EPILEPTIC SEIZURES and the EEG

Measurement, Models, Detection and Prediction

ANDREA VARSAVSKY
IVEN MAREELS
MARK COOK

CRC Press
Taylor & Francis Group

EPILEPTIC SEIZURES and the EEG

Measurement, Models, Detection and Prediction

EPILEPTIC SEIZURES and the EEG

Measurement, Models, Detection and Prediction

ANDREA VARSAVSKY
IVEN MAREELS
MARK COOK

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number: 978-1-4398-1200-6 (Hardback)

Library of Congress Cataloging-in-Publication Data

Varsavsky, Andrea.
Epileptic seizures and the EEG : measurement, models, detection and prediction / Andrea Varsavsky, Iven Mareels, Mark Cook.
p. ; cm.
Includes bibliographical references and index.
ISBN 978-1-4398-1200-6 (hardcover : alk. paper)
1. Epilepsy--Diagnosis. 2. Electroencephalography. I. Mareels, Iven, 1959- II. Cook, Mark, 1960- III. Title.
[DNLM: 1. Epilepsy--diagnosis. 2. Electroencephalography. WL 385 V325e 2010]

RC373.V27 2010
616.8'5307547--dc22 2010012764

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Contents

List of Figures **xi**

Preface **xv**

1 Introduction **1**
1.1 The Brain and Epilepsy 2
1.1.1 Micro-Scopic Dynamics: Single Neurons 5
1.1.2 Meso/Macro-Scopic Dynamics: Neural Networks . . . 8
1.1.2.1 Cortico-Cortical Projections 10
1.1.2.2 Thalamo-Cortical Projections 11
1.1.3 Neurotransmitters and Neuromodulators 12
1.1.4 Epilepsy – A Malfunctioning Brain 13
1.1.4.1 Focal Epilepsy – Failure of Meso-Scopic Networks 14
1.1.4.2 Non-Focal Epilepsy 16
1.1.4.3 Continuous Epilepsy 17
1.1.5 Diagnosis and Treatment of Epilepsy 17
1.1.5.1 Anti-Epileptic Drugs 18
1.1.5.2 Surgical Resection 18
1.1.5.3 Electrical Stimulation 19
1.2 The EEG – A Recording of the Brain 20
1.2.1 The Normal EEG 22
1.2.2 The Epileptic EEG 24
1.2.3 Detecting Changes in the EEG 27
1.3 Dynamics of the Brain 28
1.3.1 Micro- and Macro-Scopic Models 30
1.3.2 Dynamic Models of Epilepsy 32
1.4 Stochasticity in Neural Systems 33
1.5 Conclusions and Further Reading 35

2 EEG Generation and Measurement **37**
2.1 Principles of Bioelectric Phenomena 42
2.1.1 A Foreword on Notation 42
2.1.2 From Single Charges to Equivalent Dipoles 44
2.1.3 Equivalent Current Dipoles 47
2.1.4 Macro-Scopic Mean Fields – Homogeneous Media . . . 48
2.1.5 Macro-Scopic Mean Fields – Inhomogeneous Media . . 49

2.2 Current Sources in Biological Tissue 50
2.2.1 Synaptic Structure and Current Dipoles 50
2.2.2 Spatial Integration 53
2.2.2.1 Cortical Structure 54
2.2.2.2 Cortical Folds 56
2.2.3 Temporal Integration 58
2.3 Volume Conducting Properties of the Head 60
2.3.1 Head Geometry 60
2.3.2 Capacitive Effects of Tissue 63
2.3.3 Estimating Conductivities 65
2.3.3.1 Brain 66
2.3.3.2 CSF 66
2.3.3.3 Skull 67
2.3.3.4 Scalp 67
2.4 The EEG: A Macro-Scopic View of the Brain 67
2.4.1 EEG Measurement 68
2.4.1.1 Cortical (Intra-Cranial) Recordings 72
2.4.1.2 Scalp Recordings 72
2.4.1.3 The Search for an Ideal Reference 73
2.4.1.4 Spatial Filtering Properties of the Skull . . . 75
2.4.2 EEG Dynamics 78
2.4.3 Epilepsy and the EEG 80
2.5 Conclusions 81
2.A Units of Electric Quantities 84
2.B Volume Conductor Boundary Conditions 84
2.C Capacitance in RC Circuits 86

3 Signal Processing in EEG Analysis **89**
3.1 Mathematical Representation of the EEG 91
3.2 Preprocessing 93
3.3 Feature Extraction 96
3.3.0.1 Computing Statistics: Averages vs. Instances 96
3.3.0.2 Noise 99
3.3.0.3 Stationarity and Windowing 100
3.3.0.4 Linearity, Non-Linearity, Determinism and Stochasticity 103
3.3.0.5 Normalization 104
3.3.1 Time Domain Analysis 105
3.3.1.1 Signal Amplitude (Energy) and Variance (Power) 106
3.3.1.2 Periodicity (Auto-Correlation) 108
3.3.1.3 Synchronization 115
3.3.2 Frequency Domain Analysis 120
3.3.3 Time-Frequency Analysis 131
3.3.4 Non-Linear Analysis 136

3.3.4.1 Embedding Theory . . . 137
3.3.4.2 Dimension – How Complex is a System? . . 139
3.3.4.3 Lyapunov Exponents – How Predictable is a System? . . . 141
3.3.4.4 Entropy – How Random is the System? . . . 142
3.3.4.5 Non-Linear Dynamics and Analysis of the Epileptic EEG . . . 147
3.4 Detection and Prediction of Seizures in Literature . . . 149
3.5 Conclusions . . . 153

4 Classifying the EEG **155**
4.1 Types of Classifiers . . . 156
4.1.1 Association Rules . . . 157
4.1.2 Artificial Neural Networks . . . 158
4.1.3 Support Vector Machines . . . 163
4.2 Expert System . . . 166
4.2.1 Processing Decisions . . . 167
4.2.2 Spatio-Temporal Context . . . 169
4.2.3 Patient Specificity . . . 170
4.3 Conclusions . . . 171

5 Seizure Detection **173**
5.1 The Problem of Seizure Detection . . . 175
5.1.1 The EEG Database . . . 176
5.1.1.1 Group 1 – Scalp EEG Data (< 6 Seizures per Patient) . . . 178
5.1.1.2 Group 2 – Scalp EEG Data (6 − 10 Seizures per Patient) . . . 178
5.1.1.3 Group 3 – Scalp EEG Data, Non-Epileptic Patients . . . 178
5.1.1.4 Group 4 – Intra-Cranial EEG Data . . . 179
5.1.2 Performance Evaluation Metrics . . . 180
5.2 Evaluation of Classification Methods . . . 184
5.2.1 Feature Extraction . . . 185
5.2.2 ANN Training and Testing . . . 185
5.2.3 SVM Training and Testing . . . 188
5.2.4 Results and Comparisons . . . 188
5.3 Evaluation of Patient Un-Specific Seizure Detectors . . . 190
5.3.1 Algorithm 1: `Monitor` . . . 191
5.3.1.1 Algorithm Description . . . 191
5.3.1.2 Results . . . 193
5.3.2 Algorithm 2: `CNet` . . . 194
5.3.2.1 Algorithm Description . . . 194
5.3.2.2 Results . . . 196
5.3.3 Algorithm 3: `Reveal` . . . 197

5.3.3.1 Algorithm Description 197
5.3.3.2 Results 197
5.3.4 Algorithm 4: `Saab` 199
5.3.4.1 Algorithm Description 199
5.3.4.2 Results 201
5.3.5 Comparisons and Conclusions 202
5.4 Evaluation of Onset Seizure Detectors 204
5.4.1 Feature Extraction 204
5.4.1.1 Cross Correlation (XCORR) 207
5.4.1.2 Power Spectral Density (PSD) 207
5.4.1.3 Wavelet Analysis (WAV) 208
5.4.1.4 Correlation Dimension (CD) 208
5.4.2 Results and Comparisons 208
5.5 Conclusions 214

6 Modeling for Epilepsy 215
6.1 Physiological Parameters of Neural Models 219
6.1.1 Parameters in Single Neurons 221
6.1.2 Parameters in Networks of Neurons 221
6.2 Micro-Scopic (Statistical) Models 223
6.2.1 Model Summary 224
6.2.2 Validation and Limitations 228
6.3 Meso-Scopic (Phenomenological) Models 230
6.3.1 Model Summary 231
6.3.2 Analysis: Linearization, Stability and Instability . . . 234
6.3.3 Validation and Limitations: Rhythms in the EEG . . 241
6.3.3.1 Simulating the Normal EEG 241
6.3.3.2 Simulating the Seizure EEG 243
6.3.3.3 Caution 245
6.3.4 Relationship to Micro-Scopic Models 246
6.4 Macro-Scopic Models (Future Outlook) 247
6.5 Practical Use of Models 249
6.5.1 Epileptic Seizure Generation 250
6.5.1.1 Seizure Initiation 250
6.5.1.2 Seizure Termination by Electrical Stimulation 252
6.5.2 Limitations of the EEG 254
6.6 Conclusions 256
6.A Physiological Parameters and Notation 257
6.B Summary of IF Model 258
6.C Summary of Phenomenological Model 259

7 On the Predictability of Seizures **263**
7.1 Predictability – Terminology Made Clear 269
7.2 How to Estimate LRD 274
7.2.1 Example Distributions 274
7.2.2 Computing α 276
7.2.3 Simulations 281
7.2.4 Results 282
7.3 Seizure Frequency Dataset 287
7.4 Analysis – Estimation of α 291
7.5 Memory and Predictability of Seizures 300
7.6 Conclusions 303

8 Concluding Remarks **305**

Glossary **309**

Bibliography **321**

Index **339**

List of Figures

1.1 Systems and signals of different scales 2
1.2 A brain system and a measurement system 3
1.3 Functional de-composition of the entire brain 4
1.4 Pyramidal neurons and neural networks 6
1.5 Neuron system, inputs and outputs 7
1.6 Cortical column 9
1.7 Thalamo-cortical loop 11
1.8 International 10-20 placement system of scalp EEG electrodes 22
1.9 Example of 'normal' EEG traces 23
1.10 Example of epileptic EEG traces 25
1.11 Intra- and inter-scale interactions in a neural model 32
1.12 Model of initiation and generalization of epileptic seizures . . 34

2.1 Effects of measurement precision 38
2.2 The different scales of the brain 40
2.3 History of the EEG 41
2.4 Vectors and scalars 43
2.5 Electric and current dipoles 46
2.6 Current loops in the cortex 51
2.7 Effects of randomly aligned vs. parallel sources 55
2.8 Effects of dipoles in the sulci of the cortex 57
2.9 Effects of synchronous vs. asynchronous sources 58
2.10 The 4-sphere approximate model of the human head 61
2.11 Spherical harmonics 62
2.12 Theoretical electrical properties of biological tissues 64
2.13 Equivalent circuit representation of linear volume conductor . 65
2.14 Contributions of dipoles on recording location $\mathbf{r}_R$ 69
2.15 Relative contributions of dipoles to cortical and scalp recordings 70
2.16 Strength and contributions of dipoles to the EEG 71
2.17 Spatial transfer function for a bipolar reference 76
2.18 Boundary conditions of adjoining volume conductors 85

3.1 The advantage of looking at averages 97
3.2 Equality in distribution 98
3.3 Arbitrary definitions of noise 100
3.4 Overlapping and non-overlapping windows 101
3.5 The effects of applying a window over an analysis window . . 102

3.6 Example time-domain analysis of the epileptic EEG, part 1 . 109
3.6 Example time-domain analysis of the epileptic EEG, part 2 . 110
3.7 Auto-correlation calculations, part 1 112
3.7 Auto-correlation calculations, part 2 113
3.8 Auto-correlation on time-varying signals 114
3.9 Synchronization calculations, part 1 116
3.9 Synchronization calculations, part 2 117
3.10 The power of using FFT 121
3.11 The effects of using different windows when computing the FFT 122
3.12 Aliasing 123
3.13 Effects of averaging on FFT calculations 124
3.14 PSD of non-stationary signals 125
3.15 Example PSDs for EEG sequences, part 1 127
3.15 Example PSDs for EEG sequences, part 2 128
3.16 Wavelet fundamentals 132
3.17 Wavelet decimation 134
3.18 Example applications of wavelet filters 135
3.19 Example correlation integral calculation 141
3.20 Example Lyapunov exponent calculation 143
3.21 Coarse-graining in the computation of entropy 144
3.22 Example entropy calculation 146

4.1 General structure of seizure detection algorithms 155
4.2 Artificial neural networks 159
4.3 SVM classification problems 164
4.4 Example expert classifier combinations 168

5.1 EEG equipment 177
5.2 Definition of TP and FP 180
5.3 A sample ROC curve 184
5.4 Evaluation of classifiers: Individual patient performance, part 1 186
5.4 Evaluation of classifiers: Individual patient performance, part 2 187
5.5 Evaluation of classifiers: Average performance 188
5.6 `Monitor` detection algorithm summary 192
5.7 `Monitor` ROC 193
5.8 `CNet` detection algorithm summary 195
5.9 `CNet` ROC 196
5.10 `Reveal` detection algorithm summary 198
5.11 `Reveal` ROC 199
5.12 `Saab` detection algorithm summary 200
5.13 `Saab` ROC 201
5.14 Onset detection ROC 209
5.15 Onset delays, Part 1 210
5.15 Onset delays, Part 2 211

6.1 Neural models at different scales 218
6.2 Pseudo-matrix representation of aggregate models 226
6.3 IF model results 229
6.4 Typical relationship between firing rate and mean membrane potential 232
6.5 Nonlinear vs. linearized models 235
6.6 Transfer function representation of the meso-scopic model . . 236
6.7 Stability of linear systems 238
6.8 Stability boundaries of the meso-scopic model 239
6.9 The importance of the delay t_0 between cortex and sub-cortex 240
6.10 Example simulated waveforms corresponding to typical EEG signals 242
6.11 Example simulated waveforms corresponding to different types of instabilities 244
6.12 Spectra of model simulated waveforms 245
6.13 Macro-scopic models derived from coupled sub-systems 248
6.14 Sample transition in and out of system stability 251
6.15 Example transitions out of seizure with the application of an external electrical stimulus 253

7.1 Measurement precision and predictability 265
7.2 Self-similarity in the Mandelbrot set 270
7.3 Simulated time series 277
7.4 Inter-event probability histograms of simulated time series . . 278
7.5 Estimates of α on simulated time series 282
7.6 Robustness of estimated α to changes in resolution 283
7.7 Robustness of estimated α to random removal of events . . . 284
7.8 Robustness of estimated α to nonstationarities 285
7.9 The inter-event times for Datasets 1-6, part 1 289
7.9 The inter-event times for Datasets 1-6, part 2 290
7.10 Inter-event probability histograms for each of the 6 analyzed datasets 292
7.11 Calculations of y_m using different wavelet order 293
7.12 Calculations of α for Datasets 1-6, part 1 295
7.12 Calculations of α for Datasets 1-6, part 2 296
7.13 Stationarity analysis for Datasets 1, 5 and 6 298
7.14 Random removal of events for Datasets 5 and 6 299
7.15 LRD in strength of events 301

Preface

> "If there is any great secret of success in life, it lies in the ability to put yourself in the other person's place and to see things from his point of view - as well as your own."
>
> *- Henry Ford, (1863-1947)*

Biomedical engineering is not the 'new' discipline it is often quoted to be. Electric currents in biology were recognized long before electric currents in wires. Medicine has relied on mechanical, electrical and more recently electronic equipment almost as a matter of course. Engineering in therapy is not a new concept either. Passive mechanical implants such as hip-joint replacement are recorded as early as the 1890s. Functional electrical stimulation can be traced back to the 1700s, and active implants such as artificial pace-makers have been delivering electrical stimuli to human hearts since the 1950s. More sophisticated systems like the bionic ear were first used to restore auditory function to the deaf not long after.

What is true is that the volume of *research* in biomedical engineering has grown tremendously over the past couple of decades. The technological developments that allowed large-scale processing of data, coupled with the political shift that sees more money available for anything bio-technology, mean that the demand for suitably qualified researchers in this field escalated very rapidly. Unfortunately the training of such researchers has not kept up with this demand. It is true that there are plenty of biologists and plenty of engineers around, but to develop biologically applicable technology, particularly in the case where devices must interface directly with the biology, it is necessary to have people that are both.

This book is a look at one particular area of biomedical engineering: understanding the brain from a 'signals' perspective. The expertise and theory necessary to make progress in understanding the brain are well established in both neurophysiology and engineering alike, but the communication link that would allow the portability of knowledge from one discipline to the other (we contend) is not. The combined field of *neuro-engineering* is in its infancy. Fast development of neuro-engineering practices is hindered by a *lack of communication* or an understanding of 'common ground'. Neurophysiologists are not trained in the mathematical tools commonly used by the electrical engineer, whilst the engineer is typically not trained in the biology of human tissue. Furthermore both disciplines have in the last century evolved almost

entirely separately, developing their own scientific practices, idiosyncrasies, terminology and, yes, even prejudices.

Sometimes misunderstandings arise simply because the scientific practices of each discipline are different. For example, imagine that both neurophysiologists and engineers are trying to gather information so that they can understand how the brain works. Neurophysiologists are interested in as complete an understanding as possible. The more detail they can get hold of, the better. In contrast engineers are interested in the amount of detail that can be thrown out. How much detail they keep depends on the level of understanding they need for a particular application. *Fundamentally different approaches are taken by each discipline to answer the same question.*

The silos of neurophysiology and electrical engineering are very real – the jargon that a medical doctor speaks so naturally can dumbfound the engineer, whilst concepts that are fundamental to training an engineer leave the biologist at a loss. *If the combined field is to progress more rapidly than the time it takes to train a new generation of multi-disciplinary minded researchers, then communication barriers must be broken down.*

As engineers we must ask ourselves: how do we make our field more accessible to the neurologist? Mathematics and computation are very useful, but what is the point if it is not understood?

As neurophysiologists we must communicate ideas by using language that is more accessible, either by selecting important information and conveying it in simpler terms, or through education of the broader community.

Communication is the key: *An inter-disciplinary experiment must be designed in an inter-disciplinary environment.* Engineers and neurophysiologists must communicate their needs to each other, and the assumptions made must be consistent.

This book subscribes to the idea that accessibility is key in a multi-disciplinary research environment. Through the study of *epilepsy*, a common neurological pathology or disorder, we present the relevant physiology and electromagnetics of biological systems. We describe the tools available (along with their limitations) for the analysis of neurophysiological data. Both engineering and neurophysiological terminology are kept to a minimum so that anyone belonging to either discipline, and perhaps people new to both, can make sense of the information. Concepts are favored above detail so that ideas are not lost in a sea of information. Suitable references are provided to explore detail.

Epilepsy is the chosen context because, aside from practical reasons (this

is the area that the authors are familiar with), research in epilepsy holds common interest for both electrical engineers and neurophysiologists alike. It is a neurological condition in which the brain behaves 'normally' most of the time but occasionally breaks down to an altered state of consciousness known as a *seizure* or *fit*. Physical convulsions are the most widely recognized form of such seizures, but consciousness can be impaired in other forms including hallucinations and black-outs. Epilepsy is particularly well suited for consideration because it holds:

1. *High social relevance*: Epilepsy occurs in all age groups, with higher incidences in infancy and senility. Causes are numerous and include genetic or developmental abnormalities, trauma and disease – the common denominator being malfunction of brain activity. It is estimated that 1% of the population suffers from an epileptic episode at some point in their lives, although epileptic people are only the 0.6% with recurring symptoms. About 25% of epileptics cannot be helped by any drug or therapy available today. Perhaps the statistics seem small but epilepsy is the most common recurrent neurological disorder today. Assuming it affects the 6.5 billion (6.5×10^9) in the world uniformly, this means that 39 million suffer recurring symptoms, and 9.25 million cannot lead normal lives because treatment is not available to them. In reality the incidence of epilepsy in developing countries is likely to be greater. The direct costs of epilepsy treatment are estimated to be about AUD$80 billion per year world-wide[1].

2. *Interest to the neurophysiologist*: Neurophysiologists have been trying to figure out the causes, mechanisms and treatment of epilepsy for millenia. Throughout history virtually everything has been blamed for epilepsy. Seizures have been attributed to benign events such as the phases of the moon or disappointment in love affairs alike. Most often, though, it was the supernatural that was blamed. In ancient Greece epilepsy became known as the 'sacred disease' because it was believed that seizures were sent from the devil and the associated visions were sent by the gods. The word epilepsy was named after the Greek *epilepsia*, meaning 'a condition of being overcome, seized or attacked'. In Roman times epilepsy was known as *passio caduca* – 'falling sickness' or 'falling evil'. The stricken were condemned as sorcerers. Treatments were frivolous or religious, and it was inappropriate preparation or impurities of the mind that were often blamed for their failure to cure the dis-

[1] This figure is based on the extrapolation of findings in [16], which estimated the direct costs (medical, drug treatment, surgery, etc) on the epileptic population of Australia based on a survey performed in 1989. Cost per patient per year was estimated to be roughly AUD$2000. It is understood that costs of medication and treatment – as well as the availability of treatment – vary significantly between countries, and this figure is only used to give an idea of the economic burden. No world estimate of the cost of epilepsy could be found, but the figure seems a reasonable average of the international comparisons found in [84]. Indirect costs such as loss of productivity are not included in this estimate.

order. It was not until the 17th century that all aspects of epilepsy were attributed to brain malfunction, initiating the change toward finding the differences between an epileptic and a healthy human brain. Even so, scientific testing and classification of the many forms of epilepsy did not eventuate until the 19th century when conclusions were finally based on controlled experiments and repeated observations [125, 180]. Today modern technology has allowed a greater understanding of the processes of epilepsy, but much remains to be discovered.

3. *Interest to the engineer*: Engineers are drawn to the study of epilepsy because much of the processing can be performed on measurements (such as the EEG, described later) without the explicit involvement of wet labs. Abundant routine medical data of this nature exist and the burden to design new experiments is lessened. Thus although a lot of new information must be absorbed before the engineer can fully grasp the problem, the learning process can be gradual and there is no need for a major shift in his/her own practices. In addition, because the brain is so complex its analysis provides the opportunity to use state-of-the-art tools. Not to be underestimated are the social aspects of the problem that provide greater motivation for those who may find typical applications such as telecommunications or computer chip development a little on the dry side. The same tools that are learned for these other problems can be used toward understanding the brain, provided the underlying assumptions are revised.

So, being an appealing project to both engineers and neurologists alike, and in addition a socially relevant one, epilepsy is neither short in funding nor interested parties. Before delving into specifics some important concepts that may help in understanding much of this book must be introduced. It may be useful to refer to Figure 1.1 and Figure 1.2, which give graphical representation of what follows.

Throughout this book we will constantly refer to *systems* and *signals*[2]. In order to focus the terminology, we start from the notion of a signal as the fundamental concept. A signal is a convenient way to summarize or point to a collection of measurements. In mathematical terms a signal is a function of time, for example an EEG (described in Section 1.2). At each instance of time, a voltage (or a collection of voltages if we use multiple electrodes) is recorded. The sequence of all the measurements is a signal. We say that the signal is *scalar* valued if we only have one measurement for each time index, and that it is *vector* valued if we have multiple measurements (that is, multiple electrodes) at each instance in time.

The EEG is used to tell us something about a brain; the brain in our terminology is a *system* from which the EEG signal is observed or derived. The collection of all possible EEGs from a brain is called its *behavior*. More

[2]We use a 'behavioral' terminology as introduced in [136]. See also [104].

generally, a system constrains the signals that may be observed from it, and a system's behavior is the collection of all the signals compatible with the system. Although hopelessly general at this level, the framework of signals and systems is very powerful to understand relationships and interdependencies.

Signals originating within a system are *internal*, but a system often needs to communicate with entities outside of itself. These *external signals* are known as its *inputs* and *outputs*, corresponding to in-coming and out-going information respectively. For example, the input to the brain as a whole is the sensory information obtained from the environment and the rest of the body (sight, hearing, pain, taste, etc). The outputs are the messages the brain sends to the body enabling it to move, speak and react.

A *model* of a system is an attempt to formalize the way the inputs and outputs interact so that behavior may be *computed*. In a mathematical model this involves the construction of equations that describe the behavior of the system. Often the model is a simplified representation of the original system because much of the detail can be omitted, in the hope that these details are un-important in the process of interest. In this way a lot of the complexity is removed from a problem whilst the relevant information is retained. To create a model of the brain requires understanding of how the components within this system work. However, if modeling the brain were an easy task this book would most likely not exist. The task can be simplified by first creating models of *sub-systems* that exist in the brain – that is, smaller systems within a larger one, like a model of the cortex or the hippocampus which are subsystems of the larger system formed by the brain.. Smaller systems may be candidates for sub-systems of the larger ones. This is valid so long as the fact that these sub-systems belong to a larger one is not forgotten.

A model may be *deterministic* or *stochastic*. In a purely deterministic model, once the current conditions are determined, everything about the past, present and future of the system is known unambiguously. This is an unrealistic situation, which in the real world exists ... well, never. A stochastic model, on the other hand, allows for fluctuations around its solutions which cannot be accounted for before they occur. These are *random* or *stochastic* elements that can make prediction of the future of this system difficult. Stochasticity can exist not only in the model but within the sources of the brain and the measurement of these sources, as discussed in more detail later.

This book is a study of the brain as a system. We measure this system using the EEG, and we use these signals to understand its behavior. We also create models that describe the brain as seen through this signal. But what can the EEG really tell us about what is happening within the brain? A typical EEG machine records up to 64 channels, 512 times a second, each with 14 bits resolution. This means that the EEG records roughly

$$64 \times 14 \text{ bits/sample} \times 512 \text{ samples/second} \approx 10^6 \text{ bits/second}.$$

Now, assuming there are 100 billion neurons (10^{11}) in the brain, and that it

can be divided into cortical columns each containing 10^5 neurons (see Chapter 1), then we have a system which can be described with roughly 10^6 states. This means that the EEG gives us $\frac{10^6 \text{ bits/second}}{10^6} = 1$ bit of information per second per state! That is 1 bit of information per second to describe the activity of the 10^5 neurons within a cortical column! This admittedly crude analysis already gives us some idea that the EEG is a very blunt measurement of a very complex system, and it should come as no surprise that what it can tell us about the brain is severely limited. The general argument here is formalized toward the end of Chapter 6.

In this book we expand on this very simple number game and explore the usefulness as well as the limitations of the EEG. We focus specifically on the problems of both seizure detection and prediction. We make an effort to provide conceptual information in all aspects of the problem, beginning with the physiology and physics involved in the generation and measurement of activity, and then use this knowledge to develop strategies to address a problem.

The text is organized as follows. Chapter 1 summarizes the physiology and the fundamental ideas behind the measurement, analysis and modeling of the epileptic brain. We introduce the EEG as a measured signal, and explain its use in the study of epilepsy.

Chapter 2 provides an explanation of the type of brain activity likely to register in EEG measurements. It expands on qualitative ideas presented in Chapter 1 by providing quantitative analysis of the populations of neurons that contribute to both scalp and cortical EEG. At the same time the limitations and the effects that choices made in the recording process have on the data are discussed.

Chapter 3 then provides an overview of how these EEG records are and have been analyzed in the past. The applicable engineering and signal processing methods are numerous. The scope is narrowed by concentrating on the mathematics relevant to the problem of classification of EEG. Chapter 4 then deals with using these extracted features to differentiate between or *classify* inter-seizure, pre-seizure and seizure EEG. This material is applicable to the detection as well as prediction of seizures.

Chapter 5 expands on Chapter 3 by concentrating only on the problem of seizure detection, that is, the differentiation between the seizure and non-seizure EEG. It is a simpler problem than seizure prediction which has been in development for over 20 years, but one that nevertheless requires further attention. Lack of standardization makes published results difficult to compare and understand. A broad scope review of present algorithms is applied to a *common EEG data set* (to our knowledge the first time this has been done).

Although suitable for the task of detection, the ideas presented in Chapter 5 are *black box* methods that are data driven and therefore do not provide information beyond what is available in the recordings. For more complex problems that require understanding of underlying mechanisms of the brain,

a survey of the physiologically based *dynamic models* of brain activity is presented in Chapter 6.

Finally, Chapter 7 addresses the fundamental question: *can seizures be predicted?* Much research assumes that seizures are predictable, yet little work has been dedicated to the validation of this assumption. It is proposed in this chapter, through the analysis of epileptic activity spanning from 3 hours to 25 years, that seizures *are* predictable but the amount of data required is far greater than previously thought (if we rely only on the EEG), and that the measurements being used are not suitable. Furthermore, it is proposed that the problem of seizure prediction is only likely to succeed if approached differently than has been done to date. For example we may find it useful to add information from a different type of measurement, or by employing physiologically based methods such as those presented in Chapter 6.

We started as a small group of research engineers at The University of Melbourne, Australia, working together with neurophysiologists at St Vincent's Hospital, Melbourne. Our question was simple: what can signal processing contribute to understanding the epileptic brain? We were naturally interested in the EEG as a measurement early on – it is so often used clinically in the diagnosis of epilepsy that it is not surprising it is also used in many aspects of research. As time passed we became increasingly interested in its limitations, and were surprised to find how little these are understood. This is how this book was conceived, and we hope it is used by engineers and neurophysiologists alike to understand this incredibly useful signal.

Andrea Varsavsky
BEng., BSc., PhD.
The University of Melbourne, Australia

Iven Mareels
BEng., PhD.
Prof. of Electrical and Electronic Engineering
The University of Melbourne, Australia

Mark Cook
MD, MBBS, FRACP
Prof. of Neurology
The University of Melbourne
St Vincent's Hospital, Melbourne, Australia

1

Introduction

The brain is a very complex *system* composed of billions of interconnected neurons. As observers we understand the brain system from the *measurements* or *signals* we obtain from it. If the spatial scale at which measurements are made is very small, the measurement could pertain to a single neural cell. In contrast, at larger scales the measurement pertains to large collection of neurons. In Figure 1.1 the issue of spatial scale is illustrated.

In this book we concentrate on the EEG as the measurement that is used to acquire a signal. The EEG records the time-evolving voltages generated by brain activity, and is described in Section 1.2. However the measured signal is not necessarily the real signal generated by the brain, but its *projection* onto the recording equipment. Thus the measurement process is a system itself, with its *inputs* arising from the *outputs* of the original system, and its output the resultant measured data. This is shown in Figure 1.2. *A system as a whole is not restricted to the generating system (the brain), but must incorporate the measurement system (the EEG).* Analysis of the measurements can provide insight into function and dysfunction of the original system, so long as the recording process itself is well understood.

This chapter is dedicated to presenting concepts necessary to understand the brain as a system. In Section 1.1 the physiology of the brain relevant to epilepsy is summarized. Section 1.2 then concentrates on the measurement and analysis of this brain activity, and Section 1.3 summarizes how physiology and measurement can be turned into a mathematical model of brain dynamics. Section 1.4 discusses how the presence of stochastic elements affects EEG sources, measurement, analysis and modeling.

Each section is introductory only because more detail is included in later chapters, with the exception of Section 1.1 where we attempt to contain all relevant (albeit simplified) physiology. All sections focus on epilepsy and the specific problem of differentiation between seizure and non-seizure activity.

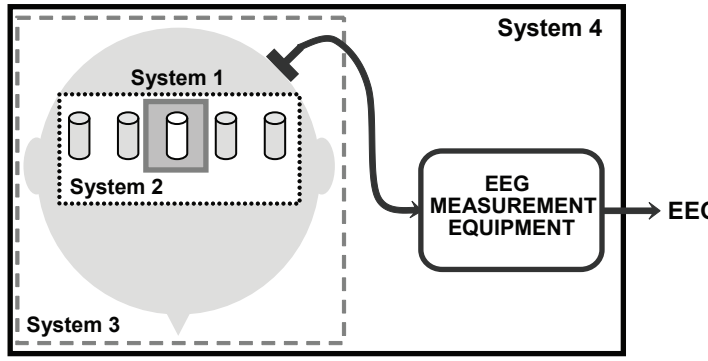


FIGURE 1.1: A *system* is defined by the *measured signals* and thus the scale of the system depends on the scales of activity that significantly affect this measurement. In this figure example systems of different scales are shown. If a measurement of the brain only involves activity of single neurons, then the smallest of these systems may be used. However in a signal such as the EEG larger scales must be incorporated, including the effects of the entire head, the recording equipment and maybe even environmental factors not shown. The smaller *sub-systems* may be used to explain the larger system.

1.1 The Brain and Epilepsy

The brain is part of the *central nervous system* (CNS) and is responsible for interpreting sensory information received from the environment so that humans can behave as humans do. Each region of the brain has its own task in this process. Figure 1.3(a) shows a functional de-composition of the *cerebral cortex* – a thin layer approximately 2-3mm thick that covers the entire surface of the brain. The different functional regions of the cortex (temporal lobe, parietal lobe, occipital lobe, frontal lobe) are responsible for motor control as well as cognitive and memory functions. Sensory information is passed on to the cortex from a subcortical system known as the *thalamus*, shown in Figure 1.3(b). The thalamus also plays an important role in regulating the interaction between different regions of the brain.

In the order of 10-100 billion densely interconnected nerve cells, called *neurons*, make up the cerebral cortex. How individual neurons work is understood quite well, but it is the complex ways in which they inter-connect and interact that determine brain function. How these *networks* function is less well understood.

The general structure of these networks is illustrated in Figure 1.3(b) and (c). All mammalian brains look roughly like this, although the details vary. For example neurons in the cerebral cortex of humans are much more densely

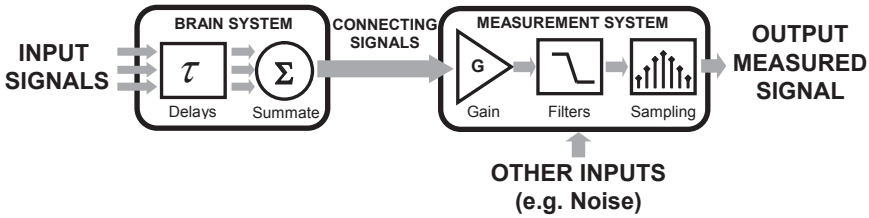


FIGURE 1.2: Above is an example of a *generating system* - the brain – and *measurement system* such as the EEG. In this over-simplistic representation of the brain system, *signals* that act as *inputs* are integrated together after some delay. The integration is the *output* of the generating system that becomes the input to the measurement system. After amplification, filtering and digitization the output of the measurement system is a record of the activity in the generating system. Shown also are the *internal signals* within each sub-system.

inter-connected than in animals, and is believed to be one reason for the more sophisticated capabilities of humans. The *gray matter* in this figure is the cortex, folded to form *gyri* and *sulci* and containing all the cortical neurons. Between cortex and subcortex is the *white matter* – a region composed mostly of connections made between different areas of the brain. The majority of these connections are between different cortical regions, but subcortical systems such as the thalamus also communicate with the cortex through the white matter. Very few neurons can be found in this region.

The way by which the brain works can be described in terms of behavior at different *spatial scales*, traditionally divided into the *micro-scopic*, *meso-scopic* and *macro-scopic*. In this book micro-scopic describes behavior at small scales (μm), encompassing a single or a few brain cells. Macro-scopic describes behavior at large scales (cm), spanning whole regions of brain. The intermediate scale, meso-scopic, describes behavior of networks of neurons spanning millimeters rather than centimeters. In particular the word is used to describe *cortical columns*, believed to be the main functional units of the cortex and described in more detail in Section 1.1.2.

Brain function can also be described in terms of different *temporal scales*. These are directly correlated to the spatial scales because of naturally occurring conditions in neural dynamics. The micro-scopic scale is generally associated with frequencies above 1000Hz because the mechanisms within single neurons are very fast. The meso-scopic scales are associated with activity between 10-1000Hz because faster events are negligible relative to the average behavior of ensembles of neurons. Activity at the macro-scopic level is associated with frequencies in the range of 1-100Hz because spatially averaging an even larger number of neurons filters out higher frequencies. These ranges are

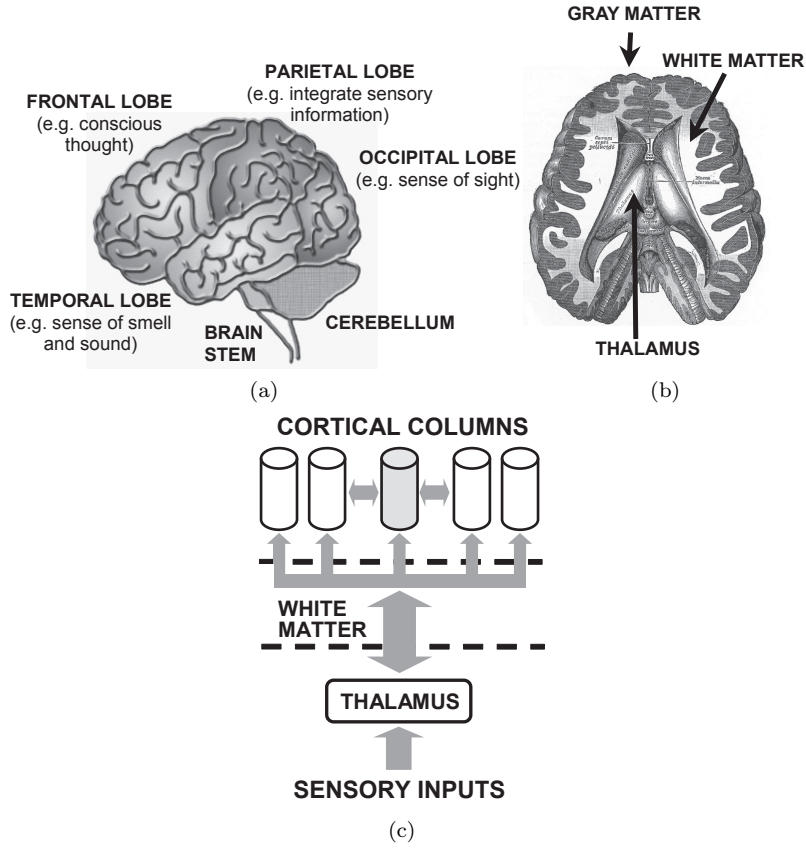


FIGURE 1.3: A functional de-composition of the entire brain. (a) shows how the *cerebral cortex* is divided into the four lobes, each responsible for different cognitive and motor functions. (b) is a cross-section of the brain showing some of the major subcortical systems, the most important for this book being the *thalamus* whose principal role is to relay sensory information onto the cortex. Also note the relative size of *gray matter* – where the neurons are – and the *white matter* – used for connections between sub-systems. In (c) the flow of information to and from one *cortical column* (described in Section 1.1.2) and other regions of the brain is shown. Notice that the inputs to this column come from other cortical columns as well as sub-cortical systems. Most connections are projected through the white matter, although nearby columns are connected more directly. The graphic in (b) is a modified reproduction of a figure taken from *Anatomy of the Human Body* ([56]), originally published in 1918 and lapsed into the public domain.

not rigid, but are used as a guideline in the expected behavior, as explained in more detail in Chapter 2.

