

**UNIVERSIDADE DE SÃO PAULO  
ESCOLA DE ENGENHARIA DE SÃO CARLOS**

**Victor Hugo Batista Tsukahara**

**Dynamic Evaluation of Induced Epilepsy in Rats:  
A Bayesian Network Perspective**

**São Carlos**

**2022**



**Victor Hugo Batista Tsukahara**

**Dynamic Evaluation of Induced Epilepsy in Rats:  
A Bayesian Network Perspective**

Tese apresentada à Escola de Engenharia de São Carlos da Universidade de São Paulo, para obtenção do título de Doutor em Ciências - Programa de Pós-Graduação em Engenharia Elétrica.

Área de concentração: Sistemas Dinâmicos

Advisor: Prof. Carlos Dias Maciel

Trata-se da versão corrigida da tese. A versão original se encontra disponível na EESC/USP que aloja o Programa de Pós-Graduação de Engenharia Elétrica.

**São Carlos  
2022**

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,  
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS  
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pela Biblioteca Prof. Dr. Sérgio Rodrigues Fontes da  
EESC/USP com os dados inseridos pelo(a) autor(a).

BT789a  
a  
Batista Tsukahara, Victor Hugo  
Avaliação Dinâmica da Epilepsia Induzida em  
Ratos: Uma Perspectiva de Rede Bayesiana / Victor Hugo  
Batista Tsukahara; orientador Carlos Dias Maciel . São  
Carlos, 2022.

Tese (Doutorado) - Programa de Pós-Graduação em  
Engenharia Elétrica e Área de Concentração em Sistemas  
Dinâmicos -- Escola de Engenharia de São Carlos da  
Universidade de São Paulo, 2022.

1. Eletroencefalografia. 2. Epilepsia. 3. Sistema  
Complexo. 4. Redes Bayesianas. 5. Volatilidade  
Estocástica. I. Título.

## FOLHA DE JULGAMENTO

Candidato: Engenheiro **VICTOR HUGO BATISTA TSUKAHARA**.

Título da tese: "Avaliação dinâmica da epilepsia induzida em ratos: uma perspectiva de rede Bayesiana".

Data da defesa: 10/10/2022.

### Comissão Julgadora

### Resultado

Prof. Associado **Carlos Dias Maciel**  
(Orientador)

(Escola de Engenharia de São Carlos/EESC-USP)

Aprovado

Prof. Titular **Jorge Alberto Achcar**

(Instituto de Ciências Matemáticas e de Computação/ICMC-USP)

Aprovado

Prof. Dr. **Rafael Naime Ruggiero**

(Faculdade de Medicina de Ribeirão Preto/FMRP-USP)

Aprovado

Prof. Dr. **Jose Manoel de Seixas**

(Universidade Federal do Rio de Janeiro/UFRJ)

Aprovado

Prof. Associado **Ricardo Zorzetto Nicolliello Vencio**

(Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto/FFCLRP-USP)

Aprovado

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica:

Prof. Dr. **João Bosco Augusto London Junior**

Presidente da Comissão de Pós-Graduação:

Prof. Titular **Murilo Araujo Romero**



*This Thesis is dedicated to my wife Cynthia  
and my sons Davi and Maria Clara*





## **ACKNOWLEDGEMENTS**

First and foremost, I thank God for all opportunities provided until today. God is truly good at all times.

I would like to express special thanks to my wife Cynthia who supported me through thick and thin and convinced me to become a PhD scholar. Her love, our life, responsiveness and kids were of paramount importance to do my doctorate.

My advisor, Carlos Maciel, is deeply appreciated for his friendship, all his precious items of advice, talks and patience with me during the whole PhD time. His friendship and lifelong teachings will never be forgotten, as they made me a whole different person.

I would also like to express my deep appreciation for my family, my parents, Walter and Gisela, my aunt Fátima, my grandparents Catharina and Setsuko, my father-in-law Hiro, my mother-in-law Shirley, my brothers Gabriell and Rennan, and my friends in need Daniel, Karen, Lilian and Luiz Paulo.

Thanks for all true friends that I made, specially Jordão and Talysson, for every talk, relaxing moment, kind pieces of advice given, and most importantly for the lifelong friendships forged during this PhD time.

I would like to thank my LPS lab fellows for all the support provided during this PhD time.



*“One of the difficulties in understanding the brain is that  
it is like nothing so much as a lump of porridge.”*

*Richard Gregory*

*“In mathematics you don’t understand things.  
You just get used to them.”*

*Johann von Neumann*



## RESUMO

TSUKAHARA, V.H.B. **Avaliação Dinâmica da Epilepsia Induzida em Ratos: Uma Perspectiva de Rede Bayesiana**. 2022. 124p. Tese (Doutorado) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2022.

A epilepsia é uma das doenças neurológicas mais comuns em todo o mundo. Considerando o cérebro um sistema complexo, estudos tem utilizado esta abordagem para realizar análise de conectividade funcional para indivíduos saudáveis, bem como acometidos pela patologia. A tese apresenta a aplicação das Redes Bayesianas (Dinâmicas) (DBN) para modelar os registros dos Potenciais de Campo Locais (LFP) de ratos induzidos a convulsões epiléticas, avaliando a influência da variável tempo para as análises. Os resultados mostraram que a análise DBN captou a natureza dinâmica da conectividade cerebral através da ictogênese com uma correlação significativa com a neurobiologia derivada de estudos pioneiros que empregavam técnicas de manipulação farmacológica, lesão e optogênese moderna. A avaliação dos arcos sob a abordagem proposta foi consistente com a literatura anterior, propôs novos entendimentos, como a descontinuidade entre o mioclonia de membros inferiores e a dinâmica generalizada da convulsão tônico-clônica (GTCS). Após a incorporação de atrasos entre os registros eletroencefalográficos, houve a indicação de melhor aderência do conjunto de sinais ao modelo DBN.

**Keywords:** Eletroencefalografia, Epilepsia, Sistema Complexo, Redes Bayesianas, Volatilidade Estocástica.



## ABSTRACT

TSUKAHARA, V.H.B. **Dynamic Evaluation of Induced Epilepsy in Rats: A Bayesian Network Perspective**. 2022. 124p. Tese (Doutorado) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2022.

Epilepsy is one of the most common neurological disorders worldwide. Recent findings on it suggest that the brain is a complex system based on a network of neurons whose interactions result in an epileptic seizure, which is currently considered an emergent property. Based on such a modern view, network physiology has emerged to address how brain areas coordinate, synchronize and integrate their dynamics during sound health and afflicted conditions. The objective of this thesis is to present an application of (Dynamic) Bayesian Networks (DBN) to model Local Field Potentials (LFP) based on recordings of rats induced to epileptic seizures and arcs evaluated using an analytical threshold approach. A dynamic network model was constructed from data using the Bayesian Network method, either by considering the delay of communication among brain areas recorded in this study or not. To such an end, the Multivariate Stochastic Volatility method was employed to identify the lag among Local Field Potentials and K2 Score so as to compare the models. Results also showed that the DBN analysis has captured the dynamic nature of brain connectivity across ictogenesis, and that there is a significant correlation to neurobiology derived from pioneering studies employing techniques of pharmacological manipulation, lesion, and modern optogenetics. The arcs evaluation under the proposed approach was consistent with previous literature. Moreover, it provided exciting novel insights, such as a discontinuity between forelimb clonus and generalized tonic-clonic seizure (GTCS) dynamics. Dynamic Bayesian Network depicted the evolution of rats' brains from resting-state until the generalized tonic-clonic seizure. Multivariate Stochastic Volatility captured the lag among brain areas, and better results were yielded after its application on the DBN model.

**Keywords:** Electroencephalography, Epilepsy, Complex System, Bayesian Networks, Stochastic Volatility.





## LIST OF FIGURES

Figure 1 – Example of a detrended Local Field Potential and histogram of rolling variance from a Local Field Potential. . . . .	32
Figure 2 – DAG - D-separation . . . . .	35
Figure 3 – DAG - Fork and collider structure . . . . .	35
Figure 4 – A generic network . . . . .	36
Figure 5 – DAG - a hidden cause . . . . .	40
Figure 6 – DBN example - non-stationary and non-linear dynamics of variables . . . . .	40
Figure 7 – A well-behaved DBN . . . . .	41
Figure 8 – Survival Analysis methodology . . . . .	45
Figure 9 – LFP signal example . . . . .	47
Figure 10 – DBN applied methodology . . . . .	49
Figure 11 – DBN input generation . . . . .	50
Figure 12 – DBN with MSV methodology part 1 . . . . .	54
Figure 13 – DBN with MSV methodology part 2 . . . . .	55
Figure 14 – Lag analysis . . . . .	55
Figure 15 – Kaplan-Meier estimation for MYO-CTRL and NPS-MYO groups . . . . .	59
Figure 16 – Dynamic Bayesian Network developed from the LFP dataset . . . . .	62
Figure 17 – The frequency of main arcs from the DBN presented in Figure 16 . . . . .	64
Figure 18 – PTZ d-separation . . . . .	66
Figure 19 – Dynamic Bayesian Network of rat R052 . . . . .	69
Figure 20 – Gaussian distribution against residual distribution values . . . . .	70
Figure 21 – Example of Markov Chain board. . . . .	94



## LIST OF TABLES

Table 1 – Time registered for each rat from MYO-CTRL and NPS-MYO groups. The MYO column indicates whether the rat had myoclonic behaviour (1) or not (0). CTRL and NPS represents the columns indicating to which group each rat belongs (control or treatment. . . . .	57
Table 2 – Time recorded for each rat from GTC-CTRL and NPS-GTC groups. GTC column indicates whether the rat presented generalized tonic-clonic seizure behaviour (1) or not (0). CTRL and NPS represents the columns indicating to which group each rat belongs (control or treatment). . . . .	58
Table 3 – The strongest arcs identified in developed Dynamic Bayesian Network. All of them have had their minimum frequency calculated through the analytical threshold defined in (GROSS <i>et al.</i> , 2019). Due to the fact that the GTCS time slice is observed only in for the GTCS group, analytical threshold was compared with the frequency observed only for this group. For Basal, Infusion and MYO time slices, the global frequency was used to perform such a comparison. Gray edges were not considered in the table due to the fact that they did not pass in the analytical threshold analysis (Figure 16). . . . .	63
Table 4 – The table depicts the K2 score obtained from the Bayesian Network method. <i>DBN column</i> presents the values of scores calculated from Dynamic Bayesian Networks using the data frames disregarding the lag among brain areas. <i>DBN considering lag column</i> reports the values of scores calculated from Dynamic Bayesian Networks considering the Multivariate Stochastic Volatility lag results, revealing that there is a communication delay among TH, HP, and CX rats brain areas. The score results suggest that the lag implies better DBN model adherence with the dataset - a difference on average of 17.00%. . . . .	68
Table 5 – The table reports the lag among brain areas obtained from the Multivariate Stochastic Volatility method, provided in samples. Due to the frequency of sampling adopted to perform Local Field Potentials acquisition, 1kHz, each sample represents the time of 1ms. . . . .	68
Table 6 – The table reports the K2 scores obtained from the Bayesian Network method considering the best DBN model (rat R052). Score results demonstrate that the R052 Dynamic Bayesian Network model improved the fit with the dataset. Once again, the lag enhanced the mean of the K2 score by approximately 6.00%. . . . .	69



## LIST OF ABBREVIATIONS AND ACRONYMS

AIC	Akaike Information Criterion
AR	Autoregressive
BIC	Bayesian Information Criterion
BN	Bayesian Networks
CDF	Cumulative Density Function
CPU	Central Process Unit
CTRL	Control Group
CX	Cortex brain area
DAG	Directed Cyclic Graph
DBN	Dynamic Bayesian Networks
DMI	Delayed Mutual Information
EEG	Electroencephalography
GABA	$\gamma$ -Aminobutyric acid
GTCS	Generalized Tonic-Clonic Seizure
HP	Hippocampus brain area
KDE	Kernel Density Estimation
LFP	Local Field Potential
LINN <i>ce</i>	Laboratory of Neuroengineering and Neuroscience
MAR	Multivariate autoregressive
MCMC	Markov Chain Monte Carlo
MYO	Myoclonic Seizure
MSV	Multivariate Stochastic Volatility
NIH	National Institutes of Health
NINDS	National Institute of Neurological Disorders and Strokes

NPS	Nonperiodic Stimulation
OMC	Ordinary Monte Carlo
PDC	Partial Directed Coherence
PhD	Philosophy Doctor
PTZ	Pentylentetrazole
RAM	Random Access Memory
TH	Thalamus brain area
UFSJ	Federal University of São João Del-Rei
USA	United States of America
USP	Universiy of São Paulo

## CONTENTS

<b>1</b>	<b>INTRODUCTION</b>	<b>23</b>
<b>1.1</b>	<b>Objectives</b>	<b>27</b>
1.1.1	Specific objectives	27
1.1.2	Organization	28
<b>2</b>	<b>FUNDAMENTAL CONCEPTS</b>	<b>29</b>
<b>2.1</b>	<b>Epilepsy</b>	<b>29</b>
<b>2.2</b>	<b>Functional Connectivity Analysis</b>	<b>30</b>
<b>2.3</b>	<b>Epilepsy Animal Model</b>	<b>30</b>
<b>2.4</b>	<b>Multivariate Stochastic Volatility</b>	<b>31</b>
2.4.1	Model definition	31
<b>2.5</b>	<b>Bayesian Networks</b>	<b>33</b>
2.5.1	Bayes' Theorem	33
2.5.2	Markov Condition	33
2.5.3	d-separation	34
2.5.4	Bayesian Networks Definition	36
2.5.5	Conditional Probabilities	38
2.5.6	Causality and Bayesian Networks	39
2.5.7	Dynamic Bayesian Networks Definition	39
<b>3</b>	<b>EXPLORATORY ANALYSIS METHODOLOGY</b>	<b>43</b>
<b>3.1</b>	<b>Survival Analysis Methodology</b>	<b>43</b>
3.1.1	Kaplan-Meier estimation and Logrank test	43
3.1.2	Animal Experiments and Database	43
3.1.3	Data Analysis	44
<b>4</b>	<b>DYNAMIC BAYESIAN NETWORK METHODOLOGY</b>	<b>47</b>
<b>4.1</b>	<b>Experimental protocol</b>	<b>47</b>
<b>4.2</b>	<b>Algorithms for Dynamic Bayesian Network Methodology</b>	<b>49</b>
<b>4.3</b>	<b>Algorithms for Dynamic Bayesian Network combined with Multivariate Stochastic Volatility Methodology</b>	<b>51</b>
<b>5</b>	<b>EXPLORATORY ANALYSIS RESULTS AND DISCUSSION</b>	<b>57</b>
<b>5.1</b>	<b>Survival Analysis</b>	<b>57</b>
5.1.1	Discussion	58
5.1.1.1	Strengths and limitations	60

<b>6</b>	<b>DYNAMIC BAYESIAN NETWORK RESULTS AND DISCUSSION . . .</b>	<b>61</b>
<b>6.1</b>	<b>Dynamic Bayesian Network . . . . .</b>	<b>61</b>
6.1.1	Discussion . . . . .	62
6.1.1.1	Temporal evolution of DAGs reflect neurodynamics of ictogenesis . . .	62
6.1.1.2	Graph evaluation and analytical threshold to identify DBNs arcs direction	65
6.1.1.3	Other approaches for Functional Connectivity analyses and limitations	65
<b>6.2</b>	<b>Delayed Dynamic Bayesian Network combined with Multivariate Stochastic Volatility . . . . .</b>	<b>67</b>
6.2.1	Discussion . . . . .	70
<b>7</b>	<b>CONCLUSION . . . . .</b>	<b>73</b>
	<b>REFERENCES . . . . .</b>	<b>75</b>
	<b>APPENDIX . . . . .</b>	<b>89</b>
	<b>APPENDIX A – DISSEMINATION ACTIVITIES . . . . .</b>	<b>91</b>
	<b>APPENDIX B – MARKOV CHAIN MONTE CARLO (MCMC) . . . . .</b>	<b>93</b>
	<b>APPENDIX C – PARTIAL DIRECTED COHERENCE . . . . .</b>	<b>99</b>
	<b>APPENDIX D – DELAYED MUTUAL INFORMATION . . . . .</b>	<b>115</b>



## 1 INTRODUCTION

The brain is a complex organ (OLDHAM; BALL; FORNITO, 2022) that is not yet fully understood (THEODONI *et al.*, 2022). There has been interest in better understanding neuroanatomy and the functioning of its circuitry for quite some time. The very first brain studies date back to ancient Egypt (1600 BC), in which the first categorization of brain areas was performed by considering their morphological features. It was unearthed by Edwin Smith (surgical papyrus) in 1862 (OGAWA; PARHAR, 2022), and it has been evolving ever since. An area emerging from all this evolution is network neuroscience in which the concept of a complex system was introduced in order to understand the brain's functional connectivity (GROSS, 2022).

From this perspective, the basic premise is that elements of the system (brain) do not behave in isolatedly, but they interact with each other, therefore giving rise to a repertoire of patterns arising from their linkage (BETZEL, 2022). Thus, the brain is a complex network consisting of nonlinear interactions between neuronal populations from various brain areas (ZHAO *et al.*, 2021), and it is a natural result derived from the concept that real-world phenomena are nonlinear and stochastic (SIGTERMANS, 2020). As examples of emergent properties (JIANG; KUMAR, 2020), it is worth mentioning dynamic memory (EL-GABY *et al.*, 2021; ZHUANG *et al.*, 2021), creative thinking (TURKHEIMER *et al.*, 2021; SCOTT, 2021), behavior (TAKANO, 2022), and brain disorders (BOARETTO *et al.*, 2021; STOJANOVIĆ; KUHLMANN; PIPA, 2020).

Brain disorders represent a global burden issue, since they have been increasingly identified as causes of death and disability (FEIGIN *et al.*, 2019). Epilepsy is a central brain disorder affecting children and adults alike, which can seriously compromise their health and quality of life (OSTENDORF; GEDELA, 2017). It is the second most common neurological disease (ORGANIZATION *et al.*, 2017) afflicting approximately 70 million people worldwide (SPICIARICH *et al.*, 2019), thence representing a public health concern (NIRIAYO *et al.*, 2019). It is estimated that about seventy percent of them can live seizure-free using low-cost and effective antiepileptic drugs (WHO, 2019). It is commonly associated with social difficulties (BEGHI, 2019) and can cause health loss such as premature mortality and residual disability (BEGHI *et al.*, 2019). Temporal lobe epilepsy (TLE), one of its most common forms, is often refractory (30% of the cases) to pharmacological treatments (BORGER *et al.*, 2021; DENG *et al.*, 2021).

Many studies on epilepsy as a network disease have been published in literature (SCOTT, 2021; LEE *et al.*, 2020). FOIT; BERNASCONI; BERNASCONI (2020) outlines the present body of knowledge about epilepsy as a network disorder and indicates the potential benefits of

a network analysis approach to perform preoperative assessment and plan resection surgeries. LIGNANI; BALDELLI; MARRA (2020) used the concept of network disease to emphasize the critical role of homeostatic plasticity in epilepsy therapeutic strategies. The gold standard method to detect sub-clinical seizures and the most prevalent exam in epilepsy-based studies (IBRAHIM *et al.*, 2019; PARVIZI *et al.*, 2018; SHRESTHA; SHRESTHA; THAPA, 2019) is electroencephalography, in which signals (Local Field Potentials) can be acquired via invasive or non-invasive techniques (BIASIUCCI; FRANCESCHIELLO; MURRAY, 2019), nevertheless electrical brain activity is mapped with better spatial resolution through the use of electrodes directly in brain tissues (BARTOLOMEI *et al.*, 2017).

From the standpoint of a complex system, epilepsy is an emergent property resulting from a dynamical coupling of neural oscillators (BOARETTO *et al.*, 2021). Seizures are caused by an imbalance in excitatory and inhibitory synaptic signaling, which results in qualitative changes in physiological parameters manifested as dynamic transitions (TURKHEIMER *et al.*, 2021). From this modeling, epilepsy is understood as a pathologic hyper-synchronization of different brain areas leading to epileptic episodes (BERGLIND; ANDERSSON; KOKAIA, 2018). A related work in literature uses such an approach, Devinsky *et al.* (2018) (DEVINSKY *et al.*, 2018), who have reported hyper-synchronization to discuss epilepsy epidemiology and pathophysiology. Olatam & Akan (2017) (OLAMAT; AKAN, 2017) developed a nonlinear synchronization analysis on Local Field Potential epileptic data. Weiss *et al.* (2019) (WEISS *et al.*, 2019) employed modeling to understand seizure genesis and spreading in human limbic areas.

These definitions were used to develop new therapies, and an electrical Stimulation (ES) of the nervous system arises as a promising alternative to treat refractory epilepsy (MIHÁLY *et al.*, 2020; STRENG; KROOK-MAGNUSON, 2020; OLIVEIRA *et al.*, 2019). The use of ES is an attempt to restore the brain's natural resonance dynamics (COTA *et al.*, 2019). In this context, a novel method has been developed which involves a non-standard form of ES with randomized inter-pulse intervals called non-periodic stimulation (NPS), initially proposed by COTA *et al.* (2009). OLIVEIRA *et al.* (2019), MORAES *et al.* (2019) and COTA *et al.* (2019) discuss it in literature, however, a comparative analysis to evaluate the time to develop an epileptic seizure (using the NPS therapy or not) has never been performed. Thus, the Survival Analysis should be applied due to its extensive use in medicine to evaluate treatments. ROSSELLO; GONZÁLEZ-DEL-HOYO (2021) discuss its use in cardiovascular research, WANG *et al.* (2021) develop a system which performs a survival analysis of real-world datasets of patients suffering from IB-IIA stage lung cancer, and ZHAN *et al.* (2021) perform a survival analysis of patients with hepatocellular carcinoma and distant metastasis.

Understanding the dynamics of epileptogenic phenomena might aid in enlightening their neurobiological mechanisms (NELSON; BONNER, 2021). Epileptic processes have spatiotemporal dynamics (EL-GABY *et al.*, 2021; YANG *et al.*, 2021; TURKHEIMER *et al.*, 2021) whose

characteristics were gleaned from observing the evolution of state variables in a dynamical system known as the human brain (YANG *et al.*, 2021). The time series analysis is commonly used to assess the dynamics and features of interactions, in which a set of univariate time series - Local Field Potentials as variables - is employed to characterize the complex system under study (JIA *et al.*, 2020). Glomb *et al.* (2021) discuss computational LFP models to obtain a comprehensive understanding of mechanisms underlying the Local Field Potential signal (GLOMB *et al.*, 2021). Snyder *et al.* (2021) used LFP to study the resting state networks in patients who had a stroke (SNYDER *et al.*, 2021), and Nobukawa *et al.* (2020) utilized LFP signals to discuss classification methods based on their complexity and synchronization to study Alzheimer's disease (NOBUKAWA *et al.*, 2020). Such investigation permeates the existence, directionality and strength of influences exerted by interactions, and subjects arising from causal inference (SIGTERMANS, 2020). The dynamics of epileptogenic phenomena can be captured through the tradeoff between pure statistical dependencies (functional connectivity) and the pattern of direct causal connections (effective connectivity) among variables (KOŘENEK; HLINKA, 2020). In this sense, a set of issues to be handled springs to mind, such as variables that change over time (MANSOURI *et al.*, 2020), state space reconstruction (JIA *et al.*, 2020) and conditional dependence (MUKHERJEE; ASNANI; KANNAN, 2020).

Spatial and temporal aspects of brain disorders, particularly epilepsy, and their functional connectivity analyses are crucial for more in-depth investigations. The patterns of brain communication found during seizures are distinct from those observed in healthy conditions (COURTIOL *et al.*, 2020). A synchronization of the brain network, also known as a hyper synchronization phenomenon (BOARETTO *et al.*, 2021) to describe epileptic seizures, is a proper manner of time coordination of the brain's network states and the dynamics of its elementary components, i.e. the neurons (NOWAK *et al.*, 2017). Additionally, it is crucial to realize that communication across brain areas requires some time delay (PARIZ *et al.*, 2021), which is necessary for maintaining the brain's health or keeping the pathological network synchronization intact (PETKOSKI; JIRSA, 2019). Based on the physician's observed temporal delay pattern, it is possible to determine the seizure initiation zone during ictal activity (MYERS *et al.*, 2020). Time delay is an inherent property of natural systems such as human or animal physiology (ZHONG *et al.*, 2018)

Different analyses had been proposed to understand the dynamic evolution of epileptic seizures based on autoregressive models such as the Partial Directed Coherence (PDC) (CIARAMIDARO *et al.*, 2018), or non-parametric approaches as the Mutual Information (MI) (GRIBKOVA; IBRAHIM; LLANO, 2018) and Bayesian Networks (MICHIELS; LARRAÑAGA; BIELZA, 2021). However, the objective of this thesis is to understand these relationships and structure from data, which has been reached using a Dynamic Bayesian Network. A Bayesian Network (BN) is a compact representation of statistical dependencies among variables (MICHIELS; LARRAÑAGA; BIELZA, 2021; BIELZA; LARRANAGA, 2014; KOLLER; FRIEDMAN, 2009; NEAPOLITAN *et al.*, 2004). BNs are probabilistic models defined by a Directed Acyclic Graph (DAG) and conditional probabilities tables (CPT) representing the probabilistic dependence over

signals. Dynamic Bayesian Networks (DBNs) can model signals as BN in successive time-slices (LEÃO *et al.*, 2021; ROBINSON; HARTEMINK, 2010; MURPHY, 2002). One of its main advantages is that it allows probabilistic reasoning under uncertainty to uncover relationships within functional connectivity analysis (BENJUMEDA *et al.*, 2021; BIELZA; LARRANAGA, 2014).

The use of BNs and DBNs in neuroscience is verified in literature for different purposes (BIELZA; LARRANAGA, 2014). ELDAWLATLY *et al.* (2010) performed a study to find the dynamic connectivity between cortical neurons. SMITH *et al.* (2011) inferred a non-linear communication association between regions of the brain; ESCH *et al.* (2020) used a Bayesian method to evaluate the effective connectivity of brain networks in order to detect the Mozart effect; SIP *et al.* (2021) developed a data-driven Bayesian Inference method to detect the seizure propagation patterns in an epileptic brain through intracranial electroencephalography.

Bayesian Networks are fundamental tools in a multidisciplinary analysis, since their graph output can be easily interpreted by specialists of different areas (MOREIRA *et al.*, 2021; CHEN *et al.*, 2020) and, if combined with the multivariate statistical dependence performed by the model, it is a good and reliable way of unifying the algorithmic knowledge from data acquired from a specialist's knowledge and an interpretation of results.

According to CHOWDHURY *et al.* (2021), the direction of associations between the nodes formed in areas of the brain during epileptic seizures is still an unresolved issue (COLMERS; MAGUIRE, 2020; GIL *et al.*, 2020). TRACY *et al.* (2021) showed that this direction could change during the seizure. As a matter of fact, FOIT; BERNASCONI; BERNASCONI (2020) showed that these directions are associated with two critical processes: the generation and expression of the seizure and the maintenance of the epileptogenic phenomenon (LIGNANI; BALDELLI; MARRA, 2020). This thesis considers the evolution of brain communication during a seizure pattern (or ictogenesis) from a basal state until a generalized tonic-clonic seizure (GTCS), and the DBN method is applied to elucidate the link among brain areas for each stage of the process. To clarify the communication among brain areas during each time slice, as well as elucidating the link between states, such as the relationship of basal and GTCS intervals, rats were exposed to a pentylenetetrazole (PTZ) pro-convulsant drug to induce ictogenesis while their brain's local-field activity was recorded and later analyzed by a DBN model in this study.

A detailed model from a set with few instances possessing many attributes, such as time slices and measured variables, is not expected. However, it does not mean that the collected dataset lacks essential and helpful information, e.g. association trends among the main variables. As reported by KOLLER; FRIEDMAN (2009), a strategy to overcome the problem of scarce data and develop a reliable structure of arcs among the nodes of a BN can be the apprenticeship of many high-score structures followed by a consolidation of results. This paper follows this rationale and suggests to use the analytical threshold proposed by GROSS *et al.* (2019) throughout the dataset of all experiments, thus highlighting the importance of a sampling approach. The

method captures the expected associations among nodes and also achieves better prediction performance than the BNs learned from neighbors' thresholds to computed thresholds GROSS *et al.* (2019), such as the method to identify significant arcs proposed by SCUTARI; NAGARAJAN 2013.

As mentioned previously, there are several studies in literature on the use of Bayesian Networks (BN) and Dynamic Bayesian Networks (DBN). TSUKAHARA *et al.* (2022) used Dynamic Bayesian Networks to model Local Field Potentials (LFP) recordings of rats induced to epilepsy and the arcs were evaluated using an analytical threshold approach. BLASI; CAMPAGNA; FINAZZI (2021) used the DBN method to predict organ failure associations in the absence of predefined outcomes. RUIZ-PÉREZ *et al.* (2019) used BN to investigate the relationship between several neuromuscular performance parameters and dynamic postural control. However, none of which considered minor time delays between variables within the BN or DBN time-slices analysis. The variables are simultaneously instantaneous prior to being incorporated into BN or DBN methods. This argument contradicts the spatiotemporal property observed in realistic systems.

By simulating the delays found in the brain, we hypothesize that a more representational model can be created using a dynamic bayesian network. We might incorporate these delays into a DBN by evaluating them through Multivariate Stochastic Volatility (MSV) (ROBINSON; HARTEMINK, 2010; ZHONG *et al.*, 2018), as the volatility of Local Field Potentials exhibits a stochastic time series behavior (MURPHY, 2002). The method's strength is its use of latent variables to analyze time series, based on the premise that time series cannot be effectively predicted just on the basis of past signal value analysis (BIELZA; LARRANAGA, 2014; BENJUMEDA *et al.*, 2021). Additionally, the method utilizes a multivariate model to analyze the dependencies and time delays between the Local Field Potentials time series (MURPHY, 2002), which is critical for analyzing epileptic disease dynamics.

## 1.1 Objectives

The objective of this thesis is to combine the Dynamic Bayesian Network method with Multivariate Stochastic Volatility aiming to develop a dynamic network model from a dataset of Local Field Potentials of rats. A functional connectivity analysis of three brain areas (hippocampus, thalamus and cortex) considers the effect of slight delays on model results. The main contribution of this study is to incorporate minor delays inside each DBN time slice so as to support the neurophysiology of epilepsy seizure dynamics considering an essential feature in the suggested methodology to analyze the temporal evolution of the phenomenon.

### 1.1.1 Specific objectives

The specific objectives of the thesis are reported as follows:

- Performing an exploratory analysis of Local Field Potentials dataset through the development of a Survival Analysis to evaluate the effectiveness of non-periodic stimulation in order to treat refractory epilepsy;
- Using Dynamic Bayesian Network to study the functional connectivity among brain areas during the evolution of epileptic seizures, as well as evaluating arcs from the DBN through an analytical threshold;
- Considering the natural minor delays that are present in the functional connectivity among brain areas, it is proposed to use MSV method in order to identify the lags of communication prior to its incorporation inside the DBN method;
- Based on results, it is performed a comparison between the DBN obtained from the LFP signals, both with and without minor delays of communication.

### 1.1.2 Organization

A brief theory about fundamental concepts are presented in Section 2: epilepsy, functional connectivity analysis, epilepsy animal model, multivariate stochastic volatility and Bayesian Networks. Section 3 presents the methodology for each specific objectives and Section 4 reports the results and discussions. Section 5 depicts the dissemination activities promoted during the PhD program, and Section 6 presents the conclusions and suggested studies.

## 2 FUNDAMENTAL CONCEPTS

The chapter describes the fundamental concepts of the present PhD thesis. A brief description about epilepsy is made to understand the context of signals database and the purpose of analyses of study cases. Afterwards, the methods to perform functional connectivity analyses are described: Bayesian Networks and Multivariate Stochastic Volatility.

### 2.1 Epilepsy

Epilepsy is a brain disorder characterized by a sustainable predisposition to generate epileptic seizures, and social, psychological, cognitive and neuro-biological consequences induced by the disease (FISHER *et al.*, 2014). It is one of the most prevalent neurological disorders (SOUZA *et al.*, 2018), based on an estimation that about 70 million people worldwide suffer from this condition (SPICIARICH *et al.*, 2019), and approximately 80% of cases are found in developing countries (DISEASES *et al.*, 2005). The prevalence rate in low and middle-income countries is 139 out of 100 thousand inhabitants, and 49 out of 100 thousand citizens in high-income countries (FIEST *et al.*, 2017).

The National Institute of Neurological Disorders and Strokes (NINDS) is a branch of National Institutes of Health (NIH) and the major source of funding medical research in the United States. (MEADOR *et al.*, 2011) published a neurological disease prevalence study in 2011 and revealed that epilepsy rate is about 7.1/1000 people. In Brazil, there are sparse statistics about epilepsy (SAMPAIO *et al.*, 2010), although a study published in 2007 has revealed an estimated prevalence rate of 9.2/1000 people (NORONHA *et al.*, 2007).

There are many mechanisms concerning ictogenesis ((BLAUWBLOMME; JIRUSKA; HUBERFELD, 2014)). Nonetheless, a well-accepted theory is related to the unbalance between excitatory and inhibitory neural tonus, in which the former is prevalent (FISHER *et al.*, 2005; KAILA *et al.*, 2014; STALEY, 2015). Furthermore, another commonly discussed approach in literature is synchronization (COTA *et al.*, 2016; MEDEIROS; MORAES, 2014; LEHNERTZ *et al.*, 2009). This perspective models the brain as a system governing other systems, which implies understanding the organ as a set of subsystems (brain areas) interacting among themselves - there is a synchronization from a particular brain region and other areas - thus giving rise to emergent properties (i.e. synchronization) (COTA *et al.*, 2016; AVENA-KOENIGSBERGER BRATISLAV MSIC, 2017). Thus, epileptic seizure is perceived as hyper-synchronization phenomena under this scope (COTA *et al.*, 2016).

In neuroscience literature, there are many studies on using a synchronization approach (JIRUSKA *et al.*, 2012; OLAMAT; AKAN, 2017; DHAMALA; RANGARAJAN; DING, 2008; QIN; WANG, 2008), some of which are specifically aimed to comprehend brain dynamics during

seizures using neural signals - Local Field Potentials (LFP) - processing (PHAN *et al.*, 2019). Statistics is applied to grasp the interactions between brain areas and model phenomena. It is important due to the fact that there might be statistical inference, which supports the development of new strategies and treatments for epilepsy.

The neural substrates used for electrofisiological mapping were hippocampus (HP), thalamus (TH), and cortex (CX). These structures were chosen for their critical role in the generation, propagation, and maintenance of epileptic crises (BERTRAM, 2014). The thalamus has connections and interactions with various areas of the brain, including the amygdala and the hippocampus (YANG *et al.*, 2022). He also has reciprocal connections with all areas of the cortex and an uncontrollable tendency to fire rhythmic bursts of potential action, acting as a marker for epileptic seizures (BERTRAM *et al.*, 1998).

In terms of epilepsy symptomatology, clinical manifestations are directly related to the region affected by the abnormal activity and may thus present in a variety of forms, including motor, sensory, autonomic, and even psychic disorders (BENICZKY *et al.*, 2022). An increase or decrease in muscle contraction is the motor manifestation. The increase can be tonic (continuous contractions), clonic (oscillations between contractions and relaxations), or myoclonic (very brief muscle contractions that characterize shock-like concussions). Atonic seizures are characterized by a decrease in muscle contraction and an abrupt loss or decrease in muscle tone (BENICZKY *et al.*, 2022).

## 2.2 Functional Connectivity Analysis

(ELLENBROEK; YOUN, 2016) defines functional connectivity as the statistical relationship among a given physiologic set of signals considering the time domain capable of being acquired using techniques, such as electroencephalography (EEG), among others. Considering the context of this Phd thesis, functional connectivity is the temporal relationship among a set of signals from different brain areas (STEPHAN; FRISTON, 2009).

To access such feature, two important methods are going to be used in this PhD thesis, Bayesian Networks and Multivariate Stochastic Volatility. The following section is going to discuss all aforementioned methods to deepen understanding about the presently proposed methodology.

## 2.3 Epilepsy Animal Model

The epilepsy model chosen for this thesis was acute seizures induced by controlled pentylenetetrazole infusion (PTZ). In experimental animals, the gradual infusion of intravenous PTZ provides a model of the transition of susceptibility states to different levels of neural excitability/synchronism, triggering anatomically-dependent convulsive behaviors. PTZ intravenous infusion causes drug levels in the brain to rise evenly, resulting in less variability in



the onset of seizure behaviors among animals (SAMOKHINA; SAMOKHIN, 2018). This medication is a central GABAergic antagonist that can be used to test antiepileptic drugs or non-pharmacological therapies. The rat can experience two kinds of crises: minimal and maximal. The former are induced by doses as high as 40 mg/kg and manifest as spasms, facial automatisms, masticatory movements, and clonias of the head and forelimbs (OLIVEIRA, 2017). This type of seizure involves structures in the forebrain (EELLS *et al.*, 2004). Doses greater than 60 mg/kg cause maximal seizures, which manifest as generalized tonic-clonic (GTC) seizures (VELISEK *et al.*, 1992). These involve structures in the forebrain and brainstem and are always preceded by minor seizures (EELLS *et al.*, 2004).

