



HAL
open science

Analysis and Extraction of Complexity Parameters of Biomedical Signals

Amira Zaylaa

► **To cite this version:**

Amira Zaylaa. Analysis and Extraction of Complexity Parameters of Biomedical Signals. Bioengineering. François-Rabelais University of Tours, 2014. English. NNT: . tel-01116466

HAL Id: tel-01116466

<https://inserm.hal.science/tel-01116466>

Submitted on 13 Feb 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



FRANÇOIS-RABELAIS UNIVERSITY OF TOURS
DOCTORAL SCHOOL OF ENERGY, MATERIAL, EARTH AND UNIVERSE
SCIENCES

LABORATORY AND RESEARCH TEAM : Imaging and Brain Laboratory, Imaging
and Ultrasound Team, National Institute of Medical and Health Sciences Unit 930

DISSERTATION presented by :

Amira ZAYLAA

Defended on : December 15th 2014

in partial fulfilment of the requirements for the degree of: **Doctor of Philosophy**
From François-Rabelais University of Tours, France

Discipline/ Speciality : Life and Health Sciences/ Biomedical Signal Processing

**Analysis and Extraction of Complexity Parameters of
Biomedical Signals**

DISSERTATION SUPERVISED BY :

Jean-Marc GIRAULT	Associate Professor QRS	François-Rabelais University of Tours, France
Jamal CHARARA	Professor	Lebanese University, Beirut, Lebanon

COMMITTEE MEMBERS :

Tuan PHAM	Professor	University of Aizu, Japan
Mohamad KHALIL *	Professor	Lebanese University, Tripoli, Lebanon
Catherine MARQUE	Professor	University of Technology of Compiègne, France
Jamal CHARARA	Professor	Lebanese University, Beirut, Lebanon
Jean-Marc GIRAULT	Associate Professor QRS	François-Rabelais University of Tours, France
Oussama BAZZI	Professor	Lebanese University, Beirut, Lebanon

* *committee president*

*There is behind every distinguished work a group of extraordinary supporters,
to my family, friends and every person fond of science*

Acknowledgements

This dissertation has never been possible without the will and blessing of God, the most gracious and the most merciful.

First, I would like to express my gratitude to my supervisors Dr. Jean-Marc Girault and Prof. Jamal Charara for their guidance and sincere advice. I would like also to express my profound gratitude to Prof. Tuan Pham from the University of Aizu in Japan and Prof. Catherine Marque from the University of Technology of Compiègne for being in the committee of my PhD and for their valuable comments. Special thanks to Profs. Mohammad Khalil and Oussama Bazzi for being in the committee of my dissertation and for their supportive comments.

The interdisciplinary nature of my work has allowed me to have great collaborators. I acknowledge Dr. Sébastien Ménigot from the University of Paris-10 and Miss Faten Khatib from Sharp Chula Vista Medical Center in U.S.A.. They have provided much guidance and assistance for my research and gave me very helpful proofreading of my published articles. The same gratitude goes to my colleagues in the Imaging and Brain Laboratory, Drs. Liviu Chira, Fatima Sbeity and Souad Oudjemia, and Mr. Maroun Geryes for their assistance in my project. My gratefulness extends to Prof. Denis Guilloteau the director of the Imaging and Brain Laboratory in Tours, and thanks to Prof. Ayache Bouakaz the director of research for his accessibility and advices. Special thank is due to Prof. Charles Tabet from the National Council for Scientific Research (Lebanon) for his support and accessibility.

I am deeply grateful to my parents, my mother Ahlam Dalle who formed part of my vision and taught me the good things that really matter in life, my sister Alaa Zaylaa, my brother Omar Zaylaa, my sister-in-law Afnan Qudus and my cousin Faten Khatib for keeping me optimistic and inspired, and for their support and encouragement that are a constant source of power for me.

I acknowledge those who gave me the possibility to complete this thesis: Dr. Keinana

ACKNOWLEDGEMENTS

Muhrez, Zainab Zahereddine, Mohammad Chehimi, Dr. Mohammad Al Akhras, Dr. Darine Abi-Haidar, Yann Sindakli, Michèle Françoise and her family, Dominique and Erica and my friends in France.

Abstract

The analysis of biomedical time series derived from nonlinear dynamic systems is challenging due to the chaotic nature of these time series. Only few classical parameters can be detected by clinicians to opt the state of patients and fetuses. Though there exist valuable complexity invariants such as multi-fractal parameters, entropies and recurrence plots, they were unsatisfactory in certain cases. To overcome this limitation, we propose in this dissertation new entropy invariants, we contributed to multi-fractal analysis and we developed signal-based (unbiased) recurrence plots and unbiased recurrence descriptors based on the dynamic transitions of time series.

Principally, we aim to improve the discrimination between healthy and distressed biomedical systems, particularly fetuses by processing the time series using our techniques. These techniques were either validated on Lorenz systems, logistic maps or fractional Brownian motions which model chaotic and random time series. Then the techniques were applied to real fetus heart rate signals recorded from patients in the third trimester of pregnancy. Statistical measures comprising the relative error, standard deviation, sensitivity, specificity, precision and accuracy were employed to evaluate the performance of detection.

Elevated discernment outcomes were realized by the high-order entropy invariants developed. Multi-fractal analysis using a structure function and coarse-graining enhanced the detection of the medical states of the fetuses. Unbiased cross-determinism invariant developed amended the discrimination process. The significance of our techniques lies behind their post-processing codes which could build up cutting-edge portable machines offering advanced discrimination and detection of Intrauterine Growth Restriction prior to fetal death. This work was devoted to Fetal Heart Rates but time series generated by alternative nonlinear dynamic systems should be further considered.

Keywords : Multi-fractal Analysis, Entropy Quantification, Maximum Entropy, N-Order Entropy, Recurrence Plots, Unbiased Recurrence Plots, Recurrence Quantification Analysis, New Invariants, Discrimination, Detection, Diagnosis, Doppler Ultrasound Fetal Heart Rates, fractional Brownian Motion, Lorenz System, Logistic Map, Correlation Sum, Complexity Analysis, Statistical Signal Processing.

Résumé

L'analyse de séries temporelles biomédicales chaotiques tirées de systèmes dynamiques non-linéaires est toujours un challenge difficile à relever puisque dans certains cas bien spécifiques les techniques existantes basées sur les multi-fractales, les entropies et les graphes de récurrence échouent. Pour contourner les limitations des invariants précédents, de nouveaux descripteurs peuvent être proposés. Dans ce travail de recherche nos contributions ont porté à la fois sur l'amélioration d'indicateurs multi-fractals (basés sur une fonction de structure) et entropiques (approchées) mais aussi sur des indicateurs de récurrences (non biaisés).

Ces différents indicateurs ont été développés avec pour objectif majeur d'améliorer la discrimination entre des signaux de complexité différente ou d'améliorer la détection de transitions ou de changements de régime du système étudié. Ces changements agissant directement sur l'irrégularité du signal, des mouvements browniens fractionnaires et des signaux tirés du système du Lorenz ont été testés. Ces nouveaux descripteurs ont aussi été validés pour discriminer des fœtus en souffrance de fœtus sains durant le troisième trimestre de grossesse. Des mesures statistiques telles que l'erreur relative, l'écart type, la spécificité, la sensibilité ou la précision ont été utilisées pour évaluer les performances de la détection ou de la classification.

Le fort potentiel de ces nouveaux invariants nous laisse penser qu'ils pourraient constituer une forte valeur ajoutée dans l'aide au diagnostic s'ils étaient implémentés dans des logiciels de post-traitement ou dans des dispositifs biomédicaux. Enfin, bien que ces différentes méthodes aient été validées exclusivement sur des signaux fœtaux, une future étude incluant des signaux tirés d'autres systèmes dynamiques non-linéaires sera réalisée pour confirmer leurs bonnes performances.

Mots clés : Analyse Multi-Fractale, Quantification d'entropie, Entropie Maximale, Entropie d'Ordre-N, Le graphe de Récurrence, Le graphe de Récurrence Nonbiaisés, Analyse de Quantification de Récurrence, Nouveaux Invariants, Détection, Discrimination, Diagnostiquer, Transition Dynamique, Suite Logistique, Système du Lorenz, Mouvements Browniens Fractionnaires, Coefficient de Corrélation, Analyse de Complexité, Traitement Statistique du Signal, Rythme Cardiac Fœtal Doppler Ultrasonore.

Contents

Abstract	7
Résumé	9
List of Tables	16
List of Figures	21
List of Acronyms	23
I GENERAL INTRODUCTION	25
1 DISSERTATION	27
1.1 Context	27
1.2 Statement of Problem	29
1.3 Objectives	31
1.4 Structure	31
2 MATERIALS	33
2.1 Fractional Brownian Motion	33
2.2 Lorenz System	33
2.3 Logistic Map	34

2.4	Real Fetal Heart Rate Time Series	35
2.4.1	Multichannel Doppler Ultrasound Device and Fetal Heart Rates	35
2.4.2	Case Study and Protocol	38
II COMPLEXITY ANALYSIS AND APPLICATION TO BIOMEDICAL SIGNALS		39
3	ENTROPY ANALYSIS	41
3.1	Introduction	41
3.2	Major Existing Entropy	41
3.2.1	Approximate Entropy	42
3.2.2	Sample and Multi-Scale Sample Entropy	43
3.2.3	Similarity and Multi-Scale Similarity Entropy	45
3.2.4	Fuzzy Entropy	46
3.3	High Order Developed Entropy	48
3.3.1	Maximum Entropy	50
3.3.2	N-Order Entropy	50
3.3.3	Results of Fractional Brownian Motion and Lorenz System	52
3.4	Application to Fetal Heart Rates	60
3.4.1	Results of Fetal Heart Rate Complexity Detection	60
3.4.2	Discussion	63
3.5	Conclusion	65
4	MULTI-FRACTAL ANALYSIS	67
4.1	Introduction	67
4.2	Multi-Fractal Descriptors	69
4.2.1	Singularity Spectrum	69
4.2.2	Hurst Exponent	70
4.2.3	Holder Spectrum	70
4.3	Contribution to Multi-Fractal Analysis	71

CONTENTS

4.4	Example On Fractional Brownian Motion	74
4.5	Application To Fetal Heart Rates	75
4.5.1	Results Discriminating Healthy from Distressed Fetuses	76
4.5.2	Discussion	83
4.6	Conclusion	83
5	RECURRENCE PLOTS AND ANALYSIS	85
5.1	Introduction	85
5.2	Recurrence Analysis	87
5.2.1	Recurrence Matrix	87
5.2.2	Recurrence Quantification Analysis and Our Descriptors	87
5.2.3	Sojourn Points Demonstration	88
5.3	Existing Recurrence Plots	91
5.3.1	Unembedded Recurrence Plots	91
5.3.2	Embedded Recurrence Plots	91
5.4	Developed Recurrence Plot Methods	92
5.4.1	Embedded Recurrence Plot With Specific Settings	92
5.4.2	Derivative-Based Recurrence Plot	93
5.4.3	M -Time Pattern Recurrence Plot	93
5.4.4	Example on Sine Wave	94
5.4.5	Results of Logistic Map	97
5.5	Application To Fetal Heart Rates	103
5.5.1	Results of Diagnosing Fetal Heart Rate Distress	103
5.5.2	Discussion	109
5.6	Conclusion	111
III	DISSERTATION SUMMARY AND PERSPECTIVES	113
6	CONCLUSION AND PERSPECTIVES	115
6.1	Comparison of Complexity Analysis Techniques	115

6.2	Conclusions	117
6.3	Limitations	118
6.4	Outlook	118
	References	132
	Annexes	135
A	Pseudocodes of Existing Entropy Parameters	135
B	Pseudocodes of Developed Unbiased RPs	139
C	Ph.D. Activities	141
D	List of Publications	143
E	Journal Papers	145

List of Tables

3.1	Average \pm standard deviation of specific values of $\mathcal{E}_p(n, m)$ for two fBms with $i = 1$, $\mathcal{H} = 0.07$ and $i = 2$, $\mathcal{H} = 0.3$. Values in parentheses represent the standard deviation normalized by the average value.	60
3.2	Average \pm standard deviation of specific values of discrimination function $\Psi_{\mathcal{E}_p}(n, m)$ for two fBms with $\mathcal{H} = 0.07, 0.3$. Values in parentheses represent the standard deviation over the average value and the asterisk represents the maximum.	60
3.3	Discrimination functions $\Psi_{\mathcal{M}\infty}$ from non-Fuzzy estimations derived from FHR signals with k, l ranging from 1 to 7. Italic values correspond to 1-order similarity entropies. Bold values correspond to the maximum of n-order similarity entropies.	61
3.4	Discrimination functions $\Psi_{\mathcal{M}2}$ from fuzzy estimations derived from FHR signals with k, l ranging from 1 to 7. Italic values correspond to 1-order similarity entropies. Bold values correspond to the maximum of n-order similarity entropies.	61
3.5	Average values (aver.) and relative accuracy (standard deviation/ average = std./aver.) of discrimination functions $\Psi_{\mathcal{E}_p}(n, m)$ derived from FHR signals. Fuzzy estimations were of benefit for reducing estimation fluctuations (std./aver.)	62
3.6	Sensitivity of discrimination functions $\Psi_{\mathcal{E}_p}(n, m)$ calculated from FHR signals. Note that $\Psi_{\mathcal{E}_p}^*(n, 1)$ was equal to $\Psi_{\mathcal{E}\infty}^*(4, 1)$ for non-fuzzy estimations and to $\Psi_{\mathcal{E}2}^*(5, 1)$ for fuzzy estimations. Fuzzy estimations were of benefit since a slight improvement in both sensitivity and specificity can be seen. .	62

4.1	Relative errors of different multi-fractal parameters between the two groups of fetuses for different scales.	82
5.1	The five nonlinear recurrence methods used in the complexity analysis for transition detection.	95
5.2	% DET and % CDET quantifications of RP_1 , RP_2 , RP_3 and RP_4 for a logistic time series for $b = 4$	99
5.3	Performance of Detection (PD) of the logistic map dynamic transitions by means of the sensitivity and specificity measures.	103
5.4	The Relative Errors (REs) of the five RQA issued for diagnosing fetal distress and computed from the unembedded RP, embedded RP, embedded RP with specific settings, derivative based RP and 2-time pattern RP. . . .	107
5.5	The Effectiveness of Discrimination (ED) of Fetal Heart Rates (FHRs) by the sensitivity, specificity, precision and accuracy measures in percentages of the Recurrence Quantification Analysis (RQA).	109

List of Figures

1.1	Schematic illustration of the phases of the time series analysis, the origin of the three different complexity analysis techniques and the common point between them.	29
1.2	Schematic illustration of the three different complexity analysis techniques.	32
2.1	Fractional Brownian Motion. (top) A negatively correlated fractional Gaussian noise. (center) Brownian motion. (bottom) Positively correlated fractional Gaussian noise.	34
2.2	The bifurcation diagram and different dynamic regions of the Logistic Map.	35
2.3	Schematic illustration of the application setup from Doppler Ultrasound Fetal Heart Rate detection to Fetal Heart Rate extraction.	36
2.4	Fetal Heart Rate (FHR) recordings from two different fetal groups for 30-minutes at Bretonneau Hospital in Tours (France). (a) A Healthy Fetal Heart Rate (H-FHR) recording and (b) a Distressed Fetal Heart Rate (D-FHR) recording.	37
3.1	Schematic Illustration of Entropy computation from a chaotic time series using a pattern of $m = 2$	44
3.2	Schematic Illustration of an example of coarse graining for $s = \{1, 2\}$ and $k = \{1, \dots, M\}$	45
3.3	The membership function.	47

- 3.4 Complexity measurement for different fBm with \mathcal{H} ranging from 0.1 to 0.9 ($r = 0.2$). (a) Non-fuzzy estimations with $\Phi_\infty(m)$. (b) Non-fuzzy estimations with $\mathcal{E}_\infty(1, m)$. Red dots correspond to maxima. (c) Fuzzy estimations with $\Phi_2(m)$. (d) Fuzzy estimations with $\mathcal{E}_2(1, m)$. Red dots correspond to maxima. 53
- 3.5 1-order entropies (non-fuzzy and fuzzy) of 100 fBms with \mathcal{H} ranging from 0.1 to 0.9 ($r = 0.2$). (a) $C_\infty^*(\mathcal{H})$. Dots correspond to the mean estimation and red lines correspond to the standard deviation. Arrows represent a limit between anti-correlation and correlation. (b) $C_\infty^*(1, m^*)$. Lozenges represent the standard deviation of m^* and C_∞^* . (c) $C_2^*(\mathcal{H})$. Dots correspond to the mean estimation and the red lines correspond to the standard deviation. Arrows represent a limit between anti-correlation and correlation. (d) Fuzzy estimation with $C_2^*(1, m^*)$. Lozenges represent the standard deviation of m_2^* and C_2^* . (e) Zoom of (b). From the zoom it is easier to discriminate between high complexity and low complexity. (f) $\mathcal{H}(m^*)$. (g) Zoom of (d). From the zoom it is easier to discriminate high complexity from low complexity. (h) $\mathcal{H}(m^*)$ 54
- 3.6 1-order entropies for Lorenz time series ($r = 0.2$). (a) Non-fuzzy estimations with $\mathcal{E}_\infty^{(1)}$ for chaotic behavior ($\rho = 28$) (represented in blue) and $\mathcal{E}_\infty^{(2)}$ for periodic behavior ($\rho = 215$) (represented in green). (b) Fuzzy estimations with $\mathcal{E}_2^{(1)}$ for chaotic behavior ($\rho = 28$) (represented in blue) and $\mathcal{E}_2^{(2)}$ for periodic behavior ($\rho = 215$) (represented in green). 55
- 3.7 Complexity measurements for two different fBms with $\mathcal{H} = 0.07$ and $\mathcal{H} = 0.3$, ($r = 0.2$). (a) Non-fuzzy estimations with $\Phi_\infty^{(1)}(m)$ (in blue) for $\mathcal{H} = 0.07$ and $\Phi_\infty^{(2)}(m)$ (in red) for $\mathcal{H} = 0.3$ Deviation between $\Phi_\infty^{(1)}(m)$ and $\Phi_\infty^{(2)}(m)$ represented by the yellow area. (b) Non-fuzzy estimations with $|\mathcal{E}_\infty^{(1)}(1, m)|$ (in blue), $|\mathcal{E}_\infty^{(2)}(1, m)|$ (in red). Deviation between $|\mathcal{E}_\infty^{(1)}(1, m)|$, $|\mathcal{E}_\infty^{(2)}(1, m)|$ represented in the blue area. (c) Non-fuzzy estimation with $\Psi_{\mathcal{E}_\infty}(n, 1)$ (in black) and $\Psi_{\mathcal{E}_\infty}(1, m)$ (in magenta). Deviation between $\Psi_{\mathcal{E}_\infty}(n, 1)$ and $\Psi_{\mathcal{E}_\infty}(1, m)$ represented by the green area. (d) Fuzzy estimations with $\Phi_2^{(1)}(m)$ (in blue) for $\mathcal{H} = 0.07$ and $\Phi_2^{(2)}(m)$ (in red) for $\mathcal{H} = 0.3$. (e) Fuzzy estimations with $|\mathcal{E}_2^{(1)}(1, m)|$ (in blue), $|\mathcal{E}_2^{(2)}(1, m)|$ (in red). (f) Fuzzy estimations with $\Psi_{\mathcal{E}_2}(n, 1)$ (in black) and $\Psi_{\mathcal{E}_2}(1, m)$ (in magenta). Deviation between $\Psi_{\mathcal{E}_2}(n, 1)$ and $\Psi_{\mathcal{E}_2}(1, m)$ represented by the green area. 57

LIST OF FIGURES

3.8	Representation of the matrix $\mathcal{M}_p^{(i)}$ for two different fBms. (a) Non-fuzzy estimations $\Phi_\infty^{(1)}(m)$ (blue line) and $\Phi_\infty^{(2)}(m)$ (red line) with $\mathcal{H} = 0.07$ and $\mathcal{H} = 0.3$, respectively. (b) Non-fuzzy estimations with $\mathcal{M}_\infty^{(1)}$. (c) Non-fuzzy estimations with $\mathcal{M}_\infty^{(2)}$. (d) Non-fuzzy estimations with difference between $\mathcal{M}_\infty^{(1)}$ and $\mathcal{M}_\infty^{(2)}$. (e) Fuzzy estimations with $\Phi_2^{(1)}(m)$ (blue line) and $\Phi_2^{(2)}(m)$ (red line) with $\mathcal{H} = 0.07$ and $\mathcal{H} = 0.3$, respectively. (f) Fuzzy estimations with $\mathcal{M}_2^{(1)}$. (g) Fuzzy estimations with $\mathcal{M}_2^{(2)}$. (h) Fuzzy estimations with difference between $\mathcal{M}_2^{(1)}$ and $\mathcal{M}_2^{(2)}$	59
3.9	Boxplots of 1-order entropies and n-order entropies calculated from FHR signals. (left) Non-fuzzy estimations. Boxplot of 1-order entropies $\mathcal{E}(1, 1)$, $\mathcal{E}^*(1, 2)$, and n-order entropy $\mathcal{E}(4, 1)$ for healthy (in green) and distressed fetuses (in blue). (right) Fuzzy estimations. Boxplot of 1-order entropies $\mathcal{E}(1, 1)$, $\mathcal{E}^*(1, 2)$ and n-order entropy $\mathcal{E}(5, 1)$ for healthy (in green) and distressed fetuses (in blue).	63
4.1	Schematic illustration of the common multi-fractal analysis techniques.	68
4.2	The Block Diagram of the proposed multi-fractal analysis.	72
4.3	Time and spectral representations of a Brownian motion. (a) Original time series superimposed on the coarse-grained time series ($s= 8$) and the re-sampled coarse-grained time series ($s= 8$). (b) Spectrum of each time series depicted in (a).	73
4.4	Effects of coarse-graining on multi-fractal descriptors for different fBm of Hurst exponents $\mathcal{H} = \{0.1, 0.5, 1\}$ with different scale factor $\alpha = \{1, 3, 6\}$. (a) Singularity spectrum \mathbf{D} versus q , (b) Singularity spectrum \mathbf{D} versus Holder spectrum h , (c) Scaling exponent τ versus q and (d) Holder spectrum of healthy and distressed FHRs.	74
4.5	Multi-fractal parameters for a healthy fetus (in blue) and a distressed fetus (in green). (a) Structure function $Q(q, \epsilon)$ versus scale. (b) Scaling exponent η versus q . (c) Scaling exponent τ versus q . The two scaling exponents were more nonlinear for a healthy fetus than for a distressed fetus.	76
4.6	Multi-fractal parameters for a healthy fetus (shown in blue) and a distressed fetus (shown in green). (a) Singularity spectrum D versus q . (b) Singularity spectrum $D(q)$ versus Holder spectrum $h(q)$. (c) Holder spectrum h versus q	77

4.7	Boxplot of Hurst exponents versus scale for healthy (shown in blue) and distressed (shown in red) fetuses.	79
4.8	Boxplot of $\Delta_h = \max(h) - \min(h)$ versus the scale for healthy (shown in blue) and distressed (shown in red) fetuses.	80
4.9	Boxplot of D_{mean} versus the scale for healthy (shown in blue) and distressed (shown in red) fetuses.	81
4.10	Boxplot of Δ_D versus the scale. Healthy fetus (in blue) and distressed fetus (in red).	82
5.1	Schematic illustration of sojourn points and recurrence test computation of (a) a sine wave $x(t)$ and $y = x(t + \tau_{optimal})$, (b) a sine wave $x(t)$ and $y = \dot{x}(t)$ and (c) an m -time pattern sine wave.	89
5.2	Schematic illustration of sojourn points for a chaotic time series (aperiodic). (left) Chaotic time series. (right) Computation of the corresponding Recurrence Plot.	90
5.3	The cost-function and optimization process of the embedding dimension and time delay (left) for a sine wave and (right) for the healthy FHR. . . .	94
5.4	Centered Caption beside Object	96
5.5	Example of a simulated sine wave and its corresponding standard and developed recurrence plots. (a) Sine wave and its corresponding (b) unembedded recurrence plot (RP ₁), (c) embedded recurrence plot with specific settings (RP ₃), (d) derivative-based recurrence plot (RP ₄) and (e) 2-time pattern recurrence plot (RP ₅).	97
5.6	The standard and developed recurrence plots, originating from a simulated Logistic time series for $b = 4$. (a) the unembedded recurrence plot (RP ₁), (b) embedded recurrence plot (RP ₂), (c) embedded recurrence plot with specific settings RP ₃ , (d) derivative-based recurrence plot (RP ₄) and 2-time recurrence plot (RP ₅).	100

LIST OF FIGURES

5.7 Qualitative detection of dynamic transitions of the logistic map through thresholding the determinism versus the dynamic parameter, originating from both the standard and signal-based recurrence plots. (a) DET computed from the unembedded recurrence plot (RP₁), (b) the embedded recurrence plot (RP₂), (c) the embedded recurrence plot with specific settings (RP₃), (d) the derivative-based recurrence plot (RP₄) and (e) the 2-time recurrence plot (RP₅). 102

5.8 The Standard Recurrence Plots of FHRs. (a) The unembedded recurrence plot RP₁ of the H-FHR. (b) The unembedded recurrence plot RP₁ of the D-FHR. (c) The 2 - d embedded recurrence plot RP₂ of the H-FHR. (d) The 2 - d embedded recurrence plot RP₂ of the D-FHR. 104

5.9 The Unbiased Recurrence Plots of FHRs. (a) The 3 - d embedded recurrence plot with specific settings RP₃ of H-FHR. (b) The 3 - d embedded recurrence plot with specific settings RP₃ of D-FHR. (c) The derivative-based recurrence plot RP₄ of the H-FHR. (d) The derivative-based recurrence plot RP₄ of the D-FHR. (e) The m-time pattern recurrence plot RP₅ of the H-FHR. (f) The m-time pattern recurrence plot RP₅ of the D-FHR. 105

5.10 The Recurrence Quantification Analysis (RQA) relative to the standard and unbiased Recurrence Plots of a healthy and distressed FHRs shown in gray and orange colors, respectively. (a) The mean DET(%) as a function of the unembedded RP, embedded RP in 3 - d, embedded RP with specific settings, derivative-based RP and 2-time pattern RP, *RP*₁, *RP*₂, *RP*₃, *RP*₄ and *RP*₅, respectively. (b) The mean CDET(%) relative of the five RP techniques. (c) The average Percentage Reduced Sojourn Points (PRSP(%)) as a function of the five Recurrence Plots (RPs). (d) The average mean RR(%) as a function of the five RPs. (e) The mean En(%) as a function of the five RPs. 108

6.1 Schematic illustration of the primary results and major findings provided by our dissertation. 116

LIST OF FIGURES

List of Acronyms

<i>ApEn</i>	Approximate Entropy
<i>bmp</i>	beats per minute
<i>CHRU</i>	Centre Hospitalier Régional Universitaire
<i>CDET</i>	Cross-Determinism
<i>D – FHR</i>	Distressed-Fetal Heart Rate
<i>RP₄</i>	Derivative-Based Recurrence Plot
<i>DET</i>	Determinism
Δ	Difference between multi-fractal parameters
ϵ	Increment in multi-fractal analysis
<i>ECG</i>	Electrocardiogram
<i>EHG</i>	Electrohysterogram
<i>RP₂</i>	Embedded Recurrence Plot
<i>d</i>	Embedding dimension
<i>FHR</i>	Fetal Heart Rate
<i>fBm</i>	fractional Brownian Motion
<i>FuzzyEn</i>	Fuzzy Entropy
<i>FNN</i>	False Nearest Neighbour
<i>GeoEn</i>	Geostatistic Entropy
<i>HRV</i>	Heart Rate Variability
<i>H – FHR</i>	Healthy-Fetal Heart Rate
<i>IUGR</i>	Intrauterine Growth Restriction
<i>LLE</i>	Largest Lyapunov Exponent
<i>M</i>	Length of the time series

<i>LOI</i>	Length Of Identity
<i>LZ</i>	Lempel-Ziv
\mathcal{E}	Mathematical Symbol of Entropy
<i>RP₅</i>	m-time pattern Recurrence Plot
<i>MCDUS</i>	Multi-channel Doppler Ultrasound
<i>MI</i>	Mutual Information
<i>n – orderFuzzySimEn</i>	n-order Fuzzy Similarity Entropy
<i>N</i> or <i>n</i>	order of computation
<i>m</i>	Pattern Length
<i>PRSP</i>	Percentage Reduced Sojourn Point
ρ	Rayleigh Number
<i>RE</i>	Relative Error
<i>RP</i>	Recurrence Plot
<i>RR</i>	Recurrence Rate
<i>RP₃</i>	Recurrence Plot with Specific settings
<i>RQA</i>	Recurrence Quantification Analysis
<i>SampEn</i>	Sample Entropy
<i>SimEn</i>	Similarity Entropy
<i>SD</i>	Standard Deviation
<i>C</i>	Correlation Sum
Ψ	The difference between two entropy values
τ	Time Delay
<i>r</i>	Tolerance
<i>URP</i>	Unbiased Recurrence Plot
<i>RP₁</i>	Unembedded Recurrence Plot
<i>d'</i>	Virtual dimension

Part I

GENERAL INTRODUCTION

Chapter 1

DISSERTATION

▮ *Often complexity challenges our lives and makes its analysis worthwhile three years and more of hard-working* ▬ *Amira Zaylaa.*

1.1 Context

The study of signals and systems, particularly in the biomedical field, is of concern in the most developed countries in the world. However, their study is not simple as biomedical systems building up living things are nonlinear processes. These nonlinear processes exhibit complexity which makes linear analysis techniques unsatisfactory.

Complexity analysis is an alternative approach to the linear time series analysis. In the 50's, Kolomogorov was undoubtedly the first to address and characterize nonlinear dynamical systems with a new insight based on seeking novel invariants. Kolomogorov joined Sinai and both proposed metric invariants. Later, they proposed side-by-side the well-known KS Entropy which was also called the correlation entropy [Sinai, 1959]. In the 2000's, a renewed interest in the characterization of nonlinear systems led researchers to adopt KS entropy as their starting point.

Nowadays, chaotic systems are characterized by complexity parameters that practically reveal a part of their identity either based on the information theory (Entropy analysis), the geometrical theory (multi-fractal dimensions) or the chaotic theory (Recurrence Plots RPs and their quantification parameters). The previous analysis methods include, but not limited to, the Approximated Entropy (ApEn) [Pincus, 1991], the Sample Entropy (SampEn) [Richman and Moorman, 2000], Similarity Entropy (SimEn) [Lake et al., 2002], Fuzzy Entropy (FuzzyEn) [Liu et al., 2013], Permutation Entropy [Bandt and Pompe, 2002],

Geostatistical Entropy (GeoEn) [Pham, 2010], Lempel-Ziv (LZ) [Ziv and Lempel, 1978, Hu et al., 2006], Largest Lyapunov Exponent (LLE) [Perron, 1930], Multi-fractal parameters [Peters, 1996], Recurrence Plots [Eckmann et al., 1987] and Recurrence Quantification Analysis (RQA) [Marwan et al., 2002, Zbilut et al., 2002].

Instead of considering KS entropy the starting point of any research work, it is apparent from the bibliography that the correlation sum [Grassberger and Procaccia, 1983] $C(r)$ could be considered the starting point, as it is commonly included in most existing complexity descriptors. From our point of view, it seemed judicious to follow the correlation sum from the onset of the research work and consider it as our guideline.

To introduce the correlation sum, recall that $C(r)$ measures the mean probability that two time series states at two different times exist within a tolerance r , this condition is governed by the Heaviside function [Grassberger and Procaccia, 2004]. The correlation sum $C(r)$ was the core of both the entropy paradigm and certain descriptors based on the recurrence paradigm. The correlation dimension D_2 , which is related to $(\log(C(r))/\log(r))$, is a special case of the generalized dimension spectrum D_q for a set of data points [Akay, 2000]. D_2 is used for deciding whether the object is a multi-fractal and whether we could apply the multi-fractal analysis on a certain time series or not.

To illustrate the three analysis techniques we set out Fig. 1.1. Fig. 1.1 represents the major phases of the time series (TS) analysis, it shows the common point between the three analysis methods highlighted in a dashed orange rectangle.

$C(r)$ was a must to decide whether one should use a fractal analysis or a multi-fractal analysis, since the condition that should be satisfied in order to use the multi-fractal analysis depends on the computation of the box-counting dimension D_0 , information dimension D_1 and correlation dimension D_2 . If the three latter dimensions are not equal then the multi-fractal analysis has to be used [Akay, 2000]. Moreover, $C(r)$ was required to fill the Recurrence matrix in the Recurrence Plot analysis, and $C(r)$ was needed to compute the entropy parameters (see Fig. 1.1).

Although our research work could be applied to any nonlinear dynamical system, we decided to test and evaluate our different contributions on fetal heart rate (FHR) time series, as FHR analysis is of major concern in the two laboratories we are attached to. This dissertation focuses on the study of Entropy analysis, multi-fractal analysis and RPs to analyse chaotic biomedical signals. We emphasize on our motivation in improving the discrimination between different chaotic signals to detect certain problems, such as

1.2. STATEMENT OF PROBLEM

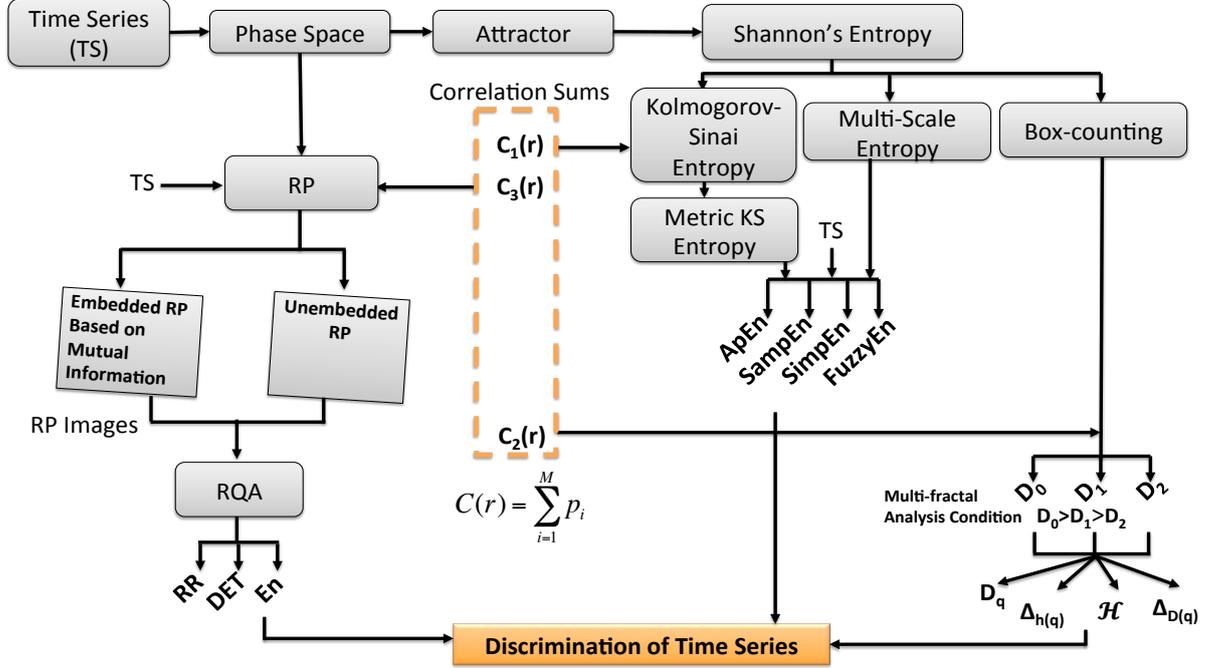


Figure 1.1 – Schematic illustration of the phases of the time series analysis, the origin of the three different complexity analysis techniques and the common point between them.

Intrauterine Growth Restriction (IUGR) inferred from FHR recordings, and to detect intrinsic dynamic transitions which could amend the knowledge of biomedical systems.

1.2 Statement of Problem

The direct computation of correlations and autocorrelations from dynamic system's time series were insufficient and ambiguous. Fourier transform was insufficient to distinguish chaos from noise [Trauth et al., 2010]. This follows that the great challenge is to have a deep look on the chaotic time series and extract its complexity in order to analyze it vividly and predict future states. Until now, there is no optimal and robust tool to analyse chaotic signals, as well as not all the aspects of recurrence plots and entropy parameters are analysed to detect anomalies. Based on what preceded, time and frequency analysis techniques in hands, the comparison of each recording with either a control provided by hospitals and current programmed monitors are not sufficient [Trauth et al., 2010, Yang et al., 2011]. Furthermore, rare are the cases were clinicians can make up their

diagnostic or therapeutic decisions in medicine when the signal is chaotic and random, and rare are the cases where the classical parameters such as the acceleration can achieve the discrimination of medical states of patients [Voicu and Girault, 2012a], here, fetuses.

As FHR signals are nonstationary and nonlinear in nature, many cutting-edge studies have used nonlinear processing tools in order to extract the complexity from such signals [Oudjemia et al., 2013, Zaylaa et al., 2014a, Zaylaa et al., 2014b]. Consequently, complexity analysis has become popular for FHR signals. Moreover, as preterm deliveries for instance, that is giving birth before the 37th week of pregnancy, remains a major problem in obstetrics, and as it has been reported that children born before the third trimester present a high risk of mortality as well as health and development hardships [Goldenberg et al., 2008], recent studies were keen on applying nonlinear techniques to FHRs.

For instance, Dima et al. have used frequent classification techniques by introducing both linear and nonlinear features (Lyapunov exponent, SampEn and Variance Entropy) to optimize the selection of features that could lead to the best classification accuracy of uterine signals [Alamedine et al., 2013]. The latter study was tailored by the detection of uterine contractions and sensing its resulting electrohysterogram (EHG).

Regardless of the sensed biomedical signal, clinicians can merely extract few features such as the beats-per-minute (bpm), acceleration in addition to heart rate variability (HRV) for diagnosing the fetus heart. They proposed a common criteria to decide when the fetus is distressed and when it is in a healthy state. Clinicians consider that the normal FHR amplitude level ranges between 110-160 bpm [Voicu et al., 2014].

The study of FHR time series could be further accomplished by a model capable of displaying the chaotic signals and extracting the complexity parameters out of these signals. The model includes RQA, determinism (DET), recurrence rate (RR) and other invariant parameters.

From a computational prospect, analysis of biomedical time series could lead to the following question: why are we keen on finding new invariant parameters or improving existing invariants ? From a medical prospect, analysis of biomedical time series could lead to the following questions: what could be the result if electrocardiograms (ECGs), electroencephalograms (EEGs) and Doppler US FHR machines are meant to classify habitual signals from strange ones ? What if strange chaotic signals were obtained ? Is it sufficient to observe temporal representation of the signals on a monitor to write a medical

report ? Are available monitors capable of analysing all chaotic signals ? Particularly, to which extent are the current recurrence tests and their corresponding plots accurate ?

The former and latter questions caused the problem to be more critical and made us keen on treating it. We hypothesize that a lot of information hidden in such complex time series can be extracted by entropy, multi-fractal and recurrence analyses.

1.3 Objectives

This dissertation presents both theoretical and numerical investigations of chaotic biomedical systems, particularly fetal heart, by means of Entropy and Multi-fractal Analysis and Recurrence Plots. Our aim is to improve complexity techniques available in the literature and propose new paradigm, complexity parameters and novel clean techniques to improve the diagnosis and detect the fetus medical state. Generally, we hypothesized that improved complexity parameters and novel complexity invariants improve the diagnosis of fetal Intrauterine Growth Restriction (IUGR) due to hypoxia.

In this dissertation we propose new methods capable of extracting and displaying the complexity hidden in chaotic signals. We have introduced the theory of our techniques, applied them on mathematical systems which model and resemble biomedical systems, and finally applied the novel techniques on real fetus heart Doppler US signals. These techniques were both programmed and validated on the fractional Brownian Motion (fBm), Lorenz System in the chaotic regime and Logistic Map. Our novel codes were written to build up cutting-edge portable machines capable of analysing chaotic biomedical signals in the field of medicine.