Epilepsy is a problem of scale. On the one hand it is a macro-scopic phenomenon that encompasses a large portion if not the whole of the brain. On the other hand, epilepsy's root cause must be found in the chemokinetic processes that may be associated with meso- or micro-scopic cellular biology. In addition, understanding the macro-scopic may not be possible without knowledge of some of the micro-scopic. The relevant information must be retained. For example, knowing that gating mechanisms that transfer information between neurons is important for the understanding of macro-scopic brain activity, but knowing the individual complex protein structures involved in the different gate types is unlikely to help us understand macro-scopic recordings, and this complexity can be omitted for the purposes of this problem.

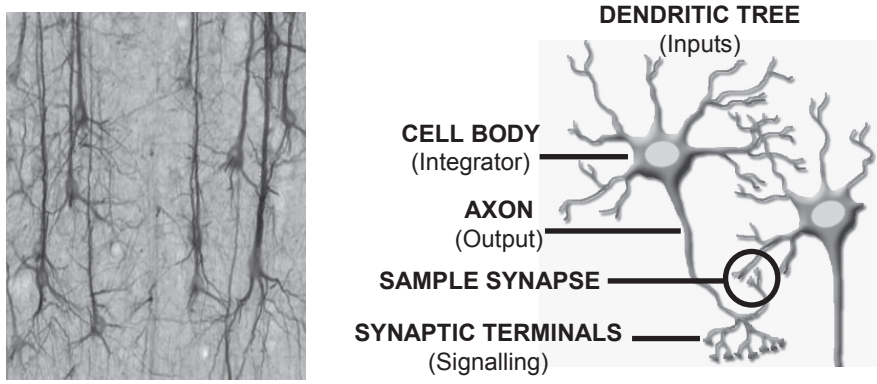
The remainder of this section provides a crash course into how the brain works at each scale, limited to the mechanisms that are *believed* to contribute to the understanding of epilepsy at the macro-scopic scale. 'Believed' is used here because epilepsy is not completely understood, and elements that are at this point thought more or less irrelevant may become relevant in the future. Wherever possible efforts are made to avoid superfluous use of medical jargon so that important points are not lost in the translation process. Of course some terminology is always necessary. In any case, as this introduction is by necessity brief and limited to the bare minimum needed to proceed with an understanding of the EEG system, readers are encouraged to obtain a detailed understanding of cellular mechanisms. An excellent introductory text is [80], although any other basic physiology book may be used.

1.1.1 Micro-Scopic Dynamics: Single Neurons

It is the *neurons* in the CNS that are responsible for the processing and transmission of information, but they are not the only type of cell present. *Glial* cells in the cortex outnumber neurons by just under 4 to 1, but their role is not as clearly understood. They are believed to be responsible for support roles such as provision of structure, insulation and maintenance [80]. More recent studies reveal that the role of glial cells may not be so passive, but in any case their contributions to neural function are more or less ignored. Since it is likely that their complete function will continue to be unknown for some time they are largely ignored in this text, although one should remember that they exist.

Figure 1.4(a) shows a picture of a stained cortical slice in which many neural cells are visible. These are known as *pyramidal neurons* and are the most common nerve cells found in the cortex. A structural decomposition of a typical pyramidal neuron is shown in (b). Neurons come in many different shapes and sizes, but they are composed of four basic structures – *dendrites*, *soma* or *cell body*, *axon* and *synaptic terminals*.

The *inputs* to a neuron are *chemical currents* (electrically charged) that



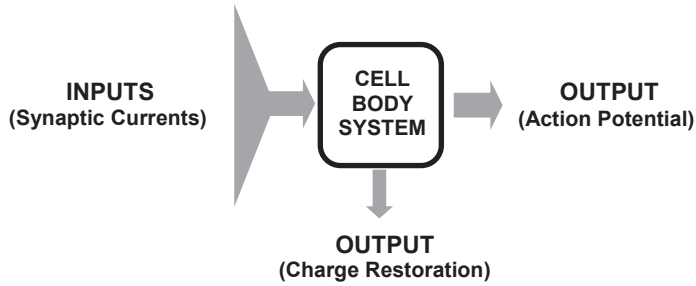
(a) (b)

FIGURE 1.4: A typical pyramidal neuron found in the cerebral cortex. In (a) is a stain obtained from cortical tissue which shows about 1% of neurons in the region, taken from `brainmaps.org` and reproduced freely using the *Creative Commons Attribution 3.0 license* (`http://creativecommons.org/licenses/by/3.0/`). (b) is a structural decomposition of neurons, showing the input regions at the *synapses* on the *dendritic tree*. These inputs are integrated at the *soma* or *cell body* and result in an output on the *axon*.

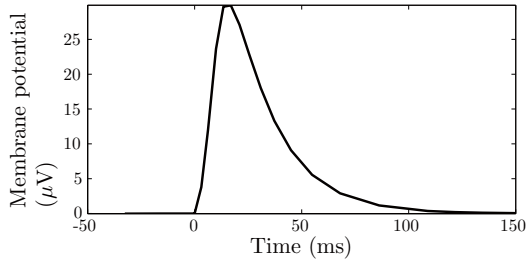
occur at the synaptic terminals spread across the dendritic tree. These currents are transmitted down to the cell body, which then accumulates all these inputs together. If the integration reaches a threshold voltage, then an *action potential* – another chemical current much larger in magnitude than any single synaptic input – is fired and travels through the axon. If this threshold is not reached, no action potential occurs. Action potentials are the *output* of the neuron, which in turn closely approach or *synapse* to the dendrites of other neurons, and thus become the input to the receiving cells. A system representation of a single neuron is shown in Figure 1.5. Here the role of the cell body is more than that of a pure integrator of dendritic inputs as it is also responsible for restoring charge concentrations in the extra-cellular fluid.

Repeated crossing of the threshold voltage at the soma can generate multiple action potentials at a maximum rate of about 1000Hz. Action potentials travel at about $5 - 10$ meters per second [146] and do not deteriorate along the axon due to built in regenerative processes. The shape of the voltage differences created by a typical action potential between the inside and outside of the axon is shown in Figure 1.5(c).

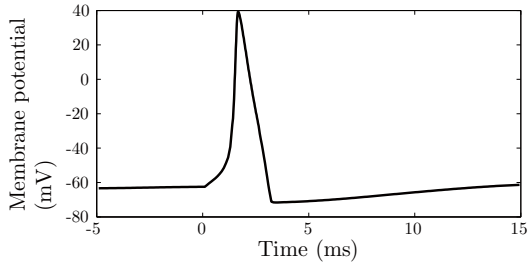
Synaptic transmission is the process by which action potentials arriving at the end of the axon of a transmitting (*pre-synaptic*) neuron are interpreted by the dendrites of the receiving cell (*post-synaptic* neuron). The pre-synaptic cell releases chemicals known as *neurotransmitters* that control the response



(a)



(b) Neuron input: Simulated EPSP



(c) Neuron output: Simulated action potential

FIGURE 1.5: In (a) the features of a real neuron shown in Figure 1.4(b) are used to construct an equivalent system model in which the inputs and the outputs are defined. The functionality of the cell body system can be described by mathematical equations that dictate how the outputs (e.g., action potentials, restoring charge concentration in extra-cellular fluid) react to the given inputs. In (b) and (c) are examples of how a Hodgkin-Huxley model equations (a system in the form of (a)) can be used to simulate input and output signals. These are only representative waveforms; much variability exists in their shape. The input (in this case an *excitatory post-synaptic potential* or EPSP, in (b)) is much longer in duration than the output (an *action potential*, in (c)), but also much smaller in amplitude. Many PSPs must be integrated at the soma for an action potential to be fired.

of the post-synaptic cell. The type of neurotransmitter released varies but these chemicals are responsible for generating a *post-synaptic potential* (PSP) that can fall in one of two categories:

- *Excitatory post-synaptic potential (EPSP)*: The response at the synapse is much like a miniature version of an action potential in the post-synaptic cell, except longer in duration and smaller in amplitude. An EPSP is shown relative to the action potential in Figure 1.5. The generation of an EPSP brings the soma closer to threshold and thus increases the likelihood that an action potential is fired.
- *Inhibitory post-synaptic potential (IPSP)*: Opposite in effect to an EPSP. The probability of a resultant action potential decreases.

A single IPSP typically has a larger effect than a single EPSP because inhibitory synapses tend to form closer to the soma. However the total number of EPSPs is greater than IPSPs and thus the effects level out. The balance of incoming IPSPs and EPSPs on a single neuron determines whether the post-synaptic cell fires an action potential – it is *relative* numbers of active EPSPs versus IPSPs that matter[1].

Action potentials and PSPs are possible only because of the *cell membrane* – the material that separates internal fluids from the *cerebrospinal fluid* (CSF) surrounding the cell. Under normal conditions the cell membrane is impermeable to fluids. Action potentials and PSPs occur because of the presence of *gates* that selectively allow the transmission of ions through this membrane. These gates open via different mechanisms. One example are the *voltage-gated ion gates* that react to differences in voltages between the inside and the outside of the neuron. When threshold is reached the gates open to allow ions to flow, and the generation of an action potential ensues. At the synapses the gating triggers are chemical and it is the presence of a neurotransmitter that allow the gates to open. Many other different types of gates exist, including in-built mechanisms that restore chemical balance after a gate is opened. These types are omitted here for simplicity, and synaptic transmission is simplified to the promotion or inhibition of a resultant action potential. More information can be found in [80].

1.1.2 Meso/Macro-Scopic Dynamics: Neural Networks

In the cerebral cortex two types of neurons form 90% of the total population: the *pyramidal cells* shown in Figure 1.4 and others known as *inter-neurons*. The axons of pyramidal cells form only excitatory synapses with other neurons, and the inter-neurons only inhibitory ones. The massive interconnectivity between these means that a single neuron may be transmitting and receiving

[1] This is a simplified picture of how neurons fire. In reality, depending on the type of neurons, many different behaviors can be observed in response to EPSPs and IPSPs, including single, multiple, or burst firing.

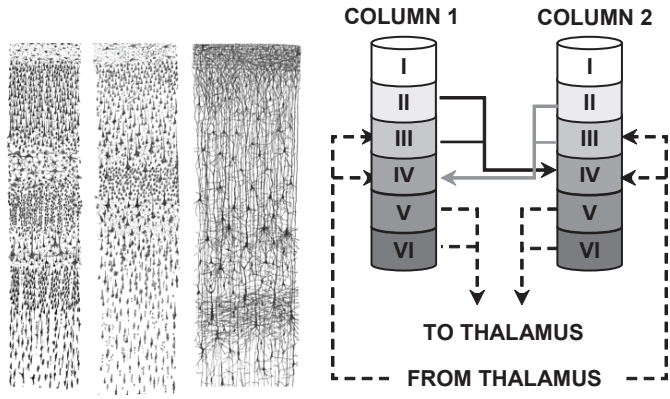


FIGURE 1.6: The layered structure of cortical columns. (a) shows a reproduction of a stain demonstrating the stratified manner in which a typical cortical column is organized. Each layer has different neural densities and types. In (b) the cortico-cortical and thalamo-cortical projections are shown. The layered structure is believed responsible for the organization of the inputs and outputs of the cortex. Notice the high levels of redundancy built in – the thalamus may communicate with a column directly, or indirectly through another column. Strong connections also exist between the layers of the column but these are not shown here. In practice, a cortical column is often treated as a structurally uniform system with the same inputs and outputs as shown, but without specification as to where along the column these signals connect to.

information to and from thousands of other neurons. These may be in the order of $1,000$-$10,000$ synapses in the dendritic tree of a pyramidal neuron, its axon projecting to a similar number [168]. For this reason, no single EPSP or IPSP is the determining factor for a post-synaptic action potential. The cell body integrates all incoming signals.

Because the connections determine behavior in the brain it is worthwhile understanding a little about its structure. Looking at the cortex, 1mm^3 of tissue contains $\sim 50,000$ neurons and in the order of $\sim 3 \times 10^8$ synapses, 84% of them excitatory and the rest inhibitory [168]. About 70% of cortical neurons are excitatory pyramidal neurons, 20% inhibitory inter-neurons and the remainder other types that are most often also excitatory. The under-representation of inhibition in the histology of a typical neural tissue sample should not be interpreted as a lower level of inhibition because the effects of IPSPs can be much larger than EPSPs. In any case the statistics are approximate and vary with location.

The numbers alone do not draw a complete picture – the organization of the cortex is not at all random. The volume of $1mm^3$ used above is representative size of what is known as a *cortical column* – a compartment within which all neurons are connected to almost every other neuron, but with relatively few connections projecting outside of it. A cortical column is believed to be the basic functional unit of the brain, although in reality they are not discrete volumes with sharp boundaries.

Connections between cortical columns are called *cortico-cortical projections*. To a lesser extent connections are also made to subcortical networks, most importantly the *thalamus* responsible for relaying sensory information to and from the cortex, among other things. These are called *thalamo-cortical projections*. Cortico-cortical projections outnumber thalamo-cortical projections by a ratio of 100 to 1. Figure 1.7 shows the possible connections made between a cortical column and thalamus.

Functionally the existence of cortical columns that can act relatively autonomously allows parallel and therefore faster processing of information [149]. Simulations suggest that one of the central roles of the thalamus is to synchronize and de-synchronize the activity between multiple cortical columns so that they can work together or independently on any given task. This synchronization is achieved through rhythmic inputs to the relevant columns. In this way the functionality of the brain can be re-organized dynamically depending on the need [68].

On a global scale, the resultant network formed by thalamo- and cortico-cortical projections *must maintain a close balance between excitation and inhibition*. Too much inhibition and the brain cannot work, too little and neural activity gets out of control. Epilepsy, discussed in Section 1.1.4, is one form of such loss of control. Because projections between regions of the brain play an important role in brain function they are discussed next.

1.1.2.1 Cortico-Cortical Projections

Even within a cortical column the neurons are not organized randomly. Figure 1.6 shows that 6 layers can be identified in the 2mm thick cortex. The type and number of connections within each layer differ. The overall structure, though, can be said to be roughly responsible for the organization of inputs and outputs of the column [80]. A rule of thumb is that cortico-cortical projections occur predominantly in the top layers, while subcortical projections occur mostly in the deeper layers. Although the structural de-composition of a column is relatively well understood, this level of detail is not necessary in macro-scopic dynamics and cortical columns are treated in terms of their inputs and outputs, irrespective of where these occur.

Both excitatory and inhibitory projections form the cortico-cortical connections, but in general inhibitory projections remain relatively local. Cortico-cortical projections are most dense between columns that are nearby, but an incredibly large number of long-range and intra-hemispheric projections also

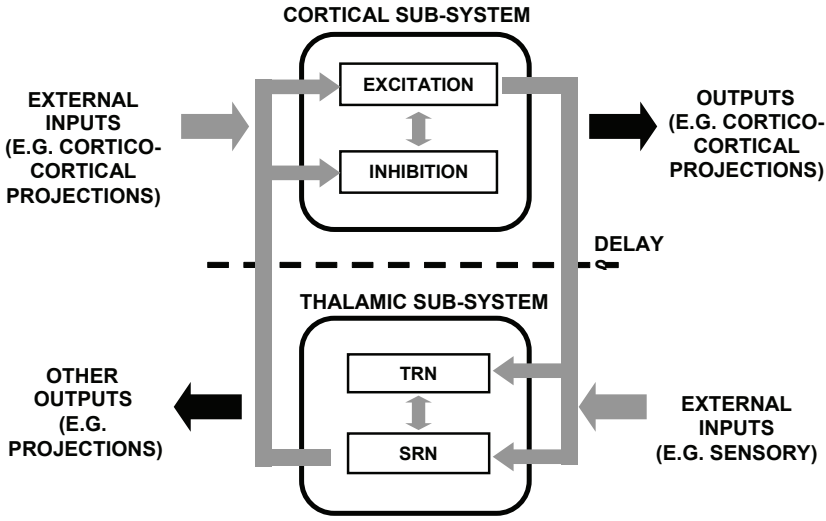


FIGURE 1.7: Thalamo-cortical loop. The figure shows the connections that are possible within two sub-systems – a cortical column and the sub-cortical thalamus – as well as the connections between them. Notice that each sub-system contains one inhibitory and one excitatory mechanism (TRN and SRN respectively in the thalamus) and that it is the excitatory parts that project out of each sub-system. Inputs, on the other hand, connect to both excitatory and inhibitory components. These long-range projections must traverse the white matter and thus experience delays in the order of $20 - 50$ms that do not affect local internal signals of each sub-system. External inputs are also shown in this diagram.

exist. The main constituents of the white matter in Figure 1.3(b) are the fibers that form cortico-cortical connections. A single axon extending from a pyramidal neuron synapses with multiple other cortical columns.

Most cortical projections are reciprocal, that is, if a column has synapses that connect to neurons in another column, then more than likely this second column also projects axons to the first. Thus the major input to any cortical area is from other cortical areas, and 75% of all synapses in the cortex are from one pyramidal neuron to another [21, 168], ocurring locally or through long-range connections.

1.1.2.2 Thalamo-Cortical Projections

Only about 1% of the fibers in the white matter are projections to and from subcortical structures; thus the input to a cortical column from the thalamus is far smaller than that from other columns [168]. However the manner in

which these connections occur suggests that this small input is amplified by local feedback mechanisms.

In any case the role of the thalamus in the functionality of the cortex is not to be under-estimated. If the thalamus were simply a way to relay sensory information to the cortex then its presence would be somewhat redundant. Amongst other things, the thalamus is also responsible for the *regulation* of information to the cortex, possible because of the structure of the thalamo-cortical loop shown in Figure 1.7. Again it is only excitatory synapses that project to and from the cortex. The excitatory synapses projecting from cortex to thalamus also synapse with inhibitory neurons in the *thalamic reticular nucleus* (TRN). The TRN is reciprocally connected to the *sensory relay nucleus* (SRN), responsible for interpreting and transmitting sensory input. This inhibitory loop is a way that the cortical activity itself can be used to regulate how much of the information incoming into the thalamus from the environment is relayed onto the cortex [169].

Signals traveling along thalamo-cortical projections experience delays in the order of $20 - 50$ms [146].

1.1.3 Neurotransmitters and Neuromodulators

The chemistry in the brain is a lot more complex than has been alluded to in previous sections. Chemical interaction with the physiology has been reduced to either excitatory or inhibitory on the premise that it is the effects that are important rather than the names and particular mechanisms related to each substance. However a brief word on the different types of chemicals involved in brain function is necessary.

The *neurotransmitters* in the brain can be either fast acting or slow acting. Fast acting neurotransmitters are the ones responsible for local inhibition and excitation, and their concentrations are important in the generation of signaling at the local neural network level. Complex mechanisms that regulate the balance of chemicals around the cells exist – the level of chemical has to be such that once an ion gate opens the neurotransmitters are appropriately re-absorbed. If insufficient neurotransmitter is released, the gate will not open. If it is not re-absorbed the gate will not close. The neurotransmitter is restricted locally to a particular synapse because if it is allowed to spread too far then other synapses may be activated.

In the human brain, the neurotransmitters most commonly responsible for excitation and inhibition are glutamate and gamma-aminobutyric acid (GABA) respectively. However GABA is known to also cause excitation dependent on location in the brain [27, 89], again emphasizing that in order to understand the brain it is best to talk about the effects rather than the chemicals themselves.

The slow acting neurotransmitters are referred to as *neuromodulators* because they are responsible for modulating the activity in a given brain region. They control the background activity or background state in this region rel-

ative to nearby regions. Neurotransmitters then work *locally* to generate activity in the context of the state provided by the more *diffuse* mechanisms of neuromodulators. Furthermore the release and re-uptake of neurotransmitters at the synaptic clefts occur on a shorter timescale than neuromodulators.

Although the complex mechanisms of neuromodulators are not clearly understood one of their roles is believed to control the level of coupling between cortical columns. When these columns behave similarly the brain activity is more global, slower and the overall parallelism in processing is reduced. Sleep is one example where the brain works in a more global way because the necessity for complex processing is reduced. When columns act more independently more tasks are possible at once. This is associated with levels of increased alertness.

Example neuromodulators include noradrenaline and dopamine, which decouple cortico-cortical interactions through inhibitory mechanisms and promote local interactions to allow increased parallel processing. Serotonin, on the other hand, suppresses local activity (again through inhibitory mechanisms) and promotes interaction between cortical columns. Noradrenaline and dopamine are found in highest concentration during alertness, whilst the opposite is true for serotonin [120].

If *any* of the mechanisms that regulate the concentration, release and uptake of chemicals breaks down then neurons and networks do not function correctly. Thus it is not only the physiology of the neuron that must remain intact, but also the mechanisms that maintain chemical balance. Malfunction in any of these can cause pathological behavior at the micro- or macro-scopic scale, depending on the type and severity of the malfunction.

1.1.4 Epilepsy – A Malfunctioning Brain

Epilepsy is a condition resulting from underlying physiological abnormalities in which seizures represent, to varying degrees, an infrequent phenomenon[2] [37]. The symptoms of epilepsy are very diverse and these depend on the region of the brain that is affected. An abnormality may cause seizures via the same mechanisms, but affect the epileptic person differently depending on its location [166]. What these mechanisms are, or what causes the onset and subsequent termination of seizures, is not well understood.

Physiologically speaking seizures are, spatially, a *large-scale phenomenon* in which a network rather than a single neuron is necessary to sustain abnormal activity. Most knowledge to date is derived from experiments on animals. Although the applicability of such knowledge to humans is not 100% understood [37] these experiments have identified the minimum conditions under

[2]The word 'seizure' is often replaced with *ictal*, so that pre-seizure and pre-ictal, seizure and ictal, post-seizure and post-ictal are all equivalent terms. This use of terminology is avoided throughout this text. Furthermore, the onset of a seizure must not be confused with the terms *epileptogenesis* which refers to how the epileptic disorder began (e.g., malformation, trauma, aging) as opposed to how individual seizures start.

which a network of neurons can sustain seizure-like activity. The network must [73]

1. Be part of a sufficiently large population (at least thousands of neurons),
2. Be relatively densely inter-connected, and
3. Have excitatory synapses that for the most part function correctly.

These conditions are *not* a result of the epilepsy itself, but instead exist naturally to sustain everyday activity. Seizures are the result of an abnormality in this network that allows neural activity to become unchecked. Neurons become *hyper-active* and *synchronous*, that is, they fire action potentials at a much higher rate than is normal and at the same time as nearby neurons. This abnormal activity spreads or *generalizes* to other regions of the brain via the same mechanisms that allow normal function. Once a relatively large proportion of the brain is involved the hyper-active synchronous firing continues in a waxing-waning manner resulting from resonances that occur due to the spatial distribution of network interconnections in the brain. These resonances can be linked to the geometry of the head.

The offset or post-seizure period is also important for a complete understanding of the epileptic brain – how the brain is capable of transitioning out of a seizure is from a dynamical systems perspective just as relevant as the transition into seizure. Most common is the belief that a seizure ends because oxygen supply to the neurons is depleted, but other mechanisms such as changes in chemical concentrations (e.g., accumulation of adenosine) or where a chemical imbalance responsible for initiating the seizure restores itself are also possible.

There are so many different types of epilepsies with different causes, symptoms and effects, that it is difficult to categorize them. A system proposed in 1981 (the International Classification of Seizure Types, [1]) is based on observation of clinical phenomena (explained in Section 1.2) as opposed to the underlying pathology. The system involves more than 15 different sub-categories, but only three broad classifications are made: *partial* or *focal* seizures, *generalized* or *non-focal* and *continuous*. The basic assumption in this book is that the epilepsies within each group are generated by similar mechanisms, and differences are nuances easily explained within each paradigm. The three categories are discussed next.

1.1.4.1 Focal Epilepsy – Failure of Meso-Scopic Networks

Focal or *partial* epilepsy is caused by an abnormality in a specific part of the brain, usually in the form of a group of damaged or abnormal neurons [80]. Seizures do not necessarily start at the focus, but are the result of its presence and its influence on the network as a whole. If a focal seizure spreads to a large proportion of the brain it is said to become *secondarily generalized*. The initiation and spread of such seizures are believed to be a progressive (as

opposed to an abrupt) transition. This transition may sometimes seem abrupt because the spread occurs very quickly.

A review of *histology*, that is, how the cellular and network properties of neurons in the focus vary from 'normal' brain, is presented in [12]. They discuss that the abnormal tissue can vary in morphology (e.g., enlarged neurons), connectivity (e.g., abnormal distribution of dendritic connections, often reduced inhibitory synapses) and excitability (e.g., abnormal neurons are more prone to fire action potentials, or their responses are larger than in 'normal' neurons). However the work they have presented is limited to cases for which samples are available from surgical removal, restricted to those epilepsies that are both resistant to medication and operable (see Section 1.1.5). Often it is not possible to know if the observed abnormalities are the *cause* or the *effect* of recurrent seizures. Furthermore the acquisition of 'normal' tissue samples from healthy humans is unethical and controlled comparisons are scarce. The abnormalities listed here are thus only examples and possibly un-representative of all epilepsies.

In any case when the brain is not seizing these neurons cannot usually participate in normal activity and are maintained in control by strong inhibition provided by surrounding neural networks [37, 166]. The abnormal levels of inhibitory activity suggest that the epileptic brain behaves differently from a 'normal' brain even between seizures [91, 37]. Seizures are the result of a breakdown in the mechanisms that maintain the epileptic neurons in check [118]. Once control of these abnormal neurons is lost they become hyperactive, firing massive bursts of action potentials that can recruit and entrain otherwise healthy neurons from nearby or remote areas.

To understand focal epilepsy two processes must be explained:

1. *Initiation*: What makes the focus become epileptic? Or in other words, what causes the inhibitory mechanisms to fail and allow runaway activity? Is it a network phenomena, a chemical imbalance or perhaps a network phenomena that leads to a chemical imbalance? The abnormalities in the focus must have something to do with this, probably in the balance between their increased excitability and the failure of inhibitory mechanisms. Alternatively it has been suggested that the inhibitory mechanisms themselves can cause synchronization in the abnormal neurons, thus the inhibitory process may, at least in some epilepsies, be responsible for both the control *and* the cause of the seizure [12]. Understanding the initiation of a seizure may lead to the ability to *predict* its onset.

2. *Generalization*: How does the seizure spread? It is known that the spread is the result of the inter-connectivity that allows the brain to function normally, but understanding the spreading mechanisms could potentially be used to provide more appropriate treatment. Other questions ensue, for example: is the role of the thalamus passive in that it simply mediates faster spreading of activity, or is it actively involved in

the seizure? From a clinical perspective it is important to understand how much a seizure must spread before it is *detectable*[3].

The effects of focal epilepsy can vary. *Simple seizures* are those that cause relatively mild cognitive, psychic, sensory or autonomic symptoms and no interruption to consciousness, whilst *complex seizures* are those that spread to cause altered states of consciousness, complex automatic or convulsive behavior. The location of the focus affects the experience of the seizure. For example if the focus is centered on the region of the brain responsible for smell, then this sense is affected during a seizure. Temporal lobe seizures are known to produce mystical experiences.

Some focal epilepsies appear non-focal because traditional imaging cannot locate the abnormalities. However others are known to be non-focal, as discussed next.

1.1.4.2 Non-Focal Epilepsy – Failure of Macro-Scopic Networks

Primarily generalized or *non-focal* epilepsies have seizures whose onset manifest across the entire brain immediately[4]. The time at which the sudden change from 'normal' to epileptic activity occurs can be likened to a *bifurcation point* because at this time the system's behavior becomes, qualitatively, very different[5]. The mechanisms that explain how the brain reaches this bifurcation point are even less understood than for focal seizures. Whereas the focus is known to be responsible for these seizures, the explanations sought for non-focal epilepsy place emphasis on understanding the origin of the activity that leads to the actual seizure:

1. *Initiation*: What activity in the brain is responsible for the initiation of a seizure? More than likely it is a combination of (a) balance of chemistry (neuro-transmitters and neuro-modulators) and (b) physiology (histology, network connectivity, delays and resonances).

2. *Spread and Generalization*: If the processes responsible for the generation of a seizure are known, what leads the brain to the bifurcation point? Is it changes within the brain or external stimulus that drives this rapid change? Does the thalamus play an active role? Some believe the origin of generalized seizures to be thalamic because it is very well connected to all parts of the brain and thus allows for an almost instantaneous spread [80]. Nevertheless these questions remain inadequately answered.

[3]Many measurements/observables can be use to detect seizures, but in this book we limit ourselves to detectability by EEG measurements, that is, a seizure is detectable when it is observed in the EEG measurement. The EEG is described in Section 1.2.

[4]Focal seizures may sometimes appear primarily generalized but it is believed that in these cases the spread is very fast so that they seem instantaneous.

[5]Here the term 'bifurcation' is used loosely. A more rigorous definition exists in dynamical systems theory. However, at least qualitatively, it is the same as that in the mathematics.

The variety of generalized seizures suggests that their cause is due to intrinsic properties of the brain as a whole rather than a consequence of a subset of neurons. For example, absence seizures usually cause a very short interruption of consciousness and they appear as a period of vacancy on the epileptic. These type of seizures are most common in children and to a large extent disappear with adolescence. What changes are experienced by the brain during this maturation period? Changes in hormones? Re-structuring of connections? Which of these could be responsible for abating the seizures? On the other hand, other generalized seizures *develop* with maturation – is it the same processes that in some brains stops epilepsy that allow others to develop it? Can a brain be said to be *predisposed* to epilepsy given a histological difference to a 'normal' brain? If so, what are these differences, and how different does a brain need to be to inherit this predisposition?

1.1.4.3 Continuous Epilepsy

Also known as *status epilepticus*, continuing seizures describe a state in which there is no observable recovery between seizures. This is a very dangerous form of epilepsy that can become life-threatening even 5 minutes after convulsions have started because the inability to sustain metabolic demand (oxygen, sugars) can cause brain damage. More than 30 minutes of convulsions can lead to death [80].

Since status epilepticus is thought of as a continuing form of either focal or non-focal seizures it is not discussed in any more detail here.

1.1.5 Diagnosis and Treatment of Epilepsy

Despite the many imaging techniques that allow visualization of brain structure and functionality (see, for example, Section 1.2) these alone are not sufficient to diagnose epilepsy. The presence of abnormal focal tissue does not necessarily cause seizures if it is located in a region of the brain that is insufficiently connected to other parts of the brain. Activity that resembles seizures may not be epileptic in origin. Such data must be coupled with a patient's medical history, including genetic predisposition as well as the nature of the symptoms so that a diagnosis may be made. The symptoms themselves (e.g., smell, dizziness, motor impairment) may provide physicians with information about the location of a seizure focus. Often it is necessary to induce a seizure – in these cases they are provoked by applying stress to the patient in the form of hyperventilation, sleep deprivation or, when appropriate, flashing lights.

Once diagnosed treatment can begin[6]. The most common form of treat-

[6]Throughout history treatments have ranged from the mystical (black hellebore, oak mistletoe, valerian) to the strange (swallowing the heart of a rattlesnake, wearing the head of a cuckoo around the neck) to the downright ridiculous (sleeping over a cow stable) [125]. Today treatments are based on scientific experimentation although a good element of experience remains in the way that it is administered.

ment is pharmacological. Fixed doses of *anti-epileptic* medication administered daily are used to control seizures. This works to varying degrees for two-thirds of the epileptic population. *Surgical removal* of the damaged neurons in focal epilepsy may be suitable for a further 8% of patients but can result in irreparable damage to other brain functions. This leaves roughly a quarter of sufferers with no viable options and for whom other forms of treatment such as *electrical stimulation* are being investigated.

1.1.5.1 Anti-Epileptic Drugs

Anti-epileptic drugs (AEDs) attempt to curtail epilepsy by affecting the chemistry driving the cellular processes underpinning epilepsy. The problem with AEDs is that it is difficult to tailor drugs to suit the particular circumstances of the individual case when the actual chemical origin of the epileptic condition is not well understood. Moreover drugs necessarily affect the entire brain, and have an impact beyond the brain. Their side effects are many, and varied, and cannot always be predicted [133]. To overcome the many side effects, AEDs that can be localized in administration and are fast acting with a limited time window of effectiveness are preferable. However at present such AEDs are purely experimental.

The effects of most drugs are complex and vary with age, medical history, genetic background as well as the type of epilepsy. As a result, achieving the correct dosage level is often based on empirical knowledge and administered through trial and error (under tight medical supervision). Drugs that have shown success on a wide scope of epilepsies are tried first, with increasing dosages. If a drug proves inadequate a new drug is tested, at different concentrations, and so the process goes on. Each patient requires a tailor-made solution, and dosage adjustments should only be made when the clinical need exists [133]. The correct cocktail of AEDs must balance the reduction of seizures versus the resultant side effects (e.g., dizziness, weight changes, cognitive impairment, rashes, etc) which can sometimes impair lifestyle almost as much as the seizures themselves. Until the mechanisms of epilepsy are more clearly understood, or fast-acting drugs can be delivered only to the regions of brain that are involved, the administration of AEDs continues to be a complicated process.

1.1.5.2 Surgical Resection

If more than three AEDs fail to lessen the number of seizures significantly the patient is medically *refractory* and more drastic measures are considered. For focal epilepsies this could mean the surgical removal or *resection* of the epileptic focus so that the damaged neurons are no longer involved in the network activity. This may appear like a radical concept but it is certainly not a new one – stone age cave paintings have been found in France that suggest that primitive forms of cranial surgery were used as treatment even then. Clearer records exist for medieval times where burning of the back of the

head (*cauterization*) was sometimes applied to healthy children to prevent the disorder [125, 180]. Today mesial temporal lobe epilepsy is the most common adult epilepsy known to be particularly resistant to drugs, and thus is the epilepsy for which most surgeries are performed.

It is amazing that such a large amount of redundancy is built into the brain so that removing a significant proportion can have relatively minor effects. Some may argue that epileptic foci are not used by the brain; thus their removal most likely alleviates the burden of maintaining control, but there are severe cases in which an entire hemisphere is resected and motor control is re-learned by the remaining half. In any case, even the smaller surgical procedures are risky and are only carefully considered when no other option exists.

1.1.5.3 Electrical Stimulation

When neither AEDs or surgery can help it is a desperate situation for the patient and alternative forms of management are necessary. Attention is shifting to the application of *electrical stimuli* to *control* or *abort* seizures. This may seem un-intuitive given the well known effects of electricity to *promote* rather than to reduce seizures [32], but certain cases have proven contrary to this.

Vagal nerve stimulation (VNS) is the most common procedure to date, used to control epileptic activity. Again a recurring theme is that the mechanisms of operation are not understood, and the rationale behind VNS is that since 90% of nerve fibres in the vagus nerve project toward the brain then activity can be *indirectly* modulated with the applied stimulus [9, 116]. The reason why these changes may be anti-epileptic is not known, but correlates between blood flow to the thalamus and efficacy have been observed [116]. Speculation as to the indirect activation of inhibition through the release of seratonin also exists [9]. No significant changes are observed on the electrical activity in the brain when the VNS is active, even though these changes are obvious in animal experiments [116].

The first human VNS implant occurred in 1988, and since then hundreds of thousands of procedures have been performed. How effective the implant is remains vague – double blind tests with control groups comparing seizure frequency before and after implantation vary significantly. An observed trend is that in the short term a 50% reduction of seizures is observed in 20-30% of cases, and this number increases to 40-50% in the long term [9, 116]. The implant remains effective years after activation [20]. Seizures cease completely in only 1-2% of cases [9], whilst no effects are observed in about 35% of cases [107]. These figures must be interpreted carefully because the epilepsies under test are by definition those that have proven particularly difficult to treat.

VNS stimulation is an 'always on' treatment that may affect a large region of the brain, not necessarily the required part. The long-term clinical consequences of this are seemingly inconsequential, but unknown nonetheless. The negative side-effects of VNS (e.g., sore throat, headache) are relatively minor

and decrease significantly with time. Positive side-effects such as mood and alertness elevation [38, 116] and quality of life improvements [107] are seeing VNS used in the treatment of other disorders including depression. The major drawback of VNS is that it is not known how it works let alone *if* it will work. In any case VNS is a form of seizure reduction rather than elimination.

The *direct* application of electrical current to subcortical structures of the brain, also known as *deep-brain stimulation* (DBS), is being reviewed for clinical application today. The structures under study include the cerebellum because it is known to promote inhibitory activity, the thalamus because of its wide-spread connectivity and the hippocampus for its involvement in temporal lobe epilepsy [116]. Studies show comparable performance to VNS.

The question of how the stimulus is applied is open for debate. Both VNS and DBS traditionally use continuous stimuli in an on-off manner (e.g., slow mode: 30 seconds on, 3-5 minutes off, rapid mode: 7 seconds on, 12 seconds off [116]). Studies considering the frequency and shape of the stimulus are rare, most likely because tests are again empirical rather than based on an understanding of the epilepsy that allows optimal selection of parameters.

A recent shift in research tries to *abort* seizures rather than control them by applying the stimulus only once the seizure has begun. The main advantage of this is that continuous stimulation is not necessary, side effects are minimized and battery life is significantly extended. The problem is that reliable detection or prediction of these seizures to initiate the treatment process is very difficult, as discussed in later sections, and that it is not known whether it is possible to abort the seizure once it has begun. Promising studies show that seizure duration can be altered, but much research must be completed before this can be clinically evaluated [189].