## 2.4 Multivariate Stochastic Volatility

Parametric methods, such as Multivariate Stochastic Volatility (MSV), increase statistical power by simultaneously representing the spatial and temporal structure of data (PHAN *et al.*, 2019). Volatility is a measure of variance that changes dynamically. In a multivariate case, a multivariate autoregressive (MAR) model is used to accommodate the average of joint processes, and variance is evaluated by considering whether it is a model of Granger causality or not (WANG; CHAN, 2021). In this context, the coupling matrix is determined for the best delays by capturing the short time delay between them, which is critical for understanding hyper brain synchronization during epileptic seizures (MCCAUSLAND; MILLER; PELLETIER, 2021; PETKOSKI; JIRSA, 2019). The MSV method is a nonlinear and non-Gaussian state-space model (GONG; STOFFER, 2021), which corresponds to the concept of nonlinearity in natural phenomena (SIGTERMANS, 2020). Local Field Potentials (LFPs) are nonlinear signals with kurtosis and heteroscedasticity (KIPINSKI; KORDECKI, 2021; XIANG *et al.*, 2020). The MSV model's multivariate nature enables the estimation of dependencies between LFPs (PHAN *et al.*, 2019).

The MSV method models the volatility of Local Field Potentials as a latent variable vector autoregressive process by converting the original signals to a log return time series suitable for inclusion in the present approach (BOUSSAHA; HAMDI, 2018). Figure 1.a depicts a detrended Local Field Potential used to calculate its volatility and then the log return series is used as input of the MSV method. Figure 1.b shows the histogram of the rolling variance of a Local Field Potential using 128 bins. It should be observed that log-normal distribution fits the second-moment distribution of the signal.

### 2.4.1 Model definition

This study is based on the MSV model proposed by Phan et al. (2019) (PHAN *et al.*, 2019) which is a generalization of the Multivariate Stochastic Volatility (MSV) model proposed by Harvey et al. (1994) (HARVEY; RUIZ; SHEPHARD, 1994). Such an approach is necessary to build a full coefficient matrix and achieve the purpose of a functional connectivity analysis

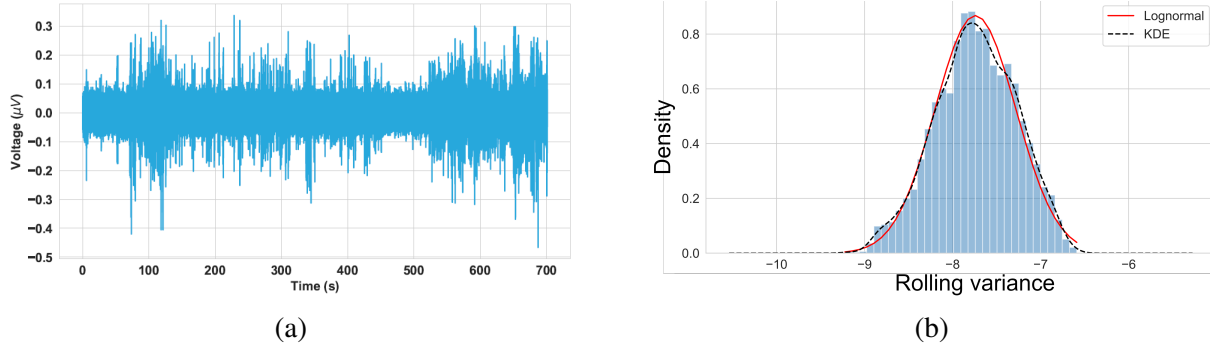


Figure 1 – (a) Example of a detrended Local Field Potential. It is used to calculate the log return from the rolling variance, the input of MSV model. (b) Histogram of rolling variance from a Local Field Potential. The signal is detrended, then it is performed the calculation of a rolling variance. It should be observe the fit with the empirical method of Kernel Density Estimation (KDE). There is also good adherence while employing Lognormal distribution fit. The results are in accordance with Phan et al. (2019) (PHAN *et al.*, 2019) study results.

among brain areas covered in this study, particularly to observe their lag of communication.

Basically, the model represents a set of Local Field Potentials at a given time  $t$  ( $q_\gamma[t]$ ) as a vector of dimension  $1 \times \gamma$ , in which  $\gamma$  is a set comprising each brain area involved in this study -  $\gamma = [1, 2, \dots, \Lambda]$ , where  $\Lambda$  is the  $\gamma^{th}$  brain area for instance. Therefore,  $q_\gamma[t] = [q_1[t], q_2[t], \dots, q_\gamma[t]]$ . Each  $q_\gamma[t]$  element can be modeled as

$$q_\gamma[t] = \exp\left(\frac{r_\gamma[t]}{2}\right) \eta_\gamma^q[t] \quad (2.1)$$

where  $\eta_\gamma^q[t]$  is an error term modeled as an independent normal distribution,  $r_\gamma[t]$  is the log-volatility of the  $n$ th brain area at time  $t$  defined as

$$r_\gamma[t] - \mu_\gamma = \sum_{j=1}^{\zeta} \sum_{k=1}^{\Lambda} \beta_{\gamma,k} \circ (r_k[t-j] - \mu_k) + \eta_\gamma^r[t] \quad (2.2)$$

$\eta_\gamma^r[t]$  is the error term associated with log-volatility modeling, which is independent and normally distributed. A summation indicates a  $\zeta^{th}$ -order autoregressive (AR) terms -  $j = 1, 2, \dots, \zeta$  - referring to the present volatility of  $r_\gamma[t]$  with its previous past and also relating to the past volatility values associated with other brain areas.  $\mu_k$  and  $\mu_\gamma$  are unconditional average volatilities associated with a given brain area.  $\beta_{\gamma,k}$  is the variable related to the model of how past values of other brain areas ( $k$ ) affect a given area  $\gamma$  at time  $t$ , a correlation matrix ( $\beta$ ). This model can be built based on the assumption that  $q_\gamma[t]^T = [q_1[t], \dots, q_\Lambda[t]]^T$  - the Local field Potentials - are conditionally independent given their log-volatilities  $r_\gamma[t]^T = [r_1[t], \dots, r_\Lambda[t]]^T$  based on the following definitions:

- $\eta_\gamma^q[t] = [\eta_1^q[t], \dots, \eta_\Lambda^q[t]] \sim N(0, I_\gamma)$ , where  $I_\gamma$  is the identity matrix of rank  $\Lambda$ ;

- $\eta_\gamma^r[t] = [\eta_1^r[t], \dots, \eta_\Lambda^r[t]] \sim N(0, \Psi)$ , where  $\Psi = \text{diag}(\sigma_1, \dots, \sigma_\Lambda)$ ;
- $\mu_\gamma = [\mu_1, \dots, \mu_\Lambda]$ .

It is possible to rewrite Equation 2.2 in the form of a matrix as follows

$$r_\gamma[t] - \mu_\gamma = \beta(r_\gamma[t-1] - \mu_\gamma) + \eta_\gamma^r[t] \quad (2.3)$$

All model parameters must be estimated through simulation since they are initially presumed as unknown.

## 2.5 Bayesian Networks

### 2.5.1 Bayes' Theorem

The Bayesian approach is a statistical inference technique that uses the Bayes theorem to update the probability of occurrence as new data or information becomes available. Formally, a random variable  $x_i$  is determined by the sample space's mutually exclusive events  $A_1, A_2, \dots$ , and  $A_k$ .  $\Omega$ , i.e.  $\bigcup_{j=1}^k A_j = \Omega$  and  $A_i \cap A_j = \emptyset$  when  $i \neq j$ . As a result,  $P\left(\bigcup_{j=1}^k A_j\right) = \sum_{j=1}^k P(A_j) = 1$ . According to Bayes' theorem, for any event  $y$ ,  $P(y)$  must equal zero and  $P(A_i)$  must equal zero for all  $i$ :

$$P(A_i|h) = \frac{P(h|A_i) P(A_i)}{\sum_{j=1}^k P(h|A_j) P(A_j)}, \quad (2.4)$$

for all  $i$  values between 1 and  $k$  (PUGA; KRZYWINSKI; ALTMAN, 2015). In other words, before receiving any information on event  $h$ ,  $P(A_i)$  is the presumed prior probability for  $A_i$ . The probability of  $A_i$  is updated when event  $y$  occurs, i.e.  $P(A_i|h)$  is the probability of  $A_i$  conditional on the occurrence of event  $h$ . The term "conditioned probability" or "posterior probability" refers to this modified probability.

When there are more variables, it becomes impossible to directly compute conditional probabilities using Bayes' theorem (NEAPOLITAN *et al.*, 2004). An approach for computing conditional probability values that takes into account the possibility of all variables being related demands exponential space. To do probabilistic inference using conditional probabilities, it is important to identify features that are associated with a direct influence. Bayesian networks (BN) were created to address these issues.

### 2.5.2 Markov Condition

A Directed Graph (DG) is a pair  $(V, E)$  in which  $V$  is a finite and non-empty collection of vertices and  $E$  is a set of distinct pairings  $(X, Y)$  in which  $X \rightarrow Y$  are both in  $V$  and there is an edge connecting  $X$  and  $Y$ . In this situation,  $X$  and  $Y$  are considered as being close, with  $X$  serving as the parent of  $Y$  and  $Y$  serving as the descendant of  $X$ . Nodes  $\{X_2, X_3, \dots, X_n\}$

are referred to as interior nodes in the set  $\{X_1, X_2, \dots, X_n\}$ , given the association rule ( $X_{i-1} \rightarrow X_i, I \geq 2$ ). The sub-path from  $X_i$  to  $X_j$  in  $X_1, X_2, \dots, X_n$  is  $X_i, X_{i+1}, \dots, X_j$ . A directed cycle is a path connecting two nodes. A Directed Acyclic Graph (DAG)  $G$  is a directed graph in which there are no directed cycles (NEAPOLITAN *et al.*, 2004).

**Definition 1** *Let us Suppose that we have a joint probability distribution  $P$  of random variables in a given set  $V$  and DAG  $G = (V, E)$ . Then,  $(G, P)$  satisfies the Markov condition for each variable  $X \in V$ , and  $X$  is independently conditional to the set of all its non-descendants  $ND_X$ , given the set of all its parents  $PA_X$ , i.e.*

$$X \perp\!\!\!\perp ND_X | PA_X \quad (2.5)$$

**Theorem 2.5.1** *If  $(G, P)$  satisfies the Markov condition, then  $P$  is equal to the product of all nodes' conditional distributions based on their parents' values and whenever these conditional distributions exist, i.e. if the set of nodes  $X_1, X_2, \dots, X_n$  is ordered in ancestral ordering (if  $X_k$  is a descendant of  $X_j$ , it appears later in the ordering).*

$$P(x_1, x_2, \dots, x_n) = P(x_n | pa_n) P(x_{n-1} | pa_{n-1}) \dots P(x_1 | pa_1) \quad (2.6)$$

in which the set of values  $\{x_1, x_2, \dots, x_n\}$  represents the states of variables  $\{X_1, X_2, \dots, X_n\}$ , and  $pa_i$  are the subsets of these values containing values of  $X_i$ 's parents.

### 2.5.3 d-separation

d-Separation (*d* for *directional*) is a connection criterion for a set  $X$  and a set  $Y$  of nodes in a DAG that is conditional on a set  $Z$  of nodes. It is fundamental to any algorithm attempting to determine the structure of a DAG considering real data, and it has significant implications for a causality analysis. There are three primary forms of a d-separation analysis, and each of them can be used to construct a more generic analysis.

The first of the main structures to analyse d-separation is the *causal chain* (Figure 2). In this structure, it is true that

$$X \perp\!\!\!\perp Z | Y \quad (2.7)$$

i.e., even though  $X$  is ancestral of  $Z$ , thus there is dependence between these variables, the presence of  $Y$  as a parent of  $Z$  and a descendant of  $X$  makes them conditionally independent.

The next case is the *common cause* or *fork* (Figure 3a). Then, it is still true that:

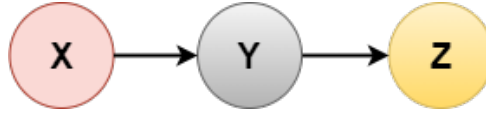


Figure 2 – In the causal chain structure, the middle node d-separates the other two.

$$X \perp\!\!\!\perp Z|Y \quad (2.8)$$

i.e. Despite the fact that  $X$  and  $Z$  share a common ancestor, which results in statistical dependence, they are conditionally independent from their common parent  $Y$ .

The *collider* is the final major structure (Figure 3b). This is a more complex case:  $X$  and  $Z$  are unambiguously independent.  $Y$ , on the other hand, d-connects  $X$  and  $Z$ , i.e.:

$$X \not\perp\!\!\!\perp Z|Y \quad (2.9)$$

Thus, a collider d-connects its parents when they are self-contained. This results in the so-called *Berskon's paradox*, which states that if  $P(X)$  represents the probability of event  $X$  and  $P(Z)$  represents the probability of event  $Z$ , then:

$$P(X|Z) = P(X) \Rightarrow P(X|Z, X \cup Z) < P(X|X \cup Z) \quad (2.10)$$

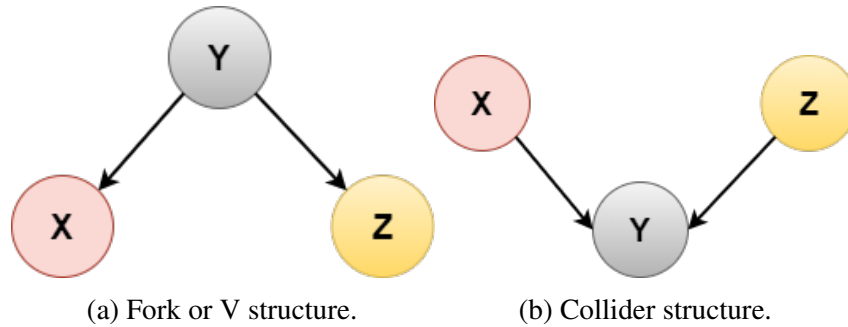


Figure 3 –  $Y$  denotes the d-separation of  $X$  and  $Z$  in a fork structure (3a). It d-connects  $X$  and  $Z$  for the collider 3b ( $Y$ ). The following experiment clarifies this counter-intuitive fact: assuming that  $X$  and  $Z$  are random binary sequences and that  $Y = f(X, Z)$ . For instance, given that  $Y = X \text{OR} Z$  and  $k$  equal the  $k$ -th entry in either of the three lists,  $X$  and  $Z$  are clearly independent, as both sequences are produced randomly; however, when both sequences are conditional on the knowledge of  $Y$ , they become dependent, as  $x_k \text{OR} z_k = 1$  if  $y = 0$  and  $x_k = z_k = 0$  if  $y_k = 0$ .

The d-connectivity of any node which is conditionally connected to a third set can be generalised using these three constructs. For instance, in Figure 4, various instances of analysis are provided:

- $Z$  and  $Y$  are d-separated and conditional on  $\emptyset$ , since they are independent;

- $Z$  and  $Y$  are d-connected and conditional on  $W$ , since they are independent,  $W$  d-connects  $X$  and  $Z$ , and  $X$  is d-connected with  $Y$ ;
- $Z$  and  $Y$  are d-connected and conditional on  $U$ , since they are  $U = f(W)$ ;
- $Z$  and  $Y$  are d-separated and conditional on  $\{X, W\}$ , since they are d-connected by  $X$  when conditional on  $W$ ;
- $U$  and  $Y$  are d-separated and conditional on  $W, X$  or  $\{X, W\}$ , since both  $X$  and  $W$  act as a fork between  $Y$  and  $U$ .

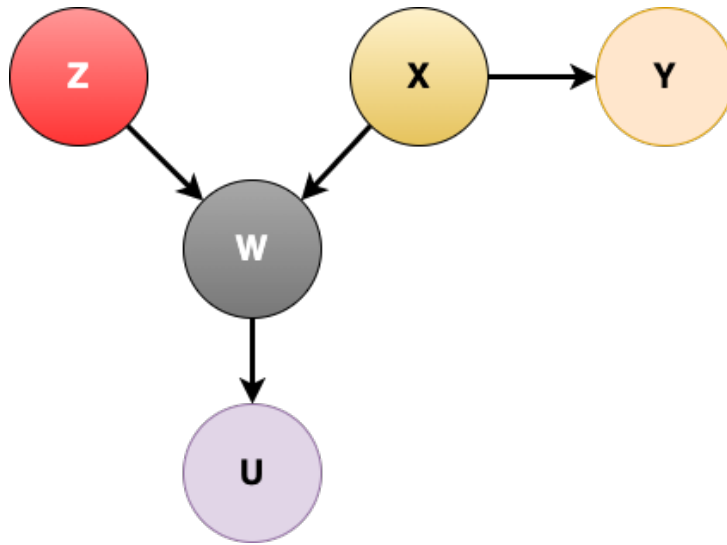


Figure 4 – A generic network to show how to use the three main structures, i.e. fork, collider and causal chain, to analyse d-separation

#### 2.5.4 Bayesian Networks Definition

Once again, let  $P$  be the joint distribution function of random variables in the set of graph vertices  $V$ . If  $(G, P)$  satisfies the Markov requirement,  $(G, P)$  is considered a Bayesian Network (BN). For historical and algorithmic reasons, a BN is most commonly referred to as the pair  $(G, \Theta)$ , where  $G = (V, E)$  is the matching DAG and  $\Theta = \theta_1, \theta_2, \text{etc}, \theta_n$  denotes a collection of conditional probability distributions encoded in DAG  $G$ . For every BN with a joint distribution  $P$  and vertices  $V$  on a collection of random variables  $X_1, X_2, \dots, X_n$ , a factorization from Equation 2.6 can be restarted as follows:

$$P = \prod_{x \in \{X\}} p(x|\text{pa}_x) \quad (2.11)$$

Bayesian Networks are used in a variety of fields consisting of knowledge-based systems, including speech recognition (NASEREDDIN; OMARI, 2017; ZWEIG, 2003; ZWEIG; RUSSELL,

1998), climate studies (LEE *et al.*, 2020; MOE; HAANDE; COUTURE, 2016), population studies (CAI, 2017; MCCANN; MARCOT; ELLIS, 2006), machine learning (JIANG; CAO; DENG, 2019; CHATURVEDI *et al.*, 2018), and neuroscience (BLASI; CAMPAGNA; FINAZZI, 2021; RUIZ-PÉREZ *et al.*, 2019). It has the advantage of embedding information about multivariate system interaction in dependencies indicated by its edges. However, there are several difficult issues arising from the application of the BN technique using real-world data.

Given a dataset with variables  $\{X_1, X_2, \dots, X_n\}$ , the most likely structure  $G$  is found by a *learning algorithm*. The number of possible DAGs with  $n$  vertices grows in a super-exponential way, given the recurrence relationship (ROBINSON, 1977):

$$G(n) = \sum_{k=1}^n (-1)^{k+1} \binom{n}{k} 2^{k(n-k)} G(n-k) \quad (2.12)$$

There are also lower and upper bound (STANLEY, 1973):  $G_{lower} = 2^{\frac{n(n-1)}{2}}$  and  $G_{upper} = n! 2^{\frac{n(n-1)}{2}}$ . Such enormous expansion (superexponential) in the number of possible DAGs as a function of the number of vertices suggests that an exhaustive search is impossible. There are various strategies for determining a reasonably good fit to the data. These algorithms are based on scoring the edges of probable DAGs and employing an evolutionary algorithm, such as Particle Swarm Optimization (KOUZIOKAS, 2020), Hill-Climbing Search (TSAMARDINOS; BROWN; ALIFERIS, 2006), or Genetic Algorithms (CONTALDI; VAFEE; NELSON, 2019; LARRANAGA *et al.*, 1996). Thus, a learning algorithm consists of two parts: a scoring measure and a search technique. This raises the question of how to construct a score-based system capable of selecting a super-exponential DAG universe using any search technique.

Cooper (1992) (COOPER; HERSKOVITS, 1992) created a score-function for learning BN models from data based on four assumptions: the database variables are discrete, cases occur independently, there are no circumstances in which variables have missing values prior to observing database  $q_\gamma$ , and we are agnostic about the numerical probabilities to assign to the belief-network structure.  $B_S$ .

The pair  $B = (G, \Theta)$  represents the graph  $G$  and the set of conditional independence  $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ . We wish to determine the pair  $B$ 's maximal a posteriori (MAP) probability, given the data  $s$  (HANKIN *et al.*, 2010):

$$\begin{aligned} B &= \arg \max_B P(B|q_\gamma) \\ &= \arg \max_B P(q_\gamma|B)P(B) \\ &= \arg \max_{G,\Theta} P(q_\gamma|G, \Theta)P(\Theta|G)P(G) \end{aligned} \quad (2.13)$$

and we call  $P(q_\gamma|G, \Theta)$  the likelihood of data  $q_\gamma$  given the pair  $B = (G, \Theta)$ . From the Bayes

Theorem

$$P(q_\gamma|G) = \int_{\Theta} P(q_\gamma|G, \Theta)P(\Theta|q_\gamma)d\Theta \quad (2.14)$$

is the marginal likelihood. Assuming  $P(\Theta|G)$  is a Dirichlet distribution (HANKIN *et al.*, 2010), as it is the prior conjugate distribution of a categorical distribution and multinomial distribution (the distribution over observed counts of each possible category in a set of categorically distributed observations), (HECKERMAN, 1994) defined the *Bayesian-Dirichlet* score, BD:

$$p(q_\gamma|G) = \prod_{i=1}^n \prod_{j=1}^{q_i} \frac{\Gamma(N'_{ij})}{\Gamma(N'_{ij} + N_{ij})} \prod_{k=1}^{r_i} \frac{\Gamma(N'_{ijk} + N_{ijk})}{\Gamma(N'_{ijk})} \quad (2.15)$$

where  $n$  denotes the number of nodes,  $r_i$  denotes the number of states associated with node  $i$ , and  $q_i$  is the number of potential instantiations of the node's parents. The Gamma function is  $\Gamma(\cdot)$ ,  $N_{ijk}$  is the number of times  $x_i$  took the value of  $k$  given the parent configuration  $j$ ,  $N_{ij} = \sum_{k=1}^{r_i} N_{ijk}$ , and  $N'_{ij} = \sum_{k=1}^{r_i} N'_{ijk}$ . Given that the quantisation level of variables,  $r$ , is contained within a product, it is evident that the learning method can become prohibitively expensive if quantisation is set to a high value. Nonetheless, if quantisation is insufficient, the discrete dataset is going to lose its attributes and behave differently, resulting in an inaccurate DAG presumed as being the best BN for fitting the original data.

A BD function in the position of a scoring function to find the best BN is called *Bayesian Dirichlet equivalence with uniform prior metric*, *BDeu*. BDeu scores DAGs equally, which entails the same conditional independencies, i.e. are Markov-equivalent structures (BIELZA; LARRANAGA, 2014). There are other scoring methods, such as K2 (BEHJATI; BEIGY, 2020), the Bayesian Information Criteria (BIC) (BHAT; KUMAR, 2010) or a simple maximum likelihood estimation (BEN-GAL, 2008).

### 2.5.5 Conditional Probabilities

A BN inference is the process of obtaining the joint Conditional Probability Distribution (NEAPOLITAN *et al.*, 2004; HECKERMAN; WELLMAN, 1995). This procedure is carried out after restructuring the BN structure in order to establish all dependencies between variables. The joint distribution,  $P$ , is given by:

$$P(x_1, x_2, \dots, x_n) = \prod_{x \in \{X\}} p(x|\text{pa}_x), \quad (2.16)$$

It is evident that computational complexity and memory consumption of the exact inference algorithm are proportional to the product of the number of parents and quantisation level of each variable. Due to the fact that it can quickly become unfeasible, some approximate algorithms have been developed, for example, Pearl's message-passing algorithm (MURPHY,



1999; PEARL, 1994), and some special cases of DAGs, such as trees, having a simple inference method (NEAPOLITAN *et al.*, 2004; HUANG; DARWICHE, 1996). In this work, the selected inference technique is referred to as *Variable Elimination* (COZMAN *et al.*, 2000), which makes use of structure dependencies to marginalize particular variables, thereby removing them from computation. COOPER (1990), on the other hand, has proved that all known inference techniques in BN are NP-hard computationally. The discretisation technique employed in this thesis was the adaptive discretisation proposed by (DARBELLAY; VAJDA, 1999).

### 2.5.6 Causality and Bayesian Networks

The arrow orientations in a BN do not always correspond to the direction of causation. (NEAPOLITAN *et al.*, 2004). However, if we define a DAG  $G$  with vertices  $V$  and adopt the following rule: let  $X, Y \in V$ ;  $X$  has an edge directed to  $Y$  if and only if  $X$  causes  $Y$ , then  $G$  is said to be a *causal DAG*. Establishing causality between two variables is extremely difficult, as it requires a thorough understanding of the process depicted by variables. (PEARL, 2012). However, for actual analyses, it is critical to ensure two properties of causal networks: fidelity and causal sufficiency. (PEARL, 1994).

Supposing we have a joint probability distribution  $P$  of random variables in a given set  $V$  and DAG  $G = (V, E)$ , it is said that  $(G, P)$  satisfies a faithfulness condition if, based on the Markov condition,  $G$  entails all and only conditional independencies in  $P$ . I.e., the following two conditions hold:

- $G = (V, E)$  satisfies the Markov condition;
- All conditional independencies in  $P$  are described by  $G$ .

According to the principle of causal sufficiency, the set of measured variables  $V$  must include all common causes of pairs in  $V$ . To ensure the model's fitness, it is critical to address both of these constraints in order to prevent the difficulty mentioned in Figure 5. Hidden causes are problematic as they fundamentally alter the structures of dependencies inside a graph (as in the case of Figure 5, variables  $X$  and  $Y$  are originally independent, thus being conditionally dependent on  $Z$ ). The entire structure is destroyed by the exposure of  $H$ .

### 2.5.7 Dynamic Bayesian Networks Definition

Dynamic Bayesian Networks (DBNs) are Bayesian Networks in which distributions exhibit time-dependence (GHAHRAMANI, 1997). Simple DBNs include Kalman Filters (CHEN *et al.*, 2003) and Hidden Markov Models (RUSSELL; NORVIG, 2002).

DBNs were introduced by (DAGUM; GALPER; HORVITZ, 1992) as a technique for dealing with multivariate time series. The objective was to apply the BN in several categorization models in order to create a forecast model for those time series. For such a purpose, Dagum

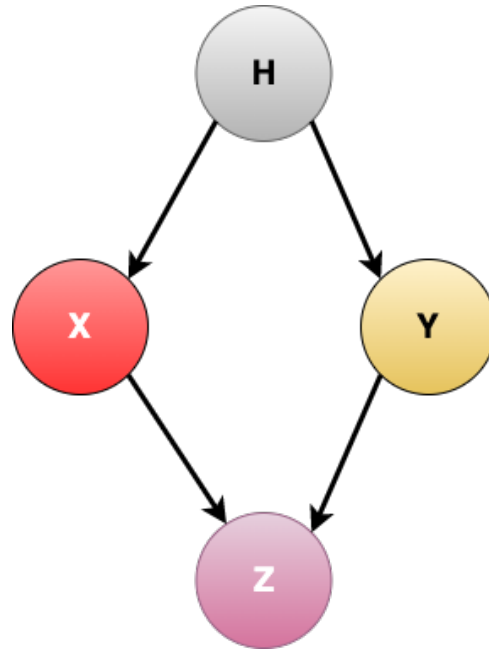


Figure 5 – A hidden cause, not observed, could change the dependencies between variables in an unpredictable way (NEAPOLITAN *et al.*, 2004).

defined *basal network* as the interaction between variables contained inside a time slice (Figure 7). Once the belief network is not updated, the basal model can be extended over several continuous time slices (until it becomes obsolete and loses adherence).

However, when dealing with more complex real-world multivariate time series data, the basal network frequently does not exist, since dependencies between variables alter between time slices. In such a circumstance, it is vital to estimate the number of time slices required for forecasting; also, some variables may become completely detached from the network.

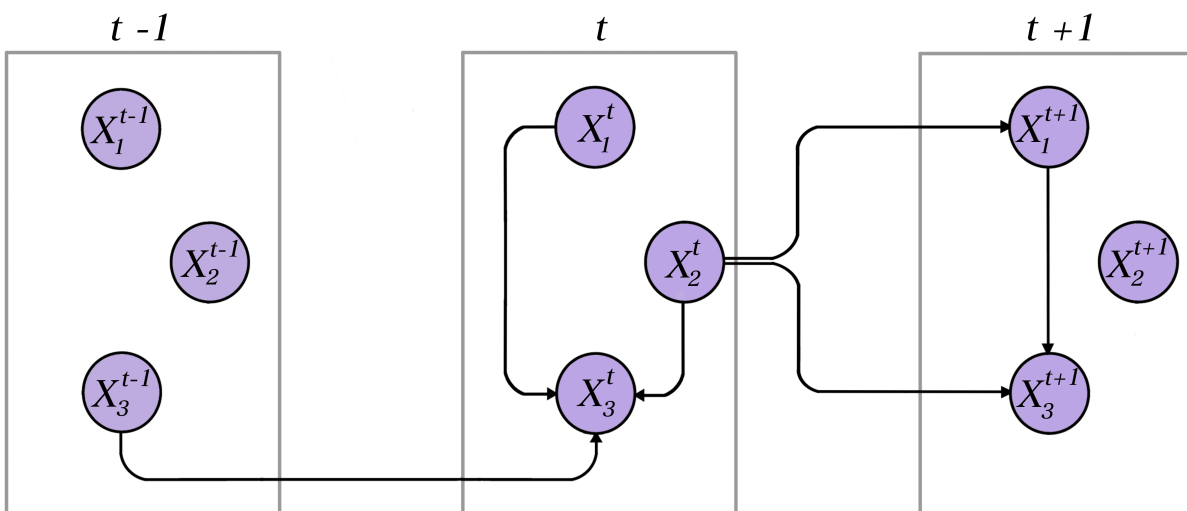


Figure 6 – A more complex DBN in which the basal structure and time-slices connections become obsolete over time due to non-stationarity and non-linear dynamics of variables.

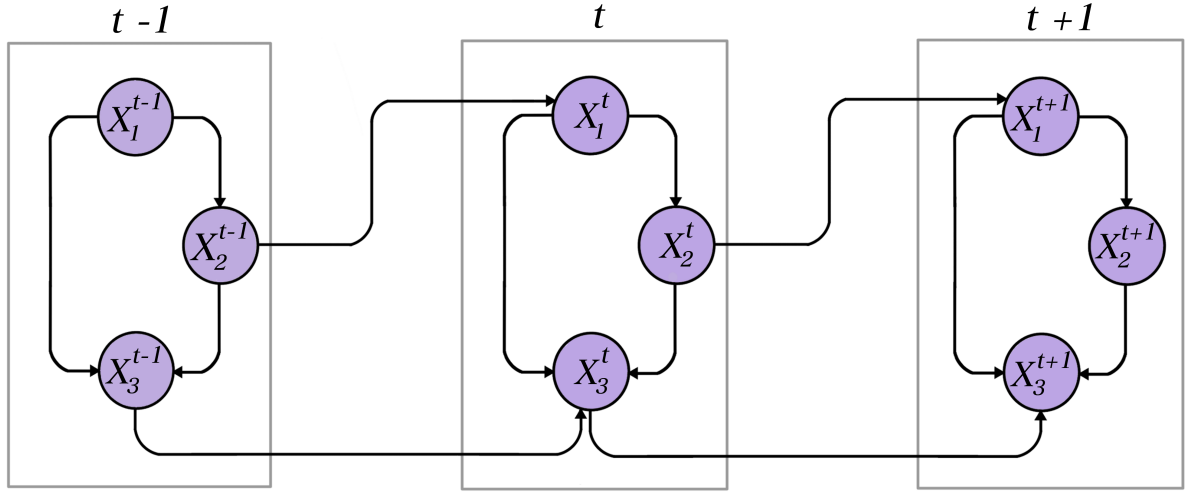


Figure 7 – A well-behaved DBN with 3 time-slices in which the basal structure and connections between time-slices are preserved over time.

A DBN mimics the effect of changes in a BN over time (LI *et al.*, 2021), which requires the addition of temporal data to the analysis (RAMOS *et al.*, 2021). Figures 7 and 6 illustrate DBNs. By definition,  $B_0$  is a prior BN defining the joint distribution of all variables in timestamp  $t = 0$ , i.e.  $B_0 = P(\mathbf{X}[0])$ , and  $B_{\rightarrow}[0 : t]$  is the set of all transition networks such as (LEãO *et al.*, 2021):

$$B_{\rightarrow}[0 : t] = P(\mathbf{X}[t]|\mathbf{X}[0 : t - 1]) \quad (2.17)$$

Applying the chain rule to both  $B_0$  and  $B_{\rightarrow}$ , a DBN defines the joint distribution of all possible trajectories given the variables (LEãO *et al.*, 2021) in the following equation:

$$P(\mathbf{X}[0 : T]) = P(\mathbf{X}[0]) \prod_{t=1}^T P(\mathbf{X}[t]|\mathbf{X}[0 : t - 1]) \quad (2.18)$$



### 3 EXPLORATORY ANALYSIS METHODOLOGY

The following section is going to describe the presently applied methodology. It was performed a survival analysis. Appendices C and D contain the methodology, results and discussions from complementary analyses performed during the doctorate but not related to the main goal of the thesis.

#### 3.1 Survival Analysis Methodology

##### 3.1.1 Kaplan-Meier estimation and Logrank test

The Kaplan-Meier estimation is a suitable alternative to measure the number of subjects that have lived for a specific time after treatment (GOEL; KHANNA; KISHORE, 2010). A statistical representation for the aforementioned definition is:

$$S(t) = P(T > t) = 1 - F(t) \quad (3.1)$$

where  $S(t)$  is the survival function, a complement of the cumulative density function (CDF)  $F(t)$ .  $T$  represents the *time of death* or *failure*. Survival in this thesis means not having developed myoclonic or generalized tonic-clonic seizures (events of interest). Therefore, according to equation 3.1,  $S(t)$  means the probability of being seizure-free (behaviorwise) after time  $t$ . The Kaplan-Meier method is an empirical non-parametric approach to estimate the survival function, mathematically defined as (KARTSONAKI, 2016):

$$\hat{S}(t) = \prod_{j:t_j \leq t} \frac{n_j - d_j}{n_j} \quad (3.2)$$

where  $\hat{S}(t)$  is the Kaplan-Meier estimation function,  $t$  ( $t_1 < t_2 < t_3 < t_k$ ) represents the times at which seizures were observed,  $d_j$  are the subjects affected by the event at time  $t_j$  ( $j=1,2,3,\dots,k$ ), and  $n_j$  is the number of subjects under risk prior to  $t_j$ . The Logrank test is a statistical tool aimed to evaluate Kaplan-Meier estimations by comparing them under a null hypothesis in which there is no difference in population survival probabilities at any analysis times (ALTMAN; BLAND, 2004).

##### 3.1.2 Animal Experiments and Database

The present experimental procedures involving rats are described in full detail in De Oliveira et al. (OLIVEIRA *et al.*, 2014) and their theoretical framework has been reviewed by Cota et al. (COTA *et al.*, 2016). The whole database was acquired at the Laboratory of Neuroengineering and Neuroscience from the Federal University of São João Del Rei (ethics committee protocol

31/2014). A total of 49 Wistar male rats weighing 250-350 g were collected from the main UFSJ vivarium, kept under a light-dark cycle of 12 h (lights on at 7 a.m.), with food and water ad-libitum, and then divided into four groups, i.e. two control groups and two others that were submitted to the NPS treatment. Briefly, all animals underwent intravenous controlled infusion of convulsant drug pentylenetetrazole (PTZ), an unspecific GABAergic antagonist, at a rate of 1 ml/min and dilution of 10 mg/ml (thus 10 mg/min) as a model of acute ictogenesis and seizure induction. Infusion was interrupted at the onset of forelimb clonus (myoclonic seizures) in both NPS-MYO (n = 12) and MYO-CTRL (n = 14) groups. Similarly, PTZ infusion was interrupted only after the occurrence of generalized tonic-clonic seizures in animals belonging to NPS-GTC (n = 14) and GTC-CTRL (n = 9) groups. Although animals of the latter two groups also displayed forelimb clonus (as well as several other convulsive behaviors) prior to generalized seizures, such time data points were not included in the survival analysis of myoclonic seizures (Table 1 and figure 2A) in order to maintain a separation between groups and allow better conformity to results of future analyses. NPS-MYO and NPS-GTC animals were submitted to a concomitant NPS treatment, but those of CTRL-MYO and CTRL-GTC were used as controls for comparison purposes. NPS was delivered as biphasic square pulses of constant current ranging from 150 to 400  $\mu$ A (according to each animal's susceptibility) lasting 100  $\mu$ s of duration in each phase at an average of four pulses per second using an off-the-shelf equipment (8-channel stimulator with isolation units, models 3500 and 3800 from A-M Systems, Sequim, WA, USA). Bipolar electrodes made of twisted pairs of Teflon-coated stainless-steel microwires (num. 7914, A-M Systems, Sequim, WA, USA) were surgically implanted in the amygdalae of both brain hemispheres with the aid of a stereotaxic apparatus (coordinates AP = 2.8 mm, ML =  $\pm$ 5.0 mm from Bregma and 7.2 mm from Dura mater). Amygdala was selected as target for the NPS given its central role in epileptic phenomena and its widespread mono- or polysynaptic connectivity with structures in the forebrain, midbrain, and hindbrain. Considering a fixed rate of PTZ infusion and very small variability with no significant difference in animals' weights across groups (CTRL-MYO:  $306 \pm 10$  g ; CTRL-GTC:  $316 \pm 10$  g; NPS-MYO:  $311 \pm 6$  g; NPS-GTC:  $314 \pm 7$  g), we measured the latency to trigger a convulsive behavior as input parameter for the present survival analysis and as a means to assess the NPS therapeutic effect. A summary of the experimental methodology is presented in the leftmost panel of Figure 8.

### 3.1.3 Data Analysis

A summary of the methodology applied herein is presented in Figure 8. All analyses were developed using the whole data collected from the following pre-clinical trial groups: MYO-CTRL (n=12), NPS-MYO (n=14), GTC-CTRL (n=9) and NPS-GTC (n=14), comprising a total of 49 rats. The first procedure was the tabulation of survival times observed for each group, namely, the time taken for rats to develop myoclonic and generalized tonic-clonic seizures (events of interest).

