantly, the overall scalp connectivity shows a positive correlation between the network segregation and the maturation of the infants. Consequently, the development of the infant decreases the cortical connectivity and the multivariate analysis can extract new biomarkers for the development of the infants.

Chapter 5 examined the relationship between the EEG fractality and maturation as well as the relationship between fractal features and vigilance states. Similarly to other research reported in the literature, the neonatal EEG shows a decreasing fractality and regularity with the development of the infant. The changes in nonlinear properties of the EEG with the infants' development show that the brain-age model based on EEG fractality can also reliably predict the age of the infant. On top of that, the fractal properties change throughout the sleep-wake cyclicity: the quiet sleep state is normally characterized by higher regularity and higher number of singularities due to it discontinuous trace, while the other sleep states are less "multifractal". However, the sleep-state classification

performance are strongly correlated to the infant's development: all sleep states present discontinuous traces at very-low postmenstrual age, which makes quiet sleep detection difficult.

The quantitative assessment of the autonomic nervous system maturation

Chapter 6 describes the development of the infants looking at the neurovegetative information. Based on spectral, temporal and fractal features of the heart-rate variability, it is possible to have a reliable estimate of the age of infants. The regression models presented in this chapter show that prediction performance are comparable to the Brain-Age models described in the Chapters 4 and 5. Furthermore, it shows that HRV has also a decreasing regularity and an increasing power in all major frequency frequency bands with the development of the infants. Eventually, it should be highlighted that the best performance is achieved if the disruptive effect of bradycardias is taken into account and if different methods to investigate the sympathovagal balance are considered.

11.1.2 Perinatal stress quantification

This part focuses on the general assessment of stress conditions of premature infants. The reported algorithms implement two binary classifiers based on the different physiological biomarkers to discriminate the stress load of the neonates. The contributions will be highlighted for each chapter.

Chapter 7 examined the possibility to discriminate stress by means of bradycardia events with concomitant desaturations. The dataset was comprised of 136 patients with recorded pain score and physiological data at different time points. This study revealed that the patients that undergo stress or procedural pain the day before the recording may present deeper desaturation, a more discontinuous EEG and tighter brain-to-brain and brain-to-heart interaction, especially at younger age. As anticipated in other studies, the management of hypoxic events might be a possible key to investigate stress conditions in the NICU and a possible connection with the future neurodevelopmental outcome of those patients.

Chapter 8 focuses on the background physiological activity during sleep states to assess stress conditions. Unlike the previous chapter, the intensity of experienced pain or stress was also investigated. Throughout their stay in the NICU, patients under stress conditions show a more discontinuous EEG and higher EEG connectivity and brain-to-heart communication. The classification results indicate a moderate association between the physiological features and the stress

outcome. It should be noted that the current study is the first unobtrusive study to investigate pain, perinatal stress and their effect.

11.1.3 Later development outcome: the effect of stress and the importance of early-life EEG

Chapter 9 provides further insights of the Resilience study dataset and the results in Chapters 7 and 8. Based on the subset of the dataset of the previous chapter, the brain-age maturation and autonomic-age maturation was investigated under the conditions of high number and low number of skin-breaking procedures in the first 5 days of life. The findings suggest that a high level of procedural pain is associated with a more dysmature EEG and more variable heart-rate throughout the development of the infant. Those results persist even after chronological or gestational age correction or even if a subset of patients at low gestational age are considered.

Chapter 10 examines the power of the quantitative analysis of early-life EEG to predict the developmental disorders, such as autism spectrum disorder. The automatic detection of EEG dysmaturity by means of fractality, entropy and connectivity does not only show a moderate association between the EEG in the first months of life and the autism diagnosis at 2 years, but it can also predict and discriminate autism from other developmental abnormalities. Furthermore, the regression models to estimate the Bayley score at 24 months confirm that EEG dysmaturity features shows a significant correlation with a poor cognitive, language and motor outcome of the patients later on in life.

11.2 Future directions

In this study, different future trajectories can be envisaged to further extend the current analysis. The next steps may concentrate on the algorithmic aspects related to the different physiological biomarkers that were used in this thesis, but they can also focus on new insights related to stress detection and the impact of pain on maturation. In the next paragraph, a list of possible refinements and extensions is proposed.

11.2.1 Graph theory

The graph theory was mainly used to investigate the topology of the EEG channel network or the brain-heart networks. However, a fundamental aspect

in graph theory relates to the modularity of the network or the identification of community structures [36]. Different connectivity studies do not simply investigate if the network presents a random vs a small-world topology or a dominant hub vs smaller hubs, but they try to identify specific groups of nodes, investigate if they are associated to a specific anatomic function (e.g. the relationship between temporal networks and the auditory cortex) and if the structure of the communities remain stable over time [213], [44], [193]. However, this type of analysis may require high-density EEG or fMRI, which are monitoring tools able to investigate the brain at more refined spatial scale. Consequently, the number of nodes in the network significantly expands because of the higher number of electrodes or because of the higher number of brain regions investigated with the magnetic resonance [185],[44].

The natural extension of the modularity analysis would be the investigation of network dynamics, which investigates how the interaction of a single node evolves over time [24],[93],[162]. Furthermore, the statistical significance of the topological indices derived via network dynamics may require a specific set of surrogate methods, known as surrogate network analysis [11]. It normally consists of a generation of random networks and a comparison between the properties of those surrogate networks and the lattice under investigation [250].

Another fundamental aspect to be investigated relates to the definition of network resilience. In this thesis, resilience was defined as a trade-off between the entropy of the graph and the squared difference (or distance) between the adjacency matrix with some of the links that were removed and the original matrix. Although the number of superfluous connections is clearly an innovative metric, its interpretation and application can be improved. The meaning of this new index can be investigated by means of a correlation analysis with other topological measure (such as the path length) or by the assessment of the number of superfluous connections in simulated graphs. Additionally, one can further tune the definition of this resilience index by the use of different distance metrics between the original matrix and the adjacency matrix with removed links or different type of function to define the entropy of the graph.

11.2.2 Network-Physiology age model

The current maturation models exploit either the cortical information or the autonomic information. The natural extension will be the description of the Network Physiology growth charts [22], [23]. They will aim to describe how the communication between brain and heart or other physiological modalities evolve over time during the stay in the NICU. In addition, those biomarkers

can be used to predict the future development outcome of the patients (such as a specific diagnosis or the Bayley score of the patient).

Interestingly, Bartsch et al. applied the framework of time-delay stability to investigate how the interaction among EEG, EMG and heart-rate change in the different sleep states of healthy adults [23]. One can also envisage a pilot study to assess the infant's network physiology during different behavioral states in the NICU.

11.2.3 Stress classification

Chapters 7 and 8 focused on the use of EEG and HRV to classify stress by means of the Leuven Pain Score. Chapter 10 described the evolution of those physiological signals under different stress loads expressed by the skin breaking procedures. Although the use of benchmark biomarkers such as cortisol are not reliable in infants [100], the current findings must be confirmed by a multicenter and multiscore study. The first fundamental future step is to prove that the relationship between EEG dysmaturity or chronic abnormalities and stress is supported by a variety of pain scores. Therefore, the current classifier can be redesigned to embody other scales (such as the skin breaking procedure or the premature infant pain profile score) in either early-integration or late integration fashion [242]. The latter normally consists of an array of classifiers for each outcome variable (in this case the different pain scores), which might lead to either a combination of the different latent variables or the predicted outcomes. The early-integration simply consists of a combination of the different pain scores in one single variable, which is mathematically far simpler than a late integration. However, the design should take extra care. In the different trials of the reported investigation, a combination of different days of the Leuven Pain Score and the use of other scales, such as the CRIB score, was also considered. Such combination might be influenced by sleep-wake cyclicity or other factors that are not necessarily related to stress. Therefore, the combination of pain scores might require a weight scheme, which should either be optimized or designed with prior clinical knowledge. Eventually, the final validation that stress enhances EEG dysmaturity should require further proof which is not simply given by a variety of scores, but it also needs to be investigated by other hospitals (multicenter analysis).

The analysis of stress was performed with age at birth and at moment of the recording as the main confounding factors. However, there are other factors that could play a role in stress classification, such as the effect of analgesic drugs or non-pharmacological treatments Some preliminary results have shown that skin-to-skin contact or Kangaroo care might decrease the classification

performance of stress detection [148]. The possible speculation is that Kangaroo care might reduce the effect of perinatal stress, but further studies are needed to support this theory. Furthermore, the parent's influence and other NICU environment factors, such as light or sound noise, could also be considered in the next step of this research.

Chapters 7 and 8 also give a first insight on the use of Network Physiology for pain and stress classification. There are multiple studies that investigated the classification of pain in a multimodal sense, with the inclusion of modalities such as near-infrared spectroscopy, EMG and skin conductance [249], [214]. Nonetheless, it is also important to investigate how the pain or stress can effectively impact the interaction among the different physiological system and whether the couplings increase or decrease variability under different stress or pain conditions.

11.2.4 Stress Maturation

The results reported in Chapter 9 clearly show that a higher load of skin breaking procedure is associated with a higher EEG discontinuity, independently from the gestational age. The next step of this research is to investigate the relationship between the early-life EEG under stress conditions and the neurodevelopment outcomes of the patients involved in the resilience study. The fundamental research question is to prove that dysmature EEG is the fundamental link between skin-breaking procedure and a lower cognitive outcome. A validation in this direction will further support the necessity to monitor perinatal stress in the NICU by means of physiological signals. Furthermore, the level of stress-related EEG or other signals dysmaturity might also give a first prognosis of the future Bayley score or development outcome of the patients. The horizon of the current research can also be expanded to a more composite definition of stress. The present findings are linked to the pain-related stress, but other factors may play a role such as maternal separation and sensory stimuli. A retrospective study on the physiological signals based on the detachment levels or the investigation of the relationship between sensory processing disorders and the physiological maturation might provide new branches for the development of the premature brain research [100],[222].

In conclusion, this thesis showed multiple ways to quantify the level of dysmaturity of EEG and HRV and a variety of approaches that link this automatic assessment to perinatal stress. The pivotal finding was that dysmature EEG and a strong EEG and HRV connectivity are related to pain-related stress. These results might indicate a possible impact on the future development of the infant. The whole research wanted to prove that data-driven strategies in

the NICU might help to preserve the maturation and wellbeing of the patients. Therefore, future studies to develop systems and monitors for stress assessment at cot-side are welcome. Most importantly, extensive longitudinal studies that link stress events during neonatal intensive care with later childhood might also shed light on the necessary preventive strategies to improve the overall cognitive outcome of the preterm population.

Appendix A

Supplementary Tables of autonomic maturation trends

This appendix provides the Chapter 6 supplementary tables, which reports the trends for all features in three age groups (PMA ≤ 32 wks, $(32-36]$ wks, > 36 wks)

Table A.1: The main temporal, spectral and fractal features are reported for the post-bradycardia periods. The results are reported as median(IQR). IQR stands for *interquartile range*. The fractal indices are reported for $f_s = 8\ Hz$. The symbol ρ stands for the Pearson correlation coefficient. The symbol ** represents a significant correlation with $p \leq 0.01$, and * is used for a significant correlation with $p \leq 0.05$. $^{n.s.}$ is used to indicate a non-significant correlation.