1.2 The EEG – A Recording of the Brain

To obtain much of the information required to diagnose epilepsy the *electroencephalogram* (EEG), literally meaning 'an electrical recorder of what is inside the head', has proved invaluable. It provides a measurement of the electric activity in the brain, translating the chemical currents into voltage recordings. It has high temporal resolution in that it is able to characterize fast changes in current flows, but poor spatial resolution because measurements are limited by the number of electrodes, their placement and properties of the head. Voltage recordings were first demonstrated on monkeys in 1875 by British neurophysiologist Richard Caton, but the practice did not become clinically viable until the 1920s when un-invasive practice became possible [103].

Measurements can be made at different spatial scales. Macro-scopic records are obtained (relatively) un-invasively from the scalp, or through surgical pro-

cedures that allow recording from within the head, that is, *intra-cranially*. Each has its advantages and disadvantages:

1. *Scalp EEG*: The signal must propagate through several layers of non-neural tissue – namely cerebrospinal fluid, skull and scalp – each affecting it in different ways. Recordings at the scalp are heavily attenuated and much larger regions of the brain must be actively synchronized for an EEG signal to register. The procedure is easy and inexpensive and used as a diagnostic tool, sometimes capable of providing sufficient information, but at other times as a preliminary step to more detailed intra-cranial records. Standard electrode positioning systems exist to make records more comparable within and between patients. The international 10-20 system for electrode placement is explained in Figure 1.8(a).

2. *Intra-cranial EEG*: Records can be taken from *cortical electrodes* (those placed *on* the cortex) or *depth electrodes* (those that penetrate to subcortical systems such as the thalamus). Smaller spatial scales are recorded by the intra-cranial EEG, but the phenomena is still macro-scopic or at least meso-scopic[7]. Intra-cranial records are often obtained for pre-surgical analysis to determine regions of the brain to be resected. Such procedures are relatively rare and data of this nature are more difficult to obtain. Standardization of electrode placement is also more difficult because decisions are made on a patient-by-patient basis.

In all cases the analog signals generated by the brain are passed to interfacing machinery responsible for amplifying, filtering and digitizing the data before relaying it to a computer for storage and analysis (see Figure 1.2). The technical specifications of the machinery play a role in the integrity of the recorded signal. One recording system is presented in Chapter 5. The positioning of the electrodes and the referencing system used play a critical role in determining the EEG. Two referencing systems for scalp-recorded EEG are shown in Figure 1.8.

For the most part, *the EEG measures the potential differences induced by currents flowing due to EPSPs and IPSPs – not action potentials.* This is because the total field potential of a group of neurons more or less equals the sum of the field potentials of individual neurons. Action potentials are of shorter duration and do not overlap as much as EPSPs and IPSPs [118, 122]. *Many thousands of neurons* $(10^4 - 10^7)$ *must behave in synchrony to generate large enough signals that register on the EEG.* This is particularly true for scalp recording where propagation through to the scalp attenuates and filters the signal.

Even so, it is reasonable to expect EEG signals to be non-random because

[7] The procedures for micro-scopic recordings involve very different considerations and are not discussed in this book. In any case these are no longer referred to as EEG, since EEG always measures *ensembles* of neurons.

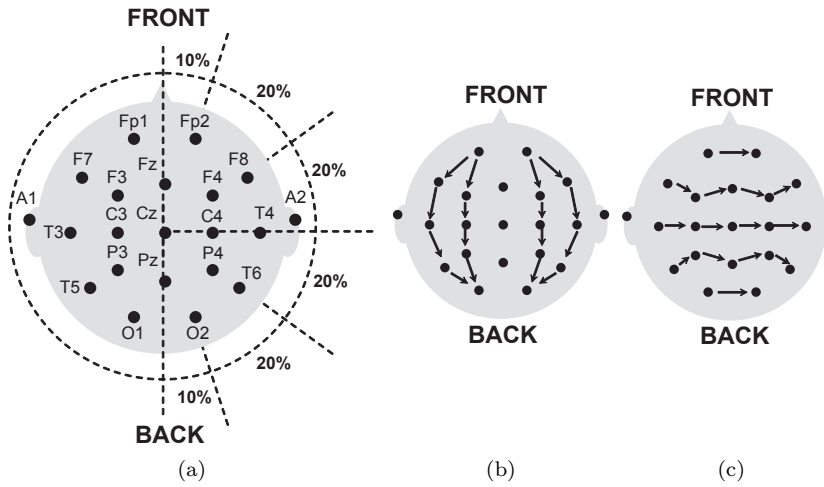


FIGURE 1.8: International 10-20 system for placement of scalp EEG electrodes. In (a) the positions and standard naming of the electrodes are shown. The advantage of this placement system is that the locations of the electrodes are computed as percentages of standard distances. In this way records are comparable between patients. Figures (b) and (c) show two standard ways in which the measured signals are referenced to each other. These referencing systems highlight relative activity between the front and back of the head (in (b)) or between the left and right hemispheres (in (c)).

neurons in functionally related areas, particularly in the cerebral cortex, generate activity that is similar to other nearby neurons. Sub-cortical networks are assumed to have little effect on scalp EEG recordings [118]. An EEG is useful in reflecting the global dynamics of electrical activity of large populations of neurons, that is why it is so useful in the diagnosis of epilepsy. A more in-depth analysis of EEG measurement for cortical and scalp records is presented in Chapter 2. Following is an overview of what 'typical' EEG records look like.

1.2.1 The Normal EEG

There is both art and science involved in the interpretation of EEG. *No global definition of what an EEG looks like (or should look like) exists.* Changes are evident at different stages of human life, different levels of awareness (sleep, awake) and different modes of behavior (eyes open, eyes closed). Electroencephalography defies standardization even within these states and hence it is difficult to train an expert (human or computer).

Even with the extensive documentation available the lack of a 'normal'

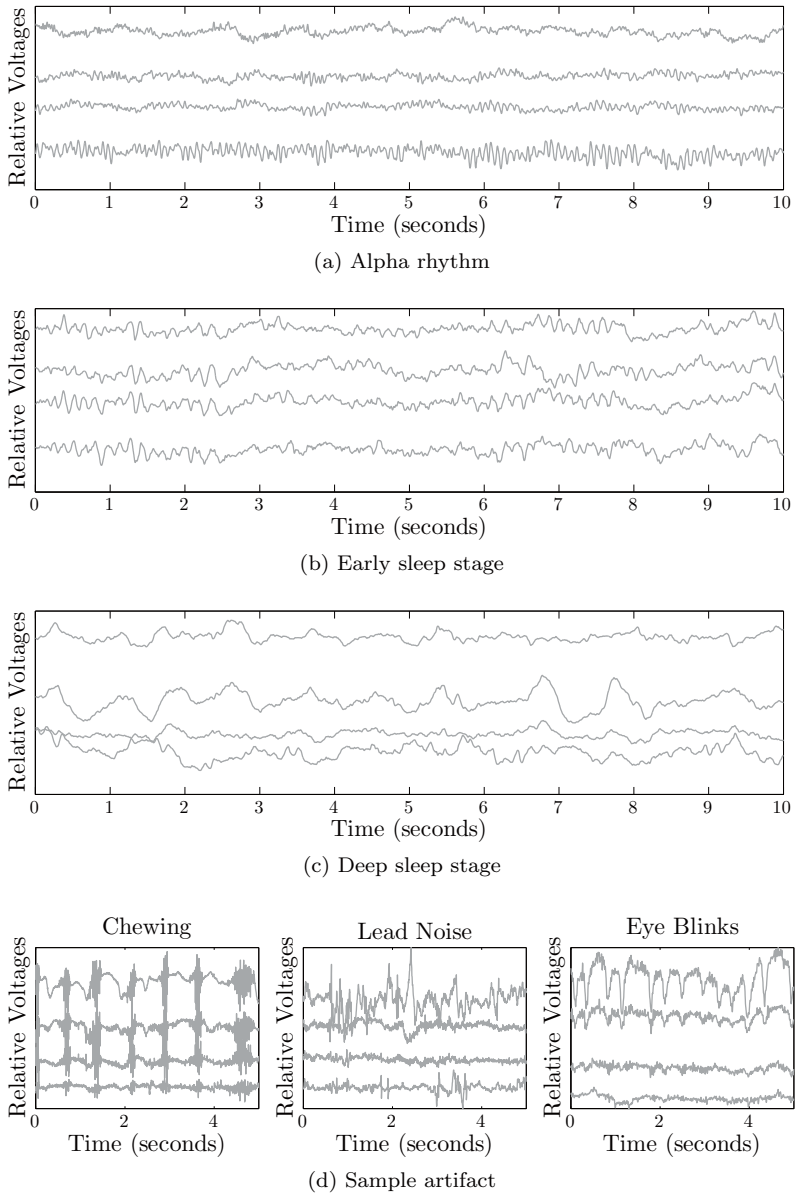


FIGURE 1.9: Example of 'normal' EEG traces. The voltage magnitude in each channel is shown relative to each other. (a-c) highlight the differences between different states of alertness. In (a) an example of an awake alpha rhythm shows the 10Hz activity present in only some (posterior) channels. In contrast to the slow waveforms of sleep stages in (b) and (c), the awake EEG shows a lot more variability between channels, demonstrating the more global nature of sleep versus awake states. Sample *artifact* are shown in (d).

EEG means that the most an electroencephalographer can do is to visually recognize general patterns that exist consistently in the majority of the population. It is important to remember that the absence of such patterns does not necessarily imply abnormality. This is why the EEG alone is not sufficient for diagnosis of epilepsy, and other forms of observation are necessary.

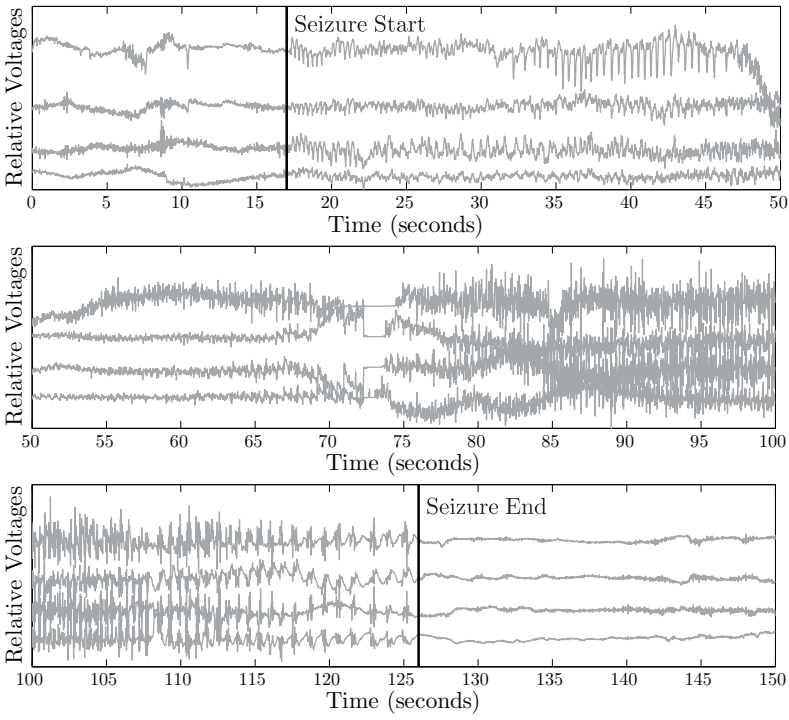
A variety of representative EEG segments can be found in Figure 1.9, each showing at least four electrodes so as to demonstrate the global nature of the EEG measurement. Typical rhythms or patterns have been identified in the 'normal' EEG to install some order in an otherwise seemingly random environment. The most common of these, the *alpha rhythm* in (a), is an 8-13Hz waveform occurring during wakefulness over the posterior regions of the head. It is best seen with eyes closed and is present in most healthy adults, and often interpreted as an indication of mental maturation and health. But even within this commonly observed phenomenon there are large variabilities in voltages, spread and quality. Perfectly healthy adults have been found with no demonstrable alpha rhythm [118]. The alpha rhythm is a global phenomena because it requires large networks to exist, but there is debate as to whether it originates due to resonances in the cortico-cortical or thalamo-cortical loops.

The different stages of sleep produce an EEG that is very different than the awake EEG. Examples are shown in Figure 1.9(b) and (c). The waveforms are slower and more global in nature, that is, the channels resemble each other more, although there are cases in which fast activity and sudden spikes are observed during sleep. These are normal phenomena but interfere with the identification of the different stages.

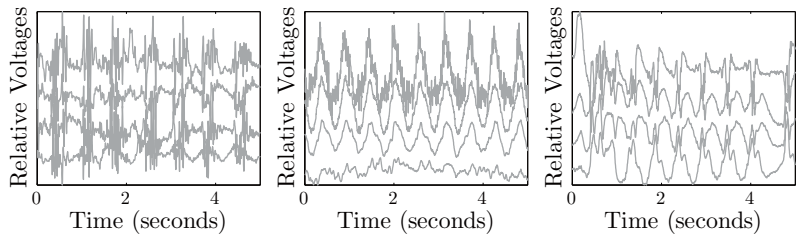
Artifacts are noise on the EEG, particularly strong in scalp recordings, caused by activity that does not originate in the brain but that cannot be removed at the time of recording. These can be external – e.g., ambient electromagnetic interference (50Hz for Australia, 60Hz in USA), improper electrode-scalp junction – or physiological – e.g., eye blinks, chewing, muscle movement of scalp musculature (EMG), and, less commonly, heart beats [118]. Examples of some such phenomena can be found in Figure 1.9(d). Although an electroencephalographer can be quite easily trained to recognize and categorize these artifacts, their removal presents a challenge in digital analysis because it is difficult to separate the artifact without also affecting the measurement of true neural activity. Appropriate techniques to separate true neurological activity from such interference are the subject of much research but are not yet sophisticated enough to be clinically applicable (e.g., [150]).

1.2.2 The Epileptic EEG

Again the variability between epilepsies means that there is no single 'epileptic' EEG. For example, inter-seizure periods can be as short as a few seconds or as long as years. The general trends that occur during seizure have been identified as



(a) Evolution of a 'typical' seizure



(b) Sample epileptiform waveforms

FIGURE 1.10: Sample epileptiform EEG. In (a) is an example of a complete seizure of approximately 110 seconds duration, with its start and end as marked. The magnitude of the EEG during a seizure is much larger than that preceding it. Notice that the seizure *evolves* over time, with changes in morphology as well as fundamental frequency. Notice also the artifact that occurs at about 73 seconds. This is an example of electrodes becoming temporarily disconnected (probably due to convulsive movements). In (b) are some sample waveforms to demonstrate the different ways in which a seizure can manifest. These are examples only and not representative of all the possibilities in the different epilepsies. In fact the above are cases in which the seizures generalize to all channels and are very easy to distinguish from 'normal' EEG. Other seizures may involve only a subset of channels and be very difficult to differentiate from background EEG.

- *Synchronization*: The EEG channels behave more like each other because the neural activity is similar across a larger area of cortex. This can occur at small scales and at large scales, that is, synchronization can be observed in a few or in many EEG channels.
- *Large amplitude*: After seizure onset the EEG may become much larger in amplitude than prior to a seizure. This can be because the cortex becomes hyper-excitable (more neurons are active) or simply a consequence of the increased synchronization between channels.
- *Oscillations*: Each channel often becomes more oscillatory, in contrast to the examples of normal EEG shown in Figure 1.9. The oscillations are typically associated with larger scales – that is, once epileptic activity has spread sufficiently – because interacting regions of the brain are necessary to sustain oscillations.

These trends are representative only, and often recognized only because of a sharp change from background activity. A seizure does not necessarily display all, or in fact any, of these features. It is difficult to come up with a single characteristic that is common to all types of epilepsies. For example, although many seizures involve large amplitude oscillations there are others whose amplitude are no larger than previous activity; channels become synchronized during seizures but not all channels are necessarily involved in an episode; seizure onset and offset are usually abrupt but some are capable of appearing more slowly; after the seizure there may or may not be a period of 'silent' EEG, where there is noticeably decreased activity and during which time the patient is in a state much like deep sleep. This variation in observables occurs between patients, between seizures in the same patient and *within* a single seizure.

Figure 1.10 shows sample EEG of 'typical' seizures. Of most importance is the *evolution* that occurs, shown in (a), where the seizure changes in fundamental frequency as well as morphology over time. In (b) are some examples of different seizure EEG waveforms. In all cases the activity in each channel is more oscillatory than normal, although the shape of the oscillations vary, and the seizure has generalized to involve all channels. Notice also how similar the seizure EEG in the leftmost panel is to the chewing artifact shown in Figure 1.9(d). This highlights the difficulties with differentiation between what is 'normal' and what is 'epileptic'.

Inter-seizure epileptiform discharges, known as spikes, can sometimes be observed. Characteristically these are short bursts of high amplitude, synchronized and multi-phasic activity (in which a change in polarity occurs several times) that manifest themselves at or around the epileptic focus and stand out from the background EEG. Some believe that full blown seizures may simply be prolonged versions of these spikes [118, 158], although this is contentious.

Given the large variability in observations the correct classification of the epileptic EEG is a difficult if not impossible task. Classification of the epileptic EEG is discussed next.

1.2.3 Detecting Changes in the EEG

The reason that EEG is so useful in the study of epilepsy is that it can quantitatively show the changes of brain activity over time. Because seizures represent a rare phenomenon, detecting these changes is important for the diagnosis and treatment of the disorder. Reliable automated detection of changes that lead up to a seizure is the major theme in this book.

The purpose of the detection of changes varies. It may be used solely to distinguish between seizure and inter-seizure EEG, useful for diagnosis. Alternatively it may be *predictive* in that it informs when a seizure is imminent before it happens – detection of the *pre-seizure* state. The ability to predict leads to the possibility to manage the disorder better, and it allows a complete re-evaluation of pharmacological treatment [91]. Certainly the social benefits are indisputable – it is the apparent randomness of seizures that make epilepsy so debilitating [96].

More than twenty years of research have demonstrated that selection of *features* that best distinguish between inter-seizure, pre-seizure and seizure EEG is a difficult task. Most likely this is because time fragments of EEG cannot in general be simply labeled either 'normal' or 'epileptic' [193], that is, the decision space is much more complex and these two 'categories' may not be separable in general. The complexity may be gleaned from the observation that a 1,200 page book edited by Ernst Niedermeyer and Fernando Lopes da Silva ([118]) is often used as the handbook by electroencephalographers, and is solely dedicated to the description of EEG.

The difficulties have not stopped the hunt for the optimal features. Epilepsy is, after all, the most common recurrent neurological condition in the world. The features themselves are extracted from the EEG through *signal processing*, a process by which a signal, in this case the EEG, is transformed to a quantitative form that is more compact than the raw data so that it can be understood more easily. To detect the onset of seizures it is logical to target quantitatively what human experts target qualitatively. Human experts base their decisions on information such as

1. *Spatial and temporal information:* A seizure is usually reported when its duration is long (short epileptic bursts, e.g., 1 second, may sometimes be classified as spikes) and not local. Oscillations in a single channel are unlikely to be epileptic in nature because epileptic events have a field which nearly always involves nearby electrodes.

2. *Background state:* A clear and well defined period of different EEG pattern to the background activity is sought when classifying the epileptic EEG. Of most importance is the difference between the awake and asleep EEG, the latter being slower and more global in nature.

3. *Expert-knowledge rejection:* Bursts that comply to criteria associated with (1) and (2) may still be artifactual in nature. At this stage, the

expert uses accumulated knowledge, gathered through experience, to distinguish between an epileptic event, common artifact or rhythms.

The guidelines that must be used to predict seizures are less clear because the human eye is often incapable of detecting the changes that lead to a seizure. Current predictors perform poorly and are only tested on very specific cases (e.g., [71], [113]). The question *'can seizures be predicted?'* has not yet been adequately answered in literature.

In the identification of the seizure or pre-seizure state the ability to quantify rules that differentiate them from the inter-seizure EEG is imperative to the performance of the detector. Current detectors are data driven, also known as *black box* models, and do not infer any knowledge beyond that which is presented by the EEG data. A review of the features commonly used and how they perform in the context of seizure detection is given in Chapter 3 and Chapter 5.

An alternative strategy is to develop detectors based on an understanding of how the brain behaves. This involves translating the physiology (Section 1.1) and its measurements (Section 1.2) into a mathematical *dynamic model*. Unlike black box methods, dynamic models can be used to infer beyond the data that is available, and may be useful when solving difficult problems such as prediction of seizures. Important concepts in the construction of these physiologically based dynamic models are introduced next.

1.3 Dynamics of the Brain

The *dynamics* of the brain describe how activity in this system evolves over time. In engineering a *model* is typically a set of mathematical equations that explain this behavior. In neuroscience it is typical to use the word to refer to *animal models* – a condition in an animal, often induced artificially, that is similar to that of a human. Both types of models are designed to describe and understand activity found in the brain, and both can *simulate* the behavior they are trying to replicate. But whilst animal models are most often used to understand the mechanisms of a specific pathology, mathematical dynamic models are in theory capable of describing both normal and pathological behavior. More importantly, a mathematical model allows *computation* of expected system behavior, a much cheaper and more reproducible alternative to animal models! For example, a suitable model of brain dynamics can be used to simulate the expected activity within the brain. In conjunction with information about the properties of the head it can then be used to compute the expected EEG waveforms at the scalp. In some cases computation allows prediction of future behavior.

The data found through an animal model are helpful in the development

of its mathematical counterpart. In particular, animal models have proven very useful to help understand the mechanisms of epilepsy – they can be used to generate epileptic data reproducibly which for ethical reasons would be impossible to obtain from human experiments. However, *a mathematical model of brain activity is not the same as a mathematical model of epilepsy*: one is a description of how the normal brain behaves and the latter is a description of a pathology. In theory a model of normal activity should also be capable of describing epilepsy if the *parameters* of that model – the numbers that describe the physiology of the brain – are changed accordingly. The animal models provide information as to what the relevant parameters are and how their behavior deviates from the 'normal' brain.

The process of constructing a mathematical model can be divided into three steps:

1. *Creating a mathematical dynamic model structure*: By using knowledge of the mechanisms of the brain the set of mathematical equations that best describe this system is written down. Identification of the relevant physiology depends on the scale of interest, which for epilepsy should be at least larger than micro-scopic. The model structure contains a number of the parameters that should be tuned to select a particular model from the collection of models that are represented by the model structure.

2. *Gathering experimental data*: Both animal models as well as routine medical measurements such as the EEG give information about the behavior of the brain. For epilepsy such data involves periods of 'normal' as well as pathological activity.

3. *Validating the model*: Using the experimental data gathered in Step 2, the dynamic model must be validated through simulations to see if activity of this type, both normal and abnormal, is reproducible when relevant changes are made to the parameters. The dynamic model may equally be *in*-validated, in which case further detail or changes in assumptions are needed in Step 1.

The type of data that will be used to validate the model is important when considering Step 1. For example, if the data reflect the electrical activity of the brain then it cannot be used to validate a model that describes the metabolic processes. In-depth knowledge of the physical *as well as the measurement* processes involved in generating the validation data is necessary so that the models can reflect this activity (see Figure 1.2). Since EEG

is the most common available data that describes changes over time, it is the global electrical activity of the brain that is most often used to generate global dynamic models. It is important to understand both the *generating system*, the brain, as well as the *measurement system* itself in order to understand the EEG signal. A detailed explanation of EEG measurement is presented in Chapter 2.

Simplifying assumptions are necessary wherever possible because the brain involves in the order of 100 billion neurons. This is not because greater complexity makes the model worse (although it is possible that too much complexity can deteriorate performance!), but because current technology cannot cope with such high computational demand. The simplifying assumptions make the problem simpler at the expense of detail, justifiable only when such detail is not essential. For example, if it is known that the activity within one column is roughly uniform, then modeling can shift from single neurons to single columns. This is only valid at the macro-scopic scales, suitable for EEG measurements, because small differences within a column would be averaged out by a large-scale measurement. Micro-scopic recordings, on the other hand, would be affected even by these small differences. Other relevant assumptions for macro-scopic modeling are listed in Chapter 6.

Once the scale of interest and corresponding assumptions are made, then the appropriate free parameters that remain in the model must be identified. For an arbitrary network this involves the identification of *structural components* such as neuron population and network topology, as well as the *dynamic parameters* including the inputs, outputs, internal signals and delays in the system.

1.3.1 Micro- and Macro-Scopic Models

In a model of the micro-scopic system, say one neuron, the most important process in its dynamics is the integration of incoming PSPs and the consequent firing of action potentials. The activity at the synapses is the input. The action potential is the output. Internal dynamics must account for the integration of the various PSPs at the soma, the propagation mechanism of PSPs from their synapse to the soma as well as the generation (or not) of the action potential itself. The dynamics of this system are governed by the strength of the inputs, the structure of the neuron as well as *delays* in activity. In a neuron delays are caused by propagation times from synapse-to-soma and soma-to-synapse, which differ depending on which part of dendritic structure connections are made. Delays also occur in the time it takes for inputs and outputs to turn 'on' and 'off', corresponding to how synaptic gates open and close. Opening and closing times are not necessarily the same, as seen in Figure 1.5(b) – a PSP is not symmetric (in time).

At larger scales it is not the single neuron that is important but the average activity of the ensemble. If our system is a cortical column in which most neurons are assumed to work similarly, the inputs to this system are incoming

signals from cortico-cortical and thalamo-cortical connections. The model output depends on what is being simulated. If it is a measurement such as the EEG that is being modeled, then the output is the electrical activity projected through to the measurement system. If instead it is the neural interactions that are important, then the outputs are simply the projections of activity to unseen cortical and subcortical regions. Propagation delays from the micro-scopic scale are relatively unimportant here, but the capacitive effects caused by gating mechanisms can shape the electrical activity of the EEG. The network topology, that is, the number of neurons and their average connections, is relevant to the generation of the internal signaling in this meso-scopic system.

Larger models can incorporate networks in which many cortical columns or subcortical systems interact with one another. Because such sub-systems are approximately independent from one another [120], each can be modeled separately and a larger system constructed by inter-connecting the smaller ones together. In this case the inputs and the outputs remain the same (cortico-cortical and thalamo-cortical projections), but additional delays corresponding to the time that a signal takes to travel between sub-systems are important in the generation of rhythms observed in the EEG. An example of a combination of sub-systems can be found in Figure 1.7.

In the last three paragraphs it is apparent that some of the dynamics at the micro-scopic scale can shape activity at the meso- and macro-scopic scales. The converse is also true – macro-scopic activity can affect the dynamics at the micro-scopic scale because average electrical activity over a large ensemble dictates how a single neuron reacts to an input. This bi-directional relationship between scales is depicted in Figure 1.11.

A model of a complete system must account for how smaller scales affect the dynamics of larger scales, and, in turn, how the activity at these larger scales affect the behavior at the smaller scales.

In the creation of the macro-scopic EEG model it is not always obvious what physiological processes transcend scales. Experimental evidence shows that of most importance are the capacitive effects and strengths of PSPs occurring at micro-scopic scales. Changes in these can affect the EEG frequencies of ensemble systems [42]. Other micro-scopic activity such as in-neuron propagation delays and dendritic structure are often ignored because they seemingly have little impact on EEG global dynamics [73]. Exactly which processes are thought important are described in further detail in Chapter 6, but recall the relationship between spatial and temporal scales states that EEG activity at macro-scopic scales is slower than micro-scopic. A rule of thumb is that dy-

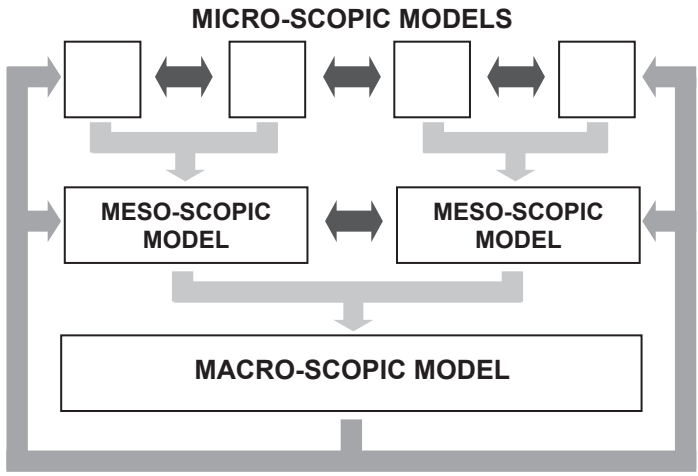


FIGURE 1.11: Representation of the intra- and inter-scale interactions in the development of a model. Within each chosen scale there is interaction between sub-systems of the same scale. Activity can also transcend scales, as shown, because the small scale dynamics can affect the larger scale dynamics. In turn, large-scale behavior can affect smaller-scale behavior. Appropriate assumptions must be made in the process of developing a model so that, once the scale of interest is determined, these interactions are not ignored. The identification of the parameters that can transcend scales is a key challenge.

namic processes at smaller scales affect the EEG at larger scales only if the activity is comparable in timescale. Synaptic signals are slow and overlap at slow timescales, thus are incorporated, whereas action potentials that are fast are omitted (see Figure 1.5).

1.3.2 Dynamic Models of Epilepsy

Epilepsy is often treated as a special case easier to model than an entire brain because it is a state in which a large population of neurons behave similarly, thus simplifying the problem. However many of the mechanisms of epilepsy arise from 'normal' processes and even a healthy brain is capable of seizing. Thus an understanding of epilepsy requires an understanding of the 'normal' brain.

In this book the approach is to focus on a single model that can describe the electrical activity of a 'normal' brain and in which suitable changes are representative of the many types of epilepsies. These changes can differ significantly depending on the pathology but should be supported by evidence found in animal models and be reflected by the appropriate parameters in the mathematical model. For example, in some types of focal epilepsy the focus

contains abnormally large pyramidal neurons [12]. A model of macro-scopic electrical activity may not contain parameters that describe the size of a typical neuron. This information must be captured by other parameters that describe how these neurons behave differently from healthy neurons rather than how differently they look.

In Section 1.1.4 two key questions were outlined in the understanding of epilepsy – *initiation* and *generalization*. A useful model should at a minimum be capable of replicating each of these, in the hope that understanding the parameter changes that lead to a seizure may shed light on the physical processes themselves. Two approaches are possible:

1. *Focus on initiation*: If normal inputs are capable of driving the brain into seizure then what are the model parameters responsible for this? Animal models are important in pinpointing relevant abnormal physiology and processes in epilepsy. The mathematical model can be validated if they are able to replicate observed experimental results. Conversely, because the mathematical models are based on physiology they can themselves be used to validate the mechanisms of epilepsy obtained from the animal models.

2. *Focus on spread and generalization*: This approach is only suitable for focal epilepsies in which the origin of epileptic activity is known. Non-focal epilepsies spread so fast that research is limited to the study of initiation of a seizure. Here we assume that the focus is behaving badly, in that hyper-synchronous and hyper-active behavior feeds into nearby cortex and subcortex. This output of the abnormal focus becomes the input to other normal regions of the brain. How does this result in the spread of activity?

The two approaches, illustrated in Figure 1.12, are not mutually exclusive but treating them as such can simplify the problem. If these models are to be used for the treatment of epilepsy the question of which approach to take is then whether the therapy is administered to prevent the initiation or the spread of epileptic activity. The former requires *prediction* of a seizure, whilst the latter focuses on *detection*, as discussed in Section 1.2.

1.4 Stochasticity in Neural Systems

The word *random*, often used interchangeably with *stochastic*, must not be confused with the colloquial understanding that 'nothing can be said about a random event'. A stochastic process can be described to a certainty dictated by the expected distribution of the process itself. Think for example of flipping a coin. Assuming no bias, each toss is an *instance* of the process with 50%

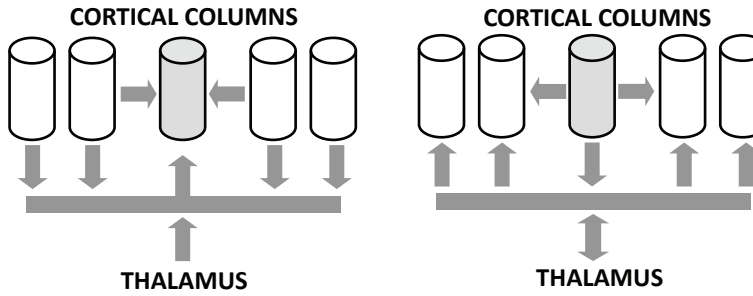


FIGURE 1.12: Two questions that must be answered in a mathematical model of focal epilepsy – *initiation* and *generalization*. In (a) the method is to concentrate on how the epileptic focus begins to behave incorrectly, that is, how the seizure is initiated. (b) looks at how this activity spreads once the focus is out of control. In both pictures above the gray cortical column represents the epileptic focus. The difference between the two methods lies in the flow of information. Approach (a) looks at how *inputs* to the abnormal neurons fail to control runaway activity, whereas (b) looks at how connections that project *out* of the focus entrain the remainder of the brain.

chance of being heads and 50% tails. For each toss we cannot say what the outcome will be until it actually happens. However on average, and given a sufficient number of trials, we expect half the tosses to be heads and half tails, as dictated by the *probability distribution* of this experiment.

A *stochastic* or *random* event cannot be predicted unambiguously – its value is only certain once it has occurred. However if the *event space* (the set of all possible values the event may take) and its *probability distribution* are known then it is expected that the *average* or *ensemble* behavior is as described by this distribution. If this distribution does not change over time this is known as a *stationary process*.

In the brain events are not disconnected from each other like the successive tosses of a coin. If the EEG is measured at 100mV in one sample then it is very unlikely that the next sample will read −100mV. *Information about the past limits the possibilities of the future.* If possible future events are limited to a sufficiently small range then it is a *prediction* of what is likely to happen.

In this book we describe many systems with stochastic elements:

1. *Brain sources (see Section 1.1)* : Even a deterministic brain has behavior dependent on stimuli that drive the brain (e.g., sensory information) which cannot be predicted and are thus best modeled as stochastic. Within the brain itself the behavior of the various cells, gates and chemicals is not known exactly because experiments provide information *averaged* over many trials. Thus the behavior of the brain can be said to be stochastic, either because the experiments themselves introduce stochasticity to the system or because the source is truly a stochastic process.

2. *Measurements (see Section 1.2)*: Determinism in principle means that there is zero uncertainty left. For a measurable system this is an illusion – measurements can only be made to a limited precision and thus there is never certainty, only shades of uncertainty. In this context stochasticity is the only way of describing reality and determinism becomes a mathematical abstraction.

3. *Dynamic models (see Section 1.3)*: Meso- and macro-scopic models assume uniform behavior of cells: given an equivalent input two pyramidal cells react in exactly the same way. This is not true because slight differences in the sizes of pyramidal neurons or chemical concentrations may result in differences in the output. Models of ensembles of neurons where the detail of each individual neuron is lost can be said to be true only *on average*, thus they are stochastic representations of the real system. Despite this 'variability', modeled as stochasticity, higher brain function is reliable and repeatable. The brain can thus cope with, or perhaps even rely on, a certain level of random behavior.

1.5 Conclusions and Further Reading

This chapter presented rather briefly the physiology of the brain, and served primarily to introduce the central topic of the book: How epilepsy may be perceived from an EEG measurement. To probe further in the fascinating world of brain structure and to learn more about the chemo-kinetic processes underpinning brain behavior one may refer to excellent textbooks such as [80].

The primary questions that will be addressed in the remainder of this book relate to how the EEG (in its various forms) relates to epilepsy:

- How well can we label epochs of EEG as epileptic or not?
- Can we explain how epileptic behavior generalizes and evolves over time?
- Can we predict transitions into epileptic episodes?

These questions do not have simple solutions, and are discussed in Chapters 5-7 in the context of epileptic seizure detection, mathematical models of epileptic activity, and the predictability of epileptic seizures respectively. However some theory must first be covered. We begin by presenting a more thorough explanation of the measurement system and its limitations in Chapter 2, and follow this by how we can make sense of the measurements using signal analysis and classification tools in Chapters 3-4.

2

EEG Generation and Measurement

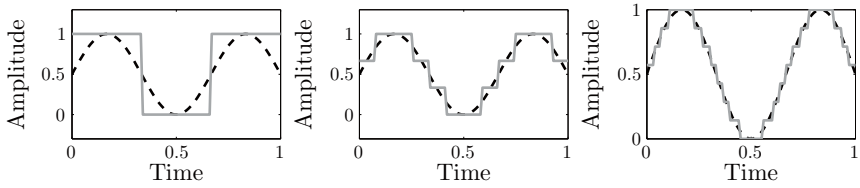
> "Since every piece of matter in the Universe is in some way affected by every other piece of matter in the Universe, it is in theory possible to extrapolate the whole of creation – every sun, every planet, their orbits, their composition and their economic and social history from, say, one small piece of fairy cake."
>
> *- 'The Restaurant at the End of the Universe', Douglas Adams (1952-2001)*

A blatant misrepresentation of the truth? Yes. Can we tell everything about the universe from a piece of fairy cake? Of course not. What *can* we tell about the universe from a piece of fairy cake? That depends on what is being measured. This tongue-in-cheek oversimplification of a complex phenomenon (true to form, the Douglas Adams way) may have been written to make us laugh, but does pose an important question – what can observations tell us about a system? The *electroencephalogram*, EEG for short, provides us with measurements of the temporal distribution of electrical activity in different parts of the head, generated by potentials in the hundred billion or so neurons in the brain. How can these measurements be used to tell us about the behavior of these neurons?

The aim of this chapter is to address the 'simpler' question: *what can the EEG recordings tell us about the underlying brain activity?* Or just as importantly *what cannot it tell us?*

To address these the neurophysiologist will reference the anatomy and neural structure in the brain, the physicist will begin by deriving a set of partial differential equations to describe charge distributions and their interactions, while the electrical engineer's instinct will be to draw equivalent circuits and derive linear transfer functions that describe the human head. None of these are wrong, all are relevant and valid methods that must be considered to determine how the measurement of EEG is affected by its unique environment. But in order to obtain a complete picture the relationship between physiology, charge distribution and electrical properties of materials must be understood to determine what the EEG, our proverbial piece of fairy cake, is capable of telling us about the underlying brain activity.