The second procedure was the Kaplan-Meier estimation so as to draw survival graphs

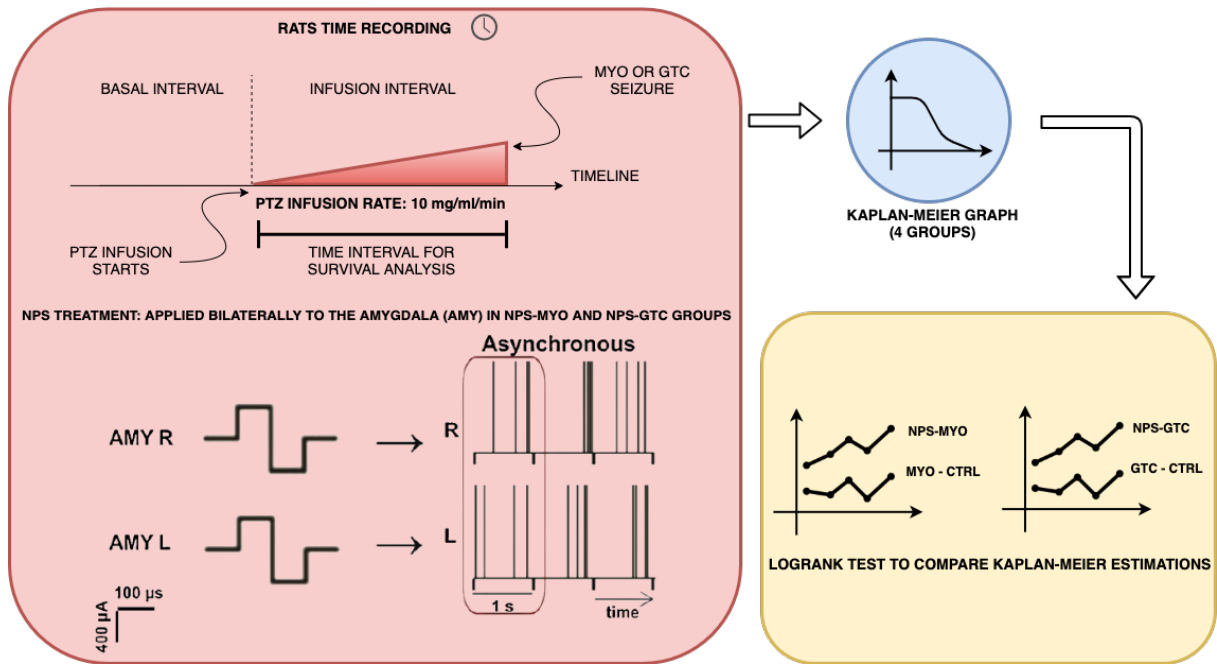


Figure 8 – Applied methodology. At first, the times at which all rats from the four study groups (MYO-CTRL, NPS-MYO, GTC-CTRL and NPS-GTC) developed seizure were recorded. Control groups underwent PTZ infusion to induce seizures and the times until observing myoclonic and generalized tonic-clonic behaviour were recorded. NPS-MYO and NPS-GTC groups followed the same protocol as that observed for control groups, but the difference is that they were submitted to the NPS therapy until developing seizures. NPS treatment was applied bilaterally to the amygdala of animals subjected to the pentylenetetrazole continuous infusion (10 mg/ml/min) model. Afterwards, the Kaplan-Meier estimator was performed for each group. Finally, the respective groups were compared through the Logrank test.

for each group. After plotting graphs for each group, they were compared according to their respective groups - MYO-CTRL with NPS-MYO and GTC-CTRL with NPS-GTC - through the Logrank statistical test.

All survival and statistical analyses were developed using Python programming through the Lifelines package and a computer environment with a 5th generation Intel i7 processor, 8GB RAM and MacOS 10.14 operational system.





## 4 DYNAMIC BAYESIAN NETWORK METHODOLOGY

This chapter is going to describe the applied methodology used to perform DBN, MSV and their combination. It is going to be depicted how DBN and MSV are used to develop the statistical assessment and perform a functional connectivity analysis. Also, the computer environment and database used to carry out the case studies are going to be described .

### 4.1 Experimental protocol

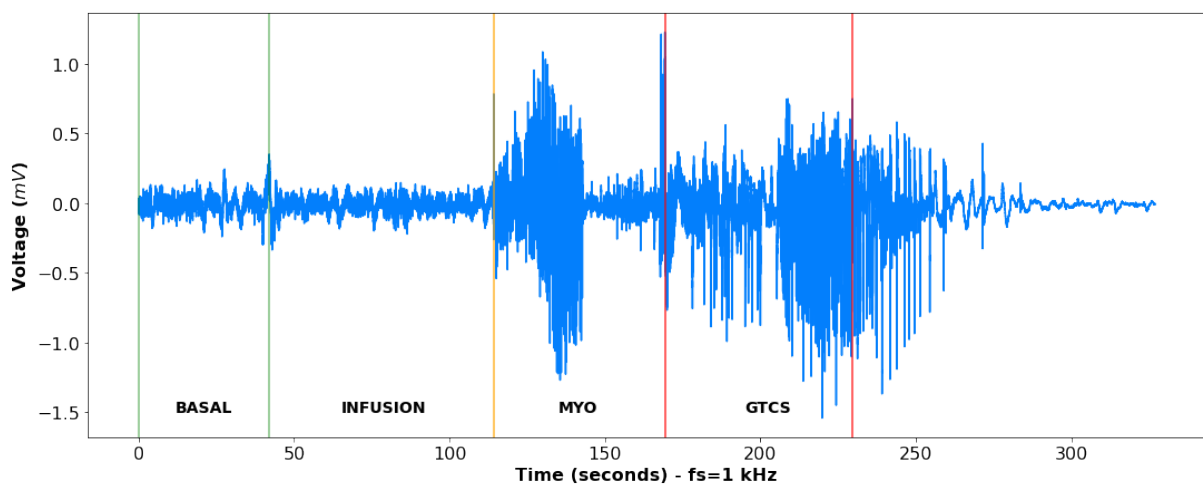


Figure 9 – An example of the signal from one of the rats involved in the study. There are four different time slices for that signal. The first one is the time between the green lines, which indicates resting state, i.e. the basal state of the animal. From the first green line until the orange line, there is an interval of PTZ infusion for epileptic seizure induction. Infusion only ceases when the animal develops generalized tonic clonic seizure. From the orange line until the first red line, there is an interval of myoclonic seizure, and the interval of generalized tonic clonic seizure is represented from the first red line until the second one . After the second red line, it is the interval after GTCS, and it is visually clear that this period does not represent a basal state but a refractory period.

The present experimental procedures with rats are described in better detail in (OLIVEIRA *et al.*, 2018) and its theoretical framework has been reviewed in (COTA *et al.*, 2016) and (COTA *et al.*, 2019). Local Field Potential (LFP) is derived from the database of Laboratory of Neuro-engineering and Neuroscience (*LINNce*) at the Federal University of São João Del Rei. Male Wistar rats weighing between 250 and 350 grams and kept under a light-dark cycle of 12 h (lights on at 7 a.m.), with food and water ad-libitum, were selected from the University's Central Vivarium. All described procedures follow the highest ethical standards for the usage of animals in research and have the previous approbation by the institutional committee (protocol 31/2014, CEUA/UFSJ). The signal recording used monopolar electrodes consisting of Teflon-coated

stainless-steel wires (#791600, A-M Systems, Sequim, WA, USA). They were placed directly into the right thalamus (TH) and right dorsal hippocampus (HP) of the animals' brains through a stereotactic surgery (COTA *et al.*, 2016). The assistance of positioning the electrodes and screws followed Paxino's neuroanatomic atlas, which were AP: 2.8 mm, ML: +1.5 mm, DV: 3.3 mm for HP, and AP: 3.0 mm, ML: +2.6 mm and DV: 6,0 mm for TH (PAXINOS; WATSON, 2013). Additionally, two microsurgical screws (length 4.7 mm, diameter 1.17 mm, Fine Science Tools, Inc., North Vancouver, Canada) were visually implanted into their right hemisphere parietal region for cortical (CX) recording and the frontal bone was used as reference. Leads soldered to copper wires were crimped in an RJ-45 jack fixed onto the skull using polymerizing dental acrylic.

Animals were filmed simultaneously to collect LFP recordings in order to perform a behavioral analysis and assess occurrence and latency to stereotypical behaviors of the selected models, such as facial automatisms, myoclonic jerk, head clonus, hind and forelimb clonus, generalized tonic-clonic seizure, and others such as rearing and falling, Straub tail. It allowed a correlation with LFP data and detection of electrophysiological events, i.e. periods of interest in this study.

Amplification of signals used a 2000 V/V gain, filtered from 0.3 to 300 Hz using an A-M Systems (model 3500) pre-amplifier, and then digitized at 1 KSample/sec using an A/D converter board (model PCI 6023E, National Instruments) controlled by a built-in LabView virtual instrument developed at *LINNce*. Shielded twisted cables and a Faraday cage were required to eliminate the power grid noise at 60 Hz.

All animals underwent intravenous controlled infusion of convulsant drug pentylenetetrazole (PTZ, Sigma Aldrich, São Paulo, SP - Brazil), an unspecific GABAergic antagonist, at a rate of 1 ml/min and dilution of 10 mg/ml (thus 10 mg/min) as a model of acute ictogenesis and seizure induction. This approach results in a gradual increase of neural excitability and, consequently, in a gradual recruitment of neural circuitry (VELISEK *et al.*, 1992), which are expressed both behaviorally and electrographically in a correlated manner. Initially, the animals displayed minor seizures, including facial automatisms, strong mastication, myoclonic jerks, forelimb, and head clonus. These are all behaviors directly related to aberrant recruitment of limbic circuitry, including areas such as the amygdala, hippocampus, and thalamus (EELLS *et al.*, 2004). It is followed by significant seizures, with or without a tonic phase, such as generalized myoclonus and generalized tonic-clonic seizures. It results from an involvement of large territories in the forebrain or also of structures in the midbrain and hindbrain, respectively (EELLS *et al.*, 2004). This gradual recruitment of areas and circuits makes a controlled infusion of PTZ an exciting model for screening new drugs or other non-pharmacological treatments. It also investigates neurodynamical processes underlying ictogenesis, as it is the case for the present study.

The time slices used in this Bayesian Networks analysis were established based on periods of interest of the previously described experimental protocol. The time slices defined

to apply the algorithm are basal state, infusion, myoclonic seizure (MYO), and generalized tonic-clonic seizure (GTCS) - Figure 9.

## 4.2 Algorithms for Dynamic Bayesian Network Methodology

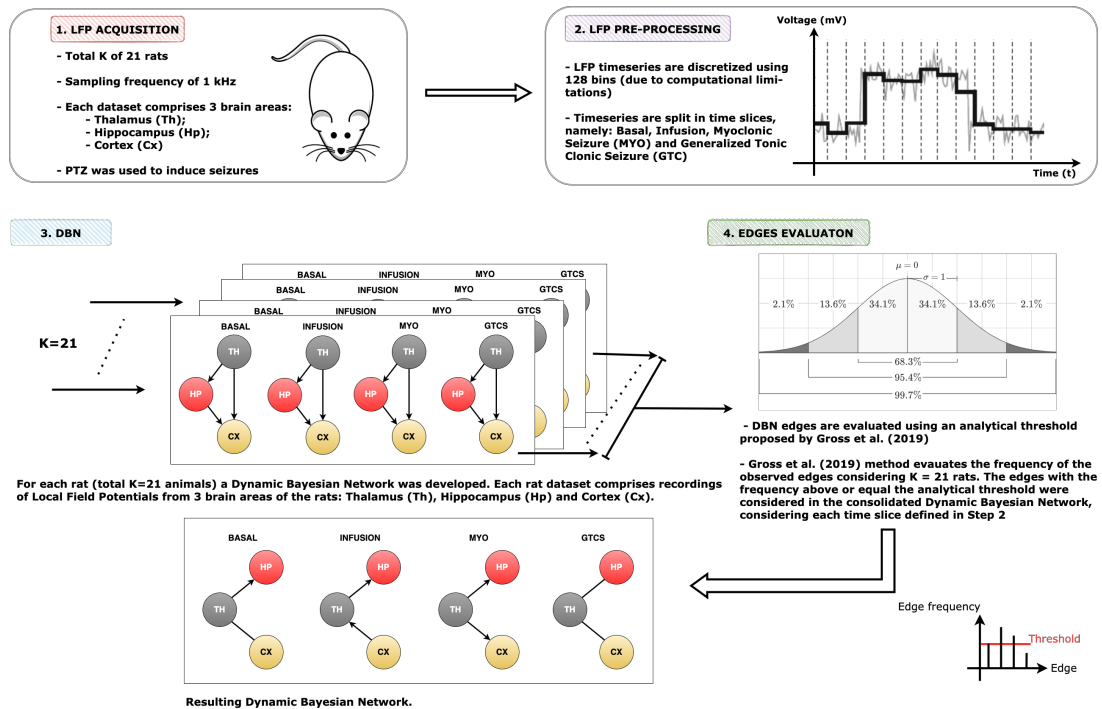


Figure 10 – Applied methodology. The first step is the LFP acquisition from rats involved in the pre-clinical trial. After discretizing time series and splitting them into time slices, the following point-in-time is depicted: basal state, PTZ infusion, myoclonic seizure (MYO) and generalized tonic clonic seizure (GTCS). The third step is to use a DBN algorithm to observe the functional connectivity among time slices during the temporal evolution of rats, i.e. from basal state until GTCS. The arcs from the developed Dynamic Bayesian Networks are evaluated using the analytical threshold proposed by (GROSS *et al.*, 2018) and described in full detail in (GROSS *et al.*, 2019).

Figure 10 presents the applied methodology. Initially, a data frame with three columns (thalamus, hippocampus, and cortex) represents each rat. Afterwards, each of these columns are divided by the samples and then regrouped, which resulted in a new data frame with 12 columns: thalamus, hippocampus, and cortex for each time-slice basal state, infusion, MYO, and GTCS (Figure 11). Since each time slice has a different duration, there was a pre-processing of all of them using numeric interpolation so that all of them would have the same size as the longest time-slice, resulting in a data frame with 12 columns and *size\_of\_longest\_time-slice* rows. The dependencies among the variables used a completely graphical and non-parametric strategy. The representation of functional connectivity networks among brain areas used a BN structure learned from the discretized dataset. The quantization followed the adaptive bins algorithm exhibited in (GENCAGA; KNUTH; ROSSOW, 2015) using a maximum of 128 bins (7 bits),

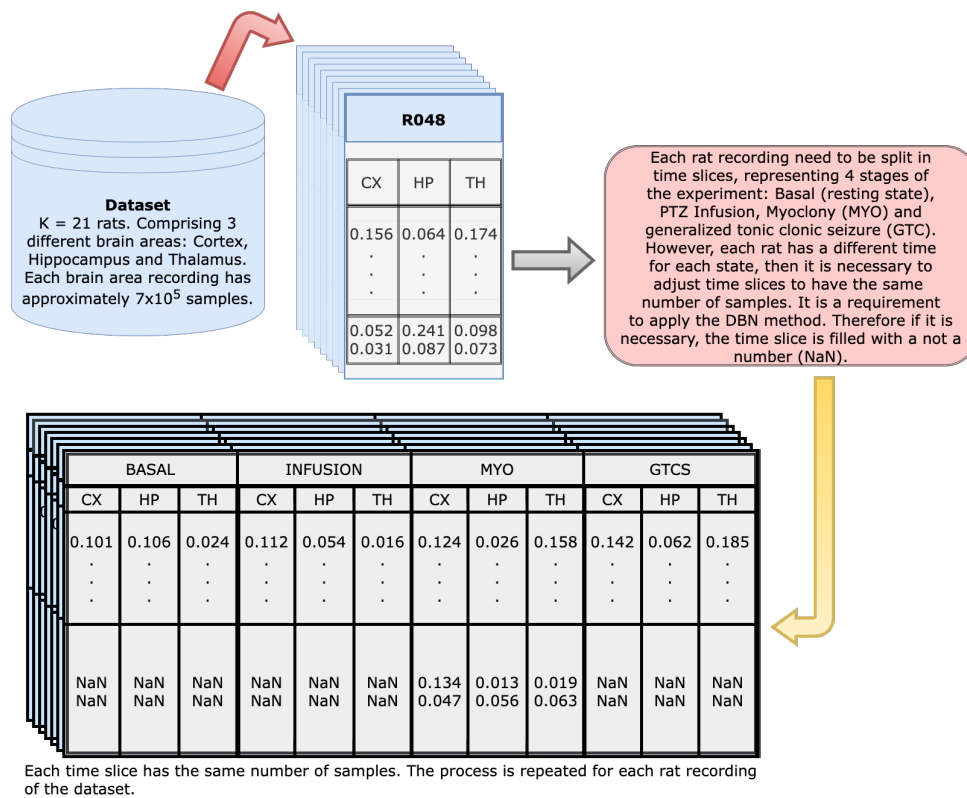


Figure 11 – Dynamical Bayesian Network input generation: for each rat, the initial dataset contains 700 thousand observations for each of the three brain regions; thus, a new dataset is generated with 12 columns representing each of the regions on all four timestamps (Basal, Infusion, MYO and GTCS). Since all of them have different duration, the resultant table has 12 columns and as many rows as the longest time-slice. The remaining columns are filled with NaN to keep the table structure intact.

since it was the maximum amount of bins supported by available computational resources. The Hill Climbing search algorithm from Python *pgmpy* package<sup>1</sup> was used to learn the DAG from the dataset, and the BDeu function is applied as a scoring method (see Appendix), once the task may be too complex or even impossible for a human being (VILLANUEVA; MACIEL, 2014).

During the experiment, there was a building of a set of  $K = 21$  DAGs by running the Hill Climbing search algorithm twenty-one times. The different data used in each of these runs represents a dataset acquired from one of the rats involved in this pre-clinical study. The underlying idea is that there is less uncertainty regarding the arcs that are still induced when collecting data from a different animal. The diversity of structures is due to the fact that data was acquired from different rats of the same species and approximately having the exact weights, and the Hill Climbing search process itself, once its initialization is always random and local optimizations performed during a run are also non-deterministic. As stop criterion, one million iterations are performed for each complete run of the Hill Climbing search algorithm. Afterwards, the set of DAGs was reduced to a single consensus DAG, as a process called the model-averaging

<sup>1</sup> <http://pgmpy.org/index.html>

approach. In this reduction, the number of times that each of the three possible connections (i.e., " $\leftarrow$ ", " $\rightarrow$ ", and "absent") occurred is counted considering every pair of nodes in the obtained 21 graphs. Only the directed arcs occurring at least at the percentage ( $f$ ) provided by equation  $f = (1/3) + \sqrt{2/K}$  were accepted. It is the analytical threshold method to evaluate the arcs proposed by (GROSS *et al.*, 2018) and described in full detail in (GROSS *et al.*, 2019); moreover, specialist analysis was taken into account in the final evaluation of the resulting network, by only the edges resulting from the analytical threshold evaluation were considered.

The entire algorithm, including the BN generation and pre-processing, takes about 2 to 3 minutes for each rat database, containing three local field potential time series with approximately 700 thousand samples in a personal computer equipped with 12 GB RAM and a 4-core/4-thread Intel(R) Core(TM) i7-4500U CPU @ 1.80GHz. Therefore, the total time spent on processing all databases, constituted of 21 rats ( $K=21$ ), was about 48 minutes. Among all databases, half of which comprises rats in the myoclonic group, through which their evolution from the basal state until the myoclonic seizure was recorded. The observation of other rats included a temporal evolution from the basal state until generalized tonic-clonic seizure, i.e. those belonging to the GTCS group.

### 4.3 Algorithms for Dynamic Bayesian Network combined with Multivariate Stochastic Volatility Methodology

Figure 12 presents the first three steps of the applied methodology. The first step is signal acquisition, depicted in detail in the subsection above. The second step is the discretization process of Local Field Potentials, following the adaptive bins algorithm described by Gencaga et al. (2015) (GENCAGA; KNUTH; ROSSOW, 2015) and using 128 bins of quantization due to a limited availability of computational resources. The aim was to keep the best possible resolution since the Local Field Potentials sampling frequency used to acquire the signals was 1 kHz.

A data frame with three columns (thalamus, hippocampus, and cortex) represents each rat. After Local Field Potentials discretization, each of these columns are divided by the samples and then regrouped, which resulted in a new data frame with 12 columns: thalamus, hippocampus, and cortex for each time-slice basal, infusion, MYO, and GTCS (Figure 12). Since each time slice has a different duration, there was a pre-processing of all using numeric interpolation so that all of them would have the same size as the longest time-slice, resulting in a data frame with 12 columns and *size\_of\_longest\_time-slice* rows.

The dependencies among variables used a graphical and non-parametric strategy. The representation of functional connectivity networks among brain areas used a BN structure learned from the discretized dataset. There is the use of Hill Climbing search algorithm from Python *pgmpy* package<sup>II</sup> to learn the DAG from the dataset, and K2 was used as scoring method.

<sup>II</sup> <http://pgmpy.org/index.html>

Two other commonly available scoring methods are the use of *Akaike Information Criteria (AIC)* or *Bayesian Information Criteria (BIC)* (AKÇA; YOZGATLIGIL, 2020; DZIAK *et al.*, 2020; HELD; BOVÉ, 2020).

However, there are concerns about the evaluation procedure, especially the asymptotic theory considered for both criteria calculation. This assumption means the use of central limit theory approximating data as a normal distribution and, usually, real data is not normally distributed (KOH; AHAD, 2020). Even in the case of normality, the asymptotic theory still requires a large sample size for a reliable analysis (MCEWAN, 2020), which is not possible at times due to the complexity to acquire them, e.g. in the case of Local Field Potential recordings (DAS; MENON, 2020; WANG *et al.*, 2020).

The second issue is regarding the likelihood function, especially when the model is not regular, and it usually results in intractable likelihood functions (ROSSI *et al.*, 2020). *AIC* and *BIC* are criteria based on likelihood penalization (OGASAWARA, 2020). Therefore, the issue may make their calculation difficult or impossible at times. During the experiment, a set of  $K = 9$  DAGs was built by running the Hill Climbing search algorithm nine times without any lag among brain areas (Figure 13). The different data used in each run represents a dataset acquired from one of the rats involved in this pre-clinical study. The underlying idea is that there is less uncertainty regarding the arcs that are still induced while collecting data from a different animal. The diversity of structures is due to the fact that data was acquired from different rats of the same species having approximately the exact same weights, and to the Hill Climbing search process itself, once its initialization is always random and the local optimizations performed during a run are also non-deterministic. As the stop criterion, one million iterations is the limit for each complete Hill Climbing search algorithm execution. The K2 score is calculated for each DBN developed (Figure 13).

Afterwards, the Multivariate Stochastic Volatility was applied to investigate the lag among brain areas. The Thalamus is used as a reference to identify the samples delay among LFP signals - Figure 14. The model described by Phan *et al.* (2019) (PHAN *et al.*, 2019) is used for such a purpose, considering flat prior information as follows:

- $\mu \sim N(0, 1000I_n)$ , the unconditional average volatility modeled as a multivariate normal distribution;
- Elements of the persistence matrix  $\beta$  follow a beta distribution function,  $\beta_{n,k} \sim B(20, 1.5)$ , to establish its values between the limits of  $-1$  and  $1$ . It is important to certify the volatility process stationarity;
- Elements of the  $\Gamma$  matrix related with the error associated with volatility model (Equation 2.2) are parameterized as  $N(0, 10)$ .

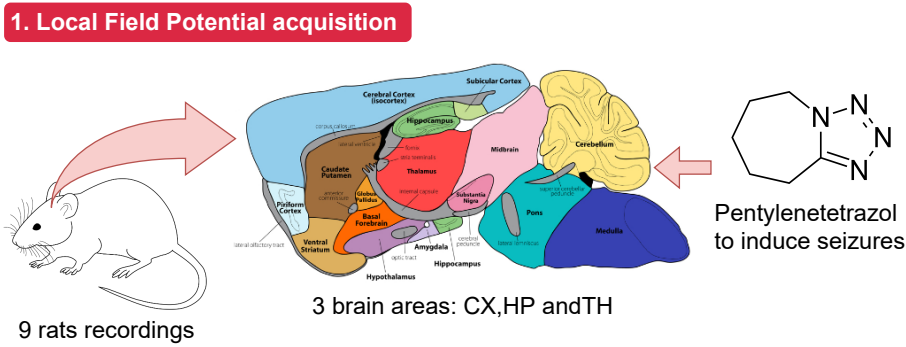
The Markov Chain Monte Carlo (MCMC) simulation was necessary due to the intractable analytical calculation of the model. Appendix B presents a more detailed theory of this methodology. The adopted strategy uses two simulation steps: the first one makes use of the Gibbs sampling method to facilitate model convergence, followed by the second step, Metropolis-Hastings sampling. The aim was to estimate the MSV model parameters through MCMC simulation due to the difficulty of using a frequentist approach, especially to calculate the likelihood function, which is almost intractable considering nonregular models such as MSV (ROSSI *et al.*, 2020).

A common approach used to capture the lag among brain areas is the mutual information, an Information Theoretical method (ENDO *et al.*, 2015). However, it is a type of univariate method, which consequently means that the spatiotemporal feature of Local Field Potentials is overlooked to perform the statistical inference among brain areas (PHAN *et al.*, 2019).

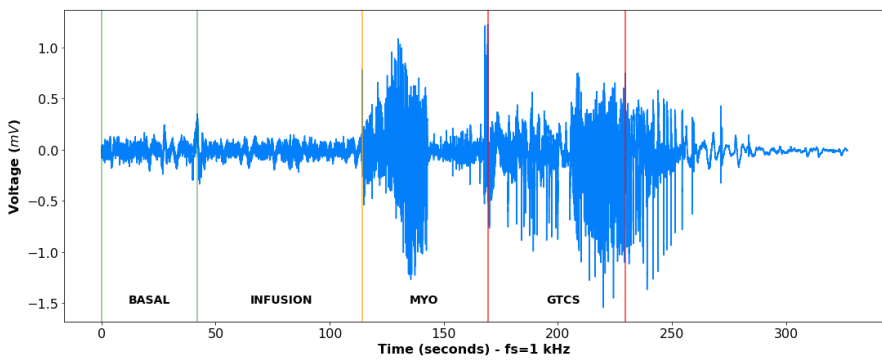
The use of MSV allows the simultaneous modeling of the spatial and temporal structures using a parametric and multivariate method (PHAN *et al.*, 2019). Once the brain is a complex system, a spatiotemporal approach can assist the analysis of epileptogenesis (hyper synchronism), once the phenomenon results from interactions among brain areas, in which communications may happen at different times. Epilepsy is an emergent property of the brain (complex system), not an isolated process.

The lag results investigated from the Multivariate Stochastic Volatility method are consolidated and used to apply the DBN method once again. Each network has its K2 score calculated to make a comparison with previous results.

Finally, it is used the network with the best fit with all rats data frames. There is the calculation of the K2 score to compare with obtained results of the applied methodology. Each dataset contains three Local Field Potential time series with approximately 700 thousand samples. The entire algorithm, including the lag investigation through Multivariate Stochastic Volatility, the generation of the BN, and the pre-processing, takes about 333 minutes for each rat database. Therefore total time spent processing all databases, constituted of 9 rats ( $K=9$ ), was about 50 hours, considering a personal computer equipped with 12 GB RAM and a 4-core/4-thread Intel(R) Core(TM) i7-4500U CPU @ 1.80GHz.



**2. Pre-Processing**



**3. Time slice adjustment**

R048	BASAL			INFUSION			MYO			GTCS		
	CX	HP	TH	CX	HP	TH	CX	HP	TH	CX	HP	TH
1	0.101	0.106	0.024	0.112	0.054	0.016	0.124	0.026	0.158	0.142	0.062	0.185
2	0.156	0.064	0.174	.	.	.	.	.	.	.	.	.
3	0.052	0.241	0.098	NaN	NaN	NaN	0.134	0.013	0.019	NaN	NaN	NaN
4	0.031	0.087	0.073	NaN	NaN	NaN	0.047	0.056	0.063	NaN	NaN	NaN

Each rat has a different time for each state, so it is necessary to adjust time slices to have the same number of samples. It is a requirement to apply the DBN method. Therefore, if necessary, the time slice is filled with a not a number (NaN).

Figure 12 – The first three steps of the applied methodology: initially, Local Field Potentials are acquired from nine rats, comprising three brain areas - Cortex (CX), Hippocampus (HP), and Thalamus (TH), at a sampling frequency of 1kHz. Induction of seizures was made through PTZ administration. After time series was discretized using 128 bins and split into four time-slices: Basal, Infusion, Myoclonic Seizure (MYO) and Generalized Tonic-Clonic Seizure (GTCS). Finally, due to length differences among time-slices, they are adjusted.



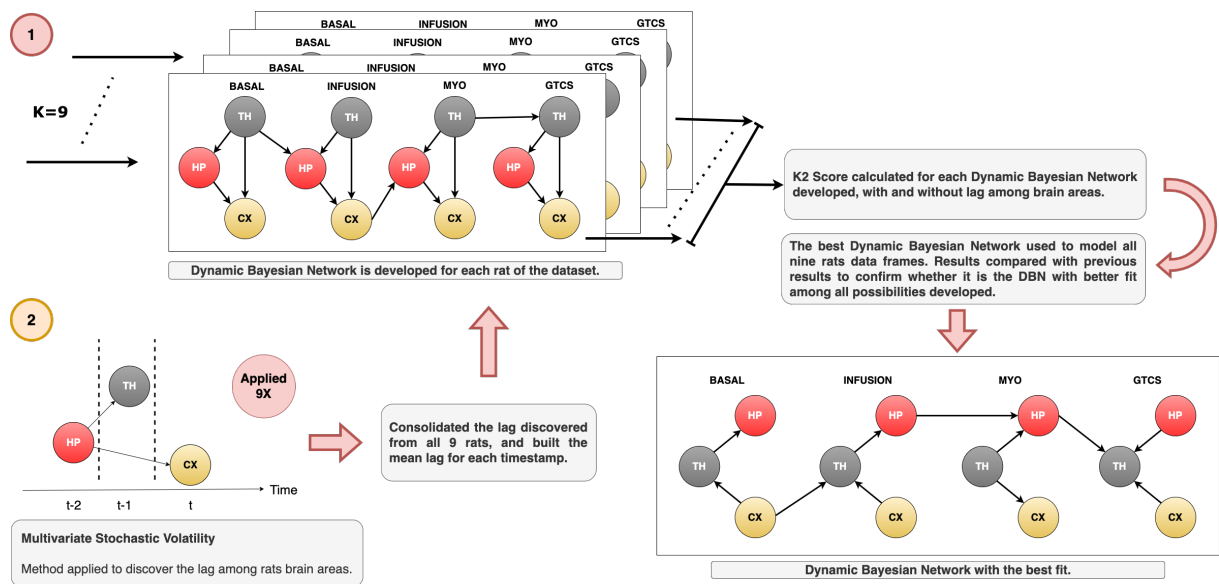
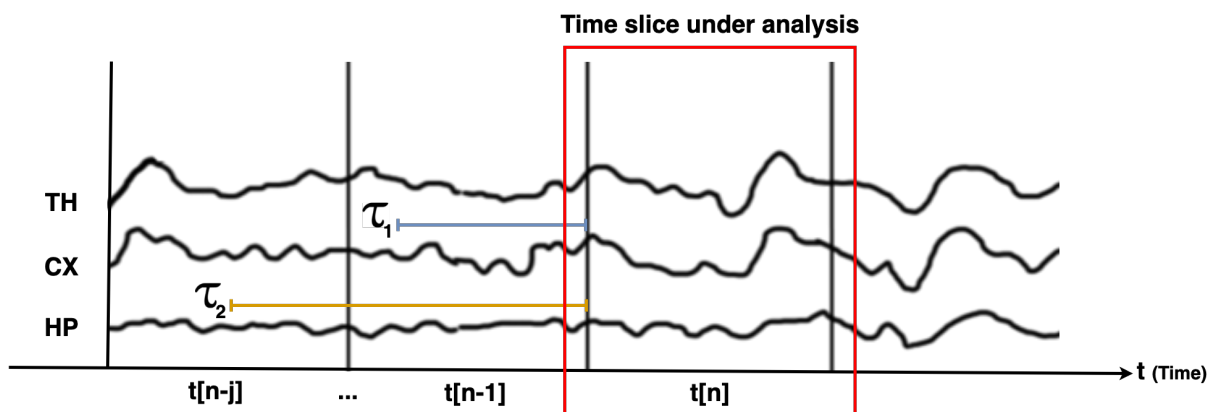


Figure 13 – (1) The method initially uses all data frames to develop the DBN of each of the nine rats without considering the lag among brain areas. K2 Score was calculated for each Bayesian Network. (2) Multivariate Stochastic Volatility was used to study the lag among rats brain areas. Consolidated lags obtained for each timestamp were used to develop the nine DBN for each rat. K2 Score was once again calculated for each network. Therefore, K2 Scores were compared to investigate which is the best DBN. The graph with better fit was applied to all data frames, the K2 Scores were calculated to make a comparison with previous results. The aim is to investigate the DBN with the best fit.



The LFP signal is the reference for the delay among brain areas. Based on MSV analysis, there is a displacing (in samples) of the CX and HP LFP signals before applying the DBN method again.

Figure 14 – The lag analysis. The Thalamus (TH) is used as a reference to identify the sample delay among LFP signals. The reasoning is the same to observe the lag during each time slice: Basal, Infusion, MYO, and GTCS. Based on MSV analysis, there is the displacing of the CX and HP LFP signals in samples of  $\tau_1$  and  $\tau_2$ , respectively. There is the repetition of the process for each time slice and rat data used in this study. It is performed as data frame preparation before applying the DBN method once more.



## 5 EXPLORATORY ANALYSIS RESULTS AND DISCUSSION

### 5.1 Survival Analysis

The time at which each rat has developed seizures can be observed in Tables 1 and 2 - which has been recorded in seconds. By comparing the results in Table 1 (MYO-CTRL and NPS-MYO groups), it is possible to verify that there is a time difference between groups, which is greater for the NPS-MYO group on average, thus suggesting treatment efficacy against seizures. In Table 2 (GTC-CTRL and NPS-GTC groups), it can also be checked that the NPS treatment achieved the desired effect by delaying the time length of rats' seizures.

Table 1 – Time registered for each rat from MYO-CTRL and NPS-MYO groups. The MYO column indicates whether the rat had myoclonic behaviour (1) or not (0). CTRL and NPS represents the columns indicating to which group each rat belongs (control or treatment).

Rat	Time(s)	MYO	CTRL	NPS
1	82.10	1	Yes	-
2	91.20	1	Yes	-
3	41.50	1	Yes	-
4	94.10	1	Yes	-
5	79.50	1	Yes	-
6	68.00	1	Yes	-
7	86.40	1	Yes	-
8	61.50	1	Yes	-
9	92.00	1	Yes	-
10	85.30	1	Yes	-
11	92.30	1	Yes	-
12	53.10	1	Yes	-
13	86.00	1	Yes	-
14	88.00	1	Yes	-
15	82.60	1	-	Yes
16	88.90	1	-	Yes
17	82.40	1	-	Yes
18	42.40	1	-	Yes
19	102.40	1	-	Yes
20	128.80	1	-	Yes
21	106.90	1	-	Yes
22	102.70	1	-	Yes
23	118.80	1	-	Yes
24	94.10	1	-	Yes
25	100.90	1	-	Yes
26	90.70	1	-	Yes

The Kaplan-Meier estimation for MYO-CTRL and NPS-MYO groups can be observed in Figure 15a. There is a difference between control and treatment groups, i.e. the survival time for the NPS-MYO group is longer and survival estimation curves started becoming different from

around 40 seconds time onward. The Kaplan-Meier estimation for GTC-CTRL and NPS-GTC groups shown in Figure 15b once again reveals a noticeable difference between treatment and control groups. The survival graph is almost the same until reaching approximately 125 seconds time, and then it becomes different, as it is higher for the NPS-GTC group. Graphically, it is possible to observe that the plot starts becoming different at an earlier time for the NPS-MYO group.

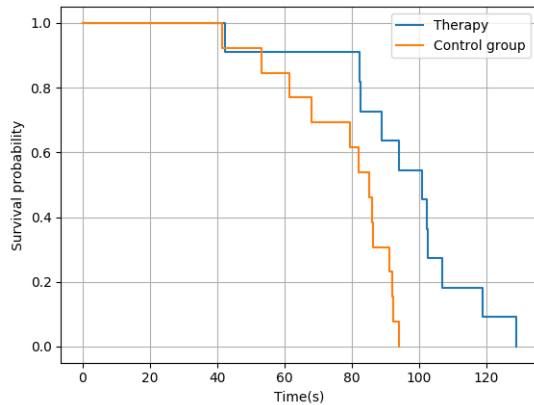
Table 2 – Time recorded for each rat from GTC-CTRL and NPS-GTC groups. GTC column indicates whether the rat presented generalized tonic-clonic seizure behaviour (1) or not (0). CTRL and NPS represents the columns indicating to which group each rat belongs (control or treatment).