The three major complexity analysis techniques realized to study the random and biomedical signals throughout this dissertation were demonstrated in Fig. 1.2. The first technique is a direct quantification analysis in which we have developed two novel entropy parameters (methods). The second technique is a geometric quantification analysis and the third is a geometric qualitative technique. In this dissertation we also present new extensions and definitions of recurrence plots and their application especially on chaotic FHR signals.

1.4 Structure

The dissertation is divided into three major parts, part-I general introduction and materials used, part-II complexity analysis and application to biomedical signals and part-

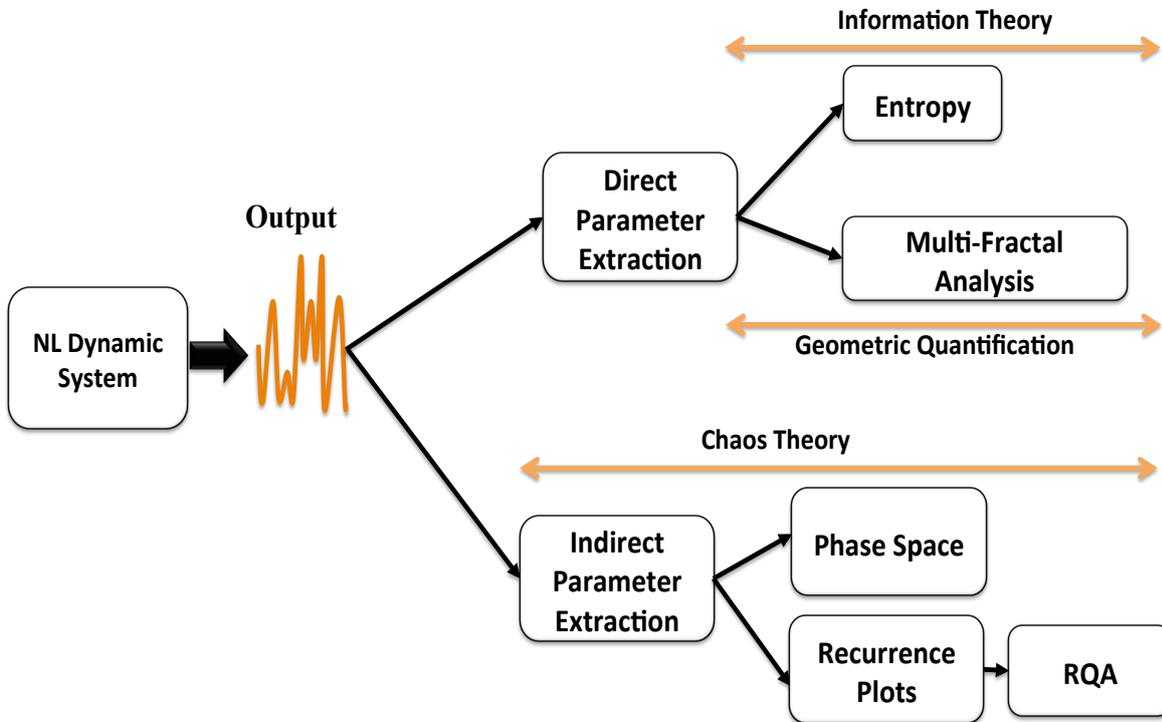


Figure 1.2 – Schematic illustration of the three different complexity analysis techniques.

III summary of the dissertation, conclusion and perspectives. A general introduction is provided in part-I including a methodical overview of existing works (chapter 1), and the materials used are introduced in chapter 2. The three theoretical and numerical contributions to the complexity analysis techniques are introduced in part-II. Chapter 3 in Part-II constitutes the entropy, chapter 4 the multi-fractal analysis and chapter 5 recurrence analysis. Each of the previous chapters in part-II was made up of a theoretical part followed by numerical applications on both simulated and real systems for classification, detection and diagnosis purposes. Finally, a general comparison of complexity analysis techniques, the conclusion and perspectives are given in chapter 6 in part-III to sum up the whole dissertation results and open the door for future works and inventions.

The existing entropy pseudocodes are provided in Appendix A. The pseudocodes of recurrence plots developed are provided in Appendix B, my Ph.D. activities in Appendix C, the List of Publications in Appendix D and the first page of the three articles that were written are supported in Appendix E.

Chapter 2

MATERIALS

The fractional Brownian motion, Lorenz and logistic times series are good archetype of biomedical signals [Marwan et al., 2002, Marwan et al., 2007, Oudjemia et al., 2013]. These well-known systems were simulated and used to validate the developed complexity parameters and Unbiased Recurrence Plots.

2.1 Fractional Brownian Motion

To illustrate fBms, three different fBm time series are depicted for different \mathcal{H} parameters (called Hurst index) [Mandelbrot and Van Ness, 1968] in Fig. 2.1. Fig. 2.1 represents (a) a negatively correlated fractional Gaussian noise as $\mathcal{H} < 0.5$, (b) the Brownian motion as $\mathcal{H} = 0.5$ and (c) a positively correlated fractional Gaussian noise as $\mathcal{H} > 0.5$.

2.2 Lorenz System

Lorenz System originally founded in (1963) is given below:

$$\left\{ \begin{array}{l} \frac{dx}{dt} = a(y - x) \\ \frac{dy}{dt} = \rho x - y - xz \\ \frac{dz}{dt} = xy - cz \\ x(0) = 8, y(0) = 9, z(0) = 25 \end{array} \right. \left\{ \begin{array}{l} a = 10 \\ \rho = 28 \\ c = 8/3 \end{array} \right. \quad (2.1)$$

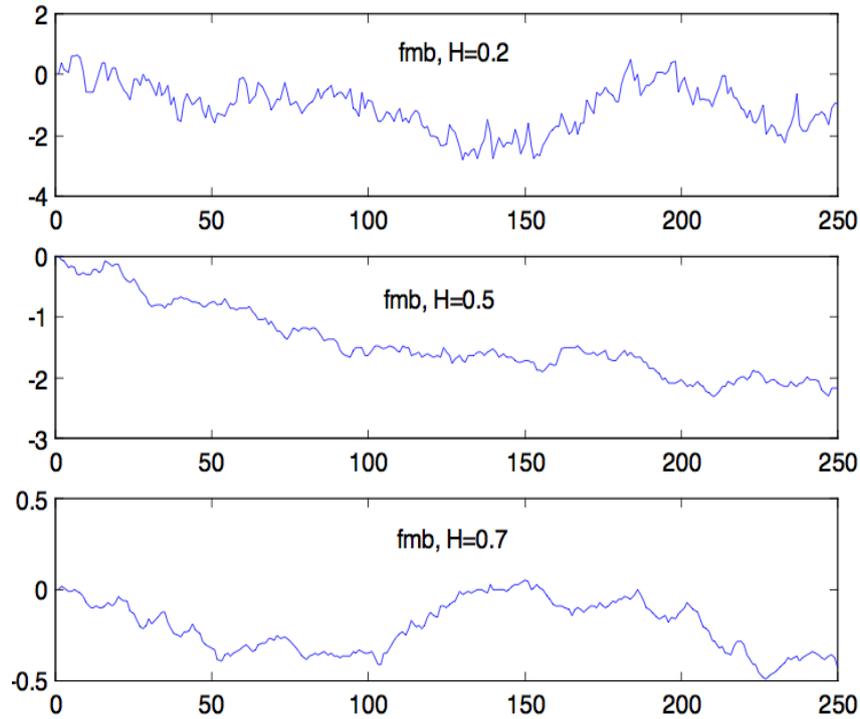


Figure 2.1 – Fractional Brownian Motion. (top) A negatively correlated fractional Gaussian noise. (center) Brownian motion. (bottom) Positively correlated fractional Gaussian noise.

where a , ρ and c are constants to be set. In order to achieve chaotic Lorenz system the parameters should be chosen as shown in Eq. 2.1. The initial conditions were chosen as $x(0) = 8$, $y(0) = 9$, $z(0) = 25$.

This system was used to validate the developed high order entropy parameters prior to their application to FHRs and it served as a chaotic system.

2.3 Logistic Map

In order to mimic dynamic biomedical systems, the five RP techniques used in chapter 5 were applied to the simulated nonlinear biologically inspired system, known as the logistic map [Trulla et al., 1996]. This type of system is characterized by periodic and chaotic dynamic regimes, it was defined as:

2.4. REAL FETAL HEART RATE TIME SERIES

$$x_{\gamma+1} = b * x_{\gamma} * (1 - x_{\gamma}). \quad (2.2)$$

where b stands for the control parameter of the logistic map and γ the iteration number.

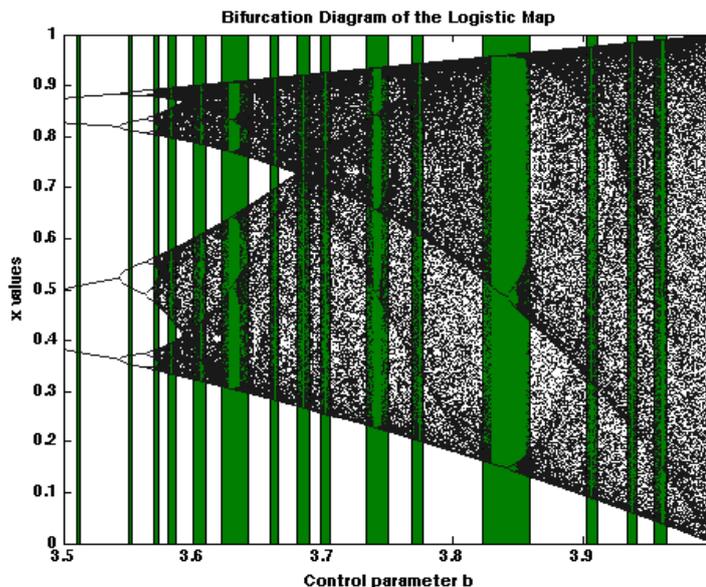


Figure 2.2 – The bifurcation diagram and different dynamic regions of the Logistic Map.

Fig. 2.2 represents the famous bifurcation diagram of the logistic map for $b \in [3.5, 4]$. The different dynamic regimes were highlighted in a green color. The idea of using such type of system to model fetal heart rates was inspired from the works of N.Marwan et al. who made use of the logistic map to model adult heart rates [Marwan et al., 2002, Marwan et al., 2007].

2.4 Real Fetal Heart Rate Time Series

2.4.1 Multichannel Doppler Ultrasound Device and Fetal Heart Rates

Multi-channel Doppler Ultrasound (MCDUS) device was the system used for monitoring FHRs. MCDUS (by ALTHAÏS Technologies, University of Tours in France) is a portable

device shown in Fig. 2.3. MCDUS in Fig. 2.3 was connected to a personal computer (PC). MCDUS comprised a Doppler acquisition board and three groups of four transducers (12 probes \times 5 gates). The operating functions of the acquisition board (1-6) were explained in details in [Rouvre, 2006]. This system which transmits ultrasound waves of 2.25 MHz for an acoustic power limited to 1 mW/cm² [Voicu and Girault, 2012b, Voicu et al., 2014], was developed to measure both the FHR and fetal movements (pseudo-breathing and limb movements).

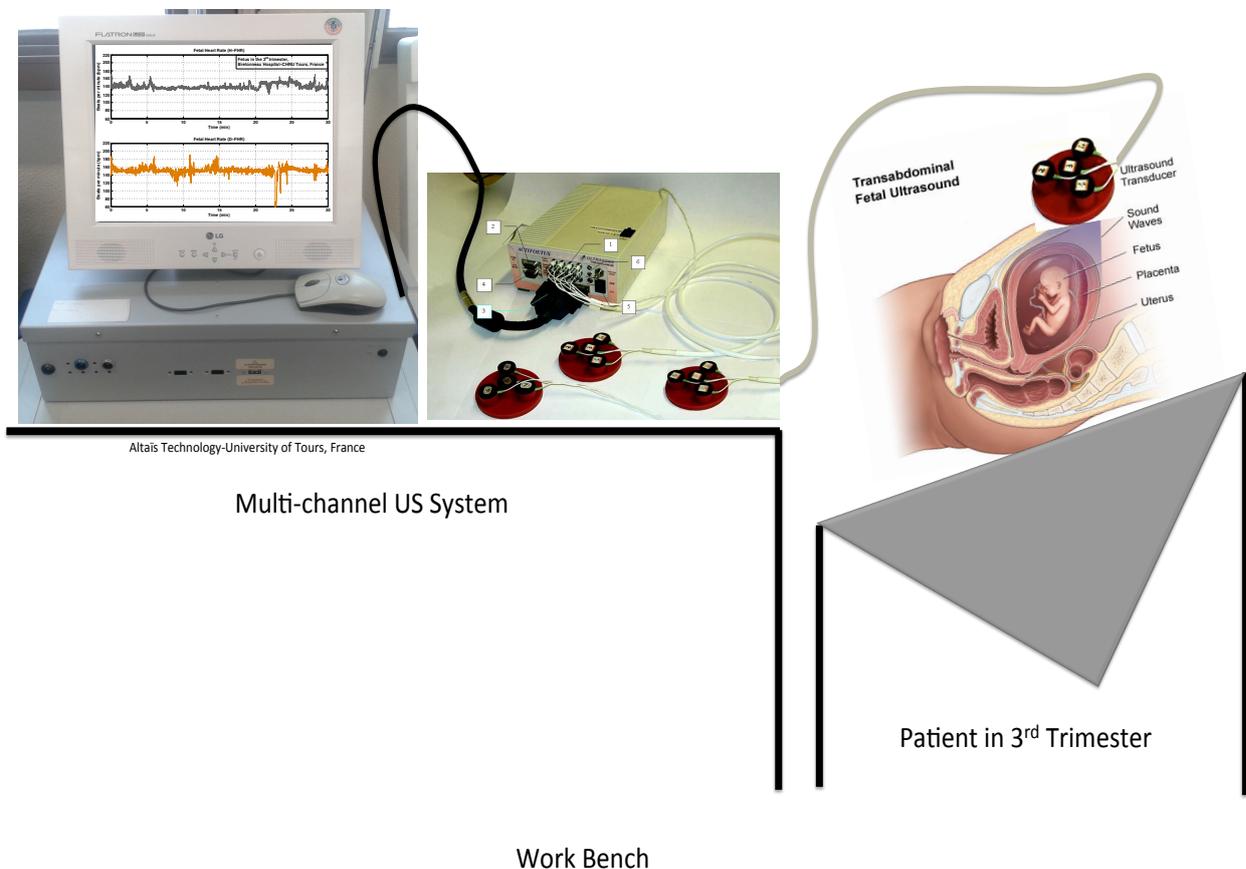


Figure 2.3 – Schematic illustration of the application setup from Doppler Ultrasound Fetal Heart Rate detection to Fetal Heart Rate extraction.

Two FHR recordings are displayed in Fig. 2.4. Fig. 2.4 represents (a) a healthy fetal heart rate (H-FHR) shown in orange and (b) a distressed fetal heart rate (D-FHR) shown in gray. H-FHR ranges from 110 to 160 beats per minute during the third trimester of pregnancy [Voicu et al., 2014]. These FHR recordings were then windowed and 3-minute sub-FHRs were produced.

2.4. REAL FETAL HEART RATE TIME SERIES

The fetal distress was due to hypoxia i.e. placental insufficiency which causes the reduction of oxygenated blood from the uterus to the placenta, thereby reducing the amount of oxygenated blood reaching the fetus [Maslova et al., 2003, Habek et al., 2001]. The interest of studying such medical problem was to reduce the death of fetuses and decide for caesarean section prior to Fetal death. As the data base was provided by Bretonneau Hospital (CHRU) in Tours, the obstetric department was able to decide that fetuses were suffering from hypoxia by chemically analyses. However according to our knowledge, this medical problem has not been fully understood and detected by the direct analysis of FHR recordings.

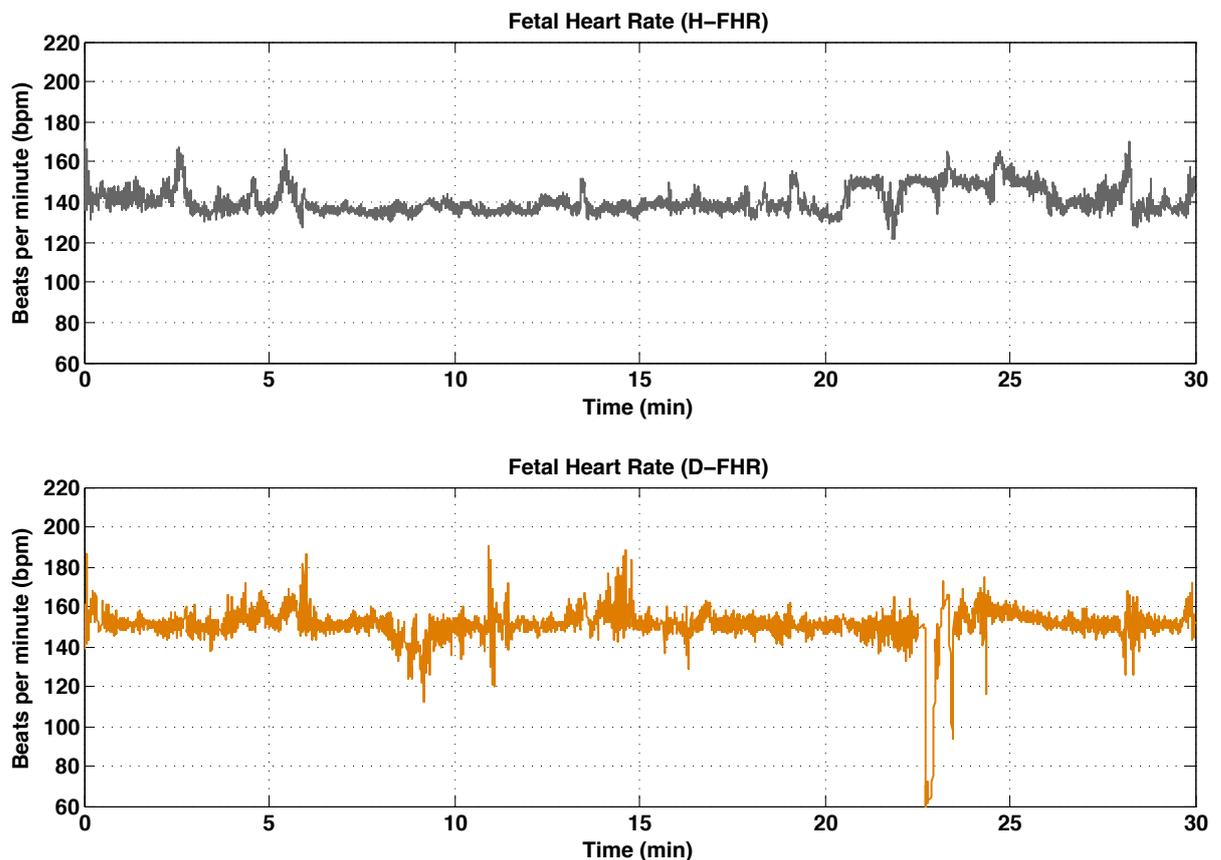


Figure 2.4 – Fetal Heart Rate (FHR) recordings from two different fetal groups for 30-minutes at Bretonneau Hospital in Tours (France). (a) A Healthy Fetal Heart Rate (H-FHR) recording and (b) a Distressed Fetal Heart Rate (D-FHR) recording.

2.4.2 Case Study and Protocol

One group of Doppler system electrodes was dedicated to explore the fetus heart. The transducers exploring the fetus heart were non-focused and mono-element, they were also circular with a diameter of 13.5 mm each. Geometrically, each electrode comprised transducers located at both the top of an equilateral triangle of side 40.7 mm and at its center.

Transducers were placed on the mother's abdomen and were adjusted to transmit a sinusoidal pulse at 2.25 MHz with a pulse repetition frequency (PRF) of 1 KHz. US was transmitted to the fetus heart through the mother's abdomen. During its transmission, US interacted with the moving structures of the heart such as the walls and valves. This interaction has modified the transmitted frequency of the pulse. The frequency content of the backscattered signal carried the heart signature when recorded from a volume containing the fetus heart. In order to divide the volume, the backscattered signal was recorded from five different depths.

The echo was converted into an electrical signal and amplified to compensate for the attenuation of 1 dB/cm/MHz. The signal was then demodulated [Jensen, 1996] and digitized. The digital output of the converter represented the digital Doppler signal. Once the digital signal was obtained, Doppler signals were processed numerically to find FHR [Voicu et al., 2009, Voicu et al., 2010]. FHR was computed every 250 ms, the interval corresponding to the average of consecutive heart beat intervals.

Our database included 80 FHR recordings collected by our MCDUS from 80 pregnant women (40 healthy recordings and 40 distressed ones). The women were in their third trimester (fetus age ranged from 25 to 39 weeks). The consent of each patient was obtained and the study was approved by the ethical committee of the Clinical Investigation Center for Innovative Technology of Tours (CIC-IT 806 CHRU of Tours).

Part II

COMPLEXITY ANALYSIS AND APPLICATION TO BIOMEDICAL SIGNALS

Chapter 3

ENTROPY ANALYSIS

「 *As words uttered are temporal events, it is crucial to balance these events and distinguish whether they are regular or irregular prior to uttering them* 」
Amira Zaylaa.

3.1 Introduction

Entropy which is the measure of the degree of irregularity and disorder of analyzed time series is the core of this chapter. Entropy was realised by evaluating the probability of finding m -similar patterns, and various Entropy descriptors were derived from the latter concept. It is a nonlinear processing tool originating from the Information theory and characterized by its prompt application to time series [Ash, 1990]. Serving as an algebraic quantification parameter, entropy was used in this dissertation to characterize the complexity of the biomedical-like random signals such as the fractional Brownian Motion (fBm) and later Fetal Heart Rates (FHR). Approximate Entropy (ApEn), Sample Entropy (SampEn), multi-scale Entropy, Similarity Entropy (SimEn) and the Fuzzy Entropy (FuzzyEn) were involved as alternative techniques. However, a high order entropy in which two novel entropy parameters were developed, Maximum Entropy and n-order Entropy was used to improve the discrimination of biomedical signals in general and Fetal heart Rates in particular. In addition, the problem of setting the pattern value for entropy computation has been optimally carried out.

3.2 Major Existing Entropy

Entropy, originally developed by Shannon [Shannon, 1948], measures the randomness of data and is given as:

$$\mathbf{En} = - \sum_{i=1}^M \Omega_i \log \Omega_i, \quad (3.1)$$

where M is the length of the time series and Ω_i is the probability of finding templates that constitute the dictionary of events. This entropy was then generalized by Rényi [Rényi, 1961]. Shannon's entropy, which was applied to analyze chaotic time series in the phase space, and the formula given by Eq. 3.1 were not immediately useful and depended on two factors. First, Shannon's entropy depended on the box size chosen to compute the probability through, and second, it depended on the intrinsic properties of the attractor [Cross, 2000]. To make the formula useful scientists proposed i) scaling the entropy by reducing the box size and/or ii) using the algorithm of Kolmogorov and Sanai and thus KS entropy [Kolmogorov, 1959, Sinai, 1959].

Kolmogorov and Sanai thought that by KSEn profound knowledge of the complexity of systems can be acquired. Then another advanced forms based on i) entropy showed up such as ApEn, SamEn, SimEn, FuzzyEn and Permutation En [Bandt and Pompe, 2002, Riedl et al., 2013].

GeoEn has been proposed by Pham [Pham, 2010] and the tolerance r was set to a value inferred from the computation of the difference in the range of the distorted signal (variance of the signal) and its advanced version. However, in our work the value of r was set at $10\% \times SD(\text{signal})$. GeoEn refers to the metric distance between distorted signals, additional conditions has been applied on this distance computation. It employs another concept of computation involving the semi-variance of semi variogram model.

To take the full advantage of both SimEn and FuzzyEn, the maximum and n-order ideas were introduced. We hypothesized that the two novel entropy descriptors enhance the differentiation between distressed and healthy fetuses, i.e. healthy fetuses and fetuses subjected to severe Intrauterine Growth Restriction (IUGR). ApEn, SampEn, SimEn and FuzzyEn also used by Xie et al. [Xie et al., 2008] were computed as alternative methods.

3.2.1 Approximate Entropy

ApEn is the quantification of regularity in data. A deterministic signal is usually composed of few patterns whose occurrence is very high, thus its entropy is low. A random signal is composed of a huge number of patterns whose occurrence is low, therefore its entropy is very high. A chaotic signal is composed of patterns whose occurrence is medium, hence its entropy is medium.

It is applied to relatively short and noisy data [Behnia et al., 2008], it takes into account the length of the signal and self-matches and it requires a heavy implementation.

3.2. MAJOR EXISTING ENTROPY

A time series x_n of length M is divided into a set of m -length vectors $u_m(i)$. Then the number of vectors $u_m(i)$ and $u_m(j)$, close to each other, in distance $\Gamma[u_m(i), u_m(j)] \leq r$, is expressed by the number $P_i^m(r)$. This number is used to calculate the probability of vectors being close $C_i^m(r) = \sum P(r, \Gamma_{i,j}^m) / (M - m + 1)$ [Pincus, 1991, Spilka et al., 2012]. The distance $\Gamma(i, j, m)$ between the two m -patterns $\mathbf{X}(i, m)$ and $\mathbf{X}(j, m)$ is defined as follows:

$$\Gamma(i, j, m) = \Gamma(\mathbf{X}(i, m), \mathbf{X}(j, m)) = \max_{k \in (0, m-1)} |x(i+k) - x(j+k)|. \quad (3.2)$$

Define the following state function:

$$\phi^{(m)}(r) = -\frac{1}{M - (m - 1)} \sum_{i=1}^{M-(m-1)} \log C_i^m(r), \quad (3.3)$$

and the consecutive state function:

$$\phi^{(m+1)}(r) = -\frac{1}{M - m} \sum_{i=1}^{M-m} \log C_i^{m+1}(r). \quad (3.4)$$

$C_i^{(m)}(r)$ accounts whether a vector $X(i, m)$ and another vector $X(j, m)$ exist within a tolerance r . Consequently ApEn is written as:

$$ApEn(m, r) = \phi^{(m+1)}(r) - \phi^{(m)}(r). \quad (3.5)$$

Fig. 3.1 represents a schematic illustration of Entropy computation for a pattern of 2 samples. ApEn measures the sum of the probability that patterns existing within r , such as, $[u(4), u(5)]$ and $[u(16), u(17)]$ match. Note that ApEn considers the template self-match, i.e. takes into account that $[u(4), u(5)]$ matches itself. The pseudocode of this entropy is found in Appendix A.

3.2.2 Sample and Multi-Scale Sample Entropy

SampEn is designed to reduce the bias of ApEn and has a closer agreement with theory for dataset with known probabilities. SampEn overcomes shortcomings of ApEn, i.e. self-matches calculated by ApEn, and it is computed to quantify the complexity of short heart rate time series [Richman and Moorman, 2000, Javorka et al., 2008].

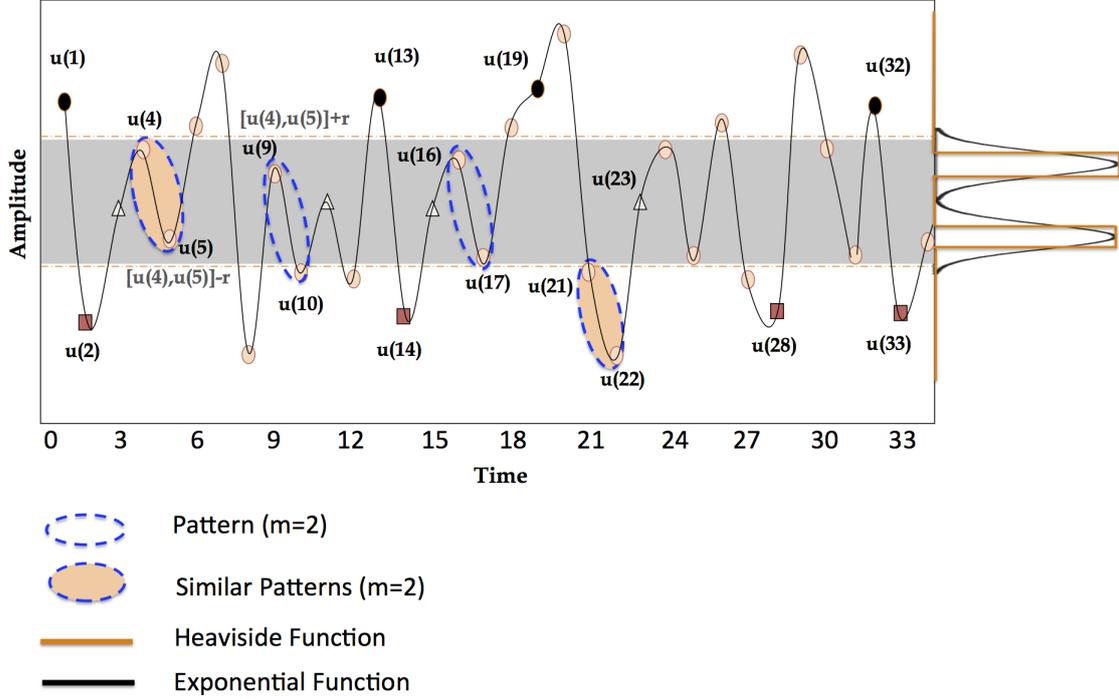


Figure 3.1 – Schematic Illustration of Entropy computation from a chaotic time series using a pattern of $m = 2$.

SampEn is the negative logarithm of the probability that one template m finds a match at $m + 1$ [Spilka et al., 2012] simultaneously within tolerance r . SampEn does not depend on the length of the signal and exhibit a trouble-free implementation. Fig. 3.1 demonstrates, for instance, how SampEn considers that both $[u(4), u(5)]$ and $[u(16), u(17)]$, and both $[u(3), u(4), u(5)]$ and $[u(15), u(16), u(17)]$ are matching for a choice of $m = 2$ and $m = 3$, respectively. However, it does not take into account that $[u(4), u(5)]$ or $[u(3), u(4), u(5)]$ templates match themselves.

The multi-scale SampEn possesses the same formulation of SampEn except that it is applied on scales produced from the original time series. It requires generating scales of a signal x by applying the following equation:

$$y_j^s = \frac{1}{s} \sum_{k=(j-1)s+1}^{js} x_k, \quad (3.6)$$

where s stands for the scale, j is the index of the sample of the scaled time series and y_j^s the scaled templates of x [Baumert et al., 2012, Costa et al., 2005], the first two scales

3.2. MAJOR EXISTING ENTROPY

are illustrated in Fig. 3.2.

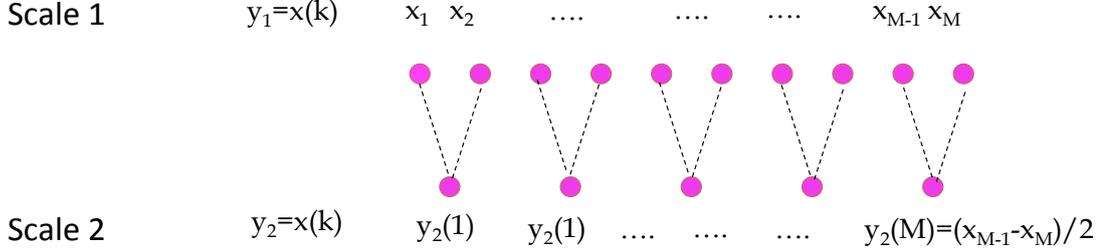


Figure 3.2 – Schematic Illustration of an example of coarse graining for $s = \{1, 2\}$ and $k=\{1, \dots, M\}$.

Both SampEn and ApEn were used to measure the disorder of the heart rate time series [Richman and Moorman, 2000]. Moreover, multi-scale SampEn was recently applied to FHRs to detect acidemia [Costa et al., 2014]. The detailed pseudocode of this entropy type is found in Appendix A. The formula of SampEn is the same as Eq. 3.5, however:

$$\phi^{(m)}(r) = -\log \left(\frac{1}{M-m} \sum_{i=1}^{M-m} C_r^m(i) \right), \quad (3.7)$$

and

$$\phi^{(m+1)}(r) = -\log \left(\frac{1}{M-m} \sum_{i=1}^{M-m} C_r^{m+1}(i) \right). \quad (3.8)$$

3.2.3 Similarity and Multi-Scale Similarity Entropy

SimEn is based on the computation of centered patterns which are similar within tolerance r [Lake et al., 2002]. Two patterns of m samples were assigned similar after:

- o Subtracting the mean of the patterns (centering process);
- o Testing whether the new samples of the centered patterns were enclosed within a tolerance r or not.

Fig. 3.1 demonstrates how similar patterns with an orange background appear in the time series. SimEn considers both $[u(4), u(5)]$ and $[u(21), u(22)]$ similar even-though they do not exist within the same tolerance r .

Multi-scale SimEn is a modified version of SampEn, however, taking into account patterns at all amplitude levels [Girault et al., 2013]. Therefore, a centered vector sequence $\mathbf{X}(m, i)$ is formed from a time series vector composed of M points:

$$\mathbf{X}(m, i) = \{x(i), x(i + 1), \dots, x(i + m - 1)\} - \bar{\mathbf{X}}(i, m), \quad (3.9)$$

with

$$\bar{\mathbf{X}}(i, m) = \frac{1}{m} \sum_{l=0}^{m-1} x(i + l),$$

$$SimEn = \mathcal{E}_p(1, m) = \phi_p(m + 1) - \phi_p(m). \quad (3.10)$$

where $\phi_p(m)$ is similar to Eqs. 3.7 and 3.8 and $p = \infty$ when the Heaviside function is considered, and the 1 in $\mathcal{E}_p(1, m)$ corresponds to the measure of the difference between two consecutive state functions. The necessary steps used to compute multi-scale SimEn were identical to those used to compute SampEn applied on the scaled time series Eq. 3.6, however, after removing the baseline (i.e. subtracting the mean). The computation of SimEn was based on the work of [Voicu and Girault, 2012b] and its pseudocode is summarized in Appendix A.

3.2.4 Fuzzy Entropy

In contrary to ApEn, SampEn, and SimEn in which the degree of similarity between two vectors is strictly constrained by the Heaviside function depicted on the right of Fig. 3.1 and shown in an orange color, FuzzyEn was constrained by a fuzzy membership function. The fuzzy membership function was set out on the right of Fig. 3.1 and shown in a black color, it bounds the similarity of two vectors and ensures statistical stability [Liu et al., 2013]. Using a Heaviside function, the boundary is rigid: the contributions of all the data points inside it are treated equally, whereas the data points outside it are discarded.

The strict boundary nature causes discontinuity, which may lead to abrupt changes in entropy values when r changes slightly, moreover, it may fail to find a SampEn value if no template match was found using a small tolerance.

On the other hand, a fuzzy membership function imposes a smooth boundary and continuity for different values of tolerance r . All data points, where the membership

3.2. MAJOR EXISTING ENTROPY

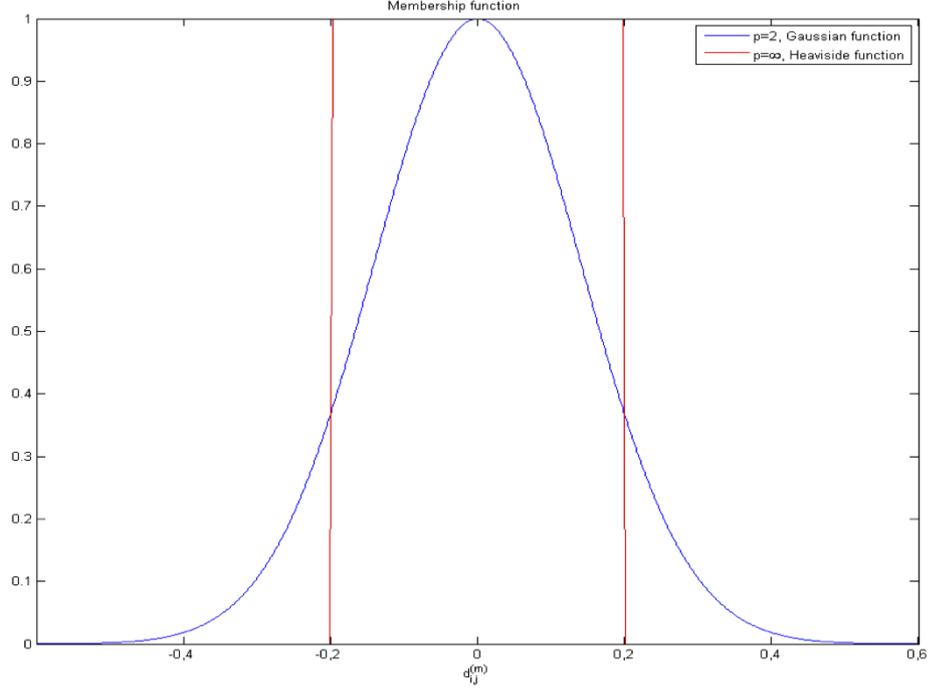


Figure 3.3 – The membership function.

function and the tolerance r are satisfied, are considered to be members of the function. For a vector X_i^m , similarity is indicated by the fuzzy membership function value: the closer the neighbouring vector X_j^m , the more similar X_j^m to X_i^m . Furthermore, the degree of similarity between X_j^m and X_i^m is almost zero when X_j^m is distant from X_i^m . Using a fuzzy membership function, FuzzyEn is continuous and does not change dramatically with a slight change in r . This type of entropy was applied on heart rate variability and took high values for healthy cases compared to distressed cases [Liu et al., 2013].

The degree of similarity $P_p(i, j, m)$ between the $\mathbf{X}(i, m)$ and $\mathbf{X}(j, m)$ vectors within a fixed tolerance r and Eq. 3.10 can be calculated through an exponential fuzzy function as follows:

$$P_p(i, m, r) = e^{-\left(\frac{F(i, j, m)}{r}\right)^p}. \quad (3.11)$$

When the family of exponential functions is selected as a membership function, setting the parameter p becomes a necessity. Fig. 3.3 represents the membership function when $p = \infty$ and $p = 2$. For $p = \infty$ the membership function is said to be the Heaviside function, and for $p = 2$ the membership function is equivalent to the Gaussian function.

The pseudocode of FuzzyEn is found in Appendix A.

3.3 High Order Developed Entropy

One of the most promising ways for detecting different dynamical states lies in quantifying the complexity of time series. Among the possible complexity descriptors, entropy [Pincus, 1991, Ferrario et al., 2006, Chen et al., 2009, Liu et al., 2011, Liu et al., 2013] indicators are undoubtedly effective.

However, entropy descriptors are highly dependent on the setting parameters (m, r) [Restrepo et al., 2014], and the choice of these parameters is critical, especially for moderately sized signal lengths [Boskovic et al., 2012]. Chon et al. [Chon et al., 2009] claimed that entropy descriptors such as ApEn and SampEn are not accurate in assessing signal complexity using the recommended values in the literature [Pincus, 1991].

Though it has been suggested by [Pincus and Keefe, 1992] to set the tolerance r empirically between 0.1 and 0.2 times the standard deviation of the time series, recent studies focusing on improving the detection of transitions [Lu et al., 2008, Restrepo et al., 2014] have shown the interest of setting it at another value. This suggested value, adapted for the detection of transitions, has been proposed to maximize entropy $\mathcal{E}(m, r)$:

$$\begin{cases} \mathcal{E}_* = \mathcal{E}(r^*, m), \\ r^* = \arg \max_r (\mathcal{E}(r, m)), \end{cases}$$

where r is the tolerance and m is the pattern size.

On the other hand, for setting the pattern size m , few studies have focused on finding the optimal value of m , and the only studies published focused on the reconstruction of the original phase space^a [Kennel et al., 1992, Cao, 1997] rather than the detection of transitions. However, the recent study by Restrepo et al. [Restrepo et al., 2014] tested several sizes of pattern m . Basically, the most commonly used values of m were set at 1, 2 or 3 [Pincus and Goldberger, 1994, Ferrario et al., 2009]. These low values of m were proposed because a too high value of m leads to a poor estimation of the entropy $\mathcal{E}(m, r)$ [Restrepo et al., 2014] or to a poor reconstruction of the system dynamics.

^aThe principle of reconstruction consists of finding the minimum embedding dimension d that corresponds to a sudden change in the nearest false neighbors.