An important distinction that is made throughout is between EEG *dynamics* and EEG *measurement*. The former refers to how charges, currents, and activation patterns are *generated* by inbuilt mechanisms of the brain. This in-



(a) One bit (b) Two bits (c) Three bits

FIGURE 2.1: The aim of a measurement is to capture a real signal. However the precision of measurement affects how closely the measured signal (solid line) approximates the real signal (dashed line). The more *bits* or information used the closer the measurement approximates the real signal, as shown in (a), (b) and (c). The differences between the measured and the real signal are *errors* that are often modeled as *stochastic* or random variables.

cludes the gate-triggered synaptic activity, the subsequent initiation of action potentials and the network organization that maintains activity in the brain. On the other hand a measurement is a *record* of this activity in an attempt to quantify the charge distribution, currents and activation patterns described by the dynamics. The relationship between measurement and dynamics is dependent on the scale that is being measured as well as how the measurement is performed.

In any case a measurement is limited by the quality of the recording equipment. This includes its *sensitivity* imposed by physical constraints (e.g., how sensitive an electrode is) as well as how the data is represented. For example if the measured signal can itself only take two values then one bit of information (that is, a number that can take two values) can be used to store a perfect representation of the measured sample. However if one bit is used to represent a signal that can take four values then information is lost in the measurement process. Increasing the number of bits reduces the *measurement error*, as shown in Figure 2.1, but only down to the sensitivity of the equipment. *Measurement error is unavoidable.* If we have well calibrated equipment so that no consistent errors are made, measurement error can be modeled as a *stochastic* process.

EEG measurements at the scalp, cortex or elsewhere in the head are also affected by the electrical properties of the surrounding medium, known as the *volume conductor*. This is because the many biological tissues each alter the electric fields generated within the brain differently. The active dynamics of the brain can be treated separately from the volume conductor because at EEG frequencies its effects are predominantly passive, as is explained in more detail later. However understanding how the volume conductor changes the measurement is important for understanding the EEG, and is the central focus of discussion throughout this chapter.

The existence of electrical activity in the brain was first discovered in the late 1800s largely due to Richard Caton's experiments on rabbits and monkeys, although it was not until 1924 that the German physicist Hans Berger reported the first EEG for the human brain. Other forms of bioelectric events have been recognized for much longer – as early as 4000BC Egyptian hieroglyphs were found depicting a catfish capable of generating electric pulses – but were for a long time thought to be the result of 'animal spirits' and not linked to electrical activity. The word *electricity* only came about in 1600 when William Gilbert named it so after the Greek word for amber ($\epsilon\lambda\epsilon\kappa\tau\rho o\nu$). Gilbert devised the first instrument to measure the attractive force generated by the static electricity in the stone (Figure 2.3(a)). Since then Johan Salemo Christopf Schwigger's development of a galvanometer in 1821, Faraday's induction coil in 1831, Maxwell's equations of electromagnetism in 1861 and a myriad of other discoveries have led to the development of our understanding, and subsequently enabled the first human EEG recordings by Hans Berger (Figure 2.3(b)). The early traces bear little resemblance to today's EEG, but the ability to record the electrical activity of the brain has revolutionized clinical diagnosis and treatment of neurological conditions, and in particular epilepsy. A much more detailed history of electricity and its relevance to bioelectromagnetics and EEG can be found in [103] and [118].

Other quantitative descriptions of brain activity exist and are in clinical use. The *magnetoencephalogram* or MEG is a relatively recent development that tracks the magnetic activity generated by the same current mechanisms that generate the electrical activity measured by the EEG. Both EEG and MEG have very good temporal but poor spatial resolution. The magnetic fields of the MEG are examined in [122] and [103], and are not discussed in further detail here. Better spatial resolution but zero temporal resolution is possible with *Magnetic Resonance Imaging* (MRI) which estimates topographic maps of the structure of the brain. *Functional MRI* (fMRI) provides limited spatial and temporal resolution of the metabolic rather than electric processes in the brain. Co-registration of EEG and fMRI may be useful and has been studied in [121]. MRI may also be used to provide geometric information for EEG analysis and is in the process of being integrated as a complement to EEG.

There are in the order of $10^{10} - 10^{11}$ interconnected neurons in the human cortex [168]. The organization of these neurons is such that the analysis of the electrical activity generated ranges significantly depending on the scale of interest. Up to 100 neurons grouped together form a *mini-column* and produce *micro-scopic* fields. A cubic millimeter of cortical tissue, a *macro-column*[1], may contain 1000s of interacting mini-columns that produce *meso-scopic* fields (see Figure 2.2). Scalp and intra-cranial EEG both measure the *macro-scopic* fields generated by averaging the behavior of many of these macro-columns.

[1] Other definitions of a cortical column exist, sometimes with as few as 100 mini-columns.

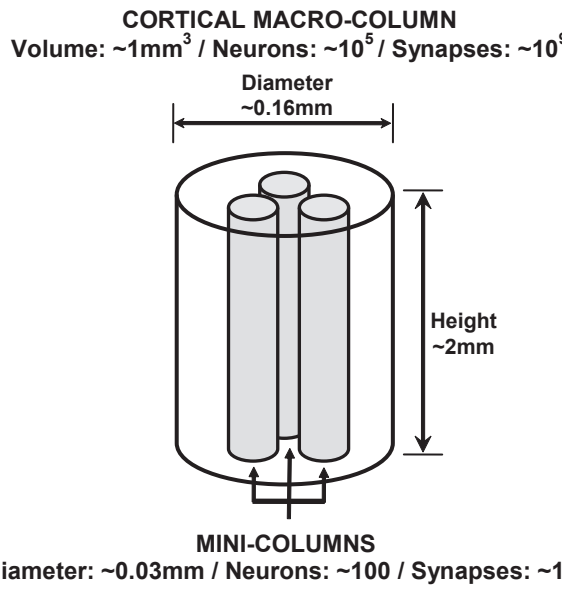
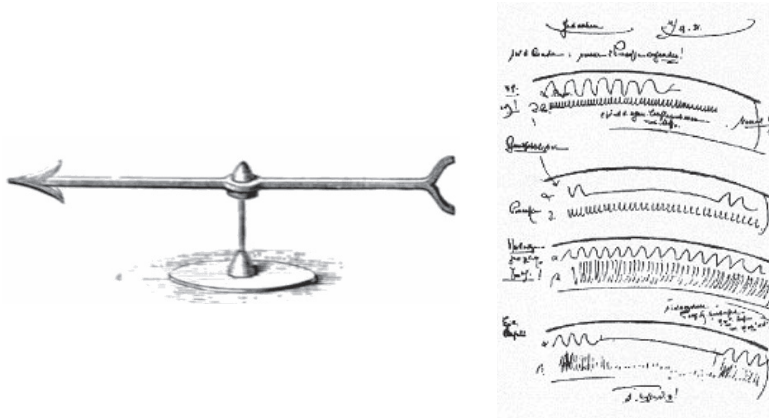


FIGURE 2.2: A pictorial representation of the different scales found in the cortex. There are in the order of 100,000 *macro-columns* in the human cortex, each occupying roughly a 1mm^3 volume of the cortex, as shown. Each macro-column has in the order of 10,000 neurons and about 10^9 synapses. A macro-column can be sub-divided into roughly 1000 *mini-columns*, each with ~ 100 neurons and $\sim 10^6$ synapses.

By convention, the potentials recorded at the micro-scopic and meso-scopic level are denoted by ϕ, whilst macro-scopic EEG scales are represented by Φ.

Macro-scopic potentials measured intra-cranially are predominantly dependent on the location and activity of mesosources. Macro-scopic potentials measured at the scalp are also dependent on source location and activity, but must consider the *volume conducting* properties of the tissues between source and recording site. The behavior of the head volume conductor is determined by the geometry and electrical properties of the brain, cerebrospinal fluid (CSF), skull and scalp. Much larger areas of cortical activity are involved at the scalp EEG because of these volume conducting effects. Calculation of the potentials generated at any scale, given source locations and activity, is known as the *forward problem*. The *inverse problem*, on the other hand, is the determination of source location and activity level given a measurement, and is a much more difficult task because there are many more degrees of freedom.

The majority of this chapter is dedicated to exploring the forward problem given the volume conducting properties of the human head. This then determines what the EEG measures so that its limitations, and what it is capable



(a) (b)

FIGURE 2.3: (a) The *electroscope*, an instrument designed by William Gilbert (1600) to detect static electricity. This image is reproduced from [182] originally published in 1902. (b) An early record of the human EEG, taken from Hans Berger's notes circa 1924. Both images have lapsed into public domain.

of telling us about the sources in the brain, may be understood. The *electric fields* and *electric potentials* generated in different parts of the head are explored. We consider typical source distributions found in those materials (Section 2.2), the properties of the materials in the human head (Section 2.3), and the fields generated when typical sources exist in this medium (Section 2.4). This chapter for the most part ignores the effects of the recording equipment and the noise or *artifact* present in EEG records. The dynamics of EEG behavior are introduced also, but details of its mechanisms are discussed in greater detail in Chapter 6.

The reader that is more interested in analysis of epileptic EEG rather than details of signal generation and measurement may skip through the majority of this chapter, although they may still find it worthwhile to browse through the boxed comments throughout the text that outline important observations.

Before moving on the specifics of the head as a volume conductor a succinct introduction to the electric field generation and measurement is provided. Even the electrical engineer or physicist that is familiar with this material may find it instructive to read this section because considerations relevant to 3-dimensional biological systems, which greatly differ from the conventional 1-dimensional analysis of electrical circuits, are outlined.

2.1 Principles of Bioelectric Phenomena

This story, like any other in electromagnetism, begins with single charges. Positive and negative charges are the fundamental units in electromagnetics – it is their position, movement and interaction that determine the fields and potentials in any medium. Although special considerations apply in biological systems because of the geometry, charge distributions and properties of the materials involved, bioelectromagnetism can be explained using the same principles as those used to explain behavior of a circuit in electrical engineering.

Moving charges produce *electric fields* as well as *magnetic fields* that interact with one another at high frequencies but are approximately independent at low frequencies. Because electrophysiological applications (including the EEG) involve the low frequency spectrum biological systems can be treated as *quasi-static* and the fields produced can be approximated by the equations in *electrostatics*. These are mathematically simpler because they ignore the direct contributions of magnetic fields. Electrostatics also implies that at any time the generated field is a result of a static distribution of charges and the explicit dependence of time on any source or consequent field can be removed. The time dependence in the EEG thus results from the dynamics of the system which evolve too slowly to affect the electrical measurement system.

This section, and the remainder of this chapter, gives an overview of the equations of relevance in EEG measurement of biological systems, that is, the *bioelectric* properties of the human head. Magnetic fields are largely ignored – to completely cover the field of bioelectromagnetic phenomena one would have to study the full set of Maxwell's equations that uniquely define interactions of charges. EEG can be understood quite well without this level of detail, but the interested reader can find excellent and complete explanations of traditional electromagnetism as well as bioelectromagnetism in texts such as [29], [76], [109] and [122].

After some preliminary words on notation, the fields generated from meso-scopic sources formed by grouping single charges together, denoted ϕ, are looked at first. This is followed by solutions to how groups of meso-scopic fields summate to generate a macro-scopic field Φ in a homogeneous volume conductor. Finally a brief outline of how Φ is modified when the volume conductor is not homogeneous is given.

2.1.1 A Foreword on Notation

In this book all quantities are presented in a 3-dimensional Cartesian coordinate system with values relative to an origin, shown in Figure 2.4 with axes labels x, y and z. Two types of quantities are referenced here: *scalars* and

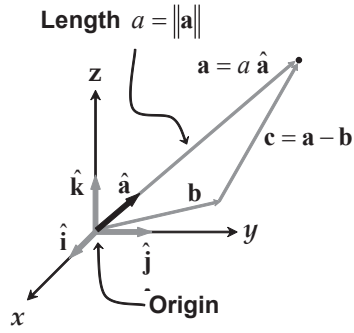


FIGURE 2.4: Vectors are quantities that have direction and magnitude whilst scalars only have magnitude. Notationally, vectors are differentiated from scalars by bold non-italic font. In this diagram $\mathbf{a}$ and $\mathbf{b}$ are example vectors defined in terms of an origin in 3-dimensional space. $\mathbf{c} = \mathbf{a} - \mathbf{b}$ is a vector defined in terms of other vectors so that this origin can be arbitrarily located. Vectors can also be re-defined in terms of a unit vector, which denotes direction, and a scalar magnitude. For example $\mathbf{a} = a\ \hat{\mathbf{a}}$, where $a = \|\mathbf{a}\|$ is the scalar length of $\mathbf{a}$, and $\hat{\mathbf{a}}$ is a vector of unit length in the same direction as $\mathbf{a}$.

vectors. A scalar gives only magnitude[2], whilst a vector has both magnitude and direction. For example, think of a broken down car that must be moved off the road. The magnitude of the force exerted to push it is a scalar quantity. However the direction of the pushing is also important because the car must be moved in a particular way. The direction of the force (e.g., forward) combined with its magnitude is a vector.

In this text vector quantities are differentiated from scalars by a bold non-italic font notation. For example, in Figure 2.4, $\mathbf{a}$ is a line that points from the origin to a point in 3-D space. Because $\mathbf{a}$ is directional it is a vector. Its length $a = \|\mathbf{a}\|$, where $\|\cdot\|$ denotes magnitude, is a scalar quantity because it is independent of the direction of $\mathbf{a}$. Vectors can also be defined in terms of each other, e.g., $\mathbf{c} = \mathbf{a} - \mathbf{b}$ as shown, so that the origin in the co-ordinate system can be made arbitrary.

A vector (e.g., $\mathbf{a}$ in Figure 2.4) of unit length, denoted $\hat{\mathbf{a}}$, can be used together with a scalar to describe any vector in the direction of $\hat{\mathbf{a}}$. It can be re-written as $\hat{\mathbf{a}} = a\ \hat{\mathbf{a}}$ where a is the scalar magnitude. Unit vectors $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ are commonly used to describe directions along the x, y and z axes, also shown.

[2]Allowing for negative and zero values also.

2.1.2 From Single Charges to Equivalent Dipoles

A single (and now static) charge[3] q suspended at location $\mathbf{r}_S$ in a medium with properties that are uniform in space (homogeneous) and direction (isotropic) produces an *electric field*, $\mathbf{E}(\mathbf{r}_R)$, that exerts a force on other charged particles. $\mathbf{E}(\mathbf{r}_R)$ is a vector quantity measured in force per unit charge (Newtons per Coulomb) pointing in the direction of this force location $\mathbf{r}_R$. It is given by

$$\mathbf{E}(\mathbf{r}_R) = \frac{q}{4\pi\epsilon_0 \|\mathbf{r}_R - \mathbf{r}_S\|^2}\hat{\mathbf{r}}, \tag{2.1}$$

where $\|\mathbf{r}_R - \mathbf{r}_S\|$ is the distance from source location $\mathbf{r}_S$ to recording location $\mathbf{r}_R$. $\hat{\mathbf{r}}$ is the unit vector pointing from the charge source to the measurement location which indicates that the electric field is also in this direction. ϵ_0 is the permittivity of the medium, measured in Farads per meter (F/m, described later). Because it describes a single charge Equation 2.1 represents as a *monopole* electric field. A summary of the SI units of each quantity and their relationships can be found in Appendix 2.A.

A quantity related to electric field is the *electric potential* $\phi(\mathbf{r}_R)$, also known as *voltage* and measured in Volts. Electric potential can be conceptually related to hydraulic pressure: if a pressure difference exists between two points in a connected water pipe then the water flows from a point of high pressure to one of low pressure. How fast the water flows and thus its potential ability to do work depends on the difference in pressure. The electric potential (difference) between two points in space is equal to the amount of work required to move a unit of charge between these two points. To be well defined the amount of work has to be independent of the path taken. The unit is work per charge, or J/C (Joules per Coulomb, see Appendix 2.A).

For charges, this 'pressure' is a consequence of the forces that exist in the form of electric fields defined by Equation 2.1. The potential ability to move charges and create electric currents, i.e., the electric potential, is determined by this field.

$\phi(\mathbf{r}_R)$ depends on recording location but is itself a scalar quantity because it is independent of the path. It is related to $\mathbf{E}(\mathbf{r}_R)$ at low frequencies as

$$\mathbf{E}(\mathbf{r}) = -\nabla\phi(\mathbf{r}) = -\left(\frac{\partial\phi}{\partial x}\hat{\mathbf{i}} + \frac{\partial\phi}{\partial y}\hat{\mathbf{j}} + \frac{\partial\phi}{\partial z}\hat{\mathbf{k}}\right), \tag{2.2}$$

where ∇ is an operator that calculates the gradient of a signal in the direction of maximum change. The electric field at a particular location can thus be interpreted to point in the direction of steepest descent of the electric potential, with magnitude equivalent to the steepness of this descent. Because of this relationship the electric field in Equation 2.1 has units of Volts per meter.

[3]Charges can have both positive or negative quantity. By convention electrons have negative charge and protons positive. Ions can have either. Charges with the same polarity repel and charges with opposite polarity attract.

The electric potential is not essential to explain electromagnetic phenomena and its definition makes sense only under the quasi-static assumption that magnetic fields can be ignored at low frequencies. However *electric potential is the quantity that EEG measures* and must be used to explain these records. The remainder of this text uses 'fields' to describe both electric and potential fields, related to each other as above.

The electric potential produced by a monopole located at $\mathbf{r}_S$ can be computed from Equations 2.1 and 2.2 as

$$\phi(\mathbf{r}_R) = \frac{q}{4\pi\epsilon_0 R}, \tag{2.3}$$

with $R = \|\mathbf{r}_R - \mathbf{r}_S\|$ the distance between the source and the point of interest[4].

Now consider a collection of N charges in space, each generating potentials $\phi_1(\mathbf{r}_R)...\phi_N(\mathbf{r}_R)$. Since the electric field equations are linear the total field can be calculated by a linear combination of the individual charges – this is known as the *principle of superposition.* Thus,

$$\phi(\mathbf{r}_R) = \sum_{n=1}^{N} \phi_n(\mathbf{r}_R) = \frac{1}{4\pi\epsilon_0} \sum_{n=1}^{N} \frac{q_n}{R_n}. \tag{2.4}$$

When there are only 2 charges ($N = 2$, as in Figure 2.5(a)) with opposite polarities then it can be shown that at location $\mathbf{r}_R$ sufficiently distant from the sources

$$\phi_{dipole}(\mathbf{r}_R) = \phi_1(\mathbf{r}_R) + \phi_2(\mathbf{r}_R) \approx \frac{q}{4\pi\epsilon_0} \frac{cos(\theta)}{R^2}, \tag{2.5}$$

where θ is the angle between the direction of the charges and $\mathbf{r}_R - \mathbf{r}_S$, as illustrated in Figure 2.5(a). Here sufficiently distant refers to a distance R large compared to the separation of the two charges d – that is $R >> d$. The above configuration of charges is known as an *electric dipole*, and it is important for bioelectric applications for reasons explained later. The potential is strongest in the direction of the two charges, and weakest perpendicular to it.

Equation 2.5 also shows that (at large distances) the potential generated by an electric dipole decreases with distance as $1/R^2$, whereas a monopole in Equation 2.3 decays slower as $1/R$. This is because there is a significant cancellation effect when charges of equal magnitude but opposite polarity are placed close to one another. Similar cancellation effects occur with configurations involving more than two charges. A quadrupole (containing 4 unit charges with zero nett charge) has a field that decays as $1/R^3$, an octupole (8 charges) as $1/R^4$, etc.

[4]It may be useful at this stage to look at the co-ordinate system in Figure 2.5(a) and (b). These figures are not for a single charge but the vectors involved are the same.

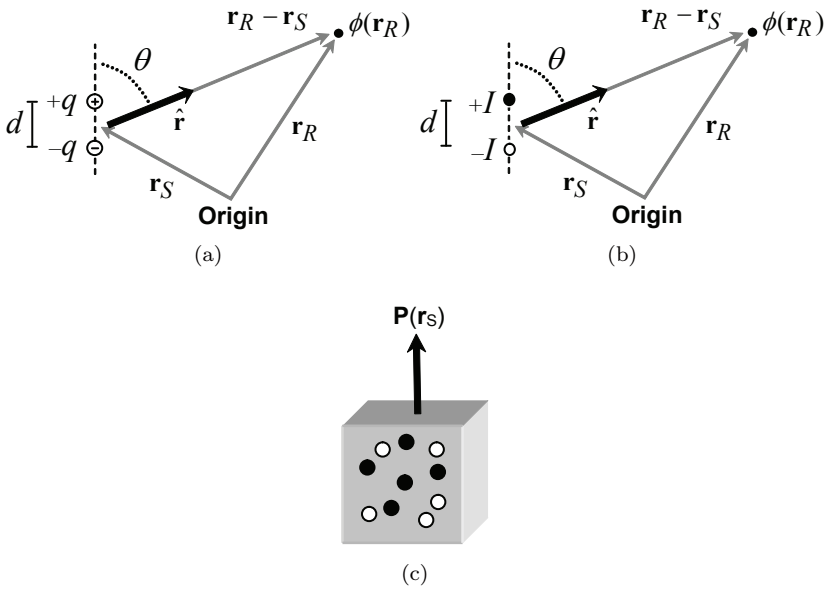


FIGURE 2.5: Representative dipoles – (a) electrical, (b) current and (c) equivalent electrical or current dipole $\mathbf{P}(\mathbf{r}_S)$, under the assumption of small volumes measured at large distances from $\mathbf{r}_S$, with equal number of positive and negative monopoles within.

Since all configurations of charges may be expressed as the sum of monopoles, dipoles, quadrupoles, octupoles and an infinite number of higher order terms, the electric potential for N charges at a large distance can be expressed as the summation over all these contributions. Furthermore, *because terms of higher order than a dipole decay very fast with distance, the field at large distance may be approximated by the monopole and dipole contributions alone.*

Under the assumption that the volume is electrically neutral (i.e., no nett charge) then the potential does not contain a monopole term, and at a distance away from this volume the potential is well approximated by the dipole term alone. The latter is called the *equivalent dipole* (see Figure 2.5(c)).

The equivalent dipole is an *approximation* to the potential $\phi(\mathbf{r}_R)$ generated

by a volume with many charges when R is much larger than the diameter of the volume in question. The unit volume covered by an equivalent dipole is known as a *voxel*. The arrow in Figure 2.5(c) indicates the direction of the nett field generated by the volume.

An equivalent dipole is important in calculations of electric and magnetic fields because it allows the simplification of a very complex system involving many billions of charges to one involving a finite number of elements. A voxel may contain a large number of charges, but as a whole it can be represented as an equivalent dipole with direction and magnitude representative of the configuration of the charges within the volume. Again this is valid only at large distances relative to volume size.

The problem with using equivalent *electric* dipoles in biological systems is that the assumption that positive and negative charges balance each other may not be valid. Monopole contributions cannot necessarily be ignored. Instead, it is more practical to consider *current dipoles*, discussed next.

2.1.3 Equivalent Current Dipoles

A *current dipole* is defined in terms of *current sources* and *sinks* rather than positive and negative charges, as shown in Figure 2.5(b). By convention, a current source is a positive monopole, whereas a sink is a negative monopole.

Mathematically current monopoles are analogous to Equation 2.3. The principle of superposition implies that many sources and sinks contribute to the overall electric potential linearly

$$\phi(\mathbf{r}_R) = \sum_{n=1}^{N} \phi_n(\mathbf{r}_R) = \frac{1}{4\pi\sigma} \sum_{n=1}^{N} \frac{I_n}{R_n}, \tag{2.6}$$

where the current I_n determines the strength and polarity of each current monopole flowing into a medium with conductivity σ, measured in Siemens per meter (S/m, described later).

The combination of one positive and negative current monopole yields a current dipole; the voltage now decreases in strength as $1/R^2$ at large distances

$$\phi_{dipole}(\mathbf{r}_R) \approx \frac{Id}{4\pi\sigma} \frac{cos(\theta)}{R^2}. \tag{2.7}$$

The strength is proportional to the separation d.

Figure 2.5(c) is true also for a volume containing multiple current dipoles with approximately equal number of current sinks and sources. Unit volumes can be described by an *equivalent current dipole*, where the strength and orientation is determined by the distribution of the monopoles within. For a current dipole the arrow in this figure indicates the nett direction of current flow in the volume. It is equivalent to the electric dipole because this current occurs in the direction of the nett field formed by the charges.

At this point electrical engineers will wonder *'How can a current source*

or sink exist? Currents require closed loops and charge must always be conserved.' This is true, and also in this case charge does not magically appear or disappear. However the properties of the biology allow approximations by which current sources and sinks can be used. These approximations are discussed in Section 2.2.1. For now, regardless of the feasibility of current dipoles, the important result is that current dipoles are mathematically equivalent to charge dipoles from the voltage point of view.

2.1.4 Macro-Scopic Mean Fields – Homogeneous Media

Macro-scopic potentials $\Phi(\mathbf{r}_R)$ at low frequencies, recorded at location $\mathbf{r}_R$ and generated by many voxels, each represented by an equivalent current dipole $\mathbf{P}(\mathbf{r}_S)$, are calculated by integrating the contribution of all sources contained within the volume conductor.

$$\Phi(\mathbf{r}_R) = \sum_{volume} \sum \mathbf{P}(\mathbf{r}_S) \,.\, \mathbf{G}(\mathbf{r}_S, \mathbf{r}_R). \tag{2.8}$$

$\mathbf{G}(\mathbf{r}_S, \mathbf{r}_R)$ is known as the *Green's function* of the volume conductor. It contains all the information about the electrical properties of the material and its geometry. Each equivalent dipole $\mathbf{P}$ located at $\mathbf{r}_S$ is projected (by the dot product, denoted .) to recording location $\mathbf{r}_R$ through the volume conductor. The effects of the volume conductor on $\mathbf{P}$ between $\mathbf{r}_S$ and $\mathbf{r}_R$ are given by $\mathbf{G}(\mathbf{r}_S, \mathbf{r}_R)$. The total potential $\Phi(\mathbf{r}_R)$ is the linear superposition of the projection of all dipoles $\mathbf{P}$ onto location $\mathbf{r}_R$.

For an ideal case where all sources are current dipoles and there is only one homogeneous material in the volume conductor, the geometrical dependence of Green's function is only on distance and direction.

$$\mathbf{G}(\mathbf{r}_S, \mathbf{r}_R) = \frac{\mathbf{r}_R - \mathbf{r}_S}{4\pi\sigma\|\mathbf{r}_R - \mathbf{r}_S\|^2}. \tag{2.9}$$

The electrical properties of the material (σ) are also part of the above equation, but play a minor role in the case of homogeneity.

For non-magnetic materials two quantities can unambiguously describe its electrical properties

1. *Permittivity* $\epsilon(f)$ describes the ability of a material to store charge. Permittivity (or dielectric constant) is measured in Farads per meter (F/m).

   Permittivity is a property of the material that does *not* depend on its geometry, whereas *capacitance*, measured in Farads, does.

2. *Conductivity* $\sigma(f)$ is a measure of the material's ability to conduct

electric charge, that is, how easy it is for charge to move through it. Higher conductivities allow charges to flow more freely. It is measured in Siemens (S= $1/\Omega$) per meter. Its inverse is called *resistivity* $\nu(f)$ and may be used instead.

Conductivity and resistivity are properties of the material that do *not* depend on its geometry, whereas *conductance* and *resistance* will.

When the material is not homogeneous and isotropic its geometry also plays a role.

It is shown in Section 2.3 that for EEG frequencies capacitive effects can be ignored, and conductivity is independent of frequency as implied in Equation 2.9. *Conductivity σ is the most important quantity in determining the potentials $\Phi(\boldsymbol{r}_R)$ at EEG frequencies.*

2.1.5 Macro-Scopic Mean Fields – Inhomogeneous Media

When a volume is not homogeneous, that is, the properties of the material are not uniform throughout, the differences in material properties become important in estimating $\Phi(\mathbf{r}_R)$. Potentials between two boundaries of materials with different conductivities (σ_{m1} and σ_{m2}) are expected to change, but must comply to the following set of *boundary conditions*

$$\sigma_{m1}\frac{\partial\Phi_{m1}}{\partial\mathbf{n}} = \sigma_{m2}\frac{\partial\Phi_{m2}}{\partial\mathbf{n}} \tag{2.10}$$

$$\frac{\partial\Phi_{m1}}{\partial\mathbf{t}} = \frac{\partial\Phi_{m2}}{\partial\mathbf{t}}, \tag{2.11}$$

where $\mathbf{n}$ refers to directions normal to the surface and $\mathbf{t}$ to components tangential to the surface.

Equation 2.10 says that the change of the potential in the direction normal to the boundary between two media is inversely proportional to the conductivity in that medium. It is a consequence of conservation of charge – it describes the necessity for current density to be continuous across the surface. Current cannot appear and disappear at the boundary.

Equation 2.11 says the change of the potential in a direction tangential to the boundary between two media is the same on either side of this boundary. It is a consequence of conservation of work – it dictates that potentials must be continuous across a boundary. To see how the above equations are derived, refer to Appendix 2.B.

The *geometry* of the materials is also important. A volume conductor taken as a whole can be inhomogeneous and possibly anisotropic even when each region is approximated as both homogeneous and isotropic. Equation 2.9

must be adapted to incorporate all this information. As a consequence, the corresponding Green's function may become complex and numerical methods must be employed to compute Equation 2.8.

2.2 Current Sources in Biological Tissue

This section focuses on the generators of electrical activity within the brain. It is accepted that the charge (or current) distribution in the brain is responsible for generating the EEG waveforms. No sources of interest exist in CSF, skull or scalp. In fact, it is the *purpose* of the EEG to measure activity in the brain. Activity arising from any other source is responsible for the artifacts, for example muscle from the scalp, eye blinks, as well as others described in Chapter 1.

This section explores the type of activity (charge and current distributions within the brain) that is expected to affect EEG the most.

The human brain is composed of excitable neural tissue interacting with itself. The fundamental unit of the brain is the *neuron*, which integrates positive and negative inputs from dendritic currents produced by incoming action potentials generated by other neurons. When a threshold of activity is reached the neuron in turn generates its own action potential, thereby continuing the transmission of information. The resulting macro-scopic EEG recordings are predominantly influenced by the micro-scopic synaptic structure (Section 2.2.1), meso-scopic and micro-scopic cortical structure (Section 2.2.2) as well as the temporal distribution of activity (Section 2.2.3).

2.2.1 Synaptic Structure and Current Dipoles

In the EEG the charges responsible for generating electric fields are those contained in the cortical layers of the brain (as explained in detail in Section 2.2). Ideally the distribution of *all* charges contained within the cortex should be used to estimate these fields. This is a computationally intractable problem that is made simpler by approximating small discrete volumes or *voxels* as one equivalent electric dipole, as discussed in Section 2.1.2. The total potential at the EEG recording site is then estimated by the principle of superposition.

However representing each voxel as an equivalent electric dipole requires knowledge of the charge distribution *within* it. No such knowledge exists. Furthermore, because charge concentrations vary across the cortex the assumptions that allow the division into discrete voxels approximated as a single electric dipole do not hold.

Instead electrophysiologists like to use a *current dipole* to describe each voxel. As discussed in Section 2.1.3 real current dipoles do not exist because any current always requires a return path – it never appears or disappears

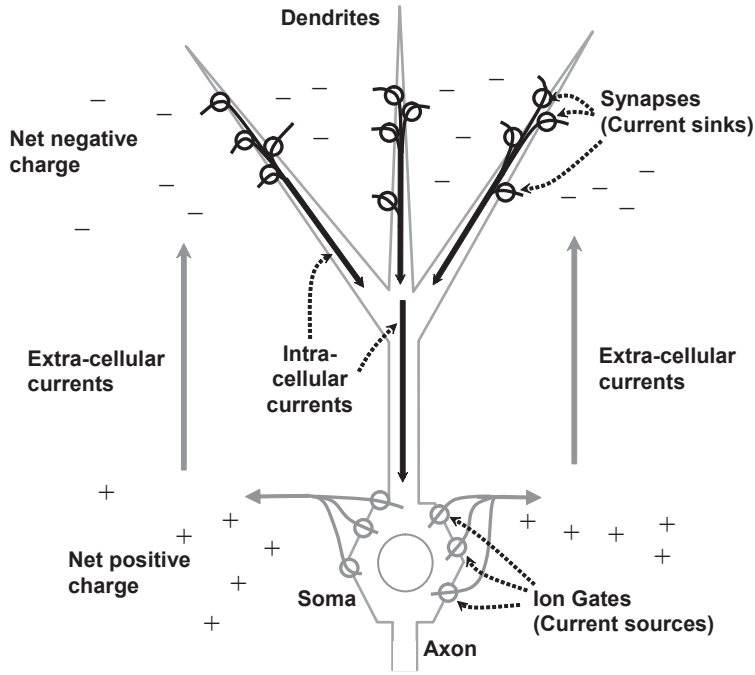


FIGURE 2.6: An example of the current loops formed by a single neuron. Current (in this case positive ions) flows into and out of the cell only at ion gates and synapses. This trans-membrane flow creates charge differentials: a net negative charge at the current sinks, denoted −, and a net positive charge at the current sources, denoted +. Extra-cellular currents are caused by this difference in charge distribution. In the cortex these extra-cellular currents are normal to the surface of the cortex because the neurons are aligned in this direction. The return intra-cellular currents can be ignored for the purposes of describing the EEG because of the highly resistive cell membrane. Because ion gates are very small relative to the size of the neuron the flow of ions into and out of the cell can be approximated as *current sources* and *current sinks*. This approximation allows the use of current dipoles instead of electric dipoles to model sources within the cortex.

from nowhere. However when estimating the EEG certain physiology in the brain can be *approximated* as current dipoles.

To see this, refer to Figure 2.6 where an example closed current loop generated by a single neuron is shown. Positive and negative charges (ions) in the brain can flow in the intra-cellular or extra-cellular space. Very little current (on average) flows through the membrane of neurons because they are highly impermeable to ions under normal conditions. At certain points (e.g., synapses and ion gates) the neuron membranes can adapt their permeability to specific ions in response to incoming stimuli, resulting in a sudden influx or efflux of current into the neuron cell. These cause a charge differential that generates extra-cellular currents. Inside the neuron the intra-cellular currents form a return path so that charge is conserved. Both intra-cellular and extra-cellular currents are shown in this figure.

Because the membrane is highly resistive, and as a direct result of the boundary conditions listed in Section 2.1.5, the potentials generated by the intra-cellular currents can be ignored at macro-scopic scales. As far as the EEG measurement is concerned these currents do not exist – *it is only the extra-cellular currents that contribute to the EEG*. Furthermore, because the physical size of synapses and gates is very small (relative to the size of the neuron), the *flow of current across the membrane can be approximated as a current source or a sink*.

This approximation allows current dipoles to be used as a model for the brain generators, helped along by the availability of a significant volume of knowledge on the distribution of gates in the neocortex. Equivalent information about charge distributions is not as readily interpretable. Furthermore, the structure of the cortex is such that discrete voxels can be defined so long as these are large relative to the size of a neuron. Within a voxel there is an equal number of current sources and sinks because each current source or sink is compensated by current sources and sinks distributed on other regions of the dendritic and somatic structure. Thus monopole contributions can be ignored and using a single current dipole $\mathbf{P}(\mathbf{r}_S)$ to describe a voxel is valid.

At macro-scopic scales the sudden flow of positive and negative ions across the cell membrane, caused by gating mechanisms that include EPSPs, IPSPs, voltage-triggered gates and ion pumps, can be approximated as current *sources* or *sinks* respectively. This is possible because

- These gates are small relative to the cell body, and
- Intra-cellular return currents do not affect the macro-scopic EEG.

Because statistical information of the distribution of these gates exists current rather than electric dipoles are a better choice

to model EEG generation – current sources and sinks exist naturally in the brain.

Only extra-cellular currents affect the EEG. A single dipole $\mathbf{P}(\mathbf{r}_S)$ oriented in the direction of this current flow can be used to describe a discrete volume or *voxel* of cortex because an equal number of current sources and sinks exist within it.

A suitable scale for a voxel depends on the composition of the cortex. Each cortical neuron has between $10^3 - 10^4$ synapses. A minicolumn is composed of about 100 pyramidal cells and about 10^6 synapses and covers a ~ 0.03mm diameter of cortex (see Figure 2.2). In turn each mm^3 (macro-column) of cortex contains in the order of 10^5 neurons and 10^9 synapses [168]. To effectively model both intra-cranial and scalp EEG, whilst maintaining computational requirements at a reasonable level, a scale in between that of a minicolumn and a macrocolumn has been suggested in [121]. Larger scales may be effectively used to model scalp EEG with only small errors [114].

2.2.2 Spatial Integration

In order to solve the forward problem the location and orientation of these dipoles is important. In this section it is shown that significant simplifications can be made based on the organization of the brain.

Imagine a volume conductor in which N equivalent dipole sources with unit magnitude are present. Each source is in the form of $\mathbf{P}(\mathbf{r}_S)$ and the total potential recorded at a sufficiently distant location $\mathbf{r}_R$ in a homogeneous medium is given by Equation 2.8. If all N sources are oriented in a random manner then the total expected recorded potential is, on average, equal to zero, because the potentials formed by the dipole sources destructively interfere. The energy of the recorded potential for each trial fluctuates randomly but is proportional to $\sqrt{N}$, the standard deviation of the distribution. A simulation to verify this is presented in Figure 2.7(a). The normalized voltage is computed for 1000 trials, each with $N = 50,000$ randomly oriented sources, and the overall probability distribution of resultant voltages is shown. On average, the expected voltage has a mean of 0 and a standard deviation (energy) proportional to $\sqrt{N} \approx 220$.

In contrast, if all N sources are aligned the resulting field is proportional to N because of constructive interference. Figure 2.7(b) simulates the number of N parallel sources required to generate an equivalent field to M randomly aligned sources. With $N = 200$ fields in the same order as $M = 40,000$ can be created, or equivalently, less than 1% of the field is due to the $M = 40,000$.