Rat	Time(s)	GTC	CTRL	NPS
1	127.50	1	Yes	-
2	133.60	1	Yes	-
3	109.00	1	Yes	-
4	146.00	1	Yes	-
5	147.70	1	Yes	-
6	140.10	1	Yes	-
7	139.40	1	Yes	-
8	159.70	1	Yes	-
9	166.00	1	Yes	-
10	189.20	1	-	Yes
11	130.60	1	-	Yes
12	198.20	1	-	Yes
13	122.40	1	-	Yes
14	145.70	1	-	Yes
15	148.10	1	-	Yes
16	163.90	1	-	Yes
17	163.00	1	-	Yes
18	135.60	1	-	Yes
19	168.00	1	-	Yes
20	114.00	1	-	Yes
21	129.60	1	-	Yes
22	212.60	1	-	Yes
23	107.20	1	-	Yes

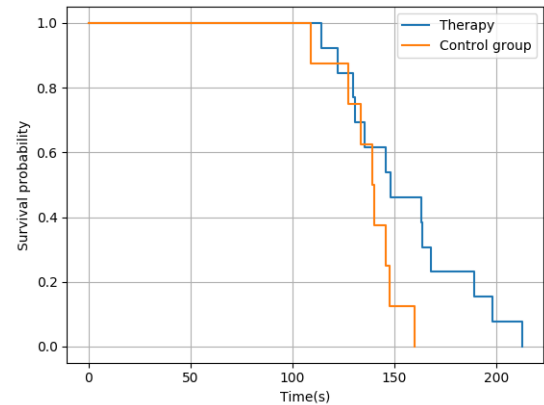
The Logrank test revealed statistic of 3.36 and a p-value of 7% when comparing the MYO control with treatment groups. Statistic was 8.59 and a p-value of less than 0.5% was found when comparing the GTC control with treatment groups. Results suggest a more pronounced difference between survival plots for the NPS-MYO group. As regards computational time, it was spent about one hour to finish all Python analysis.

### 5.1.1 Discussion

The NPS treatment is based on the concept that the brain comprises oscillators that couple together by means of a natural synchronization mechanism (BOARETTO *et al.*, 2021). Therefore, an



(a) Kaplan-Meier estimation (MYO-CTRL and NPS-MYO)



(b) Kaplan-Meier estimation (GTC-CTRL and NPS-GTC)

Figure 15 – (a) Kaplan-Meier estimation for MYO-CTRL and NPS-MYO groups. It is possible to observe in the survival probability graph that the NPS-MYO group achieved better results than the control group. It takes longer to observe the beginning of myoclonic behaviour in the NPS-MYO group. (b) Kaplan-Meier estimation for GTC-CTRL and NPS-GTC groups. It is possible to observe a better survival probability estimation for the NPS group once more. It takes longer to observe the beginning of generalized tonic-clonic seizure in the NPS-GTC group. For both comparisons, the Logrank test is used to statistically demonstrate the NPS treatment effectiveness.

epileptic seizure is a result of a hyper-synchronization phenomenon that can be suppressed by means of desynchronizing its temporal pattern (ZHOU; LIN, 2021; BOARETTO *et al.*, 2021; REIS *et al.*, 2021).

THIVIERGE; CISEK (2008) contributed to the understanding of a temporal pattern in a network of neurons by suggesting the possibility of a non-periodic synchronization of heterogeneous networks of spiking neurons, which would aid in elucidating epilepsy dynamics and provide valuable information aimed to develop new types of treatments, such as a non-periodic stimulation.

Furthermore, COTA *et al.* (2016) and OLIVEIRA *et al.* (2014) performed pre-clinical trials involving rats and achieved promising results from the NPS treatment. However, a more in-depth statistical analysis aimed to prove its effectiveness had never been performed. Therefore, a Survival Analysis was carried out in addition to the Kaplan-Meier estimator and Logrank hypothesis test in order to compare survival plots, thus further demonstrating that the NPS treatment is significantly efficacious.

Tables 1 and 2 displayed time differences for treatment and control groups. Nonetheless, it is neither possible to state that there is a marked difference among them, nor to suggest any difference between NPS-MYO and NPS-GTC treatments.

The Kaplan-Meier graph (Figure 15) shows survival estimations for each group, which more clearly indicates the difference between treatment and control groups, although it is still

impossible to definitively assert that the NPS treatment has achieved a substantial difference. However, it can be observed that, for the NPS-MYO group, the survival graph becomes different from MYO-CTRL from around 40 seconds time onward. Survival time observed is until 120 seconds or so, and the survival probability for the control group is substantially different from that achieved by the treatment group, thus suggesting effectiveness. The same was observed when comparing GTC-CTRL and NPS-GTC groups, nonetheless, the difference between groups is closer to the survival time length. This indicates that the treatment is more efficient and presents better results for myoclonic seizures.

The Logrank test was useful to suggest the NPS treatment effectiveness. For MYO groups, it was observed a difference with a p-value of less than 0.5%, but it was observed a difference with a p-value of 7% for GTC groups. It is found through the Logrank test that the NPS treatment increases the time taken to develop an epileptic seizure. Furthermore, it is once again feasible and more effective when NPS is used to treat myoclonic seizures, which is evidenced by observing the Kaplan-Meier survival plots.

#### 5.1.1.1 Strengths and limitations

In this study, we carried out a Survival Analysis using behavioral data only. On the other hand, electrophysiological data is of major importance to better understand epileptic phenomena and the underlying mechanisms of NPS. Such investigation is currently being carried out in the Laboratory of Neuroengineering and Neuroscience at the Federal University of São João Del Rei with promising results, and a few of which have already published (OLIVEIRA *et al.*, 2019). Nonetheless, the goal of this thesis is within the boundaries of a limited scope, which is to add further and more solid statistical evidence to that provided by the original study on NPS (COTA *et al.*, 2009), which was carried out by solely using behavioral data.

Different values of PTZ rates of infusion have distinct impacts on the latency to convulsive behavior results due to myriad factors (including pharmacokinetics), but many are still unknown (OLIVEIRA *et al.*, 2018). Be that as it may, considering ethical issues on animal usage, the present infusion rate (10 mg/ml/min) has remained the same across many studies performed by this group and collaborators, including this thesis.

It is also worth mentioning that, although a drug assessment of novel treatments is of great importance, the PTZ experiment is a model of controlled ictogenesis and acute seizures only, thus bearing considerable differences from clinical trials scenarios. A stronger extrapolation of results to seizures occurring in human patients spontaneously would be better supported if such an analysis were performed using data from animal models of epileptogenesis (e.g. fast amygdalar kindling, late-phase after pilocarpine-induced status epilepticus, etc.). However, for similar reasons, this is also beyond the scope of this thesis.

## 6 DYNAMIC BAYESIAN NETWORK RESULTS AND DISCUSSION

### 6.1 Dynamic Bayesian Network

Figures 18 and 16 show the use of the LFP database to build the Dynamic Bayesian Network. Figure 18 represents the Markov Chain evaluation among the employed time slices, and Figure 16 represents the Dynamic Bayesian Network built from rats LFP database. Arcs were evaluated using the method suggested by (GROSS *et al.*, 2019), thus providing two thresholds: a first one from the Basal state until reaching the MYO time slice due to more data availability, which provided a value of 0.62 and standard deviation of 0.10, which yields a minimum value of 11 arcs. Threshold during GTCS time was 0.71 with standard deviation of 0.16, representing a minimum value of 5 arcs. In Figure 16, it is possible to observe gray edges, but none of which have passed through the analytical threshold.

Table 3 reports significant arcs, i.e. those achieving the minimum frequency threshold for each time slice. The table also depicts the number of Dynamic Bayesian Networks in which connection appeared. There is also a separation between the two groups of rats used in this thesis so as to present the global value used to develop a comparison with the analytical threshold calculated, only except for the values of GTCS connections, which were compared to those found in the GTCS group column, since the recording of this time slice was performed exclusively on this experimental group.

Figure 17 presents a figure of significant arcs by comparing them with analytical threshold values. It was possible to check that there is a single direction from one node to another, such as in the case of Thalamus and Hippocampus during the GTCS time slice in which it was verified the arc TH→HP. However, TH←HP did not pass through the analytical threshold. There is only one exception during Basal and Infusion time slices in which the relationship between Thalamus and Cortex provided the same probability for both directions.

It is possible to observe from Figure 18 that the GTCS time slice is independent of the Basal time slice. Also, there is a suggestion that Infusion and MYO time slices connect them. According to Figure 16, the most robust connections patterns were: from Basal until MYO time slice, the interconnection structure remained the same which comprises the Thalamus as central node connecting to the Hippocampus and Cortex, which are in turn independent. The structure changed during the GTCS time slice, and Hippocampus now became the primary node. Cortex connects with Thalamus, which in turn appears linked with Hippocampus. The Dynamic Bayesian Network identified other connections represented in Figure 16, but the analytical threshold has not been validated (grey color).

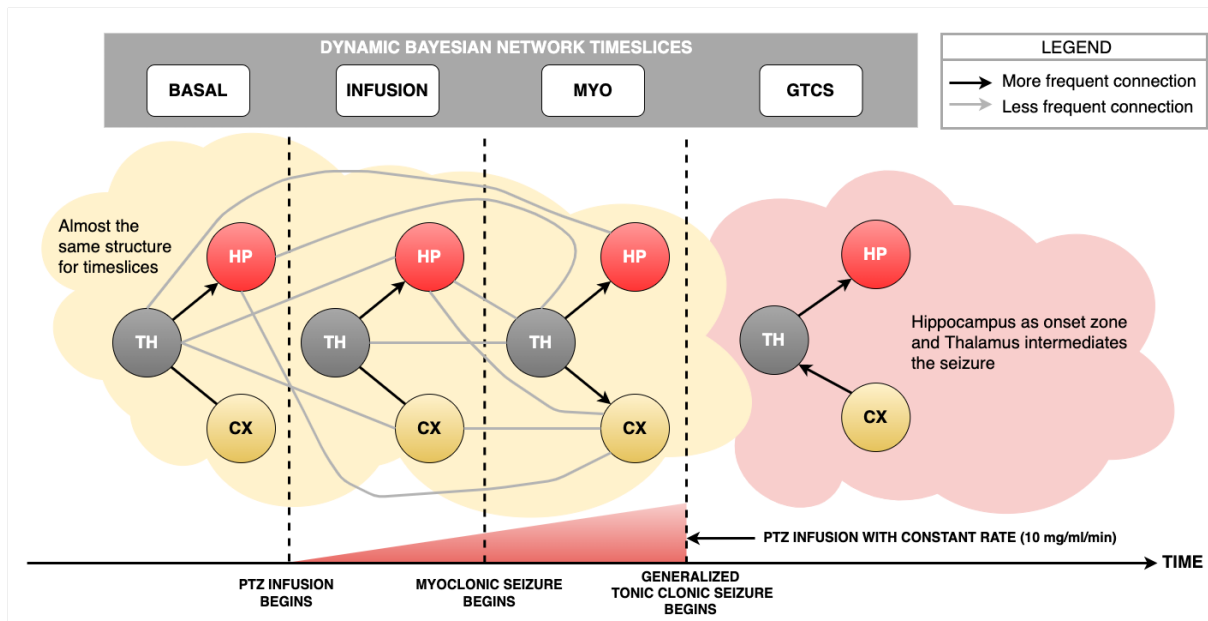


Figure 16 – Dynamic Bayesian Network developed from the LFP dataset. Black arcs represent the strongest connections provided by the analytical threshold method proposed by (GROSS *et al.*, 2019). Gray arcs represent connections that were not validated by the analytical threshold, but provided by the DBN method. From Basal state until MYO time slice, there is a common pattern, with Thalamus distributing information to the Hippocampus and Cortex. This behavior is reported in literature due to knowledge about the role of Thalamus as an important communication lane to distribute information among brain areas. The novelty lies in the fact that the same structure of communication is found during the MYO time slice. It was expected a transition structure closer to GTCS, but it was not so. During the GTCS time slice, the connection among HP, TH and CX change. It is possible to check that Hippocampus might be the onset zone by observing information reinforcing synchronization from Thalamus. There is another important path observed and reported in literature that information stemming from the Cortex reverberates in the Thalamus to be transmitted to the Hippocampus.

### 6.1.1 Discussion

#### 6.1.1.1 Temporal evolution of DAGs reflect neurodynamics of ictogenesis

The DBN results found in this study have clearly shown distinct connectivity patterns during ictogenesis induced by a controlled infusion of PTZ - Figure 16. Present findings corroborate the dynamic nature of functional neural connectivity along the time course of epileptic phenomena, while at the same time providing novel insights.

An initial interesting result is that the DAG remains unaltered during the whole PTZ infusion period, which is the same as that in the basal state - Figure 16, Basal and Infusion time slices. This connectivity pattern observed in both pre-ictal time slices is perfectly understandable and supported by well-understood information flow within neural circuitry in homeostasis. Notably, there is a recognition of the thalamus as the central relay for both incoming sensory



Table 3 – The strongest arcs identified in developed Dynamic Bayesian Network. All of them have had their minimum frequency calculated through the analytical threshold defined in (GROSS *et al.*, 2019). Due to the fact that the GTCS time slice is observed only in for the GTCS group, analytical threshold was compared with the frequency observed only for this group. For Basal, Infusion and MYO time slices, the global frequency was used to perform such a comparison. Gray edges were not considered in the table due to the fact that they did not pass in the analytical threshold analysis (Figure 16).

From	Arc (→)	Frequency		
		MYO group	GTCS group	Global
Thalamus Basal	Cortex Basal	7	4	11
Cortex Basal	Thalamus Basal	7	4	11
Thalamus Basal	Hippocampus Basal	8	3	11
Thalamus Infusion	Cortex Infusion	8	3	11
Cortex Infusion	Thalamus Infusion	7	4	11
Thalamus Infusion	Hippocampus Infusion	6	6	12
Thalamus MYO	Cortex MYO	8	4	12
Thalamus MYO	Hippocampus MYO	8	7	15
Cortex GTCS	Thalamus GTC	0	7	7
Thalamus GTCS	Hippocampus GTCS	0	6	6

information on passage to multiple primary cortices, and also for motor output from the motor cortex - Figure 16, Basal, Infusion, and MYO time slices. Thus, the observed bidirectional link between those nodes is consistent with ongoing sensory and motor function - Figure 16, Basal and Infusion time slices. Additionally, directed arcs from TH to HP are probably related to the communication from the thalamus to the hippocampus underlying neural plasticity and acquisition of novel memory traces during wakefulness (CASSEL; VASCONCELOS, 2015) - Figure 16, Basal, Infusion, and MYO time slices. Such activity is relayed by the thalamus and fed into the hippocampus for a future conversion into long-term memories during sleep (KLINZING; NIETHARD; BORN, 2019).

Next, a myoclonic seizure starts, and there is a fundamental change in the DAG: the thalamus becomes the primary driver of both the hippocampus and the cortex (notice that the TH to CX arc is now preponderant) - Figure 16, MYO time slice. It is strikingly consistent with the motor expression of partial seizures originating in the limbic system, such as the observed forelimb clonus recorded at this moment during which the thalamus assumes the role of a central synchronization hub for both the cortex and the hippocampus ((BERTRAM *et al.*, 1998; BERTRAM, 2014) for reviews). In fact, for this reason, nuclei within the thalamus are the primary targets for neuromodulation strategies in the treatment of epilepsy (VLIS *et al.*, 2019).

Finally, two essential changes occur when crossing the generalized seizure onset - Figure 16, GTCS time slice. First, the connectivity pattern changes to include a preponderant communication - signal transmission among rats brain areas - from the cortex to the thalamus and thus to

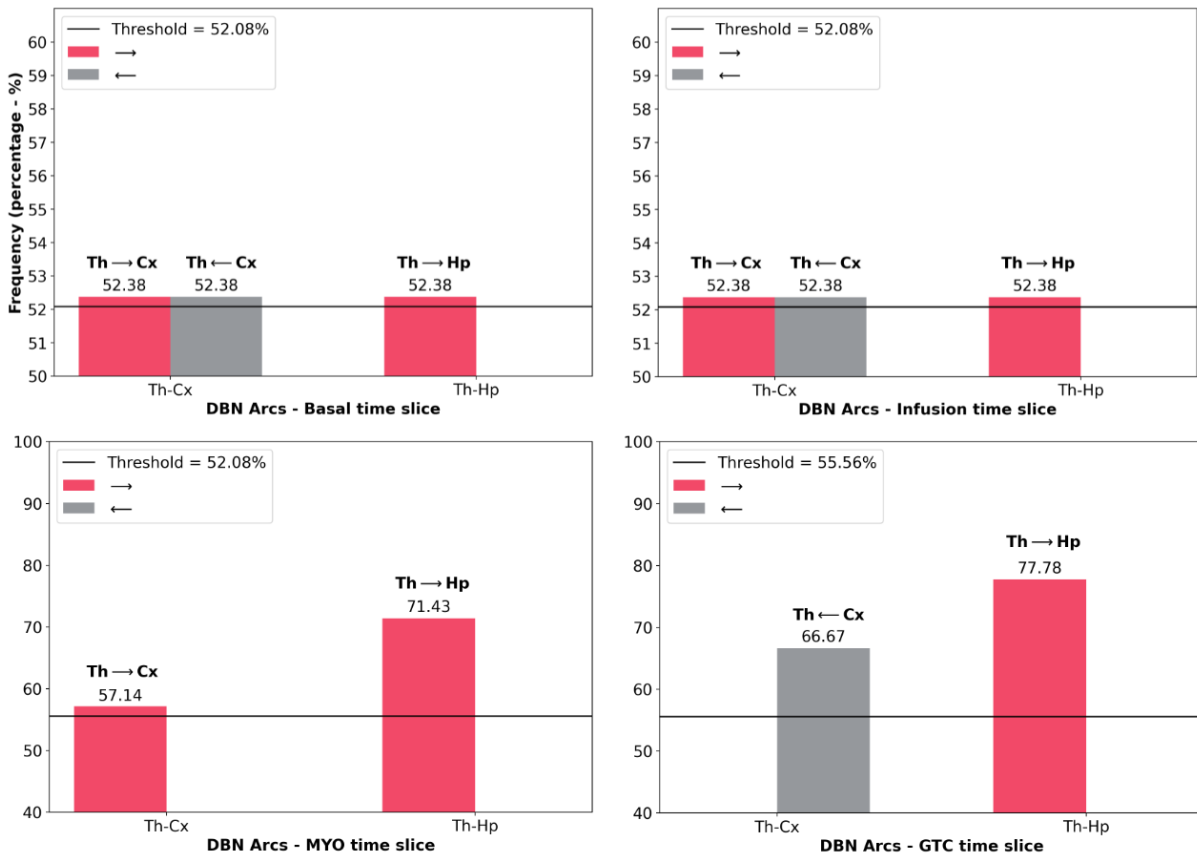


Figure 17 – Figure of the frequency of main arcs from the DBN presented in Figure 16. They were separated according to each time slice: Basal, Infusion, MYO and GTCS. Table 3 shows only the arcs that were over the threshold value (black line in figure) provided by the analytical threshold. It is also possible to observe that an opposite connection is sometimes unchecked, such as TH-CX during MYO time slice (TH→CX passed through the threshold, but TH←CX did not). The exception is during Basal and Infusion time slices in which is possible to observe the same frequency for TH-CX arcs.

the hippocampus. Once again, this is in agreement with previous literature showing a recruitment of vast neocortical territories and communication from these areas to the thalamus and other forebrain structures during secondary generalization ((BRODOVSKAYA; KAPUR, 2019) for a review). Although there is a canonical understanding of the importance of the thalamocortical neural circuit, which is a reverberant loop, it may imply a bi-directional connection between these two areas, other processes are also crucial for the onset of generalized seizures. These include a unidirectional cortex to thalamus drive through polysynaptic connections involving the basal ganglia (striatum, globus pallidus, and substantia nigra reticulata). Electrode geometry used for cortical recording may also play a unique role in this context. Given the larger dimensions of electrodes made out of surgical screws, when compared to the microwires used for deep brain recording, signals indeed represent contributions from much larger brain areas. Thus, an aberrant recruitment of vast neocortical territories and their powerful drive onto thalamic nuclei may also be a primary contributor to the preferential direction of the CX to TH arc observed in our

results. A second significant change after the onset of generalized seizures lies in the absence of DAG arcs crossing this temporal limit. Epileptiform activity during partial seizures bears some neurodynamical correlation with base-level tracings. Meanwhile, generalized tonic-clonic seizures have dynamics of their own that cannot be correlated with those of other time slices across ictogenesis. The clarity of reasons for such a result is unclear, but this intriguing result may probably have important implications on strategies for neuromodulations, particularly those involving responsive close-loop systems capable of detecting ongoing seizures.

#### 6.1.1.2 Graph evaluation and analytical threshold to identify DBNs arcs direction

Figure 18 shows a key finding: after building networks, it was found that Basal intervals do not help to explain what happens during the GTCS period. Also, Infusion and MYO timeslices connect them. It represents that an infusion of PTZ drug disconnects timeslices, which means that it is a different ongoing process since the beginning of drug administration. Table 3 substantiated the discussion about the Dynamic Bayesian Network developed herein, suggesting structures found in neuroscience literature. The Dynamic Bayesian Network method brought many possible connections, as expected. The analytical threshold was essential to analyze their significance, screening the most important arcs and enabling a better interpretation of results. Observing Figure 17, the direction of significant arcs is clear, such as the Cortex and Thalamus during GTCS time in which  $CX \rightarrow TH$  had a frequency above the threshold (7 against the minimum of 5). However,  $CX \leftarrow TH$  frequency did not pass through the threshold (3). According to the present methodology, the arcs that did not pass through the threshold frequency represent connections whose effective existence is not certain. Nonetheless, they may be essential tracks for further studies. The most important was an alignment between DBN combined with threshold analysis and neurobiological phenomena. It is essential to confirm the approach as feasible to investigate epilepsy dynamics. However, according to (BERTRAM, 2013), (LOSI *et al.*, 2019) and (HEYSIEATTALAB; SADEGHI, 2021), there is a need for further knowledge about the illness dynamics, such as the causal relationship among brain areas.

#### 6.1.1.3 Other approaches for Functional Connectivity analyses and limitations

In literature, it is possible to verify other papers performing Functional Connectivity Analyses considering other approaches:

- (TSUKAHARA *et al.*, 2020a) combined Partial Directed Coherence and Mutual Information to study the connectivity and transmission rate between brain areas, considering the same dataset used in this thesis and only the Basal and Infusion times on this study. However, it only identified connections among all brain areas, but no novel findings regarding ictogenesis were identified.
- (TSUKAHARA *et al.*, 2020b) applied Delayed Mutual Information to identify associations among brain areas, considering the lag of communication and the same dataset used in this

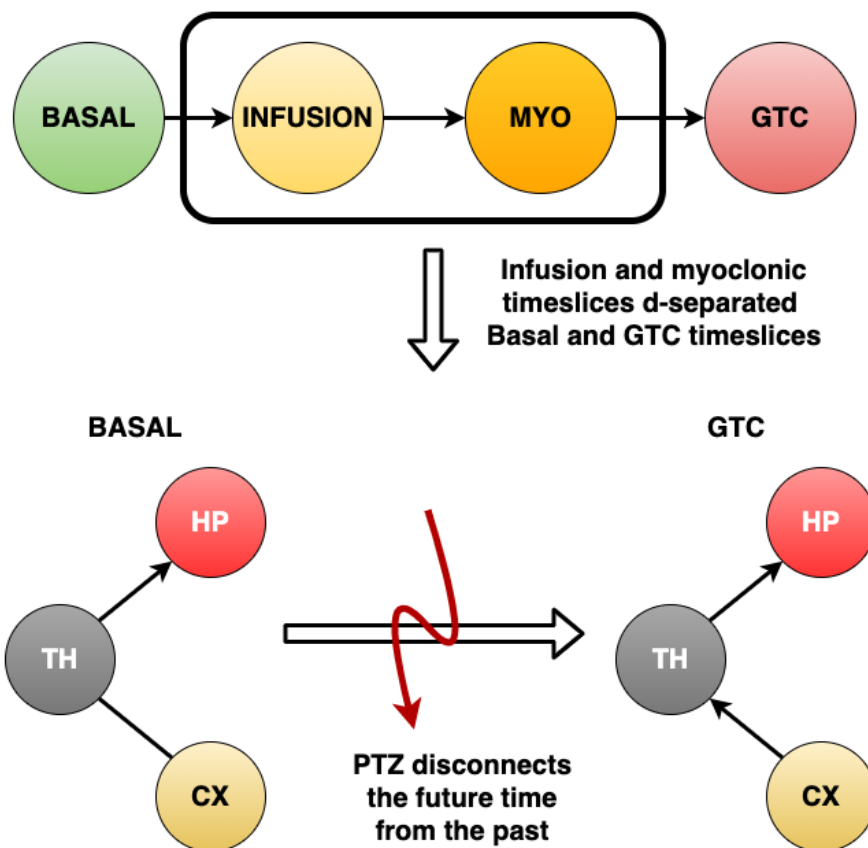


Figure 18 – The figure represents the four time slices used to develop the (Dynamic) Bayesian Network for each rat: Basal, Infusion, MYO and GTCS. After building networks, it was observed that Basal interval does not help to explain what happens during the GTCS period. Moreover, it was verified that Infusion and MYO time slices connect them. This represents that an infusion of PTZ drug disconnects timeslices, which means that it is a different process ongoing since the very beginning of drug administration.

thesis to develop the analysis. The method did not enable an identification of any novel finding regarding ictogenesis.

As it can be observed, the Information Theoretic approach, as well a linear approach in the frequency domain, are both common approaches used in neuroscience (CIARAMIDARO *et al.*, 2018; VAROTTO *et al.*, 2018; GRIBKOVA; IBRAHIM; LLANO, 2018), but no novelty regarding ictogenesis phenomena was found. There are some concerns to be considered in this thesis regarding its referenced works: the dataset used to perform analysis. A possible problem that can interfere with results is the acquisition of Local Field Potentials using a sampling frequency of 1 kHz. Discretization using the maximum number of bins was required to ensure a better resolution of signals. (ENDO *et al.*, 2015) performed a similar study as observed in (TSUKAHARA *et al.*, 2020b), however, it was used 32 bins to discretize signals sampled in 24 kHz. There is a sharp difference in signals resolution that resulted in different findings. In

(ENDO *et al.*, 2015) it was possible to identify the lag among neurons communications, while in (TSUKAHARA *et al.*, 2020b) there was no lag identification initially.

Another significant limitation to consider is the volume of data required to apply the Theoretical Information approach (ENDO *et al.*, 2015). The dataset used to perform the analysis is restricted due to the availability of rats to perform this study. Inside this scenario, a Bayesian approach may be favorable, given that initial information provided by a specialist helps to deal with smaller datasets, thus providing results as it can be observed in this study.

Partial Directed Coherence is a linear approach to perform Functional Connectivity analysis, and real-world problems are usually nonlinear, as it is the case of Local Field Potentials, animal physiology (PHAN *et al.*, 2019). Therefore, in addition to the fact that PDC brought some insights about ictogenesis, it revealed no novelty. Another significant limitation is the requirement of stationarity to apply the method, which can be a problem when studying ictogenesis phenomena.

Despite these limitations, the adoption of a Dynamic Bayesian Network approach revealed findings that are in accordance with neuroscience literature and brought some new knowledge to the light. There are some other questions to be answered considering the subject. However, it is the subject of further studies.

Finally, due to data availability to perform the analysis, only three areas were considered in this thesis: Thalamus, Hippocampus, and Cortex. Future studies should consider more than three areas aiming at a broad scope to study epileptic seizures. However, only three brain areas require less computational resources to apply the DBN analysis and provided significant findings regarding epileptogenesis.

## 6.2 Delayed Dynamic Bayesian Network combined with Multivariate Stochastic Volatility

An application of the DBN method in the dataset resulted in the K2 scores reported on Table 4, column *DBN (K2 Score)*. Afterwards, as mentioned in the *Algorithms* section, there was the Multivariate Stochastic Volatility method was employed to study the lag of communication among brain areas: CX, HP, and TH, studied in this thesis. Table 5 depicts the method results using TH as a reference to consolidate the lag.

The residual analysis was carried out in order to verify the MSV model adherence and observe the normality of the error variable of the MSV volatility model, Equation 2.2,  $\eta_m^x[t]$ . Figure 20 shows a plot of residuals as a function of normal distribution. Analytically, Durbin Watson test was carried out for heteroscedasticity and resulted in a statistic value of 2.76, and the Anderson-Darling test was performed for normality and yielded a statistic value of 0.28, with the following critical values for each significance level: 0.55 for 15.00%, 0.66 for 10.00%, 0.79 for 5.00%, 0.92 for 2.50% and 1.09 for 1.00%.

Rat	DBN (K2 score)	DBN considering lag (K2 score)	Difference (%)
R048	-3162608.86	-3330484.61	167875.75 (5%)
R052	-2274075.70	-2383758.00	109682.30 (5%)
R064	-2275717.19	-2306802.05	31084.86 (1%)
R065	-2980149.97	-3107298.32	127148.35 (4%)
R066	-3235200.40	-3281317.54	46117.14 (1%)
R067	-4427973.96	-4495950.86	67976.90 (2%)
R068	-2514562.51	-2828294.41	313731.90 (12%)
R070	-4580385.43	-4590411.94	10026.51 (0%)
R072	-2213137.01	-2808300.10	595163.09 (27%)
<b>Mean</b>	<b>-3073756.78</b>	<b>-3236957.54</b>	<b>163200.76 (17%)</b>

Table 4 – The table depicts the K2 score obtained from the Bayesian Network method. *DBN column* presents the values of scores calculated from Dynamic Bayesian Networks using the data frames disregarding the lag among brain areas. *DBN considering lag column* reports the values of scores calculated from Dynamic Bayesian Networks considering the Multivariate Stochastic Volatility lag results, revealing that there is a communication delay among TH, HP, and CX rats brain areas. The score results suggest that the lag implies better DBN model adherence with the dataset - a difference on average of 17.00%.

Table 5 – The table reports the lag among brain areas obtained from the Multivariate Stochastic Volatility method, provided in samples. Due to the frequency of sampling adopted to perform Local Field Potentials acquisition, 1kHz, each sample represents the time of 1ms.

Rat	Time slices lags among brain areas (samples)							
	Basal		Infusion		MYO		GTCS	
	TH→CX	TH→HP	TH→CX	TH→HP	TH→CX	TH→HP	TH→CX	TH→HP
R048	12.00	4.00	3.00	29.00	5.00	27.00	11.00	5.00
R052	0.00	0.00	33.00	21.00	1.00	0.00	1.00	1.00
R064	0.00	0.00	0.00	0.00	16.00	1.00	29.00	2.00
R065	0.00	3.00	6.00	2.00	0.00	2.00	0.00	17.00
R066	17.00	2.00	15.00	2.00	25.00	3.00	12.00	6.00
R067	29.00	2.00	33.00	1.00	0.00	1.00	40.00	5.00
R068	15.00	2.00	11.00	2.00	45.00	1.00	0.00	2.00
R070	0.00	3.00	0.00	29.00	0.00	25.00	0.00	28.00
R072	24.00	1.00	39.00	0.00	24.00	1.00	31.00	1.00
<b>Mean</b>	<b>11.00</b>	<b>2.00</b>	<b>16.00</b>	<b>10.00</b>	<b>13.00</b>	<b>7.00</b>	<b>14.00</b>	<b>7.00</b>

There was a different lag for each time slice, and there were higher values among TH → CX than TH → HP communication. An incorporation of the obtained lag into the MSV method resulted in the K2 scores reported in Table 4, column *DBN considering lag (K2 Score)*. A consideration of lag communication resulted in a difference of 17% of the obtained K2 score, on average.

Figure 19 shows the best Dynamic Bayesian Network developed from the dataset, assigned to rat R052, which resulted in the best K2 score. The Bayesian Network considered the lag of communication among brain areas.

There was a test of the R052 DBN with the dataset to verify its fit. Table 6 presents

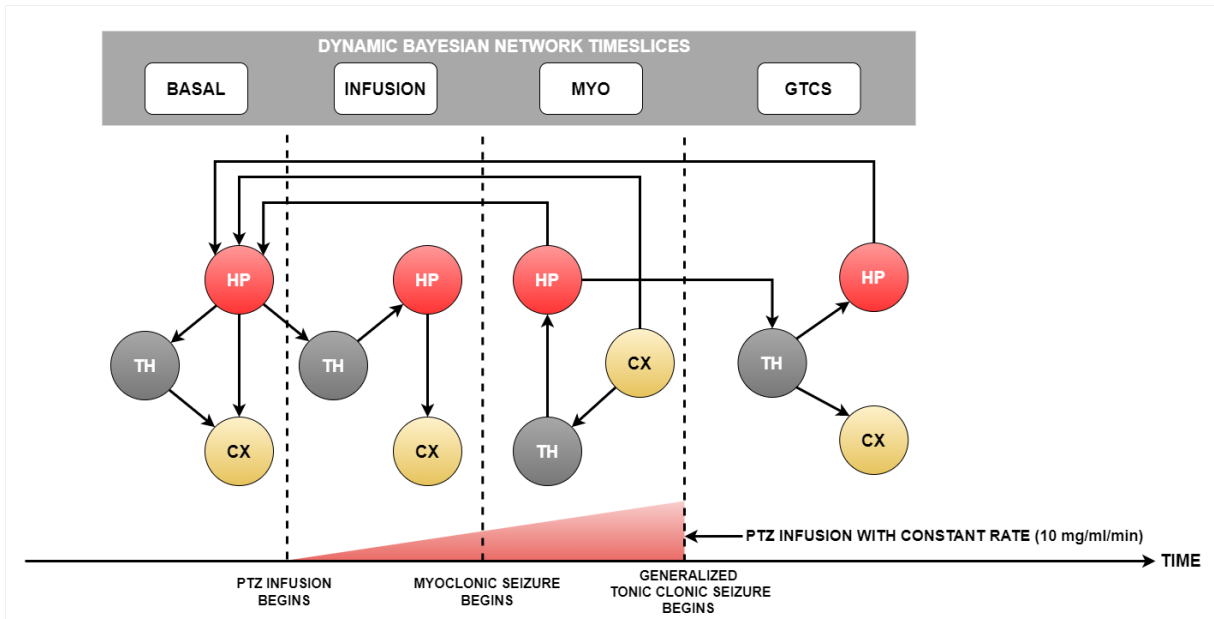


Figure 19 – The figure shows the Dynamic Bayesian Network of rat R052. It was the best DBN model developed for all rats. It is essential to mention that Bayesian Network presents a probabilistic relationship among variables. It should be considered that the DBN model is causal only after an expert evaluation to point out a correlation of results with neurophysiology. The study considers this analysis to discuss the results.

Rat	DBN (K2 Score)	DBN considering lag (K2 Score)	Difference (%)
R048	-3475044.20	-3703528.36	228484.16 (7%)
R052	-2274075.70	-2383758.00	109682.30 (5%)
R064	-2377899.37	-2403113.46	25214.09 (1%)
R065	-3148276.70	-3269655.75	121379.05 (4%)
R066	-3318412.33	-3367455.86	49043.53 (1%)
R067	-4544873.06	-4609396.26	64523.20 (1%)
R068	-2653081.49	-2937418.11	284336.62 (11%)
R070	-5318334.26	-5327615.12	9280.86 (0%)
R072	-2296311.02	-2893821.11	597510.09 (26%)
<b>Mean</b>	<b>-3267367.57</b>	<b>-3432862.45</b>	<b>165494.88 (6%)</b>

Table 6 – The table reports the K2 scores obtained from the Bayesian Network method considering the best DBN model (rat R052). Score results demonstrate that the R052 Dynamic Bayesian Network model improved the fit with the dataset. Once again, the lag enhanced the mean of the K2 score by approximately 6.00%.

the results considering two situations, either with or without the lag of communication among brain areas. Results highlighted that R052 DBN presented better K2 scores considering both simulation situations. Once more, the lag of communication presented a K2 score difference of 6%, on average.

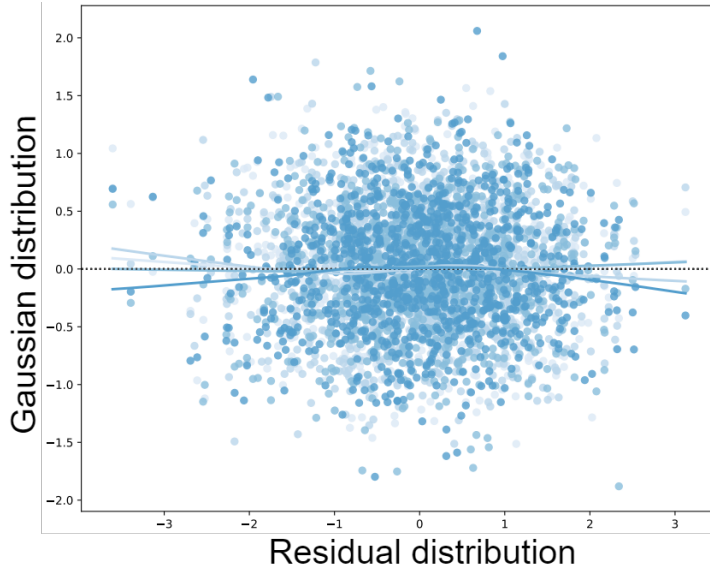


Figure 20 – Graph of Gaussian distribution against residual distribution values ( $\eta_n^x[t]$  from Equation 2.2). The resulting mean is approximately zero, indicating normality. Anderson-Darling statistical test provides analytically proof of MSV adherence.

### 6.2.1 Discussion

Rats of the same species with approximately similar weights submitted to the same experimental protocol might be physiologically different from each other (KRUBITZER; CAMPI; COOKE, 2011). This statement could be checked through the results of the DBN methodology to model rats database. The K2 score reported in Table 4 shows a different fit of the data frame from different rats with the model. However, there is an explosive growth (superexponential) of the number of possible DAGs given the number of vertices, which indicates that it is impossible to do an extensive search. There are several approaches to find a reasonably good data fit. The state-space of the Hill-Climbing algorithm allows a variety of results, once it is a probabilistic approach. Both issues ensure the variability of obtained results, as expected.

The use of Multivariate Stochastic Volatility to investigate the lag among recorded rats' brain areas seemed a suitable alternative. Local Field Potentials of epileptic animals are signals with kurtosis, nonlinearity, and heteroskedasticity, following the rationale that real-world problems are nonlinear. The method could deal with the signals, providing communication lags among brain areas.