Median(IQR) - PMA weeks	≤ 32	$(32-36]$	> 36	$\rho(\%)$
Temporal features in the post-bradycardia (PB) group				
μ_{RR}	374.65(366.38-391.36)	377.07(364.33-393.69)	387.2(374.98-416.44)	0.39^{**}
σ_{RR}	16.71(12.02-22.05)	25.5(21.65-31.1)	28.47(24.03-32.08)	0.49^{**}
Spectral features in the post-bradycardia (PB) group				
$P(VLF)_{Welch}$	106.24(63.14-156.22)	250.27(180.65-408.05)	287.59(219.04-454.15)	0.38^{**}
$P(VLF)_{SPWD}$	643.16(434.61-1029.4)	1787.69(1081.86-2475.95)	2013.8(1252.61-2653.49)	0.39^{**}
$P(VLF)_{Wavelet}$	36.25(19.22-56.58)	82.2(46.77-119.8)	89.58(62.29-139.78)	0.34^{*}
$P(LF)_{Welch}$	10.91(5.35-16.13)	28.98(13.41-48.24)	50.5(19.82-67.56)	0.63^{**}
$P(LF)_{SPWD}$	70.65(29-94.04)	141.99(84.7-219.13)	450.23(119.78-574.82)	0.69^{**}
$P(LF)_{Wavelet}$	2.14(0.96-3.68)	4.17(2.16-8.9)	17.74(4.94-25.51)	0.69^{**}
$P(HF)_{Welch}$	7.85(3.88-9.53)	9.99(6.08-13.7)	11.84(9.28-24.15)	$0.22^{n.s.}$
$P(HF)_{SPWD}$	38.05(22.98-68.8)	76.75(42.2-106.64)	124.07(87.74-210.87)	0.39^{**}
$P(HF)_{Wavelet}$	0.69(0.36-1.3)	1.61(0.67-2.28)	3.57(1.55-6.4)	0.61^{**}
$\frac{VLF}{LF}_{Welch}$	12.22(7.92-24.59)	9.9(5.8-18.72)	5.45(4.72-8.03)	$0.06^{n.s.}$
$\frac{VLF}{LF}_{SPWD}$	10.76(7.7-13.27)	7.68(4.98-15.57)	4.5(3.11-4.82)	$-0.08^{n.s.}$
$\frac{VLF}{LF}_{Wavelet}$	20.17(12.68-34.29)	19.6(7.22-26.37)	6.98(4.53-10.08)	-0.36^{**}
$\frac{LF}{HF}_{Welch}$	1.47(1.05-1.96)	2.33(1.54-3.36)	3.91(1.9-4.94)	0.56^{**}
$\frac{LF}{HF}_{SPWD}$	1.38(1.14-1.79)	1.91(1.71-2.74)	3.2(1.46-3.81)	0.57^{**}
$\frac{LF}{HF}_{wavelet}$	2.27(1.87-3.02)	3.2(2.76-3.75)	4.56(3.04-5.24)	0.48^{**}
$\frac{LF}{LF+HF}_{Welch}$	59.45(48.87-65.64)	69.97(60.1-76.75)	79.55(65.51-83.17)	0.45^{**}
$\frac{LF}{LF+HF}_{SPWD}$	57.89(53.12-63.96)	65.63(63.1-73.15)	76.13(59.34-79.22)	0.48^{**}
$\frac{LF}{LF+HF}_{Wavelet}$	69.44(65.16-74.67)	76.18(73.35-78.44)	81.96(75.27-83.65)	0.37^{**}
$\frac{LF}{LF+VLF}_{Welch}$	9.11(5.32-12.28)	9.17(5.2-15.27)	15.99(13.78-17.47)	0.48^{**}
$\frac{LF}{LF+VLF}_{SPWD}$	9.14(7.5-11.74)	11.53(6.04-16.74)	19.79(17.18-21.91)	0.56^{**}
$\frac{LF}{LF+VLF}_{Wavelet}$	5.38(3.45-8.7)	4.9(3.67-12.17)	14.06(11.77-18.09)	0.57^{**}
Fractal features in the post-bradycardia (PB) group				
$H_{exp,[j_1,j_2=5,12]}$	0.61(0.52-0.7)	0.55(0.45-0.59)	0.5(0.44-0.56)	-0.47^{**}
$C_{2,[j_1,j_2=5,12]}$	-0.2(-0.26–0.17)	-0.19(-0.21–0.13)	-0.14(-0.15–0.11)	0.45^{**}
$H_{exp,[j_1,j_2=3,12]}$	0.67(0.6-0.71)	0.66(0.59-0.69)	0.62(0.58-0.65)	-0.33^{*}
$C_{2,[j_1,j_2=3,12]}$	-0.14(-0.16–0.1)	-0.11(-0.14–0.08)	-0.09(-0.11–0.09)	$0.2^{n.s.}$

Table A.2: The main temporal, spectral and fractal features are reported for the between-bradycardias (BB) periods. The results are reported as median(IQR). IQR stands for *interquartile range*. The fractal indices are reported for $f_s = 8$ Hz. The symbol ρ stands for the Pearson correlation coefficient. The symbol ** represents a significant correlation with $p \leq 0.01$, and * is used for a significant correlation with $p \leq 0.05$. $^{n.s.}$ is used to indicate a non-significant correlation.

Temporal features in the between-bradycardias (BB) group				
Median(IQR) - PMA weeks	≤ 32	$(32-36]$	> 36	$\rho(\%)$
Fractal features in the three age groups				
μ_{RR}	370.51(359.96-388.36)	377.42(363.11-389.25)	394.93(370.01-427.45)	0.47**
σ_{RR}	13.89(10.97-18.49)	19.81(15.72-23.82)	29.1(21.99-30.66)	0.64**
Spectral features in the between-bradycardias (BB) group				
$P(VLF)_{Welch}$	68.99(46.83-128.31)	156.67(100.26-217.63)	320.47(184.27-388.76)	0.54**
$P(VLF)_{SPWD}$	358.71(321.19-673.85)	991.87(498.71-1346.25)	2332.8(901.57-2915.13)	0.58**
$P(VLF)_{Wavelet}$	19.23(16.42-40.32)	58.93(35.74-75.05)	108(55.74-147.03)	0.63**
$P(LF)_{Welch}$	8.27(3.78-13.34)	14.99(5.95-25.28)	35.69(26.87-49.37)	0.66**
$P(LF)_{SPWD}$	46.5(25.63-88.72)	106.67(56.24-154.71)	273.11(213.86-373.31)	0.73**
$P(LF)_{Wavelet}$	1.3(0.86-3.38)	4.23(1.93-6.28)	11.34(8.4-15)	0.71**
$P(HF)_{Welch}$	4.85(3.47-6.9)	5.2(4.06-8.12)	11.55(6.11-13.56)	0.17$^{n.s.}$
$P(HF)_{SPWD}$	24.07(19.35-49.89)	50.8(25.85-85.57)	103.73(63.59-129.27)	0.24$^{n.s.}$
$P(HF)_{Wavelet}$	0.45(0.36-1.06)	0.91(0.53-2.2)	2.79(2.09-4.65)	0.62**
$\frac{VLF}{LF}_{Welch}$	9.4(7.9-13.65)	9.22(5.58-18.69)	5.29(4.7-8.4)	-0.2$^{n.s.}$
$\frac{VLF}{LF}_{SPWD}$	7.86(6.19-10.78)	7.5(5.3-13.48)	4.02(3.45-6.87)	-0.14$^{n.s.}$
$\frac{VLF}{LF}_{Wavelet}$	13.42(10.75-19.77)	11.73(8.1-21.36)	7(6.17-10.44)	-0.3$^{n.s.}$
$\frac{LF}{HF}_{Welch}$	1.42(0.75-2.16)	2.19(1.8-3.09)	3.78(2.39-4.23)	0.57**
$\frac{LF}{HF}_{SPWD}$	1.45(0.98-1.6)	1.87(1.61-2.18)	2.8(1.61-3.25)	0.52**
$\frac{LF}{HF}_{Wavelet}$	2.32(1.8-2.84)	3.2(2.39-4.11)	3.67(2.47-4.6)	0.33*
$\frac{LF}{LF+HF}_{Welch}$	58.67(42.64-67.61)	68.69(64.25-75.48)	78.97(70.53-80.88)	0.56**
$\frac{LF}{LF+HF}_{SPWD}$	59.11(49.57-61.62)	65.19(61.76-68.51)	73.67(61.65-76.48)	0.47**
$\frac{LF}{LF+HF}_{Wavelet}$	69.86(64.13-73.93)	76.21(70.43-80.43)	78.39(71.21-82.13)	0.33*
$\frac{LF}{LF+VLF}_{Welch}$	9.71(6.91-11.29)	9.87(5.18-15.21)	16.08(10.77-17.58)	0.42**
$\frac{LF}{LF+VLF}_{SPWD}$	11.29(8.24-13.6)	11.67(6.91-15.29)	19.55(12.12-21.93)	0.44**
$\frac{LF}{LF+VLF}_{Wavelet}$	6.93(4.89-8.53)	7.9(4.63-11.05)	12.57(8.75-13.95)	0.48**
Fractal features in the between-bradycardias (BB) group				
$H_{exp,[j_1,j_2=5,12]}$	0.6(0.52-0.68)	0.54(0.5-0.59)	0.48(0.45-0.52)	-0.5**
$C_{2,[j_1,j_2=5,12]}$	-0.19(-0.23–0.14)	-0.17(-0.2–0.14)	-0.09(-0.12–0.08)	0.43**
$H_{exp,[j_1,j_2=3,12]}$	0.68(0.61-0.73)	0.65(0.6-0.67)	0.6(0.55-0.62)	-0.36*
$C_{2,[j_1,j_2=3,12]}$	-0.12(-0.15–0.1)	-0.12(-0.14–0.1)	-0.08(-0.09–0.05)	0.23$^{n.s.}$

Table A.3: The main temporal, spectral and fractal features are reported for the within-bradycardia (WB) periods. The results are reported as median(IQR). IQR stands for *interquartile range*. The temporal and spectral indices are reported for $f_s = 6\ Hz$ and the fractal index is reported for $f_s = 8\ Hz$. The symbol ρ stands for the Pearson correlation coefficient. The symbol ** represents a significant correlation with $p \leq 0.01$, and * is used for a significant correlation with $p \leq 0.05$. $^{n.s.}$ is used to indicate a non-significant correlation.

Temporal features in the within-bradycardia (WB) group.				
Median(IQR) - PMA weeks	≤ 32	$(32-36]$	> 36	$\rho(\%)$
Fractal features in the three age groups				
μ_{RR}	384.9(369.62-398.91)	384.16(369.2-397.51)	389.12(377.65-425.8)	0.37^{**}
σ_{RR}	38.31(32.22-44.62)	40.61(32.5-49.81)	35.89(28.43-40.74)	$-0.04^{n.s.}$
Spectral features in the the within-bradycardia (WB) group.				
$P(VLF)_{Welch}$	167.75(101.36-282.21)	312.85(213.81-435.07)	300.05(226.89-457.86)	0.28^{*}
$P(VLF)_{SPWD}$	1109.26(838.07-1615.84)	2025.93(1381.79-3909.12)	2168.01(1229.65-3824.98)	0.31^{*}
$P(VLF)_{Wavelet}$	82.66(46.46-169.3)	122.69(69.22-230.91)	128.03(77.63-164.54)	$0.1^{n.s.}$
$P(LF)_{Welch}$	13(6.4-18.71)	31.09(15.62-52.49)	49.95(20.66-73.54)	0.59^{**}
$P(LF)_{SPWD}$	102.62(68.87-185.63)	216.65(136.5-351.6)	502.28(152.3-734.74)	0.65^{**}
$P(LF)_{Wavelet}$	2.3(1.12-4.03)	4.68(2.8-9.86)	18.66(5.35-26.46)	0.69^{**}
$P(HF)_{Welch}$	7.73(4.31-10.01)	10.63(7.49-14.56)	12.21(9.75-25.52)	$0.19^{n.s.}$
$P(HF)_{SPWD}$	78.85(56.13-117.78)	106.08(78.08-182.35)	148.54(118.65-225.94)	0.3^{*}
$P(HF)_{Wavelet}$	0.6(0.4-1.34)	1.49(0.74-2.57)	3.75(1.58-6.34)	0.59^{**}
$\frac{VLF}{LF}_{Welch}$	13.25(9.99-25.23)	9.86(5.99-19.16)	5.62(4.53-7.99)	$0.01^{n.s.}$
$\frac{VLF}{LF}_{SPWD}$	8.3(6.56-12.13)	8.85(4.78-12.81)	4.18(2.93-4.69)	$-0.08^{n.s.}$
$\frac{VLF}{LF}_{Wavelet}$	43.62(24.21-96.77)	30.2(10.15-56.67)	9.21(4.96-13.94)	-0.47^{**}
$\frac{LF}{HF}_{Welch}$	1.59(1.2-2.07)	2.28(1.69-3.51)	3.86(1.85-4.99)	0.52^{**}
$\frac{LF}{HF}_{SPWD}$	1.3(1.19-1.51)	1.8(1.48-2.26)	2.89(1.35-3.25)	0.57^{**}
$\frac{LF}{HF}_{Wavelet}$	2.36(1.89-2.99)	3.23(2.95-3.84)	4.67(3.02-5.3)	0.48^{**}
$\frac{LF}{LF+HF}_{Welch}$	61.12(53.47-66.49)	69.5(62.38-77.32)	79.32(64.9-83.31)	0.41^{**}
$\frac{LF}{LF+HF}_{SPWD}$	56.49(54.27-60.15)	64.06(59.6-69.35)	74.27(57.44-76.58)	0.49^{**}
$\frac{LF}{LF+HF}_{Wavelet}$	70.22(65.4-74.94)	76.37(73.75-78.74)	82.34(75.1-84.11)	0.37^{**}
$\frac{LF}{LF+VLF}_{Welch}$	7.28(4.92-9.1)	9.2(4.99-14.83)	15.38(14.34-18.08)	0.55^{**}
$\frac{LF}{LF+VLF}_{SPWD}$	10.86(8.03-13.38)	10.86(7.2-17.11)	19.63(18.13-23.51)	0.52^{**}
$\frac{LF}{LF+VLF}_{Wavelet}$	2.57(1.13-3.97)	3.34(2.19-8.97)	10.96(6.69-16.78)	0.64^{**}
Fractal features in the within-bradycardia (WB) group.				
$H_{exp,[j_1,j_2=5,12]}$	0.61(0.49-0.71)	0.55(0.43-0.62)	0.49(0.43-0.52)	-0.45^{**}
$C_{2,[j_1,j_2=5,12]}$	-0.26(-0.3- -0.21)	-0.21(-0.24- -0.17)	-0.13(-0.18- -0.11)	0.54^{**}
$H_{exp,[j_1,j_2=3,12]}$	0.66(0.62-0.71)	0.64(0.58-0.68)	0.61(0.58-0.62)	-0.36^{**}
$C_{2,[j_1,j_2=3,12]}$	-0.15(-0.2- -0.12)	-0.14(-0.17- -0.11)	-0.11(-0.12- -0.09)	0.31^{*}

Bibliography

[1] Abdulla, W., and Wong, L. Neonatal EEG signal characteristics using time frequency analysis. *Physica A: Statistical Mechanics and its Applications 390*, 6 (mar 2011), 1096–1110.