3.3. HIGH ORDER DEVELOPED ENTROPY

As entropy descriptors $\mathcal{E}(m, r)$ suffer from setting problems and as there are no satisfactory solutions, the main purpose of this study was to extend the studies recently undertaken by [Lu et al., 2008, Chon et al., 2009, Restrepo et al., 2014] based on the search for parameter settings leading to maximum entropy.

The first challenge to resolving the setting problems was investigating the value of m^* to maximize entropy $\mathcal{E}(m, r)$:

$$\begin{cases} \mathcal{E}^* = \mathcal{E}(r, m^*), \\ m^* = \arg \max_m (\mathcal{E}(r, m)). \end{cases}$$

As this study was devoted only to searching for the optimal value m^* with a constant value r , the latter was fixed at $r = 0.2$ for the remainder of the study. The variable r was therefore deliberately omitted from the subsequent equations.

The study reported in this dissertation showed that examining the role of the pattern size m could provide important insights into quantifying the complexity of time series, thus leading to improved understanding of nonlinear dynamic systems. Furthermore, it was shown that the combined use of m^* and \mathcal{E}^* can provide a more consistent method to distinguish between different dynamics.

In response to the lack of a satisfactory method, the need to find a new transition detector or a discriminator is greatly needed. The second aim of this study was to establish a new paradigm to provide important insights into quantifying the complexity of time series. This new concept, for which, as suggested by [Liu et al., 2013], a membership function was introduced, encompassed the standard definition of entropy descriptors. The general framework of "n-order fuzzy entropy" on which the new paradigm is based depends on the following equation:

$$\mathcal{E}(n, m) = \phi(m + n) - \phi(m) \tag{3.12}$$

where n is the order, and ϕ is the average of the natural logarithm of the probability of finding similar patterns of size m . By setting $n = 1$ in equation 3.12, ApEn is recognized.

The concepts of maximum entropy and n-order entropy were then introduced in subsections 3.3.1 and 3.3.2 to achieve the above aims. The two new paradigms were then evaluated through simulations on real FHR recordings.

3.3.1 Maximum Entropy

The aim of this subsection is to introduce the maximum entropy $\mathcal{E}^* = \mathcal{E}(m^*)$, and demonstrate that it is effective for discrimination purposes. For this reason and for a more general framework, SimEn was used as the alternative entropy since it was recently shown that it is superior to ApEn [Liu et al., 2013]. SimEn was defined by Eq. 3.10.

The entropy in Eq. 3.10 can be calculated through the exponential fuzzy function defined by Eq. 3.11. To guarantee that $\phi_p(m)$ varies between 0 and 1, a normalized version is proposed as follows:

$$\Phi_p(m) = 1 + \frac{\phi_p(m)}{\log(M)}, \quad (3.13)$$

where M is the length of the time series.

Finally, the 1-order entropy reduces to:

$$\mathcal{E}_p(1, m) = \Phi_p(m + 1) - \Phi_p(m), \quad (3.14)$$

The optimization process described by the search for m^* maximizing the entropy $\mathcal{E}_p(1, m)$ and thus the maximum entropy \mathcal{E}_p^* leads to the following equations:

$$\begin{cases} m^* = \arg \max_m (\mathcal{E}_p(1, m)), \\ \mathcal{E}_p^* = \mathcal{E}_p(1, m^*). \end{cases}$$

with $p = \infty$ for non-fuzzy estimations and $p = 2$ for fuzzy estimations.

Note that the search for the optimal set of parameters (r^*, m^*) to optimize the entropy $\mathcal{E}_*^* = \mathcal{E}(r^*, m^*)$ is the subject of a future study.

3.3.2 N-Order Entropy

The monotonicity feature of $\Phi_p(m)$ was not used profitably in the maximum entropy, and this was the starting point of N-order entropy. To benefit from the slope-dependence of $\Phi_p(\mathcal{H})$, the discrimination function is defined by:

$$\Psi_{\mathcal{E}_p}(n) = |\Phi_p^{(1)}(1 + n) - \Phi_p^{(2)}(1 + n)|, \quad (3.15)$$

3.3. HIGH ORDER DEVELOPED ENTROPY

and is based on $\Phi_p^{(i)}(n+1)$, where $i = 1$ corresponds to fBm of $\mathcal{H} = 0.07$ and $i = 2$ corresponds to fBm of $\mathcal{H} = 0.3$, and $i = 1$ corresponds to H-FHR and $i = 2$ corresponds to D-FHR. By slightly modifying the definition of $\Psi_{\mathcal{E}_p}(n)$ as follows:

$$\Psi_{\mathcal{E}_p}(n) = \left| \left(\Phi_p^{(1)}(1+n) - \Phi_p^{(1)}(1) \right) - \left(\Phi_p^{(2)}(1+n) - \Phi_p^{(2)}(1) \right) \right|, \quad (3.16)$$

and by setting $\Phi_p^{(1)}(1) = \Phi_p^{(2)}(1)$ (made possible by the normalisation outlined in Eq. 3.13) the definition of the n-order SimEn previously presented in Eq. 3.10 is recognized with $m = 1$: $\mathcal{E}_p^{(i)}(n, 1) = \Phi_p^{(i)}(1+n) - \Phi_p^{(i)}(1)$, leading to

$$\Psi_{\mathcal{E}_p}(n) = \Psi_{\mathcal{E}_p}(n, 1) = \left| \mathcal{E}_p^{(1)}(1+n, 1) - \mathcal{E}_p^{(2)}(1+n, 1) \right|. \quad (3.17)$$

Although it is probable that the following equation $\Phi_p^{(1)}(m) = \Phi_p^{(2)}(m)$ is not always guaranteed, Eq. 3.17 can be generalized $\forall m$ as follows:

$$\Psi_{\mathcal{E}_p}(n, m) = \left| \mathcal{E}_p^{(1)}(n, m) - \mathcal{E}_p^{(2)}(n, m) \right|, \quad (3.18)$$

based on the n-order entropy $\mathcal{E}_p^{(i)}(n, m)$ is defined by:

$$\mathcal{E}_p^{(i)}(n, m) = \Phi_p^{(i)}(m+n) - \Phi_p^{(i)}(m), \quad (3.19)$$

where $p = \infty$ corresponded to non-fuzzy estimations and $p = 2$ to fuzzy estimations.

All possible values of $\mathcal{E}_p^{(i)}(n, m)$ are gathered in the symmetrical matrix $\mathcal{M}_p^{(i)}$ defined by:

$$\mathcal{M}_p^{(i)} = \begin{pmatrix} 0 & \mathcal{E}_p^{(i)}(1, 1) & \mathcal{E}_p^{(i)}(2, 1) & \mathcal{E}_p^{(i)}(3, 1) & \dots \\ \mathcal{E}_p^{(i)}(1, 1) & 0 & \mathcal{E}_p^{(i)}(1, 2) & \mathcal{E}_p^{(i)}(2, 2) & \dots \\ \mathcal{E}_p^{(i)}(2, 1) & \mathcal{E}_p^{(i)}(1, 2) & 0 & \mathcal{E}_p^{(i)}(1, 2) & \dots \\ \mathcal{E}_p^{(i)}(3, 1) & \mathcal{E}_p^{(i)}(2, 2) & \mathcal{E}_p^{(i)}(1, 3) & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} \quad (3.20)$$

where $\mathcal{M}_p^{(i)}(k, l) = |\Phi_p^{(i)}(k) - \Phi_p^{(i)}(l)|$. This matrix is symmetrical since $|\Phi_p^{(i)}(k) - \Phi_p^{(i)}(l)|$ is equal to $|\Phi_p^{(i)}(l) - \Phi_p^{(i)}(k)|$. The structure of this matrix is specific since its main diagonal is null ($\mathcal{M}_p^{(i)}(k, k) = 0$). The main diagonal corresponds to the 0-order entropies and the diagonals just above and below the main diagonal correspond to the 1-order entropies shown in blue, 2-order entropies shown in green, 3-order entropies shown in red, etc.

Finally, the discrimination function $\Psi_{\mathcal{M}_p}$ based on the Matrix $\mathcal{M}_p^{(i)}$ was defined by:

$$\Psi_{\mathcal{M}_p} = |\mathcal{M}_p^{(1)} - \mathcal{M}_p^{(2)}|. \quad (3.21)$$

3.3.3 Results of Fractional Brownian Motion and Lorenz System

In entropy analysis, 200 normalized fBMs comprised 1024 samples were simulated for discrimination purposes. 100 fBMs for $\mathcal{H} = 0.07$ and 100 for $\mathcal{H} = 0.3$ were selected since Oudjemia et al. showed that such fBMs could coarsely model fetal heart rate signals for distressed and healthy cases, respectively [Oudjemia et al., 2013].

3.3.3.1 Maximum entropy

After introducing the theoretical aspects of the maximum entropy it was then important to demonstrate the significance of determining the maximum value of entropy \mathcal{E}_p^* through simulations. These simulations were based on fBm and Lorenz time series for i) they are good archetypes of biomedical signals and ii) their degrees of irregularity can easily be varied through the Hurst exponent \mathcal{H} or the Rayleigh number ρ , respectively.

Note that $\Phi_p(m)$ were monotonic decreasing functions and their slopes depended on the Hurst exponent.

In order to show that \mathcal{E}_p^* and m^* could be related to the intrinsic features of fBm, 100 normalized fBMs of unitary energy composed of $M = 1024$ samples were simulated with Hurst exponents ranging from $\mathcal{H} = 0.1$ to 0.9. To guarantee that the maximum entropy was reached, the size of the pattern m was varied from 1 to 150 and $r = 0.2$.

Figs. 3.4 (a) and (c) represent the mean of $\Phi_\infty(m)$ and $\Phi_2(m)$. From these graphs, it can be seen that $\Phi_p(m)$ functions, that quantified the probability of finding patterns of size m were monotonic, decrease from 1 to 0 as m increases. These functions that represented a cumulative effect possess slopes depending on the Hurst exponent: the higher the Hurst exponent, the lower the slope. It can also be seen that $\Phi_\infty(m)$ and $\Phi_2(m)$ decrease as the size of the pattern increases. Therefore, the greater the sizes of patterns, the lower the probability of finding large patterns.

Figs. 3.4 (b) and (d), represent the mean of $\mathcal{E}_\infty(1, m)$ and $\mathcal{E}_2(1, m)$. From these graphs, it can be seen that $\mathcal{E}_p(1, m)$ functions, that represented discrete derivatives of $\Phi_p(m)$ functions (cumulative effect removed), are non-monotonic, reaching a maximum

3.3. HIGH ORDER DEVELOPED ENTROPY

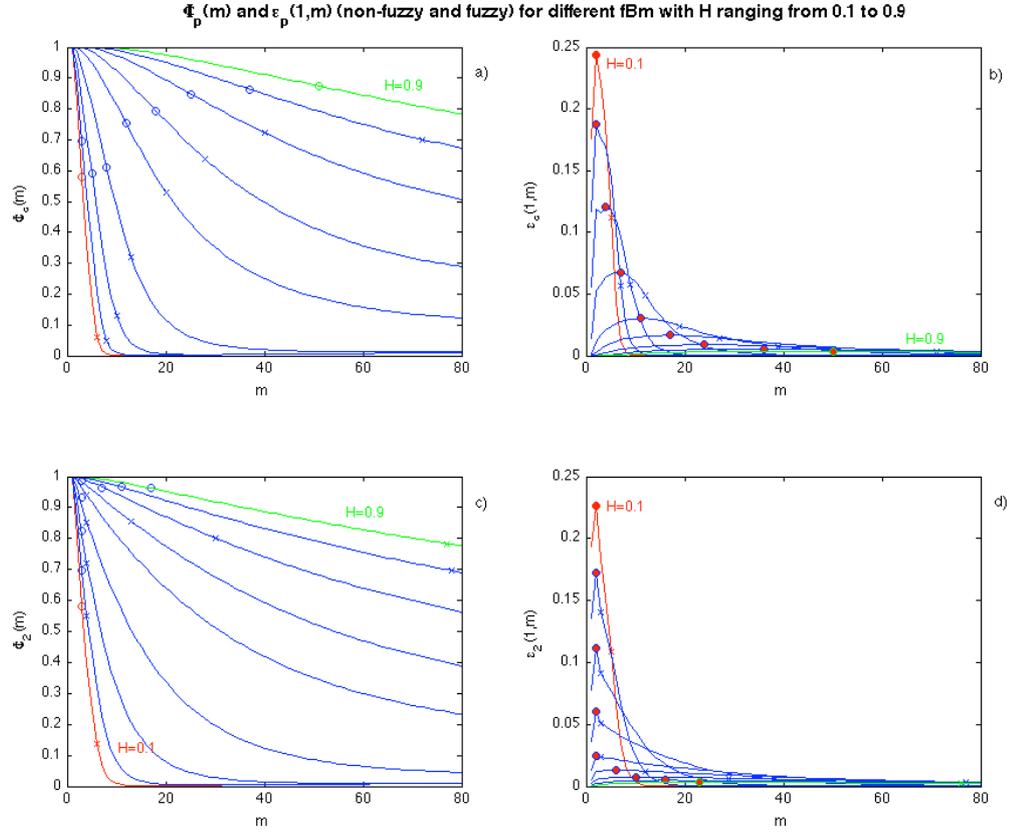


Figure 3.4 – Complexity measurement for different fBm with \mathcal{H} ranging from 0.1 to 0.9 ($r = 0.2$). (a) Non-fuzzy estimations with $\Phi_\infty(m)$. (b) Non-fuzzy estimations with $\mathcal{E}_\infty(1, m)$. Red dots correspond to maxima. (c) Fuzzy estimations with $\Phi_2(m)$. (d) Fuzzy estimations with $\mathcal{E}_2(1, m)$. Red dots correspond to maxima.

$\mathcal{E}_p(1, m^*)$ corresponding to m^* that depended on the value of the Hurst exponent. These curves start from 0 and rise to a maximum and then decrease as \mathcal{H} increases. As fBms are self-affine time series, their autocorrelations are decreasing functions and their correlation length depends on the Hurst exponent value: the lower the Hurst exponent value, the shorter the correlation length and the lower the pattern size m . Figs. 3.4 (b) and (d) show m values that are particularly interesting, the most visible are those maximizing the 1-order entropy. These m values, which are dependant on the Hurst exponent, represent points of inflection in the $\Phi_p(m)$ functions and are highlighted in the $\mathcal{E}_p(1, m)$ functions representing the discrete derivatives of $\Phi_p(m)$. As points of inflection can still be seen in the 1-order entropy $\mathcal{E}_p(1, m)$, then calculating high order derivatives of the functions $\Phi_p(m)$ should provide additional insight in the study of fBm. This will be developed in the next section.

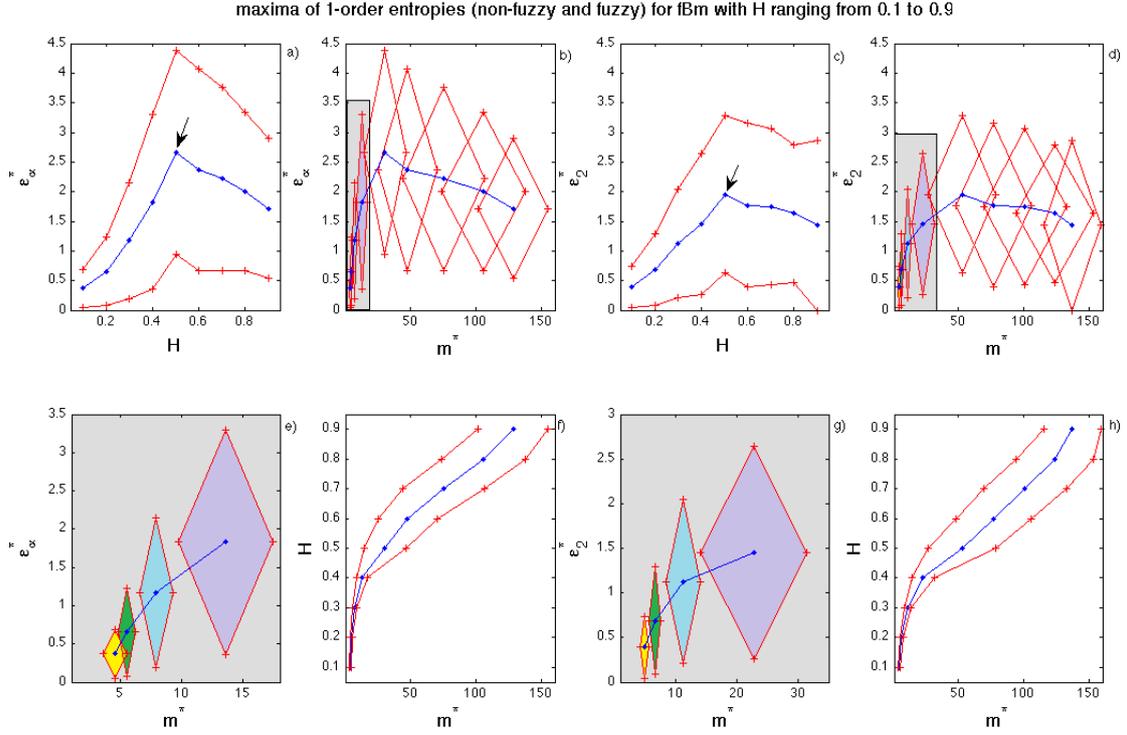


Figure 3.5 – 1-order entropies (non-fuzzy and fuzzy) of 100 fBms with \mathcal{H} ranging from 0.1 to 0.9 ($r = 0.2$). (a) $C_{\infty}^*(\mathcal{H})$. Dots correspond to the mean estimation and red lines correspond to the standard deviation. Arrows represent a limit between anti-correlation and correlation. (b) $C_{\infty}^*(1, m^*)$. Lozenges represent the standard deviation of m^* and C_{∞}^* . (c) $C_2^*(\mathcal{H})$. Dots correspond to the mean estimation and the red lines correspond to the standard deviation. Arrows represent a limit between anti-correlation and correlation. (d) Fuzzy estimation with $C_2^*(1, m^*)$. Lozenges represent the standard deviation of m_2^* and C_2^* . (e) Zoom of (b). From the zoom it is easier to discriminate between high complexity and low complexity. (f) $\mathcal{H}(m^*)$. (g) Zoom of (d). From the zoom it is easier to discriminate high complexity from low complexity. (h) $\mathcal{H}(m^*)$.

Figs. 3.5 (a) and (c) represent $\mathcal{E}_2^*(\mathcal{H})$ and $\mathcal{E}_{\infty}^*(\mathcal{H})$ derived from Figs. 3.4 (b) and (d). Similarly, Figs. 3.5 (f) and (h) represent m^* and $\mathcal{H}(m^*)$ derived from Figs. 3.4 (b) and (d). Moreover, Figs. 3.5 (e) and (g) represent a zoom of Figs. 3.5 (b) and (d) for low values of \mathcal{H} .

Figs. 3.5 (a) and (c) show that $\mathcal{E}_2^*(\mathcal{H})$ and $\mathcal{E}_{\infty}^*(\mathcal{H})$ increase up to a limit defined by $\mathcal{H} = 0.5$ and then decrease. This limit of $\mathcal{H} = 0.5$ seemed to be determined by the negatively correlated ($\mathcal{H} < 0.5$) or positively correlated ($\mathcal{H} > 0.5$) nature of the time series under study. Such behavior was not observed from estimation of $m^*(\mathcal{H})$ that grew continuously as the Hurst exponent increased. By combining the two parameters (m^* , \mathcal{E}_p^*) that seemed related to the intrinsic parameter \mathcal{H} , it was possible to discriminate different fBms. On

3.3. HIGH ORDER DEVELOPED ENTROPY

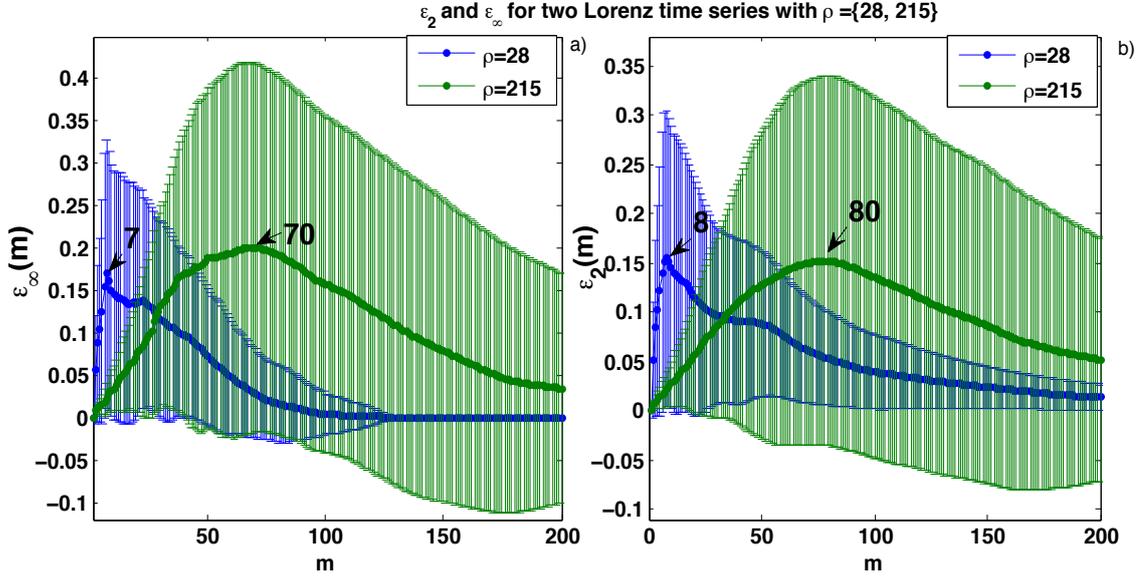


Figure 3.6 – 1-order entropies for Lorenz time series ($r = 0.2$). (a) Non-fuzzy estimations with $\mathcal{E}_\infty^{(1)}$ for chaotic behavior ($\rho = 28$) (represented in blue) and $\mathcal{E}_\infty^{(2)}$ for periodic behavior ($\rho = 215$) (represented in green). (b) Fuzzy estimations with $\mathcal{E}_2^{(1)}$ for chaotic behavior ($\rho = 28$) (represented in blue) and $\mathcal{E}_2^{(2)}$ for periodic behavior ($\rho = 215$) (represented in green).

the other hand, it was easier to make this discrimination for $\mathcal{H} < 0.5$ than for $\mathcal{H} > 0.5$ (see the zoom in Figs. 3.5 (e) and (g)). This outcome augured well for discriminating between healthy and distressed fetuses, since it was recently shown [Oudjemia et al., 2013] that the Hurst exponent of FHR signals was between 0.07 and 0.3.

In order to confirm that both m^* and \mathcal{E}_p^* also perform for other types of complexity, Lorenz^b time series were examined. Two different operating modes of the Lorenz system were examined: a chaotic behavior with $\rho = 28$ and a periodic behavior with $\rho = 215$. Two sets of 100 trials were produced for each value of the Rayleigh parameter ρ , each with 2000 points for different initial conditions randomly chosen from a uniform distribution. In order to avoid the influence of transients, the first 500 points of each trial were discarded. The resulting signals composed of 1500 points were normalized to achieve unitary energy.

Figs. 3.6 (a) and (b) represent $\mathcal{E}_\infty^{(i)}(1, m)$ and $\mathcal{E}_2^{(i)}(1, m)$, with $i = 1$ for $\rho = 28$ and $i = 2$ for $\rho = 215$. The size m ranged from 1 to 200 to guarantee the presence of maximum entropy. From these graphs, it can be seen that there are two different signatures that illustrates the chaotic and non-chaotic natures of the system. For a periodic operating

^bThe Lorenz equations were: $\dot{x} = \sigma(y - x)$, $\dot{y} = -xz + \rho x - y$ and $\dot{z} = xy - \beta z$, where $\sigma = 10$, $\rho = \{28, 215\}$, $\beta = 8/3$ and the step size was selected as 0.01.

mode, the green curves increase gradually up to the maximum and then decrease slowly (see Fig. 3.6). From Fig. 3.6, it can be seen that maxima are around $m^* = 70$ and $m^* = 80$ for $\mathcal{E}_\infty^{(2)}(1, m)$ and $\mathcal{E}_2^{(2)}(1, m)$, respectively. Although high values of m^* were not expected, the latter proved their usefulness for discrimination purposes.

For a chaotic operating mode, the blue curves increase rapidly up to the maximum and then decrease slowly (see Fig. 3.6). From Fig. 3.6, it can be seen that maxima are located at low values of $m^* = 7$ and $m^* = 8$ for $\mathcal{E}_\infty^{(1)}(1, m)$ and $\mathcal{E}_2^{(1)}(1, m)$, respectively. It can be seen that the maximum values $\mathcal{E}_\infty^{(1)*}$, $\mathcal{E}_\infty^{(2)*}$, $\mathcal{E}_2^{(1)*}$, $\mathcal{E}_2^{(2)*}$ were of the same order, whatever the operating mode of the system and whatever the algorithm version (fuzzy or not). It can therefore be asserted that the element that was sensitive to discerning the two operating modes was m^* . Finally, the level of the standard deviation of the entropy estimation (bars in Fig. 3.6) is relatively high in relation to the mean value (dots in Fig. 3.6) since on average it is about 110% for non-fuzzy estimations and 102% for fuzzy estimations. It was clear that the fuzzy estimator whose main feature was to reduce the statistical instability fulfilled its role.

To sum up this subsection, it can be asserted that the maximum amplitude \mathcal{E}_p^* and location m^* of the 1-order entropy were good descriptors, directly related to intrinsic features of the time series under study, and were also powerful indicators to discriminate different levels of complexity while keeping the tolerance fixed at $r = 0.2$. A limitation of the study of the proposed maximum entropy was the maximum value of m chosen ($m=150$ for fBm and $m=200$ for Lorenz system). Although the maximum of entropy and its location augured well for the complexity analysis over the studied range of m , not all features of the functions $\Phi_p(m)$ were used profitably. This latter point is the subject of the next subsection.

3.3.3.2 N-order entropy

The feasibility of this new concept was validated on fBm by the simulation results reported in Figs. 3.7 and 3.8 using the materials in chapter 2. For discrimination purposes, the size of m varied from 1 to 30 to guarantee that the maximum was reached, and therefore $\Phi_p(m)$ and $\mathcal{E}_p(1, m)$, as well as a discrimination function $\Psi_{\mathcal{E}_p}(n)$ were calculated for the simulated fBm.

Figs. 3.7 (a-c) represent $\Phi_\infty^{(1)}(m)$, $\Phi_\infty^{(2)}(m)$, $|\mathcal{E}_\infty^{(1)}(m)|$, $|\mathcal{E}_\infty^{(2)}(m)|$, $\Psi_{\mathcal{E}}(1, m)$ and $\Psi_{\mathcal{E}}(n, 1)$. Figs. 3.7 (d-f) represent the latter functions with fuzzy estimations. In Figs. 3.7 (a) and

3.3. HIGH ORDER DEVELOPED ENTROPY

(d) the deviation between $\Phi_\infty^{(1)}(m)$ and $\Phi_\infty^{(2)}(m)$ is in yellow and in Figs. 3.7 (c), 3.7 (f) it is in green. It can be seen from Figs. 3.7 (c), 3.7 (f) that the maximum $\Psi_{\mathcal{E}_p}(n^*, 1)$ is defined by:

$$\begin{cases} \Psi_{\mathcal{E}_p}^* = \max(\Psi_{\mathcal{E}_p}(n, 1)), \\ n^* = \arg \max_n \Psi_{\mathcal{E}_p}(n, 1). \end{cases}$$

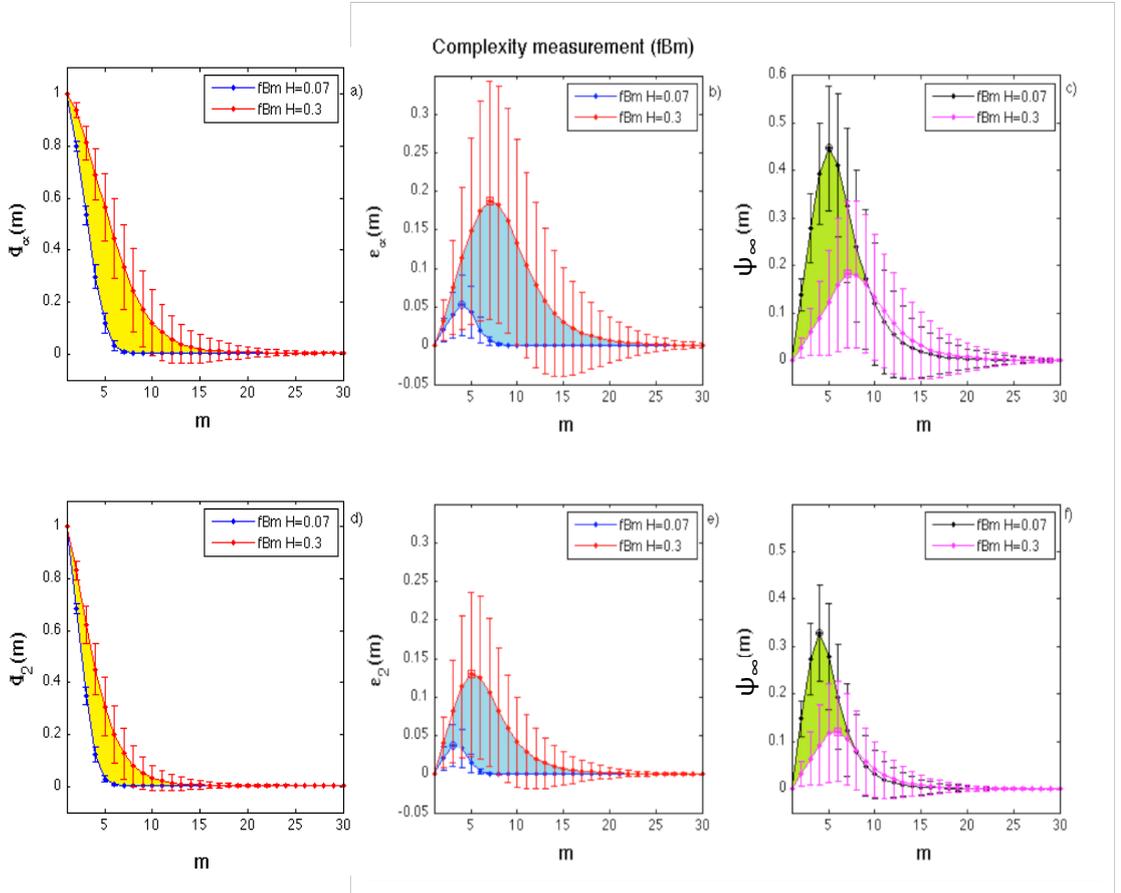


Figure 3.7 – Complexity measurements for two different fBms with $\mathcal{H} = 0.07$ and $\mathcal{H} = 0.3$, ($r = 0.2$). (a) Non-fuzzy estimations with $\Phi_\infty^{(1)}(m)$ (in blue) for $\mathcal{H} = 0.07$ and $\Phi_\infty^{(2)}(m)$ (in red) for $\mathcal{H} = 0.3$. Deviation between $\Phi_\infty^{(1)}(m)$ and $\Phi_\infty^{(2)}(m)$ represented by the yellow area. (b) Non-fuzzy estimations with $|\mathcal{E}_\infty^{(1)}(1, m)|$ (in blue), $|\mathcal{E}_\infty^{(2)}(1, m)|$ (in red). Deviation between $|\mathcal{E}_\infty^{(1)}(1, m)|$, $|\mathcal{E}_\infty^{(2)}(1, m)|$ represented in the blue area. (c) Non-fuzzy estimation with $\Psi_{\mathcal{E}_\infty}(n, 1)$ (in black) and $\Psi_{\mathcal{E}_\infty}(1, m)$ (in magenta). Deviation between $\Psi_{\mathcal{E}_\infty}(n, 1)$ and $\Psi_{\mathcal{E}_\infty}(1, m)$ represented by the green area. (d) Fuzzy estimations with $\Phi_2^{(1)}(m)$ (in blue) for $\mathcal{H} = 0.07$ and $\Phi_2^{(2)}(m)$ (in red) for $\mathcal{H} = 0.3$. (e) Fuzzy estimations with $|\mathcal{E}_2^{(1)}(1, m)|$ (in blue), $|\mathcal{E}_2^{(2)}(1, m)|$ (in red). (f) Fuzzy estimations with $\Psi_{\mathcal{E}_2}(n, 1)$ (in black) and $\Psi_{\mathcal{E}_2}(1, m)$ (in magenta). Deviation between $\Psi_{\mathcal{E}_2}(n, 1)$ and $\Psi_{\mathcal{E}_2}(1, m)$ represented by the green area.

and the maximum $\Psi_{\mathcal{E}_p}(1, m^*)$ is defined by (Figs. 3.7 (b), 3.7 (e)):

$$\begin{cases} \Psi_{\mathcal{E}_p}^* = \max(\Psi_{\mathcal{E}_p}(1, m)), \\ m^* = \arg \max_m \Psi_{\mathcal{E}_p}(1, m). \end{cases}$$

It can be inferred from these graphs that

$$\Psi_{\mathcal{E}_p}(n^*, 1) > \Psi_{\mathcal{E}_p}(1, m^*).$$

For $p = \infty$, $\Psi_{\mathcal{E}_\infty}^* = |\mathcal{E}_\infty^{(1)}(4, 1) - \mathcal{E}_\infty^{(2)}(4, 1)|$ and $\Psi_{\mathcal{E}_\infty}^* = |\mathcal{E}_\infty^{(1)}(1, 6) - \mathcal{E}_\infty^{(2)}(1, 6)|$ were obtained and for $p = 2$, $\Psi_{\mathcal{E}_2}^* = |\mathcal{E}_2^{(1)}(3, 1) - \mathcal{E}_2^{(2)}(3, 1)|$ and $\Psi_{\mathcal{E}_2}^* = |\mathcal{E}_2^{(1)}(1, 5) - \mathcal{E}_2^{(2)}(1, 5)|$ were obtained. This suggests that it is more advantageous to use n-order entropy rather than 1-order entropy. Another interesting point in favour of n-order entropy is that SD (represented by bars in Fig. 3.7) is always smaller for n-order entropy than for 1-order entropy. Several values derived from Fig. 3.7 are reported in Tables 3.1 and 3.2. The results derived from Table 3.1 showed that:

- the lowest SD is obtained for $\mathcal{E}_\infty(4, 1)$ since it represented on average $(23 + 32)/2 = 27.5\%$ of the mean value of entropy (dots in Fig. 3.7) while it represented $(82 + 105)/2 = 93.5\%$ and $(80 + 75)/2 = 77.5\%$ of $\mathcal{E}_\infty(1, 6)$ and $\mathcal{E}_\infty(1, 3)$, respectively;
- the lowest SD is obtained for fuzzy estimations was compared to non-fuzzy estimations. Reductions of the SD $(27.5 - 23) = 4.5\%$, $(93.5 - 81) = 12.5\%$ and $(77.5 - 76.5) = 1\%$ are obtained, respectively.

The results derived from Table 3.2 shows that:

- $\Psi_{\mathcal{E}_\infty}^*(4, 1) > \Psi_{\mathcal{E}_\infty}^*(1, 6)$ for non-fuzzy estimations and $\Psi_{\mathcal{E}_2}^*(3, 1) > \Psi_{\mathcal{E}_2}^*(1, 5)$ for fuzzy estimations;
- the lowest SD was obtained for fuzzy estimations since it represents 30% of the mean value of $\Psi_{\mathcal{E}_\infty}^*(4, 1)$ and 86% of $\Psi_{\mathcal{E}_\infty}^*(1, 6)$ and 29% of $\Psi_{\mathcal{E}_2}^*(3, 1)$ and 85% of $\Psi_{\mathcal{E}_2}^*(1, 5)$. It was advantageous to use fuzzy estimations since SDs were the smallest, however, it can be seen that the mean values obtained from fuzzy estimations are lower than those obtained from non-fuzzy estimations. This latter point was not in favour of fuzzy estimations when fBms were considered.

3.3. HIGH ORDER DEVELOPED ENTROPY

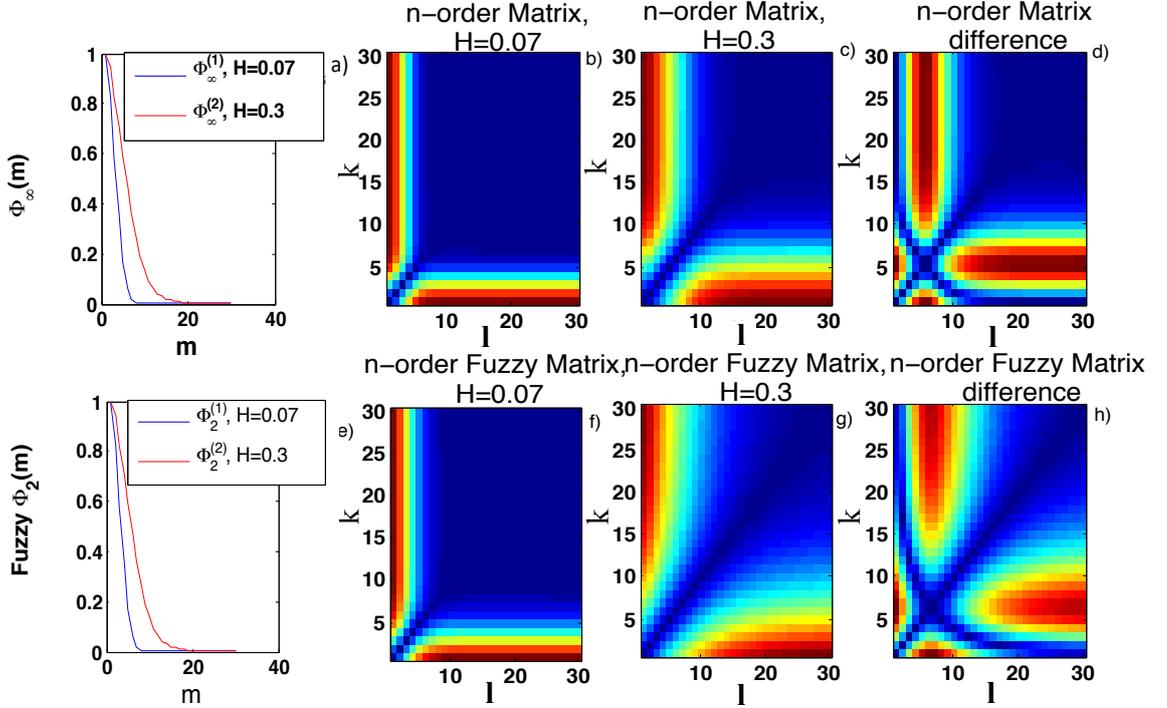


Figure 3.8 – Representation of the matrix $\mathcal{M}_p^{(i)}$ for two different fBMs. (a) Non-fuzzy estimations $\Phi_\infty^{(1)}(m)$ (blue line) and $\Phi_\infty^{(2)}(m)$ (red line) with $\mathcal{H} = 0.07$ and $\mathcal{H} = 0.3$, respectively. (b) Non-fuzzy estimations with $\mathcal{M}_\infty^{(1)}$. (c) Non-fuzzy estimations with $\mathcal{M}_\infty^{(2)}$. (d) Non-fuzzy estimations with difference between $\mathcal{M}_\infty^{(1)}$ and $\mathcal{M}_\infty^{(2)}$. (e) Fuzzy estimations with $\Phi_2^{(1)}(m)$ (blue line) and $\Phi_2^{(2)}(m)$ (red line) with $\mathcal{H} = 0.07$ and $\mathcal{H} = 0.3$, respectively. (f) Fuzzy estimations with $\mathcal{M}_2^{(1)}$. (g) Fuzzy estimations with $\mathcal{M}_2^{(2)}$. (h) Fuzzy estimations with difference between $\mathcal{M}_2^{(1)}$ and $\mathcal{M}_2^{(2)}$.

$\mathcal{M}_p^{(1)}$ and $\mathcal{M}_p^{(2)}$ are reported in Figs. 3.8 (b) and (c) for non-fuzzy estimations, and reported in Figs. 3.8 (f) and (g) for fuzzy estimations. It can be seen from Figs. 3.8 (b), (c), (f) and (g) that the more interesting values were located on the border of the matrix, i.e. for small values of n when m was fixed and vice versa.