Aligned sources produce a voltage that is much stronger than an equal number of randomly oriented ones. The potentials generated by N aligned sources become undetectable only when there are $\gg N^2$ randomly aligned ones. This relationship is shown in Figure 2.7(b).

Since EEG recorded voltages are not randomly distributed (patterns are observed) and non-zero the dipole orientations must not be random. Histological classification of different regions of the brain has revealed that although sources in deeper structures are approximately randomly oriented, the pyramidal neurons in the cortex have their dendritic structure aligned. The effects of this are discussed next.

2.2.2.1 Cortical Structure

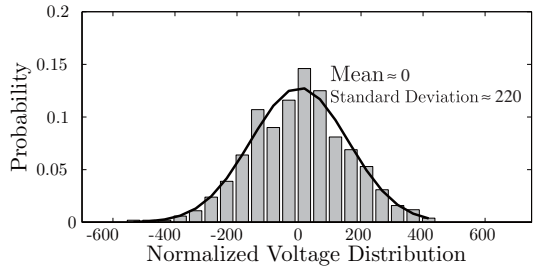
Approximately 85% of neurons in the cortex are pyramidal cells that are aligned in columnar structures normal to the cortical surface [121, 168]. The current sinks and sources in this configuration occur at different levels along this structure. One such configuration, where all current sinks are on the dendrites and all current sources at the cell body, is shown in Figure 2.6.

Regardless of the distribution of current sources and sinks, the current flow is vertical because the sinks and sources produce tiers of charge differentials, also shown in Figure 2.6. Vertical current flow is promoted by the alignment of neural fibers in this direction. Hence the meso-scopic source approximation $\mathbf{P}(\mathbf{r}_S)$ is always fixed in an orientation normal to the cortical surface, as originally depicted in Figure 2.5(c).

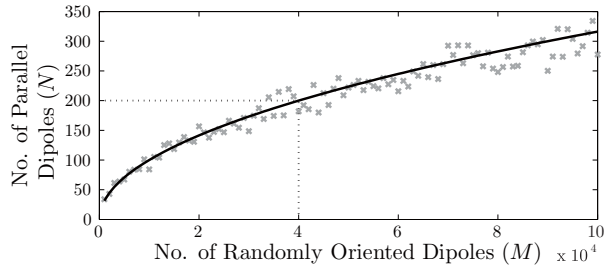
The neocortex is sufficiently thin ($\sim$ 2mm) so that the voxel represented by $\mathbf{P}(\mathbf{r}_S)$ can span the depth of the cortex. As such, and because there are no sources below the neocortex that significantly contribute to the EEG (due to the random neuron structures as well as their distances from recording sites) the cortex ($1500 - 3000\text{cm}^2$ in its entirety, including fissures) may be modeled as a folded sheet rather than a volume of dipoles [114, 122]. The potential at an arbitrary recording site $\mathbf{r}_R$ originally defined as a volume integral in Equation 2.8 can instead be simplified to a surface integral

$$\Phi(\mathbf{r}) = \sum_{cortical} \sum_{surface} \mathbf{P}(\mathbf{r}_S) \cdot \mathbf{G}(\mathbf{r}_S, \mathbf{r}_R), \tag{2.12}$$

where in this case the orientation of $\mathbf{P}(\mathbf{r}_S)$ is always perpendicular to the cortical surface. This dipole sheet is seen to be the major contributor to scalp recorded EEGs.



(a) Voltage distribution – randomly oriented dipoles



(b) Equivalent voltage of parallel vs. random sources

FIGURE 2.7: (a) The probability distribution of voltages generated by 1000 trials, where in each trial $N = 50,000$ randomly oriented dipoles in a volume conductor are simulated. A source generates a normalized voltage with magnitude of 1, but with random orientation. Mean $\mu \approx 0$ and energy (standard deviation) is proportional to $\sqrt{N} \approx 220$. (b) Relationship between the number of random and parallel orientated sources that can generate equivalent fields. Simulations are averaged over 1000 trials. An example is drawn where 200 parallel sources generate a field proportional to $40,000$ randomly oriented dipoles, that is, less than 1% of the field is due to the $40,000$ randomly oriented sources. Hence parallel sources are most likely the generators of EEG.

A current dipole $\mathbf{P}(\mathbf{r}_S)$ on the cortex has orientation normal to the surface because extra-cellular current flow occurs in this direction. This alignment makes cortical activity the single-most important contributor to EEG measurements.

A sheet rather than a volume of dipoles can be used to model the total potential $\Phi(\mathbf{r})$ because the cortex is thin.

2.2.2.2 Cortical Folds

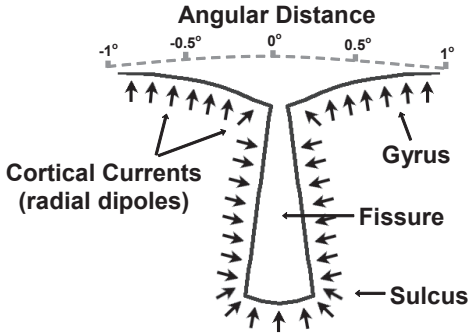
A dipole sheet is appropriate to model the cortex in its entirety, although special care must be taken with the interpretation of these dipoles. $\mathbf{P}(\mathbf{r}_S)$ is normal to the cortical surface, *but not necessarily radially oriented in the spherical head* because the cortex is a folded structure. The crown of a fold is called a *gyrus* and dipoles here are predominantly radial. The valley is called a *sulcus* and dipoles on the walls of the sulci are predominantly tangential. Both are shown in Figure 2.8(a), where each arrow is representative of the current flow (or the equivalent current dipole) of a small volume of cortical tissue. The folds in the cortex are another reason that the diameter of the voxel represented by $\mathbf{P}(\mathbf{r}_S)$ must be small – for the orientation of $\mathbf{P}(\mathbf{r}_S)$ to remain constant over this volume then the pyramidal structures must be aligned within it.

The activation of large areas of the cortex simultaneously implies that frequently the entire sulcus and gyrus are simultaneously and synchronously active. In this case, tangential dipoles in the fissures tend to cancel out. This is shown in Figure 2.8(b), where the simulated potential generated by the entire structure in (a) varies very little from that generated only by the gyrus.

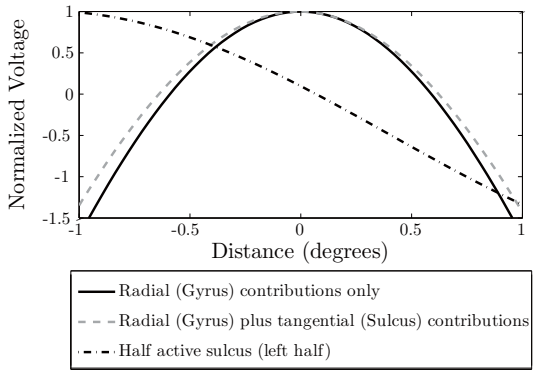
When the entire sulcus and gyrus is synchronously active the tangential dipoles on the fissures affect recorded potentials very little and can be ignored. Only radial contributions are important.

If only part of the sulcus is active this cancellation effect does not occur and the tangential dipoles do contribute to EEG recordings. This is also shown in Figure 2.8(b) where simulation is repeated when only the left half of the sources in (a) are active. However, it is worth noting that the tangential dipoles are generally located deeper within a structure, thereby contributing less than a radial dipole of the same magnitude. This is discussed in more detail in Section 2.4.

The simulations presented in Figure 2.8 concern very regular fissures and sulci that are not typical of the irregular foldings observed in a real cortex. However, the trends indicate that tangential dipoles do not affect the intracranial EEG because the electrodes record local activity and are not sensitive to relatively distant sulci. Scalp electrodes record larger areas but the contribution of the sulci remains minimal because of the distances involved (see Section 2.4). Thus in modeling the EEG the cortical folding is typically ignored and the assumption that all dipoles are radial is enforced.

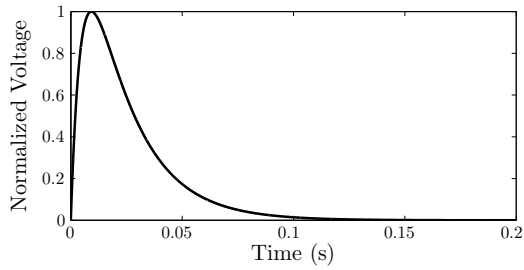


(a)

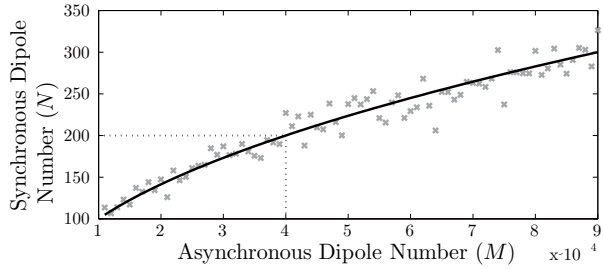


(b) Effects of including Dipoles in a sulcus

FIGURE 2.8: (a) An idealized configuration of current dipole sources present in a sulcus, used for simulation. Each arrow represents a compartment or voxel on the cortex. Each voxel is assumed to be approximately self-contained in that charge is conserved and a current loop such as that shown in Figure 2.6 is present. Only extra-cellular currents contribute to the EEG. In (b) normalized potentials are simulated at the different angular locations shown in (a), under the assumption that all sources are synchronously active. The potential generated when all sources (gyrus and sulcus) are included in simulation is very similar to those when only the gyrus is included. This is because of the significant cancellation effects that result in minimal contributions from tangential dipoles. In this case the contributions of the sulcus can be ignored. When only half the sources are active (e.g., left half, simulated in (b)) these cancellation effects do not exist and the contributions of sulci cannot be ignored. However, at scalp locations the distances are large enough so that tangential contributions remain relatively small.



(a) Dendritic Response



(b) Equivalent voltage of synchronous vs. asynchronous sources

FIGURE 2.9: (a) Dendritic response used for simulation, with typical rise and fall times as described in Chapter 6. (b) Relationship between the number of synchronized sources N and asynchronous sources M, averaged over 1000 trials. This relationship is similar to that observed in Figure 2.7(b), where the synchronous activity of few sources generates large responses that can be, on average, only equaled by many asynchronous sources. Equal magnitude responses occur when $M \approx N^2$.

2.2.3 Temporal Integration

So far all time dependence of potentials has been ignored because of the quasi-static approximation that allows the potential field at any one time to be described by static sources. However it is necessary to consider how the potentials generated are a consequence of the temporal arrangement of dipoles. This arrangement is a result of the dynamics (discussed in Section 2.4.2 and in more detail in Chapter 6) but the resulting measured field is dependent on the *synchronicity* of the dipoles.

When the mesosources in the cortex are synchronized the neurons in this area behave similarly. They are active at the same *time* and with the same *phase* (that is, the flow of current is in the same direction). The minicolumn, comprising of about 100 neurons, has been proposed as the basic functional unit of neural circuits. Neurons within a minicolumn are expected to mostly

act in synchrony [121]. At the macro-scopic scale, however, recordings are affected by very large areas (in the order of 20cm^2 for scalp recordings, and 1cm^2 for intra-cranial recordings, see later). Even for intra-cranial recordings these areas can contain hundreds of thousands of minicolumns. How does the synchronicity between these affect the EEG?

Figure 2.9 describes this relationship. A typical dendritic tree response to incoming action potentials (shown in (a)) with random polarity (positive or negative) is distributed uniformly over time, and integrated to generate the resulting EEG. Simulations of the equivalent energy of the potentials resulting from N synchronized versus M un-synchronized sources are plotted in (b). This relationship is similar to that observed in Figure 2.7, that is, synchronized sources contribute potentials proportional to N, whilst un-synchronized sources only contribute energy in a random manner, with contributions in the order of $\sqrt{M}$. Hence very few synchronized sources can register large potentials on the EEG, and many more un-synchronized sources are necessary to generate potentials of equivalent magnitude.

EEG fields are dominated by the amount of synchronized *activity,* not *the quantity of activity.*

Given that the majority of neurons within a column are on average inactive [121], and that EEG potentials are relatively large, a natural conclusion is that large areas of cortex must be synchronized to generate the observed EEG waveforms. Fortunately, the brain structure is such that cortico-cortical fibers are capable of activating large areas of cortex at any one time. If this were not the case, the observed EEG would be meaningless.

This also explains why action potentials have been ignored in the discussions relating to the measurement of EEG. For a signal to register on the EEG there must be a high level of synchronization between sources. Action potentials are short in duration (much higher frequencies) and so they do not superimpose as often as do synaptic return currents, which are slower and have a much longer response (see Figure 1.5(b)). Action potentials *do* contribute to the EEG, but their contribution is minimal.

The synaptic return currents are the predominant contributors to the global EEG. Action potentials are important for the underlying dynamics but not to model the EEG [118].

Tissue	Conductivity (S/m)	σ/σ_{brain}
σ_{brain}	0.3	1
σ_{csf}	1.5	5
$\sigma_{skull-radial}$	0.015	1/20
$\sigma_{skull-tangent}$	0.075	1/4
σ_{scalp}	0.3	1

TABLE 2.1: Nominal head tissue conductivities.

2.3 Volume Conducting Properties of the Head

Once the sources are known an accurate estimate of the potentials $\Phi(\mathbf{r}_R)$ at the scalp or at any other location in the human head requires a sufficiently accurate description of its volume conducting properties. However not only is the geometry of the head complex, the different tissues found in it are both inhomogeneous and anisotropic, that is, their electrical properties differ with location as well as direction.

The lack of accurate knowledge of these properties means that certain problems, such as solving the inverse problem (localization of sources in the brain from scalp measurements), are prone to error. However the purpose of this chapter is to obtain a reasonably good qualitative evaluation of solutions to the *forward problem* (determining fields at the scalp generated by sources in the brain) and only rough quantitative concordance is necessary. Many simplifying assumptions make this problem easier.

The remainder of this section discusses simplifications and experimental evidence that allow a reasonable estimation of the fields at any point within the human head, given an arbitrary source distribution. Simplifications based on geometry are discussed first, followed by the electrical properties of the materials in the head. This is followed by a review of experimentally observed ranges of these electrical properties.

2.3.1 Head Geometry

The human head may be divided into approximate layers each composed of a different type of tissue. The first and most important simplification in the calculation of potentials is to approximate the geometry of the head as a set of concentric spheres shown in Figure 2.10, where each region represents one of the layers of the head. This is a fairly accurate model within a small section of the scalp, which is itself roughly spherical, but not for the entire head [103].

The number of concentric spheres used varies depending on the required detail, but can include the brain (gray and white matter), cerebrospinal fluid

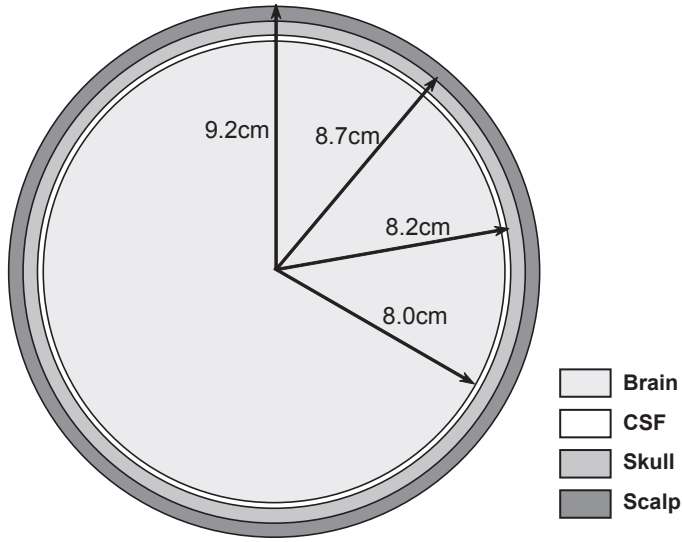


FIGURE 2.10: The 4-sphere approximate model of the human head.

(CSF), skull (two compact layers and one spongy layer) and scalp. Simulations in this chapter are performed using the 4 sphere layer model, which includes the effects of the averaged brain, CSF, averaged skull and scalp. The nominal radii chosen for these layers is depicted in Figure 2.10. Implied in this model is a fifth layer, air, that has very low conductivity so that the number of charges in the head remains constant. It should be re-iterated that these values are chosen for representative purposes only – it is well known that the radii of the skull vary with location on the head, or that the thickness of the CSF increases with age. It is not the accuracy of these measures that is important here, but rather their ability to provide qualitative descriptions. Results employing these approximations have in any case been shown to be accurate to within 10-20% when compared to more realistic head models [172, 103, 114].

An important consequence of this reduction in complexity is that Equation 2.12 for typical source distributions can be expressed as the weighted sum of basis functions known as spherical harmonics. These are applicable only to systems that can be described in a spherical volume and are the space analogy to the decomposition of time-dependent signal as the sum of sines and cosines. Equation 2.12 can be rewritten as

$$\Phi(\mathbf{r}_R) \approx \sum_{n=0}^{\infty} \sum_{m=-n}^{n} p_{n,m}\ g_{n,m}(\mathbf{r}_S, \mathbf{r}_R) Y_{n,m}(\theta, \xi), \tag{2.13}$$

where (θ, ξ) are spherical co-ordinates defined in Figure 2.11(a), n is the harmonic number that is a measure of approximate spatial frequency, m is

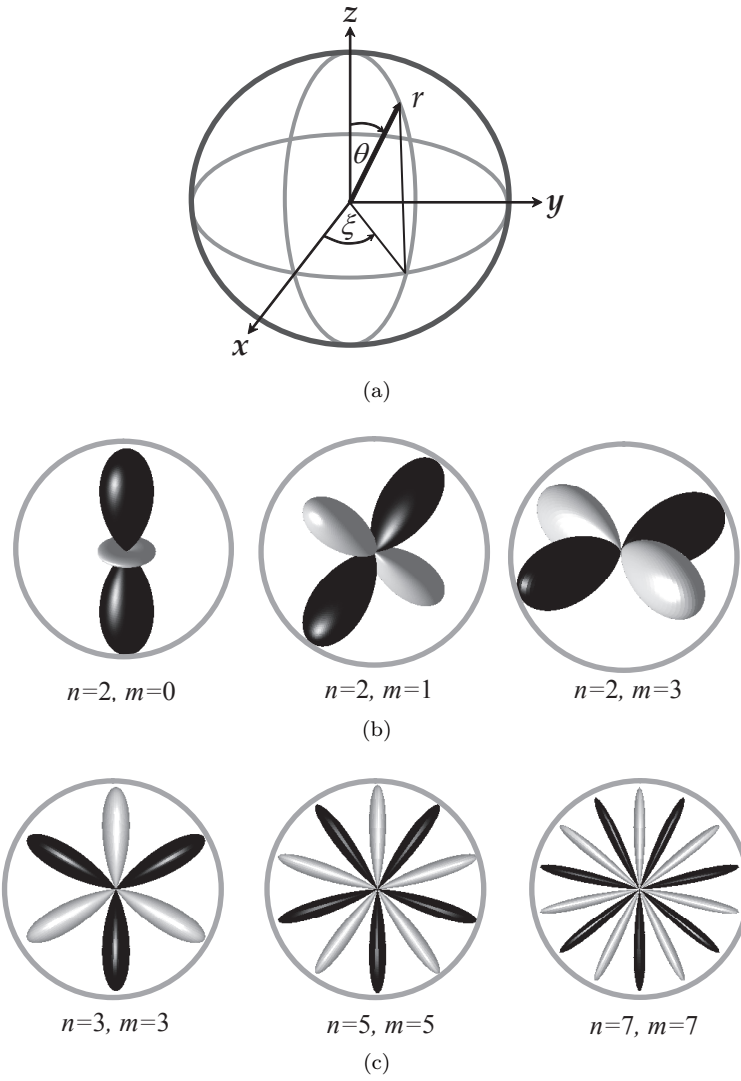


FIGURE 2.11: (a) Spherical co-ordinate system (θ, ξ, r) relative to the Cartesian co-ordinate system (x, y, z). The angles θ (elevation), ξ (azimuth) and radius r are defined as shown. In (b) and (c) are example $Y_{n,m}$ – the spherical harmonic basis functions in Equation 2.13. Black represents a negative number, and gray a positive number. Any distribution within a sphere can be defined in terms of these basis functions, summed over all possible (n, m). In (a) an example is shown in which the harmonic number n remains constant and m is varied. m changes the direction of the basis function at harmonic number n. In (b) n is varied to show that higher harmonic numbers represent higher spatial frequencies or finer spatial scales.

the order that represents the different directions of each n and $Y_{n,m}$ is the spherical harmonic basis function associated with (n, m). The coefficients $p_{n,m}\ g_{n,m}(\mathbf{r}_S, \mathbf{r}_R)$ replace $\mathbf{P}(\mathbf{r}_S)$. $\mathbf{G}(\mathbf{r}_S, \mathbf{r}_R)$ (the description of source distribution and green's function) in Equation 2.12. These quantities are now dependent on n and m.

Each spherical harmonic $Y_{n,m}$ is orthogonal (non-overlapping, independent) to all other harmonics so that Equation 2.13 has zero redundancy. Examples for $n = 3$ and $m = 0, 1, 2$ can be seen in Figure 2.11(b). Here the approximate scale of components in $Y_{n,m}$ is the same when n is constant, but its direction depends on m. Another example in (c) shows that different n relate to basis functions of different spatial frequencies.

The importance of Equation 2.13 is that it is an analytic solution, and is tractable numerically.

2.3.2 Capacitive Effects of Tissue

This section shows that at macro-scopic scales the capacitive effects of volume conduction are minimal. This is not the case for micro-scopic activity, where it is the capacitive properties of cellular membranes that allow much of the observed dynamic behavior.

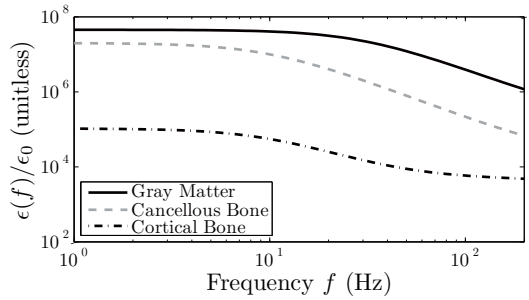
Limited research exists on the frequency dependent permittivity of biological tissues. Conductivity, on the other hand, has been widely studied and demonstrated roughly constant at EEG frequencies $(0 - 100\text{Hz})$ [5, 157]. A review of the frequency dependence of both conductivity and permittivity is presented in [46, 47, 48]. Nominal curves for the permittivity and conductivity of gray matter, cortical (compact) bone and cancellous (spongy) bone generated with the model presented in this review are shown in Figure 2.12(a) and (b). These curves are consistent with values used in Table 2.1.

If a volume conductor is linear and its electrical properties are uniform across space (isotropic), then the total resistance of a path through the conductor is simply the resistivity $(\frac{1}{\sigma(f)})$ multiplied by its length (L). A linear volume conductor can be approximated as an electrical circuit consisting of a resistor in parallel with a capacitor [103] (see Figure 2.13). For such a system it is shown in Appendix 2.C that the capacitive effects can be ignored if

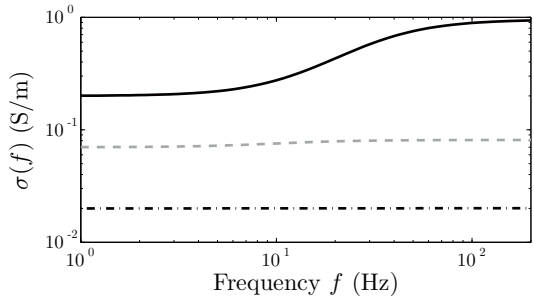
$$2\pi f \epsilon(f) \ll \sigma(f), \tag{2.14}$$

where f is the frequency of interest. The above condition indicates that it is permitted to treat the tissue as purely resistive (σ only).

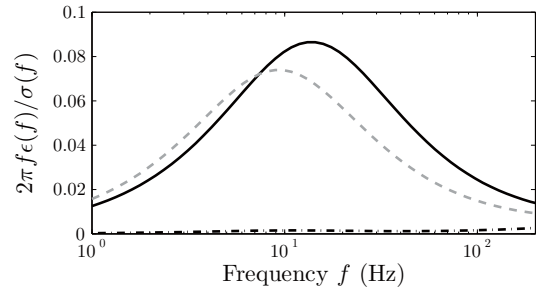
The ratio $\frac{2\pi f \epsilon(f)}{\sigma}$ for gray matter, cortical bone and cancellous bone is shown to roughly comply with this condition in Figure 2.12(c). This figure implies that the capacitive effects are small, and as a consequence there is *no temporal filtering performed by the properties of the head geometry at the frequencies of the EEG.* In other words simultaneous recordings beneath and above the skull do not experience noticeable delay or deformation, only at-



(a) Permittivity $\epsilon(f)/\epsilon_0$



(b) Conductivity $\sigma(f)$



(c) Ratio $2\pi f\epsilon(f)/\sigma(f)$

FIGURE 2.12: Theoretical frequency dependent electrical properties of biological tissues found in the head, including gray matter, cortical (compact) bone and cancellous (spongy) bone. (a) shows the permittivities $\epsilon(f)$ relative to the permittivity of free space ϵ_0, and (b) shows that conductivities $\sigma(f)$ are roughly constant for $f = 0 - 100$Hz. Both (a) and (b) are calculated using models in [48] and agree to experimental observed values, including those in Table 2.1. (c) shows the ratio described by Equation 2.14, indicating that capacitive effects are small and can effectively be ignored in the $0 - 100$Hz range.

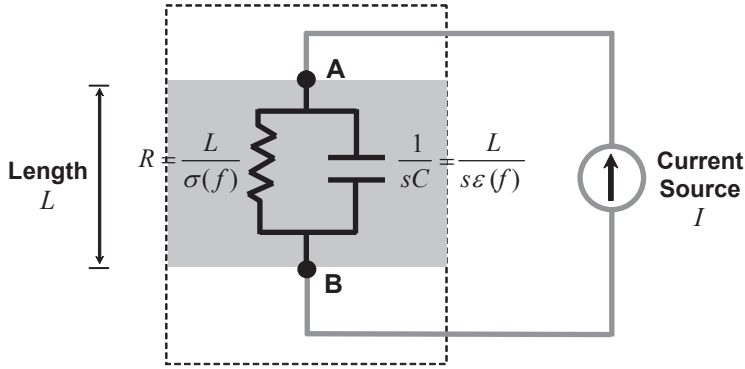


FIGURE 2.13: If a volume conductor is linear its electrical properties are uniform and both resistance and capacitance are dependent on the path length L between any two points. For points A and B shown, total resistance is $R = \frac{L}{\sigma(f)}$ and total capacitance is $C = \epsilon(f)L$. Here $s = 2\pi fj$ where f is frequency in Hz and $j = \sqrt{-1}$. This circuit is used to derive the approximate condition in Equation 2.14 for which the material's capacitance can be ignored. The derivation is shown in Appendix 2.C.

tenuation. These observations have been supported by experiments dating as far back as 1964 in [31].

Given this information one may wonder why it is that frequencies higher than about 100Hz are not present in the scalp EEG. If no filtering of the intra-cranial activity takes place by the volume conducting head, why are frequencies recorded near the cortex higher than the content of scalp EEG? The answer is that scalp EEG *is* filtered, but this is a consequence of the underlying synaptic dynamics combined with the spatial filtering properties of the head, discussed in Section 2.4.

2.3.3 Estimating Conductivities

Estimates of conductivities for regions of inhomogeneous media grouped together must always be space averages over large tissue volumes [122]. This is because smaller volumes may contain irregularities that cause estimates to be unrepresentative of the average properties. The conductivity can be estimated *in vitro* by analyzing dead tissue in an environment that simulates live conditions, or *in vivo*. The latter has recently been demonstrated as possible, non-invasively, by first determining the head geometry using MRI imaging and then using the principle of reciprocity[5] to estimate conductivities

[5]The principle of reciprocity states that if current is injected through two stimulating electrodes, the resulting current density or electric field through a volume conductor com-

[184, 40, 126]. Similar processes can be used anywhere on the human body, for example conductivities of muscles [99]. Estimates of conductivity of head tissue are also possible with simultaneous intra-cranial and scalp recordings [87].

Although conductivities of biological tissue depend on frequency, for the EEG this dependency has been shown to be small relative to the variations of tissue type [46, 47, 48]. Most estimates thus explicitly ignore this dependency, although it is imperative that experiments be performed in the appropriate frequency spectrum.

The conductive properties of each of these layers are summarized in Table 2.1 and are discussed in more detail below.

2.3.3.1 Brain

The brain is composed of two types of tissue: gray matter and white matter. Gray matter is where the generation of current sources occurs because it is where the pyramidal neurons are located. White matter is a collection of axon fibers that connect different regions of the brain together.

The isotropy of gray and white matter was studied by [167] with the use of MRI imaging that reflected the electric properties of different tissues in the brain relative to orientation. The authors found that whilst gray matter is largely isotropic in nature, white matter is highly anisotropic. This is because the conductivity parallel to the orientation of the axon fibers is expected to be higher than across the fibers [121]. However, because of the lack of consistency in fiber orientation no generalizations can be made about the anisotropic properties of the brain.

In any case the differences in conductivities of gray and white matter are relatively small, the latter having a slightly lower conductivity, and the difference can be ignored for simulation purposes – the principal effects are the result of the relatively highly resistive skull. Conductivity of the average brain ranges from $\sigma_{brain} = 0.12 - 0.48$S/m (see [40, 87, 126, 164]). A nominal value of $\sigma_{brain} = 0.3$S/m is selected for simulations presented here.

2.3.3.2 CSF

Cerebrospinal fluid (CSF) like any other body fluid has relatively large conductivity because it has high concentration of dissolved salts that facilitate the transfer of charge [122]. CSF has been shown to vary little between subjects and can thus be treated as a known parameter [40]. A nominal value of $\sigma_{CSF} = 5\sigma_{brain} = 1.5$S/m is chosen.

---

pletely specifies how the same electrodes used for recording potentials react to dipole sources in the volume conductor of the same magnitude, direction and location [103]. This principle can be used to estimate conductivities both *in vivo* and *in vitro*, but its usefulness for *in vivo* measurements is that it is non-invasive and effective enough to work with small currents that do not harm the subject [40].

2.3.3.3 Skull

The effects of the skull are the most important to volume conduction because it is, relative to all other tissues involved, the one with the lowest conductivity. The conductivity of the skull was originally computed by [155] in 1968 using a phantom model of a skull immersed in an electrolytic tank. An average conductivity ratio of $\sigma_{skull} = 1/80\sigma_{brain}$ was deduced and this was for many years accepted as the golden ratio of the 3-sphere model of the head that ignored the effects of CSF. More recently, this ratio has been disputed, with claims that in a 4-sphere model a more appropriate ratio is in the order of 1/20. This ratio varies from 1/15 ([126]) to 1/25 ([87]).

Another consideration is that the skull is itself composed of 3 layers – two compact layers at the inner and outer skull (cortical bone) and a spongy middle layer (cancellous bone) – yet most literature to date treats the skull as a homogeneous and isotropic medium. The reality is that (ignoring bone malformations and other anomalies) whereas each of the three layers may be roughly approximated as isotropic and homogeneous the average skull effect is neither. The conductivity of the cancellous bone is roughly 3-6 times higher than the compact bone; *shunting* or *re-direction* of currents in the tangential direction is expected. A better model of average skull conductivity is one in which the radial and tangential conductivities differ. Also worth mentioning is that the overall skull resistance does not necessarily increase linearly with skull thickness. This is because thicker skulls tend to have thicker spongy layers, largely absent in thin skulls [122].

Poor conductivity of the skull is expected to highly attenuate and spatially average sources beneath it. The work in this chapter uses an anisotropic model of the skull, with nominal conductivity ratio of 1/20 in the radial direction and 1/4 in the tangential direction.

2.3.3.4 Scalp

The scalp is composed of soft, fatty tissue with slightly higher conductivity than brain tissue, which is much the same. It is treated as isotropic, with nominal value of $\sigma_{scalp} = \sigma_{brain}$. Again, small differences between scalp and brain conductivities have been ignored because it is the skull (and not these small discrepancies) that impacts the effects of the volume conductor most.

2.4 The EEG: A Macro-Scopic View of the Brain

Both the type of sources and the volume conducting properties of the head have been discussed and we are now in a position to examine how EEG waveforms are generated. There is an important distinction that must be re-iterated when determining EEG waveform generation:

1. *Measurement*: How does a source configuration (in the form of a dipole layer) result in the recorded potentials? This is due to *volume conducting* properties, which has been the focus of most of this chapter. *Potential distributions due to volume conducting effects are a function of the properties of the source type and distribution, the electrical properties of the materials involved and the geometry of the volume conductor.* For the purpose of calculating potentials the current sources are assumed static and independent of the dynamics.

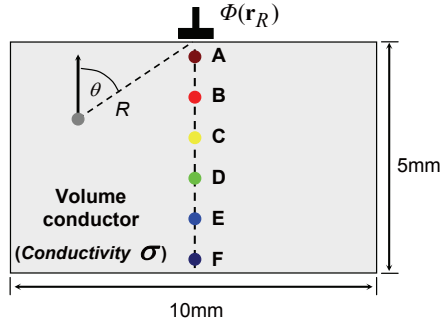
2. *Dynamics*: How is the source configuration generated? This is the study of the *dynamics* of the system, that is, how the underlying components of the systems (neurons, minicolumns, macrocolums) interact to generate a particular distribution of sources, which will vary over time (dynamics).

The volume conducting effects and the dynamics can be treated independently because the electrostatic equations presented in Section 2.1 operate at a much faster time scale than brain dynamics. Dynamic systems representative of brain activity are discussed in detail in Chapter 6. Here the features of the EEG that are a direct consequence of the dynamics and not the volume conducting properties are outlined, with particular interest given to the generation of the epileptic EEG.

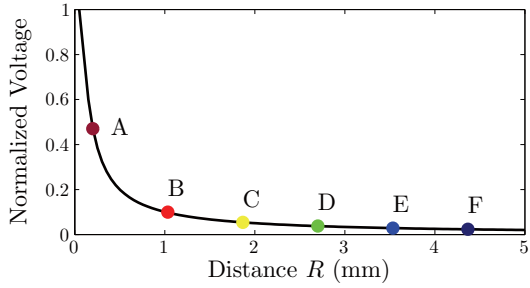
2.4.1 EEG Measurement

In this section the EEG forward problem, given sources in the brain, is simulated using the 4-layer concentric spheres model of the head (brain, CSF, skull, scalp) with conductivity and radii specified in Table 2.1 and Figure 2.10. Resulting potentials recorded at the cortex and at the scalp (corresponding to intra- and extra-cranial recordings respectively) are explored. Solutions for anatomically more realistic head models may be computed using finite or boundary element models (see, for example, [92]). However, more realistic models are best used to reduce errors in applications such as the identification of intra-cranial sources given scalp recordings (inverse problem), but do not change the qualitative observations targeted here.

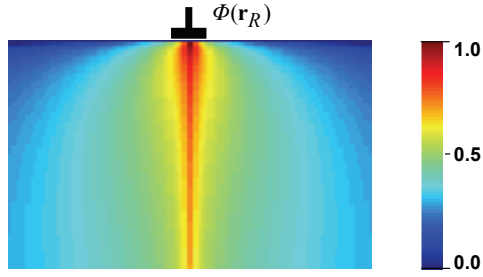
The simulations are presented in normalized color-gradient diagrams that indicate how a unit dipole at any location in the volume conductor affects the recording location $\mathbf{r}_R$. For example, in Figure 2.14(a) is a rectangular homogeneous medium with sources **A-F** located as shown. Assuming unit magnitude ($Id = 1$) and vertical orientation of dipoles then Equation 2.7 indicates that the contributions of **A-F** to $\mathbf{r}_R$ is dependent only on distance R and angle θ from source to $\mathbf{r}_R$. Thus dipole **A** contributes more than dipole **B**, followed by **C**, **D**, etc as shown in (b) and color coded in (a). The properties of the entire volume conductor can be color coded in this way by estimating the relative effects of unit dipoles at any location on the volume conductor, as shown in (c). From this diagram it can be concluded that the effects of a



(a) Dipole sources in a volume conductor

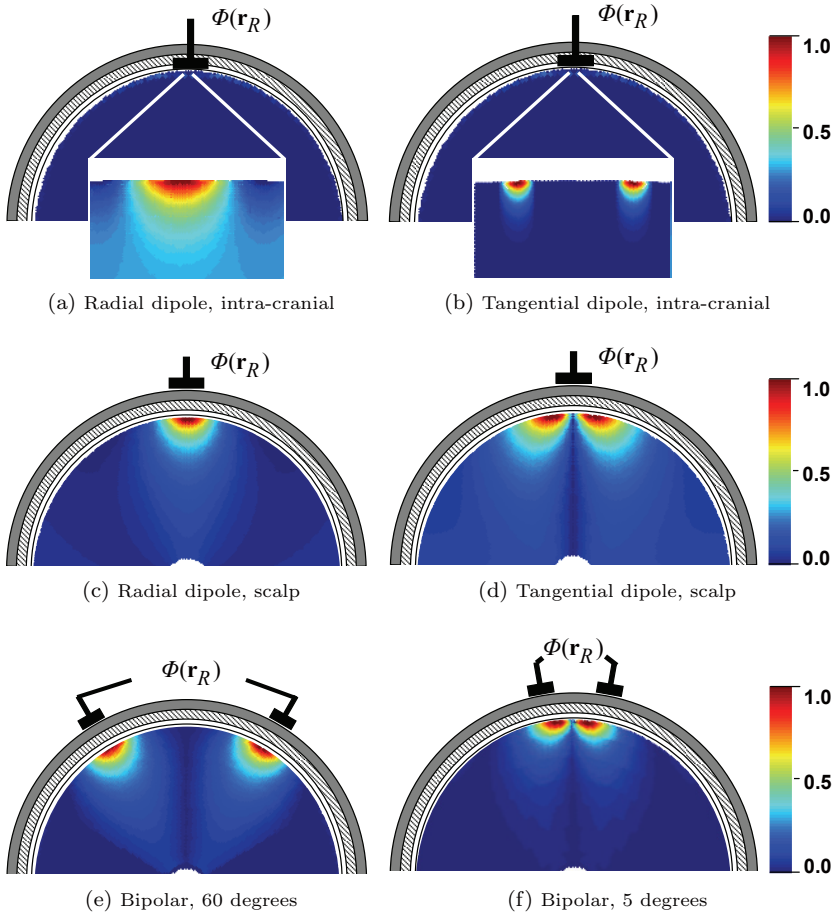


(b) Estimated dipole voltage as a function of distance



(c) Contribution of dipoles in a volume conductor

FIGURE 2.14: Example computation of the contribution that dipoles at different locations within the volume conductor have on recording location $\mathbf{r}_R$. (a) shows an example homogeneous volume conductor with sources **A-F** located as shown. In (b) the strength of **A-B** normalized relative to the strongest contributor (in this case located directly beneath the electrode) is shown. Dipole **A** contributes to $\Phi(\mathbf{r}_R)$ more than dipole **B**, followed by **C**, etc. The relative contributions can be color-coded, as in (c) which shows estimates of the relative contributions of a dipole in any location within the volume conductor. (A log-scale is used to differentiate between colors so that differences are more obvious.) Estimates of $\Phi(\mathbf{r})_R$ use Equation 2.7, with $Id = 1$ and (R,θ) defined as shown in (a).