[2] Abry, P., Wendt, H., Jaffard, S., Helgason, H., Goncalves, P., Pereira, E., Gharib, C., Gaucherand, P., and Doret, M. Methodology for multifractal analysis of heart rate variability: From LF/HF ratio to wavelet leaders. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology* (aug 2010), IEEE, pp. 106–109.

[3] Accardo, A. P., Affinito, M., Carrozzi, M., Cisint, S., and Bouquet, F. Comparison between spectral and fractal EEG analyses of sleeping newborns. *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol.20 Biomedical Engineering Towards the Year 2000 and Beyond (Cat. No.98CH36286) 20*, 3 (1998), 1569–1571.

[4] Adams-Chapman, I., and Stoll, B. J. Neonatal infection and long-term neurodevelopmental outcome in the preterm infant, jun 2006.

[5] Ahmadlou, M., Adeli, H., and Adeli, A. Improved visibility graph fractality with application for the diagnosis of Autism Spectrum Disorder. *Physica A 391*, 20 (oct 2012), 4720–4726.

[6] Allegaert, K., Naulaers, G., Vanhaesebrouck, S., and Anderson, B. J. The paracetamol concentration-effect relation in neonates. *Pediatric Anesthesia 23*, 1 (jan 2013), 45–50.

[7] Allegaert, K., Tibboel, D., Naulaers, G., Tison, D., De Jonge, A., Van Dijk, M., Vanhole, C., and Devlieger, H. Systematic evaluation of pain in neonates: effect on the number of intravenous

analgesics prescribed. *European Journal of Clinical Pharmacology 59*, 2 (jun 2003), 87–90.

[8] André, M., Lamblin, M. D., D'Allest, A. M., Curzi-Dascalova, L., Moussalli-Salefranque, F., Nguyen The Tich, S., Vecchierini-Blineau, M. F., Wallois, F., Walls-Esquivel, E., and Plouin, P. Électroencéphalographie du nouveau-né prématuré et à terme. Aspects maturatifs et glossaire. *Neurophysiologie Clinique 40*, 2 (2010), 59–124.

[9] Ansari, A. H., De Wel, O., Lavanga, M., Caicedo, A., Dereymaeker, A., Jansen, K., Vervisch, J., De Vos, M., Naulaers, G., and Van Huffel, S. Quiet sleep detection in preterm infants using deep convolutional neural networks. *Journal of Neural Engineering 15*, 6 (dec 2018), 066006.

[10] Ansari, A. H., De Wel, O., Lavanga, M., Caicedo, A., Dereymaeker, A., Jansen, K., Vervisch, J., De Vos, M., Naulaers, G., and Van Huffel, S. Quiet sleep detection in preterm infants using deep convolutional neural networks. *Journal of Neural Engineering 15*, 6 (sep 2018).

[11] Ansmann, G., and Lehnertz, K. Surrogate-assisted analysis of weighted functional brain networks. *Journal of Neuroscience Methods 208*, 2 (jul 2012), 165–172.

[12] Atkinson, E., and Fenton, A. C. Management of apnoea and bradycardia in neonates. *Paediatrics and Child Health 19*, 12 (dec 2009), 550–554.

[13] Aylward, G. P. Neurodevelopmental outcomes of infants born prematurely. *Journal of Developmental and Behavioral Pediatrics 35*, 6 (2014), 394–407.

[14] Baccalá, L. A., and Sameshima, K. Causality and Influentiability: The Need for Distinct Neural Connectivity Concepts. Springer, Cham, 2014, pp. 424–435.

[15] Bailón, R., Laouini, G., Grao, C., Orini, M., Laguna, P., and Meste, O. The integral pulse frequency modulation model with time-varying threshold: Application to heart rate variability analysis during exercise stress testing. *IEEE Transactions on Biomedical Engineering 58*, 3 PART 1 (mar 2011), 642–652.

[16] Barbeau, D. Y., and Weiss, M. D. Sleep Disturbances in Newborns. *Children 4*, 10 (oct 2017), 90.

[17] Barbeau, D. Y., and Weiss, M. D. Sleep Disturbances in Newborns. *Children 4*, 10 (oct 2017), 90.

[18] Barnett, L., and Seth, A. The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference. *Journal of Neuroscience Methods 223* (2014), 50–68.

[19] Barnett, L., and Seth, A. K. Behaviour of Granger causality under filtering: Theoretical invariance and practical application. *Journal of Neuroscience Methods 201*, 2 (2011), 404–419.

[20] Barnett, L., and Seth, A. K. The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference. *Journal of Neuroscience Methods 223* (2014), 50–68.

[21] Barrett, A. B., Barnett, L., and Seth, A. K. Multivariate Granger causality and generalized variance. *Physical Review E 81*, 4 (apr 2010), 041907.

[22] Bartsch, R. P., and Ivanov, P. C. Coexisting forms of coupling and Phase-Transitions in physiological networks. In *Communications in Computer and Information Science* (2014), vol. 438, Springer Verlag, pp. 270–287.

[23] Bartsch, R. P., Liu, K. K. L., Bashan, A., and Ivanov, P. C. Network Physiology: How Organ Systems Dynamically Interact. *PLOS ONE 10*, 11 (nov 2015), e0142143.

[24] Barzel, B., and Barabási, A. L. Universality in network dynamics. *Nature Physics 9*, 10 (sep 2013), 673–681.

[25] Bashan, A., Bartsch, R. P., Kantelhardt, J. W., Havlin, S., and Ivanov, P. C. Network physiology reveals relations between network topology and physiological function. *Nature Communications 3*, 1 (feb 2012), 1–9.

[26] Batalle, D., Hughes, E. J., Zhang, H., Tournier, J., Tusor, N., Aljabar, P., Wali, L., Alexander, D. C., Hajnal, J. V., Nosarti, C., Edwards, A. D., and Counsell, S. J. NeuroImage Early development of structural networks and the impact of prematurity on brain connectivity. 379–392.

[27] Bauer, A., Kantelhardt, J. W., Bunde, A., Barthel, P., Schneider, R., Malik, M., and Schmidt, G. Phase-rectified signal averaging detects quasi-periodicities in non-stationary data. 423–434.

[28] Beck, S., Wojdyla, D., Say, L., Betran, A. P., Merialdi, M., Requejo, J. H., Rubens, C., Menon, R., and Van Look, P. F. The worldwide incidence of preterm birth: A systematic review of maternal mortality and morbidity. *Bulletin of the World Health Organization 88*, 1 (jan 2010), 31–38.

[29] Beebe, D. W., and Gozal, D. Obstructive sleep apnea and the prefrontal cortex: towards a comprehensive model linking nocturnal upper airway obstruction to daytime cognitive and behavioral deficits. *Journal of Sleep Research 11*, 1 (mar 2002), 1–16.

[30] Berman, A., and Zhang, X. D. On the spectral radius of graphs with cut vertices. *Journal of Combinatorial Theory. Series B 83*, 2 (nov 2001), 233–240.

[31] Bezruschko, B. P., and Smirnov, D. A. *Extracting Knowledge from Time series*. Berlin - Heildeberg, 2010.

[32] Bhutta, A. T., Cleves, M. A., Casey, P. H., Cradock, M. M., Anand, K. J. S., Children, S.-a., Cleves, M. A., Casey, P. H., Cradock, M. M., and Anand, K. J. S. Cognitive and Behavioral Outcomes of School-Aged Children Who Were Born Preterm. *Journal of the American Medical Association 288*, 6 (aug 2007), 728–737.

[33] Biagioni, E., Bartalena, L., Biver, P., Pieri, R., and Cioni, G. Electroencephalographic Dysmaturity in Preterm Infants: A Prognostic Tool in the Early Postnatal Period. *Neuropediatrics 27*, 06 (dec 1996), 311–316.

[34] Billeci, L., Sicca, F., Maharatna, K., Apicella, F., Narzisi, A., Campatelli, G., Calderoni, S., Pioggia, G., and Muratori, F. On the application of Quantitative EEG for characterizing autistic brain: A systematic review, jul 2013.

[35] Blinowska, K. J. Review of the methods of determination of directed connectivity from multichannel data. *Medical & Biological Engineering & Computing 49*, 5 (may 2011), 521–529.

[36] Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment 2008*, 10 (oct 2008), P10008.

[37] Bolea, J., Pueyo, E., Orini, M., and Bailón, R. Influence of Heart Rate in Non-linear HRV Indices as a Sampling Rate Effect Evaluated on Supine and Standing. *Frontiers in Physiology 7*, NOV (nov 2016), 501.

[38] Bondy, J., and Murty, U. *Graph theory with applications*. Springer Berlin Heidelberg, 1976.

[39] Bortel, R., and Sovka, P. Approximation of statistical distribution of magnitude squared coherence estimated with segment overlapping. *Signal Processing 87* (2007), 1100–1117.

[40] Bosl, W., Tierney, A., Tager-Flusberg, H., and Nelson, C. EEG complexity as a biomarker for autism spectrum disorder risk. *BMC Medicine 9*, 1 (dec 2011), 18.

[41] Brown, J. E., Chatterjee, N., Younger, J., and Mackey, S. Towards a Physiology-Based Measure of Pain: Patterns of Human Brain Activity Distinguish Painful from Non-Painful Thermal Stimulation. *PLoS ONE 6*, 9 (sep 2011), e24124.

[42] Brummelte, S., Grunau, R. E., Chau, V., Poskitt, K. J., Brant, R., Vinall, J., Gover, A., Synnes, A. R., and Miller, S. P. Procedural pain and brain development in premature newborns. *Annals of Neurology 71*, 3 (mar 2012), 385–396.

[43] Bullmore, E., and Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience 10*, 3 (mar 2009), 186–198.

[44] Cabral, J., Hugues, E., Sporns, O., and Deco, G. Role of local network oscillations in resting-state functional connectivity. *NeuroImage 57*, 1 (2011), 130–139.

[45] Caicedo, A., Varon, C., Papademetriou, M., Tachtsidis, I., and Huffel, S. V. Decomposition of Biomedical Signals Based on Oblique Subspace Projections. *Frontiers of Physiology 13*, 9 (2014), 1–11.

[46] Camm, A., Malik, M., Bigger, J., and Günter, B. Task force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation and clinical use. *Circulation 93*, 5 (1996), 1043–1065.

[47] Chen, Y. W., and Lin, C. J. Combining SVMs with various feature selection strategies. *Studies in Fuzziness and Soft Computing 207* (2006), 315–324.

[48] Chen, Z., Venkat, P., Seyfried, D., Chopp, M., Yan, T., and Chen, J. Brain–Heart Interaction. *Circulation Research 121*, 4 (aug 2017), 451–468.

[49] Chung, F. Laplacians and the Cheeger Inequality for Directed Graphs. *Annals of Combinatorics 9* (2005), 1–19.

[50] Clairambault, J., Curzi-dascalovab, L., Kauffmann, F., and Mkdiguea, C. Heart Rate Variability in Normal Sleeping and Preterm Neonates Full-Term. *Early Human Development 28* (1992), 169–183.

[51] Clemson, P., Lancaster, G., and Stefanovska, A. Reconstructing Time-Dependent Dynamics. *Proceedings of the IEEE 104*, 2 (feb 2016), 223–241.

[52] Cong, X., Cusson, R. M., Walsh, S., Hussain, N., Ludington-Hoe, S. M., and Zhang, D. Effects of Skin-to-Skin Contact on Autonomic Pain Responses in Preterm Infants. *The Journal of Pain 13*, 7 (jul 2012), 636–645.

[53] Cong, X., Ludington-Hoe, S. M., McCain, G., and Fu, P. Kangaroo Care modifies preterm infant heart rate variability in response to heel stick pain: pilot study. *Early human development 85*, 9 (sep 2009), 561–7.

[54] Cortes, C., and Vapnik, V. Support-vector networks. *Machine Learning 20*, 3 (sep 1995), 273–297.

[55] Costa, M., Goldberger, A. L., and Peng, C.-K. K. Multiscale Entropy Analysis of Complex Physiologic Time Series. *Physical Review Letters 89*, 068102 (2002).

[56] Cruz, M., Fernandes, A., and Oliveira, C. Epidemiology of painful procedures performed in neonates: A systematic review of observational studies. *European Journal of Pain 20*, 4 (apr 2016), 489–498.