Residuals of the autoregressive component of the MSV model seemed to follow a normal distribution pattern, i.e. an important indication of good adherence with Local Field Potential signals. Another piece of evidence is found in the plot of Gaussian distribution as a function of residual distribution values (Figure 20), and the Anderson-Darling normality test confirmed the results.

The communication delay observed in Table 5 presents values that are in accordance with reported literature, such as the case of papers published by Endo et al. (2015) (ENDO *et*



*al.*, 2015) and Bonnefond et al. (2017) (BONNEFOND; KASTNER; JENSEN, 2017). The fact suggests that the method is a suitable alternative. A similar study on the lag of communication using the same dataset and the Delayed Mutual Information (DMI) method is reported in a paper by Tsukahara et al. (2020) (TSUKAHARA *et al.*, 2020b).

However, although this investigation discussed the transmission rate among CX, HP, and TH, it could not provide communication lag information. Some hypotheses substantiate the obtained results, such as the requirement of a large sample size for a successful DMI method and the frequency of sampling that impairs this type of analysis. The use of Markov Chain Monte Carlo simulation also contributes, once it is a Bayesian Inference made from MSV model distributions, in addition to helping to reduce the dependence of reliable results on the sample size of the dataset. Nonetheless, these assumptions require more in-depth studies, even though some reported literature correlates with the cited situations (ENDO *et al.*, 2015; AKÇA; YOZGATLIGIL, 2020).

The MSV method also allows expert knowledge to make prior assumptions, which may help functional connectivity analysis and accelerate the MSV model convergence. Although the study used flat priors to make the inference noninformative, two-stage sampling methods through Gibbs and the Metropolis-Hastings samplers ensured reliable results.

In addition to the communication delay among brain areas to develop the DBN method, Table 4 reports the K2 scores obtained from the approach. By comparing the new results with previous K2 scores, it is found that the lag consideration inside the DBN model improved fit with all datasets, thus indicating the difference of 17% on average.

The DBN model, which presented the best fit when compared to all Bayesian Networks developed with the dataset is illustrated in Figure 19 regarding rat R052. There was result confirmation after testing the mentioned DBN model involving all rats' Local Field Potentials. The K2 scores obtained in this study were improved and compared with the DBN methodology's initial application. Table 6 shows all K2 scores, with and without considering the communication lag in the DBN model. Once again, the use of communication delay among brain areas resulted in a model with better fit when compared to the DBN model without communication lag. These results are in compliance with reported literature which already suggested that communication among brain areas sometimes presents a delay (PARIZ *et al.*, 2021), and this lag organizes the brain network synchronization (PETKOSKI; JIRSA, 2019).

An interpretation of the DBN model of rat R052 by a neuroscience expert was essential to transport the probabilistic relationships encountered from the method to a causal analysis of epileptogenic dynamics. It is worth keeping in mind that the DBN method provides probabilistic relationships among variables, and that there is a need to evaluate the resulting associations. In general, the results make sense considering neuroscience literature. However, there are some concerns to be considered.

Observing Figure 19, the Hippocampus driving the activity of the Thalamus and Cortex seems correct. There is no feedback from the Hippocampus in Cortex communication, but only from the Thalamus. In the case of the rat's movement in a restricted space, as it occurred during Local Field Potentials signals acquisition, the Septum-Hippocampal pathway might generate a rhythm on Theta frequency of the rat's Local Field Potential, driving the Thalamus and Cortex activities and also inducing the activity from the Thalamus to the Cortex. The Hippocampus from the Basal time slice inducing the future activity of the Thalamus makes sense in the dynamics of epileptogenic activities.

A study carried out by TAKEUCHI *et al.* (2021) supports the suggestion of the Septum-Hippocampal pathway as an important factor playing a role in abnormal oscillations, in the Theta band of Local Field Potentials inside the context of the mechanisms to disrupt epileptic seizures. Furthermore, the study suggests the area as vital to treat the disease and restore the brain to its regular patterns of brain rhythms oscillation.

During the Infusion time slice, the circuits presented in Figure 19 make sense. There was a rat manipulation, and then it got off the ground, maybe due to the fact that they were already under the PTZ GABAergic effect, suggesting thalamic influence. There is an interesting fact that there is no influence of the Thalamus on the Cortex, and the Cortex on the Thalamus. However, such a connection has an intricate relationship with the animal's state, including its reaction to the PTZ infusion. There is a need for more in-depth studies to evaluate these influences.

Myoclonic seizure is usually a phenomenon driven by limbic circuits. Thalamus and Cortex are affected during both situations: myoclonic and tonic-clonic seizures, especially at this moment. However, there is no influence of the Hippocampus on the Thalamus or Cortex during the MYO time slice. A hypothesis for these results is that cortical circuitry disinhibition leads to thalamic circuitry disinhibition. This situation leads to hippocampal circuitry disinhibition. The Hippocampus influence on the Thalamus variable from the GTCS time slice is also in agreement with neuroscience literature (PARMEGGIANI; LENZI; AZZARONI, 1974).

During the GTCS time slice, the communications among brain areas depicted in Figure 19, are in accordance with established neuroscience literature. Sherman & Guillery (2006) (SHERMAN; GUILLERY, 2006) describes the influence of the Thalamus on cortical activity. The study of Cassel & Vasconcelos (2015) (CASSEL; VASCONCELOS, 2015) reports the Thalamus as a driving area to hippocampal functions.

Even though PTZ is a drug used to induce seizures, there is no modification inside the rat's brain circuitry with the present approach. The PTZ drug disrupts the past with the events after its administration, thus starting a new process. In Figure 19 shows the influence of the GTCS and MYO Hippocampus variables on the Hippocampus of the Basal time slice. These connections among different timestamps, initially unforeseen, raise a new hypothesis to be considered in future studies, once it involves more complex mechanisms of the rat's brain physiology.

## 7 CONCLUSION

The NPS is a promising alternative for refractory epilepsy treatment. Although there are papers describing the approach and suggesting its effectiveness in assisting seizure suppression, a survival analysis to confirm its effectiveness has not been developed yet. To confirm the effectiveness of a new treatment, a survival analysis can be of assistance, given that it is widely known as a way to test new drugs and treatments in the medical area. Kaplan-Meier estimation graphically enhanced the clarity of treatment results, nonetheless, it was impossible to state that there is significant treatment efficiency, although it has been suggested that using the NPS to treat myoclonic seizure is more effective. The Logrank test evaluated and provided the statistical significance of the NPS treatment, moreover, it confirmed that its effectiveness is greater at treating myoclonic seizures. Therefore, using a survival analysis to evaluate the NPS treatment provided a more complete overview and discussion about seizure suppression performance.

The Dynamic Bayesian Network method represented an affordable approach, as it offers insights into epilepsy dynamics. It was possible to observe that an infusion of PTZ drug disconnects time slices, which means that it has been a different ongoing process since the beginning of drug administration. The DBN analysis was able to accurately capture the dynamic nature of brain connectivity across ictogenesis with a significant correlation to neurobiology derived from pioneering studies employing techniques of pharmacological manipulation, lesion, and also modern optogenetics (FORCELLI, 2017). Additionally, it provided exciting novel insights, such as the discontinuity between forelimb clonus and GTCS dynamics.

According to (COLMERS; MAGUIRE, 2020), the direction of associations between the nodes formed by the areas of the brain during epileptic seizures is still an open problem. The study aimed to address the problem and provided information in agreement with (TRACY *et al.*, 2021), showing that basal and infusion time slices present a different pattern of communication than that observed during MYO and GTCS time slices. It is essential to observe that MYO and GTCS time slices present different communication patterns, and provide information about the crossing from both states. Arising from a more focal seizure (MYO) into a tonic-clonic seizure (generalized in GTCS), it suggests evidence of a work performed by (LIGNANI; BALDELLI; MARRA, 2020). The study stated that changes in communication direction are associated with two critical processes: the generation and expression of the seizure and the maintenance of the epileptogenic phenomenon.

It was possible to observe the temporal evolution of variables across time and determine other communications according to a transition from resting-state until the offset of generalized tonic-clonic seizure. Epileptiform activity during partial seizures bears some neurodynamical correlation with base-level tracings. Meanwhile, generalized tonic-clonic seizures have a dynamic of their own that cannot correlate with those of other time slices across ictogenesis. However,

this intriguing result may probably have important implications on neuromodulations strategies, particularly those involving responsive close-loop systems capable of detecting ongoing seizures. For these reasons, DBN might be an excellent tool for investigating brain circuitry and their dynamical interplay in both homeostasis and dysfunction. Analytical threshold results supported all this discussion, allowing an evaluation of the arc's significance and identifying the connections observed in the developed DBN.

Computationally, the applied methodology demonstrated to be an appropriate alternative. Each rat data frame took about three minutes to be run and providing a suggestion of a DBN model. All databases were run in about fifty minutes, which is rather fast, mainly on account of the fact that the DBN method is an NP-hard problem. There was a study development using a personal computer without any adaptations, which is also interesting since it contributes to a reproducibility of results. Thus, the algorithms presented two features supporting its availability to perform a functional connectivity analysis: good computational time of processing and reproducibility. Therefore, the approach demonstrated feasibility to investigate epilepsy dynamics by identifying important insights reported in literature and new findings.

Furthermore, the Dynamic Bayesian Network method represented an affordable approach, giving insights into epilepsy dynamics. The use of LFP signals from the Basal state until GTCS time depicted the temporal evolution of the rat's brain, starting from the basal state and finishing on the generalized tonic-clonic seizure.

The Multivariate Stochastic Volatility model captured the lag of communication among the Cortex, Hippocampus, and Thalamus rat's brain areas, and presented good adherence with the model. Incorporating the delay inside the DBN model allowed an improvement in the results. This fact can be evidenced based on the obtained K2 scores calculated for the Bayesian Networks developed, either with and without lag results.

The best DBN model presented in the results section is in agreement with neuroscience literature and suggests future in-depth studies. Therefore, a combination of both methods represented a suitable alternative to perform a functional connectivity analysis within the context of studies involving Local Field Potentials recordings.

Further work should include the assessment of results and applications of survival analyses in different experimental scenarios, supporting a stronger extrapolation of results to spontaneously occurring seizures in human patients.

Another suggestion for further studies is still the need for further and more in-depth knowledge about the illness dynamics, such as using more brain areas to increase the scope of observation of the epileptogenic dynamics. Also, the use of Local Field Potentials might be applied with more sampling frequency to make signal representation more precise, thus enhancing the applied methodology. Finally, using the proposed methodology to study other types of brain disorders, like Parkinson's disease, sounds quite illuminating and insightful.

## REFERENCES

- AKÇA, E.; YOZGATLIGIL, C. Mutual information model selection algorithm for time series. **Journal of Applied Statistics**, Taylor & Francis, p. 1–16, 2020.
- ALTMAN, D. G.; BLAND, J. The log rank test. **BMJ**, v. 328, p. 1073, 2004.
- AVENA-KOENIGSBERGER BRATISLAV MSIC, O. S. A. Communication dynamics in complex brain networks. **Nature Reviews Neuroscience**, v. 19, p. 17 – 33, 2017.
- BARTOLOMEI, F. *et al.* Defining epileptogenic networks: Contribution of seeg and signal analysis. **Epilepsia**, v. 58, n. 7, p. 1131–1147, 2017. Available at: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/epi.13791>>.
- BEGHI, E. Social functions and socioeconomic vulnerability in epilepsy. **Epilepsy & Behavior**, Elsevier, 2019.
- BEGHI, E. *et al.* Global, regional, and national burden of epilepsy, 1990–2016: a systematic analysis for the global burden of disease study 2016. **The Lancet Neurology**, Elsevier, v. 18, n. 4, p. 357–375, 2019.
- BEHJATI, S.; BEIGY, H. Improved k2 algorithm for bayesian network structure learning. **Engineering Applications of Artificial Intelligence**, Elsevier, v. 91, p. 103617, 2020.
- BEN-GAL, I. Bayesian networks. **Encyclopedia of statistics in quality and reliability**, Wiley Online Library, v. 1, 2008.
- BENICZKY, S. *et al.* Seizure semiology: Ilae glossary of terms and their significance. **Epileptic Disorders**, Wiley Online Library, v. 24, n. 3, p. 447–495, 2022.
- BENJUMEDA, M. *et al.* Patient specific prediction of temporal lobe epilepsy surgical outcomes. **Epilepsia**, Wiley Online Library, v. 62, n. 9, p. 2113–2122, 2021.
- BERGLIND, F.; ANDERSSON, M.; KOKAIA, M. Dynamic interaction of local and transhemispheric networks is necessary for progressive intensification of hippocampal seizures. **Scientific reports**, Nature Publishing Group, v. 8, n. 1, p. 5669, 2018.
- BERTRAM, E. H. Neuronal circuits in epilepsy: do they matter? **Experimental neurology**, Elsevier, v. 244, p. 67–74, 2013.
- BERTRAM, E. H. Extratemporal lobe circuits in temporal lobe epilepsy. **Epilepsy & Behavior**, Elsevier, v. 38, p. 13–18, 2014.
- BERTRAM, E. H. *et al.* Functional anatomy of limbic epilepsy: a proposal for central synchronization of a diffusely hyperexcitable network. **Epilepsy research**, Elsevier, v. 32, n. 1-2, p. 194–205, 1998.
- BETZEL, R. F. Network neuroscience and the connectomics revolution. *In*: **Connectomic Deep Brain Stimulation**. [S.l.: s.n.]: Elsevier, 2022. p. 25–58.
- BHAT, H. S.; KUMAR, N. On the derivation of the bayesian information criterion. **School of Natural Sciences, University of California**, v. 99, 2010.

BIASIUCCI, A.; FRANCESCHIELLO, B.; MURRAY, M. M. Electroencephalography. **Current Biology**, Elsevier, v. 29, n. 3, p. R80–R85, 2019.

BIELZA, C.; LARRANAGA, P. Bayesian networks in neuroscience: a survey. **Frontiers in computational neuroscience**, Frontiers, v. 8, p. 131, 2014.

BLASI, R. A. D.; CAMPAGNA, G.; FINAZZI, S. A dynamic bayesian network model for predicting organ failure associations without predefining outcomes. **Plos one**, Public Library of Science San Francisco, CA USA, v. 16, n. 4, p. e0250787, 2021.

BLAUWBLOMME, T.; JIRUSKA, P.; HUBERFELD, G. Chapter seven - mechanisms of ictogenesis. *In*: JIRUSKA, P.; CURTIS, M. de; JEFFERYS, J. G. (ed.). **Modern Concepts of Focal Epileptic Networks**. Academic Press, 2014, (International Review of Neurobiology, v. 114). p. 155 – 185. Available at: <<http://www.sciencedirect.com/science/article/pii/B9780124186934000078>>.

BOARETTO, B. *et al.* The role of individual neuron ion conductances in the synchronization processes of neuron networks. **Neural Networks**, Elsevier, v. 137, p. 97–105, 2021.

BONNEFOND, M.; KASTNER, S.; JENSEN, O. Communication between brain areas based on nested oscillations. **eneuro**, Society for Neuroscience, v. 4, n. 2, 2017.

BORGER, V. *et al.* Resective temporal lobe surgery in refractory temporal lobe epilepsy: prognostic factors of postoperative seizure outcome. **Journal of Neurosurgery**, American Association of Neurological Surgeons, v. 1, n. aop, p. 1–10, 2021.

BOUSSAHA, N.; HAMDI, F. On periodic autoregressive stochastic volatility models: structure and estimation. **Journal of Statistical Computation and Simulation**, Taylor & Francis, v. 88, n. 9, p. 1637–1668, 2018. Available at: <<https://doi.org/10.1080/00949655.2017.1401626>>.

BRODOVSKAYA, A.; KAPUR, J. Circuits generating secondarily generalized seizures. **Epilepsy & Behavior**, Elsevier, v. 101, p. 106474, 2019.

BROOKS, S. *et al.* **Handbook of markov chain monte carlo**. [S.l.: s.n.]: CRC press, 2011.

CAI, H. Assessing and modeling community resilience to coastal hazards using a bayesian network. 2017.

CASELLA, G.; GEORGE, E. I. Explaining the gibbs sampler. **The American Statistician**, Taylor & Francis, v. 46, n. 3, p. 167–174, 1992.

CASSEL, J.-C.; VASCONCELOS, A. P. de. Importance of the ventral midline thalamus in driving hippocampal functions. **Progress in brain research**, Elsevier, v. 219, p. 145–161, 2015.

CHATURVEDI, I. *et al.* Bayesian network based extreme learning machine for subjectivity detection. **Journal of The Franklin Institute**, Elsevier, v. 355, n. 4, p. 1780–1797, 2018.

CHEN, J. *et al.* Towards interpretable clinical diagnosis with bayesian network ensembles stacked on entity-aware cnns. *In*: **Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics**. [S.l.: s.n.], 2020. p. 3143–3153.

CHEN, Z. *et al.* Bayesian filtering: From kalman filters to particle filters, and beyond. **Statistics**, v. 182, n. 1, p. 1–69, 2003.

- CHOWDHURY, F. A. *et al.* Localisation in focal epilepsy: a practical guide. **Practical Neurology**, BMJ Publishing Group Ltd, v. 21, n. 6, p. 481–491, 2021.
- CIARAMIDARO, A. *et al.* Multiple-brain connectivity during third party punishment: an eeg hyperscanning study. **Scientific Reports**, v. 8, 12 2018.
- COLMERS, P. L.; MAGUIRE, J. Network dysfunction in comorbid psychiatric illnesses and epilepsy. **Epilepsy Currents**, SAGE Publications Sage CA: Los Angeles, CA, v. 20, n. 4, p. 205–210, 2020.
- CONTALDI, C.; VAFAEE, F.; NELSON, P. C. Bayesian network hybrid learning using an elite-guided genetic algorithm. **Artificial Intelligence Review**, Springer, v. 52, n. 1, p. 245–272, 2019.
- COOPER, G. F. The computational complexity of probabilistic inference using bayesian belief networks. **Artificial intelligence**, Elsevier, v. 42, n. 2-3, p. 393–405, 1990.
- COOPER, G. F.; HERSKOVITS, E. A bayesian method for the induction of probabilistic networks from data. **Machine learning**, Springer, v. 9, n. 4, p. 309–347, 1992.
- COTA, V. R. *et al.* The epileptic amygdala: toward the development of a neural prosthesis by temporally coded electrical stimulation. **Journal of neuroscience research**, Wiley Online Library, v. 94, n. 6, p. 463–485, 2016.
- COTA, V. R. *et al.* Distinct patterns of electrical stimulation of the basolateral amygdala influence pentylentetrazole seizure outcome. **Epilepsy & Behavior**, Elsevier, v. 14, n. 1, p. 26–31, 2009.
- COTA, V. R. *et al.* Nonperiodic stimulation for the treatment of refractory epilepsy: Applications, mechanisms, and novel insights. **Epilepsy & Behavior**, Elsevier, p. 106609, 2019.
- COURTIOL, J. *et al.* Dynamical mechanisms of interictal resting-state functional connectivity in epilepsy. **Journal of Neuroscience**, Soc Neuroscience, v. 40, n. 29, p. 5572–5588, 2020.
- COVER, T. M.; THOMAS, J. A. **Elements of information theory**. [S.l.: s.n.]: John Wiley & Sons, 2012.
- COZMAN, F. G. *et al.* Generalizing variable elimination in bayesian networks. *In*: CITeseer. **Workshop on probabilistic reasoning in artificial intelligence**. [S.l.: s.n.], 2000. p. 27–32.
- DAGUM, P.; GALPER, A.; HORVITZ, E. Dynamic network models for forecasting. *In*: ELSEVIER. **Uncertainty in artificial intelligence**. [S.l.: s.n.], 1992. p. 41–48.
- DARBELLAY, G. A.; VAJDA, I. Estimation of the information by an adaptive partitioning of the observation space. **IEEE Transactions on Information Theory**, IEEE, v. 45, n. 4, p. 1315–1321, 1999.
- DAS, A.; MENON, V. Spatiotemporal integrity and spontaneous nonlinear dynamic properties of the salience network revealed by human intracranial electrophysiology: A multicohort replication. **Cerebral Cortex**, 2020.
- DENG, N. *et al.* Indoleamine-2, 3-dioxygenase 1 deficiency suppresses seizures in epilepsy. **Frontiers in Cellular Neuroscience**, Frontiers, v. 15, p. 28, 2021.

DEVINSKY, O. *et al.* Epilepsy. **Nature Reviews Disease Primers**, v. 4, 2018. Available at: <<https://www.nature.com/articles/nrdp201824#supplementary-information>>.

DHAMALA, M.; RANGARAJAN, G.; DING, M. Analyzing information flow in brain networks with nonparametric granger causality. **NeuroImage**, v. 41, p. 354–62, 07 2008.

DISEASES, O. M. de la Salud. Programme for N. *et al.* **Atlas: Epilepsy Care in the World**. [S.l.: s.n.]: Programme for Neurological Diseases and Neuroscience, Department of Mental Health and Substance Abuse, World Health Organization, 2005. (Academic Search Complete). ISBN 9789241563031.

DOBROW, R. P. **Introduction to Stochastic Processes with R**. [S.l.: s.n.]: John Wiley & Sons, 2016.

DZIAK, J. J. *et al.* Sensitivity and specificity of information criteria. **Briefings in Bioinformatics**, Oxford University Press, v. 21, n. 2, p. 553–565, 2020.

EELLS, J. *et al.* Comparative fos immunoreactivity in the brain after forebrain, brainstem, or combined seizures induced by electroshock, pentylenetetrazol, focally induced and audiogenic seizures in rats. **Neuroscience**, Elsevier, v. 123, n. 1, p. 279–292, 2004.

EL-GABY, M. *et al.* An emergent neural coactivity code for dynamic memory. **Nature neuroscience**, Nature Publishing Group, v. 24, n. 5, p. 694–704, 2021.

ELDAWLATLY, S. *et al.* On the use of dynamic bayesian networks in reconstructing functional neuronal networks from spike train ensembles. **Neural computation**, MIT Press, v. 22, n. 1, p. 158–189, 2010.

ELLENBROEK, B.; YOUN, J. Chapter 9 - schizophrenia. *In*: ELLENBROEK, B.; YOUN, J. (ed.). **Gene-Environment Interactions in Psychiatry**. San Diego: Academic Press, 2016. p. 289 – 322. ISBN 978-0-12-801657-2. Available at: <<http://www.sciencedirect.com/science/article/pii/B9780128016572000094>>.

ENDO, W. *et al.* Delayed mutual information infers patterns of synaptic connectivity in a proprioceptive neural network. **Journal of computational neuroscience**, Springer, v. 38, n. 2, p. 427–438, 2015.

ESCH, R. J. van *et al.* A bayesian method for inference of effective connectivity in brain networks for detecting the mozart effect. **Computers in Biology and Medicine**, Elsevier, v. 127, p. 104055, 2020.

FEIGIN, V. L. *et al.* Global, regional, and national burden of neurological disorders, 1990–2016: a systematic analysis for the global burden of disease study 2016. **The Lancet Neurology**, Elsevier, 2019.

FIEST, K. M. *et al.* Prevalence and incidence of epilepsy. **Neurology**, Wolters Kluwer Health, Inc. on behalf of the American Academy of Neurology, v. 88, n. 3, p. 296–303, 2017. ISSN 0028-3878. Available at: <<http://n.neurology.org/content/88/3/296>>.

FISHER, R. S. *et al.* Ilae official report: A practical clinical definition of epilepsy. **Epilepsia**, v. 55, n. 4, p. 475–482, 2014. Available at: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/epi.12550>>.



- FISHER, R. S. *et al.* Epileptic seizures and epilepsy: Definitions proposed by the international league against epilepsy (ilae) and the international bureau for epilepsy (ibe). **Epilepsia**, v. 46, n. 4, p. 470–472, 2005. Available at: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0013-9580.2005.66104.x>>.
- FOIT, N. A.; BERNASCONI, A.; BERNASCONI, N. Functional networks in epilepsy presurgical evaluation. **Neurosurgery Clinics**, Elsevier, v. 31, n. 3, p. 395–405, 2020.
- FORCELLI, P. A. Applications of optogenetic and chemogenetic methods to seizure circuits: Where to go next? **Journal of neuroscience research**, Wiley Online Library, v. 95, n. 12, p. 2345–2356, 2017.
- GEMAN, S.; GEMAN, D. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *In: Readings in computer vision*. [S.l.: s.n.]: Elsevier, 1987. p. 564–584.
- GENCAGA, D.; KNUTH, K. H.; ROSSOW, W. B. A recipe for the estimation of information flow in a dynamical system. **Entropy**, Multidisciplinary Digital Publishing Institute, v. 17, n. 1, p. 438–470, 2015.
- GHAHRAMANI, Z. Learning dynamic bayesian networks. *In: SPRINGER. International School on Neural Networks, Initiated by IIASS and EMFCSC*. [S.l.: s.n.], 1997. p. 168–197.
- GIL, F. *et al.* Beyond the epileptic focus: functional epileptic networks in focal epilepsy. **Cerebral Cortex**, Oxford University Press, v. 30, n. 4, p. 2338–2357, 2020.
- GLOMB, K. *et al.* Computational models in electroencephalography. **Brain topography**, Springer, p. 1–20, 2021.
- GOEL, M. K.; KHANNA, P.; KISHORE, J. Understanding survival analysis: Kaplan-meier estimate. **International journal of Ayurveda research**, Wolters Kluwer–Medknow Publications, v. 1, n. 4, p. 274, 2010.
- GONG, C.; STOFFER, D. S. A note on efficient fitting of stochastic volatility models. **Journal of Time Series Analysis**, Wiley Online Library, v. 42, n. 2, p. 186–200, 2021.
- GRIBKOVA, E. D.; IBRAHIM, B. A.; LLANO, D. A. A novel mutual information estimator to measure spike train correlations in a model thalamocortical network. **Journal of neurophysiology**, American Physiological Society Bethesda, MD, v. 120, n. 6, p. 2730–2744, 2018.
- GROSS, C. Epilepsy research now in 3d: Harnessing the power of brain organoids in epilepsy. **Epilepsy Currents**, SAGE Publications Sage CA: Los Angeles, CA, p. 15357597211070391, 2022.
- GROSS, T. J. *et al.* Dependence between cognitive impairment and metabolic syndrome applied to a brazilian elderly dataset. **Artificial intelligence in medicine**, Elsevier, v. 90, p. 53–60, 2018.
- GROSS, T. J. *et al.* An analytical threshold for combining bayesian networks. **Knowledge-Based Systems**, Elsevier, v. 175, p. 36–49, 2019.
- HANKIN, R. K. *et al.* A generalization of the dirichlet distribution. **Journal of Statistical Software**, v. 33, n. 11, p. 1–18, 2010.

HARVEY, A.; RUIZ, E.; SHEPHARD, N. Multivariate stochastic variance models. **The Review of Economic Studies**, Wiley-Blackwell, v. 61, n. 2, p. 247–264, 1994.

HASTINGS, W. K. Monte carlo sampling methods using markov chains and their applications. **Biometrika**, Oxford University Press, v. 57, n. 1, p. 97–109, 1970.

HECKERMAN, D. Learning bayesian networks: The combination of knowledge and statical data. **Proceedings of Uncertainty in Artificial Intelligence, 1994**, 1994.

HECKERMAN, D.; WELLMAN, M. P. Bayesian networks. **Communications of the ACM**, Association for Computing Machinery, Inc., v. 38, n. 3, p. 27–31, 1995.

HELD, L.; BOVÉ, D. S. Model selection. *In: Likelihood and Bayesian Inference.* [S.l.: s.n.]: Springer, 2020. p. 221–245.

HEYSIEATTALAB, S.; SADEGHI, L. Dynamic structural neuroplasticity during and after epileptogenesis in a pilocarpine rat model of epilepsy. **Acta Epileptologica**, Springer, v. 3, n. 1, p. 1–9, 2021.

HUANG, C.; DARWICHE, A. Inference in belief networks: A procedural guide. **International journal of approximate reasoning**, Elsevier, v. 15, n. 3, p. 225–263, 1996.

IBRAHIM, F. *et al.* A statistical framework for eeg channel selection and seizure prediction on mobile. **International Journal of Speech Technology**, v. 22, n. 1, p. 191–203, Mar 2019. ISSN 1572-8110. Available at: <<https://doi.org/10.1007/s10772-018-09565-7>>.

JIA, Z. *et al.* Refined nonuniform embedding for coupling detection in multivariate time series. **Physical Review E**, APS, v. 101, n. 6, p. 062113, 2020.

JIANG, P.; KUMAR, P. Bundled causal history interaction. **Entropy**, Multidisciplinary Digital Publishing Institute, v. 22, n. 3, p. 360, 2020.

JIANG, W.; CAO, Y.; DENG, X. A novel z-network model based on bayesian network and z-number. **IEEE Transactions on Fuzzy Systems**, IEEE, v. 28, n. 8, p. 1585–1599, 2019.

JIRUSKA, P. *et al.* Synchronization and desynchronization in epilepsy: controversies and hypotheses. **The Journal of Physiology**, v. 591, n. 4, p. 787–797, 2012. Available at: <<https://physoc.onlinelibrary.wiley.com/doi/abs/10.1113/jphysiol.2012.239590>>.

KAILA, K. *et al.* Gaba actions and ionic plasticity in epilepsy. **Current Opinion in Neurobiology**, v. 26, p. 34 – 41, 2014. ISSN 0959-4388. SI: Inhibition: Synapses, Neurons and Circuits. Available at: <<http://www.sciencedirect.com/science/article/pii/S0959438813002146>>.

KARTSONAKI, C. Survival analysis. **Diagnostic Histopathology**, Elsevier, v. 22, n. 7, p. 263–270, 2016.

KIPINSKI, L.; KORDECKI, W. Time series analysis of trial-to-trial variability of meg power spectrum during rest state, unattended listening and frequency-modulated tones classification. **bioRxiv**, Cold Spring Harbor Laboratory, 2021.

KLINZING, J. G.; NIETHARD, N.; BORN, J. Mechanisms of systems memory consolidation during sleep. **Nature neuroscience**, Nature Publishing Group, v. 22, n. 10, p. 1598–1610, 2019.

KOH, K. L.; AHAD, N. A. Normality for non-normal distributions. **Journal of Science and Mathematics Letters**, v. 8, n. 2, p. 51–60, 2020.

KOLLER, D.; FRIEDMAN, N. **Probabilistic graphical models: principles and techniques**. [S.l.: s.n.]: MIT press, 2009.

KOŘENEK, J.; HLINKA, J. Causal network discovery by iterative conditioning: Comparison of algorithms. **Chaos: An Interdisciplinary Journal of Nonlinear Science**, AIP Publishing LLC, v. 30, n. 1, p. 013117, 2020.

KOUZIOKAS, G. N. A new w-svm kernel combining pso-neural network transformed vector and bayesian optimized svm in gdp forecasting. **Engineering Applications of Artificial Intelligence**, Elsevier, v. 92, p. 103650, 2020.

KRUBITZER, L.; CAMPI, K. L.; COOKE, D. F. All rodents are not the same: a modern synthesis of cortical organization. **Brain, behavior and evolution**, Karger Publishers, v. 78, n. 1, p. 51–93, 2011.

LARRANAGA, P. *et al.* Structure learning of bayesian networks by genetic algorithms: A performance analysis of control parameters. **IEEE transactions on pattern analysis and machine intelligence**, IEEE, v. 18, n. 9, p. 912–926, 1996.

LEE, S.-h. *et al.* Multi-risk assessment of heat waves under intensifying climate change using bayesian networks. **International Journal of Disaster Risk Reduction**, Elsevier, v. 50, p. 101704, 2020.

LEHNERTZ, K. *et al.* Synchronization phenomena in human epileptic brain networks. **Journal of Neuroscience Methods**, v. 183, n. 1, p. 42 – 48, 2009. ISSN 0165-0270. BrainModes: A Principled Approach to Modeling and Measuring Large-Scale Neuronal Activity. Available at: <<http://www.sciencedirect.com/science/article/pii/S016502700900274X>>.

LEVIN, D. A.; PERES, Y. **Markov chains and mixing times**. [S.l.: s.n.]: American Mathematical Soc., 2017. v. 107.

LEÃO, T. *et al.* Learning dynamic bayesian networks from time-dependent and time-independent data: Unraveling disease progression in amyotrophic lateral sclerosis. **Journal of Biomedical Informatics**, v. 117, p. 103730, 2021. ISSN 1532-0464. Available at: <<https://www.sciencedirect.com/science/article/pii/S1532046421000599>>.

LI, X. *et al.* Health state prediction and performance evaluation of belt conveyor based on dynamic bayesian network in underground mining. **Shock and Vibration**, Hindawi, v. 2021, 2021.

LIGNANI, G.; BALDELLI, P.; MARRA, V. Homeostatic plasticity in epilepsy. **Frontiers in Cellular Neuroscience**, Frontiers, v. 14, p. 197, 2020.

LOSI, G. *et al.* Cellular and molecular mechanisms of new onset seizure generation. **Aging clinical and experimental research**, Springer, p. 1–4, 2019.

MANSOURI, M. *et al.* Heidegger: Interpretable temporal causal discovery. *In: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. [S.l.: s.n.], 2020. p. 1688–1696.

MCCANN, R. K.; MARCOT, B. G.; ELLIS, R. Bayesian belief networks: applications in ecology and natural resource management. **Canadian Journal of Forest Research**, NRC Research Press Ottawa, Canada, v. 36, n. 12, p. 3053–3062, 2006.

MCCAUSLAND, W.; MILLER, S.; PELLETIER, D. Multivariate stochastic volatility using the hessian method. **Econometrics and Statistics**, Elsevier, v. 17, p. 76–94, 2021.

MCEWAN, B. Sampling and validity. **Annals of the International Communication Association**, Taylor & Francis, p. 1–13, 2020.

MEADOR, K. J. *et al.* Disparities in nih funding for epilepsy research. **Neurology**, Wolters Kluwer Health, Inc. on behalf of the American Academy of Neurology, v. 77, n. 13, p. 1305–1307, 2011. ISSN 0028-3878. Available at: <<http://n.neurology.org/content/77/13/1305>>.

MEDEIROS, D. de C.; MORAES, M. F. D. Focus on desynchronization rather than excitability: A new strategy for intraencephalic electrical stimulation. **Epilepsy & Behavior**, v. 38, p. 32 – 36, 2014. ISSN 1525-5050. SI: NEWroscience 2013. Available at: <<http://www.sciencedirect.com/science/article/pii/S1525505013006653>>.

METROPOLIS, N. *et al.* Equation of state calculations by fast computing machines. **The journal of chemical physics**, AIP, v. 21, n. 6, p. 1087–1092, 1953.

MICHIELS, M.; LARRAÑAGA, P.; BIELZA, C. Bayesuites: An open web framework for massive bayesian networks focused on neuroscience. **Neurocomputing**, Elsevier, v. 428, p. 166–181, 2021.

MIHÁLY, I. *et al.* Amygdala low-frequency stimulation reduces pathological phase-amplitude coupling in the pilocarpine model of epilepsy. **Brain sciences**, Multidisciplinary Digital Publishing Institute, v. 10, n. 11, p. 856, 2020.

MOE, S. J.; HAANDE, S.; COUTURE, R.-M. Climate change, cyanobacteria blooms and ecological status of lakes: a bayesian network approach. **Ecological Modelling**, Elsevier, v. 337, p. 330–347, 2016.

MORAES, M. F. D. *et al.* Epilepsy as a dynamical system, a most needed paradigm shift in epileptology. **Epilepsy & Behavior**, Elsevier, p. 106838, 2019.

MOREIRA, C. *et al.* Linda-bn: An interpretable probabilistic approach for demystifying black-box predictive models. **Decision Support Systems**, Elsevier, p. 113561, 2021.

MUKHERJEE, S.; ASNANI, H.; KANNAN, S. Ccmi: Classifier based conditional mutual information estimation. *In: PMLR. Uncertainty in Artificial Intelligence*. [S.l.: s.n.], 2020. p. 1083–1093.

MURPHY, K. **Pearl’s algorithm and multiplexer nodes**. [S.l.], 1999.

MURPHY, K. P. **Dynamic bayesian networks: representation, inference and learning**. 2002. Dissertação (Mestrado) — University of California, Berkeley Berkeley, CA, 2002.

MYERS, M. H. *et al.* Seizure localization using eeg analytical signals. **Clinical Neurophysiology**, Elsevier, v. 131, n. 9, p. 2131–2139, 2020.

NASEREDDIN, H. H.; OMARI, A. A. R. Classification techniques for automatic speech recognition (asr) algorithms used with real time speech translation. *In: IEEE. 2017 Computing Conference*. [S.l.: s.n.], 2017. p. 200–207.

NEAPOLITAN, R. E. *et al.* **Learning bayesian networks**. [S.l.: s.n.]: Pearson Prentice Hall Upper Saddle River, NJ, 2004. v. 38.

NELSON, C. J.; BONNER, S. Neuronal graphs: A graph theory primer for microscopic, functional networks of neurons recorded by calcium imaging. **Frontiers in Neural Circuits**, v. 15, p. 38, 2021. ISSN 1662-5110. Available at: <<https://www.frontiersin.org/article/10.3389/fncir.2021.662882>>.

NIRIAYO, Y. L. *et al.* Medication belief and adherence among patients with epilepsy. **Behavioural neurology**, Hindawi, v. 2019, 2019.

NOBUKAWA, S. *et al.* Classification methods based on complexity and synchronization of electroencephalography signals in alzheimer's disease. **Frontiers in psychiatry**, Frontiers, v. 11, p. 255, 2020.

NORONHA, A. L. A. *et al.* Prevalence and pattern of epilepsy treatment in different socioeconomic classes in brazil. **Epilepsia**, v. 48, n. 5, p. 880–885, 2007. Available at: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1528-1167.2006.00974.x>>.