(a) Radial dipole, intra-cranial

(b) Tangential dipole, intra-cranial

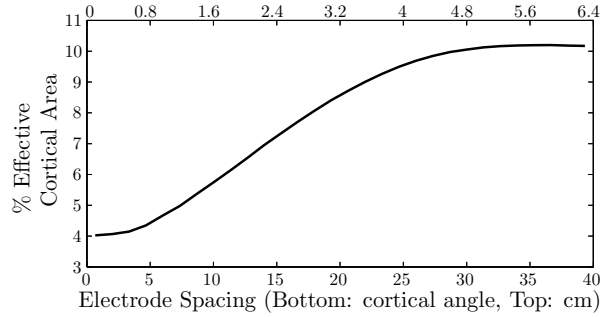
(c) Radial dipole, scalp

(d) Tangential dipole, scalp

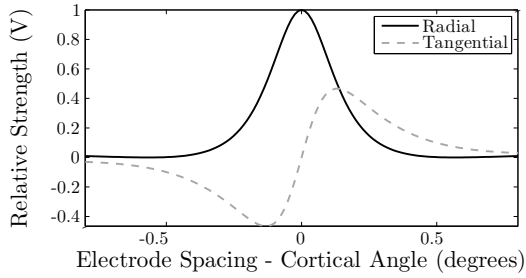
(e) Bipolar, 60 degrees

(f) Bipolar, 5 degrees

FIGURE 2.15: Relative contributions of radial and tangential dipoles to the EEG at the cortex (a-b), and at the scalp (c-d). The areas affecting EEG at the scalp are much greater than those at the cortex due to the averaging effects of the skull. Bipolar referencing effectively doubles this area if the electrodes are distant (as in (e)) but significant cancellation effects can improve the spatial resolution if electrodes are sufficiently close (as in (f)). This is highlighted in Figure 2.16(a). Note that only relative absolute strengths are given in (a-f).



(a) Cortical area versus angular separation



(b) Relative strength of directional dipoles

FIGURE 2.16: The contributing area is defined as the area that provides 50% of the response at the electrode. Figure (a) illustrates that electrodes must have a significant angular separation before their contributing areas show no overlap. (b) illustrates that radial dipoles are relatively more important than tangentially oriented dipoles.

dipole are strongest when $\theta = 0^o$, negligible when $\theta = 90^o$, and decrease with distance R.

For the simulations in the remainder of this section the volume conductor is the 4-layer spherical head and the computation is performed using the methodology described in [115]. Only radial and tangential dipoles are analyzed. Even though cortical folding allows dipoles in all directions any one dipole can be expressed as the sum of 3 components (one radial and 2 tangential) in a spherical head.

By definition voltages are never a result of a single electrode, but rather the *difference* between the recording and reference site. The location of this reference is a problematic issue in EEG recordings, and is discussed separately.

2.4.1.1 Cortical (Intra-Cranial) Recordings

Figure 2.15(a) and (b) show the effects of radial and tangential dipoles on an intra-cranial recording site $\mathbf{r}_R$. Because of symmetry 2-D diagrams are presented, although computations were performed in 3-D. Regions have been enlarged to the area of interest, as potentials recorded on the cortex further than a few mm decrease to zero rapidly.

From these diagrams it is estimated that the surface area[6] recorded by a point electrode on the cortex is in the order of 4-5mm^2. Intra-cranial recordings are very localized, and potentials beyond 2mm contribute less than 10% to the total signal. Electrodes placed a few cm apart from each other are capable of producing very different waveforms (unless both regions are acting similarly due to the dynamics). This has been observed experimentally in, for example, [123] where coherence between electrodes further than 2cm showed almost zero correlation to one another.

Tangential sources have minimal contribution to the intra-cranial EEG because recording strength falls very fast with respect to distance. In any case, electrodes on the cortex can be assumed to be affected only by sources that are roughly normal to the surface because of the small areas that are involved in recordings. Tangential sources can thus be ignored.

The potentials calculated here are for ideal point sized electrodes that do not exist in practice. In reality the electrode size has a significant effect on the area that affects cortical recordings, given that an integration over the entire volume of the electrode must take place. Electrode sizes in intra-cranial recordings are rarely greater than 1mm in diameter, with newer electrodes in the order of 0.01-0.1mm being fabricated. Example electrode dimensions can be found in Figure 5.1(b).

2.4.1.2 Scalp Recordings

Figure 2.15(c) and (d) show the effects of radial and tangential dipoles on a scalp recording site $\mathbf{r}_R$. The scalp EEG is affected by much larger areas of cortical activity. Let us define the *effective cortical area* as that in which dipoles affect the recording site by at least 50% of the contribution of the radial dipole located directly beneath $\mathbf{r}_R$ [7]. In this case a measurement is affected by all dipoles in an area of $\sim 20\text{cm}^2$ containing billions of neurons!

An important observation is that the commonly accepted notion that brain activity directly beneath the scalp electrode is the largest contributor to the recording is not necessarily true.

[6] Surface area is a more appropriate measure than volume due to the previously proposed model of a dipole sheet to represent cortical activity. Although radial dipoles of a particular depth are shown in Figure 2.15(c) and (d), there are no generators in the white matter directly beneath the cortex, and deeper sources are not radially aligned, so their contribution is minimal.

[7] This is known to engineers as the -6dB point, at which energy is halved.

The same recording can arise from one dipole directly beneath the electrode, one stronger dipole located some distance R from the electrode or many weaker ones at distant locations. The increased sensitivity of scalp electrodes to radial dipoles directly beneath them does not mean that large potentials arise from sources that are located there.

This averaging over larger surfaces takes place because of the relatively low conductive skull which tends to spread currents tangentially to the surface. Skull conductivities (as well as the inclusion of anisotropy) affect the calculated figures but do not qualitatively affect results. A skull with lower conductivity tends to spread potentials more, and the area in question increases.

Radial dipoles affect the scalp EEG more than tangential dipoles, with its largest contribution arising directly beneath the electrode. A tangential dipole of the same magnitude contributes only $1/3-1/2$ of an equivalent radial dipole (as pictured in Figure 2.15(h)) [122], with its maximum occurring at an angle dependent on the properties of the volume conductor. As a consequence of this and the fact that tangential dipoles are generally located deeper in the cortex, the effects of the tangential dipoles are relatively small and largely ignored in models of head volume conductors. The level of attenuation by the skull to each of these dipoles is not shown in these diagrams; only relative values are given.

Spatial averaging also implies that, for reasonable values, electrode size on scalp recorded EEG does not significantly affect these observations.

Scalp EEG measures average over cortical areas in the order of 20cm^2, whereas intra-cranial records are responsive to areas in the order of 4mm^2. Radial dipoles are the major contributors to both – tangential sources are largely ignored because of their more distant location as well as their smaller contributions.

2.4.1.3 The Search for an Ideal Reference

Discussions so far have ignored the contributions of a referencing electrode even though all recorded potentials depict *differences* between two points. That is

$$V(\mathbf{r}_{recorded}) = \Phi(\mathbf{r}_A) - \Phi(\mathbf{r}_B), \tag{2.15}$$

where $\mathbf{r}_B$ is called the *reference* electrode site and $\mathbf{r}_A$ is the site of interest. In an ideal situation $\Phi(\mathbf{r}_B)$ has no impact on $V(\mathbf{r}_{recorded})$ so that $V(\mathbf{r}_{recorded}) \approx \Phi(\mathbf{r}_A)$. This can be achieved provided $\mathbf{r}_B$ is in an electrically neutral location ($\Phi(\mathbf{r}_B) = 0$) that is at a large electrical distance from $\mathbf{r}_A$, in which case referencing issues become unimportant. Electrically distant in this case refers to a location sufficiently far away so that volumes contributing to $\Phi(\mathbf{r}_A)$ do not affect $\Phi(\mathbf{r}_B)$.

In reality $\mathbf{r}_B$ also needs to be somewhere where current can flow between $\mathbf{r}_A$ and $\mathbf{r}_B$ because of conditions imposed by the recording equipment. This means that when recording the EEG $\mathbf{r}_B$ cannot be, say, on a distant wall – it must lie on the subject's body. *This makes it impossible to achieve electroneutrality in the reference* because the distances involved for the scalp are not large enough. The same is not true for intra-cranial recordings because the volumes that affect each electrode are much smaller. It is well accepted that no quiet reference exists for scalp EEG, so it is important to understand the merits of different configurations, some of which are listed below.

- *Linked ears/mastoids*: Performed by physically or mathematically linking the subject's ears is historically the most well known yet fundamentally flawed method of referencing. Referencing to the linked mastoids introduces a third site whose activity can influence the EEG. Recordings of this nature are rarely un-biased.

- *Average reference*: Removes the average of all recorded electrodes from recording site. For sufficiently fine electrode sampling this is appropriate because the large number of electrodes can accurately estimate the field distribution of the entire head, and thus remove it from the recording site of interest. However for a relatively low spatial sampling (such as the typical standard 10-20 configuration system in Figure 1.8(a)) the subtraction of the average introduces a bias due to imperfect estimation of the field. Average referencing should be used for high-resolution EEG only[8].

- *Bipolar referencing*: Refers to the potential difference between two sites on the head[9] generally used for electrodes close to one another, in specific configurations so as to emphasize voltage differences between hemispheres of brain activity (see Figure 1.8(b) and (c)) [118]. Figure 2.15(e) and (f) demonstrate that bipolar recordings can both increase and decrease the cortical surface area affecting recordings significantly.

[8]Scales at this point might be getting confusing: *If the effective cortical area of an electrode is* $\sim 20cm^2$*, is this not adequate sampling?* No - 20cm^2 relates to roughly a 4-5cm circular diameter. The circumference of the head is in the order of 30cm, which means the electrodes in the 10-20 configuration are at best about 6-7cm apart. This is not adequate sampling for a correct estimate of average activity.

[9]Note that *all* recordings are bipolar in that they measure the difference in potentials between two sites. In EEG bipolar referencing refers to the electroencephalographer's explicit acknowledgment of this fact.

For electrodes far away as in (e) the cortical area affecting the EEG doubles, but as electrodes come closer as in (f) the effective area of measurement decreases dramatically because of cancellation effects between recording sites, as described in (g). Bipolar recordings closer than ∼2cm on the scalp do not improve the resolution further [31], also shown in (g). Bipolar recordings still involve large areas of cortical activity, but cancellation effects reduce contributions from the first spherical harmonics, effectively localizing the resulting potential. For the scalp EEG with few channels *bipolar recordings are the best way to improve spatial resolution.*

In practice all EEG electrodes are referenced to a single electrode at the time of acquisition, typically located on top of the head or behind the neck. Re-referencing is performed digitally post-acquisition. The effects of measurement noise must be considered when re-referencing takes place.

A measurement can only be recorded to the accuracy of the measurement equipment. Some level of *noise* is expected, say

$$V(\mathbf{r}_A) = \tilde{V}(\mathbf{r}_A) + \eta_A, \tag{2.16}$$

where $\tilde{V}(\mathbf{r}_A)$ is the real signal and η_A is the noise introduced by the recording equipment. This noise is expected to have a particular distribution determined by the specifications of the equipment as well as the measurement environment.

Recorded signals can be re-referenced, in which case

$$V(\mathbf{r}_{AB}) = (\tilde{V}(\mathbf{r}_A) - \tilde{V}(\mathbf{r}_B)) + (\eta_A - \eta_B). \tag{2.17}$$

If η_A and η_B have similar distributions but there is no overlap between them the noise of $V(\mathbf{r}_{AB})$ can as much as double. In practice some overlap does exist because both signals $V(\mathbf{r}_A)$ and $V(\mathbf{r}_B)$ are recorded referenced to the same electrode. A typical EEG measurement system limits its noise to $\leq 2\mu$V, thus a maximum of 4μV for bipolar re-referencing when signals are in the order of 1mV is acceptable.

For a more detailed discussion on the effects of referencing in EEG, refer to [122].

2.4.1.4 Spatial Filtering Properties of the Skull

It is evident from Figure 2.15 that spatial averaging of sources occurs to a large degree on scalp recorded EEG. A way to explain this is by examining how the *spherical harmonics* are affected by the volume conductor. It was mentioned in Section 2.3.1 that the resulting potential at the scalp due to volume conducting effects can be expressed as a sum of these spherical harmonics or *basis functions*, as in Equation 2.13. In the spatial domain, this is analogous to a temporal signal being represented as the sum of sinusoids. Examples of these basis functions are described in Figure 2.11.

When the distribution of source activity is broken down into spherical

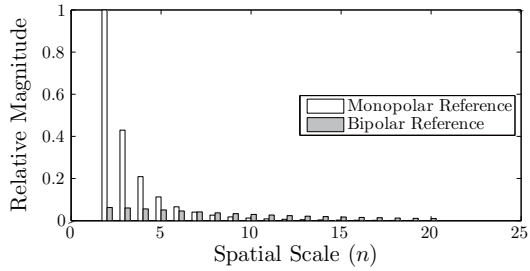


FIGURE 2.17: Spatial filter transfer function for scalp EEG for a bipolar reference, with electrodes far apart (in white), and a bipolar reference with electrodes close together (in gray). The latter improves spatial resolution by providing more attenuation to larger spatial scales and less to smaller ones.

harmonics, those events that are synchronous over large areas can in general be described with fewer terms in Equation 2.13. If many smaller scale events exist then more terms are necessary because finer spatial resolution is required to describe these events. More specifically, the harmonic number n is related to *spatial frequency* as

$$k \approx \frac{n}{2\pi r_{scalp}}, \tag{2.18}$$

where n is the harmonic number, k is the spatial frequency and r_{scalp} is the radius of the sphere. This can be verified in Figure 2.11(c) where higher n produce basis functions that describe the spatial domain at higher resolutions – spatial distances are inversely related to spatial frequencies. In a sphere with radius $r_{scalp} = 9.2$cm, as is the head, $n = 1, 2, 3, 4$ represents spatial scales of approximately $58, 29, 19, 14$cm respectively.

A spatial transfer function can be derived, that, analogous to how a transfer function in the frequency domain describes the attenuation of each sinusoid, in the spatial domain corresponds to how events occurring at the spatial scales of each spherical harmonic are attenuated. Note that a spatial transfer function does *not* tell us how much of the signal is a result of a particular spatial scale. Instead it says that *if there exists* an event that covers an area of a particular spatial scale then it is attenuated according to the relationship shown by the transfer function.

The spatial transfer functions for bipolar referencing related to Figure 2.15(e) and (f) are shown in Figure 2.17. These were calculated as described in [172], and can be interpreted as follows:

- Bipolar references with distant electrodes (denoted monopolar in this figure) emphasize global scales much more than local scales – that is, events that span a larger area of the cortex contribute more to the EEG

than events that are local in nature. Events for spherical harmonic number $n > 12$, that is, 4cm or less in spatial scales contribute minimally to the scalp EEG. This is supported experimentally in [31].

- Bringing bipolar references close together removes energy from global events. Furthermore, higher spherical harmonic numbers (approximately $n > 8$, or spatial scales of less than 7cm) contribute more to the EEG than with a distant referencing electrode. This means that bringing pairs of electrodes closer together *improves* the spatial resolution.

The above observations are qualitatively captured in Figure 2.15(e) and (f), where the recording with the electrodes in (e) is affected by a greater brain area than the spatially closely aligned electrodes in (f). Spherical harmonic numbers of approximately $n < 3$ are unlikely to be relevant in the EEG because they correspond to events occurring in more than a 20cm radius of cortex. Events of this magnitude are not frequently observed in the brain, except in special cases such as epileptic seizures where the entire brain is recruited. In these instances, both methods of bipolar recording have very large contributions from global events, and show large waveforms, although the magnitude of these waveforms are much larger for bipolar referencing with distant electrodes than nearby ones. This phenomenon is observed in real EEG waveforms.

The CSF, skull and scalp act as a spatial filter between cortex and scalp. Synchronous activity in the brain that spans large cortical areas is attenuated less than activity that affects smaller areas. Scalp recordings are influenced more by global events. Bipolar reference can improve spatial sensitivity.

An important byproduct of the different transfer functions in Figure 2.17 is the amount of spatial sampling required to accurately represent the potential distribution in the head. Engineers will be familiar with the Nyquist sampling theorem for the temporal domain, which states that to accurately represent the frequency content in the signal the sampling rate must be at least twice that of the largest frequency present. If this is not done then the recording signal is *aliased* – a type of distortion that cannot be undone by any signal processing technique. The same is true in the spatial domain.

To accurately represent the potential distribution on the head the spatial sampling frequency must be at least twice that of the largest spatial frequency. Enough electrodes need to be placed on the head to give an accurate description of it.

In the temporal domain it is easy to ensure that the sampling frequency is appropriate by filtering the relevant frequencies out of the signal prior to sampling. This cannot be done in the spatial domain. However the highly resistive skull ensures that small scales do not influence the EEG significantly and the number of electrodes required to accurately represent the potential distribution on the human head is still reasonable. This is not so at the cortex, where no such pre-filtering is supplied by nature. Dense arrays with electrodes every cm (if not closer) are required to accurately represent the potential distribution at the cortical level.

For sufficiently sampled scalp it is possible to reduce the effects of the spatial averaging and the effects of the reference site that takes place by using the inverse spline laplacian. This is a method by which the effects of the skull are removed by *spatial de-convolution*, reducing the effects of large spatial scales and improving the resolution of small ones. The resulting *dura image* is then an estimate of potentials at a spatial resolution in between that of physical cortical recordings and scalp EEG, but recorded without the need for surgery. However, studies have shown that the standard 10-20 electrode placement system is spatially under-sampled [121], which does not mean that the signals are unrepresentative of the activity in the brain but does mean that spatial de-convolution is not possible due to spatial aliasing. High resolution EEG is necessary, where 64-128 electrodes are appropriate. For details on this type of work including limitations not mentioned here refer to, for example, [172].

2.4.2 EEG Dynamics

This chapter has explicitly separated the concepts of EEG measurement (approximated by a dipole sheet model in a volume conductor) and EEG source generation. The latter is a consequence of the *dynamics* of the brain dictated by the network configuration of interacting neurons, synaptic action and reactions and external stimulus provided by sensory information. An alternative approach to that presented thus far is to first develop a dynamic system model, calculate the potentials generated by the model and then calculate the resulting fields using the volume conduction theory provided herein. Although this is the ultimate goal – a complete understanding of brain dynamics and potential field generation together – to date even the most developed models are

incapable of representing the dynamics of the brain as a whole, an issue that is discussed extensively in Chapter 6.

Because dynamics can be separated from volume conducting effects the task of explaining the EEG is greatly simplified. Volume conducting effects are well understood whilst brain dynamics is in its infancy.

Whereas temporal dependence of typical network topology and behavior have been omitted in Section 2.2 they are the essence of the brain dynamics. This is not to say that temporal aspects are not important in EEG measurement, for example synchronous activation of many neurons is required to generate fields of consequence in the EEG. However it is the *dynamic properties* of the system that allow temporal synchronization to eventuate.

Phase shifts are another phenomenon that can be explained through dynamics. Neither volume conducting model properties or experimental observation have demonstrated any phase shifts (or delays) in measurements taken simultaneously intra-cranially and at the scalp [31]. However, phase shifts at different scalp locations exist [122]. Again, this is a consequence of the dynamics of the system rather than the volume conducting properties. The phase shifts occur due to time required for waves to propagate from the dynamics in one region of the brain to another.

High frequencies ($\sim$ 1kHz) are not observed in the EEG of the cortex because the EEG captures the dendritic response, which is much slower than the time variations associated with action potentials. Moreover, the EEG only records a meaningful response when a large number of synchronous events come together – the faster the phenomena, the less likely their positive reinforcement

Perhaps the most misunderstood concept is that of temporal filtering occurring between cortex and scalp EEG. There is nothing in the volume conductor model that does not allow higher frequencies to be observed at the scalp, provided enough of the cortex is active at that frequency (see Section 2.3.2), yet the frequencies observed at the scalp are lower than those observed intra-cranially. Where does this temporal filtering occur? Again it is a consequence of the interaction between dynamics and volume conductor. The spatial filter between the cortex and scalp selectively transmits activity that is synchronous and active in a sufficiently large region of cortex. However, high frequency activity does not synchronize over large areas as often because it is more difficult for short events to phase-lock and constructively interfere. Hence a consequence of the spatial filter is to temporally filter the EEG between cortex and scalp. The dynamics play a role in that they also suppress

high frequency behavior over large cortical areas due to propagation delays in axon fibers.

EEG dynamic modeling is a very active area of research and will continue to be so for a long time to come. The most developed models to date (described in Chapter 6) are capable of describing activity in the spatial scale of the macrocolumn. Larger scale dynamics have been addressed but models remain tentative. As such, for the co-registration or validation brain dynamic models using EEG waveforms it does not make sense to use recordings measured at the scalp since it is intrinsically global in nature. Instead, research of this nature should concentrate on the use of intra-cranial recordings that take measurements in the millimeter rather than the centimeter scale.

2.4.3 Epilepsy and the EEG

Since our focus is on epilepsy it is instructive to determine the configurations of sources that could create the epileptic EEG waveforms. The trends observed during a seizure are synchronous activity over many channels, a sudden or progressive onset of this activity and oscillations that may be large or small (but well pronounced). Some of these observables are shown in Figure 1.10. The first two can only be described by the system dynamics – no mechanism of the volume conductor can affect the synchronicity over large regions of cortex, nor can the perceived sudden onset and offset of epileptic activity. The resonances that result both temporally and spatially are a function of the neuronal network topology. These issues are addressed in more detail in Chapter 6.

However, the volume conducting effects play a role in the large amplitude oscillations. What is the underlying activity of the neurons, given that epilepsy is accepted as a global phenomenon that affects large spatial scales? Large voltages at the scalp can occur for several reasons:

1. Events covering large spatial scales contribute larger voltages to scalp EEG, as discussed in Section 2.4.1.4. Epilepsy is a global event.

2. Synchronous activity contributes high voltages to the scalp EEG, as discussed in Section 2.2.3. The oscillatory nature of the EEG suggests that many neurons fire in synchrony in an on-off manner at a frequency characteristic of the seizure.

3. Increased random (asynchronous) activity of EEG may also increase the average recorded potential.

So what are the neurons doing during an epileptic seizure? The answer is probably a combination of the above. fMRI studies have shown that the metabolic activity increases throughout the seizure [134], although the relatively low temporal resolution of fMRI equipment makes it impossible to discern if this happens in an on-off nature.

The regularity of epileptic EEG waveforms as well as the fact that large amplitudes are recorded in both intra-cranial and scalp EEG alike suggests that increased activity cannot be the sole factor. The range of amplitudes that can occur in an epileptic seizure is more easily explained by the level of *synchrony* in the neural population. Smaller amplitude seizures simply have a smaller proportion of neurons firing in synchrony. This is important in a neural model; it is not necessary to propose that many neurons are firing, only that a higher proportion is synchronously active.

It is well known that for epilepsies with gradual onset the cortical EEG registers epileptic activity near the focus well before any can be seen on the scalp EEG. *A significant area of cortex must be recruited to the seizure before scalp recordings show a change.* This may mean that scalp recordings are unsuitable for analysis of onset phenomena and seizure spread, particularly if generalization of the seizure is fast.

2.5 Conclusions

The aim of this chapter is to discuss what the EEG is capable of saying about the underlying brain activity. This has been addressed by looking at what the EEG measures, that is, the forward problem.

> *Radially oriented, parallel* current dipoles determine the measured EEG, both at scalp and intra-cranial locations. Sources must be *synchronously active* for EEG magnitudes to be large.

EEG predominantly observes cortical sources because the cortex is the only structure in the brain capable of generating similarly oriented, synchronous activity in a consistent manner. Dynamic properties related to the interaction between cortical and sub-cortical structures need to employ different recording strategies – *no information about the sub-cortex is given in the scalp or cortical EEG.* Furthermore, the cortical regions that affect the EEG the most are those on the gyrii because these are where parallel sources are most typically radially oriented in the head.

The EEG is responsive to synchrony rather than activity. Low magnitudes do not necessarily imply little activity. Techniques responsive to activity rather than synchrony include fMRI [121].

The EEG recorded at the scalp is a spatial average of a large area of cortical activity. A consequence of this and previous observations is that equal waveforms can arise from

1. A few dipoles (current sources) near the electrode, or
2. Many dipoles at distal locations, or
3. A few synchronously active dipoles, or
4. Many dipoles that are asynchronous, or
5. A few dipoles of large magnitude, or
6. Many dipoles of small magnitude,

or any combination of the above.

Hence recordings at the scalp give a very poor representation of the spatial resolution of the underlying brain activity, since so many different configurations can potentially generate the same waveforms. Points (1) and (2) are not as important for intra-cranial EEG because greater spatial resolution is provided. It is the intra-cranial EEG that will likely prove most useful for modeling applications. However, the invasiveness of the recording procedure means that healthy human data are unlikely to ever be available and care must be taken that models are not representative of a condition rather than normal brain activity.

Theoretically it is always possible for a single dipole alone to generate the potentials recorded at the EEG, regardless of its orientation or synchronicity. However, the magnitude of this dipole needs to be large enough to mask the activity of all other sources that are synchronized and are similarly oriented. Research suggests that potentials of the magnitudes required to mask all other activity are much higher than those found in the brain [31]. It is much more likely that moderately low but synchronous activity over large areas are the major contributors to EEG.

The scalp EEG provides an estimate of the spatially filtered activity in the cortex. Any perception of temporal filtering between cortex and scalp is a consequence of this spatial filtering and dynamics of the brain, and not due to the volume conductor.

Dynamic models based on scalp EEG should be careful to differentiate between global and local phenomena. The dynamic variables used should be appropriate for the spatial scales involved. This is not as important for the description of global phenomena such as epilepsy or the alpha rhythm, but models designed at the meso-scopic level that claim to describe the scalp EEG should be interpreted with care.

Spatial resolution can be improved with high spatial sampling, as discussed in Section 2.4.1.4, but data of this nature are frequently unavailable.

The spatial resolution of EEG signals acquired with low spatial sampling can be improved by using bipolar recordings between electrodes close to one another.

EEG traces should *always* be interpreted as bipolar; sources that are in both locations (be they synchronous, parallel, large spatial scales, etc) affect the EEG. Digital EEG has enabled re-referencing of EEG signals post-acquisition. The most appropriate referencing system depends on the application.

We have seen that the EEG signal is limited in what it can tell us because it is a crude image of what is happening in the brain. This supports our original rough estimate, in the Preface, that 1 bit of information is gained every second for every 10^5 neurons in the brain. The spatial average removes much of the detail, particularly in the scalp EEG where we can at most only get information about every 20cm^2 of cortex at a time. Coupled with our understanding of what EEG signals measure we can safely say that the EEG cannot reveal much detail about the micro- or even meso-scopic activity in the brain. However, from a purely phenomenological point of view the EEG is very useful because patterns *do* exist in recorded traces. *Signal processing* methods that allow us to extract statistical information about the EEG signal are discussed next.

2.A Units of Electric Quantities

Below is a list of many of the quantities related to bioelectromagnetism that have been used in this chapter, along with some of the units typically used to measure them. All units listed for each quantity are equivalent, and can be used to draw relationships between these quantities.

Quantity	Symbol	Units
Electric Field	$\mathbf{E}$	Newtons per Coulomb (N/C)
		Volts per meter (V/m)
Electric Potential	ϕ, Φ	Volts (V)
		Joules per Coulomb (J/C)
		Newton-meter per Coulomb (Nm/C)
Electric Current	I	Amperes (A)
		Coulombs per second (C/s)
Resistance	R	Ohm (Ω)
		Volts per ampere (V/A)
Conductivity	σ	Siemens per meter (S/m)
		Inverse electric resistance per meter ($1/\Omega$m)
Resistivity	$\nu = \frac{1}{\sigma}$	Volt-meter per ampere (Vm/A)
Capacitance	C	Farads (F)
		Seconds per ohm (S/Ω)
		Coulombs per Volt (C/V)
Permittivity	ϵ	Farads per meter (F/m)

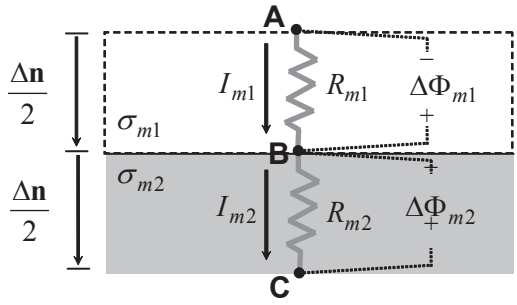
2.B Volume Conductor Boundary Conditions

This is a derivation of the boundary conditions in Equations 2.10 and 2.11.

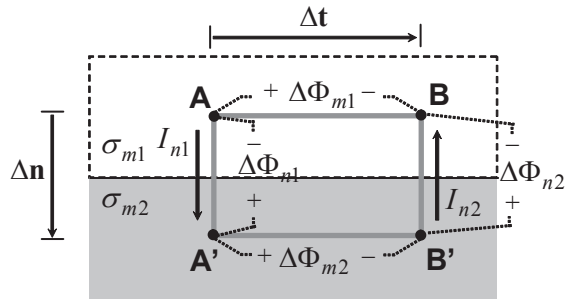
First let us look at Equation 2.10. An inhomogeneous volume conductor shown in Figure 2.18(a) is composed of two homogeneous materials with conductivity σ_{m1} and σ_{m2} respectively. Because each material is linear, that is, the electrical properties are uniform within, the effective resistance of each is a function of the length of the material, in this case $R_x = \frac{\Delta \mathbf{n}}{2}\frac{1}{\sigma_x}$. By Ohm's law,

$$I_{m1} = \frac{\Delta\Phi_{m1}}{R_{m1}} = \sigma_{m1}\frac{2\Delta\Phi_{m1}}{\Delta\mathbf{n}}, \quad \text{and} \tag{2.19}$$

$$I_{m2} = \frac{\Delta\Phi_{m2}}{R_{m2}} = \sigma_{m2}\frac{2\Delta\Phi_{m2}}{\Delta\mathbf{n}}, \tag{2.20}$$



(a)



(b)

FIGURE 2.18: Diagrams to help derive Equations 2.10 and 2.11. The boundary between two homogeneous volume conductors with conductivity σ_{m1} and σ_{m2} is shown. In (a), the effective resistances through the media, by Ohm's law, are $R_{m1} = \frac{\Delta \mathbf{n}}{2}\frac{1}{\sigma_{m1}}$ and $R_{m2} = \frac{\Delta \mathbf{n}}{2}\frac{1}{\sigma_{m2}}$. $\Delta\Phi_{m1} = (\Phi_B - \Phi_A)$ and $\Delta\Phi_{m2} = (\Phi_C - \Phi_B)$. Assume currents originate from a closed circuit, not shown. In (b), a loop around the boundary is drawn. By Kirchoff's law the potentials around this loop sum to zero. Here $\Delta\Phi_{n1} = (\Phi_{A'} - \Phi_A)$, $\Delta\Phi_{m1} = (\Phi_A - \Phi_B)$, $\Delta\Phi_{n2} = (\Phi_{B'} - \Phi_B)$, and $\Delta\Phi_{m2} = (\Phi_{A'} - \Phi_{B'})$. Refer to Appendix 2.B for derivations.

where $\Delta\Phi_{m1} = (\Phi_B - \Phi_A)$ and $\Delta\Phi_{m2} = (\Phi_C - \Phi_B)$.

As a consequence of conservation of charge (charge cannot appear or disappear) the currents flowing across the boundary must be equal. Thus,

$$I_{m1} = I_{m2} \tag{2.21}$$

$$\sigma_{m1}\frac{\Delta\Phi_{m1}}{\Delta \mathbf{n}} = \sigma_{m2}\frac{\Delta\Phi_{m2}}{\Delta \mathbf{n}}. \tag{2.22}$$

By taking the limit as $\Delta \mathbf{n} \to 0$

$$\sigma_{m1}\frac{\partial\Phi_{m1}}{\partial\mathbf{n}} = \sigma_{m2}\frac{\partial\Phi_{m2}}{\partial\mathbf{n}}, \tag{2.23}$$

as in Equation 2.10.

For Equation 2.11 refer to Figure 2.18(b) where a path (loop) is drawn across the boundary. This loop has length $\Delta\mathbf{n}$ and width $\Delta\mathbf{t}$. By the principle of conservation of energy Kirchoff's voltage law states that the voltage sum around any loop in an electrical circuit is equal to zero. Thus,

$$\Delta\Phi_{n1} + \Delta\Phi_{m1} - \Delta\Phi_{n2} - \Delta\Phi_{m2} = 0, \tag{2.24}$$

where $\Delta\Phi_{n1} = (\Phi_{A'} - \Phi_A)$, $\Delta\Phi_{m1} = (\Phi_A - \Phi_B)$, $\Delta\Phi_{n2} = (\Phi_{B'} - \Phi_B)$ and $\Delta\Phi_{m2} = (\Phi_{A'} - \Phi_{B'})$, as shown.

To observe the change of potential with respect to $\Delta\mathbf{t}$, the tangent, this can be re-written as

$$\frac{\Delta\Phi_{n1}}{\Delta\mathbf{t}} + \frac{\Delta\Phi_{m1}}{\Delta\mathbf{t}} - \frac{\Delta\Phi_{n2}}{\Delta\mathbf{t}} - \frac{\Delta\Phi_{m2}}{\Delta\mathbf{t}} = 0. \tag{2.25}$$

Next take the limit $\Delta\mathbf{t} \to 0$. By symmetry and conservation of charge $I_{n1} = -I_{n2}$, resistances of both paths normal to the boundary are the same and thus $\Delta\Phi_{n1} = \Delta\Phi_{n2}$. Terms 1 and 3 in the above equation cancel. As a consequence,

$$\lim_{\Delta\mathbf{t}\to 0}\frac{\Delta\Phi_{m1}}{\Delta\mathbf{t}} = \lim_{\Delta\mathbf{t}\to 0}\frac{\Delta\Phi_{m2}}{\Delta\mathbf{t}} \tag{2.26}$$

$$\frac{\partial\Phi_{m1}}{\partial\mathbf{t}} = \frac{\partial\Phi_{m2}}{\partial\mathbf{t}}, \tag{2.27}$$

as in Equation 2.11.

2.C Capacitance in RC Circuits

This is a worked example of the conditions under which capacitive effects in a circuit such as that shown in Figure 2.13 can be ignored. This relates to the electrical properties of biological tissue discussed in Section 2.3.2. Basic knowledge of circuit analysis is assumed in this example.

In Figure 2.13 it is shown that the frequency dependent capacitance is given by $C(f) = \frac{\epsilon(f)}{L}$ and frequency dependent resistance by $R(f) = \frac{L}{\sigma(f)}$, where L is the linear distance from a point A to point B in the tissue and f is the frequency of interest. Ignoring frequency dependence of R and C for the moment, the equivalent impedance Z of the circuit in Figure 2.13 is given by

$$Z = R \,||\, \frac{1}{sC} = \frac{R}{1 + sRC}, \tag{2.28}$$

where $s = 2\pi f j$ with $j = \sqrt{-1}$ and $||$ denotes that R and C are in parallel. The voltage between points A and B is, by Ohm's law,

$$V_{BA} = ZI = \frac{RI}{1 + sRC}, \tag{2.29}$$

where I is the current passing through the medium, shown as a current source in Figure 2.13. This voltage can be re-written as

$$V_{BA} = RI\left(\frac{1 + sRC - sRC}{1 + sRC}\right) \tag{2.30}$$

$$= RI\left(1 - \frac{sRC}{1 + sRC}\right). \tag{2.31}$$

This says that voltage V_{BA} is approximately equal to a purely resistive circuit ($V_{AB} \approx RI$) if the magnitude of this second term $\|\frac{sRC}{1+sRC}\|$ is small. This quantity is thus representative of the percentage error in V_{AB} when the assumption that the medium is purely resistive is made.

To make this error small it is sufficient to make $\|sRC\| = \|2\pi f RCj\|$ and consequently $2\pi f RC$ small. Substituting $R = \frac{L}{\sigma(f)}$ and $C = \frac{\epsilon(f)}{L}$ results in the condition that $2\pi f\epsilon(f) \ll \sigma(f)$ be true if capacitive effects are to be ignored.

3

Signal Processing in EEG Analysis

> "It is easy to lie with statistics. It is hard to tell the truth without it."
>
> *- Andrejs Dunkels, Swedish mathematics teacher (1939-1998)*

The EEG is a collection of recorded signals that represent the electrical activity in the brain. From Chapter 2 we have an idea about how these signals are generated and how they are recorded. We have seen examples of what they look like, and we have seen that they can be used to tell us something about what is happening in the brain system. This chapter presents some of the mathematical tools available to interpret these records, that is, it explores the *signal processing* commonly used to analyze EEG records.