[57] Curatolo, P. Mechanistic target of rapamycin (mTOR) in tuberous sclerosis complex-associated epilepsy, mar 2015.

[58] Curzi-Dascalova, L. Development of sleep and autonomic nervous system contorl in premature and full-term newborns. *Archives de pediatrie 2*, I 995 (mar 1994), 255–262.

[59] David, M., Hirsch, M., Karin, J., Toledo, E., and Akselrod, S. An estimate of fetal autonomic state by time-frequency analysis of fetal heart rate variability. *Journal of Applied Physiology 102*, 3 (2007).

[60] De Brabanter, K., De Brabanter, J., Suykens, J., and De Moor, B. Optimized fixed-size kernel models for large data sets. *Computational Statistics & Data Analysis 54*, 6 (2010), 1484–1504.

[61] De Ridder, J., Verhelle, B., Vervisch, J., Lemmens, K., Yepiskoposyan, L., Jozwiak, S., Kotulska, K., Curatolo, P., Benvenuto, A., Moavero, R., Aronica, E., Jansen, A. C., Jansen, F., David, K., Weschke, B., Riney, K., Kresk, P., Feucht, M., Nabbout, R., and Lagae, L. EEG in infants with Tuberous Sclerosis Complex predicts developmental outcome at the age of 2 years. In *13th European Paediatric Neurology Society Congress* (2019), p. 32.

[62] De Souza Filho, L. F. M., De Oliveira, J. C. M., Ribeiro, M. K. A., Moura, M. C., Fernandes, N. D., De Sousa, R. D., Pedrino, G. R., and Rebelo, A. C. S. Evaluation of the autonomic nervous system by analysis of heart rate variability in the preterm infants. *BMC Cardiovascular Disorders 19*, 1 (aug 2019), 198.

[63] De Tommaso, M., Trotta, G., Vecchio, E., Ricci, K., Van de Steen, F., Montemurno, A., Lorenzo, M., Marinazzo, D., Bellotti, R., and Stramaglia, S. Functional connectivity of EEG signals under laser stimulation in migraine. *Frontiers in Human Neuroscience 9*, NOV (nov 2015), 640.

[64] De Wel, O., Lavanga, M., Caicedo, A., Jansen, K., Naulaers, G., and Van Huffel, S. Decomposition of a Multiscale Entropy Tensor for Sleep Stage Identification in Preterm Infants. *Entropy 21*, 10 (sep 2019), 936.

[65] De Wel, O., Lavanga, M., Dorado, A., Jansen, K., Dereymaeker, A., Naulaers, G., and Van Huffel, S. Complexity Analysis of Neonatal EEG Using Multiscale Entropy: Applications in Brain Maturation and Sleep Stage Classification. *Entropy 19*, 10 (sep 2017), 516.

[66] Dereymaeker, A., Pillay, K., Vervisch, J., De Vos, M., Van Huffel, S., Jansen, K., and Naulaers, G. Review of sleep-EEG in preterm and term neonates. *Early Human Development 113* (oct 2017), 87–103.

[67] Dereymaeker, A., Pillay, K., Vervisch, J., Van Huffel, S., Naulaers, G., Jansen, K., and De Vos, M. An Automated Quiet Sleep Detection Approach in Preterm Infants as a Gateway to Assess Brain Maturation. *International Journal of Neural Systems 27*, 06 (sep 2017), 1750023.

[68] Deviaene, M., Testelmans, D., Buyse, B., Borzée, P., Van Huffel, S., and Varon, C. Automatic Screening of Sleep Apnea Patients Based on the SpO Signal. *IEEE Journal of Biomedical and Health Informatics 00*, 00 (2018), 1–1.

[69] Dick, O. E. Multifractal analysis of the psychorelaxation efficiency for the healthy and pathological human brain. *Chaotic Model Simul. 1*, December 2011 (2012), 219–227.

[70] Doesburg, S. M., Chau, C. M., Cheung, T. P., Moiseev, A., Ribary, U., Herdman, A. T., Miller, S. P., Cepeda, I. L., Synnes, A., and Grunau, R. E. Neonatal pain-related stress, functional cortical activity and visual-perceptual abilities in school-age children born at extremely low gestational age. *Pain 154*, 10 (2013), 1946–1952.

[71] Doret, M., Spilka, J., Chudáček, V., Gonçalves, P., Abry, P., and Stanley, H. Fractal Analysis and Hurst Parameter for Intrapartum Fetal Heart Rate Variability Analysis: A Versatile Alternative to Frequency Bands and LF/HF Ratio. *PLOS ONE 10*, 8 (aug 2015), e0136661.

[72] Dubois, J., Benders, M., Cachia, A., Lazeyras, F., Ha-Vinh Leuchter, R., Sizonenko, S. V., Borradori-Tolsa, C., Mangin, J. F., and Huppi, P. S. Mapping the Early Cortical Folding Process in the Preterm Newborn Brain. *Cerebral Cortex 18*, 6 (jun 2008), 1444–1454.

[73] Duerden, E. G., Grunau, R. E., Guo, T., Foong, J., Pearson, A., Au-Young, S., Lavoie, R., Chakravarty, M. M., Chau, V., Synnes, A., and Miller, S. P. Early procedural pain is associated with regionally-specific alterations in thalamic development in preterm neonates. *Journal of Neuroscience 38*, 4 (jan 2018), 878–886.

[74] Duffy, F. H., and Als, H. A stable pattern of EEG spectral coherence distinguishes children with autism from neuro-typical controls - a large case control study. *BMC Medicine 10*, 1 (dec 2012), 64.

[75] Duncan, A. F., Bann, C., Boatman, C., Hintz, S. R., Vaucher, Y. E., Vohr, B. R., Yolton, K., and Heyne, R. J. Do currently recommended Bayley-III cutoffs overestimate motor impairment in infants born <27 weeks gestation? *Journal of Perinatology 35*, 7 (jul 2015), 516–521.

[76] El-Sheikh, M., and Buckhalt, J. A. Vagal regulation and emotional intensity predict children's sleep problems. *Developmental Psychobiology 46*, 4 (2005), 307–317.

[77] Estrada, E., and Hatano, N. Communicability Angle and the Spatial Efficiency of Networks. *SIAM Review 58*, 4 (jan 2016), 692–715.

[78] Fabrizi, L., Slater, R., Worley, A., Meek, J., Boyd, S., Olhede, S., and Fitzgerald, M. A shift in sensory processing that enables the

developing human brain to discriminate touch from pain. *Current Biology 21*, 18 (sep 2011), 1552–1558.

[79] Fabrizi, L., Verriotis, M., Williams, G., Lee, A., Meek, J., Olhede, S., and Fitzgerald, M. Encoding of mechanical nociception differs in the adult and infant brain. *Scientific Reports 6* (jun 2016).

[80] Faes, L., Marinazzo, D., Jurysta, F., and Nollo, G. Linear and non-linear brain–heart and brain–brain interactions during sleep. *Physiol. Meas 36* (2015), 683.

[81] Fagiolo, G. Clustering in complex directed networks. *Physical Review E 76*, 2 (aug 2007), 026107.

[82] Fair, D. A., Dosenbach, N. U. F., Church, J. A., Cohen, A. L., Brahmbhatt, S., Miezin, F. M., Barch, D. M., Raichle, M. E., Petersen, S. E., and Schlaggar, B. L. Development of distinct control networks through segregation and integration. *Proceedings of the National Academy of Sciences of the United States of America 104*, 33 (aug 2007), 13507–12.

[83] Fallani, F. D. V., and Babiloni, F. The Graph Theoretical Approach in Brain Functional Networks: Theory and Applications. *Synthesis Lectures on Biomedical Engineering 5*, 1 (jan 2010), 1–92.

[84] Feldman, R. From Biological Rhythms to Social Rhythms: Physiological Precursors of Mother-Infant Synchrony. *Developmental Psychology 42*, 1 (2006), 175–188.

[85] Fitzgerald, M. The development of nociceptive circuits, jul 2005.

[86] Fitzgerald, M. The development of nociceptive circuits, jul 2005.

[87] Florin, E., Gross, J., Pfeifer, J., Fink, G. R., and Timmermann, L. The effect of filtering on Granger causality based multivariate causality measures. *NeuroImage 50*, 2 (apr 2010), 577–588.

[88] Floyd, S., and Warmuth, M. Sample compression, learnability, and the Vapnik-Chervonenkis dimension. *Machine Learning 21*, 3 (dec 1995), 269–304.

[89] Franke, K., Luders, E., May, A., Wilke, M., and Gaser, C. Brain maturation: predicting individual BrainAGE in children and adolescents using structural MRI. *Neuroimage 63*, 3 (2012), 1305–1312.

[90] Friston, K. J. Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping 2*, 1-2 (jan 1994), 56–78.

[91] Friston, K. J. Functional and Effective Connectivity: A Review. *Brain Connectivity 1*, 1 (jan 2011), 13–36.

[92] Gee, A. H., Barbieri, R., Paydarfar, D., and Indic, P. Predicting Bradycardia in Preterm Infants Using Point Process Analysis of Heart Rate. *IEEE Transactions on Biomedical Engineering 64*, 9 (sep 2017), 2300–2308.

[93] Ghosh, A., Rho, Y., McIntosh, A. R., Kötter, R., and Jirsa, V. K. Cortical network dynamics with time delays reveals functional connectivity in the resting brain. *Cognitive Neurodynamics 2*, 2 (jun 2008), 115–120.

[94] Goldenberg, D., and Galván, A. Developmental Cognitive Neuroscience The use of functional and effective connectivity techniques to understand the developing brain. 155–164.

[95] González, J. J., Mañas, S., De Vera, L., Méndez, L. D., López, S., Garrido, J. M., and Pereda, E. Assessment of electroencephalographic functional connectivity in term and preterm neonates. *Clinical Neurophysiology 122*, 4 (2011), 696–702.

[96] González, J. J., Mañas, S., De Vera, L., Méndez, L. D., López, S., Garrido, J. M., and Pereda, E. Assessment of electroencephalographic functional connectivity in term and preterm neonates. *Clinical Neurophysiology 122*, 4 (2011), 696–702.

[97] Granger, C. W. J. Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica 37*, 3 (aug 1969), 424.

[98] Gretton, A., Bousquet, O., Smola, A., and Schlkopf, B. Measuring statistical dependence with Hilbert-Schmidt norms. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (2005), vol. 3734 LNAI, Springer, Berlin, Heidelberg, pp. 63–77.

[99] Grieve, P. G., Isler, J. R., Izraelit, A., Peterson, B. S., Fifer, W. P., Myers, M. M., and Stark, R. I. EEG functional connectivity in term age extremely low birth weight infants. *Clinical neurophysiology 119*, 12 (dec 2008), 2712–20.

[100] Grunau, R. E. Neonatal pain in very preterm infants: long-term effects on brain, neurodevelopment and pain reactivity. *Rambam Maimonides medical journal 4*, 4 (2013), e0025.

[101] Grunau, R. E., Whitfield, M. F., Petrie-Thomas, J., Synnes, A. R., Cepeda, I. L., Keidar, A., Rogers, M., Mackay, M., Hubber-Richard, P., and Johannesen, D. Neonatal pain, parenting stress and interaction, in relation to cognitive and motor development at 8 and 18 months in preterm infants. *Pain 143*, 1-2 (may 2009), 138–146.

[102] Gruss, S., Treister, R., Werner, P., Traue, H. C., Crawcour, S., Andrade, A., and Walter, S. Pain Intensity Recognition Rates via Biopotential Feature Patterns with Support Vector Machines. *PLOS ONE 10*, 10 (oct 2015), e0140330.

[103] Hartley, C., Berthouze, L., Mathieson, S. R., Boylan, G. B., Rennie, J. M., Marlow, N., and Farmer, S. F. Long-range temporal correlations in the EEG bursts of human preterm babies. *PLoS ONE 7*, 2 (2012).

[104] Haufe, S., Nikulin, V. V., Müller, K.-R., and Nolte, G. A critical assessment of connectivity measures for EEG data: A simulation study. *NeuroImage 64* (jan 2013), 120–133.

[105] Hayakawa, F., Okumura, A., Kato, T., Kuno, K., and Watanabe, K. Dysmature EEG pattern in EEGs of preterm infants with cognitive impairment: Maturation arrest caused by prolonged mild CNS depression. *Brain and Development 19*, 2 (mar 1997), 122–125.

[106] Hellström-Westas, L. Continuous Electroencephalography Monitoring of the Preterm Infant, sep 2006.

[107] Hendrikx, D., Smits, A., Lavanga, M., De Wel, O., Thewissen, L., Jansen, K., Caicedo, A., Van Huffel, S., and Naulaers, G. Measurement of Neurovascular Coupling in Neonates. *Frontiers in Physiology 10* (feb 2019), 65.

[108] Ho, T. K. The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence 20*, 8 (1998), 832–844.