NOWAK, A. *et al.* Functional synchronization: The emergence of coordinated activity in human systems. **Frontiers in psychology**, Frontiers, v. 8, p. 945, 2017.

OGASAWARA, H. An asymptotic equivalence of the cross-data and predictive estimators. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 49, n. 3, p. 755–768, 2020.

OGAWA, S.; PARHAR, I. S. Heterogeneity in gnRH and kisspeptin neurons and their significance in vertebrate reproductive biology. **Frontiers in neuroendocrinology**, Elsevier, v. 64, p. 100963, 2022.

OLAMAT, A. E.; AKAN, A. Synchronization analysis of epilepsy data using global field synchronization. In: **2017 25th Signal Processing and Communications Applications Conference (SIU)**. [S.l.: s.n.], 2017. p. 1–4.

OLDHAM, S.; BALL, G.; FORNITO, A. Early and late development of hub connectivity in the human brain. **Current opinion in psychology**, Elsevier, v. 44, p. 321–329, 2022.

OLIVEIRA, J. C. de. Investigação do sincronismo intra e inter substrato como mecanismo da supressão de crises por estratégias espaço-temporais de estimulação elétrica. Universidade Federal de São Joao Del-Rei, 2017.

OLIVEIRA, J. C. de *et al.* Temporally unstructured electrical stimulation to the amygdala suppresses behavioral chronic seizures of the pilocarpine animal model. **Epilepsy & Behavior**, Elsevier, v. 36, p. 159–164, 2014.

OLIVEIRA, J. D. *et al.* Asynchronous, bilateral, and biphasic temporally unstructured electrical stimulation of amygdalae enhances the suppression of pentylenetetrazole-induced seizures in rats. **Epilepsy research**, Elsevier, v. 146, p. 1–8, 2018.

OLIVEIRA, J. de *et al.* Seizure suppression by asynchronous non-periodic electrical stimulation of the amygdala is partially mediated by indirect desynchronization from nucleus accumbens. **Epilepsy research**, Elsevier, v. 154, p. 107–115, 2019.

ORGANIZATION, W. H. *et al.* **Atlas: country resources for neurological disorders 2004**. Geneva: World Health Organization; 2004. 2017.

OSTENDORF, A. P.; GEDELA, S. Effect of epilepsy on families, communities, and society. *In: ELSEVIER. Seminars in pediatric neurology*. [S.l.: s.n.], 2017. v. 24, p. 340–347.

PARIZ, A. *et al.* Transmission delays and frequency detuning can regulate information flow between brain regions. **PLoS computational biology**, Public Library of Science San Francisco, CA USA, v. 17, n. 4, p. e1008129, 2021.

PARMEGGIANI, P. L.; LENZI, P.; AZZARONI, A. Transfer of the hippocampal output by the anterior thalamic nuclei. **Brain research**, Elsevier, v. 67, n. 2, p. 269–278, 1974.

PARVIZI, J. *et al.* Detecting silent seizures by their sound. **Epilepsia**, Wiley Online Library, v. 59, n. 4, p. 877–884, 2018.

PAXINOS, G.; WATSON, C. **The Rat Brain in Stereotaxic Coordinates 7th edition**. [S.l.: s.n.]: Elsevier, 2013.

PEARL, J. Belief networks revisited. **Artificial intelligence in perspective**, p. 49–56, 1994.

PEARL, J. **The causal foundations of structural equation modeling**. [S.l.], 2012.

PETKOSKI, S.; JIRSA, V. K. Transmission time delays organize the brain network synchronization. **Philosophical Transactions of the Royal Society A**, The Royal Society Publishing, v. 377, n. 2153, p. 20180132, 2019.

PHAN, T. D. *et al.* Multivariate stochastic volatility modeling of neural data. **eLife**, eLife Sciences Publications, Ltd, v. 8, 2019.

PUGA, J. L.; KRZYWINSKI, M.; ALTMAN, N. Bayes' theorem. **Nature Methods**, v. 12, p. 277 – 278, 2015.

QIN, L.; WANG, Y. Nonparametric spectral analysis with applications to seizure characterization using eeg time series. **The Annals of Applied Statistics**, Institute of Mathematical Statistics, v. 2, n. 4, p. 1432–1451, 2008. ISSN 19326157. Available at: <<http://www.jstor.org/stable/30245142>>.

RAMOS, D. *et al.* Dynamic bayesian networks for temporal prediction of chemical radioisotope levels in nuclear power plant reactors. **Chemometrics and Intelligent Laboratory Systems**, Elsevier, p. 104327, 2021.

REIS, A. S. *et al.* Bursting synchronization in neuronal assemblies of scale-free networks. **Chaos, Solitons & Fractals**, Elsevier, v. 142, p. 110395, 2021.

ROBERT, C.; CASELLA, G. A short history of markov chain monte carlo: Subjective recollections from incomplete data. **Statistical Science**, JSTOR, p. 102–115, 2011.

ROBINSON, J. W.; HARTEMINK, A. J. Learning non-stationary dynamic bayesian networks. **Journal of Machine Learning Research**, v. 11, n. 118, p. 3647–3680, 2010. Available at: <<http://jmlr.org/papers/v11/robinson10a.html>>.

ROBINSON, R. W. Counting unlabeled acyclic digraphs. *In: Combinatorial mathematics V*. [S.l.: s.n.]: Springer, 1977. p. 28–43.

ROSSELLO, X.; GONZÁLEZ-DEL-HOYO, M. Survival analyses in cardiovascular research, part i: the essentials. **Revista Española de Cardiología (English Edition)**, Elsevier, 2021.

- ROSSI, R. *et al.* Upgrading model selection criteria with goodness of fit tests for practical applications. **Entropy**, Multidisciplinary Digital Publishing Institute, v. 22, n. 4, p. 447, 2020.
- RUIZ-PÉREZ, I. *et al.* A bayesian network approach to study the relationships between several neuromuscular performance measures and dynamic postural control in futsal players. **PLoS one**, Public Library of Science San Francisco, CA USA, v. 14, n. 7, p. e0220065, 2019.
- RUSSELL, S.; NORVIG, P. *Artificial intelligence: a modern approach*. 2002.
- SAMOKHINA, E.; SAMOKHIN, A. Neuropathological profile of the pentylenetetrazol (ptz) kindling model. **International Journal of Neuroscience**, Taylor & Francis, v. 128, n. 11, p. 1086–1096, 2018.
- SAMPAIO, L. P. *et al.* Prevalence of epilepsy in children from a brazilian area of high deprivation. **Pediatric Neurology**, v. 42, n. 2, p. 111 – 117, 2010. ISSN 0887-8994. Available at: <<http://www.sciencedirect.com/science/article/pii/S0887899409004342>>.
- SCOTT, R. C. Brains, complex systems and therapeutic opportunities in epilepsy. **Seizure**, Elsevier, 2021.
- SCUTARI, M.; NAGARAJAN, R. Identifying significant edges in graphical models of molecular networks. **Artificial intelligence in medicine**, Elsevier, v. 57, n. 3, p. 207–217, 2013.
- SHERMAN, S. M.; GUILLERY, R. W. **Exploring the thalamus and its role in cortical function**. [S.l.: s.n.]: MIT press, 2006.
- SHRESTHA, S.; SHRESTHA, R. D.; THAPA, B. Implementing neural network and multi resolution analysis in eeg signal for early detection of epilepsy. **SCITECH Nepal**, v. 14, n. 1, p. 8–16, 2019.
- SIGMAN, K. **Time-Reversible Markov chains**. 2009. [Http://www.columbia.edu/~ks20/stochastic-I/stochastic-I-Time-Reversibility.pdf](http://www.columbia.edu/~ks20/stochastic-I/stochastic-I-Time-Reversibility.pdf). Accessed: 2019-05-06.
- SIGTERMANS, D. Towards a framework for observational causality from time series: when shannon meets turing. **Entropy**, Multidisciplinary Digital Publishing Institute, v. 22, n. 4, p. 426, 2020.
- SIP, V. *et al.* Data-driven method to infer the seizure propagation patterns in an epileptic brain from intracranial electroencephalography. **PLoS computational biology**, Public Library of Science San Francisco, CA USA, v. 17, n. 2, p. e1008689, 2021.
- SMITH, S. M. *et al.* Network modelling methods for fmri. **Neuroimage**, Elsevier, v. 54, n. 2, p. 875–891, 2011.
- SNYDER, D. B. *et al.* Electroencephalography resting-state networks in people with stroke. **Brain and behavior**, Wiley Online Library, v. 11, n. 5, p. e02097, 2021.
- SOUZA, P. *et al.* Knowledge about epilepsy in university health students: A multicenter study. **Epilepsy & Behavior**, v. 79, p. 112 – 116, 2018. ISSN 1525-5050. Available at: <<http://www.sciencedirect.com/science/article/pii/S1525505017304754>>.
- SPICIARICH, M. C. *et al.* Global health and epilepsy: Update and future directions. **Current neurology and neuroscience reports**, Springer, v. 19, n. 6, p. 30, 2019.

STALEY, K. Molecular mechanisms of epilepsy. **Nature Neuroscience**, v. 18, p. 367–372, 2015.

STANLEY, R. P. Acyclic orientations of graphs. **Discrete Mathematics**, Elsevier, v. 5, n. 2, p. 171–178, 1973.

STEPHAN, K.; FRISTON, K. Functional connectivity. *In*: SQUIRE, L. R. (ed.). **Encyclopedia of Neuroscience**. Oxford: Academic Press, 2009. p. 391 – 397. ISBN 978-0-08-045046-9. Available at: <<http://www.sciencedirect.com/science/article/pii/B9780080450469003089>>.

STOJANOVIĆ, O.; KUHLMANN, L.; PIPA, G. Predicting epileptic seizures using nonnegative matrix factorization. **PLoS one**, Public Library of Science San Francisco, CA USA, v. 15, n. 2, p. e0228025, 2020.

STRENG, M. L.; KROOK-MAGNUSON, E. The cerebellum and epilepsy. **Epilepsy & Behavior**, Elsevier, p. 106909, 2020.

TAKANO, T. Self-injury as a predominant challenging behavior in epilepsy: A study in a residential facility for profoundly disabled patients. **Research in developmental disabilities**, Elsevier, v. 120, p. 104149, 2022.

TAKEUCHI, Y. *et al.* The medial septum as a potential target for treating brain disorders associated with oscillopathies. **Frontiers in Neural Circuits**, 2021.

THEODONI, P. *et al.* Structural attributes and principles of the neocortical connectome in the marmoset monkey. **Cerebral Cortex**, Oxford University Press, v. 32, n. 1, p. 15–28, 2022.

THIVIERGE, J.-P.; CISEK, P. Nonperiodic synchronization in heterogeneous networks of spiking neurons. **Journal of Neuroscience**, Soc Neuroscience, v. 28, n. 32, p. 7968–7978, 2008.

TRACY, J. I. *et al.* Computational support, not primacy, distinguishes compensatory memory reorganization in epilepsy. **Brain Communications**, Oxford University Press, v. 3, n. 2, p. fcab025, 2021.

TSAMARDINOS, I.; BROWN, L. E.; ALIFERIS, C. F. The max-min hill-climbing bayesian network structure learning algorithm. **Machine learning**, Springer, v. 65, n. 1, p. 31–78, 2006.

TSUKAHARA, V. H. *et al.* Pdc-mi method for eeg functional connectivity analysis. *In*: SPRINGER. **International Joint Conference on Biomedical Engineering Systems and Technologies**. [S.l.: s.n.], 2020. p. 304–328.

TSUKAHARA, V. H. B. *et al.* Delayed mutual information to develop functional analysis on epileptic signals. *In*: **BIOSIGNALS**. [S.l.: s.n.], 2020. p. 89–97.

TSUKAHARA, V. H. B. *et al.* Data-driven network dynamical model of rat brains during acute ictogenesis. **Frontiers in Neural Circuits**, v. 16, 2022. ISSN 1662-5110. Available at: <<https://www.frontiersin.org/articles/10.3389/fncir.2022.747910>>.

TURKHEIMER, F. E. *et al.* A complex systems perspective on neuroimaging studies of behavior and its disorders. **The Neuroscientist**, Sage Publications Sage CA: Los Angeles, CA, 2021.

VAROTTO, G. *et al.* Network characteristics in benign epilepsy with centro-temporal spikes patients indicating defective connectivity during spindle sleep: A partial directed coherence study of eeg signals. **Clinical Neurophysiology**, v. 129, n. 11, p. 2372 – 2379, 2018. ISSN 1388-2457. Available at: <<http://www.sciencedirect.com/science/article/pii/S138824571831229X>>.



- VELISEK, L. *et al.* Pentylenetetrazol-induced seizures in rats: an ontogenetic study. **Naunyn-Schmiedeberg's archives of pharmacology**, Springer, v. 346, n. 5, p. 588–591, 1992.
- VILLANUEVA, E.; MACIEL, C. D. Efficient methods for learning bayesian network super-structures. **Neurocomputing**, Elsevier, v. 123, p. 3–12, 2014.
- VLIS, T. A. B. V. D. *et al.* Deep brain stimulation of the anterior nucleus of the thalamus for drug-resistant epilepsy. **Neurosurgical review**, Springer, v. 42, n. 2, p. 287–296, 2019.
- WANG, J. *et al.* Survnet: A novel deep neural network for lung cancer survival analysis with missing values. **Frontiers in Oncology**, Frontiers, v. 10, p. 3128, 2021.
- WANG, J. J.; CHAN, J. S. Stochastic modelling of volatility and inter-relationships in the australian electricity markets. **Communications in Statistics-Simulation and Computation**, Taylor & Francis, p. 1–20, 2021.
- WANG, Y. *et al.* Interictal intracranial electroencephalography for predicting surgical success: The importance of space and time. **Epilepsia**, Wiley Online Library, 2020.
- WEISS, S. A. *et al.* Interneurons and principal cell firing in human limbic areas at focal seizure onset. **Neurobiology of Disease**, v. 124, p. 183 – 188, 2019. ISSN 0969-9961. Available at: <<http://www.sciencedirect.com/science/article/pii/S096999611830682X>>.
- WHO. **Epilepsy: a public health imperative**. Geneva: World Health Organization, 2019.
- XIANG, J. *et al.* Kurtosis and skewness of high-frequency brain signals are altered in paediatric epilepsy. **Brain communications**, Oxford University Press, v. 2, n. 1, p. fcaa036, 2020.
- YANG, C. *et al.* Epilepsy as a dynamical disorder orchestrated by epileptogenic zone: a review. **Nonlinear Dynamics**, Springer, p. 1–16, 2021.
- YANG, H. *et al.* Mapping the neural circuits responding to deep brain stimulation of the anterior nucleus of the thalamus in the rat brain. **Epilepsy Research**, Elsevier, v. 187, p. 107027, 2022.
- ZHAN, H. *et al.* Correlation and survival analysis of distant metastasis site and prognosis in patients with hepatocellular carcinoma. **Frontiers in Oncology**, Frontiers, v. 11, p. 1646, 2021.
- ZHAO, C. *et al.* Importance of brain alterations in spinal cord injury. **Science Progress**, SAGE Publications Sage UK: London, England, v. 104, n. 3, p. 00368504211031117, 2021.
- ZHONG, G.-Y. *et al.* The time delay restraining the herd behavior with bayesian approach. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 507, p. 335–346, 2018.
- ZHOU, S.; LIN, W. Eliminating synchronization of coupled neurons adaptively by using feedback coupling with heterogeneous delays. **Chaos: An Interdisciplinary Journal of Nonlinear Science**, AIP Publishing LLC, v. 31, n. 2, p. 023114, 2021.
- ZHUANG, K. *et al.* Connectome-based evidence for creative thinking as an emergent property of ordinary cognitive operations. **NeuroImage**, Elsevier, v. 227, p. 117632, 2021.
- ZWEIG, G. Bayesian network structures and inference techniques for automatic speech recognition. **Computer Speech & Language**, Elsevier, v. 17, n. 2-3, p. 173–193, 2003.
- ZWEIG, G.; RUSSELL, S. Speech recognition with dynamic bayesian networks. University of California, Berkeley, 1998.



## **APPENDIX**



## APPENDIX A – DISSEMINATION ACTIVITIES

The following papers have been published as a result of this doctoral research:

- Journals:
  - Published: Victor Hugo Batista Tsukahara, Jasiara Carla de Oliveira, Vinicius Rosa Cota, Carlos Dias Maciel, Survival analysis of nonperiodic stimulation (NPS) performance, **Epilepsy & Behavior**, Elsevier, v.122, 2021.
  - Published: Victor Hugo Batista Tsukahara, Jordão Natal de Oliveira Junior, Vitor Bruno de Oliveira Barth, Jasiara Carla de Oliveira, Vinicius Rosa Cota, Carlos Dias Maciel, Data-Driven Network Dynamical Model of Rat Brain During Acute Ictogenesis, **Frontiers in Neural Circuits**, Frontiers, 2022.
- Chapters:
  - Victor Hugo Batista Tsukahara, Pedro Virgílio Basílio Jeronymo, Jasiara Carla de Oliveira, Vinicius Rosa Cota, Carlos Dias Maciel, PDC-MI method for EEG Functional Connectivity Analysis. **13th International conference on bio-inspired systems and signal processing (Biosignals)**, Communications in Computer and Information Science (CCIS), Springer International Publishing, 2021.
  - Talysson Manoel de Oliveira Santos, Victor Hugo Batista Tsukahara, Jasiara Carla de Oliveira, Vinicius Rosa Cota, Carlos Dias Maciel, Graph Model Evolution During Epileptic Seizures: linear model approach. **II Latin American Workshop on Computational Neuroscience (LAWCN)**, Communications in Computer and Information Science (CCIS), Springer International Publishing, 2019.
- Conference:
  - Victor Hugo Batista Tsukahara, Pedro Virgílio Basílio Jeronymo, Jasiara Carla de Oliveira, Vinicius Rosa Cota, Carlos Dias Maciel, Delayed Mutual Information to Develop Functional Analysis on Epileptic Signals. **13th International conference on bio-inspired systems and signal processing**, 2020.

The papers below are not related to this Thesis, but just collaborations:

- Journals:
  - Published: Jordão Natal de Oliveira Junior, Ivonete Ávila, Victor Hugo Batista Tsukahara, Marcelo Pinheiro, Carlos Dias Maciel, Entropy: from thermodynamics to signal processing, **Entropy**, MDPI, 2021.

- Conference:

- Rafael Rodrigues Mendes Ribeiro, Talysson Manoel de Oliveira Santos, Tadeu Junior Gross, Victor Hugo Batista Tsukahara, Carlos Dias Maciel, Applying PDC for the recognition of firearm's calibre. **Simpósio Brasileiro de Automação Inteligente**, 2019.
- Mateus Camargo Pedrino, Rafael Caracciolo Arone, Victor Hugo Batista Tsukahara, Carlos Dias Maciel, Neural network classification, coherence and power spectrum analysis with stress database. **II Latin American Workshop on Computational Neuroscience**, 2019.
- Vitor Bruno de Oliveira Barth, Jordão Natal de Oliveira Junior, Victor Hugo Batista Tsukahara, Fábio Viadanna Serrão, Ana Flávia dos Santos, Carlos Dias Maciel, Information theoretic analysis of EMG and kinematic data among runners with patellofemoral pain. **Congresso Brasileiro de Automação**, 2022.

## APPENDIX B – MARKOV CHAIN MONTE CARLO (MCMC)

Supposing we are performing a simple statistical study related to an unbiased coin toss. The aim is to find what is the probability of flipping the coin resulting in heads or tails. In an initial analysis, we could think that there are only two possible states and therefore we can infer 50 % probability for each state. For such a situation, an analytical probability calculation would be possible. However, another question may arise: would it be possible to use another method to perform such a calculation?

Let us consider the same example in another study. A researcher initially tossed the coin 100 times by calculating the total number of heads and tails. Verifying the frequency for each state, the researcher concluded that 47/100 were heads and 53/100 were tails. Curious about possible results, the study was repeated by tossing the coin 10.000 times and registering results. After checking, it was observed that probability was approximately 1/2, exactly as it has been analytically calculated. Therefore, it is possible to infer that the larger the number of repetitions, the closer the frequencies are to the analytical calculation results. This represents one of the most important theorems in probability calculation: the Law of Large Numbers.

The theorem may initially seem useless, once an analytical calculation is simpler. Nonetheless, in real-world situations such as in the volatilities of share prices or natural disasters probability of occurrence (like earthquakes and storms for instance), the Law of Large Numbers may be useful due to the difficulty or impossibility of performing an analytical calculation.

Therefore, for such situations, the use of sampling methods might be suitable for providing statistics with relative precision. Following the rationale for complex probability models, Markov Chains can be used to perform simulations (DOBROW, 2016). (ROBERT; CASELLA, 2011) states this as one of the ten most important algorithms developed in the 20th century. In fact, its development mainly allowed the evolution of Bayesian Inference (BROOKS *et al.*, 2011).

Initially, the chapter is going to present the main concepts for understanding Monte Carlo simulations and then the Monte Carlo Markov Chain is presented. Finally, the algorithm used in this work is going to be presented, i.e. the Hamiltonian.

### Markov Chain

To better understand the concept of Markov Chain, let us suppose that there is a board containing points numbered from 1 to 6 forming a circle (see Figure 21).

Now let us suppose that the player is in point number 4, just as in Figure 21. For each movement on the board, the player has to roll a six-sided dice numbered from 1 to 6. According to the number drawn, the player changes the point on board. Therefore, we can say that there is



Figure 21 – Example of Markov Chain board.

a space state  $\Xi = \{1, 2, 3, 4, 5, 6\}$  representing all possible states that a player has in the game. Now, let us define  $\Omega_t$  as the previous states of a player after 1 move. Then, according to the example, we can define  $\Omega_0 = 4$  (initial state). Assuming that the player rolled the dice twice and providing that the number 2 was obtained in the first movement and 3 in the second movement, we have the first three positions of the player:

$$\xi = \{\Omega_0, \Omega_1, \Omega_2\} = \{4, 6, 3\} \quad (\text{B.1})$$

Given the  $\xi$  history of movements which will be the player position  $\Omega_3$ . In that case this history will not contribute to infer the next payer position on board, however,  $\Omega_2 = 3$  is an important piece of information because the last point can provide an insight about the player's next position ( $v$ ), which may be:

$$v = \{4, 5, 6, 1, 2, 3\} \quad (\text{B.2})$$

For those positions and supposing that the dice is not biased, they have the same probability of occurrence, which means:

$$P(\Omega_3 = q | \Omega_0 = 4, \Omega_1 = 6, \Omega_2 = 3) = P(\Omega_3 | \Omega_2 = 3) = \frac{1}{6}, q = 4, 5, 6, 1, 2, 3 \quad (\text{B.3})$$

According to definitions provided until now, we can state that the future player's position on board is independent from movement history, except for the last state. Another finding is that the observed stochastic process is stationary, once a joint distribution of any subset of random variables is time invariant (COVER; THOMAS, 2012):



$$P(\Omega_1 = \omega_1, \dots, \Omega_c = \omega_c) = P(\Omega_{1+l} = \omega_1, \dots, \Omega_{c+l} = \omega_c) \quad (\text{B.4})$$

Given that the process is stochastic and the probability in any time depends only on the previous instant (equation B.3), we can define that the sequence  $\lambda$  is a Markov Chain (LEVIN; PERES, 2017). Given that probabilities of each state do not depend on time, it is homogeneous (DOBROW, 2016). Thus, we have:

$$P(\Omega_{c+1} = m | \Omega_c = \rho) = P(\Omega_1 = m | \Omega_0 = \rho), c \geq 0 \quad (\text{B.5})$$

Transition probabilities from one state to another can be organized in a matrix called Markov Matrix or Transition Matrix. Once space state  $\Phi$  contains  $r$  elements, its dimension is going to be  $r \times r$ . Then, given our example, we have the following matrix ( $m$  lines and  $\rho$  columns, dimension  $6 \times 6$ ):

$$\mathbf{S} = \begin{bmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{bmatrix}$$

The matrix is based on probabilities of state transitions, thus its values are non-negative and the sum of each line equals 1:

$$\sum_m P_{m\rho} = \sum_m P(\Omega_1 = m | \Omega_0 = \rho) = \sum_m \frac{P(\Omega_1 = m, \Omega_0 = \rho)}{P(\Omega_0 = \rho)} = \frac{P(\Omega_0 = \rho)}{P(\Omega_0 = \rho)} = 1 \quad (\text{B.6})$$

The aforementioned matrix having described properties is called a stochastic matrix (DOBROW, 2016).

Another important property of Monte Carlo Markov Chains is reversibility.  $\xi$  is the Markov Chain and  $\mathbf{S}$  is the transition states matrix. According to properties that have already been defined, the future state is independent of its history, given the present state. Therefore, based on this rationale, it is possible to state that the past is independent of the future, given the present state. Then, the process is time reversible because the result is still a stationary Markov Chain and transition probabilities would be calculated based on  $\xi$  and  $\mathbf{S}$ . Therefore, the following definition is set: a Markov Chain with transition matrix  $\mathbf{S}$  and stationary distribution  $\xi$  is reversible if the reverse Markov Chain contains the same direct distribution, which means  $P(r) = P$ ;  $P_{\rho m} = P_{m\rho}$  for all  $\rho$  and  $m$  pairs (SIGMAN, 2009). In equivalent manner, it can be defined as:

$$\xi_\rho P_{\rho m} = \xi_m P_{m\rho} \quad (\text{B.7})$$

### Ordinary Monte Carlo (OMC)

Otherwise known as independent and identically distributed Monte Carlo (i.i.d), it is a special case of MCMC where  $\Omega_1, \Omega_2 \dots \Omega_c$  are independent and identically distributed values and the Markov Chain is stationary and reversible. Supposing that we want to calculate the following expected value:

$$h = E[g(\Omega)] \quad (\text{B.8})$$

then  $g$  is a real value function in state space which cannot be calculated by any exact method. By performing a simulation of  $\Omega_1, \Omega_2 \dots \Omega_c$  assuming i.i.d. and the same distribution of  $\phi$ , it is possible to define:

$$\hat{h}_c = \frac{1}{c} \sum_{\rho=1}^c g(\Omega_\rho) \quad (\text{B.9})$$

using notation  $D_\rho = g(\Omega_\rho)$ ,  $D_\rho$  is i.i.d., mean  $h$  and variance  $\sigma^2 = \text{var}[g(\Omega)]$ . Using the central limit theorem, we can assume that, asymptotically, we have:

$$\hat{h}_c \approx N\left(h, \frac{\sigma^2}{c}\right) \quad (\text{B.10})$$

Also, according to the central limit theorem, we can estimate variance as:

$$\sigma_c^2 = \frac{1}{c} \sum_{\rho=1}^c (g(\phi_i) - \hat{h}_c)^2 \quad (\text{B.11})$$

which represents the empirical variance of  $D_\rho$ . Moreover, according to asymptotic probability theory, we have a 95% confidence interval (C.I.) for  $h$ :

$$C.I. = \hat{h}_c \pm 1.96 \frac{\hat{\sigma}_c}{\sqrt{c}} \quad (\text{B.12})$$

As observed, the method essentially uses elementary statistics to estimate the information of interest (BROOKS *et al.*, 2011).

### MCMC Theory

MCMC theory is similar to OMC theory, except for the fact that it assumes dependency on the Markov Chain by changing the standard error (BROOKS *et al.*, 2011). Using the example given in the previous section, the variance equation considering dependency is written as:

$$\sigma^2 = \text{var}[g(\Omega_\rho)] + 2 \sum_{o=1}^{\infty} \text{cov}\{g(\Omega_\rho), g(\Omega_{\rho+o})\} \quad (\text{B.13})$$

Stationary Markov Chains are not generally used in MCMC. This is due to the fact that, once a  $\Omega_1$  chain is simulated, the resulting distribution will not have any variance by definition, thus  $\Omega_2 \dots \Omega_c$  would be simulated and the OMC Theory could be applied. However, the Harris occurrence states that Central Limit Theorem remains for an initial distribution and transition probability, then it is possible to infer that it remains for other initial distributions and the same probability transitions. Therefore, it is possible to conclude that asymptotic variance is the same for all initial distributions (BROOKS *et al.*, 2011). Although the theoretic asymptotic variance comprises variances and covariances for stationary Markov Chains, it allows asymptotic variance for non-stationary Markov Chains having the same transition probability distribution (with different initial distributions) (BROOKS *et al.*, 2011).

### Metropolis-Hastings Algorithm

It is one of the most common MCMC samplers (DOBROW, 2016). It is a tribute to the algorithm's creator proposed by Metropolis (METROPOLIS *et al.*, 1953) and its scope was redefined by Hastings (HASTINGS, 1970). To understand how the algorithm works, let us define that there is a discrete probability distribution  $\gamma = (\chi_1, \chi_2 \dots \chi_c)$ . Using the stationary distribution  $\chi$ , the method then builds a Markov Chain  $\Omega_0, \Omega_1 \dots \Omega_c$  using the stationary distribution  $\gamma$ . Assuming that the user knows how to sample P, chain P is going to be used as a proposal chain, which in turn generates elements from a sequence that the algorithm will either accept or reject. To describe the transition mechanism for  $\Omega_0, \Omega_1 \dots \Omega_c$ , let us assume that the chain state is  $\rho$  during time  $c$ . Then, for the next step of the chain,  $\Omega_c + 1$ , it is determined by a process consisting of two steps:

1. Choosing the new state according to P; therefore, choosing  $m$  (state  $m$  is called proposed state) with probability  $P_{\rho m}$ ;
2. Deciding whether to accept or reject  $m$ :

$$a(\rho, m) = \frac{\lambda_m P_{m\rho}}{\lambda_\rho P_{\rho m}} \quad (\text{B.14})$$

where  $a$  is an acceptance function. Assuming  $a(\rho, m) \geq 1$ , then  $m$  is accepted as the next state of the chain. By assuming  $a(\rho, m) < 1$ , then  $j$  is accepted with probability  $a(\rho, m)$ . In the case of no acceptance of  $j$ , then  $i$  is kept as the next state of the chain. In summary, assuming that  $\Omega_c = \rho$  and  $U$  is a uniform distribution ( $U$ ) in the interval  $(0, 1)$ , we have:

$$\Omega_{c+1} = \begin{cases} m, & \text{if } u_m \leq a(\rho, m) \\ \rho, & \text{if } u_j > a(\rho, m) \end{cases} \quad (\text{B.15})$$

## Gibbs Sampling

This algorithm is a folding work of (METROPOLIS *et al.*, 1953) and (HASTINGS, 1970). The advances provided by such works ratchet up the development of new solutions, mainly regarding Physics (DOBROW, 2016). The Gibbs Sampling algorithm was one of its cornerstones, as (GEMAN; GEMAN, 1987) presented an algorithm capable of being addressed to face problems involving high dimensionality of data, which may be in the area of Bayesian statistics. Inside the method, distribution  $\lambda$  is multidimensional according to the following joint probability density function:

$$\gamma(\chi) = \gamma(\omega_1, \omega_2 \dots \omega_c) \quad (\text{B.16})$$

A multivariate Markov Chain is built based on  $\lambda$  distribution, thus its values are taken from multidimensional space. The algorithm creates elements as follows:

$$\Omega^{(0)}, \Omega^{(1)} \dots \Omega^{(c)} \quad (\text{B.17})$$

$$= (\Omega_1^{(0)}, \Omega_2^{(0)} \dots \Omega_\alpha^{(0)}), (\Omega_1^{(1)}, \Omega_2^{(1)} \dots \Omega_\alpha^{(1)}), (\Omega_1^{(c)}, \Omega_2^{(c)} \dots \Omega_\alpha^{(c)}) \quad (\text{B.18})$$

Also, it updates each component of the iteratively multidimensional conditional vector in the remaining  $\alpha - 1$  parts. It is a particular case of the Metropolis-Hastings algorithm (DOBROW, 2016; CASELLA; GEORGE, 1992).

## APPENDIX C – PARTIAL DIRECTED COHERENCE

One of the initial exploratory analyses in order to understand the Local Field Potentials dataset that were used in this thesis was the application of the Partial Directed Coherence methodology. This work was converted into an article published in congress:

- Talysson Manoel de Oliveira Santos, Victor Hugo Batista Tsukahara, Jasiara Carla de Oliveira, Vinicius Rosa Cota, Carlos Dias Maciel, Graph Model Evolution During Epileptic Seizures: linear model approach. **II Latin American Workshop on Computational Neuroscience (LAWCN)**, Communications in Computer and Information Science (CCIS), Springer International Publishing, 2019.

# Graph Model Evolution During Epileptic Seizures: linear model approach

Talysson M. O. Santos<sup>1</sup>, Victor H. B. Tsukahara<sup>1</sup>, Jasiara C. de Oliveira<sup>2</sup>,  
Vinicius Rosa Cota<sup>[0000-0002-2338-5949]</sup><sup>2</sup>, and Carlos D. Maciel<sup>1</sup>

<sup>1</sup> Signal Processing Laboratory, Dept of Electrical Engineering, University of São Paulo, São Carlos, Brazil

<sup>2</sup> Laboratory of Neuroengineering and Neuroscience, Dept of Electrical Engineering, Federal University of São João Del-Rei  
talyssonsantos@usp.br

**Abstract.** Epilepsy is a brain disorder characterized by sustained predisposition to generate epileptic seizures. According to the World Health Organization, it is one of the most common neurological disorders, affecting approximately 50 million people worldwide. A modern approach for brain study is to model it as a complex system composed of a network of oscillators in which the emergent property of synchronization arises. By this token, epileptic seizures can be understood as a process of hypersynchronization between brain areas. To assess such property, Partial Directed Coherence (PDC) method represents a suitable technique, once it allows a more precise investigation of interactions that may reveal direct influences from one brain area on another. During connectivity analysis, there may be a need to assess the statistical significance of the communication threshold and Surrogate Data, a method already applied for that purpose, can be used. Hence, the objective in this work was to carry out PDC connectivity analysis in combination with Surrogate Data to evaluate the communication threshold between brain areas and develop a graph model evolution during epileptic seizure, according to the classical EEG frequency bands. The main contribution is the threshold analysis adding statistical significance for connectivity investigation. A case study performed using EEG signals from rats showed that the applied methodology represents an appropriate alternative for functional analysis, providing insights on brain communication.

**Keywords:** Epilepsy · Seizures · Connectivity Analysis · Partial Directed Coherence · Surrogate

## 1 Introduction

Neurological disorders represent a global burden issue for the healthcare area due to their increasingly important role in death and disability causation [1]. Among them, epilepsy is a highly and prevalent disease, characterized by a sustainable predisposition to generate epileptic seizures which result in social, psychological, cognitive, and neurobiological deficits [2]. Approximately 50 million people worldwide suffer from this condition [3].

There are many mechanisms to explain this disease [4]. A well-accepted theory has to do with multifaceted unbalance between excitatory and inhibitory neural tonus [5, 6, 4]. Another important aspect discussed in the literature is synchronization. This perspective models the brain as a system of systems in which the understanding of the organ as a set of subsystems (brain areas) interacting among themselves is implied. As a consequence, there is synchronization from a particular brain region with other areas giving rise to emergent properties [7, 8]. Therefore epileptic seizures under this scope is interpreted as hypersynchronization phenomena [9, 4, 10, 8].

Following this rationale, epileptic seizures are a neural synchronization expression and usually its epileptiform activity is evidenced into electrographic (EEG) recording through high-amplitude spikes and other disturbances resulting in its prevalence to develop epilepsy-based studies [11]. EEG signals can be collected using invasive or non-invasive techniques [12]. However, using electrodes directly in brain tissue to perform acquisition could be an essential option to map measurable indicators of epileptogenicity with good enough resolution [13].

Specifically, to develop a brain functional analysis, there are many methods to be applied. On the other hand, technique selection depends on signal features and study objectives. A multivariate autoregressive (MVAR) modelling to perform connectivity analysis of electroencephalographic signals represents a suitable technique, and Partial Directed Coherence (PDC) fits in this situation [14]. It is a well-established method used in Neuroscience [15–19] proposed by [20] and allows a more precise study of the interactions [21] because of its main advantage: its capability to denote active connections exhibiting the direct influences from a given brain area to other regions [22].

Once the PDC method is carried out, there may be a need to assess the statistical significance of connectivity measures, and Surrogate Data can be used [23]. The method basically creates data based on original signals, maintaining the power spectrum and randomizing Fourier phases. Then, they are compared with original signals through a hypothesis test to check the strength of connectivities discovered during functional analysis. It was used in combination with Information Theoretical methods in [24, 25], with Granger Causality in [23], and with Directed Coherence and Partial Directed Coherence in [21, 26–28] for that aim.

Hence, the paper presents the use of the PDC method combined with Surrogate Data to develop a functional connectivity analysis of EEG epileptic signals and with monitoring of its graph model evolution. Section 2 is going to present

the theory related to Partial Directed Coherence, Surrogate Data Analysis. Section 3 presents the EEG data used and the applied methodology. Section 4 reports the results achieved, and Section 5 discusses the results. To close the paper, the Section 6 presents a conclusion.

## 2 Theory

The section presents a brief review of the required theory to comprehend the methodology application of this paper. Partial Directed Coherence and Surrogate Data are defined. Some equations are established to uniform mathematical notation of this paper.

### 2.1 Partial Directed Coherence (PDC)

Partial Directed Coherence is a frequency-domain approach of Granger Causality [29]. The method is grounded on a multivariate autoregressive model, aiming the study of direct connection among time series [20].