Signal processing is a way in which the EEG record can be converted into a numerical description of *features* in the data. For example, numbers that describe the energy, frequency content or complexity of a signal can be computed. These features are used to *unmask* information that is not visually obvious, or to *compress* vast amounts of data to a more manageable level of selected information. Compression is important for EEG analysis because many hours of recording lead to gigabytes of data.

Often the extracted features are better viewed as *statistics*. Formally, a statistic is defined as "a fact or piece of data obtained from the study of a large quantity of numerical data" [124]. Statistical analysis is easily (and therefore frequently) unintentionally misused, misinterpreted and misreported. It is important as researchers to strive for correct implementation and interpretation – *all statistical analysis techniques are limited, and these limitations must be understood and reported*. It is often just as important to know what a statistic cannot say as well as what it can say. This depends on the context of the signal, in this case the fact that the signal is an EEG.

The aims of this chapter are (1) to present common signal processing tools, and (2) to outline their limitations given that we are using EEG signals to interpret the brain system. Although the methods presented are applicable to a broad spectrum of problems, signal processing techniques relevant to epileptic seizure detection and prediction are the main focus. A broader, less specific survey of EEG signal analysis can be found in, for example, [118] or [181].

Both detection and prediction of epileptic seizures are examples of *classification* problems. The former classifies between two states, seizure versus

non-seizure, whereas the latter includes a third pre-seizure state indicative of an imminent epileptic event. Neither seizure detection nor seizure prediction are new problems. Simple seizure detection methods have been around for decades, although their sophistication increased significantly in the age of the digital EEG. The original analog amplitude detectors have evolved into complicated systems involving many levels of computation. Research in seizure detection remains relevant because of the search for 'optimal' performance that can emulate that of a human expert. The first attempt at seizure prediction using EEG can be traced back to the 1970s [70], although poor performance meant that efforts were quickly abandoned. In the 1980s spatial and temporal patterns of interseizure events were analyzed. The study of seizure prediction really flourished in the 1990s, in particular through the use of non-linear systems theory.

A complete detection and prediction system can be divided into the four stages outlined in Figure 4.1, namely *preprocessing*, *feature extraction*, *classification* and *expert system*. This chapter focuses on signal processing of EEG signals, which is most relevant to the first two. However the selection of the type of signal processing should be aligned with the remaining stages of classification and expert systems. These are discussed separately in Chapter 4.

Preprocessing is the process in which the EEG is prepared for analysis. The signal processing in this area involves the removal of unwanted aspects, such as artifact and high frequency content, and normalizing the EEG data so that it is comparable to all other data (e.g., normalize the amplitude range, sampling frequency, etc). Preprocessing is discussed in Section 3.2.

Feature extraction is the process whereby the relevant statistics or *features* are extracted from the EEG. In the case of seizure detection and prediction, relevant information refers to features that are capable of distinguishing between non-seizure, pre-seizure and seizure states. This is perhaps the most important component of the classifier given that without appropriate extraction of features the classifier cannot perform well. Feature extraction is discussed in Section 3.3.

For seizure detection and prediction it is essential that the extracted features characterize the *temporal evolution* of the EEG so that changes can be detected at the appropriate time scale[1]. This brings us to a fundamental problem in the design of a detection system: we want to capture change, but change is bad because most signal processing techniques require *stationarity*, that is, they require that the statistical distribution of the signal (such as that shown in Figure 3.2) remains constant over time.

Any EEG is an observation from an intrinsically non-stationary system. Simply looking at the very different morphologies of the signals during 'sleep' and 'awake' shown in Figure 1.9 highlights this point. The question is then how easy is it to detect these changes? How fast can we conclude that a

[1] An appropriate time scale may be seconds for detection and minutes for prediction.

transition has taken place? If an observation window of a certain duration is required for the correct computation of a statistic, but the EEG changes at a rate that is much faster than that, does the statistic become useless? How must results be interpreted? Short enough windows are necessary so that the EEG can be assumed to remain roughly stationary and temporal resolution remains high[2], yet they must also be large enough so that the signal processing techniques are useful. These trade-offs are discussed throughout.

To avoid re-writing signal processing books a certain level of mathematical knowledge is assumed of the reader, although an effort is made to maintain this at a minimum and the explanations given are as self contained as possible. For someone not so familiar with the relevant mathematics but interested in learning more, the standard textbooks [58], [105] and [86] may be useful. Otherwise reading the introductions to each section, summaries and boxed comments should be sufficient to obtain a conceptual understanding.

Before delving into descriptions at the different stages of a classification system, the EEG is introduced as a mathematical construct where the conventions, notations and assumptions used throughout the text are presented.

3.1 Mathematical Representation of the EEG

An EEG is a set of recordings taken from C_{TOT} electrodes that can be placed on the scalp, the surface of the cortex or deeper within the brain. In this text a continuous time signal $x_c(t)$ is used to denote the temporal evolution of voltages contributing to each location, where t represents time and $c = 1..C_{TOT}$ is the electrode/channel number. These are the *true* signals generated by the brain system.

The digital EEG approximates the continuous time signal $x_c(t)$ by sampling it at discrete time sampling points, spread at an interval Δ. That is

$$x_c[n] = x_c(n\Delta), \quad n = 1, 2, 3 \cdots . \tag{3.1}$$

Here $x_c[n]$ is the *recorded* discrete time signal that samples $x_c(t)$ at time intervals $t = \Delta, 2\Delta, \cdots, n\Delta$. The discrete time is enumerated as $n = 1, 2, 3 \cdots$. Once digitized we only have access to $x_c[n]$, the assumption being that the sampling interval is small enough (alternatively, a large enough sampling frequency is used) so that $x_c[n]$ closely represents the relevant properties of $x_c(t)$. Aspects of this are discussed in Section 3.2.

In Chapter 2 we saw that $x_c(t)$ and consequently $x_c[n]$ are affected by activity in networks of hundreds of thousands of neurons. The behavior of these networks over time, known as the *dynamics* of the brain, can be described by a set of mathematical equations expressed in terms of 'hidden' variables known

[2]Here we are talking about time scales much larger than those presented in Chapter 2.

as the (dynamic) state, here denoted as $\mathbf{z}[n]$. The *dimension* of the state is assumed finite[3], say $\hat{N}$. That is, we assume $\mathbf{z}[n] \in \mathbb{R}^{\hat{N}}$, where $\mathbb{R}^{\hat{N}}$ is the *phase space* of the system. The time series $x_c[n]$ is a *single dimensional* recording that is not able to describe $\mathbf{z}[n]$ in its entirety without further manipulation. Methods that allow the *reconstruction* of at least some of the properties of $\mathbf{z}[n]$ given the single time series $x_c[n]$ exist. This is discussed in more detail in Section 3.3.4.

First let us look at how the relationship between $x_c[n]$ and $\mathbf{z}[n]$ may be conceived. A discrete-time system can be described by [159, 183]

$$\mathcal{B} = \left\{ x_c[n], \quad n = 1, 2, 3.. \;\middle|\; \begin{array}{l} \exists(u[n], \kappa[n]) : \\ \\ \mathbf{z}[n+1] = P(\mathbf{z}[n], \kappa[n], u[n], n) \\ \\ x_c[n] = B_c(\mathbf{z}[n]) \end{array} \right. . \tag{3.2}$$

The above may define the *behavior* $\mathcal{B}$ of a single EEG channel $x_c[n]$ at times $n = 1, 2, 3..$. This description informs us that $x_c[n]$ is generated by a dynamical system with input $u[n]$, parameters $\kappa[n]$ and state $\mathbf{z}[n]$. The relationship between these quantities is determined by maps P and B_c. P is the state transition map and B_c is the output map for channel index c. The input is an unconstrained signal that in the brain could represent sensory information (e.g., vision, hearing), could be set to zero or could be a 'white noise' signal when describing endogenous activity.

The *map* $P : \mathbb{R}^{\hat{N}} \to \mathbb{R}^{\hat{N}}$ describes how to get from the current state at time n to the next state $\mathbf{z}[n+1]$ at time $n+1$. A time series $\mathbf{z}[n]$ for $n = 1, 2, 3..$ is generated by repeated applications of the map P beginning with some starting condition $\mathbf{z}[0]$. This time series is known as an *orbit* or *trajectory* of the system [159].

If the map P can be written as a constant matrix, then Equation 3.2 is a *linear system*, otherwise it is *non-linear*. If Equation 3.2 is used to describe a completely stochastic system then $\mathbf{z}[n]$ defines the *distribution* of the state of the system, and P is used to describe how this distribution changes over time.

P depends on a set of *parameters* ($\kappa[n]$) representative of physiology (e.g., network topology, chemical concentrations), as well as *inputs* to the system $u[n]$ (e.g., visual and auditory stimuli) that together determine the behavior of the system. The parameters, inputs and the map itself can all change over time and are functions of time n. *Stochastic* or *random* processes can be incorporated into Equation 3.2 via $u[n]$ or as a property of P itself. This, along with the nature and assumptions made about P, $\kappa[n]$ and $u[n]$, is discussed in

[3]More realistically, given that a lot of the phenomena in the brain are related to transport of chemicals, heat, energy or electrical signals, the state should be more appropriately represented as infinite dimensional.

further detail in Chapter 6. For this chapter it is sufficient to speak about the map P as an arbitrary (but appropriate) description of the dynamics, from which measurements can be taken. Here, P can be viewed as the dynamical system called the brain.

The *measured output* or *observable* taken from the system $\mathbf{z}[n]$, in our case the EEG signal $x_c[n]$, is also defined in Equation 3.2. The signal is derived from $\mathbf{z}[n]$ by the map $B_c : \mathbb{R}^{\hat{N}} \to \mathbb{R}$. Here c is the channel index that corresponds to a particular recording location on the head. This map largely depends on the geometry and the electrical properties of the materials in the head, and as described in Chapter 2 is, unlike P, assumed stationary and independent of inputs $u[n]$.

In this chapter a single recorded EEG channel is represented as a time signal $x_c[n]$. n is the time index representing $x_c(n\Delta)$, which samples a real signal $x_c(t)$ using a sampling interval Δ. $c = 1, 2..C_{TOT}$ is an index representing one of the C_{TOT} EEG channel locations.

$x_c[n]$ is a *single* discrete time measurement taken from the behavior of the brain, generated by a complex dynamical system with state $\mathbf{z}[n]$ that is $\hat{N}$ dimensional. Mathematically the behavior $\mathcal{B}$ of the resultant $x_c[n]$ can be expressed in terms of the dynamics of $\mathbf{z}[n]$ as in Equation 3.2.

The mathematical conventions introduced here are used throughout the book. However in this chapter it is only necessary to acknowledge that a brain system described by $\mathbf{z}[n]$ exists, that an EEG signal $x_c[n]$ is an observation taken from this high dimensional system and that signal processing is used as an attempt to extract properties of $\mathbf{z}[n]$ from $x_c[n]$. More than one channel ($C_{TOT} > 1$) may be analyzed at a time.

3.2 Preprocessing

Preprocessing in a classification system involves the preparation of the raw EEG data, $x_c[n]$, so that it is ready for feature extraction. This requires (1) making sure that the data are appropriately sampled, referenced and filtered, (2) *normalizing* the data so that it complies to a particular standard and (3) dealing with unwanted artifact.

Appropriate sampling is dealt with by the recording system (such as that

described in Chapter 5) at the time of acquisition. This involves abiding by the *Nyquist criterion* which states that a sampling rate of F_s Hz can only correctly represent signals that contain frequencies up to $F_s/2$. If frequencies higher than $F_s/2$ are present then distortion known as *aliasing* occurs. *The effects of aliasing cannot be removed by any form of signal processing.* This is demonstrated in Section 3.3.2.

If the true signal $x_c(t)$ contains frequencies higher than $F_s/2$ prior to sampling these components must be removed. This is known as *pre-filtering* of the signal. In practice a pre-filter removes frequencies smaller than $F_s/2.5$ to avoid edge effects [122]. *Re-sampling* of the raw signal to a new (smaller) F_s also requires pre-filtering.

Normalization here refers to the process by which data are converted to a form that can be compared to other data that have been acquired using different recording equipment or taken from different people. For example, consider two sets of EEG with similar phenomena but acquired using two separate recording systems. System 1 encodes the recorded signal in the range $[0, 10]$ and system 2 in the range $[-20, 20]$). The amplitudes of these two measurements are not directly comparable. However, if prior to comparison data are normalized to a common range (e.g., $[0, 1]$) then both signals may be compared directly.

A standard way to normalize a signal is to

1. *Remove its mean*: The EEG is a measure of *relative* voltage, and as such it is a quantity that can only be defined up to a constant (see Chapter 2). The mean of a single recording does not provide useful information and can interfere with analysis. Therefore it is best to *detrend* the signal by removing its mean prior to analysis. If there are multiple simultaneous recordings then all channels should be detrended by the same constant. This preserves their *relative* means, which may contain information. It is implicit when bipolar referencing is applied.

2. *Scale to unit variance*: To normalize a detrended signal to its unit variance it must be divided by its standard deviation, making the resulting signal *scale invariant.* Indeed the size of the signals in an EEG record is meaningless, as it depends on the gain of the measurement amplifier. This means that records measured on different equipment are comparable to one another. Both variance and standard deviation are defined in Section 3.3.1.

Normalization is an important process that applies both the raw data (in the preprocessing stage) and processed data (in the feature extraction stage). It is discussed in more detail in Section 3.3.

Once the EEG signal is acquired and normalized, the referencing scheme used can affect the localization of EEG activity, as discussed in Chapter 2. For the purpose of seizure detection and prediction bipolar referencing is most

often used to improve spatial resolution, that is, so as to emphasize the differences in activity between electrodes. Differences in activity at different sites are important for the identification of seizure activity. Re-referencing can be applied post-acquisition.

Finally, artifact must be dealt with. Several strategies exist to deal with the types of artifact outlined in Chapter 1. These include:

1. *Ignoring*: Assume that the feature extraction methods are only minimally affected by the artifact.

2. *Rejecting*: Artifact is (automatically) identified and channels or epochs that are contaminated are excluded from analysis.

3. *Removing*: Artifact is again identified and if possible removed from the EEG signal by separation methods such as independent component analysis (ICA) [69] or wavelet filtering (described later). Less data is lost than through rejection.

4. *Training*: The system is trained to identify and cope with common artifact. This shifts the responsibility over to the Classification and Expert system described in Chapter 4.

Seizure detection and prediction algorithms often employ all of the above. Most algorithms deal with simple artifact such as electromagnetic interference using strategy 3 – the frequencies at which interference occurs are removed from the signal. Weak artifacts such as that caused by heart activity are ignored because their influence is assumed small.

Ocular artifacts (EOG) are in most cases ignored because they are assumed to have a minimal effect on signal processing methods. Although this assumption is not necessarily true, it is convenient because rejection results in the loss of large amounts of data (ocular artifacts occur very often) and removal (e.g., by using ICA, as discussed in [78] and [79]) is computationally cumbersome. At the most a classifier may be trained to expect interference from EOG.

Muscle artifacts (EMG) are also very common, particularly in scalp EEG because there are muscles close to electrodes. Rejection is thus not an option. Removal is a better solution but is a difficult task because EMG frequencies overlap with normal and seizure EEG frequencies, particularly but not exclusively in the 15-20Hz range [55, 128]. The EMG separation process uses wavelet filters (described in Section 3.3.3) rather than conventional filters because these avoid distortion effects [152, 151]. Epochs of heavy muscle artifact may be rejected altogether.

The more sophisticated detection algorithms employ strategy 4 and selectively train the system to cope with artifact (see [170] and [156] as examples). This training typically includes examples of common artifact as well as alpha rhythms that although not artifactual in nature are often responsible for incorrect classification of the EEG.

3.3 Feature Extraction

Feature extraction is key to the performance of a classifier. Because the field of signal processing is enormous the theory presented here is restricted to what is most relevant to EEG analysis and epilepsy in particular. Methods valid only for long term studies are omitted – epileptic events are short and it is their temporal evolution that must be represented.

The problem under investigation in this text is: *what signal processing is able to extract a feature ζ_s from $x_c[n]$ that can be used to differentiate between seizure, non-seizure and, in the case of prediction, pre-seizure EEG?*

The choice of an appropriate *discriminating* feature requires a deep understanding of the problem [49].

Multiple features can be used together, in which case the classifier has access to a *feature set* $(\zeta_1, \zeta_2, \cdots, \zeta_S)$. More features are not necessarily better – each ζ_s must introduce additional information to improve the discriminating power of the feature set. Spurious addition of features can confuse results, degrade performance and prohibit efficient computation as shown in [185].

Before delving into the specifics of feature computation several general concerns particular to the EEG are addressed. These are discussed below.

3.3.0.1 Computing Statistics: Averages vs. Instances

Statistics are computed from large quantities of data because any one *instance* of the raw data can vary significantly. For example, no two EEG sequences will ever look the same. Even the EEG of the same person, doing the same thing, but recorded at a different time will be different. However there is an expectation that if the person is doing the same thing then the EEGs will indeed look similar. We postulate that our observations are drawn from a large ensemble of possible observations, and that this ensemble has some well defined features allowing us to infer the probability of a certain observation. More often than not we also postulate that observing a signal over a long

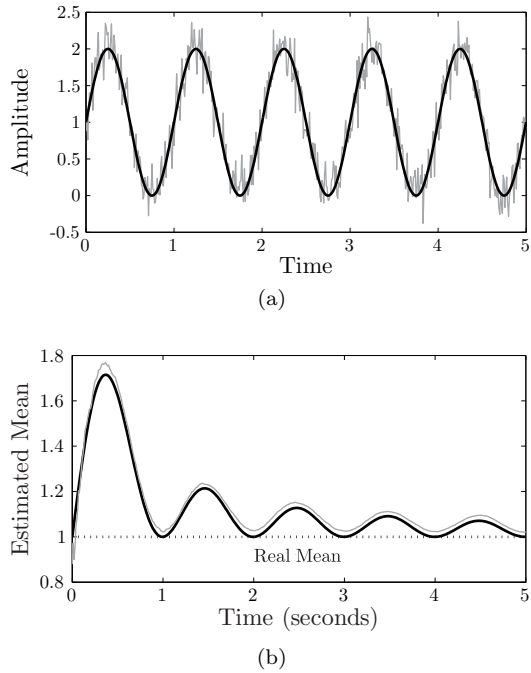
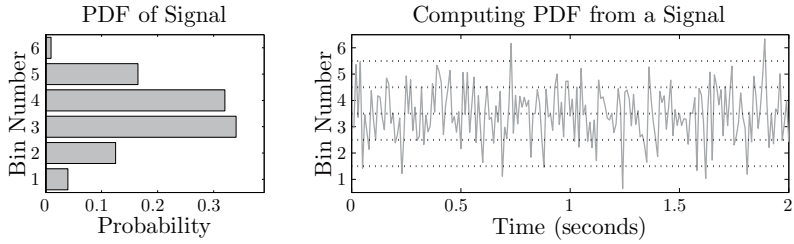


FIGURE 3.1: To observe the sinusoid in (a) it is necessary to look at *average* behavior over a suitable time frame, particularly so when noise is present. If only a few samples are available the sinusoidal pattern is not evident. When computing statistics the length of the analysis window is also important, as shown in (b), where the statistic of interest is the mean. Using only a few samples the estimated mean is nowhere near the real value. Only when the analysis window is longer does the estimate approach the real mean. In this figure the sampling frequency is $F_s = 512$Hz, so that 1 second is equivalent to $N = 512$ samples.

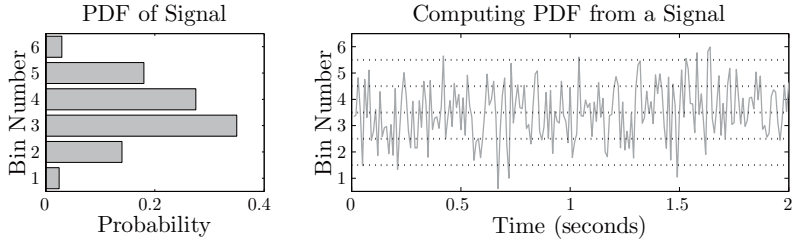
period of time will allow us to infer the properties of the ensemble. This is called *ergodicity*.

We can see this with two examples. First refer to Figure 3.1(a) where a stand-alone sinusoid is plotted. Any one *sample* or *instance* in time isolated from all other samples does not tell us much – it could belong to any signal. However if observations are made in the context of a larger timeframe the sinusoid becomes evident. This is true also when noise is added, as shown. *On average* this noisy signal is a sinusoid.

Next examine Figure 3.2 where two signals, different to one another, are plotted. Even though their raw values are not the same their *empirical distribution* of amplitude is very similar. Empirical distribution is computed by



(a) Signal 1



(b) Signal 2

FIGURE 3.2: Two signals in (a) and (b) are different because their raw values are not the same. However when we look at the *probability distribution function* (PDF) of each signal we see that their *distributions* are roughly similar. The two signals are said to be *samples* or *instances* of data sequences that belong to the same statistical distribution, on average. The PDF is computed by summing the number of samples that fall within each marked bin, and dividing by the total number of samples in the signal. Thus it is a measure of the relative frequency of occurrence. The process is explained later in Section 3.3.4, Equation 3.42. In this figure the sampling frequency $F_s = 512$Hz, so that 1 second is equivalent to $N = 512$ samples.

counting the relative number of samples that lie within one of the amplitude bins shown. This is known as the *probability distribution function* or PDF. The two signals are said to be *samples* or *instances* of data sequences that are drawn from an ensemble with a particular statistical distribution.

To say anything meaningful about a signal or system we study averages. These averages can be taken over time, over space or over a distribution.

However we are rarely in a position to compute averages precisely as we

do not know the 'true' distribution of the signal, or worse, there may not be a 'true' distribution. It is then very important to consider how we actually compute a statistic. How many data points were included? Did we make any assumptions about measurement errors and the environment? Is the instance from which the feature is extracted representative of the overall behavior? *The interpretation of a statistic must consider the specifics of the calculation.* Refer to the example shown in Figure 3.1(b), where the statistic of interest is the mean amplitude of a signal. Here the estimated mean is shown when different observation times or *window* lengths are used. The longer the observation time, the closer the estimated statistic reflects the true value. However if we do not know the true mean of the signal then we cannot know that the feature was extracted using sufficiently long observation time. What is the significance of a computed mean if it is not known that the computation is correct?

3.3.0.2 Noise

Any measurement can be expressed as the sum of the 'true signal' plus some *noise*. In the case of the EEG signal this is

$$x_c[n] = x_c(n\Delta) + \eta(n\Delta), \quad n = 1, 2, 3 \cdots \tag{3.3}$$

where $x_c(n\Delta)$ is the value of the true signal $x_c(t)$ at time $t = n\Delta$ (where Δ is the sampling interval), and $\eta(n\Delta)$ is the error of this measurement at this time. Here $\eta(n\Delta)$ is known as the *measurement noise* because it is the quantity by which the recorded signal $x_c[n]$ differs from the true signal $x_c(n\Delta)$. These differences arise from errors in the acquisition process, which cannot be removed, or from artifact – the part of the measured EEG that is not caused by the neural activity in the brain. If $\eta(n\Delta)$ is small then $x_c[n] \approx x_c(n\Delta)$.

In practice we have the freedom to determine what we call signal $x_c(n\Delta)$ and what we call noise $\eta(n\Delta)$. For example, consider once more a noisy sinusoid similar to that shown in Figure 3.1(a). If for the purposes of analysis we are only interested in frequencies below 3Hz we can assign anything higher than this frequency to noise. The signal $x_c[n]$ can be decomposed as in Figure 3.3(a), and $\eta(n\Delta)$ can be discarded prior to analysis. Similarly, imagine a situation where absolute amplitudes larger than 0.75 are of no interest. In this case $x_c[n]$ can be decomposed as in Figure 3.3(b). *Noise and signal are in the eye of the beholder and they depend on the application.*

Another type of error is that caused by manipulation of $x_c[n]$. When a statistic is estimated using sub-optimal conditions, and this estimate deviates from the true value, this is known as *computational noise*. Computational noise arises from inappropriate representation of data at the signal processing stage, for example, when insufficient data are used, when modeling errors are introduced or when the data are non-stationary and/or non-linear. Computational noise places a limitation on how a statistic may be interpreted.

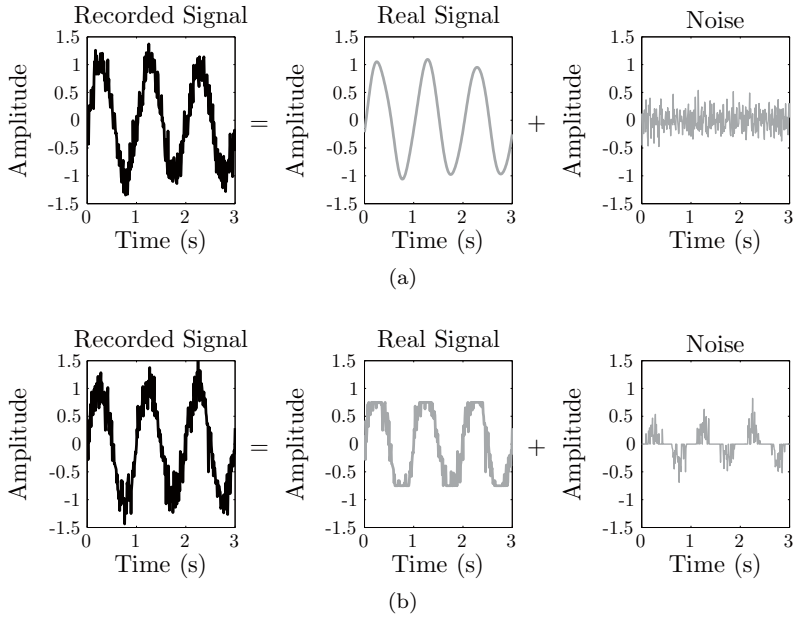
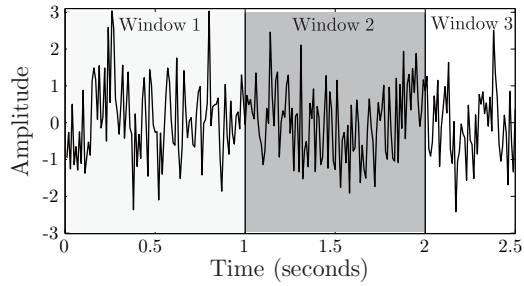


FIGURE 3.3: Noise is a quantity that can, within limits, be arbitrarily defined. Shown above is an example where the recorded signal $x_c[n]$ is decomposed into real signal $x_c(n)$ and some noise $\eta(n\Delta)$. In (a) $\eta(n\Delta)$ consists of all frequencies above 3Hz, whilst in (b) it is when absolute amplitudes are above 0.75. Once the type of noise is defined it can be removed from a signal prior to analysis. Whilst the case in (a) is more realistic in practice there is no theoretical reason why noise such as that in (b) cannot be used. In this figure the sampling frequency $F_s = 512$Hz, so that 1 second is equivalent to $N = 512$ samples.

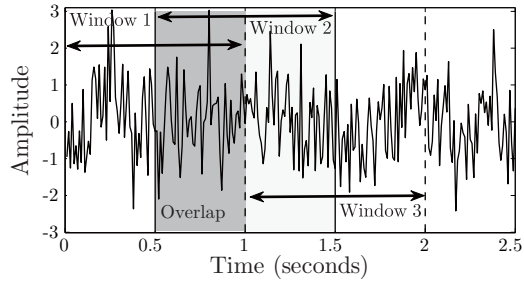
Measurement and *computational* noise both result in errors in an estimated statistic.

3.3.0.3 Stationarity and Windowing

Most mathematical tools assume *stationarity* in a signal, that is, they assume that the information in $x_c[n]$ (both statistical and dynamical) remains the same under arbitrary time shifts. However the EEG is indisputably inherently not stationary: differences in EEG time series between sleep, wakefulness, eyes open, eyes closed, etc. are evident. The very attempt to differentiate



(a) Non-overlapping window



(b) Overlapping window

FIGURE 3.4: An example of using non-overlapping one-second windows is seen in (a). In (b) the windows are shown to overlap by 0.5 seconds, thus the first window has $k = 1$ and the second $k = 1 + N/2$. In this figure the sampling frequency $F_s = 512$Hz, so that 1 second is equivalent to $N = 512$ samples.

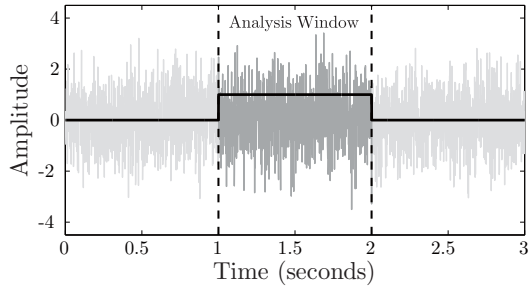
between non-seizure, pre-seizure and seizure states at different times implies non-stationarity [37].

To reflect changes over time a signal $x_c[n]$ is *windowed* by dividing the time axis into sections that may or may not overlap. Examples of overlapping and non-overlapping windows are in Figure 3.4.

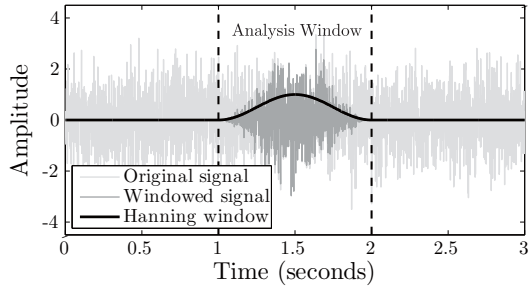
The most often used window is the *rectangular* window, defined as

$$H_k[n] = \begin{cases} 1, & \text{if } k+1 < n < k+N \\ 0, & \text{otherwise} \end{cases}, \tag{3.4}$$

for a window of length N, starting at time sample $n = k + 1$. Windowing is applied by multiplying the signal with this function, that is, $x_c[n]H_k[n]$. Between times $k+1 < n < k+N$ the signal $x_c[n]$ remains unchanged. Outside this range the windowed signal is zero. This is shown in Figure 3.5(a), where the dotted line is the original signal and the solid line the resultant windowed signal.



(a) Rectangular window



(b) Hanning window

FIGURE 3.5: The effects of applying a window over an analysis window. In (a) a rectangular window defined in Equation 3.4 is applied to the original signal by multiplying the two together. In (b) the edge effects are less pronounced using a *Hanning window* defined in Equation 3.5. In both cases all values outside the analysis window are zero. In this figure the sampling frequency $F_s = 512$Hz, so that 1 second is equivalent to $N = 512$ samples.

The problem of rectangular windows is that the sharp edges of the rectangle can affect analysis. Different types of windows exist to deal with this problem, a common one being the *Hanning* window defined as

$$H_k[n] = \begin{cases} 0.5\left(1 - cos\left(\frac{2\pi(n-k+1)}{N-1}\right)\right), & \text{if } \; k+1 < n < k+N \\ 0, & \text{otherwise} \end{cases} . \quad (3.5)$$

An example of this is shown in Figure 3.5(b).

Windowing effectively enforces an (artificial) stationarity by making the local signal within it globally valid in analysis. Of course this means that no information about signals outside the window is used. A rectangular window does not guarantee continuity at the edges, but other windows (like the Hanning) do. This goes a long way toward explaining the reason for the

use of Hanning windows, in particular if Fourier analysis is used, where the assumption is that the data within the window is being repeated outside of it.

We want to make windows short because stationarity in $x_c[n]$ cannot be assumed when long windows are used. However N must also be long enough so that the computed statistic is informative (which for some methods, particularly those presented in Section 3.3.4, may be very long). In practice rather than restricting analysis to tools only valid for non-stationary signals (a minuscule number compared to those designed for stationary signals) the issue is pragmatically resolved by assuming that windows of up to 20-30 seconds are *weakly stationary* or *almost*-stationary [37]. This is an approximation, and even on this time scale non-stationary behavior can be observed; for example short abrupt bursts are very common in the EEG.

The assumption of weak stationarity over windows of 20-30 seconds in length of EEG data underpins most signal processing tools used to extract features from EEG sequences. The computational noise introduced by non-stationarities in these windows is assumed to be negligible.

To analyze a signal over time it is sufficient to vary the time k at which the window is applied. Successive values of k can be chosen so that consecutive windows of length N overlap by $N-k$ samples. In some cases it is also possible to vary the length N of each successive window.

3.3.0.4 Linearity, Non-Linearity, Determinism and Stochasticity

The signal $x_c[n]$ can be generated by a system that is linear, non-linear, or stochastic. The difference between a linear and a non-linear system is that the former preserves the operations of addition and scalar multiplication. The behavior $\mathcal{B}$ is linear if for any two trajectories $x_{c1}[n]$, $x_{c2}[n]$ that belong to $\mathcal{B}$, $(x_{c1}[n], x_{c2}[n] \in \mathcal{B})$, then for all real constants a and b

$$ax_{c1}[n] + bx_{c2}[n] \in \mathcal{B}. \tag{3.6}$$

That is, it is linear if the scaled sum of any two trajectories also belongs to the behavior $\mathcal{B}$. If Equation 3.6 does not hold then the system is non-linear. It is safe to assume that the EEG signal is an observation from a non-linear system.

To analyze the signal generated by a complex system we have linear as well as non-linear tools. Linear tools can be applied to signals generated by non-linear systems, effectively modeling the non-linearities as a stochastic component. However linear tools are limited in what they can tell us and this approximation may miss information resulting from the non-linearities.

Non-linear signal processing tools are more effective in dealing with this type of data, at the expense of greater computational cost, and typically requiring stationarity over longer windows (more than 20-30 seconds).

Stochastic elements are often assumed to arise from noise and artifact introduced by measurement, although unknown processes within the brain are also described as stochastic in models of the generating systems (see Chapter 6). Stochasticity is dealt with by estimating its effect on a particular signal processing tool.

By far many more linear signal processing tools exist than non-linear ones. Both typically assume stationarity of the signal, although some methods exist that do not. Linear methods are described in Section 3.3.1, Section 3.3.2 and Section 3.3.3. Non-linear analysis tools are described in Section 3.3.4.

Linear signal processing tools are prevalent, and far better understood than non-linear ones. Non-linear tools typically require more data, in which case the stationarity, quantity and quality of the data must be taken into account.

3.3.0.5 Normalization

Normalization applies both to the raw data (usually in the preprocessing stage, described in Section 3.2) and to the computed features. A features ζ_s presented to a classifier *must* be normalized so that it can be compared to the same feature extracted from any other data of the same nature. Normalization depends on the type and length of the window used.

Often it is possible to incorporate the normalization in the process of estimating the statistic. In cases where this is not so (because no appropriate normalization exists) a normalized comparison can be achieved by using *relative* rather than absolute measures, for example, comparing a feature relative to background activity. *Normalization is imperative to the analysis of a system such as the EEG in which signals vary so significantly between states of awareness and between different people.* The features discussed in this section are based on normalized data sequences unless otherwise stated.

Relative and normalized statistics are important in a recording environment where large variation in signals are observed at different times or between different records, and where the length of observation time or window length needed is not always obvious.

To summarize, an EEG signal $x_c[n]$ is a *noisy measurement* of brain activity.

Features extracted from data using linear signal processing techniques may not be accurate if $x_c[n]$ has a *non-stationary* or *non-linear* origin.

For practical reasons analysis of EEG often assumes *weak stationarity* in data windows of less than 30 seconds. Non-linear signal processing requires even longer windows of stationarity to be valid.

Normalization is imperative for the interpretation and comparison of data.

The remainder of this section discusses feature extraction methods relevant to the analysis of the epileptic EEG. The tools presented are applicable to any arbitrary discrete time signal.

We define the signal $y[n]$ to be an arbitrary time signal. During analysis the signal is *windowed* using a function $H_k[n]$ of length N. This window is typically rectangular and is incorporated implicitly as part of the computation. This means that in EEG analysis a windowed signal is denoted $y_k[n] = x_c[n]H_k[n]$. Time-evolution is achieved by varying the time k at which the window is applied.

For brevity theory is limited to discrete time analysis. Analogous results for continuous time signals exist and can be found in generic signal processing books such as [64], [138] and [143].

3.3.1 Time Domain Analysis

A signal $y[n]$ is a function of time n. It is defined within the *observation window* that depends on time. It can take values in a certain range, known as the *measurement range*. Estimating features that depend on time is known as *time domain analysis*. For the classification of the epileptic EEG time domain analysis is often used to give a numerical representation to qualitative visual observations, for example the increase in amplitude, increase in regularity

and increase in synchronicity observed during epileptic events. Statistics that characterize these observables are described in this section.

3.3.1.1 Signal Amplitude (Energy) and Variance (Power)

A signal's *instantaneous* amplitude at time n is given by $|y[n]|$, where $|.|$ is the magnitude. We will call this quantity the signal's *energy*. Its square, $|y[n]|^2$, we call the signal's *power*. Both power and energy give an idea of the magnitude of $y[n]$ at time n. Because of the squared term power emphasizes changes more than energy but is consequently more affected by noise. $y[n]$, $|y[n]|$ and $|y[n]|^2$ are all time domain signals because these types of transformations preserve dependence on time.

Instantaneous energy is rarely able to say anything about the waveform that cannot already be observed in the original signal. Averages over time are more useful because emphasis is placed on mean behavior – recall that estimates of statistics are only meaningful in ensembles rather than single instances. The mean of a signal $y[n]$ is estimated as

$$\mu_y[k] = \frac{1}{N} \sum_{n=k+1}^{k+N} y[n]. \tag{3.7}$$

$\mu_y[k]$ is the mean of sequence $y[n]$ of length N starting at time k. Here $y[n]$ can be the original recorded signal or can be replaced by any other signal such as its energy $|y[n]|$, its power $|y[n]|^2$ or another transformation altogether. If $y[n]$ is stationary with real mean $\hat{\mu}_y$ then the larger the value of N, that is, the more samples used, the closer that the estimate $\mu_y[k]$ is to the real mean $\hat{\mu}_y$. This can be observed in Figure 3.1.