The matrix difference $\Psi_{\mathcal{M}_p}$ reported in Fig. 3.8 (d) for non-fuzzy estimations and in Fig. 3.8 (h) for fuzzy estimations, showed that the best values were between $n = 4$ and 8 when $m = 1$ and between $n = 25$ and 30 when $m = 6$. These results indicate that it is worthwhile using the n -order entropy for discriminating fBMs of different levels of complexity.

It was shown that it was more advantageous to use n -order entropy rather than 1 -order entropy. This was confirmed for high values of entropy and for the lowest standard

Table 3.1 – Average \pm standard deviation of specific values of $\mathcal{E}_p(n, m)$ for two fBms with $i = 1$, $\mathcal{H} = 0.07$ and $i = 2$, $\mathcal{H} = 0.3$. Values in parentheses represent the standard deviation normalized by the average value.

Entropy Type \ Hurst Exponent	Non-Fuzzy			Fuzzy		
H	$\mathcal{E}_\infty(4, 1)$	$\mathcal{E}_\infty(1, 6)$	$\mathcal{E}_\infty(1, 3)$	$\mathcal{E}_2(3, 1)$	$\mathcal{E}_2(1, 5)$	$\mathcal{E}_2(1, 3)$
0.3	0.565 \pm 0.129	0.188 \pm 0.154	0.076 \pm 0.061	0.452 \pm 0.099	0.129 \pm 0.106	0.081 \pm 0.065
(i=2)	(23%)	(82%)	(80%)	(22%)	(82 %)	(80%)
0.07	0.120 \pm 0.039	0.006 \pm 0.0063	0.040 \pm 0.030	0.123 \pm 0.03	0.014 \pm 0.0112	0.037 \pm 0.027
(i=1)	(32%)	(105 %)	(75 %)	(24 %)	(80 %)	(73 %)

Table 3.2 – Average \pm standard deviation of specific values of discrimination function $\Psi_{\mathcal{E}_p}(n, m)$ for two fBms with $\mathcal{H} = 0.07, 0.3$. Values in parentheses represent the standard deviation over the average value and the asterisk represents the maximum.

Non-Fuzzy		Fuzzy	
$\Psi_{\mathcal{E}_\infty}^*(4, 1)$	$\Psi_{\mathcal{E}_\infty}^*(1, 6)$	$\Psi_{\mathcal{E}_2}^*(3, 1)$	$\Psi_{\mathcal{E}_2}^*(1, 5)$
0.447 \pm 0.132	0.182 \pm 0.150	0.329 \pm 0.096	0.121 \pm 0.103
(30%)	(86%)	(29%)	(85%)

deviation. These results were obtained for fBms that were hypothesized as a coarse model of FHR signals. In order to confirm that n-order entropy works on biomedical signals, the examination of FHR signals is reported in the subsection below.

3.4 Application to Fetal Heart Rates

3.4.1 Results of Fetal Heart Rate Complexity Detection

From 80 clinical recordings of 30 minutes, the discrimination functions $\Psi_{\mathcal{E}_p}(n, m) = \left| \mathcal{E}_p^{(1)}(n, m) - \mathcal{E}_p^{(2)}(n, m) \right|$ based on n-order similarity entropies $\mathcal{E}_p^{(i)}(n, m)$ ($i = 1$ for healthy and $i = 2$ for distressed fetuses) were calculated through a short term analysis using centered \mathbf{X} vectors composed of 720 points and overlapping by 97%. From each recording composed of 7200 points (one point every 250 ms), 349 non-fuzzy estimations $\mathcal{E}_\infty(n, m)$ and 349 fuzzy estimations $\mathcal{E}_2(n, m)$ were evaluated.

3.4. APPLICATION TO FETAL HEART RATES

From a total number of $2 \times 2 \times (40 + 40) \times 349$ of n-order entropies, i) $\Psi_{\mathcal{M}\infty}$ were reported from non-fuzzy estimations (Table 3.3), ii) $\Psi_{\mathcal{M}2}$ were evaluated from fuzzy estimations (Table 3.4), and iii) average values and relative accuracy of discrimination functions $\Psi_{\mathcal{E}p}(n, m)$ and $\Psi_{\mathcal{E}p}(1, m)$ (Table 3.5), iv) boxplots (Fig. 3.9), and v) sensitivity and specificity (Table 3.6) were determined.

Table 3.3 – Discrimination functions $\Psi_{\mathcal{M}\infty}$ from non-Fuzzy estimations derived from FHR signals with k, l ranging from 1 to 7. Italic values correspond to 1-order similarity entropies. Bold values correspond to the maximum of n-order similarity entropies.

k/l	1	2	3	4	5	6	7
1	0.000	<i>0.068</i>	0.137	0.186	0.203	0.193	0.173
2	<i>0.068</i>	0.000	0.070	0.118	0.135	0.126	0.105
3	0.137	0.070	0.000	<i>0.049</i>	0.066	0.056	0.036
4	0.186	0.118	<i>0.049</i>	0.000	<i>0.017</i>	0.007	0.013
5	0.203	0.135	0.066	<i>0.017</i>	0.000	<i>0.009</i>	0.030
6	0.193	0.126	0.056	0.007	<i>0.009</i>	0.000	<i>0.020</i>
7	0.173	0.105	0.036	0.013	0.030	<i>0.020</i>	0.000

Table 3.4 – Discrimination functions $\Psi_{\mathcal{M}2}$ from fuzzy estimations derived from FHR signals with k, l ranging from 1 to 7. Italic values correspond to 1-order similarity entropies. Bold values correspond to the maximum of n-order similarity entropies.

k/l	1	2	3	4	5	6	7
1	0.000	0.067	0.131	0.177	0.201	0.206	0.202
2	0.067	0.000	<i>0.064</i>	0.110	0.134	0.140	0.135
3	0.131	<i>0.064</i>	0.000	<i>0.046</i>	0.070	0.076	0.071
4	0.177	0.110	<i>0.046</i>	0.000	<i>0.024</i>	0.030	0.025
5	0.201	0.134	0.070	<i>0.024</i>	0.000	<i>0.006</i>	0.001
6	0.206	0.140	0.076	0.030	<i>0.006</i>	0.000	<i>0.005</i>
7	0.202	0.135	0.071	0.025	0.001	<i>0.005</i>	0.000

From Tables 3.3 and 3.4 it can be seen that the lowest values belonged to the 1-order entropy estimations. The best values of $\Psi_{\mathcal{E}p}(n^*, 1)$ and $\Psi_{\mathcal{E}p}(1, m^*)$ are reported in bold in

Tables 3.3 and 3.4, and the maximum values of 1-order entropies corresponded to $\Psi_{\mathcal{E}\infty}^*(1, 2)$ and $\Psi_{\mathcal{E}2}^*(1, 1)$. Of all Ψ^* it was obvious that the best values did not correspond to the 1-order entropies but rather to $\Psi_{\mathcal{E}\infty}(4, 1)$ and $\Psi_{\mathcal{E}2}(5, 1)$ for non-fuzzy and fuzzy estimations, respectively. This outcome confirmed that it was advantageous to use n-order over 1-order based-estimators of complexity.

Table 3.5 – Average values (aver.) and relative accuracy (standard deviation/ average = std./aver.) of discrimination functions $\Psi_{\mathcal{E}p}(n, m)$ derived from FHR signals. Fuzzy estimations were of benefit for reducing estimation fluctuations (std/aver.)

		$\Psi_{\mathcal{E}p}(1, 1)$	$\Psi_{\mathcal{E}p}^*(1, 2)$	$\Psi_{\mathcal{E}p}(4, 1)$	$\Psi_{\mathcal{E}p}(5, 1)$
Non-fuzzy	aver.	0.07	0.07	0.20	0.19
	std/aver.	45%	50 %	50 %	54%
Fuzzy	aver.	0.07	0.06	0.20	0.21
	std/aver.	44%	49 %	48 %	49%

Results shown in Table 3.5 demonstrated that it was preferable to use fuzzy algorithms since i) the relative precision (standard deviation/average) was lower for fuzzy than for non-fuzzy estimations and ii) average values were similar for fuzzy and non-fuzzy estimations, and no bias was reported.

Table 3.6 – Sensitivity of discrimination functions $\Psi_{\mathcal{E}p}(n, m)$ calculated from FHR signals. Note that $\Psi_{\mathcal{E}p}^*(n, 1)$ was equal to $\Psi_{\mathcal{E}\infty}^*(4, 1)$ for non-fuzzy estimations and to $\Psi_{\mathcal{E}2}^*(5, 1)$ for fuzzy estimations. Fuzzy estimations were of benefit since a slight improvement in both sensitivity and specificity can be seen.

		$\Psi_{\mathcal{E}p}(1, 1)$	$\Psi_{\mathcal{E}p}^*(1, 2)$	$\Psi_{\mathcal{E}p}^*(n, 1)$
Non-fuzzy	Sensitivity	87.3%	87.7%	88.5 %
	Specificity	87.3%	87.7%	88.5 %
Fuzzy	Sensitivity	87.7%	88.5 %	90.1 %
	Specificity	87.7%	88.5 %	90.1 %

The findings shown in Fig. 3.9 demonstrated that meaningful classifications were obtained using fuzzy estimations since outliers were removed only for $\Psi_{\mathcal{E}2}(1, 2)$ and $\Psi_{\mathcal{E}2}(5, 1)$. It can be seen that the level of similarity was lower for distressed fetuses than for healthy fetuses. This outcome corroborated the results previously reported in [Ferrario et al., 2009]. Furthermore, it can be seen that the similarity measurements were fairly homogeneous. This was certainly due to the good homogeneity of the fetus groups. Note that the

3.4. APPLICATION TO FETAL HEART RATES

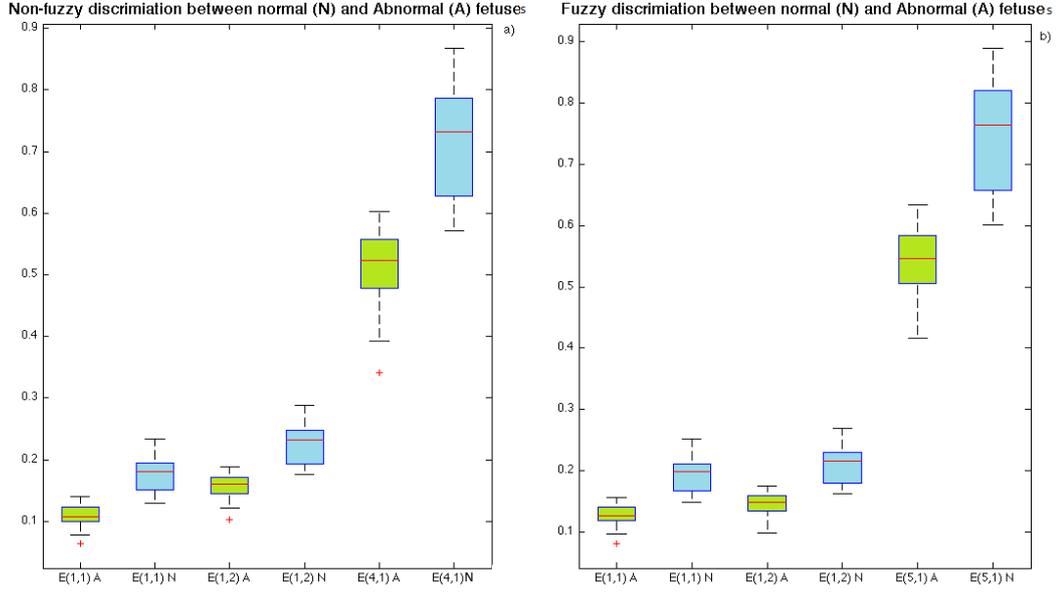


Figure 3.9 – Boxplots of 1-order entropies and n-order entropies calculated from FHR signals. (left) Non-fuzzy estimations. Boxplot of 1-order entropies $\mathcal{E}(1,1)$, $\mathcal{E}^*(1,2)$, and n-order entropy $\mathcal{E}(4,1)$ for healthy (in green) and distressed fetuses (in blue). (right) Fuzzy estimations. Boxplot of 1-order entropies $\mathcal{E}(1,1)$, $\mathcal{E}^*(1,2)$ and n-order entropy $\mathcal{E}(5,1)$ for healthy (in green) and distressed fetuses (in blue).

high level of performance of n-order entropies was confirmed by the very good values of sensitivity and specificity reported in Table 3.6.

The findings reported in Table 3.6 showed that the sensitivity and specificity were good since they were above 87%. Fuzzy estimations were preferable since slightly improved sensitivity and specificity were found. The best performance was obtained for $\Psi_{\mathcal{E}_{\infty}}^*(5,1)$, suggesting that n-order fuzzy entropy parameters were better than 1-order fuzzy entropies.

3.4.2 Discussion

In this chapter, which is dedicated to the discrimination of time series exhibiting different complexity by entropy analysis, four key points were emphasized.

The first main result was derived from the search for the best setting parameter i.e. pattern size m^* when the tolerance r was constant and fixed at $r = 0.2$. As a follow-up of previous studies by [Lu et al., 2008, Restrepo et al., 2014], and based on the research for the best setting r^* to maximize entropy $\mathcal{E}_* = \mathcal{E}(r^*)$, our findings showed that it was also of benefit to find out the pattern size m^* that maximized entropy $\mathcal{E}^* = \mathcal{E}(1, m^*)$. In

particular, it was shown that both m^* and $\mathcal{E}(1, m^*)$ were related to intrinsic features of the system, such as the Hurst exponent of fBm.

Moreover, it was shown that it was easy to discriminate fBm for Hurst exponents $\mathcal{H} < 0.5$. Surprisingly, it was even shown that $\mathcal{E}^*(\mathcal{H})$ was sensitive to the anti-correlated and correlated nature of the time series under study. In any event, both m^* and $\mathcal{E}(1, m^*)$ were sensitive to the complexity changes. This latter behavior was also observed when the operating mode of the Lorenz system changed from a periodic to a chaotic regime. This outcome augured well for detecting complexity changes of other nonlinear dynamic systems or for discriminating other kinds of signals of different complexity. One of the novel findings were the significant values for pattern size m^* since values lower than 20 were often observed [Restrepo et al., 2014]. However, what was even more surprising, the huge gap between the two m^* values obtained when the system was in periodic regime ($m^* = 70$) and in chaotic regime ($m^* = 7$). Finding the best setting of m was interesting since it provides a complementary insight into the role of m that was now no longer equivalent to the role of embedding dimension. It can therefore be confirmed that it is necessary to seek both the value of m^* that maximizes the entropy and the value of the maximum entropy $\mathcal{E}(1, m^*)$.

The second main result was linked to the development of a new paradigm that encompasses the 1-order similarity entropy. The development of this new concept took advantage of the monotonic decrease in the function $\Phi(m)$, a feature that has never been used to date. As for m^* and \mathcal{E}^* , it was shown that the n-order similarity entropy has quite easily discriminated fBms with low Hurst exponents. This outcome augured well for discrimination of biomedical signals since it had been reported in [Oudjemia et al., 2013] that such low Hurst exponents are measured for FHR signals.

The third important outcome concerned the validation of the different descriptors of complexity in the discrimination between healthy and distressed fetuses. Using the best discrimination function based on the n-order similarity entropy, it was shown that it was fairly easy to discriminate between fetuses without disorders and fetuses with IUGR. It seemed clear that the difference in complexity between the two groups of fetuses was behind the good discrimination. In this study, a considerable relative error of $(0.20 - 0.07)/0.07 = 186\%$ (see Table 3.5) between n-order similarity entropy and 1-order similarity entropy was obtained. Although the n-order similarity entropy was very appealing, it was still difficult to identify exactly what led the n-order similarity entropy to outperform the other methods. It was probably a mixture of several elements based

3.5. CONCLUSION

on i) the monotonic decrease in the probability of finding similar patterns of size m , ii) the maximum gap between similarity measurements of patterns of non-consecutive size and iii) other features that remain hidden. The high level of performance in terms of sensitivity and specificity obtained for all complexity descriptors used in this study was certainly due to the level of homogeneity of each group. Nevertheless, it is important that this study be extended to greater numbers of recordings. In any event, these findings augur well for the discrimination of other kinds of biomedical signals.

In the entropy analysis part of this dissertation the:

- problem of setting entropy descriptors was solved by varying the pattern size and fixing the tolerance;
- optimal pattern sizes that maximizes the similarity entropy were found;
- intrinsic features of the time series and systems were linked to the optimal pattern sizes;
- FHR discrimination was improved by the new developed paradigm that encompasses the standard similarity entropy: n -order fuzzy similarity entropy;
- sensitivity and specificity of the discrimination between healthy and distressed fetuses with IUGR confirmed the improvement in FHR discrimination for $n = 4$ SimEn and $n = 5$ for FuzzySimEn.

3.5 Conclusion

Profound investigations of the link between m -patterns was carried out in this chapter. By introducing the maximum and n -order concepts to entropy, it was possible to improve the discrimination between different FHRs and more robustly classify healthy and distressed fetuses as compared to the standard ApEn, SampEn, SimEn and FuzzyEn. N -order entropies combined with fuzzy frameworks are effective for diagnosing the state of the fetus. This new paradigm has an overwhelming potential and could be applied to other applications seeking the extraction of complexity invariant. The problem of setting entropy descriptors was solved by varying the pattern size leading an optimal value that maximizes the n -order FuzzySimEn. FHR discrimination was improved by the new developed paradigm that encompasses the standard SimEn. However, there is no explicit

correspondence between regularity and complexity. Predictable (i.e. periodic) time series did not necessary possess the minimum entropy, also, non-predictable signals (noisy or uncorrelated) did not possess the maximum entropy. Therefore, complexity could not be written in an explicit form, and there is no consensus on its definition. Intuitively, complexity is associated with significant incorporation in the correlated structures on multiple spatio-temporal samples. To track the complexity, it is judicious to understand the phenomena associated with it such as the spatio-temporal incorporated structures. This induced the use of the multi-fractal analysis which will be the subject of chapter 4.

Chapter 4

MULTI-FRACTAL ANALYSIS

▮ *It is popular that one should count till ten before taking any action. Why a scale of ten ? Some situations requires counting till more than this scale, multi-fractality of situations leads the correct analysis, scaling and counting* ▮
Amira Zaylaa.

4.1 Introduction

This chapter deals with the Multi-fractal analysis of signals. Such analysis was triggered by the scale invariance properties observed through the power law spectral density [Kobayashi and Musha, 1982] of heart rate variability (HRV). Multi-fractal analysis can be achieved by various significant approaches such as box-counting, Detrended Fluctuation Analysis (DFA), length techniques, and wavelet transform. The core of the technique we present in this chapter is the proposition of a structure function along with coarse-graining. The latter contribution to the field is under the category of multi-fractal analysis by length methods.

In Euclidean space, a point has a dimension 0, a line has a dimension 1 and a surface has a dimension of 2 etc. A fractal object is any object having a non-integer dimension [Akay, 2000]. This ameliorates the classification of objects into either standard or fractal. The multi-fractal analysis is a generalization of the fractal analysis, and the major multi-fractal analysis techniques are summarized in Fig. 4.1. The first and simplest method is box-counting, it is called the Kolmogorov capacity and leads to the singularity spectrum. A major problem with box-counting is to cover all the system with small sized cells [Akay, 2000]. Grassberger and Procaccia have proposed an algorithm to compute the correlation sum $C(r)$ approximately based on the Heaviside function. Then, they assumed

that the correlation dimension (D_2) is deduced from the slope of the curve $\log(C(r))$ versus $\log(r)$. However, the previous assumption is not accurate and subject to errors, and the value of r was inferred empirically [Akay, 2000].

Researchers have applied multi-fractal analysis to various applications using box-counting, differential box-counting, Renyi dimension, wavelet transform, wavelet transform modulus maxima (WTMM) and direct length methods (see Fig. 4.1) [Keller et al., 1989, Levy-Vehel, 1995, Chaudhuri and Sarkar, 1995, YUM and Kim, 2002, Arneodo et al., 2003, Popivanov et al., 2005, Xia et al., 2006, Wang et al., 2007, Hsu et al., 2007]. One of the techniques of multi-fractal analysis is the length or the direct determination method. It is highlighted in an orange color in Fig. 4.1 as it is the core method in which we built our investigation.

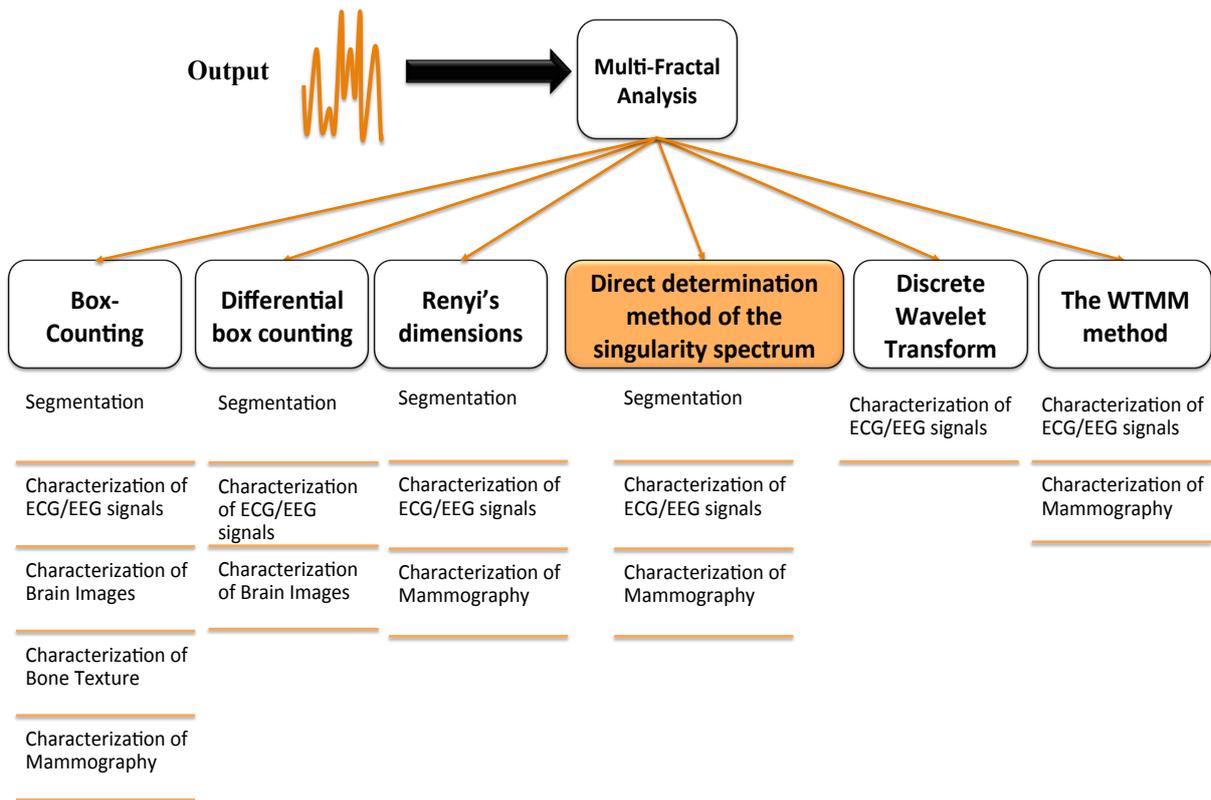


Figure 4.1 – Schematic illustration of the common multi-fractal analysis techniques.

Ivanov et al. were the first to demonstrate multi-fractality in cardiac dynamics as well as in physiological dynamics [Ivanov et al., 1996, Ivanov et al., 1999, Ivanov et al., 2001]. Recently, HRV analysis proposed by Ferrario et al. was built on multi-fractal analy-

sis [Ferrario et al., 2006, Ferrario et al., 2009]. These studies were associated multi-fractal analysis of physiological time series [Wang et al., 2003, Wang et al., 2005, Sassi et al., 2009, Humeau et al., 2010]. Since the time scale dependency of the multi-fractal features is still ambiguous, we proposed a combination of the coarse-graining approach, structure function and multi-fractal analysis and applied them on FHRs. The technique discussed in this chapter is based on two pioneer works; the first is the work of Wang et al. that combined the multi-fractal analysis of adult ECG with the coarse-graining approach proposed by Zhang [Zhang, 1991], and the second is based on a structure function after which the multi-fractal indicators are extracted [Barabasi and Vicsek, 1991, Muzy et al., 1993, Lin and Hughson, 2001, Ching and Tsang, 2007, Frisch and Parisi, 1985]. We hypothesized that the profound investigation of the time-scale dependence of the multi-fractal features of FHRs improves the discrimination between distressed and healthy FHRs.

4.2 Multi-Fractal Descriptors

Multi-fractal analysis is a geometric quantification of the output of nonlinear systems. Various multi-fractal complexity parameters can be extracted simultaneously from the time series. Multi-fractal parameters involved in our dissertation are the Hurst Exponent \mathcal{H} , Singularity Spectrum ($D(q)$), Holder spectrum ($h(q)$) and the scaling parameters $\tau(q)$ and $\eta(q)$.

4.2.1 Singularity Spectrum

The singularity spectrum $D(q)$ measures the variation of the density of multi-fractals relative to the scale length. The structure function used in this dissertation is defined [Frisch and Parisi, 1985] for $q > 0$ as follows:

$$\mathbf{Q}(q, \epsilon) = \left(\int |y_{\alpha}^{(i)}(t + \epsilon) - y_{\alpha}^{(i)}(t)|^q dt \right)^{1/q}, \quad (4.1)$$

it is a length measurement [Girault et al., 2010] where $|y_{\alpha}^{(i)}(t + \epsilon) - y_{\alpha}^{(i)}(t)|^q$ reveals a local behavior while $(\int \dots dt)^q$ reveals a global behavior and ϵ reveals the increment or tolerance.

If the structure function $Q(q, \epsilon) = k \cdot \epsilon^{\eta(q)}$ then:

$$\eta(q) = \lim_{\epsilon \rightarrow 0} \frac{\log(Q(q, \epsilon))}{\log(\epsilon)}, \quad (4.2)$$

For fBm which models stochastic signals of Hurst exponent \mathcal{H} , the scaling exponent is $\eta(q) = \mathcal{H}$.

From the previous equation, the singularity spectrum $D(q)$ is evaluated based on the previous length method and given as follows:

$$\mathbf{D}(q) = q^2 \frac{d\eta(q)}{dq} + 1, \quad (4.3)$$

where q is the number of order or the information dimension. The singularity spectrum $D(q)$ can also be obtained through a Legendre transform from the scaling parameter $\tau(q)$:

$$\mathbf{D}(q) = q \frac{d\tau(q)}{dq} - \tau(q), \quad (4.4)$$

where $\tau(q)$ is another scaling exponent defined by Eq. 4.5 and $d\tau(q)/dq$ is the approximation of the Holder coefficient h [Lopes and Betrouni, 2009].

$$\tau(q) = q\eta(q) - 1. \quad (4.5)$$

4.2.2 Hurst Exponent

It is the index reflecting the long-range dependence of states of the time series [Hurst, 1951, Hurst et al., 1965, Mandelbrot and Wallis, 1969, Mesa and Poveda, 1993]. This parameter measures the rate at which the autocorrelations of time series decrease as the time delay between values increases.

As the fractal dimension $D = 2 - \mathcal{H}$, this implies that the Hurst exponent $\mathcal{H} = 2 - D$ [Lopes and Betrouni, 2009].

4.2.3 Holder Spectrum

The Holder spectrum is a multi-fractal parameter measuring the uniformity of a time series, it is written as:

$$\mathbf{h}(q) = q \frac{d\eta(q)}{dq} + \eta(q), \quad (4.6)$$

This parameter could also be written as:

$$\mathbf{h}(q) = \frac{d\tau(q)}{dq}. \quad (4.7)$$

4.3 Contribution to Multi-Fractal Analysis

This work was based on the use of a structure function. As there are a lot of multi-fractal parameters that could be extracted from the studied time series using the structure function, we have tried to calculate the major parameters and compare their performance for discrimination purposes. The block diagram demonstrating our contribution to multi-fractal analysis and its prior requirements is set out in Fig. 4.2. The method is not applied simultaneously on the original time series, rather it is applied on the coarse-grained time series. The process of coarse-graining [Zhang, 1991] was applied as proposed by Zhang:

- For instance, each FHR time series $x(n)$ composed of M points is analyzed;
- Multi-scale analysis is introduced to capture the fluctuations present in the time series at different scales. It consisted of evaluating approximate versions of the original time series from a local average of neighbouring points;
- A new reduced time series composed of M/α samples at α scales are produced by coarse-graining using Eq. 3.6, where $\alpha = s$.

It is worthwhile highlighting the difference between the work of Wang et al. and the work proposed in this chapter. Wang et al. proposed the use of a partition function [Wang et al., 2005], while we proposed a structure function a more mathematically easier approach.

Due to the nonstationary nature of the coarse-grained time series, a short-term procedure comprised evaluating multi-fractal descriptors from $y_\alpha^{(i)}(k)$ sub-signals composed of $M = 720$ points and issued over 3-minutes.

Among all existing methods making use of multi-fractal descriptors, we used $Q(q, \epsilon)$. Although it has been theoretically demonstrated that for certain types of signals the

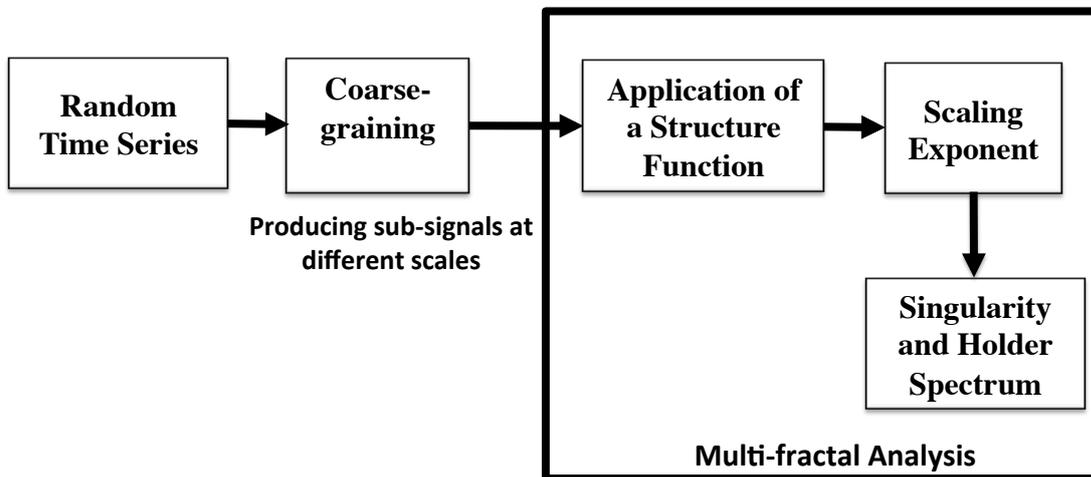


Figure 4.2 – The Block Diagram of the proposed multi-fractal analysis.

methods based on structure function of order q is limited for $q < 0$, we believe that this approach is still worth using because of the following:

- Q is by far the simplest method to implement compared to box counting and wavelet methods.
- For $q < 0$ it is valuable to analyze very small variations in time series. However, as time series were mostly corrupted with noise, clear probing of small variations in the time series was not trivial. The practical value of a negative order q was strongly limited by the presence of noise.
- The signals under consideration were not theoretical signals. This means that mathematical demonstrations operating exclusively on theoretical signals are not systematically applicable in practice.
- Several multi-fractal analyses showed that it was more possible to discriminate between healthy and distressed subjects for $q > 0$ than for $q < 0$. This was supported by the work of Ivanov et al. [Ivanov et al., 2001] where they have showed that the difference between the scaling exponent $\tau(q)$ obtained for healthy and distressed

4.3. CONTRIBUTION TO MULTI-FRACTAL ANALYSIS

subjects was greater for $q > 0$ than for $q < 0$. It was also supported by the works of Stanley et al. [Stanley, 1996, Amaral et al., 2001] where they have showed clearly that for $q > 0$ it was possible to discriminate patients better with atropine than with placebo. The structure function was defined by Eq. 4.1.

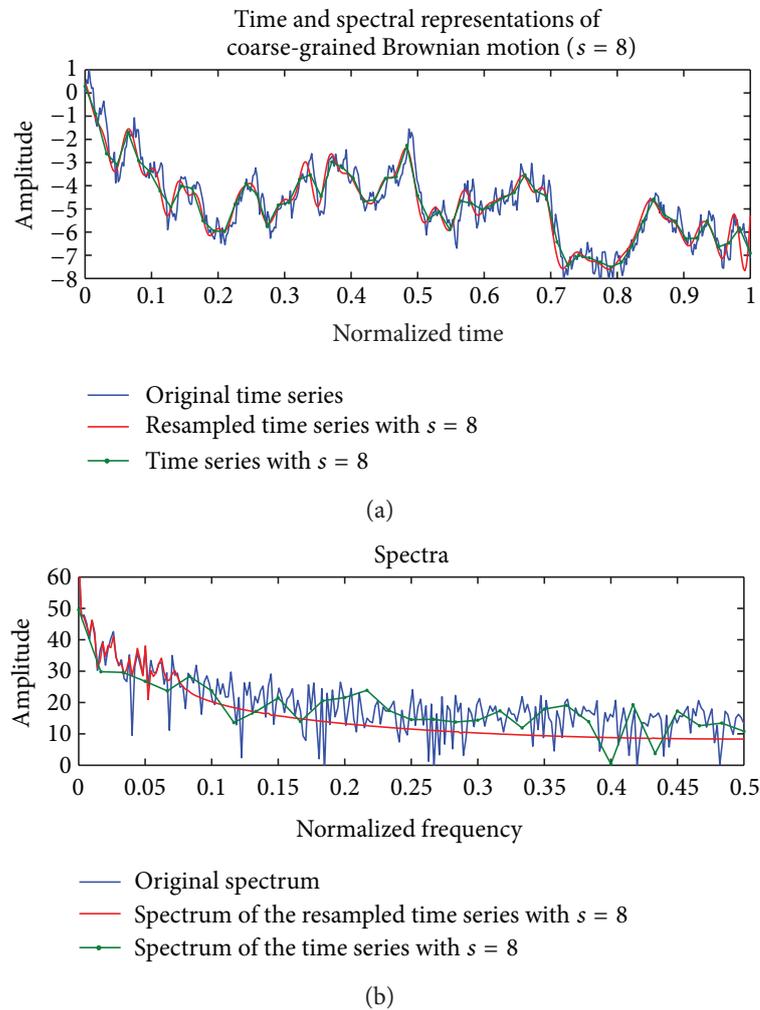


Figure 4.3 – Time and spectral representations of a Brownian motion. (a) Original time series superimposed on the coarse-grained time series ($s=8$) and the re-sampled coarse-grained time series ($s=8$). (b) Spectrum of each time series depicted in (a).

4.4 Example On Fractional Brownian Motion

Fig. 4.3 sets out the time and the spectral representations of coarse-grained fBm time series used to calculate one of the effects resulting from the coarse-grained procedure. The time and the frequency were normalized. Fig. 4.3 (a) shows the original time series superimposed on the coarse-grained time series with $s = 8$, and the re-sampled coarse-grained time series with $s = 8$. Note that a re-sampled coarse-grained time series was an interpolated and filtered time series by a factor of α . The re-sampled coarse-grained times series was composed of s samples.

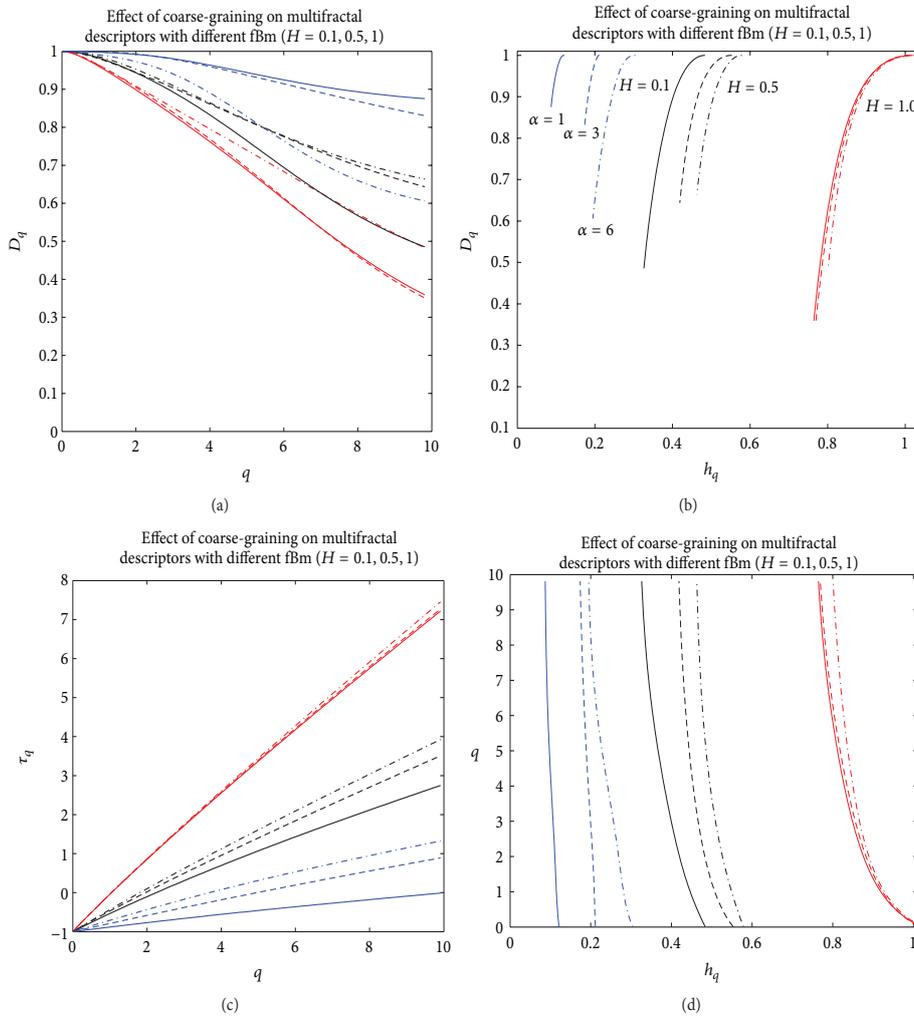


Figure 4.4 – Effects of coarse-graining on multi-fractal descriptors for different fBm of Hurst exponents $\mathcal{H} = \{0.1, 0.5, 1\}$ with different scale factor $\alpha = \{1, 3, 6\}$. (a) Singularity spectrum D versus q , (b) Singularity spectrum D versus Holder spectrum h , (c) Scaling exponent τ versus q and (d) Holder spectrum of healthy and distressed FHRs.

4.5. APPLICATION TO FETAL HEART RATES

Fig. 4.3 (a) and (b) show clearly that the coarse-grained time series were filtered time series. It can be claimed from these outcomes that the reduced duration of each coarse-grained times series is a side effect that can be avoided by re-sampling. Here, the multi-fractal descriptors were evaluated from the re-sampled coarse-grained time series. As previously shown by Girault and others [Girault et al., 2010], as the fBm is more filtered it becomes more regular. Thus, it can be inferred that as the α scaling factor increases, the Hurst exponent \mathcal{H} increases. Xu et al. [Xu et al., 2011] showed that the coarse-graining affected the anti-correlated time series ($\mathcal{H} < 0.5$) in a more pronounced way than the correlated time series ($\mathcal{H} > 0.5$). By assuming that the coarse-graining effect is equivalent to a low pass filtering, the coarse-graining filter performs as follows:

- o For $\mathcal{H} < 0.5$, fBm has several high frequency components that can be removed. The time series before and after filtering are quite different, indicating that the coarse-graining has a non-negligible effect on the time series.
- o For $\mathcal{H} > 0.5$, fBm has several low frequency components that are slightly removed by the coarse-graining effect. Time series before and after filtering were fairly similar, indicating that the coarse-graining has a negligible effect on time series.