Initially, the time series matrix can be drafted as:

$$X(t) = \sum_{r=1}^p A_r(r)X(t-r) + E(t) \quad (1)$$

being  $p$  the order of autoregressive equation,  $A_r(r)$  representing the coefficients matrix - whereby contains  $a_{ij}$  items and  $E(t)$  the noise matrix - for each time series. It is important to note  $a_{ij}$  elements, which depict the effect of  $x_j(n-r)$  towards  $x_i(t)$ . Another issue is the equation in time domain, and PDC is performed in frequency domain. In this regard, Discrete Time Fourier Transform (DTFT) is applied [30]. Therefore, the coefficient matrix  $A_r(r)$  is transformed into  $A_r(f)$ :

$$A_r(f) = \sum_{r=1}^p A_r e^{-ir2\pi f} \quad (2)$$

in equation 2,  $p$  is still the autoregressive model order. The  $i$  variable inside squared root represents the complex number unit -  $i = \sqrt{-1}$ . Thus, the PDC equation which expresses the effect of  $x_j(n-r)$  towards  $x_i(t)$  can be written as follows:

$$PDC = \pi_{ij}(f) = \frac{A'_{ij}(f)}{\sqrt{a_j'^H(f)a_j'(f)}} \quad (3)$$

the variable  $H$  represents the Hermitian matrix,  $a_j$  denotes the  $j$ th item from matrix  $A'$ , being calculated as:

$$A'(f) = I - A(f) \quad (4)$$

variable  $I$  denotes identity matrix.



## 2.2 Surrogate Data for Hypothesis Test

The investigation of connectivity between brain regions when using the PDC method may reveal issues to handle and one of the questions that may appear is the connectivity threshold. In that case, statistical methods may be helpful. In particular, when there is a lack of data, Surrogate Data can be useful [31]. Another important issue is the conclusions provided: they are sufficiently robust to provide new information on the subject under investigation [21].

For method application, there are quite a few techniques like Random Permutation (RP) Surrogates, Fourier Transform (FT), Amplitude Adjusted Fourier Transform (AAFT), and Iterative Amplitude Adjusted Fourier Transform (IAAFT). With regard to the IAAFT technique, it was proposed by Schreiber & Schmitz [32], and had the aim to overcome the AAFT technique bias [21]. It consists of generating a Surrogate Data from the original signal, keeping the same power spectrum and randomizing Fourier phases. Therefore, when comparing the statistics from original signals and Surrogate Data statistics, the null hypothesis can be accepted or rejected. Some literature can be reviewed in neuroscience related to the use of Surrogate Data to assist the strength as well as the type of interdependency among electroencephalographic signals [33, 26, 34, 21].

Related with the number of Surrogate Data to be created, there is a well established rank-order test proposed by [35] that can be used [36]. Assume  $\Psi$  is the probability of false rejection - defining the level of significance(S) as:

$$S = (1 - \Psi) \cdot 100\% \quad (5)$$

The number of Surrogate Data to be created(M) is defined as follows:

$$M = \frac{K}{\Psi} - 1 \quad (6)$$

where K is an integer number defined by the type of test - 1 if it is one-sided and 2 in the case of a two-sided test- and  $\Psi$  is the probability of false rejection. Usually, K=1 is adopted due to computational effort to generate surrogates [36].

Therefore, to evaluate the connectivity threshold between EEG signals, Surrogate Data can be created using the IAAFT algorithm. Then original signals can be compared with built data, using a hypothesis test to evaluate the hypothesis of presence or absence of connectivity.

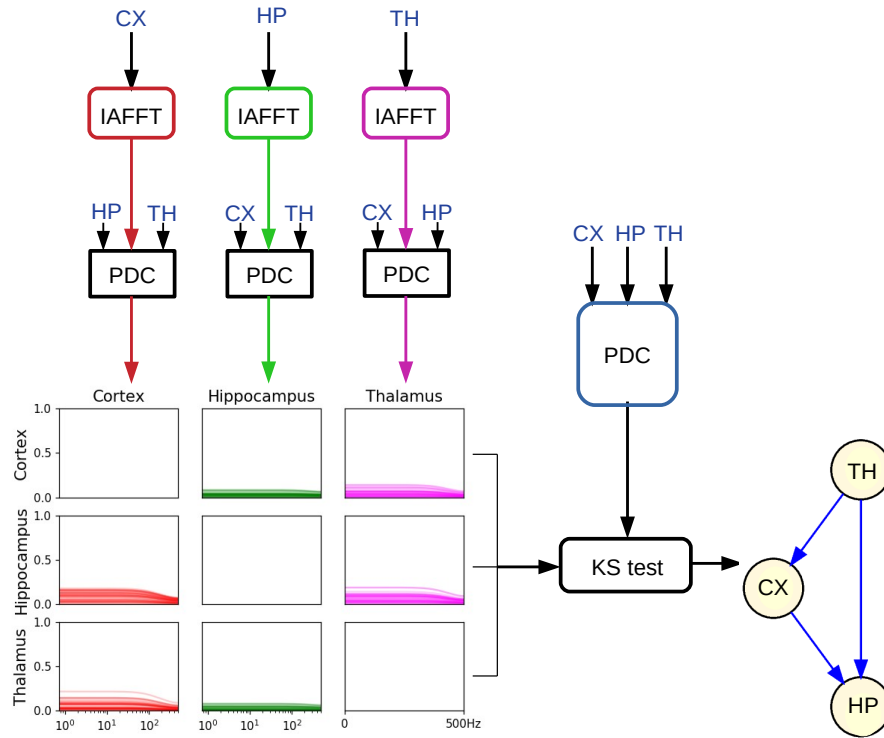
## 3 Materials and Methods

In this section information about the EEG signals used to perform the study is reported, as well as the applied methodology, including algorithms and the computational environment.

### 3.1 Applied Methodology

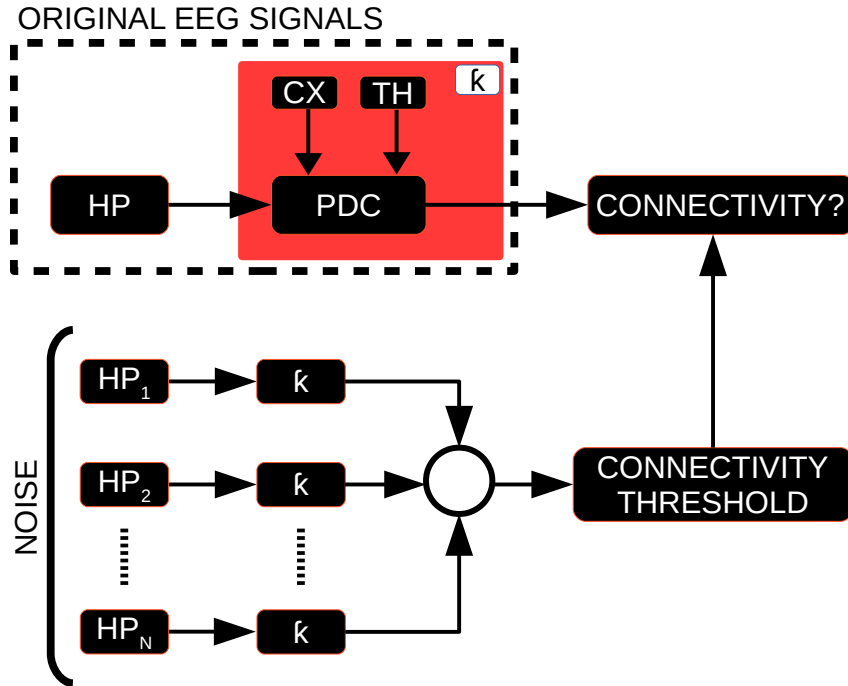
The summary of the applied methodology can be observed in Figure 1. The first step was the application of PDC method on the original EEG time series. The

second step was the creation of Surrogate Data by reference to three brain areas signal recordings: cortex (CX), thalamus (TH) and hippocampus (HP). It was defined  $\psi=3\%$  and  $K=1$ , resulting on M number of 35 for rounding purpose. Each Surrogate Data was used with the other original signals to perform PDC method. Finally, the connectivity analysis was performed, comparing the Surrogate Data PDC with original time series PDC. Details of Surrogate analysis can be observed in Figure 2.



**Fig. 1.** Methodology diagram: Using Surrogate Data, 35 signals from each area are generated to determine the connectivity threshold. Kolmogorov-Smirnov test is used for null hypothesis test comparing the PDC of the real signal with PDC obtained with the Surrogate Data.

To perform the connectivity analysis, classical EEG frequency bands were observed. For each frequency interval, Kolmogorov-Smirnov statistical test was applied, defining  $p\text{-value}=5\%$ , to validate the communication between brain areas. The acceptance of the null hypothesis represents that there is no connection.



**Fig. 2.** Surrogate analysis to check connectivity between hippocampus and other brain areas. PDC method is applied on original signals. Surrogate data for hippocampus (HP) EEG signals are built, maintaining the power spectrum. Then for each Surrogate signal PDC is performed using cortex (CX) and thalamus (TH) original signals to discover the connectivity threshold between hippocampus and cortex ( $\pi_{HP \rightarrow CX}$ ) and hippocampus and thalamus ( $\pi_{HP \rightarrow TH}$ ). Finally, threshold is compared with PDC from original signals to validate the connectivity values and strength through KS Test.

The applied methodology described and represented in Figure 1 was used to investigate two periods of EEG signals: the basal and PTZ infusion - until generalized tonic-clonic (GTCS) behaviour. The analysis was performed in the recordings of six rats.

Simulations were developed in Python language, using the packages: Connectivipy, Graphviz, Matplotlib, Nolitsa, Numpy, Pandas, Scipy, Seaborn, and Time. The code was executed on a computer with an Intel i7 processor, 4GB of RAM, running Linux Lite 4.0 operational system.

### 3.2 Simulated data

To test the python package Connectivipy, examples of PDC application from [20] and [37] were performed to reproduce their paper results. Nonlinear equations,

a situation in which the method fails, were used to verify whether the python package is working properly. After all tests the case study was carried out.

### 3.3 Database to Perform Case Study

The study uses EEG signals database from the Laboratory of Neuroengineering and Neuroscience (LINNce), from Federal University of São João Del Rei. The laboratory uses male Wistar rats weighing between 250 and 350 grams - coming from the University Central Vivarium -, to collect data and test methods of electrical stimulation. All described procedures are in according to ethics committee under protocol 31/2014.

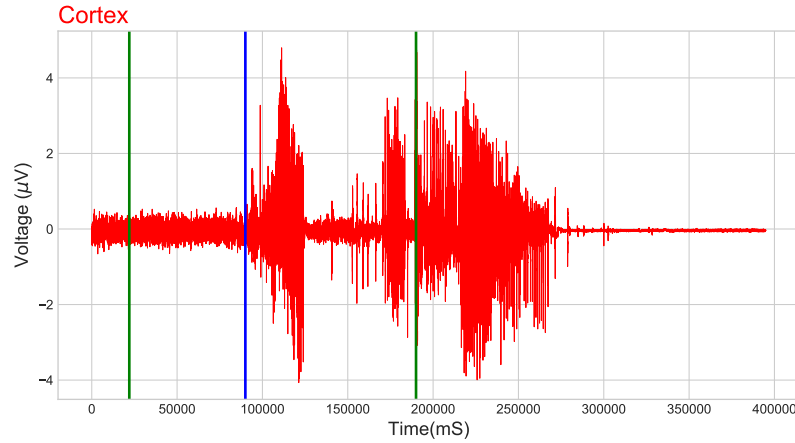
To collect the electroencephalographic signals, electrodes - monopolar type, stainless steel, covered by Teflon - were placed directly inside the rat brain through stereotactic surgery - inside right Thalamus and Hippocampus -[8]. Furthermore, two microsurgical screws (length 4.7 mm, diameter 1.17 mm, Fine Science Tools, Inc., North Vancouver, Canada) were implanted: the first aiming the cortical recording of right hemisphere, and the second to be used as the reference, positioned in frontal bone. Electrodes and screws were positioned with assistance of neuroanatomic atlas [38]. Pentylenetetrazole (PTZ; a proconvulsant drug) was used to induced acute epileptic seizures in the rats [8]. EEG signals from each rat were recorded and the rat was filmed at the same time to perform behavioural analysis (video-EEG technique) - observe classic seizure features such as facial automatisms, head myoclonus, forelimb and hindlimbs clonus, elevation and fall, generalized tonic-clonic seizure -, to enable correlation of them with the electrophysiological events verified during the brain activity recording.

EEG recording was performed using 1 KHz sampling rate. Signals were also amplified 2000 V/V, filtered (second-order Butterworth filter from 0.3 to 300 Hz band) using A-M Systems (model 3500) amplification system, and acquired on National Instruments (PCI 6023E) A/D converter controlled by LabView's Virtual Instrument developed at LINNce. Noise coming from power grid at 60Hz frequency was avoided with employment of shielded cables and Faraday cage.

## 4 Results

An example of original EEG signals can be observed in Figure 3. It is important to note the basal and infusion intervals: green vertical lines indicate the infusion interval, being the first line the end of the basal interval that starts in 0 mS. The blue vertical line represents the onset of epileptiform activity jointly with the beginning of convulsive behaviour, and the last green vertical line indicates the limit of infusion interval in which the rat did not present GTCS behaviour yet.

The PDC method was applied to EEG signals on basal and infusion intervals. After this step, Surrogate Data was created for each brain area and then used with other original signals to perform another PDC analysis. The computational



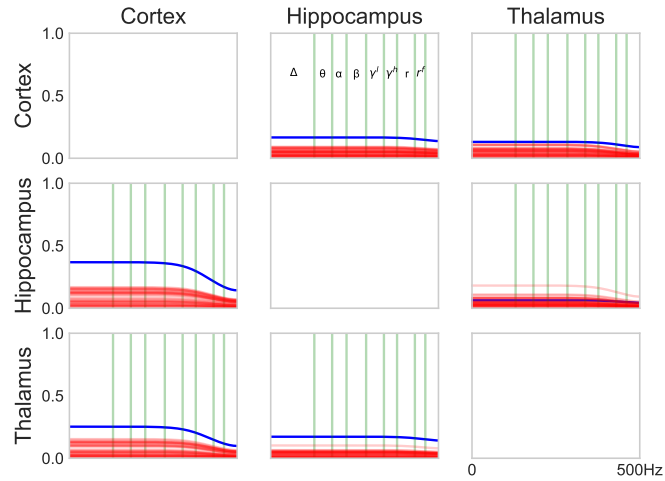
**Fig. 3.** Example of cortex signal - rat R070. The green vertical lines indicate the infusion interval - the first represent the end of basal interval which starts in 0 mS. The blue vertical line represents the limit of infusion interval in which the rat did not present GTCS behaviour yet.

time to run PDC on original signals was 32.33 seconds and to perform Surrogate Data and its PDC method was 2030.85 seconds, in a total of 2063.18 seconds - for each interval(basal and infusion). An example of PDC method, containing the original signals (blue line) and Surrogate Data (red line), and an illustration of Surrogate Data power spectrum can be seen in Figures 4-a and 4-b, respectively.

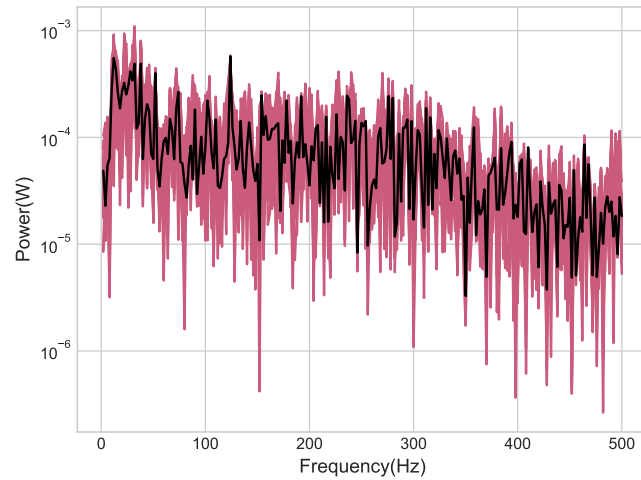
The result of the PDC method performed with Surrogate Data was used to discover the connectivity ( $\pi_{ij}$ ) threshold between brain areas, for each standard EEG frequency - Figure 4-a shows an example extracted from rat R070. Thus, the Kolmogorov-Smirnov Test was used to evaluate the PDC performed with original EEG signals, revealing the connectivity between brain regions, according to each threshold, and to compute the strength of each connection.

Figure 5 shows an example of a connectivity graph developed based on applied methodology. For each EEG frequency band, communication between cortex, hippocampus, and thalamus, during basal (blue edges), and infusion intervals (red edges) is shown.

The table 1 reports the connectivity measures which remained among most rats used to build the EEG signals database. It is possible to check the connectivity threshold between each brain area, and the strength of its communication.

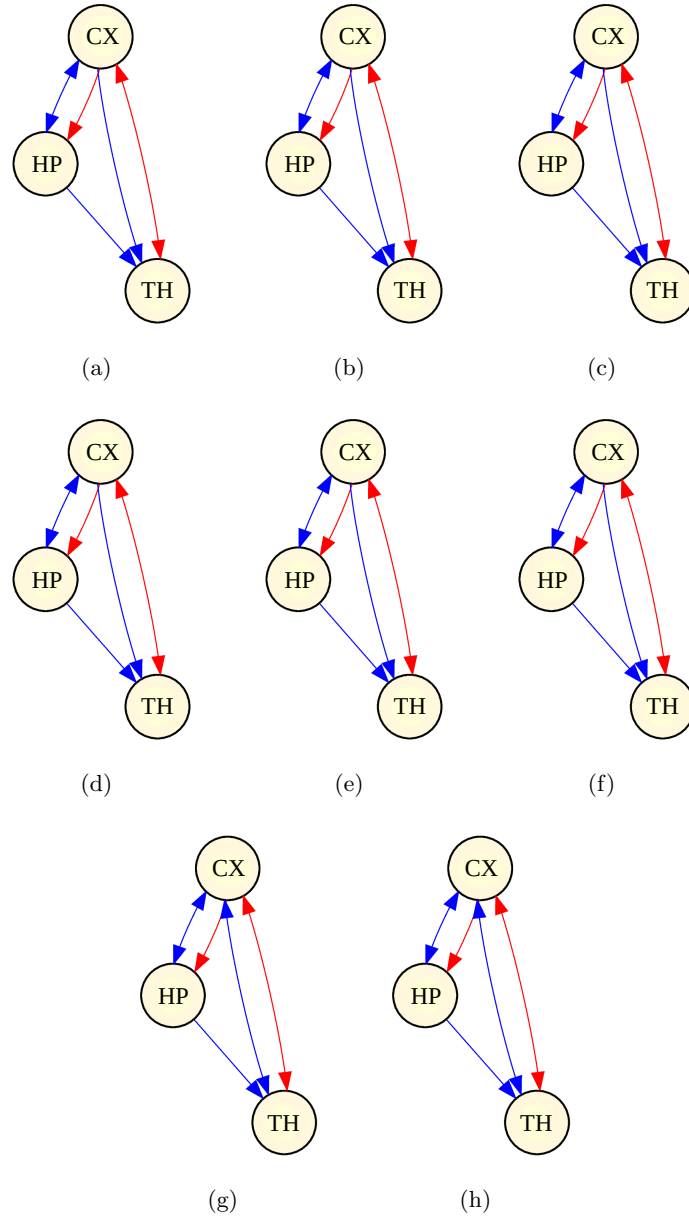


(a)



(b)

**Fig. 4.** (a) Example of connectivity threshold - extracted from rat R070, basal interval. The blue line represents the PDC for the real signal. Red lines represents the PDC for surrogate data. The x-axis is in logarithmic scale and vertical green lines represents frequency intervals. (b) Power spectrum of real data and Surrogate Data example - from cortex, rat R070. Black line represents the power spectrum for real EEG signal. Red blur represents the power spectrum for 35 surrogates data.



**Fig. 5.** Example of connectivity ( $\pi_{ij}$ ) graph for each frequency range - extracted from rat R070. a)  $\Delta$ , b)  $\theta$ , c)  $\alpha$ , d)  $\beta$ , e)  $\gamma^l$ , f)  $\gamma^h$ , g) ripple and h) fast ripple. Blue edges represents connectivity during basal interval and red edges represents connectivity during infusion interval.

**Table 1.** Connectivity measures that remained for most rats. The table is organized according to classical EEG frequency intervals ( $\Omega$ ), time series interval - basal (B) or infusion (I) -, and connectivity ( $\pi_{ij}$ ) between brain areas - cortex (CX), hippocampus (HP) and thalamus (TH). Arrows indicates the connectivity - e.g. C  $\rightarrow$  T represent the communication between cortex and thalamus. Threshold ( $\lambda$ ) and strength ( $\tau$ ) - how much connectivity value surpassed threshold - and no connectivity (-) are pointed out.

$\Omega$		$\Delta$		$\theta$		$\alpha$		$\beta$		$\gamma^l$		$\gamma^h$		$r$		$r^f$	
		B	I	B	I	B	I	B	I	B	I	B	I	B	I	B	I
H $\rightarrow$ C	Mean	0.33	-	0.33	-	0.32	-	0.30	-	0.24	-	0.18	-	0.13	-	0.09	-
	Max	0.69	-	0.68	-	0.65	-	0.56	-	0.39	-	0.24	-	0.18	-	0.14	-
	Min	0.09	-	0.09	-	0.09	-	0.09	-	0.09	-	0.09	-	0.09	-	0.06	-
	Mean	0.11	-	0.11	-	0.11	-	0.11	-	0.09	-	0.07	-	0.05	-	0.04	-
	Max	0.38	-	0.38	-	0.38	-	0.36	-	0.32	-	0.24	-	0.14	-	0.08	-
T $\rightarrow$ C	Mean	-	0.21	-	0.21	-	0.21	-	0.20	-	0.17	-	0.13	-	0.10	-	0.08
	Max	-	0.53	-	0.52	-	0.51	-	0.47	-	0.37	-	0.25	-	0.15	-	0.11
	Min	-	0.06	-	0.06	-	0.06	-	0.06	-	0.06	-	0.06	-	0.05	-	0.04
	Mean	-	0.27	-	0.27	-	0.27	-	0.27	-	0.26	-	0.23	-	0.18	-	0.14
	Max	-	0.54	-	0.54	-	0.55	-	0.55	-	0.53	-	0.50	-	0.42	-	0.32
C $\rightarrow$ H	Mean	0.25	0.11	0.25	0.11	0.24	0.11	0.22	0.11	0.18	0.11	0.15	0.09	0.11	0.07	0.08	0.06
	Max	0.45	0.16	0.44	0.16	0.40	0.16	0.30	0.16	0.23	0.16	0.19	0.14	0.16	0.11	0.12	0.08
	Min	0.14	0.07	0.14	0.07	0.14	0.07	0.14	0.07	0.14	0.07	0.10	0.06	0.06	0.05	0.04	0.04
	Mean	0.10	0.29	0.10	0.29	0.10	0.29	0.10	0.29	0.11	0.28	0.11	0.27	0.10	0.26	0.09	0.24
	Max	0.15	0.38	0.15	0.38	0.15	0.38	0.15	0.38	0.16	0.37	0.18	0.37	0.17	0.39	0.15	0.42
T $\rightarrow$ H	Mean	-	0.16	-	0.16	-	0.16	-	0.15	-	0.13	-	0.11	-	0.09	-	0.07
	Max	-	0.27	-	0.27	-	0.27	-	0.24	-	0.19	-	0.17	-	0.13	-	0.11
	Min	-	0.05	-	0.05	-	0.05	-	0.05	-	0.05	-	0.05	-	0.04	-	0.03
	Mean	-	0.15	-	0.15	-	0.14	-	0.14	-	0.13	-	0.11	-	0.07	-	0.05
	Max	-	0.49	-	0.49	-	0.48	-	0.44	-	0.33	-	0.23	-	0.18	-	0.14
C $\rightarrow$ T	Mean	-	0.12	-	0.12	-	0.12	-	0.12	-	0.12	-	0.10	-	0.09	-	0.07
	Max	-	0.19	-	0.19	-	0.19	-	0.18	-	0.18	-	0.16	-	0.13	-	0.10
	Min	-	0.08	-	0.08	-	0.08	-	0.08	-	0.08	-	0.07	-	0.06	-	0.04
	Mean	-	0.17	-	0.17	-	0.17	-	0.17	-	0.17	-	0.18	-	0.19	-	0.19
	Max	-	0.35	-	0.35	-	0.35	-	0.36	-	0.36	-	0.38	-	0.42	-	0.46
H $\rightarrow$ T	Mean	0.30	0.24	0.30	0.24	0.29	0.24	0.27	0.22	0.22	0.18	0.17	0.14	0.12	0.10	0.09	0.08
	Max	0.67	0.67	0.66	0.66	0.64	0.64	0.55	0.55	0.38	0.38	0.24	0.24	0.17	0.17	0.12	0.17
	Min	0.09	0.05	0.09	0.05	0.09	0.05	0.09	0.05	0.09	0.05	0.09	0.04	0.06	0.04	0.04	0.03
	Mean	0.17	0.2	0.16	0.20	0.16	0.19	0.14	0.17	0.13	0.14	0.10	0.10	0.07	0.06	0.06	0.04
	Max	0.40	0.52	0.40	0.52	0.39	0.51	0.36	0.49	0.29	0.42	0.20	0.31	0.12	0.20	0.14	0.13

## 5 Discussion

It is possible to observe from recorded EEG signals, the basal and infusion intervals, including the epileptic seizure interval (not analysed in this paper). An important issue when applying MVAR modelling is stationarity, a requirement to perform linear models analysis such as PDC. The intervals used to apply



methodology (basal interval and infusion interval only until GTCS behaviour) assured the conditions to apply PDC.

As it can be verified in Figure 4-a, PDC was applied to original EEG signals and Surrogate Data. Sometimes visually it is not possible to check connectivity between brain areas, such as the link between thalamus and hippocampus. Kolmogorov-Smirnov Test was essential to identify where there is genuine communication and detect the strength and threshold for each case. Figure 4-b ascertain the maintenance of power spectrum for Surrogate Data, essential to evaluate threshold between brain areas - this is assuring the same information contained in EEG signals.

For most rats from EEG database, it is possible to verify connectivity between some areas (Table 1). For basal and infusion intervals, connectivity was identified between  $\pi_{CX \rightarrow HP}$  and  $\pi_{HP \rightarrow TH}$ . It was identified only during infusion interval connectivity between  $\pi_{TH \rightarrow CX}$ ,  $\pi_{TH \rightarrow HP}$  and  $\pi_{CX \rightarrow TH}$ . Only for the basal interval, it was observed connectivity between  $\pi_{HP \rightarrow CX}$ .

Regarding hippocampus and cortex communication it was not possible to observe information flow between them during basal interval but appeared in the infusion interval. This result is not biologically plausible. Literature reveals bidirectional communication related with memory formation [39, 40]. Probably the answer for the issue stems from communication strength between brain areas: EEG noise is affecting functional connectivity analysis resulting in no information flow identification.

A significant result is observed in the threshold values found for basal and infusion intervals. It is possible to check a smaller value for the infusion interval when compared with the basal period. It may represent that a slight value during infusion interval is cooperating to hypersynchronization between brain areas, aiding the maintenance of excitatory activity, a phenomenon studied in [41] and [42]. The connectivity difference observed between  $\pi_{HP \rightarrow TH}$  is supported by [43], in which it was identified that epileptic seizure starts in the limbic system (amygdala and hippocampus), recruiting thalamus that, thus, plays as a neural substrate synchronizer. As a consequence, the thalamus is responsible for propagating the epileptic activity to the cortex and other brain areas supporting the spread of the seizure.

Regarding threshold values, there is yet a critical behaviour. For low EEG frequencies (delta to approximately low gamma), it is possible to observe more difference in basal and infusion thresholds when compared with the difference in high frequency bands. It is consistent with the concept reported by [44] that low frequencies support the communication between brain substrates. Conversely, high EEG frequencies support the local hyper synchronism, and this is also consistent with the performed methodology.

## 6 Conclusions

The use of PDC method to perform functional connectivity analysis revealed to be a suitable option to investigate direct communications between brain areas,

based on its MVAR model and linear approach. For EEG intervals where stationarity is ensured, the method works properly. The use of Surrogate Data and KS Test to implement the threshold analysis added more insights in connectivity analysis. Through this approach it was possible to discover differences between the threshold of EEG low and high frequencies, more about the communication dynamics between brain areas, consistent with neuroscience literature, allowing a thorough investigation. It demonstrates that the combination of methods proposed for this work is a suitable choice to perform functional analysis in electroencephalographic signals.

## 7 Acknowledgment

Laboratory of Neuroengineering and Neuroscience, Department of Electrical Engineering, Federal University of São João Del-Rei for partnership and EEG data used in this paper. Experimental procedures in animals were supported by the Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG) [grant number APQ 02485-15].

## References

1. V.L. Feigin, E. Nichols, T. Alam, M.S. Bannick, E. Beghi, N. Blake, W.J. Culpeper, E.R. Dorsey, A. Elbaz, R.G. Ellenbogen, et al., *The Lancet Neurology* (2019)
2. R.S. Fisher, C. Acevedo, A. Arzimanoglou, A. Bogacz, J.H. Cross, C.E. Elger, J. Engel, L. Forsgren, J.A. French, M. Glynn, D.C. Hesdorffer, B. Lee, G.W. Mathern, S.L. Moshé, E. Perucca, I.E. Scheffer, T. Tomson, M. Watanabe, S. Wiebe, *Epilepsia* **55**(4), 475 (2014). <https://doi.org/10.1111/epi.12550>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/epi.12550>
3. R.D. Thijs, R. Surges, T.J. O'Brien, J.W. Sander, *The Lancet* (2019)
4. O. Devinsky, A. Vezzani, T.J. O'Brien, I.E. Scheffer, M. Curtis, P. Perucca, *Nature Reviews Disease Primers* **4** (2018). <https://doi.org/10.1038/nrdp.2018.24>. URL <https://www.nature.com/articles/nrdp201824#supplementary-information>
5. M. Mele, R.O. Costa, C.B. Duarte, *Frontiers in Cellular Neuroscience* **13**, 77 (2019). <https://doi.org/10.3389/fncel.2019.00077>. URL <https://www.frontiersin.org/article/10.3389/fncel.2019.00077>
6. A. Oyarzabal, I. Marin-Valencia, *Journal of Inherited Metabolic Disease* **42**(2), 220 (2019). <https://doi.org/10.1002/jimd.12071>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/jimd.12071>
7. O.S. Andrea Avena-Koenigsberger, Bratislav Mšic, *Nature Reviews Neuroscience* **19**, 17 (2017). <https://doi.org/http://dx.doi.org/10.1038/nrn.2017.149>
8. V. Cota, B. Marcela Bacellar Drabowski, J. Carla de Oliveira, M. Moraes, *Journal of Neuroscience Research* **94**, 463 (2016). <https://doi.org/10.1002/jnr.23741>
9. S.A. Weiss, R. Staba, A. Bragin, K. Moxon, M. Sperling, M. Avoli, J. Engel, *Neurobiology of Disease* **124**, 183 (2019). <https://doi.org/https://doi.org/10.1016/j.nbd.2018.11.014>. URL <http://www.sciencedirect.com/science/article/pii/S096999611830682X>
10. A.E. Olamat, A. Akan, in *2017 25th Signal Processing and Communications Applications Conference (SIU)* (2017), pp. 1–4. <https://doi.org/10.1109/SIU.2017.7960194>

11. F. Ibrahim, S. Abd-Elateif El-Gindy, S.M. El-Dolil, A.S. El-Fishawy, E.S.M. El-Rabaie, M.I. Dessouky, I.M. Eldokany, T.N. Alotaiby, S.A. Alshebeili, F.E. Abd El-Samie, *International Journal of Speech Technology* **22**(1), 191 (2019). <https://doi.org/10.1007/s10772-018-09565-7>. URL <https://doi.org/10.1007/s10772-018-09565-7>
12. E.K.M. St.Louis, L.C.M. Frey, *Electroencephalography (EEG): An Introductory Text and Atlas of Normal and Abnormal Findings in Adults, Children, and Infants* (American Epilepsy Society, 2016). URL <https://www.ncbi.nlm.nih.gov/books/NBK390354/>
13. F. Bartolomei, S. Lagarde, F. Wendling, A. McGonigal, V. Jirsa, M. Guye, C. Bénar, *Epilepsia* **58**(7), 1131 (2017). <https://doi.org/10.1111/epi.13791>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/epi.13791>
14. B. Pester, T. Lehmann, L. Leistriz, H. Witte, C. Ligges, *Journal of Neuroscience Methods* **309**, 199 (2018)
15. G. Varotto, S. Franceschetti, D. Caputo, E. Visani, L. Canafoglia, E. Freri, F. Ragona, G. Avanzini, F. Panzica, *Clinical Neurophysiology* **129**(11), 2372 (2018). <https://doi.org/https://doi.org/10.1016/j.clinph.2018.09.008>. URL <http://www.sciencedirect.com/science/article/pii/S138824571831229X>
16. S. Schulz, J. Haueisen, K.J. Bär, A. Voss, *Physiological Measurement* **39**(7), 074004 (2018). <https://doi.org/10.1088/1361-6579/aace9b>. URL <https://doi.org/10.1088/1361-6579/aace9b>
17. A. Ciaramidaro, J. Toppi, C. Casper, C. Freitag, M. Siniatchkin, L. Astolfi, *Scientific Reports* **8** (2018). <https://doi.org/10.1038/s41598-018-24416-w>
18. J.A. Gaxiola-Tirado, R. Salazar-Varas, D. Gutiérrez, *IEEE Transactions on Cognitive and Developmental Systems* **10**(3), 776 (2018). <https://doi.org/10.1109/TCDS.2017.2777180>
19. L. Ning, Y. Rathi, *IEEE Transactions on Medical Imaging* **37**(9), 1957 (2018). <https://doi.org/10.1109/TMI.2017.2739740>
20. L.A. Baccalá, K. Sameshima, *Biological Cybernetics* **84**(6), 463 (2001). <https://doi.org/10.1007/PL00007990>. URL <https://doi.org/10.1007/PL00007990>
21. J.A. Adkinson, B. Karumuri, T.N. Hutson, R. Liu, O. Alamoudi, I. Vlachos, L. Iasemidis, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **27**(1), 22 (2019). <https://doi.org/10.1109/TNSRE.2018.2886211>
22. D. Huang, A. Ren, J. Shang, Q. Lei, Y. Zhang, Z. Yin, J. Li, K.M. von Deneen, L. Huang, *Frontiers in Human Neuroscience* **10**, 235 (2016). <https://doi.org/10.3389/fnhum.2016.00235>. URL <https://www.frontiersin.org/article/10.3389/fnhum.2016.00235>
23. P.L.C. Rodrigues, L.A. Baccalá, in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (2016), pp. 5493–5496. <https://doi.org/10.1109/EMBC.2016.7591970>
24. W. Endo, F.P. Santos, D. Simpson, C.D. Maciel, P.L. Newland, *Journal of computational neuroscience* **38**(2), 427 (2015)
25. F.P. Santos, C.D. Maciel, P.L. Newland, *Journal of computational neuroscience* **43**(2), 159 (2017)
26. L. Faes, A. Porta, G. Nollo, *IEEE Transactions on Biomedical Engineering* **57**(8), 1897 (2010). <https://doi.org/10.1109/TBME.2010.2042715>
27. L. Faes, A. Porta, G. Nollo, in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE, 2009)*, pp. 6280–6283
28. L. Faes, G.D. Pinna, A. Porta, R. Maestri, G. Nollo, *IEEE transactions on biomedical engineering* **51**(7), 1156 (2004)






29. R. Chopra, C.R. Murthy, G. Rangarajan, IEEE Transactions on Signal Processing **66**(22), 5803 (2018). <https://doi.org/10.1109/TSP.2018.2872004>
30. D.Y. Takahashi, L.A. Baccalá, K. Sameshima, Journal of Applied Statistics **34**(10), 1259 (2007). <https://doi.org/10.1080/02664760701593065>. URL <https://doi.org/10.1080/02664760701593065>
31. G. Lancaster, D. Iatsenko, A. Pidde, V. Ticcinelli, A. Stefanovska, Physics Reports **748**, 1 (2018). <https://doi.org/https://doi.org/10.1016/j.physrep.2018.06.001>. URL <http://www.sciencedirect.com/science/article/pii/S0370157318301340>. Surrogate data for hypothesis testing of physical systems
32. T. Schreiber, A. Schmitz, Phys. Rev. Lett. **77**, 635 (1996). <https://doi.org/10.1103/PhysRevLett.77.635>. URL <https://link.aps.org/doi/10.1103/PhysRevLett.77.635>
33. E. Pereda, R.Q. Quiroga, J. Bhattacharya, Progress in Neurobiology **77**(1), 1 (2005). <https://doi.org/https://doi.org/10.1016/j.pneurobio.2005.10.003>. URL <http://www.sciencedirect.com/science/article/pii/S030100820500119X>
34. N.P. Subramaniam, J. Hyttinen, Phys. Rev. E **91**, 022927 (2015). <https://doi.org/10.1103/PhysRevE.91.022927>. URL <https://link.aps.org/doi/10.1103/PhysRevE.91.022927>
35. J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, J.D. Farmer, Physica D: Nonlinear Phenomena **58**(1), 77 (1992). [https://doi.org/https://doi.org/10.1016/0167-2789\(92\)90102-S](https://doi.org/https://doi.org/10.1016/0167-2789(92)90102-S). URL <http://www.sciencedirect.com/science/article/pii/016727899290102S>
36. T. Schreiber, A. Schmitz, Physica D: Nonlinear Phenomena **142**(3-4), 346 (2000)
37. Y. Chen, S.L. Bressler, M. Ding, Journal of neuroscience methods **150**(2), 228 (2006)
38. G. Paxinos, C. Watson, *The Rat Brain in Stereotaxic Coordinates* (Elsevier, 2013). 7th edition
39. A.R. Preston, H. Eichenbaum, Current Biology **23**(17), R764 (2013)
40. H. Eichenbaum, Nature reviews neuroscience **1**(1), 41 (2000)
41. P.J. Uhlhaas, W. Singer, neuron **52**(1), 155 (2006)
42. F. Varela, J.P. Lachaux, E. Rodriguez, J. Martinerie, Nature reviews neuroscience **2**(4), 229 (2001)
43. E.H. Bertram, P. Mangan, N. Fountain, D. Rempe, et al., Epilepsy research **32**(1-2), 194 (1998)
44. A. Schnitzler, J. Gross, Nature reviews neuroscience **6**(4), 285 (2005)

## APPENDIX D – DELAYED MUTUAL INFORMATION

Another exploratory analyses in order to understand the Local Field Potentials dataset that were used in this thesis was the application of the Mutual Information methodology. This work was converted into an article published in congress:

- Victor Hugo Batista Tsukahara, Pedro Virgílio Basílio Jeronymo, Jasiara Carla de Oliveira, Vinicius Rosa Cota, Carlos Dias Maciel, PDC-MI method for EEG Functional Connectivity Analysis. **13th International conference on bio-inspired systems and signal processing (Biosignals)**, Communications in Computer and Information Science (CCIS), Springer International Publishing, 2021.