[109] Horne, R. S., Fung, A. C., NcNeil, S., Fyfe, K. L., Odoi, A., and Wong, F. Y. The Longitudinal Effects of Persistent Apnea on Cerebral Oxygenation in Infants Born Preterm. *The Journal of Pediatrics 182* (mar 2017), 79–84.

[110] Hoyer, D., Nowack, S., Bauer, S., Tetschke, F., Rudolph, A., Wallwitz, U., Jaenicke, F., Heinicke, E., Götz, T., Huonker, R., Witte, O. W., Schleussner, E., and Schneider, U. Fetal development of complex autonomic control evaluated from multiscale heart

rate patterns. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology 304*, 5 (mar 2013), R383–R392.

[111] Hoyer, D., Tetschke, F., Jaekel, S., Nowack, S., Witte, O. W., Schleussner, E., and Schneider, U. Fetal Functional Brain Age Assessed from Universal Developmental Indices Obtained from Neuro-Vegetative Activity Patterns. *PLoS ONE 8*, 9 (2013), 1–8.

[112] Humeau-Heurtier, A. The Multiscale Entropy Algorithm and Its Variants: A Review. *Entropy 17*, 5 (may 2015), 3110–3123.

[113] Hüppi, P. S., and Dubois, J. Diffusion tensor imaging of brain development. *Seminars in Fetal and Neonatal Medicine 11*, 6 (2006), 489–497.

[114] Iatsenko, D., McClintock, P. V., and Stefanovska, A. Extraction of instantaneous frequencies from ridges in time-frequency representations of signals. *Signal Processing 125* (aug 2016), 290–303.

[115] Imperatori, C., Farina, B., Quintiliani, M. I., Onofri, A., Castelli Gattinara, P., Lepore, M., Gnoni, V., Mazzucchi, E., Contardi, A., and Della Marca, G. Aberrant EEG functional connectivity and EEG power spectra in resting state post-traumatic stress disorder: A sLORETA study. *Biological Psychology 102*, 1 (2014), 10–17.

[116] Isler, J. R., Stark, R. I., Grieve, P. G., Welch, M. G., and Myers, M. M. Integrated information in the EEG of preterm infants increases with family nurture intervention, age, and conscious state. *PLOS ONE 13*, 10 (oct 2018), e0206237.

[117] Isler, J. R., Stark, R. I., Grieve, P. G., Welch, M. G., and Myers, M. M. Integrated information in the EEG of preterm infants increases with family nurture intervention, age, and conscious state. *PLOS ONE 13*, 10 (oct 2018), e0206237.

[118] Ivanov, P. C., Amaral, L. a., Goldberger, a. L., Havlin, S., Rosenblum, M. G., Struzik, Z. R., and Stanley, H. E. Multifractality in human heartbeat dynamics. *Nature 399*, 6735 (1999), 461–465.

[119] Ivanov, P. C., and Bartsch, R. P. Network Physiology: Mapping Interactions Between Networks of Physiologic Networks. In *Networks of Networks: The Last Frontier of Complexity*. Springer, Cham, 2014, pp. 203–222.

[120] Ivanov, P. C., Liu, K. K., and Bartsch, R. P. Focus on the emerging new fields of network physiology and network medicine. *New Journal of Physics 18*, 10 (2016).

[121] Jaffard, S., Lashermes, B., and Abry, P. Wavelet leaders in multifractal analysis. *Wavelet Analysis and Applications* (2006), 201–246.

[122] Janvier, A., Khairy, M., Kokkotis, A., Cormier, C., Messmer, D., and Barrington, K. J. Apnea Is Associated with Neurodevelopmental Impairment in Very Low Birth Weight Infants. *Journal of Perinatology 24*, 12 (dec 2004), 763–768.

[123] Javorka, K., Lehotska, Z., Kozar, M., Uhrikova, Z., Kolarovszki, B., Javorka, M., and Zibolen, M. Heart Rate Variability in Newborns. *Physiol. Res 66* (2017), 203–214.

[124] Johnson, S., Moore, T., and Marlow, N. Using the Bayley-III to assess neurodevelopmental delay: Which cut-off should be used? *Pediatric Research 75*, 5 (feb 2014), 670–674.

[125] Jones, L., Fabrizi, L., Laudiano-Dray, M., Whitehead, K., Meek, J., Verriotis, M., and Fitzgerald, M. Nociceptive Cortical Activity Is Dissociated from Nociceptive Behavior in Newborn Human Infants under Stress. *Current Biology 27*, 24 (dec 2017), 3846–3851.e3.

[126] Jones, L., Fabrizi, L., Laudiano-Dray, M., Whitehead, K., Meek, J., Verriotis, M., and Fitzgerald, M. Nociceptive Cortical Activity Is Dissociated from Nociceptive Behavior in Newborn Human Infants under Stress. *Current Biology 27*, 24 (dec 2017), 3846–3851.e3.

[127] Jurysta, F., Lanquart, J.-P., Van De Borne, P., Migeotte, P.-F., Dumont, M., Degaute, J.-P., and Linkowski, P. The link between cardiac autonomic activity and sleep delta power is altered in men with sleep apnea-hypopnea syndrome. *Am J Physiol Regul Integr Comp Physiol 291* (2006), 1165–1171.

[128] Kaffashi, F., Scher, M. S., Ludington-Hoe, S. M., and Loparo, K. A. An analysis of the kangaroo care intervention using neonatal EEG complexity: A preliminary study. *Clinical Neurophysiology 124*, 2 (2013), 238–246.

[129] Karin, J., Hirsch, M., and Akselrod, S. An estimate of fetal autonomic state by spectral analysis of fetal heart rate fluctuations. *Pediatric Research 34*, 2 (1993), 134–138.

[130] Knyazev, G. G. EEG delta oscillations as a correlate of basic homeostatic and motivational processes. *Neuroscience and Biobehavioral Reviews 36* (2011), 677–695.

[131] Kong, A. H., Lai, M. M., Finnigan, S., Ware, R. S., Boyd, R. N., and Colditz, P. B. Background EEG features and prediction of cognitive outcomes in very preterm infants: A systematic review. *Early Human Development 127* (dec 2018), 74–84.

[132] Koolen, N., Dereymaeker, A., Räsänen, O., Jansen, K., Vervisch, J., Matic, V., Naulaers, G., De Vos, M., Van Huffel, S., and Vanhatalo, S. Early development of synchrony in cortical activations in the human. *Neuroscience 322* (2016), 298–307.

[133] Koolen, N., Jansen, K., Vervisch, J., Matic, V., De Vos, M., Naulaers, G., and Van Huffel, S. Line length as a robust method to detect high-activity events: Automated burst detection in premature EEG recordings. *Clinical Neurophysiology 125*, 10 (2014), 1985–1994.

[134] Kostović, I., and Judaš, M. The development of the subplate and thalamocortical connections in the human foetal brain. *Acta Paediatrica, International Journal of Paediatrics 99*, 8 (mar 2010), 1119–1127.

[135] Kralemann, B., Pikovsky, A., and Rosenblum, M. Reconstructing effective phase connectivity of oscillator networks from observations. *New Journal of Physics 16*, 8 (aug 2014), 085013.

[136] Krueger, C., van Oostrom, J. H., and Shuster, J. A Longitudinal Description of Heart Rate Variability in 28–34-Week-Old Preterm Infants. *Biological Research For Nursing 11*, 3 (2010), 261–268.

[137] Kugiumtzis, D., and Tsimpiris, A. Measures of Analysis of Time Series (MATS): A MATLAB Toolkit for Computation of Multiple Measures on Time Series Data Bases. *Journal of Statistical Software 33*, 5 (feb 2010), 1–30.

[138] Lake, D. E., and Moorman, J. R. Accurate estimation of entropy in very short physiological time series: the problem of atrial fibrillation detection in implanted ventricular devices. *American journal of physiology. Heart and circulatory physiology 300*, 1 (2011), H319–H325.

[139] Lake, D. E., Richman, J. S., Griffin, M. P., and Moorman, J. R. Sample entropy analysis of neonatal heart rate variability. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology 283*, 3 (sep 2002), R789–R797.

[140] Lancaster, G., Iatsenko, D., Pidde, A., Ticcinelli, V., and Stefanovska, A. Surrogate data for hypothesis testing of physical systems, jul 2018.

[141] Lang, E. W., Tomé, A. M., Keck, I. R., Górriz-Sáez, J. M., Puntonet, C. G., Ez, J. M., and Puntonet, C. G. Brain Connectivity Analysis: A Short Survey. *Computational Intelligence and Neuroscience 2012* (2012), 1–21.

[142] Lavanga, M., Bollen, B., Jansen, K., Ortibus, E., Naulaers, G., Van Huffel, S., and Caicedo, A. A bradycardia-based stress calculator for the neonatal intensive care unit: a multisystem approach. *Frontiers in Physiology 11* (2020), 741.

[143] Lavanga, M., De Wel, O., Caicedo, A., Heremans, E., Jansen, K., Dereymaeker, A., Naulaers, G., and Van Huffel, S. Automatic quiet sleep detection based on multifractality in preterm neonates: Effects of maturation. In *Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (jul 2017), IEEE, pp. 2010–2013.

[144] Lavanga, M., De Wel, O., Caicedo, A., Jansen, K., Dereymaeker, A., Naulaers, G., and Van Huffel, S. Monitoring Effective Connectivity in the Preterm Brain: A Graph Approach to Study Maturation. *Complexity 2017* (2017).

[145] Lavanga, M., De Wel, O., Caicedo, A., Jansen, K., Dereymaeker, A., Naulaers, G., and Van Huffel, S. A brain-age model for preterm infants based on functional connectivity. *Physiological Measurement 39*, 4 (apr 2018), 044006.

[146] Lavanga, M., De Wel, O., Caicedo Dorado, A., Jansen, K., Dereymaeker, A., Naulaers, G., and Van Huffel, S. Linear and nonlinear functional connectivity methods to predict brain maturation in preterm babies. In *8th international workshop on biosignal interpretation* (2016), pp. 37–40.

[147] Lavanga, M., Heremans, E., Moeyersons, J., Jansen, K., Ortibus, E., Naulaers, G., Van, S., and Lavanga, M. Maturation of the autonomic nervous system in premature infants : a bradycardia perspective. *Frontiers in Physiology - Submission* (2020), 1–28.

[148] Lavanga, M., Wel, O. D., Caicedo, A., Deviaene, M., Moeyersons, J., Varon, C., Bollen, B., Jansen, K., Ortibus, E., Naulaers, G., and Van Huffel, S. The implementation of an

apnea-based perinatal stress calculator. In *Proceedings of the 41th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (Berlin, Germany, jul 2019), IEEE, pp. 6000–6003.

[149] Lavanga, M., Wel, O. D., Caicedo, A., Jansen, K., Dereymaeker, A., Naulaers, G., and Huffel, S. V. A brain-age model for preterm infants based on functional connectivity. *Physiological Measurement 39*, 4 (apr 2018), 044006.

[150] Le Bihannic, A., Beauvais, K., Busnel, A., de Barace, C., and Furby, A. Prognostic value of EEG in very premature newborns. *Archives of disease in childhood. Fetal and neonatal edition 97*, 2 (mar 2012), F106–9.

[151] Leonarduzzi, R. F., Schlotthauer, G., and Torres, M. E. Wavelet leader based multifractal analysis of heart rate variability during myocardial ischaemia. *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology* (2010), 110–113.

[152] Lerner, V. V., and A. Pattern recognition using generalized portrait method. *Automation and Remote Control 24* (1963), 774–780.

[153] Levy, J., Hassan, F., Plegue, M. A., Sokoloff, M. D., Kushwaha, J. S., Chervin, R. D., Barks, J. D., and Shellhaas, R. A. Impact of hands-on care on infant sleep in the neonatal intensive care unit. *Pediatric Pulmonology 52*, 1 (jan 2017), 84–90.

[154] Li, D., Li, X., Liang, Z., Voss, L. J., and Sleigh, J. W. Multiscale permutation entropy analysis of EEG recordings during sevoflurane anesthesia. *Journal of Neural Engineering 7*, 4 (aug 2010).

[155] Lin, A., Liu, K. K. L., Bartsch, R. P., and Ivanov, P. C. Delay-correlation landscape reveals characteristic time delays of brain rhythms and heart interactions. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences 374*, 2067 (may 2016), 20150182.

[156] Linsell, L., Johnson, S., Wolke, D., Morris, J., Kurinczuk, J. J., and Marlow, N. Trajectories of behavior, attention, social and emotional problems from childhood to early adulthood following extremely preterm birth: a prospective cohort study. *European Child and Adolescent Psychiatry 28*, 4 (apr 2019), 531–542.

[157] Linsell, L., Johnson, S., Wolke, D., O'Reilly, H., Morris, J. K., Kurinczuk, J. J., and Marlow, N. Cognitive trajectories from infancy to early adulthood following birth before 26 weeks of gestation: A prospective, population-based cohort study. *Archives of Disease in Childhood 103*, 4 (apr 2018), 363–370.