Fig. 4.4 represents different multi-fractal descriptors for different fBm of Hurst exponents $\mathcal{H} = \{0.1, 0.5, 1\}$ with different scale factors $\alpha = \{1, 3, 6\}$. It shows the effect of coarse-graining on the previous discussed multi-fractal parameters. The fBm under consideration was composed of 720 samples. The results derived from Fig. 4.4 shows that the anti-correlated fBm of Hurst exponent $\mathcal{H}= 0.1$ is more affected by the coarse-graining effect than the correlated fBm of Hurst exponent $\mathcal{H}= 1$. These results were in accordance with those reported in [Girault et al., 2010, Xu et al., 2011].

4.5 Application To Fetal Heart Rates

From our own dataset comprised 80 recordings, each time series of $M=7200$ points was coarse-grained for 6 different scales. From each coarse-grained signal, sub-signals composed of 720 points and overlapping by 97% were analyzed with multi-fractal tools.

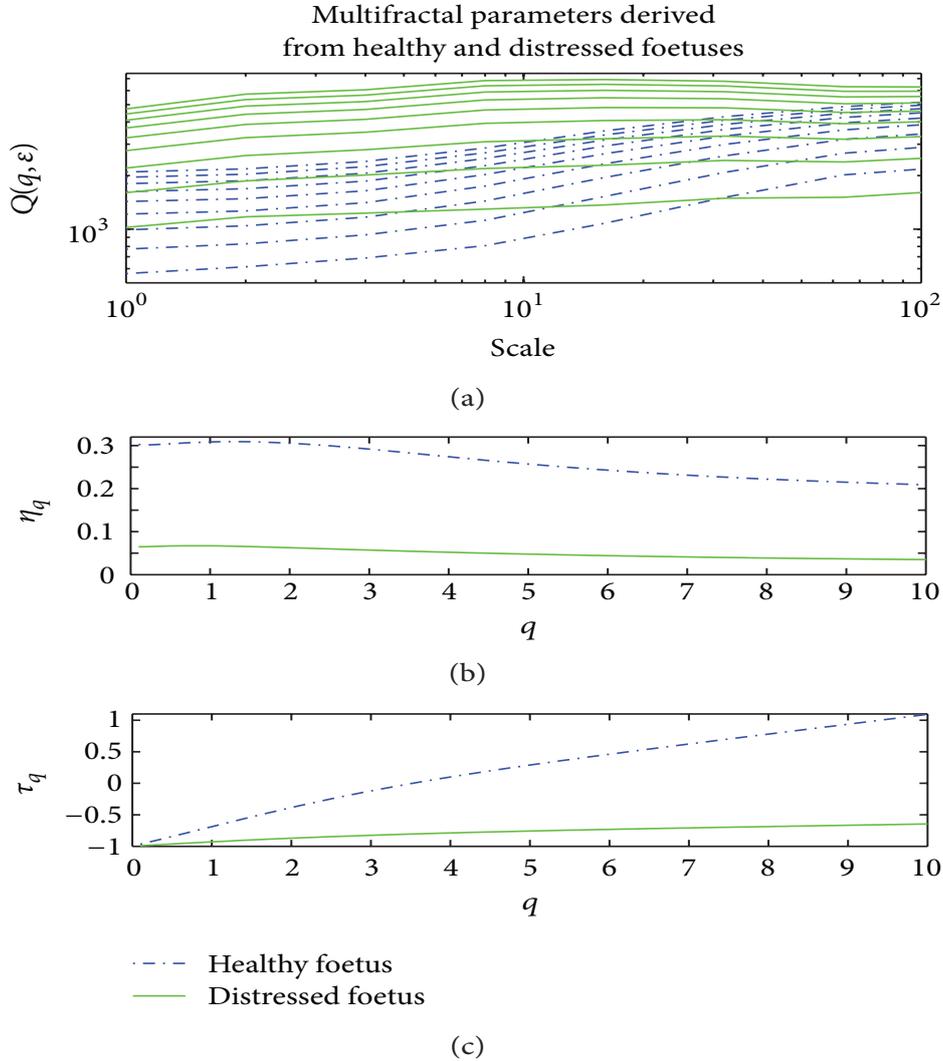


Figure 4.5 – Multi-fractal parameters for a healthy fetus (in blue) and a distressed fetus (in green). (a) Structure function $Q(q, \epsilon)$ versus scale. (b) Scaling exponent η versus q . (c) Scaling exponent τ versus q . The two scaling exponents were more nonlinear for a healthy fetus than for a distressed fetus.

4.5.1 Results Discriminating Healthy from Distressed Fetuses

By analogy to the entropy work, the same FHR database was used. The structure function $Q(q, \epsilon)$ and the scaling exponents $\eta(q)$ and $\tau(q)$ for a healthy and distressed fetuses are reported in Fig. 4.5. Fig. 4.5 (a) shows that the slopes of the curves obtained for different values of q derived from the structure function $Q(q, \epsilon)$, are similar for the healthy fetuses. By analogy to H-FHRs, similar results were derived for distressed fetuses. Figs. 4.5 (b)

4.5. APPLICATION TO FETAL HEART RATES

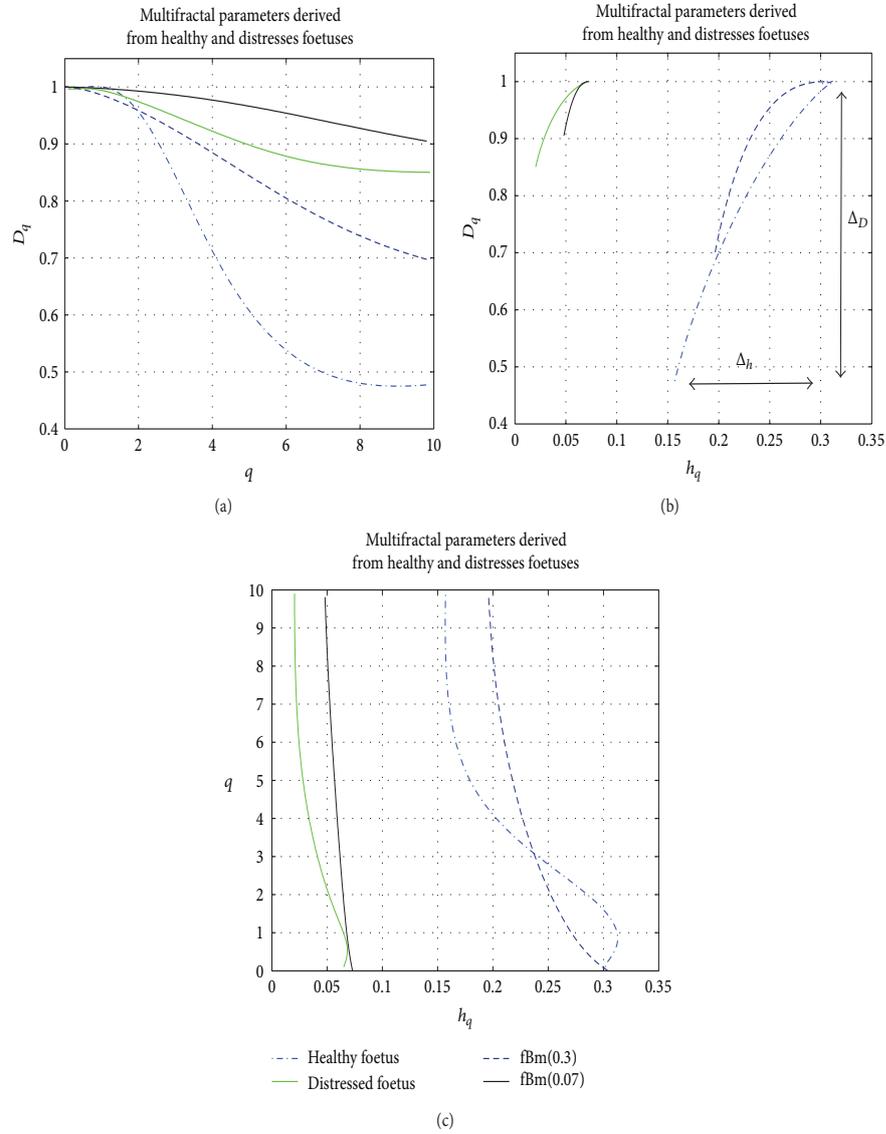


Figure 4.6 – Multi-fractal parameters for a healthy fetus (shown in blue) and a distressed fetus (shown in green). (a) Singularity spectrum D versus q . (b) Singularity spectrum $D(q)$ versus Holder spectrum $h(q)$. (c) Holder spectrum h versus q .

and (c) show that both scaling exponents $\eta(q)$ and $\tau(q)$ are more nonlinear for the healthy fetus than for the distressed one.

Other multi-fractal descriptors such as the singularity spectrum $D(q)$ and the Holder spectrum $h(q)$ are reported in Fig. 4.6. Fig. 4.6 sets out the results obtained from four different signals: a signal from a distressed fetus of an estimated Hurst exponent $\mathcal{H} = 0.07$, a H-FHR of an estimated Hurst exponent $\mathcal{H} = 0.31$, and two fBm signals of Hurst

exponents $\mathcal{H} = 0.07$ and $\mathcal{H} = 0.31$. These four signals were each composed of 720 samples. Fig. 4.6 shows that the magnitude of the dynamics of the singularity spectrum $D(q)$ and the Holder spectrum $h(q)$ is higher for the healthy fetus compared to the distressed fetus.

Similarly, the magnitude of the dynamics of $D(q)$ and $h(q)$ is higher for fetal signals than for the fBm. This corroborates the studies based on the analysis of multi-fractal HRV [Ivanov et al., 1999], where a more pronounced multi-fractal feature for healthy subjects was demonstrated than for distressed subjects. The Holder spectrum for healthy and distressed fetuses decreased with increasing values of q , thus supporting the multi-fractal nature of FHR time series. Such results are consistent with previous similar studies [Ivanov et al., 1999, Sassi et al., 2009]. Several measurements are performed in order to quantify the different trends observed in the multi-fractal indicators $D(q)$ and $h(q)$ for different α -scales.

To evaluate the performance of discrimination, the Relative Error (RE) is calculated for all the involved multi-fractal parameters. RE_1 (in %) of the Hurst exponent is defined as follows:

$$\mathbf{RE}_1(\alpha) = \frac{|\overline{\mathcal{H}}_n(\alpha) - \overline{\mathcal{H}}_d(\alpha)|}{\overline{\mathcal{H}}_n(\alpha)}, \quad (4.8)$$

where $\mathcal{H} = \eta(q)$ is the Hurst exponent for all q with $\mathcal{H} = \eta(1)$, $\overline{\mathcal{H}}_n$ and $\overline{\mathcal{H}}_d$ are the mean Hurst exponents corresponding to the average value obtained for all healthy and distressed fetuses, respectively.

The relative error RE_2 (in %) of the dynamics of $h(q)$ is defined as follows:

$$\mathbf{RE}_2 = \frac{|\overline{\Delta}_{hn} - \overline{\Delta}_{hd}|}{\overline{\Delta}_{hn}}, \quad (4.9)$$

where $\Delta_h = \max(h) - \min(h)$ stands for the dynamics of $h(q)$, $\overline{\Delta}_{hn}$ and $\overline{\Delta}_{hd}$ stand for the mean dynamics corresponding to the average value obtained for all healthy and distressed foetuses, respectively.

The relative error RE_3 of the singularity spectrum (in %) is defined as follows:

$$\mathbf{RE}_3 = \frac{|\overline{D}_n - \overline{D}_d|}{\overline{D}_n}, \quad (4.10)$$

where $\overline{D} = \text{mean}(D(q))$ is the mean value of the singularity spectrum, \overline{D}_n and \overline{D}_d are the mean values corresponding to the average values obtained for all healthy and distressed fetuses, respectively.

4.5. APPLICATION TO FETAL HEART RATES

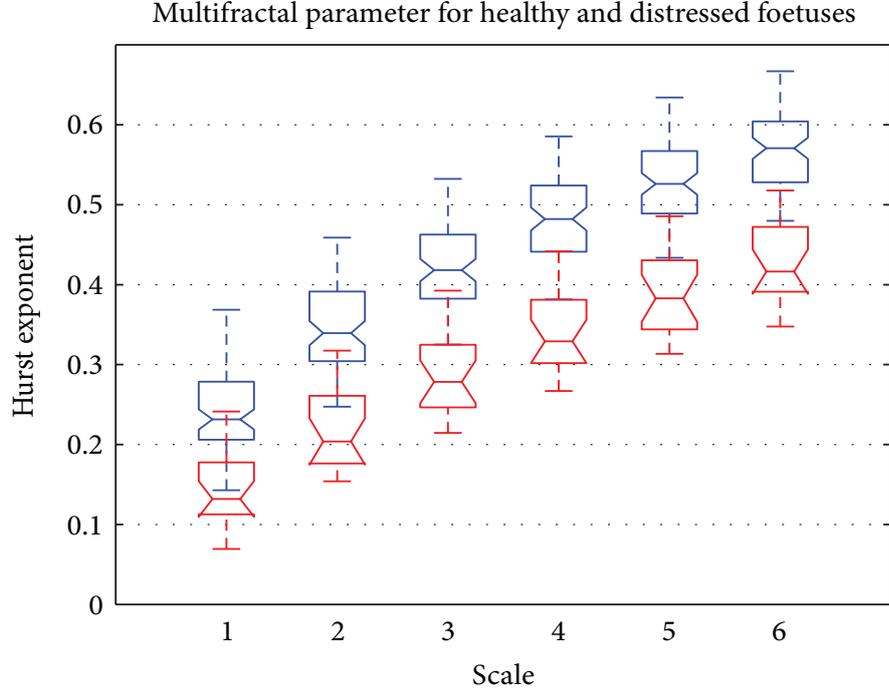


Figure 4.7 – Boxplot of Hurst exponents versus scale for healthy (shown in blue) and distressed (shown in red) fetuses.

The relative error RE_4 of the singularity spectrum difference (in %) is defined as follows:

$$\mathbf{RE}_4 = \frac{|\overline{\Delta_{Dn}} - \overline{\Delta_{Dd}}|}{\overline{\Delta_{Dn}}}, \quad (4.11)$$

where $\Delta_D = \max(D(q)) - \min(D(q))$ is the mean value of the singularity spectrum difference, $\overline{\Delta_{Dn}}$ and $\overline{\Delta_{Dd}}$ are the mean values corresponding to the average value obtained for all healthy and distressed fetuses, respectively.

Fig. 4.7 shows a boxplot representation of the mean Hurst exponent for scale values ranging from 1 to 6. Red boxplots correspond to distressed foetuses, however blue boxplots correspond to healthy fetuses. Fig. 4.7 shows that the mean Hurst exponent for healthy fetuses was higher than that obtained for distressed ones. This meant that the signatures of distressed fetuses were more irregular and complex than those obtained for healthy fetuses. Furthermore, Fig. 4.7 shows that there is a sufficient deviation between the mean Hurst exponent of H-FHRs and D-FHRs which serves well for discrimination purposes. Fig. 4.7 also shows that as the scale increases, the more regular or filtered the signal becomes (Fig. 4.7). This was in accordance with the results of [Girault et al., 2010] which

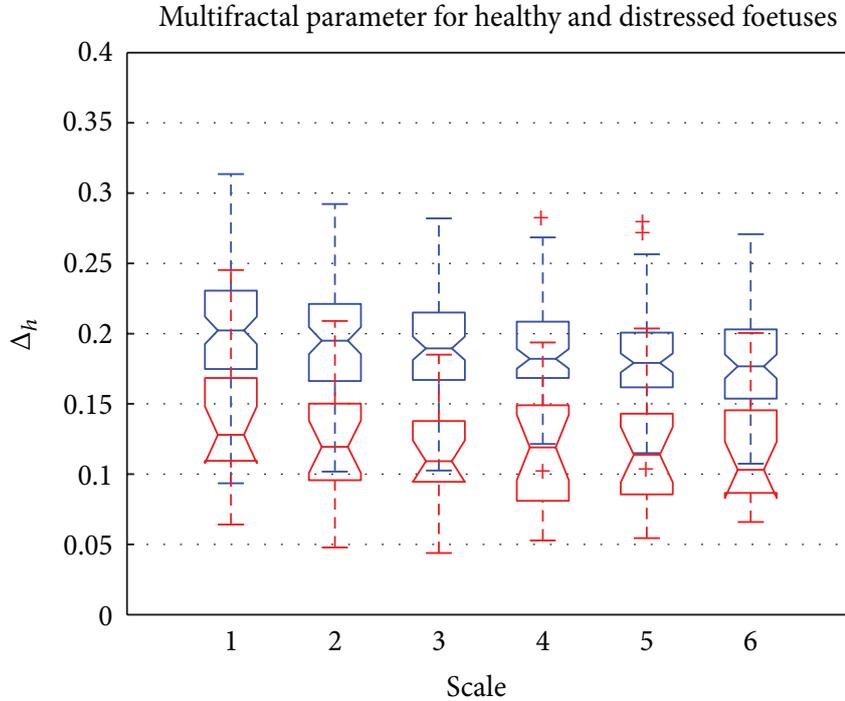


Figure 4.8 – Boxplot of $\Delta_h = \max(h) - \min(h)$ versus the scale for healthy (shown in blue) and distressed (shown in red) fetuses.

reported that as the time series increases, the Hurst exponent increases.

Fig. 4.8 shows a boxplot representation of $\Delta_h = \max(h) - \min(h)$. The parameter's behavior is represented for scale values ranging from 1 to 6. Red boxplots correspond to distressed fetuses and blue boxplots to healthy fetuses. Fig. 4.8 shows that the dynamics of Δ_h for healthy fetuses were higher than that obtained for distressed fetuses. This meant that the signatures for healthy fetuses were more multi-fractal than those obtained for distressed fetuses. This has already been reported in recent studies such as [Ivanov et al., 1999]. Furthermore, Fig. 4.8 shows that there is a sufficient deviation between the two plotted dynamics leading the discrimination between healthy and distressed fetuses by means of Δ_h .

Fig. 4.9 shows a boxplot representation of the mean singularity spectrum D_{mean} . This parameter was represented for scale values ranging from 1 to 6. Red boxplots correspond to distressed fetuses and blue boxplots to healthy fetuses. Fig. 4.9 shows that the values of \bar{D} are higher for distressed fetuses than for healthy fetuses. The signatures of healthy fetuses are more regular than those corresponding to distressed fetuses. Fig. 4.9 shows

4.5. APPLICATION TO FETAL HEART RATES

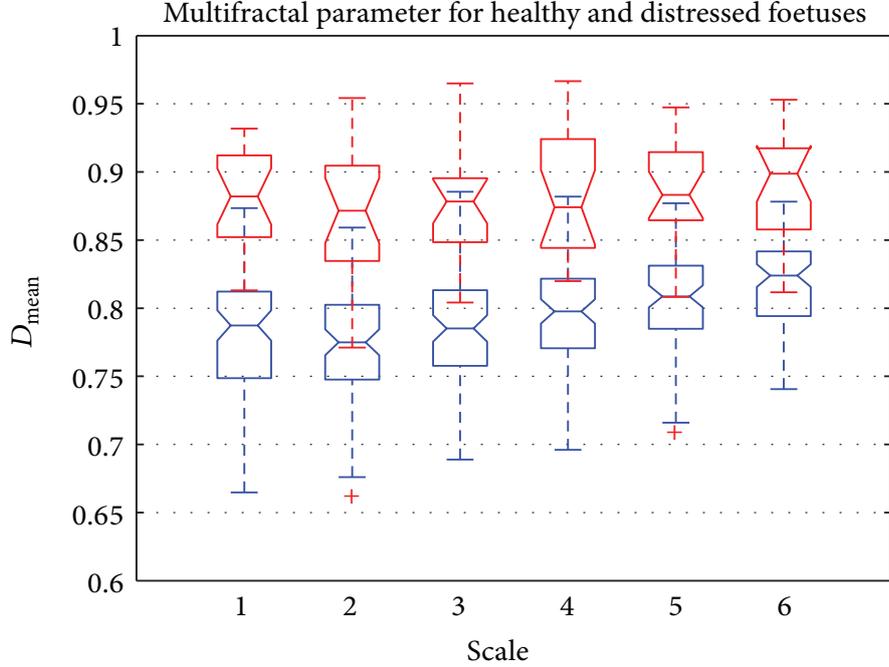


Figure 4.9 – Boxplot of D_{mean} versus the scale for healthy (shown in blue) and distressed (shown in red) fetuses.

also that it is more difficult to discriminate between healthy and distressed fetuses by means of \bar{D} . This parameter did not seem efficient for discrimination purposes. Note that the higher the scale, the lower the RE_3 .

Fig. 4.10 shows a boxplot representation of Δ_D , i.e. the dynamics of D . This dynamic parameter represented for scale values ranging from 1 to 6. Red boxplots correspond to distressed fetuses and blue boxplots to healthy fetuses. Fig. 4.10 shows that Δ_D was higher for healthy fetuses than for distressed fetuses. Hence, the signatures of healthy fetuses were more multi-fractal than those for distressed fetuses. Moreover, Fig. 4.10 shows also that there was a sufficient deviation between the dynamics of healthy and distressed fetuses. Thus Δ_D is a relevant parameter for discrimination purposes.

To sum up, Table 4.1 reports the REs of the four multi-fractal parameters. Table 4.1 shows that the best parameter permitting discrimination among fetuses was RE_4 , followed by RE_1 and then RE_2 . Furthermore, the best differentiation was obtained for a scale value of 2 by RE_4 and RE_1 and a scale value of 3 by RE_2 . This confirms the need to coarse-grain the FHR time series prior to multi-scale analysis. It is obvious from Table 4.1 that as

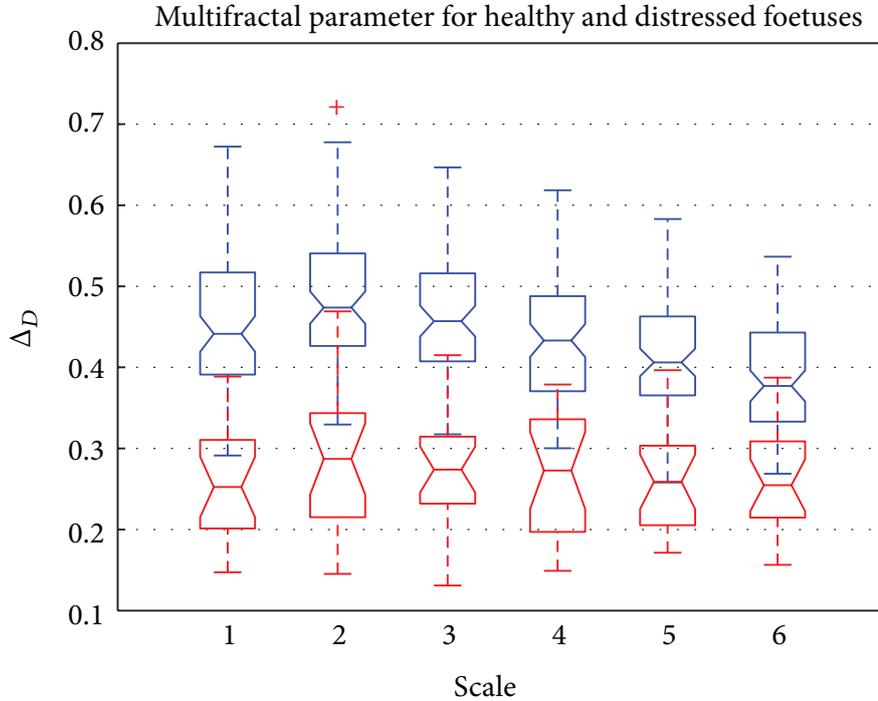


Figure 4.10 – Boxplot of Δ_D versus the scale. Healthy fetus (in blue) and distressed fetus (in red).

Table 4.1 – Relative errors of different multi-fractal parameters between the two groups of fetuses for different scales.

Relative Error \ Scale	1	2	3	4	5	6
RE_1	0.37	0.40	0.33	0.29	0.26	0.24
RE_2	0.32	0.37	0.38	0.37	0.36	0.35
RE_3	0.11	0.12	0.11	0.10	0.09	0.08
RE_4	0.41	0.42	0.41	0.38	0.36	0.33

the scale increases, RE decreases. The latter findings require choosing a maximum scale that is not too high. A value set at 2 seems sufficient regardless of the type of RE used. Furthermore, as the best RE_4 was sensitive to the multi-fractal features of the time series analyzed for a scale of 2, this finding confirms the need to analyze FHR from a coarse-grained multi-fractal point of view. However, as the second discriminative parameter was RE_1 , which is sensitive to mono-fractal features set at a scale of 2, a coarse-grained mono-fractal approach is also relevant.

4.5.2 Discussion

Although the study present in this chapter was quite similar to that presented by Wang et al. [Wang et al., 2005], our study was quite different from two aspects. First, our study was dedicated to fetal discrimination, whereas the work of Wang et al. was dedicated to adults. Second, our study was based on a much simpler structure function than the other approach that was based on a partition function.

In addition, although a large number of research studies have mainly been based on the use of partition functions (DFA, box-counting and wavelet approaches) on the pretext that structure functions do not operate for negative orders, we have shown in this dissertation that the use of such structure functions is fully justified due to the simplicity of implementation, and that structure functions completely fulfil their role in distinguishing between healthy and distressed fetuses.

4.6 Conclusion

Various works using multi-fractal analysis of chaotic and random time-series are available. The study present in this chapter is quite different as it was dedicated to fetal discrimination, and was based on a much simpler structure function than the approach of a partition function. We have shown that the use of the structure function predefined in this chapter is fully justified due to the mathematical and computational simplicity. In this study scaling the time series at 2 was enough to improve the classification, and the mean of the singularity spectrum dynamic parameter and the mean Hurst exponent insured the highest errors between fetuses of different medical states and consequently the ultimate discernment as compared to other parameters. To analyze profoundly the time series, a more general geometric but qualitative tool from which various quantifications can be extracted was the subject of the next chapter, where physical recurrence notion was the core.

Chapter 5

RECURRENCE PLOTS AND ANALYSIS

⌈ *Sweet life events make the person wishes they were continuously recurring, while bitter events make the person wishes they never possessed the physical phenomena of recurrence* ⌋ *Amira Zaylaa.*

5.1 Introduction

Nowadays, the complexity analysis of biomedical time series by means of various descriptors, including, but not limited to, fractal dimension [Gough, 1993], multi-fractal spectra and entropy [Voicu and Girault, 2012b, Pincus, 1991, Oudjemia et al., 2013] introduced in chapters 2 and 3 are quantitative and of standard practice. Recently, vast studies involving the analysis of biomedical signals use recurrence plots (RPs). RPs featured and located in 2-dimension recurring states constituting the time series [Takens et al., 1981, Eckmann et al., 1987, Gao and Cai, 2000, Thiel et al., 2004, Marwan and Kurths, 2002, Marwan et al., 2007]. In order to extract some scalar indicators from RPs, quantitative indicators named Recurrence Quantification Analysis (RQA) [Marwan et al., 2002, Zbilut et al., 2002, Balibrea et al., 2008] have been computed.

One of the significant tasks of *RQA* was to detect the various dynamic transitions of logistic maps [Marwan et al., 2002, Marwan et al., 2007, Iwanski and Bradley, 1998]. In addition, the determinism parameter (*DET*) was employed to quantify the chaotic-periodic and periodic-chaotic transitions [Marwan et al., 2002, Marwan et al., 2007]. The detection of such transitions has been achieved in two different ways. The first way was dedicated to the computation of *RQA* promptly from a single unembedded time

series [Iwanski and Bradley, 1998]. This method was appealing due to its simple mathematical formulation. However, quantification was biased with a poor transition detection rate, due to the presence of sojourn points [Marwan et al., 2002, Zbilut et al., 2002, Balibrea et al., 2008, Gao and Cai, 2000].

The second way was based on embedding the time series obtained from the system in d [Takens et al., 1981, Eckmann et al., 1987, Marwan et al., 2002, Trulla et al., 1996]. Although the previous method required heavy computation, it empirically reduced sojourn points [Marwan et al., 2002, Marwan et al., 2007] and promoted the transition detection. However, to date no one has quantified such detection rate improvement, nor identified if it was possible to enhance the transition detection rate. This leads to the question: what is the best value of the embedding dimension that will allow the detection of all dynamic transitions ?

This work describes our on-going efforts to extract and to cleanly quantify the dynamic information from nonlinear dynamic systems, such as, the logistic map and fetal heart. The goal of this work provided in this chapter is to improve the detection rate of transitions by eliminating sojourn points present in recurrence plots. For the logistic map, the transitions that are meant to be detected are periodic-chaotic and chaotic-periodic transitions, while for the fetal heart the states that are meant to be detected are healthy-distressed and distressed-healthy fetal heart rates (FHRs). The adopted solution consisted of eliminating sojourn points present in recurrence plots. To find an efficient solution to the former problems, several RPs were employed. The first empirical method was composed of two plots having different embedding dimension values. The first plot required no embedding as proposed in [Iwanski and Bradley, 1998] and was named the unembedded RP. The second recurrence plot required an embedding dimension greater than 2 as suggested by N. Marwan and others [Marwan et al., 2002, Marwan et al., 2007]. As the choice of the embedding dimension in the second method is often arbitrary, we developed a new way that consisted of finding out the optimal embedding dimension minimizing the number of sojourn points. Furthermore, we developed two additional signal-based methods using the derivative concept and the m -pattern concept, and compared them with the former plots.

To quantify the level of performances of the proposed approaches, each technique was applied to the nonlinear logistic map and to the FHR. In each case, both the sensitivity and specificity were assessed.

5.2 Recurrence Analysis

RP tool was used to display the recurrences of the states of the time series in 2-Dimensions and to calculate its corresponding RQA [Hurezeanu et al., 2013]. Several Scientists have modified the recurrence test expression by changing the nature of vectors involved, while fixing the tolerance value to ensure an unbiased analysis. Unbiased recurrence plots have been developed in our previous works [Zaylaa et al., 2013, Zaylaa et al., 2014a].

As URPs are cutting edge in FHR analysis, this work aims to evaluate and compare the performance of URPs through an experimental measurement on real Doppler Ultrasound FHR signals.

5.2.1 Recurrence Matrix

An RP is a two-dimensional squared matrix, with black and white dots and two time-axes. Each black dot at the coordinates (i,j) represents a recurrence of the system's state \mathbf{X}_i with another \mathbf{X}_j . It is expressed for a tolerance r as follows:

$$\mathbf{RP} = \Theta(r - \|\mathbf{X}_i - \mathbf{X}_j\|), \quad \mathbf{X}_i \in \mathcal{R}^d, \quad (5.1)$$

where $\mathbf{X}_i \in \mathcal{R}^d$ stands for the points in the phase space at which the system is situated at time i , $\Theta(\cdot)$ is the Heaviside function, $\|\cdot\|$ is the L_∞ norm, $i, j = \{1, \dots, M - d + 1\}$, M is the total number of points and d is the embedding dimension [Eckmann et al., 1987, Marwan et al., 2002] and recently [Marwan et al., 2007, Kantz and Schreiber, 2004]. Note that $\Theta(\cdot)$ plays the role of the correlation sum that was discussed in chapter 1.

5.2.2 Recurrence Quantification Analysis and Our Descriptors

After reproducing the qualitative RP, scalar quantitative parameters are calculated. Of all the existing RQA, *DET* [Marwan et al., 2002, Marwan et al., 2007, Zaylaa et al., 2014b] seems to be the most sensitive for transition detection. The determinism in percentage [Zbilut and Webber, 1992, Nguyen et al., 2012, Manetti et al., 1999] is given as:

$$\%DET = \left(\frac{\# \text{ of Points in diagonal lines}}{\# \text{ of recurring points}} \right) \times 100, \quad (5.2)$$

The novel complexity descriptor that we have proposed in this type of analysis is the Cross-Determinism (% CDET) defined by:

$$\%CDET = \left(\frac{\# \text{ of Points in cross - diagonal lines}}{\# \text{ of recurring points}} \right) \times 100, \quad (5.3)$$

The Percentage of Reduced Sojourn Points *PRSP* in percentage is calculated from %CDET as follows: $(100 - \%CDET)$.

RR which already exists and measures the density of recurrence points in RPs is defined [Marwan et al., 2007, Marwan, 2003, Zbilut and Webber, 1992] in percentage for a given window of size M by:

$$\%RR = \left(\frac{1}{M^2} \# \text{ of recurrent points} \right) \times 100, \quad (5.4)$$

The entropy measure computed from the RP of the time series refers to Shanon's entropy [Zbilut and Webber, 1992, Webber Jr and Zbilut, 2005, Marwan et al., 2007]. It measures the sum of the bin probability of distribution of each diagonal line for each non-zero bin.

$$\%En = - \left(\sum_{l=l_{min}}^M P_{bin} \ln P_{bin} \right) \times 100. \quad (5.5)$$

5.2.3 Sojourn Points Demonstration

In order to describe the challenge of using RPs, which comprised Sojourn Points (SPs), to both discriminate and detect intrinsic dynamic changes of biomedical time series we illustrate the problematic of SPs. SPs are demonstrated starting from a single deterministic signal (Fig. 5.1) and then on a chaotic time series (Fig. 5.2).

Fig. 5.1 presents a scheme of (a) a sine wave, $x(t)$, made up of 200 sample points, its time delayed version, $y(t) = x(t + \tau)$, their elliptical phase space (i.e. $x(t)$ versus $y(t)$) and (b) a sine wave, $x(t)$, its derivative, $\dot{x}(t)$ and their circular phase space (i.e. $x(t)$ versus $\dot{x}(t)$). Note that the size of a single chosen strip is called the tolerance r .

Fig. 5.1 (a) shows the points denoted 1 (red circles) and 3 (black cross) existing within the same strip of size r . In reality, only the points denoted by 1 are truly recurring with each other and not with points denoted by 3. According to the standard recurrence test

5.2. RECURRENCE ANALYSIS

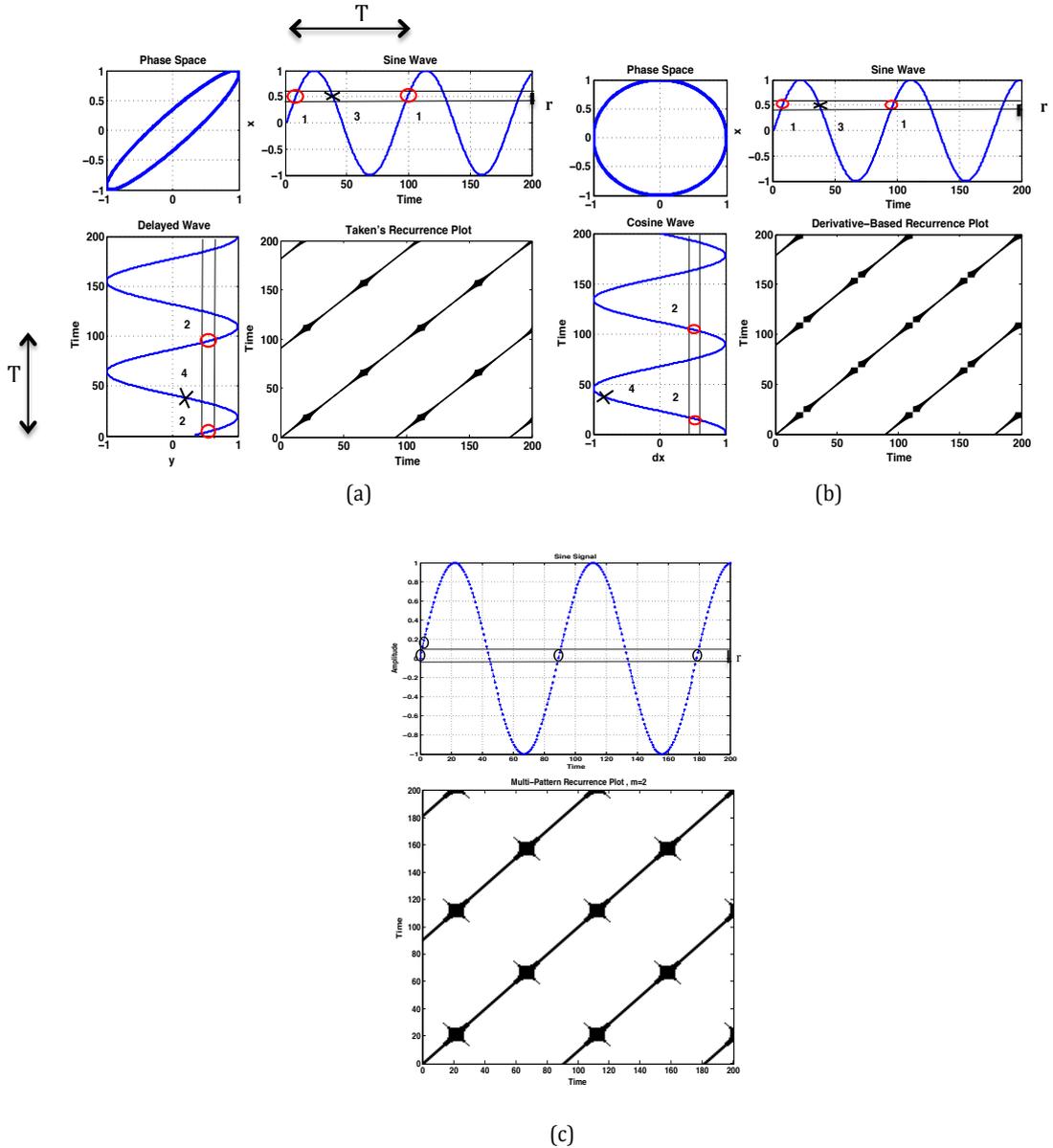


Figure 5.1 – Schematic illustration of sojourn points and recurrence test computation of (a) a sine wave $x(t)$ and $y = x(t + \tau_{optimal})$, (b) a sine wave $x(t)$ and $y = \dot{x}(t)$ and (c) an m -time pattern sine wave.

(of an unembedded time series), all 1's and 3 were recurrences. This was vivid while comparing all the points existing on the same amplitude level of $x(t)$ to those of the delayed signal and confined in tolerance r . By introducing a second signal $y(t)$ (a delayed or a derivative version of $x(t)$), it is now possible to remove sojourn points by comparing points 2 and 4 of $y(t)$ to the corresponding points 1 and 3 of $x(t)$. Point 4 of $y(t)$ which corresponded to 3 in $x(t)$ did not exist at the same amplitude level, whereas all the red

circles were within r .

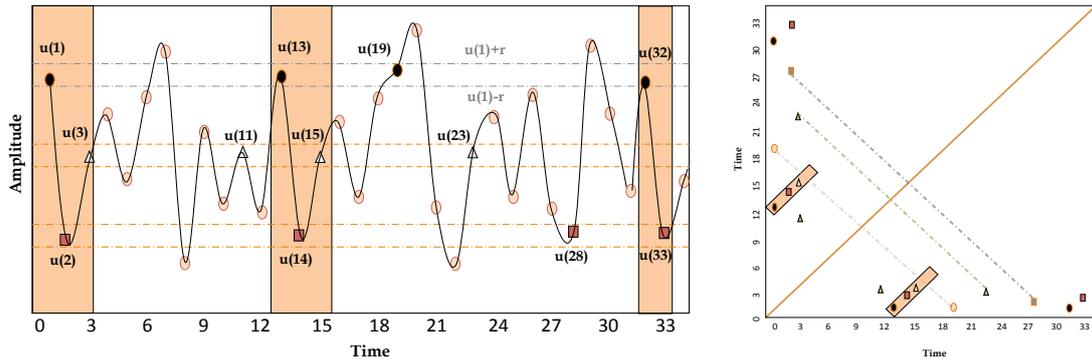


Figure 5.2 – Schematic illustration of sojourn points for a chaotic time series (aperiodic). (left) Chaotic time series. (right) Computation of the corresponding Recurrence Plot.

Consequently, point 3 was a sojourn point since it is not periodic ^a. This tolerance value r is usually 10% of the standard deviation of $x(t)$.

Fig. 5.2 represents how SPs contribute to cross-diagonal lines in the RPs of chaotic time series. Although this case does not correspond to the periodic signal in Fig. 5.1, still SPs constitute cross-diagonals lines. By choosing $m = 1$ the four black circles at $u(1)$, $u(13)$, $u(19)$ and $u(32)$ are within r thus they are classically recurring, however $u(19)$ is not a true recurring point as it correspond to different variation compared to $u(1)$, $u(13)$ and $u(32)$. Analogous to the black circles, the triangle states and the squares are computed in the same way. The resulting RP on the right of Fig. 5.2 shows how true recurrences are basically diagonal lines and false recurrences form mostly cross-diagonal lines.