# Delayed Mutual Information to Develop Functional Analysis on Epileptic Signals

Victor Hugo Batista Tsukahara<sup>1</sup><sup>a</sup>, Pedro Virgilio Basilio Jeronymo<sup>1</sup><sup>b</sup>, Jasiara Carla de Oliveira<sup>2</sup><sup>c</sup>,  
Vinicius Rosa Cota<sup>2</sup><sup>d</sup> and Carlos Dias Maciel<sup>1</sup><sup>e</sup>

<sup>1</sup>Signal Processing Laboratory, Dept of Electrical Engineering, University of São Paulo, São Carlos, Brazil

<sup>2</sup>Laboratory of Neuroengineering and Neuroscience, Dept of Electrical Engineering, Federal University of São João Del-Rei, São João Del-Rei, Brazil

vhbtukahara@usp.br, pedro.jeronymo@usp.br, jasi.oliveira@yahoo.com.br, vrcota@ufsj.edu.br, carlos.maciel@usp.br

**Keywords:** Epilepsy, Entropy, Delayed Mutual Information, Channel Capacity, Transmission Rate.

**Abstract:** Epilepsy is the second most prevalent brain disorder affecting approximately 70 million people worldwide. A modern approach to develop the brain study is to model it as a system of systems, represented by a network of oscillators, in which the emergent property of synchronisation occurs. Based on this perspective, epileptic seizures can be understood as a process of hyper-synchronisation between brain areas. To investigate such process, a case study was conducted applying Delayed Mutual Information (DMI) to perform functional connectivity analysis, investigating the channel capacity (C) and transmission rate (R) between brain areas — cortex, hippocampus and thalamus — during basal and infusion intervals, before the beginning of generalised tonic-clonic behaviour (TCG). The main contribution of this paper is the study of channel capacity and transmission rate between brain areas. A case study performed using 5 LFP signals from rodents showed that the applied methodology represents an another appropriate alternative to existing methods for functional analysis such as Granger Causality, Partial Directed Coherence, Transfer Entropy, providing insights on epileptic brain communication.

## 1 INTRODUCTION


Epilepsy is the second most common neurological disease (Organization et al., 2017) and affects approximately 70 million people worldwide (Spiciarich et al., 2019) representing a public health concern (Niriayo et al., 2019). It is a chronic disease of the central nervous system (CNS) that reaches people of all ages in which it is commonly associated with social difficulties (Beghi, 2019) and can cause health loss such as premature mortality and residual disability (Beghi et al., 2019).


Epilepsy-based studies usually uses electroencephalography (EEG) (Ibrahim et al., 2019) or local field potentials (LFP) (Biasiucci et al., 2019) to check brain electrical activity, although the use of electrodes directly in brain tissue is an important option to map


electrical activity of the brain with better spatial resolution (Bartolomei et al., 2017).


The latest approach to study epilepsy is the analysis of hyper-synchronisation of brain frequencies oscillations as a feature (Yu et al., 2019). (Olamat and Akan, 2017) performed a nonlinear synchronisation analysis in LFP epileptic data introducing this new perspective. (Weiss et al., 2019) used the concept to understand seizure genesis and spreading in human limbic areas and (Devinsky et al., 2018) reported hyper-synchronisation to discuss epilepsy epidemiology and pathophysiology. The brain is modelled as a complex system where each region represents a subsystem and synchronization is an emergent property (Andrea Avena-Koenigsberger, 2017). Changes in this feature during the occurrence of epileptic seizures are an important aspect to understand the epileptic brain network and synchronization (Mei et al., 2019). The pathologic hyper-synchronisation of frequencies oscillations give rise to seizures (Berglind et al., 2018).


There is a hypothesis that high-frequencies oscillations are related with the cortical local brain in-

<sup>a</sup> <https://orcid.org/0000-0003-0713-9067>

<sup>b</sup> <https://orcid.org/0000-0002-1468-9051>

<sup>c</sup> <https://orcid.org/0000-0002-5170-1072>

<sup>d</sup> <https://orcid.org/0000-0002-2338-5949>

<sup>e</sup> <https://orcid.org/0000-0003-0137-6678>

formation processing whereas low-frequencies have connection with larger cortical networks (Anastasiadou et al., 2019). Consequently, the brain interactions through these areas can become complex because of interactions between oscillations at different frequency bands (Anastasiadou et al., 2019). In this situation, functional connectivity may be performed to detect dependencies among neurophysiological signals (Andrea Avena-Koenigsberger, 2017). It can be assessed through different methods with the aim to infer patterns of direct influences (Andrea Avena-Koenigsberger, 2017).

To estimate the dependency between time series there are several methods (Gribkova et al., 2018) and Mutual Information is one of them. It is an Information Theoretic and nonparametric approach that measures generalized, both linear and nonlinear, interdependence between two variables (Akbarian and Erfanian, 2017). This meets the accepted vision that real world time series usually are nonlinear and non stationary (Wan and Xu, 2018).

Usually, the Information Theoretic approaches do not make any hypothesis about the dependency between time series (Nichols et al., 2005). The use of time Delayed Mutual Information (DMI) seeks to quantify the information shared between time series taking into account the previous information content as function of time (Endo et al., 2015). (Li et al., 2018) demonstrated that DMI is a suitable option to develop analysis of nonlinear systems as in the case of neuroscience data. (Kim et al., 2018) used DMI to analyse information transmission of an EEG set from groups of people with mild Alzheimer disease. (Li et al., 2017) applied DMI to characterise hippocampal theta-driving neurons. (Chapeton et al., 2017) performed a study using intracranial EEG to identify effective connections in the brain that exhibit consistent timing across multiple temporal scales.

The objective of this paper is, in performing a case study using Delayed Mutual Information, to develop functional connectivity analysis in rodents LFP signals, investigating the channel capacity (C) and the transmission rate between brain areas. In addition, Surrogate method is used to evaluate the DMI measures. In section 2, it is presented the theory related to DMI and Surrogate. Section 3 describes the LFP data used and the applied methodology. Section 4 presents the achieved results, and in Section 5 the results are discussed. Finally Section 6 brings forward paper conclusions.

## 2 THEORY

This section presents the main theory required to develop this paper. First it is introduced Mutual Information explaining the main concepts of channel capacity and transmission rate. Then the Surrogate method, used to assess statistical significance of the performed analysis, is defined.

### 2.1 Delayed Mutual Information

The measure of how deterministic is a given variable can be determined through its entropy ( $H$ ), defined by (Cover and Thomas, 2012):

$$H(X) = - \sum_{x \in \mathcal{X}} p(x) \log_a p(x) \quad (1)$$

where  $X$  is a discrete random variable,  $p(x) = P\{X = x\}$  is the probability of  $X$  equal to  $x$ ,  $x \in \mathcal{X}$ , i.e. the probability mass function of  $X$ , and  $a$  is the logarithm base that provides the entropy measure in bits in the case of  $a = 2$ . Given a signal  $X$  and another signal  $Y$ , the Mutual Information may quantify the information shared between this signals, which means how much it is possible to reduce the uncertainty of signal  $X$  given the knowledge of signal  $Y$  (Cover and Thomas, 2012).

The Delayed Mutual Information (DMI) according to (Nichols et al., 2005) is the quantification of information shared between  $X$  and  $Y^\tau$  where  $Y^\tau$  is the signal displaced by a lag  $\tau$ . It is mathematically defined as:

$$I(X; Y^\tau) = \sum_{x_n \in \mathcal{X}} \sum_{y_{n-\tau} \in \mathcal{Y}} p(x_n, y_{n-\tau}) \log_a \frac{p(x_n, y_{n-\tau})}{p(x_n)p(y_{n-\tau})} \quad (2)$$

According to (Cover and Thomas, 2012) the channel capacity (C) represents the maximum measure of Mutual Information:

$$C = \max I(X, Y) \quad (3)$$

and according to (Proakis and Salehi, 2001) the channel capacity for DMI is quantified by its peak value. Also according to (Proakis and Salehi, 2001), the transmission rate estimation (R) can be written as a function of channel capacity and signal bandwidth (BW) in Hertz:

$$R = 2.BW.C \quad (4)$$

If the entropy is measured in bits, the transmission rate is going to be measured as bits/s.

## 2.2 Surrogate for Hypothesis Test

The investigation of information sharing between brain regions sometimes require the assertion of statistical significance for confidence in the functional analysis performed. In this case Surrogate may be useful (Lancaster et al., 2018).

One of the techniques to apply this method is the IAAFT technique, proposed by (Schreiber and Schmitz, 1996). It consists of generating a surrogate data from the original signal, keeping the same power spectrum and randomizing Fourier phases, creating uncorrelated signals. Applying DMI on surrogate data represents the measures that are expected when signals do not share any connectivity. Therefore, when comparing the statistics from original signals and the surrogate data statistics, the null hypothesis can be accepted or rejected. Some literature can be reviewed in neuroscience related to the use of Surrogate to assist the type of interdependency among electroencephalographic signals (Pereda et al., 2005; Faes et al., 2010; Subramaniyam and Hyttinen, 2015; Adkinson et al., 2019).

Related with the number of surrogate data to be created, there is a well established rank-order test, proposed by (Theiler et al., 1992), that can be used (Schreiber and Schmitz, 2000). First, assume  $\Psi$  is the probability of false rejection, then, define the level of significance ( $S$ ) as:

$$S = (1 - \Psi) \cdot 100\% \quad (5)$$

The number of surrogate data to be created ( $M$ ) is defined as follows:

$$M = \frac{K}{\Psi} - 1 \quad (6)$$

where  $K$  is an integer number defined by the type of test - 1 if it is one-sided and 2 in the case of a two-sided test. Usually,  $K = 1$  is adopted due to computational effort to generate surrogates (Schreiber and Schmitz, 2000).

## 3 METHODOLOGY

The section details the methodology used to analyse the local field potential signals, in order to assess mutual information, channel capacity and transmission rate between brain areas. Furthermore, it describes the methodology employed to acquire the LFP signals and the computational environment.

## 3.1 Applied Methodology

The diagram presented in Figure 1 depicts the applied methodology. The first step — blue box in Figure 1 — is to calculate the LFP signals entropy for cortex (Cx), hippocampus (Hp) and thalamus (Th).

Then, in the second step, the optimal number of bins to apply Mutual Information is determined — orange box in Figure 1 — as follows: The rodents LFP signals are discretized with different number of bins — 2, 4, 8, 16, 32, 64 and 128 where used — and DMI is applied among brain areas. Next, the DMI measures are compared to find the best number of bins - the DMI curve for a given number of bins that is closest to highest DMI curve. Figure 2 explain better the method. Another important measure performed is the Kolmogorov-Smirnov test to compare the group of rodents to check if there is a statistical difference between groups.

The third step is to apply Delayed Mutual Information to understand the information sharing and depict the lag where there the maximum value occurs, therefore determining the channel capacity. In the fourth step, surrogate data is created with Iterative Amplitude Adjusted Fourier Transform (IAAFT) algorithm for cortex, hippocampus and thalamus signals. In this case study, 35 signals for 97% significance level were generated. In the fifth step, each surrogate is combined with other two original signals to perform DMI analysis, investigating the connectivity significance between brain areas.

The surrogate data is compared with the original DMI measures by means of Kolmogorov-Smirnov (KS) test. In this paper, p-value = 5% was chosen to evaluate the maximum mutual information and its values. The last step is to apply Fast Fourier Transform to verify the signal's bandwidth and, finally, calculate the transmission rate between brain areas.

The applied methodology described was used to investigate two periods of LFP signals: basal and convulsant drug infusion, until generalized tonic-clonic (TCG) behaviour. The analysis was performed in the recordings of five rats.

Simulations were developed in Python language, using the packages: Matplotlib, Nolitsa, Numpy, Pandas, Scipy, Seaborn and Time. The code was executed on a computer with an Intel i7 processor, 8GB of RAM, running MAC OS 10.14.6 operational system.

## 3.2 Database for Case Study

We used LFP signals database from Interventional Laboratory of Neuroengineering and Neuroscience



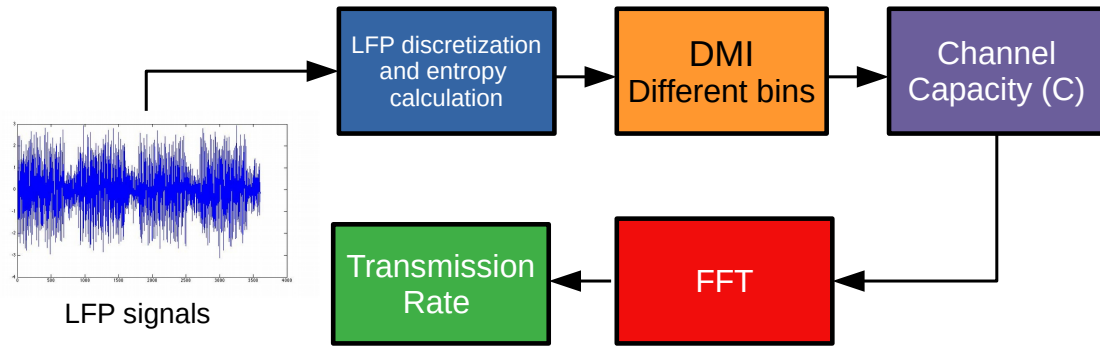


Figure 1: Summary of applied methodology: First, the LFP signals are discretized and their entropy are calculated. Then, DMI with different bin sizes are computed to determine the value that best fits the dataset. After the optimal number of bins is identified, DMI is calculated among all rodent brain signals (Cx, Hp and Th) and channel capacities are determined. To calculate the transmission rate between brain areas, Fast Fourier Transform (FFT) is applied to all LFP signals to check the signal's bandwidth. Finally, the channel capacity and signal's bandwidth are used to find the transmission rate between brain areas.

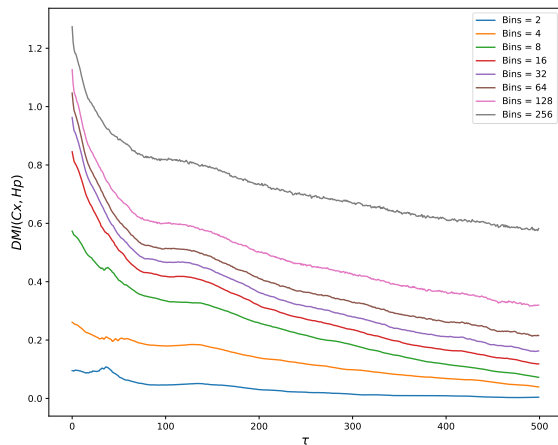


Figure 2: Delayed Mutual Information performed with different number of bins between two brain areas ( $\tau$  is given in samples). As can be observed the highest curve is presented for 256 bins. In this case the better resolution for DMI is 256 bins. However, if the 256 bins curve would not exist, only the other curves, it is possible to check three curves with almost same values (32, 64 and 128 bins). In that case it could be possible to choose 32 bits to perform analysis because the results are similar to 128 bins curve and the computational processing would be smaller.

(LINNce) from Federal University of São João Del Rei. The laboratory employs male Wistar rats weighing between 250 and 350 grams coming from the University Central Biotherm to acquire data and evaluate methods of electrical stimulation. All described procedures are in according to ethics committee under-protocol 31/2014.

The signal recording is conducted with the aid of electrodes (monopolar type and stainless steel covered by teflon) placed directly inside the right thalamus and hippocampus of the rat's brain through

stereotactic surgery (Cota et al., 2016). In addition, two microsurgical screws were implanted (length 4.7 mm, diameter 1.17 mm, Fine Science Tools, Inc., North Vancouver, Canada) aiming the cortical registration of right hemisphere and to operate as reference in frontal bone. The electrodes and screws were positioned with assistance of neuroanatomic atlas (Paxinos and Watson, 2013).

The LFP signals for each rodent was registered while the subject was simultaneously filmed, to perform behavioural analysis (observe classic seizure features such as facial automatisms, myoclonic concussion, head myoclonus, anterior and posterior limbs myoclonus, elevation and fall, generalized tonic-clonic seizure) to allow their correlation with the electrophysiological events observed during LFP recording. For all rodents the time of recording was the same with ten minutes of duration.

LFP recording was performed using 1 kHz sampling rate. Signals were amplified 2000 V/V through A-M Systems (model 3500) amplification system and digitalized on National Instruments (PCI 6023E) A/D converter controlled by developed LINNce Virtual Instrument from LabView platform. Sequentially, they were filtered using a second-order Butterworth filter (0.3 to 300 Hz band). The power grid noise at 60 Hz frequency was mitigated with use of shielded cables and Faraday cage.

## 4 RESULTS

During basal and infusion intervals, stationarity was observed, allowing the calculation of Shannon entropy during each part of the signal. The Tables 1 and

2 display the entropy values (in bits) for all rodents used in the case study during basal and infusion intervals respectively. To find the optimal number of bins for Delayed Mutual Information different numbers of bins were tested on the LFP signals and it provided the number of 256.

The Kolmogorov-Smirnov test performed with rodents groups indicated that there is no difference between the groups at the level of p value equals 10%. Next, Surrogate method was applied. Figure 3 depicts an example of the power spectrum of original signal in comparison to the power spectrum of the surrogate data.

DMI was calculated for surrogate data and original signals. An example result of  $DMI(Cx, Hp)$  for rodent R048 can be seen in Figure 4. The lag with maximum mutual information can be also observed:  $\tau = 0$  for all rodents used in this case study. The Tables 5 and 6 exhibit the signals bandwidths for each rodent during basal and infusion intervals respectively.

Tables 3 and 7 show, respectively, all channel capacities and transmission rates, simulated during basal interval. The similar results during infusion interval can be verified in Tables 4 and 8. Boxplots in Figures 5 and 6 depict the mean and standard deviation of transmission rate between each brain area considering all the five rodents used to develop the case study.

The computational time to perform this case study was approximately 3 hours for each DMI. Three DMI between each pair of brain areas were computed, resulting in a total of 9 hours. Therefore, for each interval, basal and infusion, it was spent 27 hours, reaching 54 hours total.

Table 1: Entropy calculated for all rodents used in the case study during basal interval. DMI is given in bits.

Rodent	Cx Entropy	Hp Entropy	Th Entropy
R048	15.36	15.36	15.36
R052	13.28	13.28	13.28
R064	14.96	14.97	14.97
R065	15.05	15.05	15.05
R072	15.05	15.05	15.05

## 5 DISCUSSION

The entropy for all rodents used in this case study during basal interval were similar with exception of rodent R052 which was slightly lower. During infusion interval, the entropy was more uniform and

Table 2: Entropy calculated for all rodents used in the case study during infusion interval. DMI is given in bits.

Rodent	Cx Entropy	Hp Entropy	Th Entropy
R048	17.39	17.39	17.39
R052	17.02	17.02	17.02
R064	16.62	16.62	16.62
R065	17.14	17.14	17.14
R072	17.29	17.29	17.29

Table 3: Channel capacity in bits for each rodent during basal interval.

Rodent	Cx $\rightarrow$ Hp	Cx $\rightarrow$ Th	Hp $\rightarrow$ Th
R048	1.28	1.42	1.75
R052	3.00	2.93	3.72
R064	0.61	0.65	1.48
R065	1.53	1.51	1.25
R072	2.45	2.40	2.15

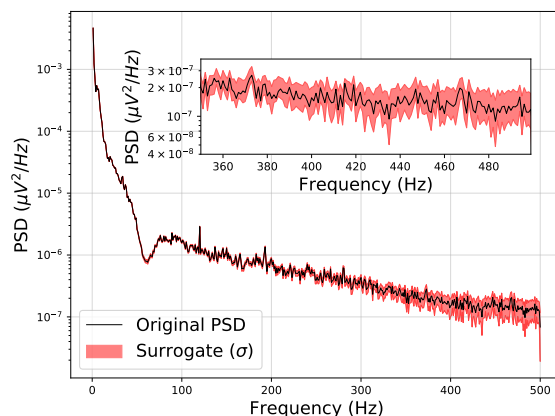


Figure 3: Power spectrum of original cortex signal and surrogate data of cortex area for rodent R048. The power density spectrum of the synthetic data is approximately the same of the original signal.

higher than calculated for basal interval. This indicates that LFP signals become more probabilistic during infusion, meeting the hypothesis of increase of information rate, which in turn is compliant with the perspective of understanding epileptic seizure as a hyper-synchronisation phenomena.

It is possible to check that signal bandwidth during infusion interval is bigger than the bandwidth during basal interval. This indicates initially larger information for each brain area during infusion, again indicating the increasing communication among cortex, hippocampus and thalamus.

The lag with maximum MI was zero, found for all rodents in this case study. Probably, this result is related with the low sampling rate of LFP signals. All surrogate data created for this case study

Table 4: Channel capacity in bits for each rodent during infusion interval.

Rodent	Cx → Hp	Cx → Th	Hp → Th
R048	0.91	1.26	1.10
R052	1.42	1.51	2.81
R064	0.70	0.71	1.47
R065	1.29	1.42	1.09
R072	2.61	2.72	2.29

Table 5: Bandwidth (BW) in Hertz for each rodent during basal interval.

Rodent	BW Cx	BW Hp	BW Th
R048	1.30	1.49	1.23
R052	1.73	1.71	1.71
R064	9.00	10.50	8.71
R065	10.00	10.00	10.00
R072	2.77	2.75	2.74

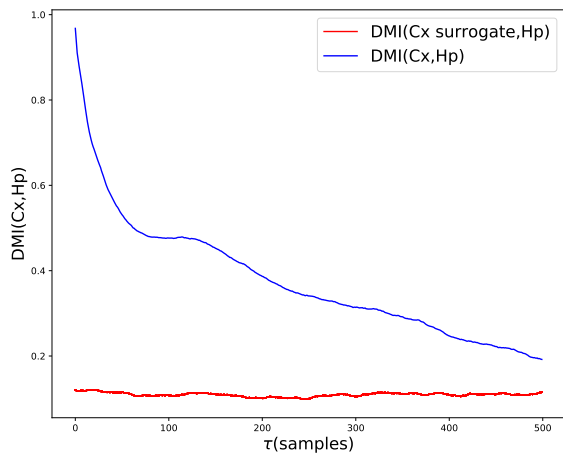


Figure 4: Delayed mutual information between cortex and hippocampus for rodent R048 using 256 bins. The blue line is the DMI performed with original signals and the red lines are the DMI with surrogate. Due to the difference between the surrogate DMI it is not possible to distinguish the difference among surrogate data DMI because it was approximately the same. It is important to note that the lag for maximum mutual information is zero, the same found for another rodents used to perform the case study.

maintained the power spectrum and randomized the Fourier phases, as expected for IAAFT algorithm, as illustrated in Figure 3. This created uncorrelated signals, with whose measures, it was possible to validate the DMI measures for original signals. For all rodents, it was possible to observe that the measures of the original signals were higher than surrogate data measures.

DMI revealed different measures for each rodent even Kolmogorov-Smirnov test indicated statistical equality among groups for 10% significance. Yet,

Table 6: Bandwidth (BW) in Hertz for each rodent during infusion interval.

Rodent	BW Cx	BW Hp	BW Th
R048	30.00	20.00	20.00
R052	20.00	13.95	10.43
R064	26.00	33.00	36.00
R065	30.00	27.00	30.00
R072	20.26	15.00	20.48

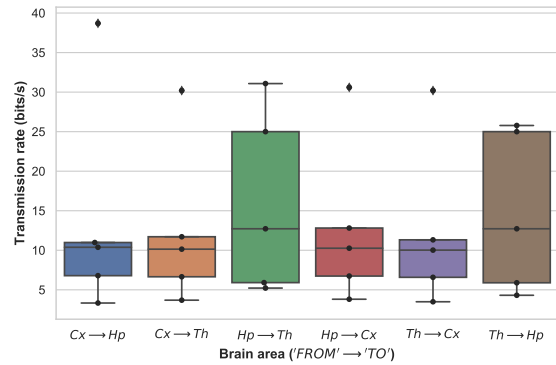


Figure 5: Boxplot denoting the transmission rate in bits/s between each brain areas, for all the 5 rodents used to perform the case study, during basal interval (black dots represent the rodents measures). It is important to observe that the highest standard deviation is verified in transmission rate between hippocampus and thalamus. The black dots represent the rodents in the boxplot.

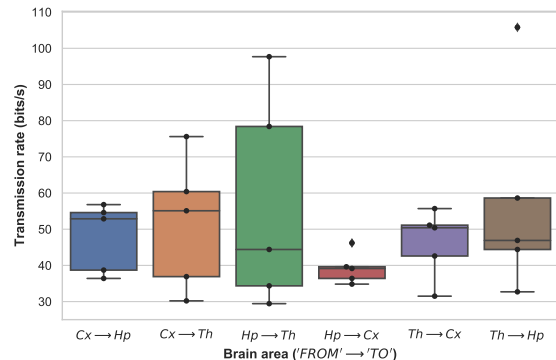


Figure 6: Boxplot denoting the transmission rate in bits/s between each brain areas, for all the 5 rodents used to perform case study, during infusion interval (black dots represent the rodents measures). The values during infusion interval are higher when compared with values measured during basal interval. A possible reason for that difference is a larger signal's bandwidth. The DMI is approximately the same among brain areas for both intervals, see Tables 3 and 4. The black dot represent the rodents in boxplot.

during basal and interval the values for each rodent was approximately the same, indicating that the channel capacity did not change during basal and infusion intervals. Nonetheless, the bandwidth was different between them, resulting in different transmission

Table 7: Transmission rate ( $R$ ) for each rodent during basal interval. Base 2 was used for DMI logarithm. Frequency was measured in Hertz, resulting in  $C$  being provided in bits and  $R$  in bits/s.

Rodent	Cx $\rightarrow$ Hp	Cx $\rightarrow$ Th	Hp $\rightarrow$ Th	Hp $\rightarrow$ Cx	Th $\rightarrow$ Cx	Th $\rightarrow$ Hp
R048	3.33	3.69	5.22	3.81	3.49	4.31
R052	10.38	10.14	12.72	10.26	10.02	12.72
R064	10.98	11.71	31.08	12.81	11.32	25.78
R065	38.70	30.2	25.00	30.6	30.2	25.00
R072	6.79	6.65	5.91	6.74	6.58	5.89

Table 8: Transmission rate ( $R$ ) for each rodent during infusion interval. Base 2 was used for DMI logarithm. Frequency was measured in Hertz, resulting in  $C$  being provided in bits and  $R$  in bits/s.

Rodent	Cx $\rightarrow$ Hp	Cx $\rightarrow$ Th	Hp $\rightarrow$ Th	Hp $\rightarrow$ Cx	Th $\rightarrow$ Cx	Th $\rightarrow$ Hp
R048	54.60	75.60	44.40	36.40	50.40	44.40
R052	56.80	60.40	78.40	39.62	31.50	58.62
R064	36.40	36.90	97.68	46.20	51.12	105.80
R065	38.70	30.20	29.43	34.83	42.60	32.70
R072	52.88	55.10	34.35	39.15	55.70	46.90

rates as can be observed in Tables 7 and 8.

It is important to make clear that surrogate analysis assured 5% p-value using Kolmogorov-Smirnov test indicating that significance level of comparison between original signals DMI and surrogates DMI for each rodent. When the results are compared between rodents the results presented 10% p-value significance level when Kolmogorov-Smirnov test was applied. That is the reason for different Kolmogorov-Smirnov p-values presented in this paper.

The increase of transmission rate, observed for all rodents during infusion interval — in Figures 5, 6 and Tables 7, 8 — is an important result, meeting the concept of hyper-synchronization phenomena that appears during epileptic seizures. DMI was able to capture the increase of communication among brain areas. Since there is more volume of information moving between brain areas during infusion, the uncertainty (entropy) accordingly grows, as expected. The channel capacity practically do not change between intervals, pointing that the main reason for hyper-synchronization is the growth of information sharing between cortex, hippocampus and thalamus.

Delayed Mutual Information represented an another appropriate alternative to existing methods because it was able to provide insights about the functional connectivity among brain areas. It is a non linear method supporting the real world condition that is not linear, do not require a large volume of data when compared to another methods such as Transfer Entropy and it is non parametric which means more flexibility to data analysis.

## 6 CONCLUSIONS

The use of DMI to perform a case study with rodents LFP presented insights about the communication among brain areas before the occurrence of an epileptic seizure. It was observed that entropy during infusion interval was higher than during basal interval, being the first indicator that the communication was increasing during infusion. The uncertainty about signals was becoming higher. The second indicator was the growth of LFP bandwidth for all signals. It was also observed a consistent lag of zero for DMI for all rodents. This last result may have being influenced by the relatively low sampling rate used to record the LFP signals. The verified channel capacity was different for each rodent, however, exhibited the same behaviour of staying approximately equal during basal and infusion intervals. Consequently, the transmission rate was different between periods mainly due to the change of signal's bandwidth. It indicates that communication is increasing essentially due to the growth of information sharing among brain areas. This is the last indicator, which is in agreement with the idea of hyper-synchronisation phenomena associated with epileptic seizure. Therefore, DMI represented a helpful method to perform functional analysis on LFP signals.

## ACKNOWLEDGEMENT

This work was supported by the Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG) [grant number APQ 02485-15] and financed in part by

the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

## REFERENCES

- Adkinson, J. A., Karumuri, B., Hutson, T. N., Liu, R., Alamoudi, O., Vlachos, I., and Iasemidis, L. (2019). Connectivity and centrality characteristics of the epileptogenic focus using directed network analysis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(1):22–30.
- Akbarian, B. and Erfanian, A. (2017). Automatic detection of ptz-induced seizures based on functional brain connectivity network in rats. In *2017 8th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 576–579. IEEE.
- Anastasiadou, M. N., Christodoulakis, M., Papathanasiou, E. S., Papacostas, S. S., Hadjipapas, A., and Mitsis, G. D. (2019). Graph theoretical characteristics of eeg-based functional brain networks in patients with epilepsy: The effect of reference choice and volume conduction. *Frontiers in Neuroscience*, 13:221.
- Andrea Avena-Koenigsberger, Bratislav Mšic, O. S. (2017). Communication dynamics in complex brain networks. *Nature Reviews Neuroscience*, 19:17 – 33.
- Bartolomei, F., Lagarde, S., Wendling, F., McGonigal, A., Jirsa, V., Guye, M., and Bnár, C. (2017). Defining epileptogenic networks: Contribution of seeg and signal analysis. *Epilepsia*, 58(7):1131–1147.
- Beghi, E. (2019). Social functions and socioeconomic vulnerability in epilepsy. *Epilepsy & Behavior*.
- Beghi, E., Giussani, G., Abd-Allah, F., Abdela, J., Abdelalim, A., Abraha, H. N., Adib, M. G., Agrawal, S., Alahdab, F., Awasthi, A., et al. (2019). Global, regional, and national burden of epilepsy, 1990–2016: a systematic analysis for the global burden of disease study 2016. *The Lancet Neurology*, 18(4):357–375.
- Berglind, F., Andersson, M., and Kokaia, M. (2018). Dynamic interaction of local and transhemispheric networks is necessary for progressive intensification of hippocampal seizures. *Scientific reports*, 8(1):5669.
- Biasiucci, A., Franceschiello, B., and Murray, M. M. (2019). Electroencephalography. *Current Biology*, 29(3):R80–R85.
- Chapeton, J. I., Inati, S. K., and Zaghoul, K. A. (2017). Stable functional networks exhibit consistent timing in the human brain. *Brain*, 140(3):628–640.
- Cota, V., Marcela Bacellar Drabowski, B., Carla de Oliveira, J., and Moraes, M. (2016). The epileptic amygdala: Toward the development of a neural prosthesis by temporally coded electrical stimulation. *Journal of Neuroscience Research*, 94:463–485.
- Cover, T. M. and Thomas, J. A. (2012). *Elements of information theory*. John Wiley & Sons.
- Devinsky, O., Vezzani, A., O'Brien, T. J., Scheffer, I. E., Curtis, M., and Perucca, P. (2018). *Epilepsy. Nature Reviews Disease Primers*, 4.
- Endo, W., Santos, F. P., Simpson, D., Maciel, C. D., and Newland, P. L. (2015). Delayed mutual information infers patterns of synaptic connectivity in a proprioceptive neural network. *Journal of computational neuroscience*, 38(2):427–438.
- Faes, L., Porta, A., and Nollo, G. (2010). Testing frequency-domain causality in multivariate time series. *IEEE Transactions on Biomedical Engineering*, 57(8):1897–1906.
- Gribkova, E. D., Ibrahim, B. A., and Llano, D. A. (2018). A novel mutual information estimator to measure spike train correlations in a model thalamocortical network. *Journal of neurophysiology*, 120(6):2730–2744.
- Ibrahim, F., Abd-Elateif El-Gindy, S., El-Dolil, S. M., El-Fishawy, A. S., El-Rabaie, E.-S. M., Dessouky, M. I., Eldokany, I. M., Alotaiby, T. N., Alshebeili, S. A., and Abd El-Samie, F. E. (2019). A statistical framework for eeg channel selection and seizure prediction on mobile. *International Journal of Speech Technology*, 22(1):191–203.
- Kim, H.-R., Go, H.-J., and Kim, S.-Y. (2018). Discrimination of mild alzheimers disease patients using cluster analysis of information transmission in eeg. *Journal of the Korean Physical Society*, 73(3):377–387.
- Lancaster, G., Iatsenko, D., Pidde, A., Ticcinelli, V., and Stefanovska, A. (2018). Surrogate data for hypothesis testing of physical systems. *Physics Reports*, 748:1 – 60. Surrogate data for hypothesis testing of physical systems.
- Li, S., Xiao, Y., Zhou, D., and Cai, D. (2018). Causal inference in nonlinear systems: Granger causality versus time-delayed mutual information. *Physical Review E*, 97(5):052216.
- Li, S., Xu, J., Chen, G., Lin, L., Zhou, D., and Cai, D. (2017). The characterization of hippocampal theta-driving neurons a time-delayed mutual information approach. *Scientific reports*, 7(1):5637.
- Mei, T., Wei, X., Chen, Z., Tian, X., Dong, N., Li, D., and Zhou, Y. (2019). Epileptic foci localization based on mapping the synchronization of dynamic brain network. *BMC medical informatics and decision making*, 19(1):19.
- Nichols, J., Seaver, M., Trickey, S., Todd, M., Olson, C., and Overbey, L. (2005). Detecting nonlinearity in structural systems using the transfer entropy. *Physical Review E*, 72(4):046217.
- Niriayo, Y. L., Mamo, A., Gidey, K., and Demoz, G. T. (2019). Medication belief and adherence among patients with epilepsy. *Behavioural neurology*, 2019.
- Olamat, A. E. and Akan, A. (2017). Synchronization analysis of epilepsy data using global field synchronization. In *2017 25th Signal Processing and Communications Applications Conference (SIU)*, pages 1–4.
- Organization, W. H. et al. (2017). Atlas: country resources for neurological disorders 2004. geneva: World health organization; 2004.
- Paxinos, G. and Watson, C. (2013). *The Rat Brain in Stereotaxic Coordinates*. Elsevier. 7th edition.
- Pereda, E., Quiroga, R. Q., and Bhattacharya, J. (2005).

- Nonlinear multivariate analysis of neurophysiological signals. *Progress in Neurobiology*, 77(1):1 – 37.
- Proakis, J. G. and Salehi, M. (2001). *Digital communications*, volume 4. McGraw-hill New York.
- Schreiber, T. and Schmitz, A. (1996). Improved surrogate data for nonlinearity tests. *Phys. Rev. Lett.*, 77:635–638.
- Schreiber, T. and Schmitz, A. (2000). Surrogate time series. *Physica D: Nonlinear Phenomena*, 142(3-4):346–382.
- Spicciarich, M. C., von Gaudecker, J. R., Jurasek, L., Clarke, D. F., Burneo, J., and Vidaurre, J. (2019). Global health and epilepsy: Update and future directions. *Current neurology and neuroscience reports*, 19(6):30.
- Subramaniam, N. P. and Hyttinen, J. (2015). Dynamics of intracranial electroencephalographic recordings from epilepsy patients using univariate and bivariate recurrence networks. *Phys. Rev. E*, 91:022927.
- Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., and Farmer, J. D. (1992). Testing for nonlinearity in time series: the method of surrogate data. *Physica D: Nonlinear Phenomena*, 58(1):77 – 94.
- Wan, X. and Xu, L. (2018). A study for multiscale information transfer measures based on conditional mutual information. *PloS one*, 13(12):e0208423.
- Weiss, S. A., Staba, R., Bragin, A., Moxon, K., Sperling, M., Avoli, M., and Engel, J. (2019). Interneurons and principal cell firing in human limbic areas at focal seizure onset. *Neurobiology of Disease*, 124:183 – 188.
- Yu, H., Zhu, L., Cai, L., Wang, J., Liu, C., Shi, N., and Liu, J. (2019). Variation of functional brain connectivity in epileptic seizures: an eeg analysis with cross-frequency phase synchronization. *Cognitive Neurodynamics*, pages 1–15.