[158] Liu, K. K., Bartsch, R. P., Ma, Q. D., and Ivanov, P. C. Major component analysis of dynamic networks of physiologic organ interactions. *Journal of Physics: Conference Series 640*, 1 (2015).

[159] Loggia, M. L., Juneau, M., and Bushnell, C. M. Autonomic responses to heat pain: Heart rate, skin conductance, and their relation to verbal ratings and stimulus intensity. *Pain 152*, 3 (mar 2011), 592–598.

[160] Longin, E., Gerstner, T., Schaible, T., Lenz, T., and König, S. Maturation of the autonomic nervous system: Differences in heart rate variability in premature vs. term infants. *Journal of Perinatal Medicine 34*, 4 (aug 2006), 303–308.

[161] Maalouf, E. F., Duggan, P. J., Counsell, S. J., Rutherford, M. A., Cowan, F., Azzopardi, D., and Edwards, A. D. Comparison of findings on cranial ultrasound and magnetic resonance imaging in preterm infants. *Pediatrics 107*, 4 I (apr 2001), 719–727.

[162] Majdandzic, A., Podobnik, B., Buldyrev, S. V., Kenett, D. Y., Havlin, S., and Eugene Stanley, H. Spontaneous recovery in dynamical networks. *Nature Physics 10*, 1 (dec 2013), 34–38.

[163] Malmivuo, P., Malmivuo, J., and Plonsey, R. *Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields.* Oxford University Press, New York, USA, 1995.

[164] Marret, S., Ancel, P. Y., Marpeau, L., Marchand, L., Pierrat, V., Larroque, B., Foix-L'Hélias, L., Thiriez, G., Fresson, J., Alberge, C., Rozé, J. C., Matis, J., Bréart, G., and Kaminski, M. Neonatal and 5-year outcomes after birth at 30-34 weeks of gestation. *Obstetrics and Gynecology 110*, 1 (jul 2007), 72–80.

[165] Maynard, D., Prior, P. F., and Scott, D. F. Device for Continuous Monitoring of Cerebral Activity in Resuscitated Patients. *British Medical Journal 4*, 5682 (nov 1969), 545–546.

[166] Mazursky, J. E., Birkett, C. L., Bedell, K. A., Ben-Haim, S. A., and Segar, J. L. Development of baroreflex influences on heart rate variability in preterm infants. *Early Human Development 53*, 1 (nov 1998), 37–52.

[167] McHugh, M. L. Interrater reliability: The kappa statistic. *Biochemia Medica 22*, 3 (2012), 276–282.

[168] Meijer, E., Hermans, K. H. M., Zwanenburg, A., Jennekens, W., Niemarkt, H. J., Cluitmans, P. J. M., Van Pul, C., Wijn, P.

F. F., and Andriessen, P. Functional connectivity in preterm infants derived from EEG coherence analysis. *European journal of paediatric neurology 18*, 6 (2014), 780–9.

[169] Mento, G., and Bisiacchi, P. S. Neurocognitive development in preterm infants: Insights from different approaches. *Neuroscience and Biobehavioral Reviews 36*, 1 (2012), 536–555.

[170] Misra, G., Wang, W.-e., Archer, D. B., Roy, A., and Coombes, S. A. Automated classification of pain perception using high-density electroencephalography data. *Journal of Neurophysiology 117*, 2 (feb 2017), 786–795.

[171] Moavero, R., Benvenuto, A., Emberti Gialloreti, L., Siracusano, M., Kotulska, K., Weschke, B., Riney, K., Jansen, F., Feucht, M., Krsek, P., Nabbout, R., Jansen, A., Wojdan, K., Borkowska, J., Sadowski, K., Hertzberg, C., Hulshof, H., Samueli, S., Benova, B., Aronica, E., Kwiatkowski, D., Lagae, L., Jozwiak, S., Curatolo, P., and on behalf of the EPISTOP Consortium. Early Clinical Predictors of Autism Spectrum Disorder in Infants with Tuberous Sclerosis Complex: Results from the EPISTOP Study. *Journal of Clinical Medicine 8*, 6 (jun 2019), 788.

[172] Moeyersons, J., Amoni, M., Huffel, S. V., Willems, R., and Varon, C. R-DECO: An open-source Matlab based graphical user interface for the detection and correction of R-peaks. *PeerJ Computer Science 2019*, 10 (oct 2019), e226.

[173] Montalto, A., Faes, L., Marinazzo, D., Porta, A., and Vicente, R. MuTE: A MATLAB Toolbox to Compare Established and Novel Estimators of the Multivariate Transfer Entropy. *PLoS ONE 9*, 10 (oct 2014), e109462.

[174] Moorman, J. R., Lake, D. E., and Ivanov, P. C. Early detection of sepsis - A role for network physiology? *Critical Care Medicine 44*, 5 (may 2016), e311–e312.

[175] Mrdalj, J., Pallesen, S., Milde, A. M., Jellestad, F. K., Murison, R., Ursin, R., Bjorvatn, B., and Grønli, J. Early and Later Life Stress Alter Brain Activity and Sleep in Rats. *PLoS ONE 8*, 7 (2013), 1–10.

[176] Mulkey, S. B., Kota, S., Swisher, C. B., Hitchings, L., Metzler, M., Wang, Y., Maxwell, G. L., Baker, R., du Plessis, A. J., and Govindan, R. Autonomic nervous system depression at term in

neurologically normal premature infants. *Early Human Development 123* (2018), 11–16.

[177] Myers, M. M., Grieve, P. G., Stark, R. I., Isler, J. R., Hofer, M. A., Yang, J., Ludwig, R. J., and Welch, M. G. Family Nurture Intervention in preterm infants alters frontal cortical functional connectivity assessed by EEG coherence. *Acta Paediatrica, International Journal of Paediatrics 104*, 7 (jul 2015), 670–677.

[178] Nakagawa, S., Johnson, P. C. D., and Schielzeth, H. The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of The Royal Society Interface 14*, 134 (sep 2017), 20170213.

[179] Natarajan, G., and Shankaran, S. Short- and Long-Term Outcomes of Moderate and Late Preterm Infants, jan 2016.

[180] Newnham, C. A., Inder, T. E., and Milgrom, J. Measuring preterm cumulative stressors within the NICU: The neonatal infant stressor scale. *Early Human Development 85*, 9 (sep 2009), 549–555.

[181] Niemarkt, H. J., Jennekens, W., Pasman, J. W., Katgert, T., Van Pul, C., Gavilanes, A. W., Kramer, B. W., Zimmermann, L. J., Oetomo, S. B., and Andriessen, P. Maturational changes in automated EEG spectral power analysis in preterm infants. *Pediatric Research 70*, 5 (nov 2011), 529–534.

[182] Nolte, G., Bai, O., Wheaton, L., Mari, Z., Vorbach, S., and Hallett, M. Identifying true brain interaction from EEG data using the imaginary part of coherency. *Clinical Neurophysiology 115*, 10 (oct 2004), 2292–2307.

[183] Novak, I., Morgan, C., Adde, L., Blackman, J., Boyd, R. N., Brunstrom-Hernandez, J., Cioni, G., Damiano, D., Darrah, J., Eliasson, A. C., De Vries, L. S., Einspieler, C., Fahey, M., Fehlings, D., Ferriero, D. M., Fetters, L., Fiori, S., Forssberg, H., Gordon, A. M., Greaves, S., Guzzetta, A., Hadders-Algra, M., Harbourne, R., Kakooza-Mwesige, A., Karlsson, P., Krumlinde-Sundholm, L., Latal, B., Loughran-Fowlds, A., Maitre, N., McIntyre, S., Noritz, G., Pennington, L., Romeo, D. M., Shepherd, R., Spittle, A. J., Thornton, M., Valentine, J., Walker, K., White, R., and Badawi, N. Early, accurate diagnosis and early intervention in cerebral palsy: Advances in diagnosis and treatment, sep 2017.

[184] Okumura, A., Hayakawa, F., Kato, T., Watanabe, K., and Kuno, K. Developmental outcome and types of chronic-stage EEG abnormalities in preterm infants. *Developmental Medicine and Child Neurology 44*, 11 (2002), 729–734.

[185] Omidvarnia, A., Fransson, P., Metsäranta, M., and Vanhatalo, S. Functional Bimodality in the Brain Networks of Preterm and Term Human Newborns. *Cerebral Cortex 24*, 10 (oct 2014), 2657–2668.

[186] OpenStax College. Introduction - Anatomy & Physiology - OpenStax CNX, 2013.

[187] Orini, M., Bailon, R., Mainardi, L. T., Laguna, P., and Flandrin, P. Characterization of dynamic interactions between cardiovascular signals by time-frequency coherence. *IEEE Transactions on Biomedical Engineering 59*, 3 (mar 2012), 663–673.

[188] O'Toole, J. M., and Boylan, G. B. NEURAL: quantitative features for newborn EEG using Matlab. *arXiv preprint arXiv:1704.05694* (apr 2017).

[189] O'Toole, J. M., and Boylan, G. B. Quantitative Preterm EEG Analysis: The Need for Caution in Using Modern Data Science Techniques. *Frontiers in Pediatrics 7*, MAY (may 2019), 174.

[190] O'Toole, J. M., Boylan, G. B., Vanhatalo, S., and Stevenson, N. J. Estimating functional brain maturity in very and extremely preterm neonates using automated analysis of the electroencephalogram. *Clinical Neurophysiology* (2016).

[191] O'Toole, J. M., Boylan, G. B., Vanhatalo, S., Stevenson, N. J., O'Toole, J. M., Boylan, G. B., Vanhatalo, S., and Stevenson, N. J. Estimating functional brain maturity in very and extremely preterm neonates using automated analysis of the electroencephalogram. *Clinical Neurophysiology 127*, 8 (aug 2016), 2910–2918.

[192] O'Toole, J. M., Pavlidis, E., Korotchikova, I., Boylan, G. B., and Stevenson, N. J. Temporal evolution of quantitative EEG within 3 days of birth in early preterm infants. *Scientific Reports 9*, 1 (dec 2019).

[193] Palliser, H. K., Bennett, G. A., Kelleher, M. A., Cumberland, A. L., Walker, D. W., and Hirst, J. J. *Prenatal and Postnatal Determinants of Development*, vol. 109. 2016.

[194] Palmu, K., Stevenson, N., Wikström, S., Hellström-Westas, L., Vanhatalo, S., and Palva, J. M. Optimization of an NLEO-based

algorithm for automated detection of spontaneous activity transients in early preterm EEG. *Physiological measurement 31*, 11 (2010), N85–93.

[195] Paolillo, P., and Picone, S. Apnea of prematurity. *Journal of Pediatric and Neonatal Individualized Medicine 2*, 2 (2013).

[196] Pavlidis, E., Lloyd, R. O., and Boylan, G. B. EEG-A Valuable Biomarker of Brain Injury in Preterm Infants. In *Developmental Neuroscience* (jul 2017), vol. 39, S. Karger AG, pp. 23–35.

[197] Pavlidis, E., Lloyd, R. O., Mathieson, S., and Boylan, G. B. A review of important electroencephalogram features for the assessment of brain maturation in premature infants, sep 2017.

[198] Pavlidis, E., Lloyd, R. O., Mathieson, S., and Boylan, G. B. A review of important electroencephalogram features for the assessment of brain maturation in premature infants. *Acta Paediatrica 106*, 9 (sep 2017), 1394–1408.

[199] Peng, H., Li, K., Li, B., Ling, H., Xiong, W., and Hu, W. Predicting Image Memorability by Multi-view Adaptive Regression. In *Proceedings of the 23rd ACM international conference on Multimedia - MM '15* (New York, New York, USA, 2015), ACM Press, pp. 1147–1150.

[200] Pereda, E., Quiroga, R. Q., and Bhattacharya, J. Nonlinear multivariate analysis of neurophysiological signals. *Progress in Neurobiology 77*, 1-2 (2005), 1–37.

[201] Perlman, J. M., and Volpe, J. J. Suctioning in the Preterm Infant: Effects on Cerebral Blood Flow Velocity, Intracranial Pressure, and Arterial Blood Pressure. *PEDIATRICS 72*, 3 (1983), 329–334.

[202] Peters, J. M., Taquet, M., Vega, C., Jeste, S. S., Fernández, I. S., Tan, J., Nelson, C. A., Sahin, M., and Warfield, S. K. Brain functional networks in syndromic and non-syndromic autism: a graph theoretical study of EEG connectivity. *BMC Medicine 11*, 1 (dec 2013), 54.

[203] Pfurtscheller, G., Schwerdtfeger, A. R., Seither-Preisler, A., Brunner, C., Stefan Aigner, C., Brito, J., Carmo, M. P., and Andrade, A. Brain–heart communication: Evidence for "central pacemaker" oscillations with a dominant frequency at 0.1 Hz in the cingulum. *Clinical Neurophysiology 128*, 1 (jan 2017), 183–193.

[204] Pfurtscheller, K., Bauernfeind, G., Müller-Putz, G. R., Urlesberger, B., Müller, W., and Pfurtscheller, G. Correlation

between EEG burst-to-burst intervals and HR acceleration in preterm infants. *Neuroscience Letters 437*, 2 (may 2008), 103–106.