^aIn 1-Dimension, sojourn points are non periodic points existing within the tolerance r . with 1's. However, in the two signal-based RPs Fig. 5.1 (a) and (b) this was overcome. The tolerance r was fixed in Fig. 5.1 to insure a fair comparison

5.3 Existing Recurrence Plots

5.3.1 Unembedded Recurrence Plots

An unembedded recurrence plot is that produced from a single time series, thus there is no need for sub-time series. Eq. 5.1 becomes:

$$\mathbf{RP}_1 = \Theta(r - \|\mathbf{x}_i - \mathbf{x}_j\|), \quad (5.6)$$

where \mathbf{x}_i stands for the time series points at time i . As reported by Iwanski and others [Iwanski and Bradley, 1998], when $d = 1$, Eq. 5.1 is called unembedded RP. This plot was denoted throughout this work by \mathbf{RP}_1 (see Table 5.4).

5.3.2 Embedded Recurrence Plots

Originally developed by Eckmann [Eckmann et al., 1987], embedded plots have been used to track recurrences of systems' states out of a reconstructed phase space of d -embedding dimension. This was fulfilled using the embedding theorem [Trulla et al., 1996, Takens et al., 1981, Marwan et al., 2002, Zbilut et al., 1998, Packard et al., 1980]. The reconstructed RP was obtained by calculating a time delay τ and embedding dimension d using the mutual information (M.I.) [Fraser and Swinney, 1986] and the false nearest neighbour (F.N.N.) [Kennel et al., 1992, Kantz and Schreiber, 2004] methods, and subsequently computing Eq. 5.1, given that $\mathbf{X}_j = \mathbf{X}_{i+\tau}$ and $d \geq 2$.

Analogous to previous investigations, d has been fixed to 3, i.e. three sub-time series produced from $x(t)$ were used to reconstruct RP [Trulla et al., 1996, Marwan et al., 2002, Marwan et al., 2007, Ahlstrom et al., 2006]. The corresponding RP was denoted by \mathbf{RP}_2 (see Table 5.4). Note that such plot has been recommended to eliminate sojourn points by using an embedding dimension of $d \geq 2$ [Gao and Cai, 2000, Ahlstrom et al., 2006]. It works well on stable and low noise systems. However, when the system is non-stationary with much noise, such as the biomedical systems, it becomes not suitable for estimating d . According to Webber [Webber and Zbilut, 1994] d is not appropriate to be set too high or too low on the biological systems.

The second parameter to be set for \mathbf{RP}_2 is the time delay τ . It should be selected for minimizing the interaction between the points of the measured time series. Fraser and Swinney reported that the first minimum of MI can provide the best available systematic

criterion for choosing time delays for phase portraits [Fraser and Swinney, 1986]. However it should not be too high, otherwise RQA cannot provide the enough information to distinguish the kinds of signals [Yan et al., 2010]. In RP_2 , the time delay was selected from the mutual information.

The other parameter in RQA which needs to be set is the minimal diagonal line. N. Marwan, who developed the software to calculate the RQA measures [Marwan, 2009], specified a default value. But, we chose in this work to set it after testing a range of values and tracking the minimum possible number of sojourn points that could show up. According to our previous works [Zaylaa et al., 2013, Zaylaa et al., 2014a] the tolerance r was set at 10% of the error of the fetal signals.

5.4 Developed Recurrence Plot Methods

5.4.1 Embedded Recurrence Plot With Specific Settings

Herein we provide our first signal-based RP that is responsible for enhancing the detection of dynamic transitions. In this technique, instead of looking for the best embedding dimension and time delay that guarantee the independence of sub-time series, as already done by Trulla et al. [Trulla et al., 1996] and later on by N. Marwan et al. [Marwan et al., 2002, Marwan et al., 2007], we looked for the best value of the embedding dimension and time delay that minimized the presence of sojourn points in an RP. Both $d=d_{optimal}$ and $\tau=\tau_{optimal}$ were obtained by minimizing a cost function ($J(d_k, \tau_k) = CDET$) characterizing sojourn points.

In contrary to DET where $P(l)$ was the number of diagonal lines, $CDET$ was defined by Eq. 5.2 for $P(l)$ being the number of cross-diagonal lines of length l . It quantifies the number of points constituting the diagonal lines perpendicular to the line of identity (LOI). Instead of developing a new algorithm to assess $CDET$, it was obtained by Eq. 5.2 computed from the 90° rotation of the recurrence matrix. The optimization problem was given as follows:

$$\varphi_k^* = \arg \min_{\varphi} (J(\varphi_k)), \quad (5.7)$$

where $\varphi_k=(d_k, \tau_k)$ is a vector composed of the k^{th} embedding dimension and the k^{th} time delay to be optimized. An example of such optimization is presented in Fig. 5.3.

The embedding dimension was chosen after minimizing a cost-function describing the number of SPs in RPs. In this dissertation $d_{optimal}$ was set at 2 as defined in our previous work [Zaylaa et al., 2013, Zaylaa et al., 2014a]. We denoted the resulting RP by **RP**₃ (see Table 5.4).

5.4.1.1 Examples of The Optimization of Both The Embedding Dimension and Time Delay

Fig. 5.3 (top left) represents the cost function versus d , while $\tau = T/4$ with T being the period of the signal as indicated theoretically. τ could be also calculated by the autocorrelation. Fig. 5.3 (center left) shows the first zero-crossing of J or $J_{minimum}$ occurring at $d = 2$. Consequently, d was set to 2 in the optimization. Fig. 5.3 (bottom left) exhibits the cost function versus the time delay, while holding d fixed at 2. The parameter τ was tested in the following range $0 \leq \tau \leq 55$. This led to an optimized range rather than a unique value, $\tau \in \{6,14\} \cup \{26,34\} \cup \{46,54\}$ time unit. By analogy to the sine wave, d and τ were optimized for the RP of both the logistic map and the FHRs. For the logistic map, d was set to 2 and τ to 4. For the healthy FHR, Fig. 5.3 (center right) shows that J_{min} occurred at $d = 2$. The tested range of d was from 2 to 7 embedding dimensions. Fig. 5.3 (bottom right) exhibits that J_{min} occurred at $\tau = 1$ time unit.

Instead of assessing the best embedding dimension and time delay, another RP has been developed and discussed in the upcoming subsection.

5.4.2 Derivative-Based Recurrence Plot

Herein we provide our second signal based RP method. Starting from a single time series, and instead of considering the independence between two sub-time series, we considered the orthogonal data by involving the derivative principle (see the example in Fig. 5.1 (b)). This approach neither required looking at the time delay nor embedding dimension. We have denoted the resulting plot by **RP**₄ (see Table 5.4).

5.4.3 M -Time Pattern Recurrence Plot

After introducing the derivative-based RPs, we provide our third nonlinear technique which is the m -time Pattern Recurrence Plot. This is a multi-pattern algorithm. However, in this study it was used as a 2-point to 2-point recurrence computation. Fig. 5.1 (c)

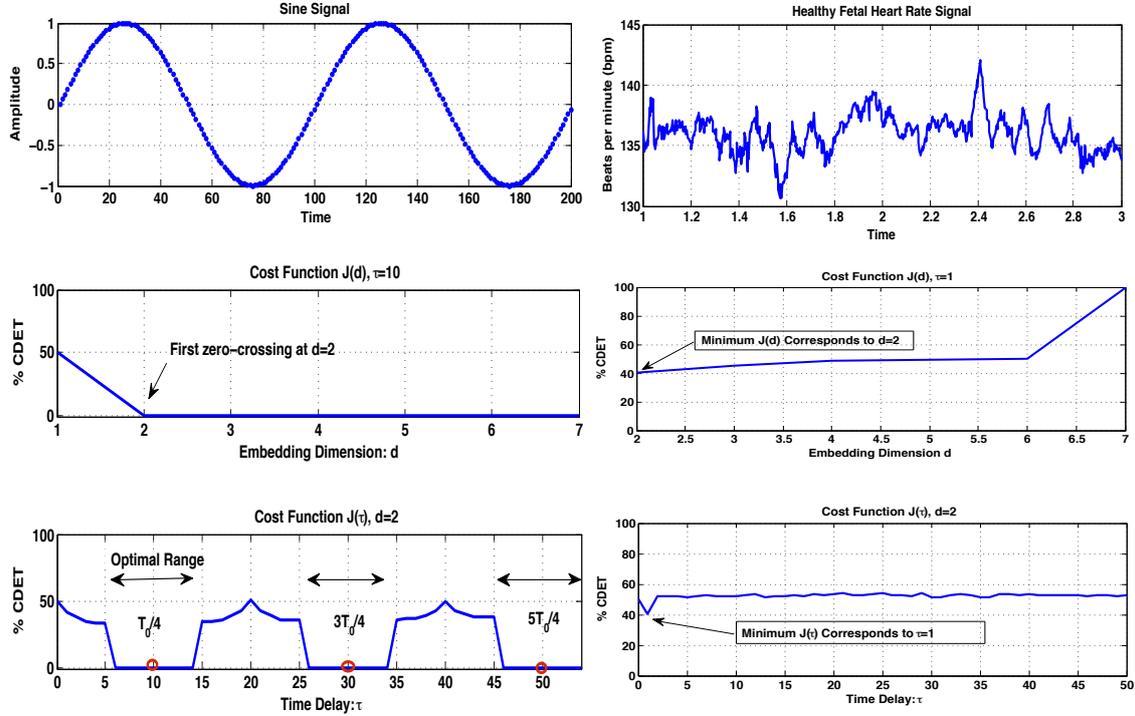


Figure 5.3 – The cost-function and optimization process of the embedding dimension and time delay (left) for a sine wave and (right) for the healthy FHR.

represented the 2-point to 2-point computations. For $m = 2$, the m -time Pattern RP, i.e. ($RP(m = 2)$) was denoted throughout this dissertation by \mathbf{RP}_5 .

Table 5.4 summarizes the vectors involved in Eq. 5.1, d the embedding dimension, m the number of pattern and d' the virtual embedding dimension used in all the recurrence plots discussed in this chapter.

Note that the flow charts of the three developed URPs are summarized in Fig. 5.4.

5.4.4 Example on Sine Wave

We illustrate in Fig. 5.5 an evaluation of the drop off in sojourn points for a sine wave displayed within a window of 200 points (a) as compared to the unembedded RP_1 (b). It is an example of a simulated sine wave and its corresponding standard and developed recurrence plots. Fig. 5.5 (a) sets out a sine wave and its corresponding (b) unembed-

5.4. DEVELOPED RECURRENCE PLOT METHODS

Table 5.1 – The five nonlinear recurrence methods used in the complexity analysis for transition detection.

Method	Notation	Time Series	d	m	d'
Unembedded RP	RP ₁	$X_j = x_j$	1	1	1
Embedded RP	RP ₂	$X_j = X_{i+\tau}$	≥ 2	1	≥ 2
Embedded RP with specific settings	RP ₃	$X_j = X_{i+\tau_{opt}}$	2	1	2
Derivative-Based RP	RP ₄	$X_j = \dot{X}_i$	-	1	2
M -Time Pattern RP	RP ₅	X_i, X_j	-	2	2

ded recurrence plot RP₁, (c) embedded recurrence plot RP₃ with specific settings, (d) derivative-based recurrence plot RP₄ and (e) 2-time pattern recurrence plot RP₅.

The plots in Fig. 5.5 are arranged as follows: RP₁, RP₃ for ($\tau = T/4$) where T is the period of the sine wave, RP₄ and RP₅. Fig. 5.5 (b) demonstrates a few diagonal and cross-diagonal lines revealing periodicity and uniformity, the crossing data being due to the presence of sojourn points. Fig. 5.5 (c) merely reveals diagonal lines in RP₃ and the cross-diagonals which showed up in RP₁ have vanished completely. This was due to embedding the time series.

Moreover, RP₄ manifested how the cross-diagonal lines totally vanished. However, junction points connecting both diagonal and cross-diagonal lines vanished (see Fig. 5.5 (d)) at $\tau = 30$ time units. This is due to the fact that the positions of crests and troughs in the sine wave (a) did not correspond to the same positions in the derived signal. Thus points constituting the crests and troughs did not exist within the same tolerance r , and hence black points disappeared. This could be due to using the approximate derivative.

In Fig. 5.5 (e) RP₅ shows how the cross-diagonal lines vanished differently. A few points lying on the cross-diagonal lines persisted, as compared to RP₃ and RP₄. This might have been due to the fact that the time delay was not set according to a formulation, but rather it was by default equal to one. The time delay was less than $T/4$ (as for RP₃). The two signals adhered, consequently little cross-diagonal information showed up.

In Fig. 5.5 (e), RP₅ shows how the cross-diagonal lines have vanished differently. Few points lying on the cross-diagonal lines persisted as opposed by RP₃ and RP₄. This could be due to the fact that the time delay was not set according to a formulation, rather it was by default equals to one. The time delay was less than $T/4$ (as for RP₃). The two

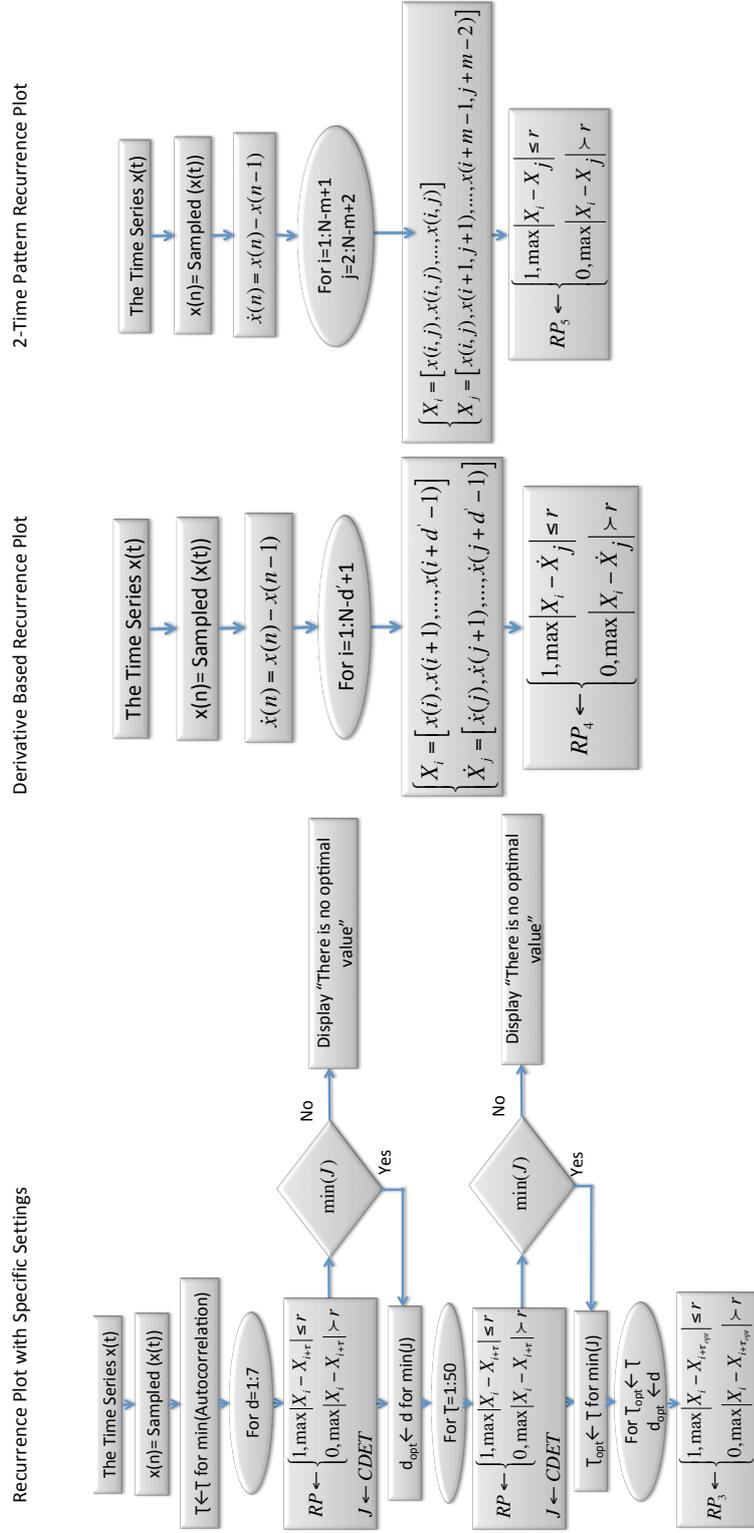


Figure 5.4 – The flowcharts of the three signal-based recurrence plots. (left) The flowchart of the Recurrence Plot with specific settings RP_3 . (center) The derivative base Recurrence Plot RP_4 . (right) The 2-time Pattern Recurrence Plot RP_5 . N here is the same as M the length of the signal.

5.4. DEVELOPED RECURRENCE PLOT METHODS

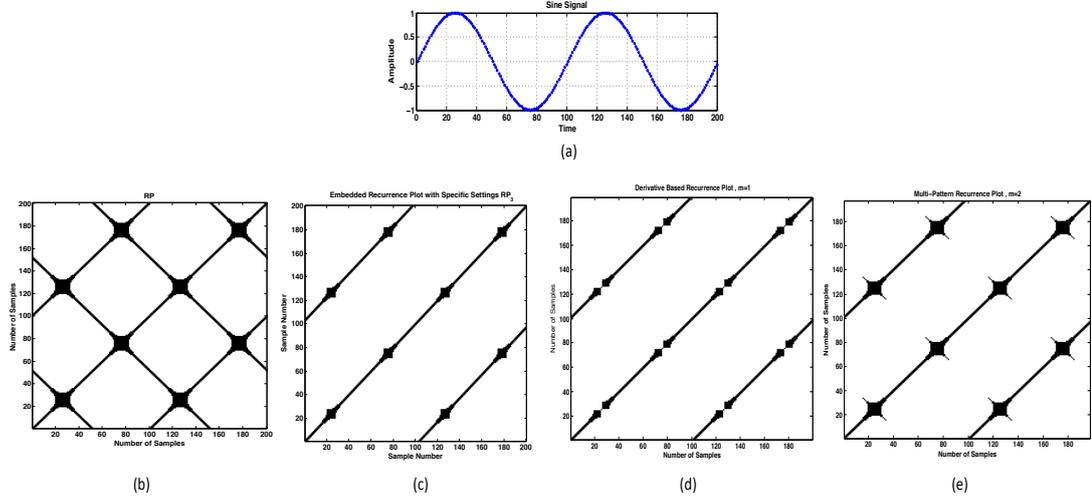


Figure 5.5 – Example of a simulated sine wave and its corresponding standard and developed recurrence plots. (a) Sine wave and its corresponding (b) unembedded recurrence plot (RP_1), (c) embedded recurrence plot with specific settings (RP_3), (d) derivative-based recurrence plot (RP_4) and (e) 2-time pattern recurrence plot (RP_5).

signals were adhering, consequently few cross-diagonal information showed up.

As a global visual evidence, sojourn points have been eliminated using RP_3 , RP_4 and RP_5 . The evaluation of the performance of the transition detector and sojourn point reduction was considered for the biologically inspired logistic map and for the real data of healthy and distressed FHRs. For the logistic map, each detector operating on the DET parameter was evaluated for the five different RPs over different values of the control parameter b . Both a qualitative RP representation and a quantitative (the rate of the reduced sojourn points, sensitivity and specificity) evaluations were obtained. For FHRs, three parameters: DET , $CDET$ and $PRSP$ were evaluated from the five RPs over 30 minutes FHRs split into 10 windows of 3 minute each. The three statistical measures: mean, Standard Deviation (SD) and the Relative Error (RE) were used to evaluate the fetal transition detection. Simulations were obtained using Matlab (Mathworks, Natick, MA, USA).

5.4.5 Results of Logistic Map

For each b value of the Logistic Map defined in chapter 2, a time series of length $M = 1000$ has been computed with a step of $\Delta b = 0.0005$.

DET values were recorded relative to b and the detection of systems state transitions has been evaluated qualitatively and quantitatively in the next subsection. Afterwards, the time series corresponding to $b = 4$, i.e. the chaotic region, was simulated as suggested by [Marwan et al., 2002, Marwan et al., 2007] and involved in the upcoming illustration of sojourn point reduction. The preference of $b = 4$ maintained the maximum rate of system's growth (i.e. chaoticity) [Marek, 1995].

5.4.5.1 Qualitative Evaluation Of Sojourn Point Reduction

As there is a major interest in reducing the number of sojourn points, N. Marwan and others [Marwan et al., 2002, Marwan et al., 2007] suggested that calculating a time series for fixed value of b is better for this purpose. Furthermore, to illustrate the advantages of our signal-based recurrence plots over existing plots, we simulated RP_1 and RP_2 for ($b = 4$) [Marwan et al., 2002, Marwan et al., 2007]. Fig. 5.6 represents the (a) unembedded RP_1 , (b) embedded RP_2 and three signal based recurrence plots arranged as follows: (c) RP_3 for $\tau = 4$, (d) RP_4 and (e) RP_5 . On each plot a zoomed portion of the RP was imposed to highlight different cross-diagonal lines, in a red color, that seemed to present the sojourn points, and were involved in the further quantitative evaluation. RP_1 of the simulated time series $x(t)$ showed various small diagonal and cross-diagonal lines (i.e diagonals perpendicular to the main diagonal). The previous features reflected the chaotic nature of the simulated system. RP_2 resembled that simulated by N. Marwan and others [Marwan et al., 2002, Marwan et al., 2007] for $b = 4$.

RP_3 showed the least number of recurrences. It revealed that the major cross-diagonal lines have vanished, also, significant vertical lines have vanished too. This could be due to the choice of $d_{opt} = 2$ based on the minimum number of sojourn points that could be attained. RP_4 elaborated how the cross-diagonal lines, which existed in the classical image RP_1 and in RP_2 , have vanished, whereas extra small structures and cross-diagonals showed up instead.

RP_5 revealed that there was a remarkable loss in the vertical information as compared to RP_4 , especially in the small cross-diagonals. Thus sojourn points were reduced as a consequence of the detection techniques, as it was obvious in Fig. 5.6 (c), (d) and (e).

5.4.5.2 Quantitative Evaluation Of Sojourn Point Reduction

As N. Marwan et al. [Marwan et al., 2002, Marwan et al., 2007] previously introduced that DET parameters extracted from RPs seemed to be a good indicator for detecting

5.4. DEVELOPED RECURRENCE PLOT METHODS

dynamic changes we made use of them.

Both *CDET* and *DET* given by Eq. 5.2 for l characterizing cross-diagonals and diagonals, respectively, have been evaluated in percentage from the RPs of the simulated Logistic time series. Quantitative results were depicted in Table 5.2. The tolerance r was initially chosen to be 10% the standard deviation of the time series as suggested in [Marwan et al., 2002, Marwan et al., 2007] to insure a noise-free situation.

Table 5.2 shows that *DET* changes from 78% by RP_2 to 73% by RP_3 , i.e. the number of points detected along the diagonal decreases slightly. This could be due to some large diagonal lines that were transformed into separate discrete points to reinforce the chaotic nature of the logistic map for ($b = 4$). Table 5.2 shows also that the number of points detected along the cross-diagonal lines increases. The straight-forward comparison of the cross-diagonal values relative to the detected points forming the cross-diagonal lines obtained through the reference method, RP_2 , shows that (i) 27% out of 22% cross-diagonal points are detected in RP_3 , (ii) 13.3% out of 22% cross-diagonal points are detected in RP_4 , however, (iii) 13% out of 22% cross-diagonal points are detected in RP_5 .

As a result, the remaining percentages of SPs (% CDET out of % CDET $_{RP_2}$) are: (i) 6%, (ii) 3% and (iii) 2.9% in RP_3 , RP_4 and RP_5 , respectively. Consequently, %PRSP are: (i) 94%, (ii) 97% and (iii) 97.1% in RP_3 , RP_4 and RP_5 , respectively. Table 5.2, shows that the performance of the elimination of SPs can be arranged as follows: $RP_5 > RP_4 > RP_3 > RP_2$.

Table 5.2 – % DET and % CDET quantifications of RP_1 , RP_2 , RP_3 and RP_4 for a logistic time series for $b = 4$.

Dynamic System	Logistic Time Series			
	% DET	% CDET	% CDET out of % CDET $_{RP_2}$	%PRSP
Unembedded RP [RP_1]	0.0%	100.0%	22.0%	0.0%
Embedded RP [RP_2]	78.0%	22.0%	-	-
Embedded RP with specific settings [RP_3]	73.0%	27.0%	6.0%	94.0%
Derivative-Based RP [RP_4]	86.7%	13.3%	3.0%	97.0%
<i>M</i> -Time Pattern RP [RP_5]	87.0%	13.0%	2.9%	97.1%

CHAPTER 5. RECURRENCE PLOTS AND ANALYSIS

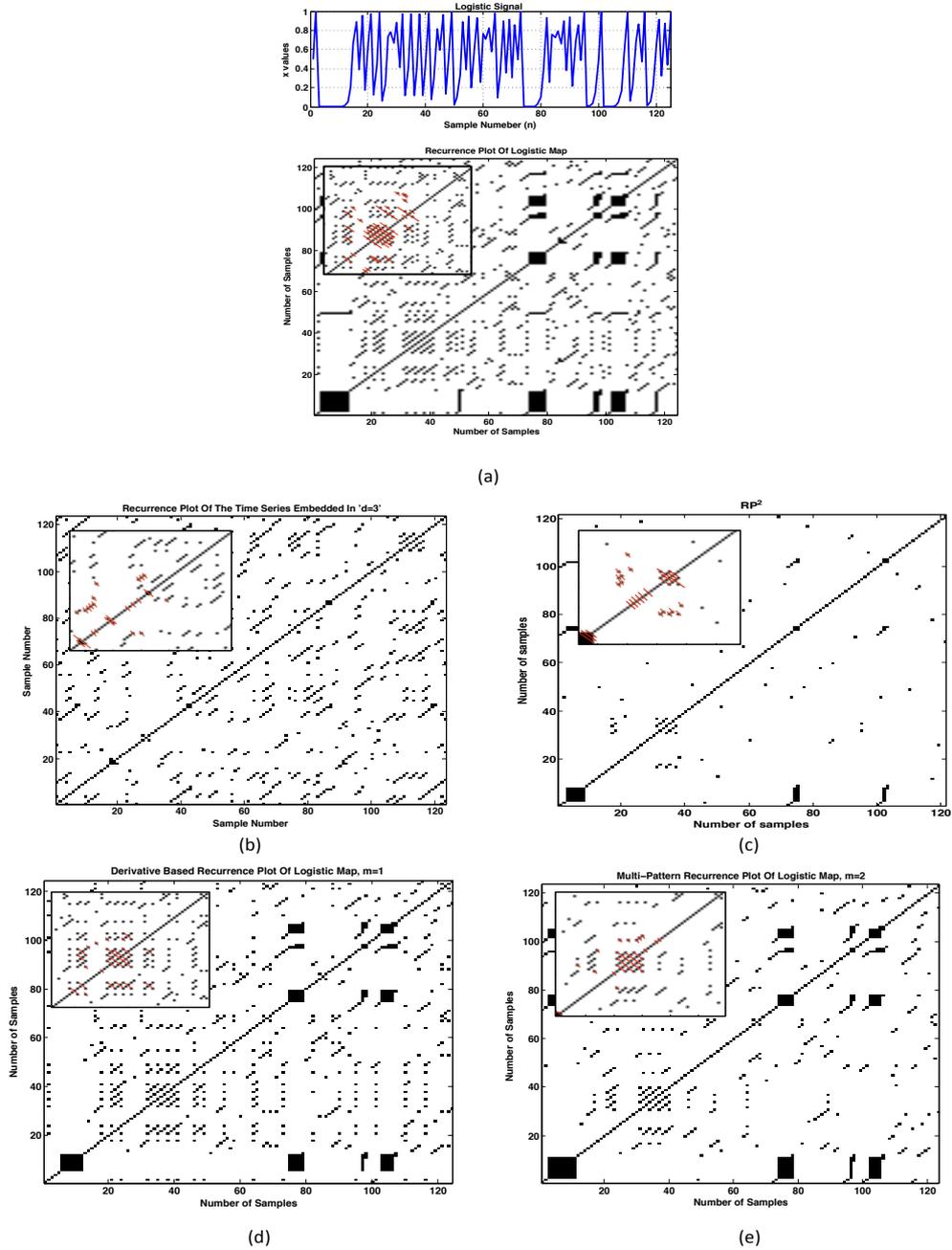


Figure 5.6 – The standard and developed recurrence plots, originating from a simulated Logistic time series for $b = 4$. (a) the unembedded recurrence plot (RP_1), (b) embedded recurrence plot (RP_2), (c) embedded recurrence plot with specific settings RP_3 , (d) derivative-based recurrence plot (RP_4) and 2-time recurrence plot (RP_5).

5.4.5.3 Detection of Dynamic Transitions

To evaluate the performance of the three signal-based recurrence plot methods in detecting dynamic changes relative to already existing techniques, simulations were performed and reported in Fig. 5.7. Fig. 5.7 depicts the variation of the normalized determinism (shown in blue) as a function of the control parameter of the logistic map. The quantification parameter was computed from (a) RP_1 (b) RP_2 , (c) RP_3 , (d) RP_4 and (e) RP_5 , respectively. The three developed methods exhibited the major peaks present in the bifurcation diagram (diagonal information has been preserved).

In Fig. 5.7, the results of DET versus b of the developed techniques imitate that of RP_2 but are not the same. For instance, the amplitudes of peaks are not strictly equal to 1 as in RP_2 . This triggered us to choose another detection criteria to discern between transitions. Although tracking unity to distinguish between transitions was effective in RP_2 , the number of peaks detected were not sufficient as compared to our developed unbiased recurrence plots and the bifurcation diagram.

Despite this evaluation, a quantitative evaluation of the dynamic transition detection was necessary to provide a solid interpretation. Statistical measures were used to evaluate the performance of detection. On each DET versus b displayed in Fig. 5.7 a constant threshold λ is applied and used as a detector. The value of λ was chosen from an empirical inference, i.e. after trying several values then choosing the one permitting the detection of the maximum number of peaks. Four cases were defined in the detection process: i) True Positive, TP , represented the number of detected peaks/transitions in the right place (where there was a real transition), ii) False positive, FP , exhibited the number of detected peaks in the wrong place (where there was no transition), iii) False negative, FN , reflected the number of undetected peaks in the right position and iv) True Negative, TN , represented the number of undetected peaks in the wrong position. Finally, both the $Sensitivity = \frac{TP}{TP+FN}$ and $Specificity = \frac{TN}{TN+FP}$ were calculated.

Based on Fig. 5.7 we carried out our quantitative evaluation of detection provided the exact position of each transition (see Fig. 2.2). Table 5.3 reports the performance of detection (PD), i.e. the percentages of both sensitivity and specificity of detecting dynamic transitions.

The values depicted in Table 5.3 were calculated for $b \in [3.64, 4]$, and the total number of transitions was 11. According to Table 5.3, both m -time pattern and embedded recurrence plots with specific settings possess both the ultimate sensitivity and specificity,

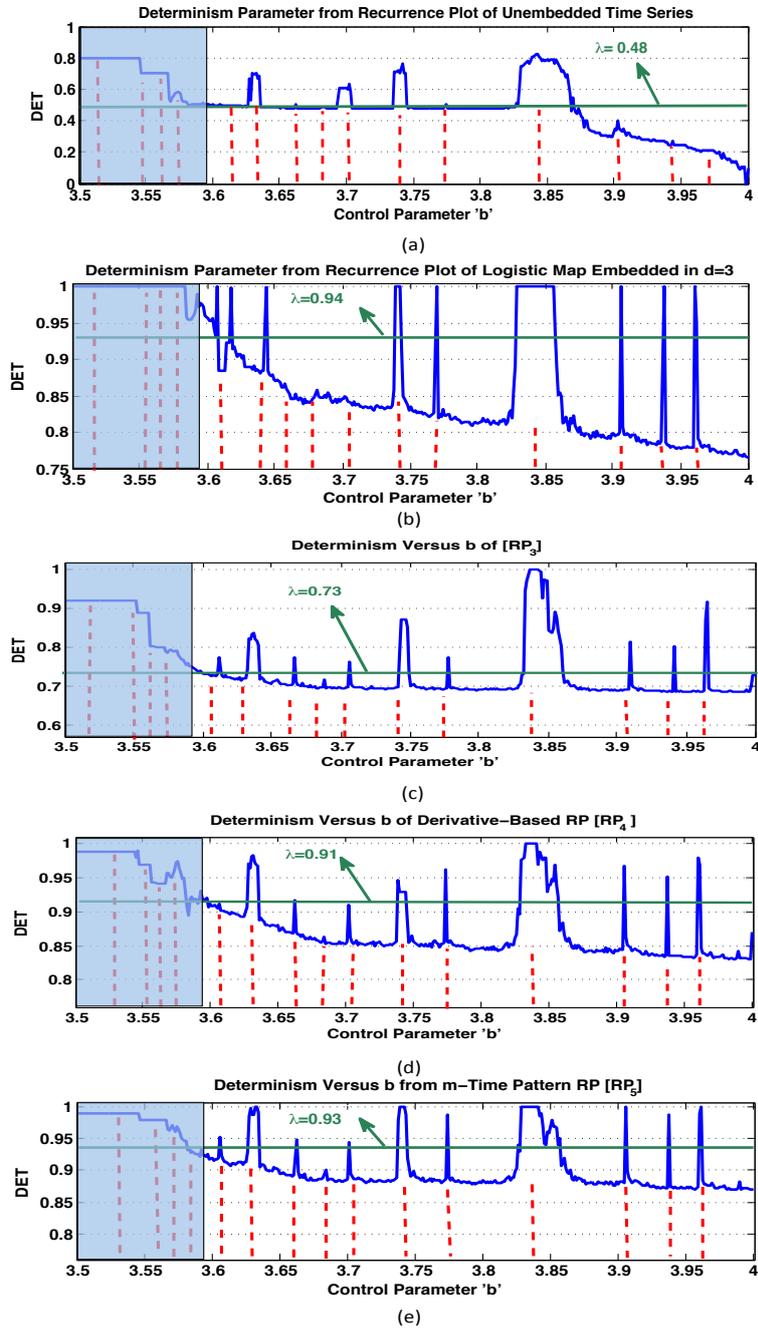


Figure 5.7 – Qualitative detection of dynamic transitions of the logistic map through thresholding the determinism versus the dynamic parameter, originating from both the standard and signal-based recurrence plots. (a) DET computed from the unembedded recurrence plot (RP_1), (b) the embedded recurrence plot (RP_2), (c) the embedded recurrence plot with specific settings (RP_3), (d) the derivative-based recurrence plot (RP_4) and (e) the 2-time recurrence plot (RP_5).

i.e. 91% and 100%, respectively. In particular, the relative percentage of differences between the sensitivities of detection was computed.

5.5. APPLICATION TO FETAL HEART RATES

For instance, an increase of 25% ($(91 - 73)/73 \simeq 25\%$) in sensitivity of detection was obtained by the m -time pattern and embedded recurrence plot with specific settings over the embedded recurrence plot (RP_2). Also, an increase of 52% ($(91 - 60)/60 \simeq 52\%$) in sensitivity of detection was obtained by the m -time pattern and embedded recurrence plot with specific settings over the unembedded recurrence plot (RP_1).

From Table 5.3, it was clear that the levels of performance, in terms of sensitivity and specificity, were arranged as follows: $RP_5 = RP_3 > RP_4 > RP_2 > RP_1$. This evidenced that we could eliminate sojourn points without the requirements of embedding in high dimension.

Table 5.3 – Performance of Detection (PD) of the logistic map dynamic transitions by means of the sensitivity and specificity measures.

Technique \ PD	Sensitivity	Specificity
RP₁	60%	100%
RP₂	73%	100%
RP₃	91%	100%
RP₄	80%	100%
RP₅	91%	100%

5.5 Application To Fetal Heart Rates

Of all the FHR database available, the URPs were applied on 25 H-FHRs and 25 D-FHRs to amend the diagnosis of IUGR.

5.5.1 Results of Diagnosing Fetal Heart Rate Distress

This subsection provides both the qualitative and the quantitative results of the evaluation of the diagnosis of IUGR.

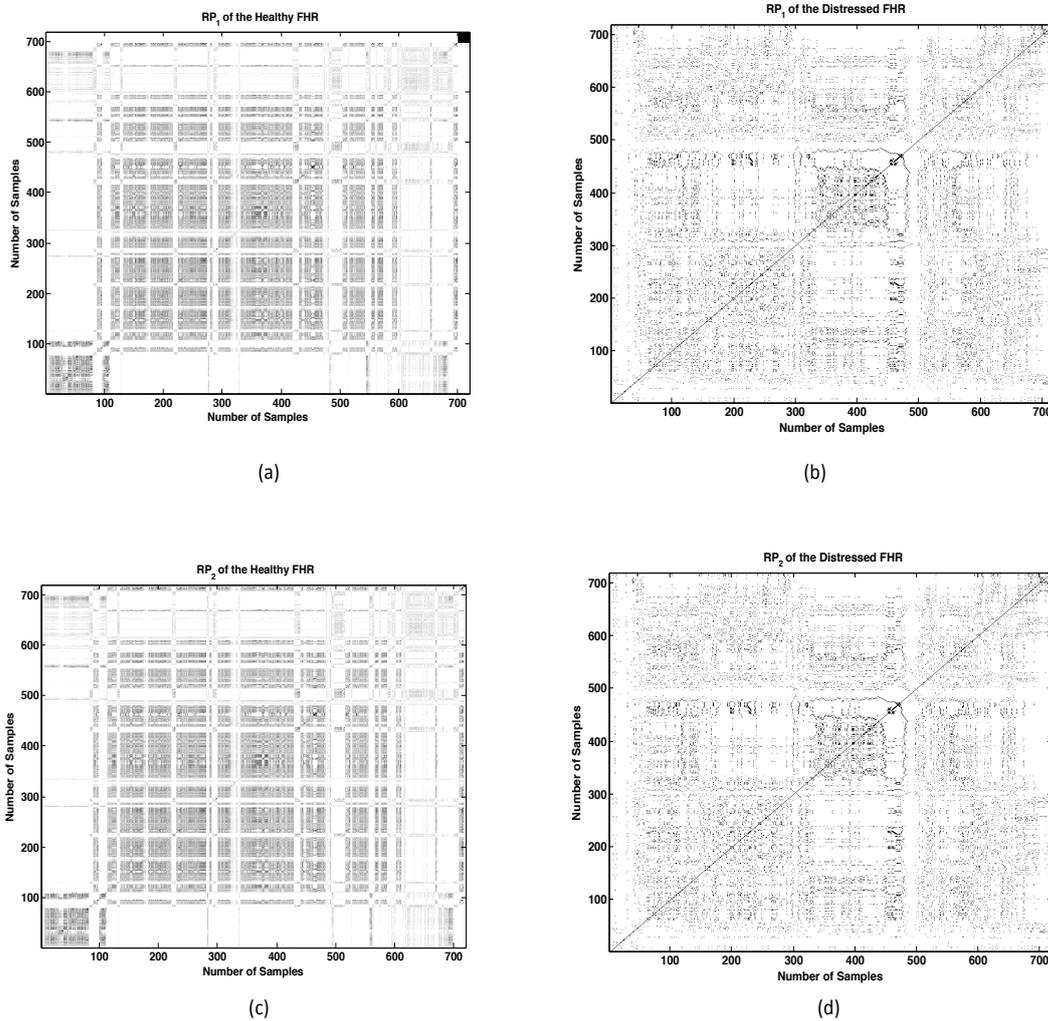


Figure 5.8 – The Standard Recurrence Plots of FHRs. (a) The unembedded recurrence plot RP_1 of the H-FHR. (b) The unembedded recurrence plot RP_1 of the D-FHR. (c) The $2-d$ embedded recurrence plot RP_2 of the H-FHR. (d) The $2-d$ embedded recurrence plot RP_2 of the D-FHR.

5.5.1.1 Qualitative evaluation of diagnosis

Figs. 5.8 (a) and (c) represent RP_1 and RP_2 of a H-FHR, respectively, and Figs. 5.8 (b) and (d) shows RP_1 and RP_2 of a D-FHR, respectively. Patterns present in Figs. 5.8 (a) and (c) differ from the patterns present in Figs. 5.8 (b) and (d).

Fig. 5.9 shows the three simulated URPs. Fig. 5.9 (a) represents 3-d embedded recurrence plot with specific settings RP_3 of H-FHR, (b) the 3-d embedded recurrence plot with specific settings RP_3 of D-FHR, (c) the derivative-based recurrence plot RP_4 of the H-FHR, (d) the derivative-based recurrence plot RP_4 of the D-FHR, (e) The m-time

5.5. APPLICATION TO FETAL HEART RATES

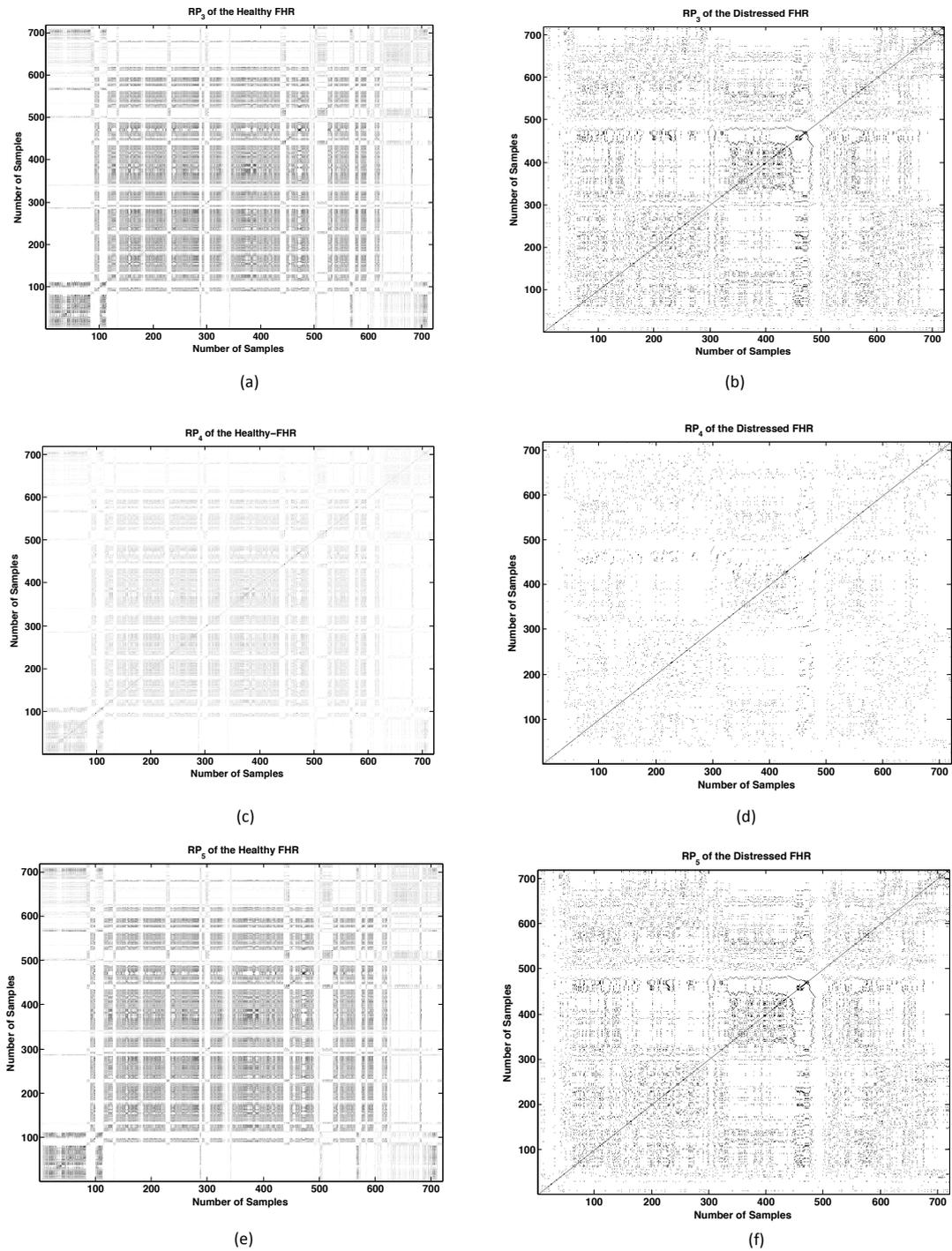


Figure 5.9 – The Unbiased Recurrence Plots of FHRs. (a) The 3 – d embedded recurrence plot with specific settings RP_3 of H-FHR. (b) The 3 – d embedded recurrence plot with specific settings RP_3 of D-FHR. (c) The derivative-based recurrence plot RP_4 of the H-FHR. (d) The derivative-based recurrence plot RP_4 of the D-FHR. (e) The m -time pattern recurrence plot RP_5 of the H-FHR. (f) The m -time pattern recurrence plot RP_5 of the D-FHR.

pattern recurrence plot RP_5 of the H-FHR and (f) the m-time pattern recurrence plot RP_5 of the D-FHR. Note that the mean diagonal length is higher in case of the distressed fetus, and LOI is more dense in case of the distressed fetus. This meant that the cardiac activity of the distressed fetus is more predictable than that of the healthy one. A greater predictability is possible when a loss in complexity of the heart system appears. The previous fact could be triggered by the fragile interaction between brain and heart of the fetus caused by hypoxia.

5.5.1.2 Quantitative evaluation of diagnosis

To evaluate the performance of diagnosis or discrimination of the fetus medical state while taking the full advantage of the nonlinearity of FHR time series, both the Relative Error (RE) and the Standard Deviation (SD) were calculated as in [Hazewinkel, 2001]. RE is given as:

$$\%RE(RQA_i) = \left(\frac{RQA_i^{(H-FHR)} - RQA_i^{(D-FHR)}}{RQA_i^{(H-FHR)}} \right) \times 100, \quad (5.8)$$

where $i = \{1, \dots, 10\}$ and RQA's are DET, CDET, PRSPs, RR and En.

SD which measures the amount of RQA_i dispersion from its average value is given as follows:

$$\mathbf{SD}(RQA_i) = \left\{ \begin{array}{l} \sqrt{\frac{1}{M} \sum_{i=1}^M \left(RQA_i^{(H-FHR)} - RQA_i^{(D-FHR)} \right)^2}, \\ \pm \delta SD(RQA_i) \end{array} \right., \quad (5.9)$$

where M is the total number of FHRs used.

Fig. 5.10 sets out the five recurrence quantification parameters computed from the recurrence plots described in sections 5.3 and 5.4. Fig. 5.10 (a) shows the average of DET(%) as a function of the unembedded RP, embedded RP in $3-d$, embedded RP with specific settings, derivative-based RP and 2-time pattern RP, RP_1 , RP_2 , RP_3 , RP_4 and RP_5 , respectively. Fig. 5.10 (b) represents the average CDET(%) relative to the five techniques. Fig. 5.10 (c) shows the average PRSP(%) as a function of the five RPs. Fig. 5.10 (d) represents the average RR(%) as a function of the five RPs. Finally, Fig. 5.10 (e) shows the average En(%) as a function of the standard and unbiased RPs. The gray color is associated with H-FHRs and the orange color is associated with D-FHRs, and REs in percentage were reported between the two different bar graphs of each RP

5.5. APPLICATION TO FETAL HEART RATES

detection. Blue and red SDs were imposed on top of the bar graphs of H-FHRs and D-FHRs, respectively.

Entropy bar plots in Fig. 5.10 (e) shows that H-FHRs are more irregular ($\sim 65\%$) compared to D-FHRs ($\sim 55\%$). This finding was in accordance with the fact that some diseases turn the FHR to be more regular than in its normal behavior.

Table 5.4 reports SDs of the five RQA issued for diagnosing fetal distress. All the recurrence quantifications were computed from the unembedded RP, embedded RP in $3-d$, embedded RP with specific settings, derivative-based RP and 2-time pattern RP. The five RQA were manipulated for both H-FHRs and D-FHRs, and the minimum recurrence parameter values were associated a blue color.

Table 5.4 – The Relative Errors (REs) of the five RQA issued for diagnosing fetal distress and computed from the unembedded RP, embedded RP, embedded RP with specific settings, derivative based RP and 2-time pattern RP.

SD(RQA) \ FHR	RP ₁		RP ₂		RP ₃		RP ₄		RP ₅	
	H-FHR	D-FHR								
SD(\overline{DET})	± 0.12	± 0	± 0.12	± 0	± 0.12	± 0	± 0.03	± 0	± 0.09	± 0.21
SD(\overline{CDET})	± 0.18	± 0.15	± 0.20	± 0.16	± 0.18	± 0.15	± 0.02	± 0.02	± 0.02	± 0.15
SD(\overline{PRSP})	± 17.93	± 14.69	± 19.60	± 15.65	± 17.94	± 14.69	± 2.37	± 1.59	± 14.45	± 14.62
SD(\overline{En})	± 10.48	± 12.08	± 10.65	± 11.99	± 10.47	± 12.08	± 16.28	± 15.26	± 0	± 12.03
SD(\overline{RR})	± 0.02	± 0.03	± 0	± 0.03	± 0.02	± 0.03	± 0.01	± 0.01	± 0	± 0.03

Table 5.5 reports the Effectiveness of Discrimination (ED) of FHRs by both the sensitivity and specificity measures in percentages of the Recurrence Quantification Analysis (RQA) in addition to the accuracy and precision. Table 5.5 reveals that the ultimate sensitivity and specificity belongs to CDET descriptor as it was proposed in our previous paper [Zaylaa et al., 2013]. Moreover, both the precision and accuracy were computed in this paper, and the best precision and accuracy corresponded to CDET descriptor

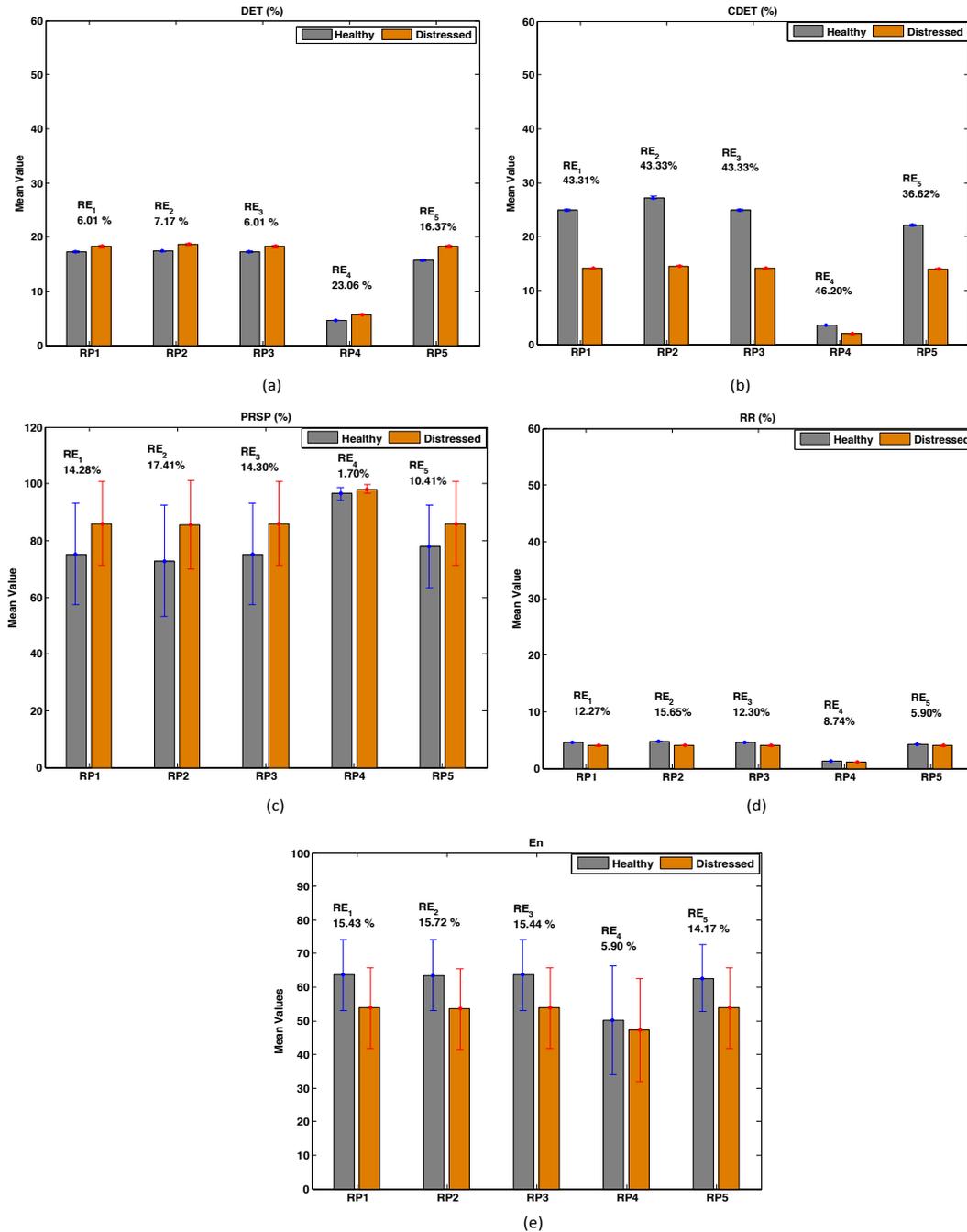


Figure 5.10 – The Recurrence Quantification Analysis (RQA) relative to the standard and unbiased Recurrence Plots of a healthy and distressed FHRs shown in gray and orange colors, respectively. (a) The mean DET(%) as a function of the unembedded RP, embedded RP in $3 - d$, embedded RP with specific settings, derivative-based RP and 2-time pattern RP, RP_1 , RP_2 , RP_3 , RP_4 and RP_5 , respectively. (b) The mean CDET(%) relative of the five RP techniques. (c) The average Percentage Reduced Sojourn Points (PRSP(%)) as a function of the five Recurrence Plots (RPs). (d) The average mean RR(%) as a function of the five RPs. (e) The mean En(%) as a function of the five RPs.

5.5. APPLICATION TO FETAL HEART RATES

Table 5.5 – The Effectiveness of Discrimination (ED) of Fetal Heart Rates (FHRs) by the sensitivity, specificity, precision and accuracy measures in percentages of the Recurrence Quantification Analysis (RQA).

RQA \ ED	Sensitivity	Specificity	Precision	Accuracy
DET	80%	20%	50%	50%
CDET*	80%	100%	100%	90%
PRSP	20%	80%	50%	50%
RR*	80%	100%	100%	90%
En*	80%	100%	100%	90%

followed by RR and En. The latter parameters are associated an asterisk in Table 5.5 to highlight their importance.

Recall that the previous statistical measures were computed after choosing empirically a threshold and imposing it on each bar-graph in Fig. 5.10. The criteria of classification is as follows: there is a total of 10 bars, (i) True Positives (TP) are healthy bar-peaks correctly detected, (ii) False positive (FP) are distressed bar-peaks incorrectly detected, (iii) False negative (FN) are healthy bar-peaks incorrectly undetected and (iv) True Negatives (TN) are distressed bar-peaks correctly undetected. The different statistical parameters used are defined as follows: $Sensitivity = \frac{TP}{TP+FN}$, $Specificity = \frac{TN}{TN+FP}$, $Precision = \frac{TP}{TP+FP}$ and $Accuracy = \frac{TP+TN}{TP+FN+FP+TN}$.

Of all the five RQA parameters, the sensitivity of healthy-distressed and distressed-healthy FHR detections is arranged as follows: $Sensitivity(CDET) = Sensitivity(DET) = Sensitivity(RR) = Sensitivity(En) > Sensitivity(PRSP)$. The specificity of discrimination is as follows: $Specificity(CDET) = Specificity(En) = Specificity(RR) > Specificity(PRSP) > Specificity(DET)$. REs(RQA) set out in Fig. 5.10 are arranged in the following order: $\overline{RE}(CDET) > \overline{RE}(En) > \overline{RE}(DET) > \overline{RE}(PRSP) > \overline{RE}(RR)$.

5.5.2 Discussion

The novel CDET descriptor that we have proposed through this dissertation has yielded the maximum RE between the two FHR classes, a minimum SD and possessed a high sensitivity and specificity values with 100% precision and 90% accuracy of diagnosis and discrimination. Although most of SD(RR) and SD(DET) parameters values were

comparable to those obtained by CDET, and both the sensitivity and the specificity of discrimination of RR were the same as that of CDET, the specificity of DET was 80% less than that of CDET.

RE values that both RR and DET have yielded were two to three times less than that obtained by CDET, still they served as moderate generic markers of fetal state transitions. Therefore, RR parameter was much more adapted for discrimination than DET but both were less effective than CDET. Moreover, $\overline{RE}(En)$ shows that En was much more adapted for discrimination than RR but still less efficient than CDET. SD(PRSP) was the highest and SD(DET) the lowest, consequently DET data are less dispersed. While comparing PRSP to DET, we noticed that the sensitivity was 60% higher in DET than in PRSP, however the specificity was 60% lower for PRSPs than for DET, while the accuracy and the precision were still 50%.

Of all RQA, CDET followed by En then RR were the most sensitive parameters to the discernment of Intrauterine Growth Restriction (IUGR). Moreover, of all the RQA used, CDET served as the ultimate generic marker for fetal state transition.

Furthermore, RR of FHRs was in general small (Fig. 5.10 (d)). Although RR values were in general small, a higher RR characterized H-FHRs compared to RR of D-FHRs. Moreover, H-FHRs were more irregular than D-FHRs (Fig. 5.10 (e)). The previous result still emphasizes on the fact that healthy biomedical signals such as H-FHRs are chaotic or pseudo-random in nature [Hurezeanu et al., 2013].

Discrimination of FHRs was not possible when the classical obstetrical parameters were used, like accelerations^b which were surprisingly more visible in case of distressed fetuses than healthy fetuses [Voicu and Girault, 2012a]. \overline{DET} remained greater for healthy fetuses compared to distressed fetuses. The previous behavior of H-FHRs was in accordance with the results reported in [Ferrario et al., 2006].

The derivative-based RP is the least RP contaminated with SPs, or the cleanest RP, followed by m-time pattern RP and RP with specific settings, then embedded RP and finally RP_1 (see Figs. 5.10 (b) and (c)). The previous result was in accordance with the alternative works that experienced the presence of sojourn points and the bias of RP_1 [Iwanski and Bradley, 1998, Zaylaa et al., 2013].

One of the limitations of RP_5 application to FHRs is the choice of $m = 2$ and the

^bFHR acceleration is an increase of the heart beats by at least 15 bmp during at least 15 seconds. This is the classical parameter with the smallest false positive rate in detecting hypoxia.

5.6. CONCLUSION

computation of RQA from the 2-time pattern recurrence plot. As it was inferred that CDET increased when d was increased, we believe that the performance of discrimination of RP_5 and its corresponding RQA's for ($m > 2$) could surpass the alternative unbiased RPs.

Moreover, by means of RP_4 it was possible to detect more dynamic transitions taking place for healthy fetal hearts compared to distressed fetal hearts. Although the difference was not huge, yet it still shows the advantage of the unbiased RP_4 not only based on the discrimination related to differences in the RQA parameters between H-FHRs and D-FHRs, but also based on the intrinsic behavior and dynamic transitions of each fetal heart group.

The intrinsic transition detection findings were in accordance with the physiological information, \overline{DET} remained greater for healthy fetuses compared to distressed fetuses. This is due to the complexity degradation of the heart rate in distressed fetuses compared to healthy fetuses. An overwhelming predictability of D-FHRs is foreseen when the complexity of the fetal heart is degraded. This complexity degradation can be due to a lower interaction between the brain and the heart of the fetus as a consequence of hypoxia, drop off in the concentration of oxygen in the fetal bloodstream.

The recurrence rate of FHRs was in general small and was in accordance with the findings of Hurezeanu et al. [Hurezeanu et al., 2013]. Although RR values were in general small, a higher RR descriptor characterized H-FHRs compared to RR of D-FHRs. Moreover, H-FHRs were more irregular than D-FHRs, this emphasizes the fact that healthy biomedical signals such as H-FHRs are more complex and random in nature.

One of the limitations of m-time pattern RP application to FHRs is the choice of $m = 2$ and the computation of RQA from the 2-time pattern recurrence plot. As our novel CDET marker increases when d increases, we believe that the performance of discrimination of m-time pattern RP and accordingly RQA for ($m > 2$) could surpass the alternative unbiased RPs.

5.6 Conclusion

By means of the unbiased derivative-based RP it is possible to detect more dynamic transitions such as healthy-distressed and distressed-healthy fetal transitions. Although the difference between URPs is not huge, it still shows the advantage of unbiased RP_4 not

only based on the discrimination offered by RQA between H-FHRs and D-FHRs, but also based on the intrinsic behavior and dynamic transitions of each fetal heart group. Both En and DET parameters served as moderate markers of fetal distress. While, the novel CDET parameter proposed in this dissertation, served as the ultimate generic marker for fetal heart rate distress.

The behaviors of similar fetal heart rates were demonstrated in this dissertation by the transitions they undergo, and could pave the way for the detection of certain diseases that encounter the fetus. This research should be extended to link certain diseases to the detected complexities and distress. The diagnosis of fetal distress by the state-of-the-art RPs was improved, which could open the door for advanced machines and advanced detection of distress to decide for immediate and preterm deliveries.

Part III

DISSERTATION SUMMARY AND PERSPECTIVES

Chapter 6

CONCLUSION AND PERSPECTIVES

6.1 Comparison of Complexity Analysis Techniques

Among the scale-based entropy parameters, there are entropy based on the probability principle and entropy based on permutation. The latter type of entropy is based on the metric distance and measures the likeliness of occurrence of patterns at different time. The former type of entropy measures the likeliness of similarly arranged patterns.

As the probability-based entropy parameters depended on m and r setting parameters, r was set at a value inferred empirically, and the pattern m was optimized by tracking the maximum entropy over a predefined range of m . Then the effect of increasing the order of entropy was studied by n -order entropy. The previous two entropy parameters have upgraded the discrimination of fetuses with IUGR from healthy fetuses.

The previous analysis was built on the probability of the correlation sum $C(r)$ which in turn depends on the distance I , and is computed directly from the time series.

The second type of analysis was a geometrical quantification based on computing the multi-fractal dimension and others. The computation of multi-fractal dimension depends on the correlation dimension D_2 (the slope of the correlation sum curve versus the tolerance r), the information dimension D_1 and the box counting dimension D_0 . Such that when $D_0 > D_1 > D_2$ is valid, the system is said to be multi-fractal, consequently the multi-fractal dimension is computed [Akay, 2000]. Our contribution to the previous method aided in the progression of FHR classification. This was justified by the amended relative error outcomes between healthy and distressed FHRs.

However, the third type of analysis realized by URPs provides a visual discrimination, and the quantification parameters are linked to the visual discrimination. It is not based on the probability of $C(r)$ but rather on statistics and the count of recurrences assigned a value of zero in the RP matrix. This analysis tool have provided overwhelming results of

detection of IUGR among the tested FHRs. This was reflected by the enhanced sensitivity, specificity, precision and accuracy outcomes. However, a pure alternative tool to quantify the complexity and discrimination could be entropy discussed in chapter 3, which is directly computed from the time series.

The main results of the application of the three different complexity analysis tools and the novel methods that we have proposed to both simulated and real data were summarized in Fig. 6.1.

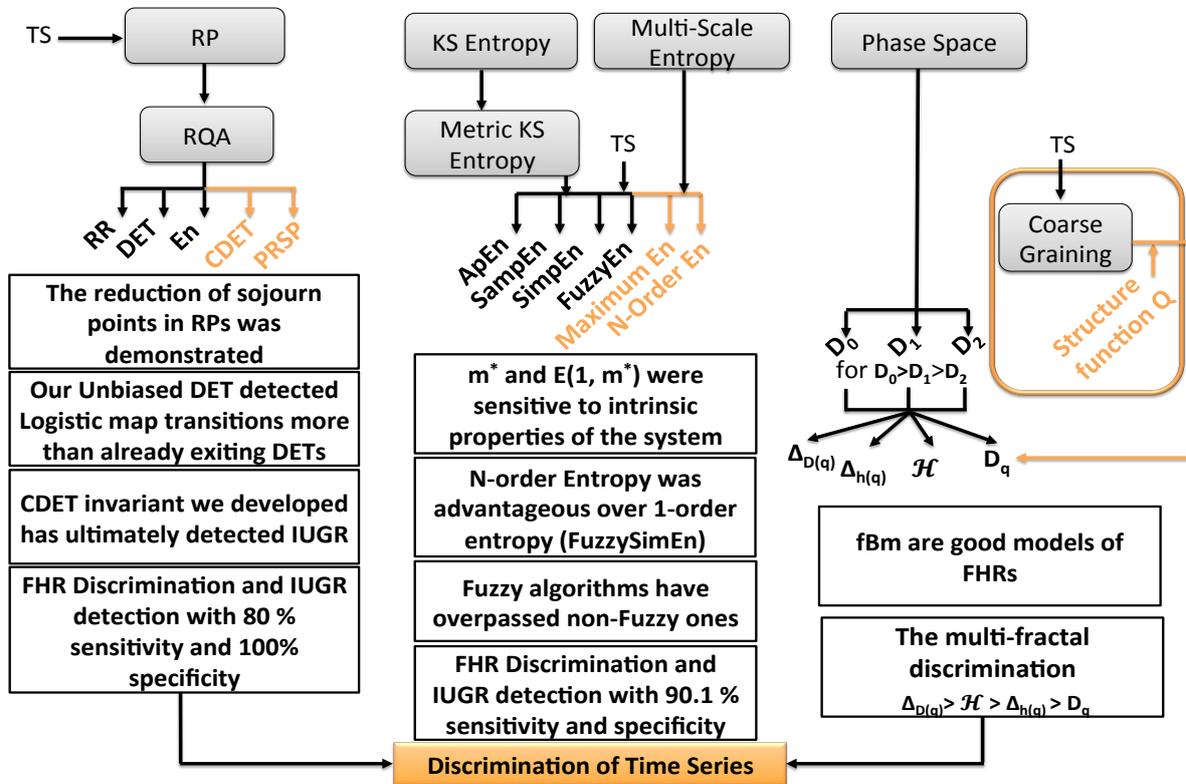


Figure 6.1 – Schematic illustration of the primary results and major findings provided by our dissertation.

The proposed processing tools of FHRs aided to detect severe IUGR which could be caused by but not limited to, respiratory, cardiovascular and digestive problems. These tools might help obstetricians opt to immediate delivery through a caesarean section. It might also provide the fundamental reason leading to fetal distress. Still, FHR processing and analysis are not trivial due to both the wide range of possible pathologies that could be encountered during the gestational period and the chaotic nature of the signals.

6.2 Conclusions

The three developed complexity analysis techniques resulting from the entropy quantification, multi-fractal analysis and recurrence plots and their quantification analysis were proven to be sensitive to the changes in the chaotic, random and/or biomedical time series compared to alternative analysis techniques.

We used several complexity descriptors for either detecting medical state transitions or discriminating between different fetuses and consequently predicting future states.

Geometric quantifications such as Hurst exponent, singularity and holder spectra were employed as fractal descriptors. They were obtained by contributing to multi-fractal analysis, i.e. coarse graining the time series and defining a specific structure function. These two steps amended the discrimination of fBm, as well as Fetal heart rates, accordingly the enhanced performance achieved by the mean of the singularity spectrum descriptor and the mean hurst exponent was reflected by the high relative errors between healthy and distressed cases as compared to alternative multi-fractal parameters.

Entropy descriptors derived from the information theory were involved as well. These included, but not limited to ApEn, SampEn, SimEn and FuzzyEn. Inspired from these latter complexity descriptors, alternative entropies descriptor were introduced in this dissertation. The first novel descriptor was the N-order En which is equivalent to N-order FuzzySimEn. It is based on permitting the selection of any pattern composed of 1-point, 2-points, or n-points for fuzzy similarity entropy computation. This new paradigm has an overwhelming potential and could be applied to other applications seeking the extraction of complexity invariant. The problem of setting entropy descriptors was solved by varying the pattern size leading an optimal pattern size that maximizes the n -order FuzzySimEn. FHR discrimination was improved by the new developed paradigm that encompasses the standard SimEn.

Another geometrical analysis tool was involved in the analysis of chaotic time series named recurrence plot. Three contributions to the already existing RPs has been done. Accordingly, three Unbiased RPs were developed to avoid the presence of sojourn points contaminating standard RPs. Such clean and URPs serve as geometrical analysis techniques from which significant recurrence quantification descriptors can be extracted. The determinism descriptor was involved due to its feasibility in detecting dynamic changes and state transitions. Two novel recurrence descriptors were extracted from RPs and introduced as a result of the development of clean RPs.

These descriptors were Cross-Determinism and the Percentage of Reduced Sojourn Points, they aided in the process of time series analysis based on URPs. The three recurrence descriptors were invariants for the fetal heart and were mainly used to detect healthy-distressed and distressed-healthy FHRs.

So far regarding Fetal Heart Rates, it was confirmed throughout this dissertation that the detection of Intrauterine Growth Restriction and indirectly the presence of hypoxia reduces the variability of Fetal Heart Rate time series and changes the values of the complexity parameters. Mainly, our novel CDET parameter is the strongest indicator of fetal distress and is powerful for discrimination purposes.

Furthermore, three multi-fractal descriptors (singularity spectrum, Hurst exponent then holder spectrum parameter) along with two entropy descriptors (n-order entropy and maximum entropy), and cross-determinism CDET parameter were mainly dedicated to discrimination purposes. They were employed to discern fetal IUGR due to hypoxia, improve the diagnosis and give hand in the therapeutic uptaking.

6.3 Limitations

- The N-order entropy paradigm was built by fixing the tolerance value and optimizing merely the number of pattern points.
- The performance of discrimination of Multi-fractal method extracted from the method proposed was compared with respect to the multi-fractal parameters extracted. However, the general technique was not compared to alternatives.
- Only a value of $m = 2$ for m-time Unbiased Recurrence Plots was simulated.
- The three complexity analysis approaches were paramountly applied to Fetal heart rate solely. Moreover, Intrauterine Growth Restriction caused by hypoxia was considered, while other physiological causes could be studied.

6.4 Outlook

The new ideas involved in the three major nonlinear analysis techniques provided an advanced knowledge of the underlying physical properties of the biomedical system under

6.4. OUTLOOK

study. It paved the way for and an early detection of distress, as well as, it gives information whether the system is predictable or not.

In regards to the recurrence analysis, as the recurrence matrix is a special case of isometric transformations, other types of transformations could be involved in the future as statistical signal processing tools to detect specific diseases reflecting certain features. Moreover, all the developed algorithms and post-processing codes will build programmable portable medical devices providing enhanced discrimination of medical states after detecting the signals, particularly FHRs and will pave the way for a degraded risk of fetal death.

Furthermore, the following points should be considered:

- Optimization of the tolerance value in the N-order entropy paradigm and apply it to biomedical signals;
- The search for the optimal set of parameters (r^*, m^*) to optimize the maximum entropy;
- Investigation and comparison the performance of discrimination of Multi-fractal method proposed were compared to alternatives;
- The evaluation of multi-fractal descriptors online in the near future, as our proposed methodology in multi-fractal analysis was based on offline investigation;
- Test the performance of m-time Recurrence Plot for m greater than 2;
- Other biomedical signals and diagnostic purposes will be considered.

REFERENCES

Bibliography

- [Ahlstrom et al., 2006] Ahlstrom, C., Hult, P., and Ask, P. (2006). Thresholding distance plots using true recurrence points. In *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006*, pages 688–691. IEEE.
- [Akay, 2000] Akay, M. (2000). *Nonlinear Biomedical Signal Processing Vol. II: Dynamic Analysis and Modeling*. Wiley-IEEE Press.
- [Alamedine et al., 2013] Alamedine, D., Khalil, M., and Marque, C. (2013). Comparison of different ehg feature selection methods for the detection of preterm labor. *Computational and Mathematical Methods in Medicine*, 2013(9).
- [Amaral et al., 2001] Amaral, L. A. N., Ivanov, P. C., Aoyagi, N., Hidaka, I., Tomono, S., Goldberger, A. L., Stanley, H. E., and Yamamoto, Y. (2001). Behavioral-independent features of complex heartbeat dynamics. *Physical Review Letters*, 86(26):6026.
- [Arneodo et al., 2003] Arneodo, A., Decoster, N., Kestener, P., and Roux, S. (2003). A wavelet-based method for multifractal image analysis: from theoretical concepts to experimental applications. *Advances in Imaging and Electron Physics*, 126:1–92.
- [Ash, 1990] Ash, R. (1990). Information theory. corrected reprint of the 1965 original.
- [Balibrea et al., 2008] Balibrea, F., Caballero, M., and Molera, L. (2008). Recurrence quantification analysis in lius attractor. *Chaos, Solitons & Fractals*, 36(3):664–670.
- [Bandt and Pompe, 2002] Bandt, C. and Pompe, B. (2002). Permutation entropy: a natural complexity measure for time series. *Physical Review Letters*, 88(17):174102.
- [Barabasi and Vicsek, 1991] Barabasi, A.-L. and Vicsek, T. (1991). Multifractality of self-affine fractals. *Physical Review A*, 44(4):2730.

- [Baumert et al., 2012] Baumert, M., Javorcka, M., Seeck, A., Faber, R., Sanders, P., and Voss, A. (2012). Multiscale entropy and detrended fluctuation analysis of qt interval and heart rate variability during normal pregnancy. *Computers in biology and medicine*, 42(3):347–352.
- [Behnia et al., 2008] Behnia, S., Akhshani, A., Mahmodi, H., and H., H. (2008). Nonlinear measure of ecg time series: Detection of cardiac diseases. *Iranian Physical Journal*, 2-1:53–62.
- [Boskovic et al., 2012] Boskovic, A., Loncar-Turukalo, T., Sarenac, O., Japundzic-Zigon, N., and Bajic, D. (2012). Unbiased entropy estimates in stress: A parameter study. *Computers in biology and medicine*, 42(6):667–679.
- [Cao, 1997] Cao, L. (1997). Practical method for determining the minimum embedding dimension of a scalar time series. *Physica D: Nonlinear Phenomena*, 110(1):43–50.
- [Chaudhuri and Sarkar, 1995] Chaudhuri, B. and Sarkar, N. (1995). Texture segmentation using fractal dimension. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(1):72–77.
- [Chen et al., 2009] Chen, W., Zhuang, J., Yu, W., and Wang, Z. (2009). Measuring complexity using fuzzyen, apen, and sampen. *Medical Engineering & Physics*, 31(1):61–68.
- [Ching and Tsang, 2007] Ching, E. S. and Tsang, Y.-K. (2007). Multifractality and scale invariance in human heartbeat dynamics. *Physical Review E*, 76(4):041910.
- [Chon et al., 2009] Chon, K. H., Scully, C., and Lu, S. (2009). Approximate entropy for all signals. *Engineering in Medicine and Biology Magazine, IEEE*, 28(6):18–23.
- [Costa et al., 2005] Costa, M., Goldberger, A. L., and Peng, C.-K. (2005). Multiscale entropy analysis of biological signals. *Physical Review E*, 71(2):021906.
- [Costa et al., 2014] Costa, M. D., Schnettler, W. T., Amorim-Costa, C., Bernardes, J., Costa, A., Goldberger, A. L., and Ayres-de Campos, D. (2014). Complexity-loss in fetal heart rate dynamics during labor as a potential biomarker of acidemia. *Early human development*, 90(1):67–71.
- [Cross, 2000] Cross, M. (2000). Physics 161: Introduction to chaos, chapter 8: Information and entropy. Online. in pdf.

REFERENCES

- [Eckmann et al., 1987] Eckmann, J.-P., Kamphorst, S. O., and Ruelle, D. (1987). Recurrence plots of dynamical systems. *Europhys. Lett*, 4(9):973–977.
- [Ferrario et al., 2009] Ferrario, M., Signorini, M. G., and Magenes, G. (2009). Complexity analysis of the fetal heart rate variability: early identification of severe intrauterine growth-restricted fetuses. *Medical & biological engineering & computing*, 47(9):911–919.
- [Ferrario et al., 2006] Ferrario, M., Signorini, M. G., Magenes, G., and Cerutti, S. (2006). Comparison of entropy-based regularity estimators: application to the fetal heart rate signal for the identification of fetal distress. *IEEE Transactions on Biomedical Engineering*, 53(1):119–125.
- [Fraser and Swinney, 1986] Fraser, A. M. and Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information. *Physical review A*, 33(2):1134.
- [Frisch and Parisi, 1985] Frisch, U. and Parisi, G. (1985). On the singularity structure of fully developed turbulence.
- [Gao and Cai, 2000] Gao, J. and Cai, H. (2000). On the structures and quantification of recurrence plots. *Physics Letters A*, 270(1):75–87.
- [Girault et al., 2010] Girault, J.-M., Kouame, D., and Ouahabi, A. (2010). Analytical formulation of the fractal dimension of filtered stochastic signals. *Signal Processing*, 90(9):2690–2697.
- [Girault et al., 2013] Girault, J.-M., Oudjemia, S., and Voicu, I. (2013). Multi-scale similarity entropy as a complexity descriptor to discriminate healthy to distress foetus. *International Journal of Systems, Control and Communications*, 5(3):276–284.
- [Goldenberg et al., 2008] Goldenberg, R. L., Culhane, J. F., Iams, J. D., and Romero, R. (2008). Epidemiology and causes of preterm birth. *The Lancet*, 371(9606):75–84.
- [Gough, 1993] Gough, N. A. J. (1993). Fractal analysis of foetal heart rate variability. *Physiological measurement*, 14(3):309.
- [Grassberger and Procaccia, 1983] Grassberger, P. and Procaccia, I. (1983). Characterization of strange attractors. *Physical review letters*, 50(5):346.

- [Grassberger and Procaccia, 2004] Grassberger, P. and Procaccia, I. (2004). Measuring the strangeness of strange attractors. In *The Theory of Chaotic Attractors*, pages 170–189. Springer.
- [Habek et al., 2001] Habek, D., Habek, J., Jugović, D., and Salihagić, A. (2001). Intrauterine hypoxia and sudden infant death syndrome. *Acta medica Croatica: casopis Hrvatske akademije medicinskih znanosti*, 56(3):109–118.
- [Hazewinkel, 2001] Hazewinkel, M. (2001). *Encyclopaedia of mathematics supplement III*, volume 13. Springer.
- [Hsu et al., 2007] Hsu, W.-Y., Lin, C.-C., Ju, M.-S., and Sun, Y.-N. (2007). Wavelet-based fractal features with active segment selection: Application to single-trial eeg data. *Journal of neuroscience methods*, 163(1):145–160.
- [Hu et al., 2006] Hu, J., Gao, J., and Príncipe, J. C. (2006). Analysis of biomedical signals by the lempel-ziv complexity: the effect of finite data size. *Biomedical Engineering, IEEE Transactions on*, 53(12):2606–2609.
- [Humeau et al., 2010] Humeau, A., Buard, B., Mahe, G., Chapeau-Blondeau, F., Rousseau, D., and Abraham, P. (2010). Multifractal analysis of heart rate variability and laser doppler flowmetry fluctuations: comparison of results from different numerical methods. *Physics in medicine and biology*, 55(20):6279.
- [Hurezeanu et al., 2013] Hurezeanu, B., Ungureanu, G. M., Digulescu, A., Serbanescu, A., Costin, H., and Strungaru, R. (2013). Fetal heart rate variability study with recurrence plot analysis. In *E-Health and Bioengineering Conference (EHB), 2013*, pages 1–4. IEEE.
- [Hurst, 1951] Hurst, H. (1951). Long-term storage capacity of reservoirs. *Trans. Am. Soc. Civil Eng.*, 116:7701–7799.
- [Hurst et al., 1965] Hurst, H. E., Black, R. P., and Simaika, Y. (1965). *Long-term storage: an experimental study*. Constable.
- [Ivanov et al., 1996] Ivanov, P., Rosenblum, M., and Peng, C.-K. (1996). Scaling behavior of heartbeat intervals obtained by wavelet-based time series analysis. *Nature*, 383(6598):323–327.