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

[206] Pillay, K., Dereymaeker, A., Jansen, K., Naulaers, G., Van Huffel, S., and De Vos, M. Automated EEG sleep staging in the term-age baby using a generative modelling approach. *Journal of Neural Engineering 15*, 3 (feb 2018), 036004.

[207] Piper, D., Schiecke, K., Pester, B., Benninger, F., Feucht, M., and Witte, H. Time-variant coherence between heart rate variability and EEG activity in epileptic patients: an advanced coupling analysis between physiological networks. *New Journal of Physics 16*, 11 (nov 2014), 115012.

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

[209] Pitson, D., and Stradling, J. Autonomic markers of arousal during sleep in patients undergoing investigation for obstructive sleep apnoea, their relationship to EEG arousals, respiratory events and subjective sleepiness. *Journal of Sleep Research 7*, 1 (mar 1998), 53–59.

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

[211] Popivanov, D., Stomonyakov, V., Minchev, Z., Jivkova, S., Dojnov, P., Jivkov, S., Christova, E., and Kosev, S. Multifractality of decomposed EEG during imaginary and real visual-motor tracking. *Biological Cybernetics 94*, 2 (2006), 149–156.

[212] Porges, S. W. Orienting in a defensive world: Mammalian modifications of our evolutionary heritage. A Polyvagal Theory. *Psychophysiology 32*, 4 (jul 1995), 301–318.

[213] Power, J. D., Fair, D. A., Schlaggar, B. L., and Petersen, S. E. The Development of Human Functional Brain Networks. *Neuron 67* (2010), 735–748.

[214] Ranger, M., Chau, C. M. Y., Garg, A., Woodward, T. S., Beg, M. F., Bjornson, B., Poskitt, K., Fitzpatrick, K., Synnes, A. R., Miller, S. P., and Grunau, R. E. Neonatal Pain-Related Stress Predicts Cortical Thickness at Age 7 Years in Children Born Very Preterm. *PLoS ONE 8*, 10 (oct 2013), e76702.

[215] Ranger, M., Johnston, C. C., and Anand, K. J. Current Controversies Regarding Pain Assessment in Neonates, oct 2007.

[216] Räsänen, O., Metsäranta, M., and Vanhatalo, S. Development of a novel robust measure for interhemispheric synchrony in the neonatal EEG: Activation Synchrony Index (ASI). *NeuroImage 69* (apr 2013), 256–266.

[217] Rawlings, J. O., Pantula, S. G., and Springer, D. A. D. Applied Regression Analysis: A Research Tool, Second Edition. Tech. rep., 2001.

[218] Reynolds, M. L., and Fitzgerald, M. Long-term sensory hyperinnervation following neonatal skin wounds. *The Journal of Comparative Neurology 358*, 4 (aug 1995), 487–498.

[219] Richards, J. E. Development and Stability in Visual Sustained Attention in 14, 20, and 26 Week Old Infants. *Psychophysiology 26*, 4 (jul 1989), 422–430.

[220] Rosenblum, M. G., Cimponeriu, L., Bezerianos, A., Patzak, A., and Mrowka, R. Identification of coupling direction: application to cardiorespiratory interaction. *Physical review. E, Statistical, nonlinear, and soft matter physics 65*, 4 (2002), 11.

[221] Rubinov, M., and Sporns, O. Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage 52*, 3 (2010), 1059–1069.

[222] Ryckman, J., Hilton, C., Rogers, C., and Pineda, R. Sensory processing disorder in preterm infants during early childhood and relationships to early neurobehavior. *Early Human Development 113* (oct 2017), 18–22.

[223] Saji, R., Hirasawa, K., Ito, M., Kusuda, S., Konishi, Y., and Taga, G. Probability distributions of the electroencephalogram envelope of preterm infants. *Clinical Neurophysiology 126*, 6 (2015), 1132–1140.

[224] Sameshima, K., Takahashi, D. Y., and Baccalá, L. A. On the statistical performance of Granger-causal connectivity estimators. *Brain Informatics 2*, 2 (jun 2015), 119–133.

[225] Schiecke, K., Pester, B., Piper, D., Benninger, F., Feucht, M., Leistritz, L., and Witte, H. Nonlinear Directed Interactions Between HRV and EEG Activity in Children With TLE. *IEEE Transactions on Biomedical Engineering 63*, 12 (dec 2016), 2497–2504.

[226] Schneider, A., Hommel, G., and Blettner, M. Lineare regressionsanalyse - Teil 14 der serie zur bewertung wissenschaftlicher publikationen, nov 2010.

[227] Schreiber, T. Measuring Information Transfer. *Physical Review Letters 85*, 2 (jul 2000), 461–464.

[228] Schreiber, T., and Schmitz, A. Surrogate time series. *Physica D 142* (2000), 346–382.

[229] Schumacher, E., Stiris, T., and Larsson, P. Effective connectivity in long-term EEG monitoring in preterm infants. *Clinical Neurophysiology 126*, 12 (2015), 2261–2268.

[230] Schwab, K., Skupin, H., Eiselt, M., Walther, M., Voss, A., and Witte, H. Coordination of the EEG and the heart rate of preterm neonates during quiet sleep. *Neuroscience Letters 465*, 3 (nov 2009), 252–256.

[231] Selig, F. A., Tonolli, E. R., Silva, E. V. C. M. D., and Godoy, M. F. D. Heart rate variability in preterm and term neonates. *Arquivos brasileiros de cardiologia 96*, 6 (2011), 443–449.

[232] Selton, D., André, M., Debruille, C., Deforge, H., Fresson, J., and Hascoet, J. M. EEG at 6 weeks of life in very premature neonates. *Clinical Neurophysiology 121*, 6 (jun 2010), 818–822.

[233] Selton, D., Andre, M., Debruille, C., Deforge, H., and Hascoët, J. M. Cognitive outcome at 5years in very premature children without severe early cerebral abnormalities: Relationships with EEG at 6 weeks after birth. *Neurophysiologie Clinique 43*, 5-6 (dec 2013), 289–297.

[234] Seth, A. K. Causal connectivity of evolved neural networks during behavior. *Network (Bristol, England) 16*, 1 (mar 2005), 35–54.

[235] Slater, R., Cornelissen, L., Fabrizi, L., Patten, D., Yoxen, J., Worley, A., Boyd, S., Meek, J., and Fitzgerald, M. Oral sucrose as an analgesic drug for procedural pain in newborn infants: A randomised controlled trial. *The Lancet 376*, 9748 (oct 2010), 1225–1232.

[236] Slater, R., Fabrizi, L., Worley, A., Meek, J., Boyd, S., and Fitzgerald, M. Premature infants display increased noxious-evoked neuronal activity in the brain compared to healthy age-matched term-born infants. *NeuroImage 52*, 2 (aug 2010), 583–589.

[237] Slater, R., Worley, A., Fabrizi, L., Roberts, S., Meek, J., Boyd, S., and Fitzgerald, M. Evoked potentials generated by noxious stimulation in the human infant brain. *European Journal of Pain 14*, 3 (mar 2010), 321–326.

[238] Smith, G. C., Gutovich, J., Smyser, C., Pineda, R., Newnham, C., Tjoeng, T. H., Vavasseur, C., Wallendorf, M., Neil, J., and Inder, T. Neonatal intensive care unit stress is associated with brain development in preterm infants. *Annals of Neurology 70*, 4 (oct 2011), 541–549.

[239] Smith, S. L., Haley, S., Slater, H., and Moyer-Mileur, L. J. Heart rate variability during caregiving and sleep after massage therapy in preterm infants. *Early Human Development 89*, 8 (2013), 525–529.

[240] Stam, C. J., Nolte, G., and Daffertshofer, A. Phase lag index: Assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. *Human Brain Mapping 28*, 11 (nov 2007), 1178–1193.

[241] Stevens, M. C. The developmental cognitive neuroscience of functional connectivity. *Brain and Cognition 70*, 1 (2009), 1–12.

[242] Suykens, J., and Vandewalle, J. Least Squares Support Vector Machine Classifiers. *Neural Processing Letters 9*, 3 (1999), 293–300.

[243] Suykens, J. A. K., Van Gestel, T., De Brabanter, J., De Moor, B., and Vandewalle, J. *Least Squares Support Vector Machines*. WORLD SCIENTIFIC, nov 2002.

[244] Szabó, Z. Information Theoretical Estimators Toolbox. *Journal of Machine Learning Research 15* (2014), 283–287.

[245] Takahashi, E., Folkerth, R. D., Galaburda, A. M., and Grant, P. E. Emerging Cerebral Connectivity in the Human Fetal Brain: An MR Tractography Study. *Cerebral Cortex 22*, 2 (feb 2012), 455–464.

[246] The International Neonatal Network. The CRIB (clinical risk index for babies) score: a tool for assessing initial neonatal risk and comparing performance of neonatal intensive care units. *The Lancet 342*, 8865 (jul 1993), 193–198.

[247] Tibshirani, R. Regression Shriknage and Selectino via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological) 58*, 1 (1996), 267–288.

[248] Torsney, C., and Fitzgerald, M. Spinal Dorsal Horn Cell Receptive Field Size is Increased in Adult Rats Following Neonatal Hindpaw Skin Injury. *The Journal of Physiology 550*, 1 (jul 2003), 255–261.

[249] Vaart, M., Duff, E., Raafat, N., Rogers, R., Hartley, C., and Slater, R. Multimodal pain assessment improves discrimination between noxious and non-noxious stimuli in infants. *Paediatric and Neonatal Pain 1*, 1 (sep 2019), 21–30.

[250] van Wijk, B. C. M., Stam, C. J., and Daffertshofer, A. Comparing Brain Networks of Different Size and Connectivity Density Using Graph Theory. *PLoS ONE 5*, 10 (oct 2010), e13701.

[251] Vanhatalo, S., and Kaila, K. Development of neonatal EEG activity: From phenomenology to physiology. *Seminars in Fetal and Neonatal Medicine 11*, 6 (dec 2006), 471–478.

[252] Vapnik, V. N. An overview of statistical learning theory, 1999.

[253] Varon, C., Alzate, C., and Suykens, J. A. K. Noise Level Estimation for Model Selection in Kernel PCA Denoising. *IEEE Transactions on Neural Networks and Learning Systems 26*, 11 (nov 2015), 2650–2663.

[254] Varon, C., Faes, L., Testelmans, D., Buyse, B., and Huffel, S. V. Information transfer between respiration and heart rate during sleep apnea. In *2016 Computing in Cardiology Conference (CinC)* (2016), pp. 845–848.

[255] Verriotis, M., Chang, P., Fitzgerald, M., and Fabrizi, L. The development of the nociceptive brain. *Neuroscience 338* (dec 2016), 207–219.

[256] Vicente, R., Wibral, M., Lindner, M., and Pipa, G. Transfer entropy—a model-free measure of effective connectivity for the neurosciences. *Journal of Computational Neuroscience 30*, 1 (feb 2011), 45–67.

[257] Vijayakumar, V., Case, M., Shirinpour, S., and He, B. Quantifying and Characterizing Tonic Thermal Pain Across Subjects From EEG Data Using Random Forest Models. *IEEE Transactions on Biomedical Engineering 64*, 12 (dec 2017), 2988–2996.

[258] Vinall, J., and Grunau, R. E. Impact of repeated procedural pain-related stress in infants born very preterm, feb 2014.

[259] Vinall, J., Miller, S. P., Bjornson, B. H., Fitzpatrick, K. P., Poskitt, K. J., Brant, R., Synnes, A. R., Cepeda, I. L., and Grunau, R. E. Invasive procedures in preterm children: Brain and cognitive development at school age. *Pediatrics 133*, 3 (mar 2014), 412–421.

[260] Vincent, J. L., Patel, G. H., Fox, M. D., Snyder, A. Z., Baker, J. T., Van Essen, D. C., Zempel, J. M., Snyder, L. H., Corbetta, M., and Raichle, M. E. Intrinsic functional architecture in the anaesthetized monkey brain. *Nature 447*, 7140 (may 2007), 83–86.

[261] Wallois, F. Synopsis de la maturation des activités cérébrales chez le nouveau-né prématuré. *Neurophysiologie Clinique 40*, 2 (2010), 125–126.

[262] Watanabe, K., Hayakawa, F., and Okumura, A. Neonatal EEG: A powerful tool in the assessment of brain damage in preterm infants, sep 1999.

[263] Watanabe, K., Hayakawa, F., and Okumura, A. Neonatal EEG: A powerful tool in the assessment of brain damage in preterm infants, sep 1999.

[264] Weeke, L. C., Dix, L. M., Groenendaal, F., Lemmers, P. M., Dijkman, K. P., Andriessen, P., De Vries, L. S., and Toet, M. C. Severe hypercapnia causes reversible depression of aEEG background activity in neonates: An observational study. *Archives of Disease in Childhood: Fetal and Neonatal Edition 102*, 5 (sep 2017), F383–F388.

[265] Welch, M. G., Stark, R. I., Grieve, P. G., Ludwig, R. J., Isler, J. R., Barone, J. L., and Myers, M. M. Family nurture intervention in preterm infants increases early development of cortical activity and independence of regional power trajectories. *Acta Paediatrica 106*, 12 (dec 2017), 1952–1960.

[266] Wendt, H., Abry, P., and Jaffard, S. Bootstrap for Multifractal Analysis. *IEEE Signal Processing Magazine 24*, 4 (2007), 38–48.

[267] West, C. R., Groves, A. M., Williams, C. E., Harding, J. E., Skinner, J. R., Kuschel, C. A., and Battin, M. R. Early low cardiac output is associated with compromised electroencephalographic activity in very preterm infants. *Pediatric Research 59*, 4 PART 1 (apr 2006), 610–615.

[268] Widjaja, D., Orini, M., Vlemincx, E., and Huffel, S. V. Cardiorespiratory Dynamic Response to Mental Stress: A Multivariate Time-Frequency Analysis. *Computational and Mathematical Methods in Medicine* (2013).

[269] Wim De Clercq, Vergult, A., Vanrumste, B., Van Paesschen, W., and Van Huffel, S. Canonical Correlation Analysis Applied to Remove Muscle Artifacts From the Electroencephalogram. *IEEE Transactions on Biomedical Engineering 53*, 12 (nov 2006), 2583–2587.

[270] Witte, H., Putsche, P., Eiselt, M., Hoffmann, K., Schack, B., Arnold, M., and Jäger, H. Analysis of the interrelations between a low-frequency and a high- frequency signal component in human neonatal EEG during quiet sleep. *Neuroscience Letters 236*, 3 (nov 1997), 175–179.

[271] Zhang, D., Ding, H. H., Liu, Y., Zhou, C., Ding, H. H., and Ye, D. Neurodevelopment in newborns: A sample entropy analysis of electroencephalogram. *Physiological Measurement 30*, 5 (2009), 491–504.

[272] Ziv, Y., Aviezer, O., Gini, M., Sagi, A., and Koren-Karie, N. Emotional availability in the mother-infant dyad as related to the quality of infant-mother attachment relationship. *Attachment and Human Development 2*, 2 (2000), 149–169.

Curriculum Vitae

Mario Lavanga was born in Rho, Italy on July 8th, 1991. Since his childhood Mario has developed an early interest in human physiology and biology. He went on to study Biomedical engineering at Polytechnic university of Milan, where he delved into the intertwined realities of physics and biology. During his studies, he had the opportunity to build a strong background in the field of biomedical signal processing. Furthermore, he joined a top-level students program organized by the two main Italian technical universities, known as Alta Scuola Politecnica (ASP), which provided opportunity to learn a multidisciplinary approach to solving problems. He graduated in 2015 with a thesis entitled *Advanced approaches for blood pressure regulation assessment. Application in a porcine model of cardiac arrest*, which focused on the dynamics of the baroreflex sensitivity after cardiac arrest in an animal model.

After graduating, he moved to KU Leuven, Belgium, where he started focusing on mental health and neurodevelopment as part of his PhD studies. In January 2017, he received a strategic basic research grant from the Flemish scientific research fund (FWO). The goal of his research is to quantify stress in premature babies by means of EEG and ECG signal processing and to understand its impact on the infants' development. He is also interested in nonlinear and chaotic time series analysis and graph theory applied to the biomedical world.

Aware of the challenges regarding public institutions transparency and general public engagement in science and research, Mario recently developed an interest in science communication in order to raise awareness about mental health research and its importance. Besides science, Mario enjoys reading, travelling in Asia and political activism.

List of Publications

Articles in internationally reviewed journals

1. **Lavanga M.**, Heremans E., Moeyersons J., Bollen B., Jansen K., Ortibus E., Van Huffel S., Naulears G., Caicedo, A. (2020). Maturation of the autonomic nervous system in premature infants: estimating development based on heart-rate variability analysis. Submitted to Frontiers in Physiology, 1-28.

2. **Lavanga M.**, Bollen B., A. Dereymaeker, Jansen K., Ortibus E., Van Huffel S., Naulaers G., Caicedo A. (2020). The effect of early procedural pain in preterm infants on the maturation of EEG and heart rate variability. Submitted to PAIN, 1–11.

3. **Lavanga M.**, Bollen B., Jansen K., Ortibus E., Van Huffel S., Naulaers G., Caicedo A. (2020). A bradycardia-based stress calculator for the neonatal intensive care unit: a multisystem approach. Frontiers in Physiology, 11, 74.

4. **Lavanga M.**, Smets L., Bollen B., Jansen, K., Ortibus E., Van Huffel S., Naulaers G., Caicedo A. (2020). A perinatal stress calculator for the neonatal intensive care unit: an unobtrusive approach. Physiological Measurement, 1–26.

5. De Ridder J., **Lavanga M.**, Verhelle B., Vervisch J., Kotulska K., Moavero R., Curatolo P., Weschke B., Riney K., Feucht M., Krsek P., Nabbour R., Jansen A.C., Wojdan K., Domanska-Pakiela D., Kaczorowska-Frontczak M., Hertzberg C., Ferrier C.H., Samueli S., Jahodova A., Aronica E., Kwiatkowski D.J., Jansen F.E., Jóźwiak S., Van Huffel S., Lagae L. (2020). Prediction of neurodevelopment in infants with Tuberous Sclerosis Complex using early EEG characteristics. Frontiers in Neurology (*in press*), 1–20.

6. De Wel O., **Lavanga M.**, Caicedo A., Jansen K., Naulaers G., Van Huffel S. (2019). Decomposition of a Multiscale Entropy Tensor for Sleep Stage Identification in Preterm Infants. Entropy, 21(10), 936.

7. Hendrikx D., Smits, A., **Lavanga M.**, De Wel O., Thewissen L., Jansen K., Jansen K., Naulaers G. (2019). Measurement of Neurovascular Coupling in Neonates. Frontiers in Physiology, 10, 65.

8. Ansari A. H., De Wel O., **Lavanga M.**, Caicedo A., Dereymaeker A., Jansen K., Van Huffel S., De Vos M. (2018). Quiet sleep detection in preterm infants using deep convolutional neural networks. Journal of Neural Engineering, 15(6), 066006.

9. **Lavanga M.**, De Wel O., Caicedo A., Jansen K., Dereymaeker A., Naulaers G., Van Huffel S. (2018). A brain-age model for preterm infants based on functional connectivity. Physiological Measurement, 39(4), 044006.

10. De Wel O., **Lavanga M.**, Dorado A., Jansen K., Dereymaeker A., Naulaers G., Van Huffel S. (2017). Complexity Analysis of Neonatal EEG Using Multiscale Entropy: Applications in Brain Maturation and Sleep Stage Classification. Entropy, 19(10), 516.

11. **Lavanga M.**, De Wel O., Caicedo A., Jansen K., Dereymaeker A., Naulaers G., Van Huffel S. (2017). Monitoring Effective Connectivity in the Preterm Brain: A Graph Approach to Study Maturation. Complexity, 2017(ID 9078541), 1–13.

12. **Lavanga M.**, Baselli G., Fumagalli F., Ristagno G., Ferrario M. (2017). The possible role of the vagal nervous system in the recovery of the blood pressure control after cardiac arrest: a porcine model study. Physiological Measurement, 38(1), 63–76.

13. **Lavanga M.**, De Ridder J., Kotulska K., Moavero R., Curatolo P., Weschke B., Riney K., Feucht M., Krsek P., Nabbour R., Jansen A.C., Wojdan K., Domanska-Pakiela D., Kaczorowska-Frontczak M., Hertzberg C., Ferrier C.H., Samueli S., Jahodova A., Aronica E., Kwiatkowski D.J., Jansen F.E., Jóźwiak S., Lagae L., Van Huffel S., Caicedo A. (2020). Results of quantitative EEG analysis are associated with autism spectrum disorder and development abnormalities in infants with TSC. Submitted to Computer Methods and Programs in Biomedicine, 1–18.

Papers in proceedings of international conferences

1. **Lavanga M.**, De Wel O., Caicedo A., Deviaene, M., Moeyersons J., Varon C., Bollen B., Jansen K., Ortibus E., Naulaers G., Van Huffel S. (2019). The implementation of an apnea-based perinatal stress calculator. *Proceedings of the 41th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, EMBS, (pp 6000–6003).

2. **Lavanga M.**, De Wel O., Caicedo A., Heremans E., Jansen K., Dereymaeker A., Naulaers G., Van Huffel S. (2017). Automatic quiet sleep detection based on multifractality in preterm neonates: Effects of maturation. *Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, EMBS, (pp. 2010–2013).

3. **Lavanga M.**, De Wel O., Caicedo Dorado A., Jansen K., Dereymaeker A., Naulaers G., Van Huffel S. (2016). Linear and nonlinear functional connectivity methods to predict brain maturation in preterm babies. *Proceedings of the 8th International Workshop on Biosignal Interpretation*, (pp. 37–40).

Abstract in proceedings of (inter)national conferences

1. **Lavanga M.**, De Wel O., Caicedo A., Deviaene, M., Moeyersons J., Varon C., Bollen B., Jansen K., Ortibus E., Naulaers G., Van Huffel S. (2019). Perinatal stress quantification in preterm infants. In *18th National Day on Biomedical Engineering.* Brussels, Belgium.

2. **Lavanga M.**, De Wel O., Caicedo A., Deviaene, M., Moeyersons J., Varon C., Bollen B., Jansen K., Ortibus E., Naulaers G., Van Huffel S. (2019). Perinatal stress quantification in premature infants. In *Second International Summer Institute of Netwrok Physiology.* Como, Italy.

3. **Lavanga M.**, Bernagie C., De Wel O., Caicedo A., Jansen K., Ortibus E., Naulaers G., Van Huffel S. (2019). A novel brain-heart interaction framework to detect stress in premature infants. In *11th International Conference on Brain Monitoring and Neuroprotection in the Newborn.* Clearwater Beach, Florida.

4. **Lavanga M.**, Heremans E., Moeyersons J., De Wel O., Jansen K., Ortibus E., Van Huffel S., Naulears G., Caicedo, A. (2019). An autonomic age

model for premature infants based on heart-rate variability. In *11th International Conference on Brain Monitoring and Neuroprotection in the Newborn.* Clearwater Beach, Florida.

5. **Lavanga M.**, Bernagie C., Caicedo A., Jansen K., Ortibus E., Naulaers G., Van Huffel S. (2019). Brain-heart interaction to assess stress in premature infants. In *7th Dutch Biomedical Engineering Conference.* Egmond aan Zee, The Netherlands.

6. **Lavanga M.**, Heremans E., Jansen K., Ortibus E., Van Huffel S., Naulears G., Caicedo, A. (2019). A heart-rate variability model for the maturation of premature infants. In *7th Dutch Biomedical Engineering Conference.* Egmond aan Zee, The Netherlands.

7. **Lavanga M.**, Bernagie C., Caicedo A., Jansen K., Ortibus E., Naulaers G., Van Huffel S. (2018). Stress detection in preterm infants based on brain-heart connectivity. In *17th National Day on Biomedical Engineering.* Brussels, Belgium.

8. **Lavanga M.**, De Wel O., Caicedo A., Bernagie C., Houben J., Jansen K., Ortibus E., Naulaers G., Van Huffel S. (2018). Stress detection in preterm infants based on brain-heart connectivity. In *10th ESGCO meeting.* Vienna, Austria.

9. **Lavanga M.**, De Wel O., Caicedo A., Jansen K., Dereymaeker A., Naulaers G., Van Huffel S. (2017). An EEG maturation study based on multifractality in preterm neonates. In *16th National Day on Biomedical Engineering.* Brussels, Belgium.

10. **Lavanga M.**, De Wel O., Caicedo A., Dereymaeker A., Jansen K., Naulaers G., Van Huffel S. (2017). An EEG maturation study based on multifractality in preterm neonates. In *10th International Conference on Brain Monitoring and Neuroprotection in the Newborn.* Killarney, Ireland.

11. **Lavanga M.**, De Wel O., Caicedo A., Jansen, K., Dereymaeker A., Naulaers G., Van Huffel S. (2017). Investigation of preterm brain development by means of directed graphs. In *Coupling and Causality in Complex Systems conference.* Cologne, Germany.

12. **Lavanga M.**, De Wel O., Caicedo A., Jansen K., Dereymaeker A., Naulaers G., Van Huffel S. (2017). An age prediction model for preterm neonates on brain connectivity. In *6th Dutch Biomedical Engineering Conference.* Egmond aan Zee, The Netherlands.



FACULTY OF ENGINEERING SCIENCE
DEPARTMENT OF ELECTRICAL ENGINEERING
STADIUS
Kasteelpark Arenberg 10 box 2446
B-3001 Leuven
mlavanga@esat.kuleuven.be
https://www.esat.kuleuven.be/stadius/