REFERENCES

- [Ivanov et al., 1999] Ivanov, P. C., Amaral, L. A. N., Goldberger, A. L., Havlin, S., Rosenblum, M. G., Struzik, Z. R., and Stanley, H. E. (1999). Multifractality in human heartbeat dynamics. *Nature*, 399(6735):461–465.
- [Ivanov et al., 2001] Ivanov, P. C., Nunes Amaral, L. A., Goldberger, A. L., Havlin, S., Rosenblum, M. G., Stanley, H. E., and Struzik, Z. R. (2001). From 1/f noise to multifractal cascades in heartbeat dynamics. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 11(3):641–652.
- [Iwanski and Bradley, 1998] Iwanski, S. J. and Bradley, E. (1998). Recurrence plots of experimental data: To embed or not to embed? *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 8(4):861.
- [Javorka et al., 2008] Javorka, M., Trunkvalterova, Z., Tonhajzerova, I., Javorkova, J., Javorka, K., and Baumert, M. (2008). Short-term heart rate complexity is reduced in patients with type 1 diabetes mellitus. *Clinical Neurophysiology*, 119(5):1071–1081.
- [Jensen, 1996] Jensen, J. A. (1996). *Estimation of Blood Velocities Using Ultrasound*. Cambridge University Press.
- [Kantz and Schreiber, 2004] Kantz, H. and Schreiber, T. (2004). *Nonlinear time series analysis*, volume 7. Cambridge university press.
- [Keller et al., 1989] Keller, J. M., Chen, S., and Crownover, R. M. (1989). Texture description and segmentation through fractal geometry. *Computer Vision, Graphics, and Image Processing*, 45(2):150–166.
- [Kennel et al., 1992] Kennel, M. B., Brown, R., and Abarbanel, H. D. I. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical review A*, 45(6):3403.
- [Kobayashi and Musha, 1982] Kobayashi, M. and Musha, T. (1982). 1/f fluctuation of heartbeat period. *IEEE Transactions on Biomedical Engineering*, (6):456–457.
- [Kolmogorov, 1959] Kolmogorov, A. N. (1959). Entropy per unit time as a metric invariant of automorphisms. In *Dokl. Akad. Nauk SSSR*, volume 124, pages 754–755.
- [Lake et al., 2002] Lake, D. E., Richman, J. S., Griffin, M. P., and Moorman, J. R. (2002). Sample entropy analysis of neonatal heart rate variability. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 283(3):R789–R797.

-
- [Levy-Vehel, 1995] Levy-Vehel, J. (1995). Fractal approaches in signal processing. *Fractals*, 3(04):755–775.
- [Lin and Hughson, 2001] Lin, D. and Hughson, R. (2001). Modeling heart rate variability in healthy humans: a turbulence analogy. *Physical review letters*, 86(8):1650.
- [Liu et al., 2013] Liu, C., Li, K., Zhao, L., Liu, F., Zheng, D., Liu, C., and Liu, S. (2013). Analysis of heart rate variability using fuzzy measure entropy. *Computers in biology and Medicine*, 43(2):100–108.
- [Liu et al., 2011] Liu, C., Liu, C., Shao, P., Li, L., Sun, X., Wang, X., and Liu, F. (2011). Comparison of different threshold values r for approximate entropy: application to investigate the heart rate variability between heart failure and healthy control groups. *Physiological Measurement*, 32(2):167.
- [Lopes and Betrouni, 2009] Lopes, R. and Betrouni, N. (2009). Fractal and multifractal analysis: a review. *Medical image analysis*, 13(4):634–649.
- [Lu et al., 2008] Lu, S., Chen, X., Kanters, J. K., Solomon, I. C., and Chon, K. H. (2008). Automatic selection of the threshold value for approximate entropy. *Biomedical Engineering, IEEE Transactions on*, 55(8):1966–1972.
- [Mandelbrot and Van Ness, 1968] Mandelbrot, B. B. and Van Ness, J. W. (1968). Fractional brownian motions, fractional noises and applications. *SIAM review*, 10(4):422–437.
- [Mandelbrot and Wallis, 1969] Mandelbrot, B. B. and Wallis, J. R. (1969). Some long-run properties of geophysical records. *Water resources research*, 5(2):321–340.
- [Manetti et al., 1999] Manetti, C., Ceruso, M.-A., Giuliani, A., Webber Jr, C. L., and Zbilut, J. P. (1999). Recurrence quantification analysis as a tool for characterization of molecular dynamics simulations. *Physical Review E*, 59(1):992.
- [Marek, 1995] Marek, M. (1995). *Chaotic behaviour of deterministic dissipative systems*, volume 1. Cambridge University Press.
- [Marwan, 2003] Marwan, N. (2003). *Encounters with neighbours: current developments of concepts based on recurrence plots and their applications*. Norbert Marwan.
- [Marwan, 2009] Marwan, N. (2009). Commandline recurrence plots.

REFERENCES

- [Marwan et al., 2007] Marwan, N., Carmen Romano, M., Thiel, M., and Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports*, 438(5):237–329.
- [Marwan and Kurths, 2002] Marwan, N. and Kurths, J. (2002). Nonlinear analysis of bivariate data with cross recurrence plots. *Physics Letters A*, 302(5):299–307.
- [Marwan et al., 2002] Marwan, N., Wessel, N., Meyerfeldt, U., Schirdewan, A., and Kurths, J. (2002). Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. *Physical Review E*, 66(2):026702.
- [Maslova et al., 2003] Maslova, M., Maklakova, A., Sokolova, N., Ashmarin, I., Goncharenko, E., and Krushinskaya, Y. V. (2003). The effects of ante-and postnatal hypoxia on the central nervous system and their correction with peptide hormones. *Neuroscience and behavioral physiology*, 33(6):607–611.
- [Mesa and Poveda, 1993] Mesa, O. J. and Poveda, G. (1993). The hurst effect: The scale of fluctuation approach. *Water Resources Research*, 29(12):3995–4002.
- [Muzy et al., 1993] Muzy, J.-F., Bacry, E., and Arneodo, A. (1993). Multifractal formalism for fractal signals: The structure-function approach versus the wavelet-transform modulus-maxima method. *Physical review E*, 47(2):875.
- [Nguyen et al., 2012] Nguyen, C. D., Wilson, S. J., and Crozier, S. (2012). Automated quantification of the synchrogram by recurrence plot analysis. *IEEE Trans. Biomedical Engineering*, 59(4):946–955.
- [Oudjemia et al., 2013] Oudjemia, S., Zaylaa, A., Haddab, S., and Girault, J.-M. (2013). Coarse-grained multifractality analysis based on structure function measurements to discriminate healthy from distressed fetuses. *Computational and Mathematical Methods in Medicine, USA*, 2013.
- [Packard et al., 1980] Packard, N. H., Crutchfield, J. P., Farmer, J. D., and Shaw, R. S. (1980). Geometry from a time series.
- [Perron, 1930] Perron, O. (1930). Die stabilitätsfrage bei differentialgleichungen. *Mathematische Zeitschrift*, 32(1):703–728.
- [Peters, 1996] Peters, E. E. (1996). *Chaos and order in the capital markets: a new view of cycles, prices, and market volatility*, volume 1. John Wiley & Sons.

-
- [Pham, 2010] Pham, T. D. (2010). Geentropy: A measure of complexity and similarity. *Pattern Recognition*, 43(3):887–896.
- [Pincus, 1991] Pincus, S. M. (1991). Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences*, 88(6):2297–2301.
- [Pincus and Goldberger, 1994] Pincus, S. M. and Goldberger, A. L. (1994). Physiological time-series analysis: what does regularity quantify? *American Journal of Physiology*, 266:H1643–H1643.
- [Pincus and Keefe, 1992] Pincus, S. M. and Keefe, D. L. (1992). Quantification of hormone pulsatility via an approximate entropy algorithm. *Am J Physiol*, 262(5 Pt 1):E741–E754.
- [Popivanov et al., 2005] Popivanov, D., Jivkova, S., Stomonyakov, V., and Nicolova, G. (2005). Effect of independent component analysis on multifractality of eeg during visual-motor task. *Signal processing*, 85(11):2112–2123.
- [Rényi, 1961] Rényi, A. (1961). On measures of entropy and information. In *Fourth Berkeley Symposium on Mathematical Statistics and Probability*, pages 547–561.
- [Restrepo et al., 2014] Restrepo, J. F., Schlotthauer, G., and Torres, M. E. (2014). Maximum approximate entropy and r threshold: A new approach for regularity changes detection. *Physica A: Statistical Mechanics and its Applications*, 409:97–109.
- [Richman and Moorman, 2000] Richman, J. S. and Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology*, 278(6):H2039–H2049.
- [Riedl et al., 2013] Riedl, M., Müller, A., and Wessel, N. (2013). Practical considerations of permutation entropy. *The European Physical Journal Special Topics*, 222(2):249–262.
- [Rouvre, 2006] Rouvre, D. (2006). *Caractérisation de l’activité foetale: mise en oeuvre d’un dispositif d’enregistrement et analyse des signaux doppler multidimensionnels*. PhD thesis, Université François Rabelais de Tours, Tours, France.
- [Sassi et al., 2009] Sassi, R., Signorini, M., and Cerutti, S. (2009). Multifractality and heart rate variability. *Chaos*.
- [Shannon, 1948] Shannon, C. (1948). A mathematical theory of communication.

REFERENCES

- [Sinai, 1959] Sinai, Y. G. (1959). On the concept of entropy of a dynamical system. In *Dokl. Akad. Nauk. SSSR*, volume 124, pages 768–771.
- [Spilka et al., 2012] Spilka, J., Chudáček, V., Koucký, M., Lhotská, L., Huptych, M., Jank, P., Georgoulas, G., and Stylios, C. (2012). Using nonlinear features for fetal heart rate classification. *Biomedical Signal Processing and Control*, 7(4):350–357.
- [Stanley, 1996] Stanley, H. E. (1996). Scaling behaviour of heartbeat intervals obtained by wavelet-based time-series analysis. *Nature*, 383:26.
- [Takens et al., 1981] Takens, F. et al. (1981). Dynamical systems and turbulence. *Lecture notes in mathematics*, 898(9):366.
- [Thiel et al., 2004] Thiel, M., Romano, M., Read, P., and Kurths, J. (2004). Estimation of dynamical invariants without embedding by recurrence plots. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 14(2):234–243.
- [Trauth et al., 2010] Trauth, M. H., Gebbers, R., Marwan, N., and Sillmann, E. (2010). *MATLAB Recipes for Earth Sciences*. Springer.
- [Trulla et al., 1996] Trulla, L., Giuliani, A., Zbilut, J., and Webber, C. (1996). Recurrence quantification analysis of the logistic equation with transients. *Physics Letters A*, 223(4):255–260.
- [Voicu and Girault, 2012a] Voicu, I. and Girault, J.-M. (2012a). Multi-scale sample entropy and recurrence plots distinguish healthy from suffering foetus. *Acoustics 2012 Nantes*.
- [Voicu and Girault, 2012b] Voicu, I. and Girault, J.-M. (2012b). Multi-scale similarity entropy as a new descriptor to differentiate healthy to suffering foetus. In *International Conference on Complex Systems, Agadir-Morocco*.
- [Voicu et al., 2010] Voicu, I., Girault, J. M., Roussel, C., Decock, A., and Kouame, D. (2010). Robust estimation of fetal heart rate from us doppler signals. *Physics Procedia*, 3(1):691–699. International Congress on Ultrasonics, Santiago de Chile, January 2009.
- [Voicu et al., 2009] Voicu, I., Kouame, D., Fournier-Massignan, M., and Girault, J.-M. (2009). Estimating fetal heart rate from multiple ultrasound signals. In *International Conference on Advancements of Medicine and Health Care through Technology*, volume 26, pages 185–190. Springer Berlin Heidelberg.

- [Voicu et al., 2014] Voicu, I., Ménigot, S., Kouamé, D., and Girault, J.-M. (2014). New estimators and guidelines for better use of fetal heart rate estimators with doppler ultrasound devices. *Computational and mathematical methods in medicine*, 2014.
- [Wang et al., 2007] Wang, G., Huang, H., Xie, H., Wang, Z., and Hu, X. (2007). Multifractal analysis of ventricular fibrillation and ventricular tachycardia. *Medical engineering & physics*, 29(3):375–379.
- [Wang et al., 2003] Wang, J., Ning, X., and Chen, Y. (2003). Modulation of heart disease information to the 12-lead ecg multifractal distribution. *Physica A: Statistical Mechanics and its Applications*, 325(3):485–491.
- [Wang et al., 2005] Wang, J., Ning, X., Ma, Q., Bian, C., Xu, Y., and Chen, Y. (2005). Multiscale multifractality analysis of a 12-lead electrocardiogram. *Physical Review E*, 71(6):062902.
- [Webber and Zbilut, 1994] Webber, C. and Zbilut, J. (1994). Dynamical assessment of physiological systems and states using recurrence plot strategies. *Journal of Applied Physiology*, 76(2):965–973.
- [Webber Jr and Zbilut, 2005] Webber Jr, C. L. and Zbilut, J. P. (2005). Recurrence quantification analysis of nonlinear dynamical systems. *Tutorials in contemporary nonlinear methods for the behavioral sciences*, pages 26–94.
- [Xia et al., 2006] Xia, Y., Feng, D. D., and Zhao, R. (2006). Morphology-based multifractal estimation for texture segmentation. *IEEE Transactions on Image Processing*, 15(3):614–623.
- [Xie et al., 2008] Xie, H., He, W., and Liu, H. (2008). Measuring time series regularity using nonlinear similarity-based sample entropy. *Physics Letters A*, 372:7140–7146.
- [Xu et al., 2011] Xu, Y., Ma, Q., Schmitt, D. T., Bernaola-Galvan, P., and Ivanov, P. C. (2011). Effects of coarse-graining on the scaling behavior of long-range correlated and anti-correlated signals. *Physica A: Statistical Mechanics and its Applications*, 390(23):4057–4072.
- [Yan et al., 2010] Yan, J., Zhou, C., Xia, C., Wang, Y., Li, F., Guo, R., and Yan, H. (2010). Recurrence quantification analysis base on wavelet packets for wrist pulse. In *3rd*

REFERENCES

- International Conference on Biomedical Engineering and Informatics, 2010*, volume 3, pages 1011–1015. IEEE.
- [Yang et al., 2011] Yang, H., Bukkapatnam, S., and Barajas, L. G. (2011). Local recurrence based performance prediction and prognostics in the nonlinear and nonstationary systems. *Pattern Recognition*, 44(8):1834–1840.
- [YUM and Kim, 2002] YUM, M.-K. and Kim, J. S. (2002). Increased intermittency and decreased nonstationarity of heart rates during the daytime in patients with neurocardiogenic syncope. *Journal of cardiovascular electrophysiology*, 13(8):788–793.
- [Zaylaa et al., 2014a] Zaylaa, A., Charara, J., and Girault, J. M. (2014a). Reducing sojourn points from recurrence plots to improve transition detection: Application to fetal heart rate transitions. *Computers In Biology and Medicine, Elsevier*.
- [Zaylaa et al., 2013] Zaylaa, A., Girault, J.-M., and Charara, J. (2013). Unbiased recurrence plot quantification of chaotic dynamic systems by eliminating sojourn points. In *2nd International Conference on Advances in Biomedical Engineering*, pages 187–190. IEEE.
- [Zaylaa et al., 2014b] Zaylaa, A., Oudjemia, S., Charara, J., and Girault, J. M. (2014b). N-order and maximum fuzzy similarity entropy for discrimination of signals of different complexity: application to fetal heart rate signals. *Computers In Biology and Medicine, Elsevier, In Revision*.
- [Zbilut et al., 1998] Zbilut, J. P., Giuliani, A., and Webber, C. L. (1998). Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification. *Physics Letters A*, 246(1):122–128.
- [Zbilut et al., 2002] Zbilut, J. P., Thomasson, N., and Webber, C. L. (2002). Recurrence quantification analysis as a tool for nonlinear exploration of nonstationary cardiac signals. *Medical engineering & physics*, 24(1):53–60.
- [Zbilut and Webber, 1992] Zbilut, J. P. and Webber, C. L. (1992). Embeddings and delays as derived from quantification of recurrence plots. *Physics letters A*, 171(3):199–203.
- [Zhang, 1991] Zhang, Y.-C. (1991). Complexity and 1/f noise. a phase space approach. *Journal de Physique I*, 1(7):971–977.

- [Ziv and Lempel, 1978] Ziv, J. and Lempel, A. (1978). Compression of individual sequences via variable-rate coding. *Information Theory, IEEE Transactions on*, 24(5):530–536.

Annexes

Appendix A

Pseudocodes of Existing Entropy Parameters

ApEn:

Initialize The System's Output $x(n)$
Divide $u_m(i) \leftarrow x(n)$
Fix $r \leftarrow$ tolerance $(0.15, 0.2) \times \text{S.D.}(x(n))$ according to the theoretical inference
Set $m \leftarrow$ length of the pattern, default is 2
Compute $P_i^m(r)$ (the number of vectors) $\leftarrow \Gamma[u_m(i), u_m(j)] \leq r$
Generate $C_i^m(r)$ (the probability of vectors) $\leftarrow P_i^m / (M - m + 1)$
Evaluate $\phi^{(m)}(r) \leftarrow -\frac{1}{M - (m - 1)} \sum_{i=1}^{M-(m-1)} \log C_i^m(r)$
Increase the pattern size m by 1
Calculate $\phi^{(m+1)}(r) \leftarrow -\frac{1}{M - m} \sum_{i=1}^{M-m} \log C_i^{m+1}(r)$
Evaluate $\text{ApEn}(m,r) \leftarrow [\phi^{(m+1)}(r) - \phi^{(m)}(r)]$
Output The Approximate Entropy.

SampEn:

Initialize The Biomedical System's Output $x(n)$
Divide $u_m(i) \leftarrow x(n)$
Fix $r \leftarrow$ tolerance $(0.15, 0.2) \times \text{S.D.}(x(n))$ according to the theoretical inference
Set $m \leftarrow$ length of the pattern, default is 2
Compute $P_i^m(r)$ (the number of vectors) $\leftarrow \Gamma[u_m(i), u_m(j)] \leq r$
Generate $C_i^m(r)$ (the probability of vectors) $\leftarrow P_i^m / (M - m + 1)$
Calculate $\phi^{(m)}(r) \leftarrow -\log \left(\frac{1}{M - m} \sum_{i=1}^{M-m} C_r^m(i) \right)$
 $\phi^{(m+1)}(r) \leftarrow -\log \left(\frac{1}{M - m} \sum_{i=1}^{M-m} C_r^{m+1}(i) \right)$
Evaluate $\text{SampEn}(m,r) \leftarrow [\phi^{(m+1)}(r) - \phi^{(m)}(r)]$
 It is equivalent to counting the number of vectors of length (m) and $(m+1)$ that exist within the tolerance r excluding self-matches.
Output The Sample Entropy.

SimEn: SimEn pseudocode is given as follows:

Initialize The Biomedical System's Output $x(n)$
Declare $\mathbf{X}_i \leftarrow [x(i), x(i+1), \dots, x(i+m-1)]$
Compute mean \bar{X}_i of the vector \mathbf{X}_i
Evaluate $\mathbf{X}_c \leftarrow [x_j(i) - \bar{X}_i, x_j(i+1) - \bar{X}_i, \dots, x_j(i+m-1) - \bar{X}_i]$
 $\mathbf{X}_c \leftarrow [x(i), x(i+1), \dots, x(i+m-1)]$
Find For each $\mathbf{X}_c(i)$ the vectors $\mathbf{X}_c(k)$ which have the same derivative sign as that of $\mathbf{X}_c(i)$
Compute the number P_i of similar vectors $(I(\mathbf{X}_c(i), \mathbf{X}_k(i)) \leq r)$
Exclude the self-similarity $(i \neq j)$ in the calculation of $P(r, \Gamma_{i,j}^m)$
For $i \leftarrow 1, \dots, M - m + 1$
Compute the probability vector $\mathbf{C}_r^m(\mathbf{i}) \leftarrow \frac{1}{M - m + 1} P_i^m$
Evaluate $\phi^{(m)}(r) \leftarrow -\log \left(\frac{1}{M - m} \sum_{i=1}^{M-m} C_r^m(i) \right)$
 $\phi^{(m+1)}(r) \leftarrow -\log \left(\frac{1}{M - m} \sum_{i=1}^{M-m} C_r^{m+1}(i) \right)$
Evaluate $\text{SimEn}(m,r,M) \leftarrow \phi_{m+1}(r) - \phi_m(r)$
Output The Similarity Entropy.

FuzzyEn:

Initialize The Biomedical System's Output $x(n)$

Declare $\mathbf{X}_i^m \leftarrow [x(i), x(i+1), \dots, x(i+m-1)] - v\bar{\mathbf{X}}_i^{(m)}$
 $\bar{\mathbf{X}}_i^{(m)} = \frac{1}{m} \sum_{i=0}^{m-1} x(i+1)$ and

$$v = \begin{cases} 0 & \text{if the baseline exists} \\ 1 & \text{if the baseline is eliminated} \end{cases}$$

Declare these m -patterns $\mathbf{X}_i^{(m)} \leftarrow$ directly formed from the original m -consecutive values extracted from the time series. This step is dedicated to remove the base line of each m -pattern in the time series

Test whether a vector $X_i^{(m)}$ is similar to $X_j^{(m)}$ within a tolerance r

Evaluate the correlation summation $C_i^{(m)}(r)$ by:

$$C_r^m(i) \leftarrow P(r, \Gamma_{i,j}^m) / (N - m + 1)$$

Compute $\Gamma_{i,j}^m \leftarrow$ the distance between the two m -patterns $X_i^{(m)}$ and $X_j^{(m)}$ defined as follows:

$$\Gamma_{i,j}^m \leftarrow \Gamma(X_i^{(m)}, X_j^{(m)}) = \max_{k \in (0, m-1)} |x(i+k) - x(j+k)|$$

Declare P the number of similar patterns:

$$P(r, \Gamma_{i,j}^m) \leftarrow e^{-\left(\frac{\Gamma_{i,j}^m}{r}\right)^p}$$

Evaluate $\phi_m(r) \leftarrow -\log\left(\frac{1}{M-m} \sum_{i=1}^{M-m} C_r^m(i)\right)$

Increase the pattern size m by 1

Calculate $\phi_{m+1}(r) \leftarrow -\log\left(\frac{1}{M-m} \sum_{i=1}^{M-m} \ln C_r^{m+1}(i)\right)$

Evaluate FuzzyEn(m, r, M) $\leftarrow \phi_{m+1}(r) - \phi_m(r)$

Output The Fuzzy Entropy.

APPENDIX A. PSEUDOCODES OF EXISTING ENTROPY PARAMETERS

Appendix B

Pseudocodes of Developed Unbiased RPs

Pseudocode of embedding recurrence plot with specific settings:

```
Initialize The Biomedical System's Output,  $x(t)$  and sample it
Fix  $\tau \leftarrow$  number inspired from a theoretical inference
Set  $d \leftarrow$  A range for the embedding dimension
Repeat
  For each value of  $d$ 
    Generate  $RP_3 \leftarrow$  Eq. 5.1, for  $\mathbf{X}_{i+\tau} \leftarrow \mathbf{X}_j$ 
    Evaluate  $J$  as in Eq. 5.7
    While ( $J \neq 0$ )
      do Optimize the cost function  $J$ 
    Select  $J_{op} \leftarrow J_{min}$ 
  Output The optimal Embedding dimension ( $d \leftarrow d_{op}$ )
Fix  $d \leftarrow d_{op}$ 
Set  $\tau \leftarrow$  A range for the time delay
Repeat
  For each value of  $\tau$ 
     $RP_3 \leftarrow \max(|\mathbf{X}_i - \mathbf{X}_j|)$ , Eq. 5.1, for  $\mathbf{X}_{i+\tau} \leftarrow \mathbf{X}_j$ ;
    While ( $J \neq 0$ )
      do Optimize the cost function  $J$ 
  Until The optimum time delay ( $\tau \leftarrow \tau_{op}$ ) is gained
Set  $\tau \leftarrow \tau_{op}$  and  $d \leftarrow d_{op}$ 
Generate  $RP_3 \leftarrow$  Eq. 5.1, for  $\mathbf{X}_{i+\tau_{op}} \leftarrow \mathbf{X}_j$ 
Output The Embedding Recurrence Plot with Specific Settings.
```

Pseudocode of the derivative-based recurrence plot:

Initialize The Biomedical System's Output $x(t)$ and sample it

Generate The derivative of the sampled time series

$$y(n) \leftarrow \dot{x}(n) \simeq x(n) - x(n-1)$$

Set $M \leftarrow$ total number of points in the time series
 $d' \leftarrow$ number of operations
 $i \leftarrow 1, \dots, N - d' + 1$

Declare $\mathbf{X}_i \leftarrow [x(i), x(i+1), \dots, x(i+d'-1)]$ and
 $\mathbf{X}_j = \dot{\mathbf{X}}_j \leftarrow [\dot{x}(i), \dot{x}(i+1), \dots, \dot{x}(i+d'-1)]$

Set $\mathbf{X}_j \leftarrow 0$ for $\mathbf{X}_j \leq 10^{-6}$

Repeat
 For $i=j \leftarrow 1, \dots, M - m + 1$
 RP \leftarrow Eq. 5.1, for \mathbf{X}_i and $\dot{\mathbf{X}}_i \leftarrow \mathbf{X}_j$

Until $RP_4 \leftarrow$ All matrix is filled

Output The Derivative-Based Recurrence Plot.

Pseudocode of the m-time pattern recurrence plot:

Initialize $x(n) \leftarrow$ Output of the Biomedical System and sample it

Set $d \leftarrow 1$
 $M \leftarrow$ total number of points in the time series
 $m \leftarrow$ number of patterns
 $i \leftarrow 1, \dots, M - m + 1$

Declare $\mathbf{X}_i \leftarrow [x(i), x(i+1), \dots, x(i+m-1)]$ and
 $\mathbf{X}_j \leftarrow \mathbf{X}_j$

Generate RP \leftarrow Eq. 5.1

Declare For $m \leftarrow 2$
 $\mathbf{X}_i \leftarrow [x(i, j), x(i, j), \dots, x(i, j)]$ and
 $\mathbf{X}_j \leftarrow [x(i, j), x(i+1, j+1), \dots, x(i+m-1, j+m-1)]$

For $i \leftarrow 1, \dots, M - m + 1$ and $j \leftarrow 2, \dots, M - m + 2$
 Generate $RP_5 \leftarrow$ Eq. 5.1, for \mathbf{X}_i and \mathbf{X}_j

End

Output The M-Time Pattern Recurrence Plot.

Appendix C

Ph.D. Activities

Awards

1. Ph.D. Scholarship for 3 Years (2011-2014). National Council For Scientific Research in Lebanon (CNRS-L) and the Lebanese University.
2. Best Oral Presentation Award in Biomedical Engineering. Rector of the Lebanese University. 3rd Doctoral Forum, Doctoral School of Sciences and Technology, Lebanese University, July 4th and 5th 2013.



Acquired Ph.D. Courses

1. Advanced Signal and Image Processing. *Responsible: Dr. Jean-Marc Girault*
2. Optical and Molecular Imaging. *Responsible: Dr. Darine Abi-Haidar*
3. Modeling and Simulation Using Matlab. *Responsible: Miss Christelle Suppo*
4. Two French Courses: Intermediate and Advanced Levels. *Responsible: Miss Magali Sabio*
5. Initiation and Processing of a Scientific Text, LateX. *Responsible: Mr. Hubert Cardot*
6. Electronic Thesis Online. *Responsible: Mr. Gerard Bruere*
7. Oral and Written Scientific Communication. *Responsible: Mrs. Patricia Volland-Nail*
8. The Public Research, Which job ? How is it prepared ? *Responsible: Mrs. Anne Cheignon Sred*
9. Promoting Research. *Responsible: Mrs. Anne Cheignon Sred*
10. Knowledge of the Enterprise, Market and Expectations of Recruiters. *Responsible: Miss Anne Cheignon Sred*
11. Mastering Technical and Scientific Information. *Responsible: Mrs. Patricia Volland-Nail*
12. Professional Risk Prevention. *Responsible: Mr. Aline Dingremont*

Appendix D

List of Publications

Journal Publications

Published Articles

A. Zaylaa, J. Charara and J.-M. Girault, "Reducing Sojourn Points From Recurrence Plots To Improve Transition Detection: Application To Fetal Heart Rate Transitions". *Computers In Biology and Medicine, Elsevier*, 2014.
doi: <http://dx.doi.org/10.1016/j.compbimed.2014.09.007>.

A. Zaylaa, J. Charara, S. Ménigot and J.-M. Girault, "Automatic Optimization of Chirp Setting Parameters in Medical Ultrasound Contrast Imaging". *Journal of Life Sciences, USA*, 2013 June, vol. 7, no. 6, pp. 592-598.

S. Oudjemia, **A. Zaylaa**, S. Haddab and J.-M. Girault, "Coarse-Grained Multifractality Analysis Based on Structure Function Measurements to Discriminate Healthy From Distressed Foetuses". *Computational and Mathematical Methods in Medicine*, 2013, vol. 2013, Article ID 152828, 9 pages.
doi:10.1155/2013/152828.

Submitted Articles

A. Zaylaa, S. Oudjemia, J. Charara and J.-M. Girault, "N-order and maximum fuzzy similarity entropy for discrimination of signals of different complexity: application to fetal heart rate signals". *Computers In Biology and Medicine, Elsevier*, 2014, *In revision*.

A. Zaylaa, J. Charara and J.-M. Girault, "Intrauterine Growth Restriction Detection by Statistical Recurrence Plots and Measurements of Fetal Heart Rates". *Physiological Measurements, Institute Of Physics*, 2014, *In preparation*.

Conference Publications

International Conferences

A. Zaylaa, J.-M. Girault and J. Charara, "Unbiased Recurrence Plot Quantification of Chaotic Dynamic Systems By Eliminating Sojourn Points", *In Proceedings 2nd International Conference on Advances in Biomedical Engineering, IEEE*, Tripoli, Lebanon, September 11-13, 2013, pp. 187-190.

S. Oudjemia, **A. Zaylaa**, J. Charara and J.-M. Girault, "Delta-Fuzzy Similarity Entropy to Discriminate Healthy from Sick Fetus", *In Proceedings 2nd International Conference on Advances in Biomedical Engineering, IEEE*, Tripoli, Lebanon, September 11-13, 2013, pp. 1-4.

A. Zaylaa, S. Ménigot, J. Charara and J.-M. Girault, "Automatic Approach Seeking Optimal Frequency Modulation Parameters In Chirp Inversion and Chirp Reversal Ultrasound Contrast Imaging", *Poster In 2nd Science Meeting, Doctoral School of Sciences and Technology, Lebanese University*, Beirut, Lebanon, 2012.

A. Zaylaa, S. Ménigot, J. Charara and J.-M. Girault, "Empirical Optimization Of Frequency Parameters In Chirp Inversion Imaging", *In Proceedings Acoustics 2012*, Nantes, France, April 23-27, 2012, pp. 2889-2893.

A. Zaylaa, J.-M. Girault and J. Charara, "Empirical Research On The Optimization Of The Frequency Parameters Of Chirp Sequences Used In Contrast Ultrasound Imaging", *In Proceedings 1st International Conference on Advances in Biomedical Engineering*, Tripoli, Lebanon, July 6-8, 2011, pp. 59-60.

Appendix E

Journal Papers

n-order and maximum fuzzy similarity entropy for discrimination of signals of different complexity: application to fetal heart rate signals

Amira Zaylaa^{1,2}, Souad Oudjemia³, Jamal Charara² and Jean-Marc Girault¹

¹*University François Rabelais of Tours, UMR Brain-Imaging, INSERM U930, Tours, France*

²*Department of Physics and Electronics, Faculty of Sciences, Lebanese University, Beirut, Lebanon*

³*University of Mouloud Mammeri, Tizzi Ouzou, Algeria*

Abstract-This paper presents two new concepts for discrimination of signals of different complexity. The first focused initially on solving the problem of setting entropy descriptors by varying the pattern size instead of the tolerance. This led to the search for the optimal pattern size that maximized the similarity entropy. The second paradigm was based on the n-order similarity entropy that encompasses the 1-order similarity entropy. To improve the statistical stability, n-order fuzzy similarity entropy was proposed. Fractional Brownian motion and Lorenz time series were simulated to validate the different methods proposed, and fetal heart rate signals were used to discriminate normal from abnormal fetuses. In all cases, it was found that it was possible to discriminate time series of different complexity such as fractional Brownian motion, Lorenz time series and fetal heart rate signals. The best performance in terms of sensitivity (90 %) and specificity (90 %) was obtained with the n-order fuzzy similarity entropy. However, it was shown that the optimal pattern size and the maximum similarity measurement were related to intrinsic features of the time series.

keyword: Maximum similarity, n-order, Fetal heart rate, Entropy, Fuzzy, Complexity, fetal distress.

Reducing Sojourn Points From Recurrence Plots To Improve Transition Detection: Application To Fetal Heart Rate Transitions

Amira Zaylaa^{1,2}, Jamal Charara² and Jean-Marc Girault¹

¹*Department of Medical Biophysics and Imaging, Signal-Imaging Group, Team-5, François-Rabelais University of Tours, France*

²*Department of Physics and Electronics, Faculty of Sciences, Lebanese University, Beirut, Lebanon*

Abstract-The analysis of biomedical signals demonstrating complexity through recurrence plots is challenging. Quantification of recurrences is often biased by sojourn points that hide dynamic transitions. To overcome this problem, time series have previously been embedded at high dimensions. However, no one has quantified the elimination of sojourn points and rate of detection, nor the enhancement of transition detection has been investigated. This paper reports our on-going efforts to improve the detection of dynamic transitions from logistic maps and fetal hearts by reducing sojourn points. Three signal-based recurrence plots were developed, i.e. embedded with specific settings, derivative-based and m-time pattern. Determinism, cross-determinism and percentage of reduced sojourn points were computed to detect transitions. For logistic maps, an increase of 50% and 34.3% in sensitivity of detection over alternatives was achieved by m-time pattern and embedded recurrence plots with specific settings, respectively, and with a 100% specificity. For fetal heart rates, embedded recurrence plots with specific settings provided the best performance, followed by derivative-based recurrence plot, then unembedded recurrence plot using the determinism parameter. The relative errors between healthy and distressed fetuses were 153%, 95% and 91%, respectively. More than 50% of sojourn points were eliminated, allowing better detection of heart transitions triggered by gaseous exchange factors. This could be significant in improving the diagnosis of fetal state.

keyword:Recurrence Plots, Signal-Based Recurrence Plots, Sojourn Points, Dynamic Transitions, Detection, Complexity Analysis, Fetal Heart Rate.

Coarse-Grained Multifractality Analysis Based on Structure Function Measurements to Discriminate Healthy From Distressed Foetuses

Souad Oudjemia^{1,2}, Amira Zaylaa^{1,3}, Salah Haddab² and Jean-Marc Girault¹

¹*University François Rabelais of Tours, UMR Brain-Imaging, INSERM U930, Tours, France*

²*University of Mouloud Mammeri, Tizzi Ouzou, Algeria*

³*Department of Physics and Electronics, Faculty of Sciences, Lebanese University, Beirut, Lebanon*

Abstract-This paper proposes a combined coarse-grained multifractal method to discriminate between distressed and normal foetuses. The coarse-graining operation was performed by means of a coarse-grained procedure and the multifractal operation was based on a structure function. The proposed method was evaluated by one hundred recordings including eighty normal foetuses and twenty distressed foetuses. We found that it was possible to discriminate between distressed and normal foetuses using the Hurst exponent, singularity and Holder spectra.

keyword: Foetal heart rate, coarse-graining, multifractal, distressed foetus, Doppler ultrasound monitor.

Amira ZAYLAA

Analysis and Extraction of Complexity Parameters of Biomedical Signals

Abstract : The analysis of biomedical time series derived from nonlinear dynamic systems is challenging due to the chaotic nature of these time series. Only few classical parameters can be detected by clinicians to opt the state of patients and fetuses. Though there exist valuable complexity invariants such as multi-fractal parameters, entropies and recurrence plots, they were unsatisfactory in certain cases. To overcome this limitation, we propose in this dissertation new entropy invariants, we contributed to multi-fractal analysis and we developed signal-based (unbiased) recurrence plots and unbiased recurrence descriptors based on the dynamic transitions of time series.

Principally, we aim to improve the discrimination between healthy and distressed biomedical systems, particularly fetuses by processing the time series using our techniques. These techniques were either validated on Lorenz systems, logistic maps or fractional Brownian motions which model chaotic and random time series. Then the techniques were applied to real fetus heart rate signals recorded from patients in the third trimester of pregnancy. Statistical measures comprising the relative error, standard deviation, sensitivity, specificity, precision and accuracy were employed to evaluate the performance of detection.

Elevated discernment outcomes were realized by the high-order entropy invariants developed. Multi-fractal analysis using a structure function and coarse-graining enhanced the detection of the medical states of the fetuses. Unbiased cross-determinism invariant developed amended the discrimination process. The significance of our techniques lies behind their post-processing codes which could build up cutting-edge portable machines offering advanced discrimination and detection of Intrauterine Growth Restriction prior to fetal death. This work was devoted to Fetal Heart Rates but time series generated by alternative nonlinear dynamic systems should be further considered.

Keywords : Multi-fractal Analysis, Entropy Quantification, Maximum Entropy, N-Order Entropy, Recurrence Plots, Unbiased Recurrence Plots, Recurrence Quantification Analysis, New Invariants, Discrimination, Detection, Diagnosis, Doppler Ultrasound Fetal Heart Rates, fractional Brownian Motion, Lorenz System, Logistic Map, Correlation Sum, Complexity Analysis, Statistical Signal Processing.

Résumé : L'analyse de séries temporelles biomédicales chaotiques tirées de systèmes dynamiques non-linéaires est toujours un challenge difficile à révéler puisque dans certains cas bien spécifiques les techniques existantes basées sur les multi-fractales, les entropies et les graphes de récurrence échouent. Pour contourner les limitations des invariants précédents, de nouveaux descripteurs peuvent être proposés. Dans ce travail de recherche nos contributions ont porté à la fois sur l'amélioration d'indicateurs multi-fractals (basés sur une fonction de structure) et entropiques (approchées) mais aussi sur des indicateurs de récurrences (non biaisés).

Ces différents indicateurs ont été développés avec pour objectif majeur d'améliorer la discrimination entre des signaux de complexité différente ou d'améliorer la détection de transitions ou de changements de régime du système étudié. Ces changements agissant directement sur l'irrégularité du signal, des mouvements browniens fractionnaires et des signaux tirés du système du Lorenz ont été testés. Ces nouveaux descripteurs ont aussi été validés pour discriminer des fœtus en souffrance de fœtus sains durant le troisième trimestre de grossesse. Des mesures statistiques telles que l'erreur relative, l'écart type, la spécificité, la sensibilité ou la précision ont été utilisées pour évaluer les performances de la détection ou de la classification.

Le fort potentiel de ces nouveaux invariants nous laisse penser qu'ils pourraient constituer une forte valeur ajoutée dans l'aide au diagnostic s'ils étaient implémentés dans des logiciels de post-traitement ou dans des dispositifs biomédicaux. Enfin, bien que ces différentes méthodes aient été validées exclusivement sur des signaux fœtaux, une future étude incluant des signaux tirés d'autres systèmes dynamiques non-linéaires sera réalisée pour confirmer leurs bonnes performances.

Mots clés : Analyse Multi-Fractale, Quantification d'entropie, Entropie Maximale, Entropie d'Ordre-N, Le graphe de Récurrence, Le graphe de Récurrence Nonbiaisés, Analyse de Quantification de Récurrence, Nouveaux Invariantes, Detection, Discrimination, Diagnostiquer, Transition Dynamique, Suite Logistique, Système du Lorenz, Mouvements Brownien Fractionnaires, Coefficient de Corrélation, Analyse de Complexité, Traitement Statistique du Signal, Rythme Cardiac Fœtal Doppler Ultrasonore.