The *variance* of a signal $y[n]$ is a statistical quantity that gives an idea of its spread and regularity by computing how much, on average, it deviates from its mean. An unbiased estimator of the variance of $y[n]$ is defined as

$$\sigma_y^2[k] = \frac{1}{N-1} \sum_{n=k+1}^{k+N} \left(y[n] - \mu_y[k]\right)^2, \tag{3.8}$$

where $\sigma_y^2[k]$ is the variance of a sequence $y[n]$ of length N and $\mu_y[k]$ is the mean as calculated by Equation 3.7. The square root of variance $\sigma_y[k]$ is known as the *standard deviation*.

Sometimes it is necessary to compare statistics extracted from signals that have significantly different mean values. In such cases a new statistic can be defined where the variance of a signal is normalized by its mean as

$$c_y[k] = \frac{\sigma_y[k]}{\mu_y[k]}. \tag{3.9}$$

Here $c_y[k]$ is known as the *coefficient of variation* (COV). This normalization is not ideal when $\mu_y[k]$ is close to zero because $c_y[k]$ becomes very

sensitive to changes in $\sigma_y[k]$. This limits the usefulness of COV to non-zero mean signals such as $|y[n]|$ or $|y[n]|^2$.

Short term non-stationarities (e.g., transient bursts) can affect the estimate of $\mu_y[k]$, and thus COV and variance, significantly. If we want to avoid this it is possible to talk instead about the *variability* of a signal rather than its variance. One way to describe variability is to count the number of times a signal changes polarity. A statistic of this nature is *total variation* $v_y[k]$ defined over the analysis window as [110]

$$v_y[k] = \frac{1}{N-1} \frac{\sum_{n=k+2}^{k+N} |y[n] - y[n-1]|}{\left(\max_{y[k]} - \min_{y[k]}\right)}. \tag{3.10}$$

It is defined only for *non-constant* signals $y[n]$ with maximum value $\max_{y[k]}$ and minimum value $\min_{y[k]}$ over the analysis window, defined as

$$\min_{y[k]} = \min_{k+1 \le n \le k+N} (y[n]) \tag{3.11}$$

$$\max_{y[k]} = \max_{k+1 \le n \le k+N} (y[n]). \tag{3.12}$$

The dividing term $(\max_{y[k]} - \min_{y[k]})$ included in Equation 3.10 normalizes results by the range of $y[n]$, thus making it comparable to estimates computed at other times. $v_y[k]$ takes on values between $\frac{1}{N-1}$ and 1, that is, $v_y[k] \in [\frac{1}{N-1}, 1]$. Slow, smooth signals have low total variation, with the lower bound of $v_y[k] = \frac{1}{N-1}$ achieved by functions that change monotonically between $\min_{y[k]}$ and $\max_{y[k]}$ over the analysis window ($n = k+1, k+2, \cdots, k+N$). Fast, large oscillations increase total variation with $v_y[k] \le 1$ and the upper bound achieved when $y[n]$ alternates between $\max_{y[k]}$ and $\min_{y[k]}$ at each consecutive n.

Let us now see how $\mu_y[k]$, $\sigma_y^2[k]$, $c_y[k]$ and $v_y[k]$ can be used for the purposes of detecting changes over time. Tracking of these quantities requires the choice of window length N used to compute each statistic, which can be problematic for mean, variance and COV. Long windows give a better estimate of long-term behavior, but short events are averaged out and cannot be identified. Conversely, short windows can provide better temporal resolution and detect short events, but good (or representative) estimates are not obtained. Total variation is more suitable for short windows of data. Fortunately all these statistics can be computed using relatively small N (in the order of a few thousand samples is often sufficient) so a suitable compromise is possible.

When tracking is specific to the detection of epileptic events ($y[n] = x_c[n]$) the mean of the raw EEG signal $x_c[n]$ is not so useful because these converge to DC offsets that do not change fast enough to characterize epileptic events. Energy and power can be better discriminators. Because DC offsets affect estimates of variance, COV and total variation $x_c[n]$ is *detrended* (the mean is removed) prior to computation. In any case the mean of a signal is

meaningless because it depends on the referencing system used. Voltage is a *relative* quantity defined only up to a constant (see Chapter 2).

Examples of mean signal energy, power, variance, COV and total variation for EEG are explored in Figure 3.6, where each is computed for a single EEG channel $x_c[n]$ using different window sizes ($N = 512$ and $N = 5120$, corresponding to 1 and 10 seconds of a signal sampled at 512Hz). Calculations are done once a second by incrementing k by 512 samples at each iteration. These estimates are computed from the same seizure shown in Figure 1.10 that occurs at approximately 620 seconds. Notice that short time windows give better representation of short events but the resulting estimates are not as smooth. All methods show some degree of change during the seizure, but the type of change is different for each method. The features in this figure are powerful in describing average trends but are not enough to discriminate between non-seizure, pre-seizure and seizure activity when a larger EEG database is studied.

Prior to analysis the EEG sequence is normalized to unit variance. Now that both mean and variance have been introduced we can formally define a detrended normalized signal (described in Section 3.2) as

$$y_{normalized}[n] = \frac{y[n] - \mu_y}{\sigma_y}, \tag{3.13}$$

where μ_y and σ_y are the mean and standard deviation of the entire signal $y[n]$, before windowing. When tracking changes it is also possible and sometimes necessary to apply this normalization only to the (shorter) analysis window.

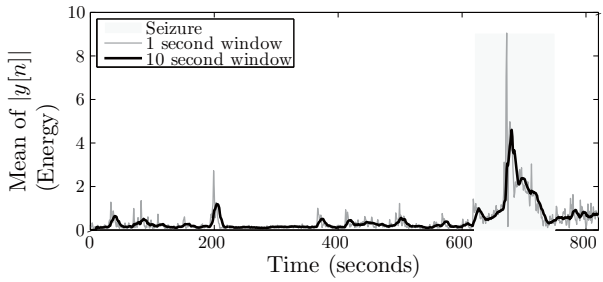
Both numerator and denominator in this equation have the same units and thus the resulting $y_{normalized}[n]$ is scale invariant. *This normalization should be applied to any signal prior to analysis.* This has been done for the examples presented in Figure 3.6, and is always implied in the work that follows.

It should also be noted that even after normalization statistics like mean energy, mean power and variance only make sense as relative measures. For example we can say that the variance in Figure 3.6(c) increases during the seizure relative to the non-seizure EEG, but its raw value does not matter. In contrast the total variation statistic is truly scale invariant because its ranges are well defined and changing the scale of $y[n]$ does not change the estimated statistic.

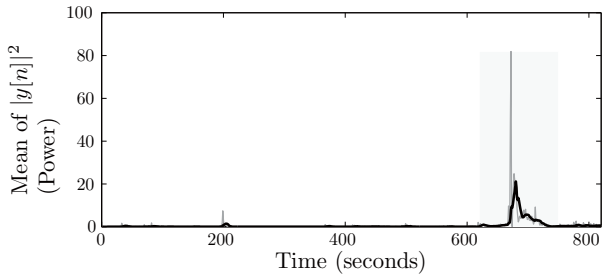
3.3.1.2 Periodicity (Auto-Correlation)

The *auto-correlation* function gives an idea of how much a signal repeats itself, and thus can be used to identify regularity. For a real signal $y[n]$ the auto-correlation function is defined as

$$\text{CORR}_y[\tau, k] = \frac{1}{N} \sum_{n=k+1}^{k+N} y[n+\tau]y[n], \quad 0 \leq \tau < N. \tag{3.14}$$

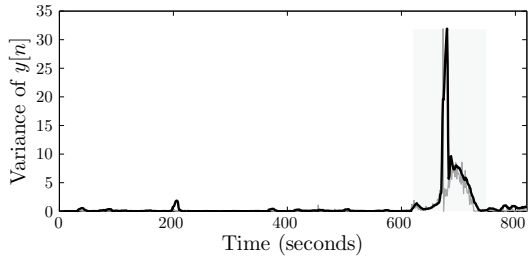


(a) Tracking mean energy

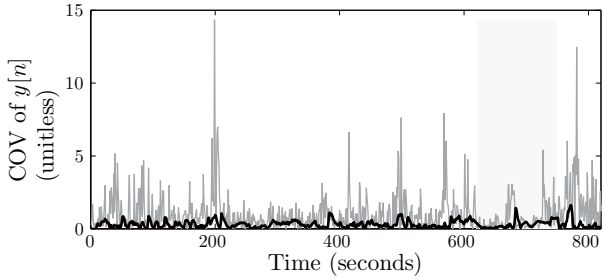


(b) Tracking mean power

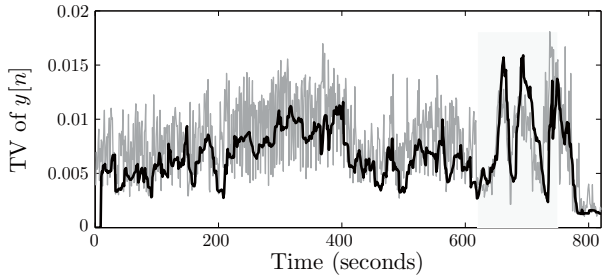
FIGURE 3.6: Computations of (a) mean energy ($|y[n]|$), (b) mean power ($|y[n]|^2$), (c) variance, (d) coefficient of variation (COV) and (e) total variation of EEG channel amplitude $|y[n]| = |x_c[n]|$. In all cases the entire data sequence was normalized as in Equation 3.13 prior to windowing. Tracking was performed using different length windows. The signal is sampled at $F_s = 512$Hz and thus 1 second windows correspond to $N = 512$ samples and 10 second windows have $N = 5120$ samples. All methods show that a degree of change is observed when the seizure occurs between 620-750 seconds, although this change is not necessarily consistent and these time domain methods may not by themselves be used to recognize seizures. In all cases, a longer window is shown to give smoother estimates, less susceptible to transient peaks and troughs in the data. (Continued)



(c) Tracking variance



(d) Tracking coefficient of variation



(e) Tracking total variation

FIGURE 3.6: (Continued)

Assuming stationarity in $y[n]$ over the analysis window, this function gives high values at delays τ for which $y[n]$ is most like itself, and values close to zero for delays for which $y[n]$ is least like itself. The maximum value occurs at $\tau = 0$ because at zero delay $y[n]$ exactly equals $y[n + \tau]$. In fact for a zero mean signal $\mathrm{CORR}_y[0, k]$ is, for large N, equivalent to the variance $\sigma_y[k]$ defined in Equation 3.8.

Equation 3.14 is defined for $-N < \tau < N$, but because it is symmetric for stationary signals it is sufficient to look at values for $0 \leq \tau < N$. For non-stationary $y[n]$ Equation 3.14 is not symmetric and the values of $\mathrm{CORR}_y[\tau, k]$ for $\tau > 0$ may be larger than $\mathrm{CORR}_y[0, k]$. When it is not known whether $y[n]$

is stationary or not, this can be used as an indicator. In practice as discussed earlier stationarity is implicitly forced onto $y[n]$ by windowing the signal with a small enough N.

Equation 3.14 is appropriate to identify regular and oscillatory behavior because these signals repeat themselves often, that is, $\mathrm{CORR}_y[\tau, k]$ is maximized for values of τ for which

$$y[n + \tau] \approx y[n]. \tag{3.15}$$

Periodicity of the signal $y[n]$ can be inferred from periodicity in the correlation function.

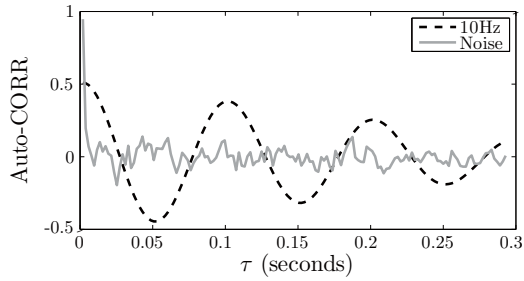
To visualize the effects of the auto-correlation observe the example described in Figure 3.7 using the two signals shown in Figure 3.9(a): $y[n]$ is random white noise and $y[n]$ is a 10Hz sinusoid. Figure 3.7(a) shows their auto-correlations. For the random noise $\mathrm{CORR}_y[\tau, k]$ is very small for all delays except $\tau = 0$, when the auto-correlation is 1. The sinusoid, however, shows maxima every 0.1 seconds because the $y[n]$ repeats itself at 10Hz. There are also minima every 0.1 seconds, corresponding to delays at which the sinusoid is a negative copy of itself. When the 10Hz sinusoid is corrupted with random phase noise, as in (b), the peaks and troughs of $\mathrm{CORR}_y[\tau, k]$ are preserved, albeit at a smaller magnitude and with more fluctuations. This is so even though the levels of noise are quite high – auto-correlation is a fairly robust measure.

In practice $y[n]$ is finite and Equation 3.14 is a *biased* estimate. For example, notice in Figure 3.7(a) and (b) that the auto-correlation at multiples of $\tau = 0.1$ seconds decreases with increasing τ, even though the sinusoid is an exact replica of itself and the auto-correlation should yield the same number. It occurs because with higher τ fewer samples are used and the dividing $1/N$ term biases values at larger τ. An *un-biased* estimate is achieved by modifying Equation 3.14 to

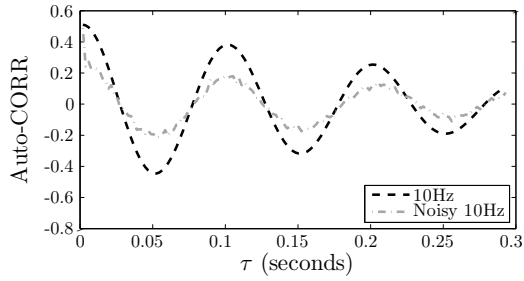
$$\mathrm{CORR}_y[\tau, k] = \frac{1}{N - |\tau|} \sum_{n=k}^{k+N} y[n + \tau]y[n], \quad 0 \leq \tau < N, \tag{3.16}$$

thereby accounting for the fewer samples used for higher τ. The effects of biased versus un-biased estimation are shown in Figure 3.7(c). Equation 3.16 is used from now on.

Finally, the comparison of the auto-correlations between functions of different magnitudes is not possible because the values of $\mathrm{CORR}_y[\tau, k]$ are not scaled to compatible ranges. *Measures such as auto-correlation should not depend on magnitude because the amount of information contained in* $y[n]$ *is exactly the same as* $ay[n]$, for any constant $a \in \mathbb{R}$. A normalized auto-correlation in which the values always lie between $[-1, 1]$ can be achieved by dividing $\mathrm{CORR}_y[\tau, k]$ by its power, $\mathrm{CORR}_y[0, k]$, so that, assuming stationarity in $y[n]$,

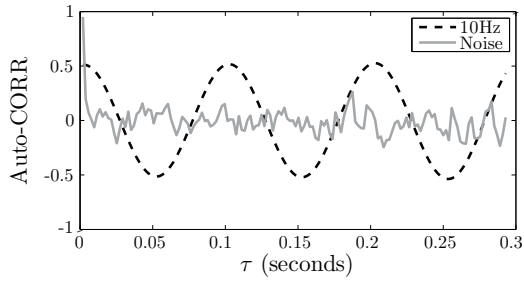


(a) Biased auto-correlation

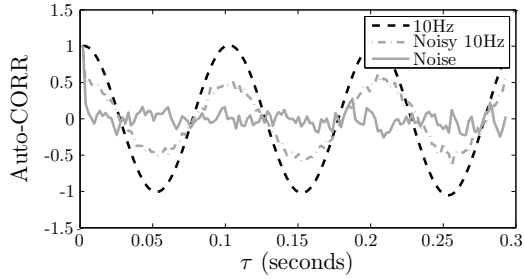


(b) Auto-correlation (biased) of noisy signals

FIGURE 3.7: The computation of auto-correlation. (a) shows a biased estimate for noise and sinusoid – the same signals shown in Figure 3.9(a). The auto-correlation gives high values for several τ only if the signal is regular. (b) shows how the addition of noise in a regular 10Hz signal can decrease performance. The auto-correlation still shows significant structure, albeit at smaller amplitudes, and is fairly resilient to noise. (c) shows an unbiased estimate of the same signals as in (a). (d) shows the same unbiased estimate as in (c), but this time normalized by signal power so that values always lie in the range $[-1, 1]$. All signals here are sampled at $F_s = 512$Hz, and so, for example, $\tau = 0.125$ seconds refers to a delay of $\tau = 512 \times 0.125 = 64$ samples. (Continued)



(c) Unbiased auto-correlation



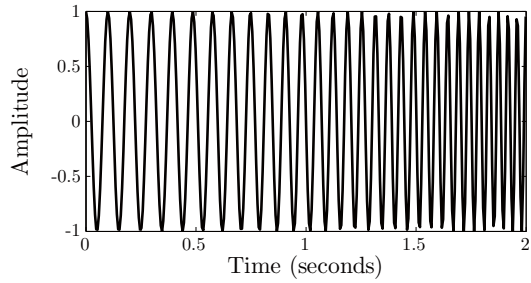
(d) Normalized unbiased auto-correlation

FIGURE 3.7: (Continued)

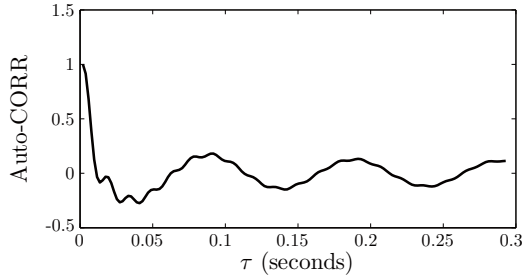
$$\mathrm{CORR}_{y-normalized}[\tau, k] = \frac{\mathrm{CORR}_y[\tau, k]}{\mathrm{CORR}_y[0, k]}. \tag{3.17}$$

Here the maximum value of 1 occurs at $\tau = 0$ independent of the magnitude of the signal. Normalized auto-correlation for the test signals is shown in Figure 3.7(d). Note also that this normalization is not required if the original signal $y[n]$ is normalized to unit variance over the analysis window, as defined by Equation 3.13, prior to analysis.

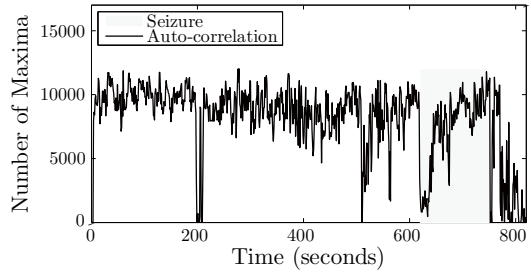
During a seizure there is often an observed increase in regularity in the EEG where the signal becomes more oscillatory. Thus when $y[n] = x_c[n]$ and we are interested in detecting seizures Equation 3.17 seems like an appropriate feature. The problem is that this estimate requires *stationarity* that cannot be assumed even within a seizure because its frequency evolves over time. The auto-correlation for non-stationary signals is not very robust – see for example Figure 3.8 where a *chirp*, a signal whose frequency evolves over time [163] (in this case from 10Hz to 25Hz), is shown in (a). Although this is a very regular signal its auto-correlation in (b) fails to yield high values because the times at which the signal is most like itself are not fixed (i.e., there is no single τ at



(a) Chirp signal



(b) Auto-correlation of a 10-25Hz chirp



(c) Auto-correlation for seizure detection

FIGURE 3.8: Effects of auto-correlation for time-varying signals. (a) shows a *chirp* – a sinusoid whose frequency evolves, in this case between 10-25Hz, over the course of time. Even though the signal is quite regular, auto-correlation fails to give high values for any τ other than $\tau = 0$, and as such auto-correlation is not suitable for non-stationary signals. Nevertheless auto-correlation may be used for seizure detection, as shown in (c), where the number of peaks and troughs observed in the auto-correlation in an EEG channel is computed over time. The assumption is that fewer peaks occur for repetitive signals. It is obvious that a change occurs when the seizure begins. All signals here are sampled at $F_s = 512$Hz. Tracking in (c) is done using 2 second ($N = 1024$ samples) non-overlapping analysis windows.

which this happens). Even the highly noisy example shown in Figure 3.7(b) yields larger auto-correlation values than does the highly regular chirp.

Shorter windows in which stationarity is more likely is unsuitable to detect lower frequency seizures because the periodicity is not obvious in a short amount of time. The applicability of auto-correlation as a detector is limited. Nevertheless in certain cases auto-correlation can be used to characterize a seizure, as shown in Figure 3.8(c), and this feature can be used as part of a feature set to detect seizures.

3.3.1.3 Synchronization

Qualitatively *synchronization* describes the amount of locking between signals. A measure of synchronicity gives an idea of how similar signals are to *each other*. Several methods exist to measure different types of synchronicity. The first is the same as the auto-correlation described by Equations 3.16 and 3.17 but applied to two different signals $y_1[n]$ and $y_2[n]$. An un-biased estimate of their linear cross-correlation is

$$\text{XCORR}_{y_1,y_2}[\tau,k] = \frac{1}{N-|\tau|}\sum_{n=k+1}^{k+N} y_1[n+\tau]y_2[n], \quad 0 \leq \tau < N. \tag{3.18}$$

The same issues as those described for auto-correlation apply and these are not repeated here. A normalized statistic suitable to compare signals of different amplitudes and lengths is given by

$$\text{XCORR}_{y_1,y_2-normalized}[\tau,k] = \frac{\text{XCORR}[\tau,k]}{\sqrt{\text{CORR}_{y_1}[0,k]\ \text{CORR}_{y_2}[0,k]}}, \tag{3.19}$$

where the power of both $y_1[n]$ and $y_2[n]$ are now used. The idea is that large values result when two signals are *lag synchronized*, that is, when

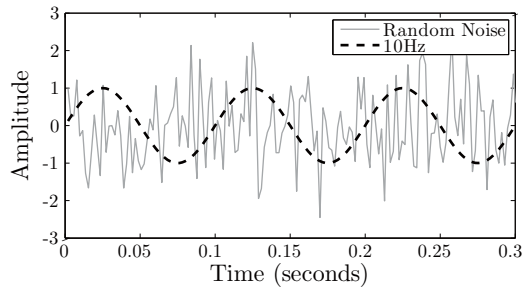
$$y_1[n+\tau] \approx y_2[n]. \tag{3.20}$$

High values result at time lags τ for which $y_1[n]$ and $y_2[n]$ have a similar course in time. Two signals are synchronized if the time lag τ for which high values of Equation 3.19 occur is similar for all time.

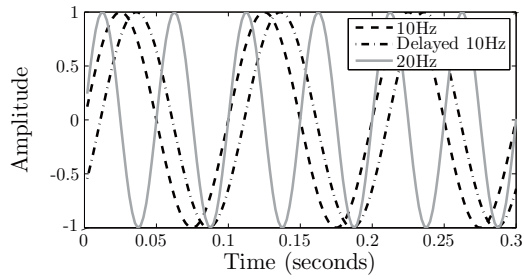
The cross-correlation computed by Equation 3.19 is linear in that it assumes that similarity between two channels is given by a linear combination of $y_1[n]$ and $y_2[n]$. Non-linear methods exist in which correlation calculations are maximized for times at which

$$y_1[n+\tau] = F(y_2[n]), \tag{3.21}$$

where $F(\cdot)$ can be a non-linear transformation. There are ways to estimate $F(\cdot)$ but these are computationally expensive, require large amounts of

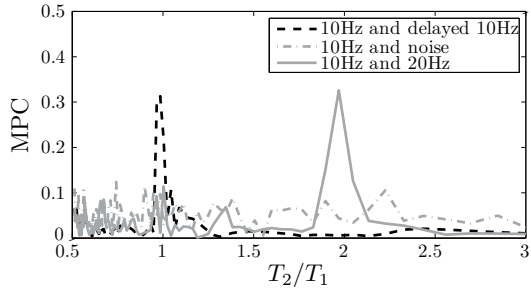


(a) Test signals 1

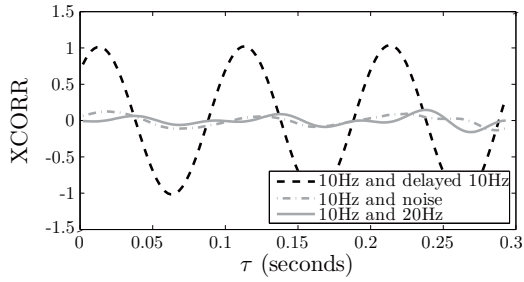


(b) Test signals 2

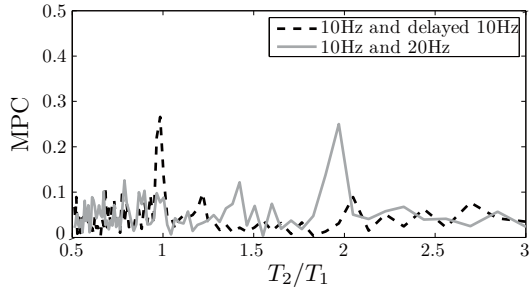
FIGURE 3.9: Synchronization calculations. (a) and (b) show test cases: noisy signal, 10Hz sinusoid, out of phase 10Hz sinusoid and a 20Hz sinusoid. (c) and (d) show the mean phase coherence (Equation 3.24) and cross-correlation (Equation 3.19) estimates for the test signals relative to a 10Hz sinusoid. Notice that the mean phase coherence can pick out both out of phase and different frequency synchronization, whilst cross correlation cannot be used when signals are of different frequency. Values are low in both cases for the noisy case, as expected. (e) shows the sensitivity of mean phase coherence when noise is introduced to signals known to be synchronized. Neither case is identified. Finally (f) shows how cross correlation between two channels may be used for the detection of epileptic seizures. The measure counts the number of peaks and troughs observed in the cross correlation between two EEG channels known to display synchronicity during seizure, under the assumption that fewer peaks occur for synchronized signals. There is an obvious change at the beginning of a seizure. All estimates are normalized to signal energy and signals are sampled at $F_s = 512$Hz. In (f) we use 2 second non-overlapping windows ($N = 2 \times F_s = 1024$ samples). (Continued)



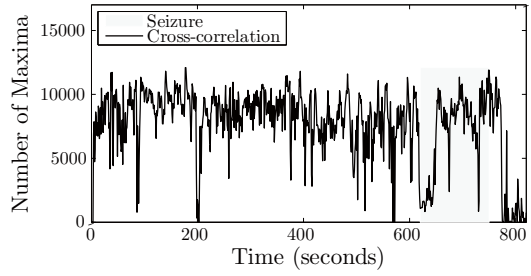
(c) Mean phase coherence of test signals



(d) Cross correlation of test signals



(e) Mean phase coherence for noisy signals



(f) Cross correlation for seizure detection

FIGURE 3.9: (Continued)

data and do not intuitively add to the material presented here. For more information, see, as an example, the work done in [28].

Another way that two signals can be synchronized is if they repeat themselves at regular intervals, but do not necessarily evolve at the same frequency. This is called *phase locking*, an example of which is shown in Figure 3.9(b). Phase locked signals are said to be $T_1 : T_2$ *synchronized*, where T_1 and T_2 are real numbers that represent the ratio between the fundamental frequencies of $y_1[n]$ and $y_2[n]$.

This type of synchronization can be measured using *instantaneous phase*, where an arbitrary signal $y[n]$ is decomposed into its instantaneous phase $\theta_y[n]$ and amplitude $A_y[n]$ as [28, 113]

$$\hat{y}[n,k] = y[n] + i\tilde{y}[n] = A_y[n]e^{i\theta_y[n]}, \quad k+1 \leq n \leq k+N. \tag{3.22}$$

$\tilde{y}[n]$ is the Hilbert transform of $y[n]$, and it ensures that $\hat{y}[n]$ has both a real and imaginary component[4] so that the phase

$$\theta_y[n,k] = \arg\left(\hat{y}[n,k]\right), \quad k+1 \leq n \leq k+N \tag{3.23}$$

exists. Here $\arg(x)$ indicates the angular component[5]of a complex number x. See, for example, [138] for the calculation of the Hilbert transform. Note that this is only one method for determining phase information, and others such as wavelet transforms and correlation coefficients can be used as a substitute. The choice of method does not impact phase estimation results perceptibly [28, 112].

Two signals y_1 and y_2 are said to be approximately $T_1 : T_2$ synchronized if $\theta_{y_1,y_2}[n,k] = T_1\theta_{y_1}[n,k] - T_2\theta_{y_2}[n,k]$ remains bounded for all n. This phase locking (or synchronization) can be quantified by taking the mean phase coherence

$$\gamma_{y_1,y_2}[T_1,T_2,k] = \left| \frac{1}{N} \sum_{n=k+1}^{k+N} e^{i\theta_{y_1,y_2}[n,k]} \right|, \tag{3.24}$$

which is an average of all the instantaneous phases over time. $\gamma_{y_1,y_2}[T_1,T_2,k]$ has a maximum value of 1 when both signals are phase locked, that is, $\theta_{y_1,y_2}[n,k] = 0$ for all n, and a minimum value of 0 when no phase locking occurs. Since most real signals are noisy it is unlikely that Equation 3.24 is ever either exactly one or exactly zero. It is expected that low values are

[4]This is called making a signal *analytic*, and is in some cases useful in the analysis and manipulation of signals.

[5]The angular component or argument of a signal is calculated as $\arg(x) = \arctan\left(\frac{\mathrm{Im}(x)}{\mathrm{Re}(x)}\right) + 2m\pi$. Because the arctan() function yields values between $[0, 2\pi)$, m is an integer used to ensure that phase differences greater than 2π can be computed in Equation 3.24.

achieved for signals that are not phase locked in general and high values for those that are.

In contrast to lag synchronization that uses both the phase and amplitude information of $y_1[n]$ and $y_2[n]$, Equation 3.24 only uses the phase information and therefore no amplitude normalization is necessary.

The effects of synchronization are explored in Figure 3.9. Test signals, including random noise, phase locked signals of the same frequency and phase locked signals of different frequencies can be seen in (a) and (b). $T_1 : T_2$ and lag synchronization are presented in (c) and (d) respectively. As expected, both types of synchronization do not respond to random noise. But whereas lag synchronization gives high values for phase locked signals of different frequencies (identifying the factor of two between frequencies), using the cross-correlation does not. This implies that mean phase coherence is a more appropriate choice for measuring synchrony in systems in which two signals are in phase but not behaving similarly.

The effects of noise can be devastating on estimates of $\gamma_{y_1,y_2}[T_1, T_2, k]$, as shown in Figure 3.9(e), whereas cross-correlation (like auto-correlation) is fairly robust to it. This trade-off, the computational expense incurred by not knowing T_1 and T_2, and the long stationary windows of data required to obtain stable estimates make mean phase unsuitable for many tasks.

Synchronization has been described in Chapter 1 as a key observable during many types of seizures, and as such is a suitable feature to extract for the detection of epilepsy. Figure 3.9(f) shows how lag synchronization may be used to detect epileptic events using EEG data, where the temporal evolution between two channels, selectively chosen to produce good results, is seen to identify a seizure.

Statistics discussed in this section describe *time domain* features of the epileptic EEG.

- *Amplitude*: Energy $|y[n]|$ and power $|y[n]|^2$ of a signal $y[n]$. Mean values of these can be computed as in Equation 3.7.
- *Variability and Regularity*: Variance $\sigma_y[k]$, coefficient of variation $c_y[k]$ and total variation $v_y[k]$ all give an idea of how variable $y[n]$ is by quantifying how much it deviates from its mean or across its range. Periodicity can be deduced from the autocorrelation function $\text{CORR}_{y-normalized}[\tau, k]$ defined in Equation 3.17.
- *Synchronicity*: How much two signals $y_1[n]$ and $y_2[n]$ are like each other can be estimated through Equation 3.18, where the cross correlation $\text{XCORR}_{y_1,y_2-normalized}[\tau, k]$ is defined, or through mean phase coherence $\gamma_{y_1,y_2}[T_1, T_2, k]$ found in Equation 3.24.

To remove dependence on mean and scale these statistics

should be applied to a signal $y[n]$ that has been normalized as in Equation 3.13 over the analysis window.

By themselves these statistics are insufficient to reliably detect seizures.

3.3.2 Frequency Domain Analysis

Frequency is a measure of how often an event occurs in a unit period of time. If a ball is bouncing once a second, then it is hitting the ground at a frequency of 1Hz, twice a second is 2Hz and so on. Many of the time domain methods presented in the previous section implicitly include the notion of frequency. For example, the auto-correlation detects repeated events at a particular frequency or period, just as the cross-correlation detects similarity in the frequencies between two signals.

Signals such as the EEG contain events that occur at different frequencies. What these are is not always obvious in the time domain because events at many frequencies interfere with one another. To make the frequency content more clear a *transformation* is applied to a signal $y[n]$ so that it is defined in terms of frequency ω rather than time n. For example suppose that a signal $y[n]$ is the sum of two sinusoids at 10 and 20Hz respectively, as shown in Figure 3.10(a) for both a noise-free and a noisy case. Here it is not so obvious what these frequencies are, particularly in the noisy case. However when described in terms of ω it is clear, as in Figure 3.10(b), that there are 10 and 20Hz components. Even fairly high noise levels do not significantly interfere with this observation.

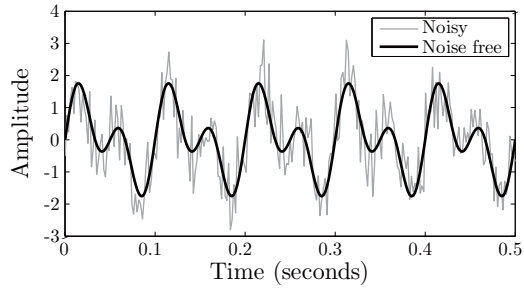
A signal can always be constructed as a linear combination of *basis functions*. In the time domain these basis functions are implicitly an identity function that isolate elements in time. That is, the basis function is defined as

$$b_n[k] = \begin{cases} 1, & \text{if } k = n \\ 0, & \text{otherwise} \end{cases} . \tag{3.25}$$

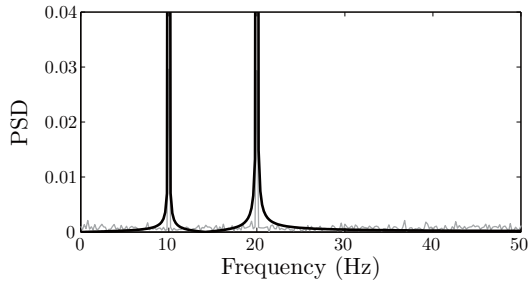
The original time-domain signal can then always be reconstructed as

$$y[n] = \sum_{k=-\infty}^{\infty} y[k] b_n[k], \quad n = 1, 2, 3 \cdots . \tag{3.26}$$

In the frequency domain the basis functions isolate the different frequency components of the signal $y[n]$ by projecting it onto *sinusoidal* basis functions. Sinusoids are chosen because they are very good at isolating components of different frequencies. The signal is then said to be described in terms of its frequency components and is defined in the *frequency domain*. Estimating features that depend on frequency is known as *frequency domain analysis*. The



(a) Test signals



(b) PSD of test signals

FIGURE 3.10: The power of using the FFT, as shown in (a) and (b), is its ability to reveal the frequency content in signals even in a noisy environment. The test signals are in (a) and the corresponding PSDs in (b). A sampling frequency of $F_s = 512$Hz is used (1 second equals $N = 512$ samples); thus the spectrum in (b) is defined between 0-256Hz, even though this figure shows only frequencies up to 50Hz. All PSDs are normalized to total PSD energy.

transformation of a time domain signal to the frequency domain is known as the *Fourier transform*, named so in honor of the French mathematician and physicist Jean Baptiste Joseph Fourier.

For discrete-time finite time domain signals[6] the *fast Fourier transform* (FFT) of an arbitrary windowed signal $y[n]$ for $n = k + 1, k + 2 \cdots k + N$ is given by

$$\mathrm{FFT}[\omega, k] = \sum_{n=1}^{N} y[n+k]\mathrm{e}^{-i\omega n}, \quad \omega = \frac{2\pi m}{N}, \quad m = 0, 1 \cdots N-1. \tag{3.27}$$

The sinusoidal basis functions $\mathrm{e}^{-i\omega n} = \cos(\omega n) - i\sin(\omega n)$ are able to

[6]The interested reader can find definitions of the Fourier transform for other types of signal, including those defined in continuous time, in signal processing books such as [138].

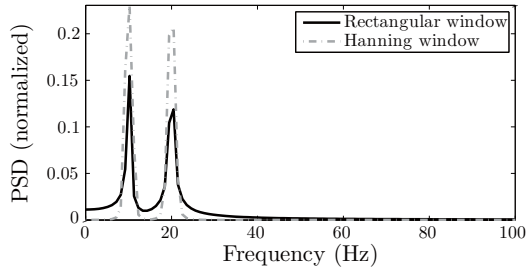


FIGURE 3.11: The effects of using different windows when computing the FFT. A test signal with 2 sinusoidal frequency components (10 and 20Hz) is used. When a rectangular window, defined in Equation 3.4 is applied to the time domain signal its PSD shows frequency content other than the 10 and 20Hz. This effect is reduced when the Hanning window defined in Equation 3.5 is used. Test signals are sampled at $F_s = 512$Hz; thus the PSD is defined for frequencies up to 256Hz, even though only the 1-100Hz range is shown.

isolate activity at different frequencies ω, measured in radians. The value of the FFT at each ω represents the relative contribution of events that occur at that frequency to $y[n]$. The FFT is defined for $0 \leq \omega = \frac{2\pi m}{N} < 2\pi$ with $m = 0, 1..N-1$. That is, the range 0 to 2π is divided in equidistant segments dependent on the number of samples in the windowed $y[n]$. To scale to correct frequency range in Hz, a conversion of $\omega = 2\pi f/F_s$ is necessary, where F_s is the sampling rate of the data and f is the frequency in Hz between zero and F_s.

Equation 3.27 implicitly uses a rectangular window to define the analysis data. However the rectangular edges of the window can affect the accuracy of the computed statistic by causing the estimated content at different frequencies to interfere with one another. This is known as *spectral leakage*. To manage this effect the windowed signal $y[n]$ is often first multiplied by a non-rectangular function such as the *Hanning window* shown in Figure 3.5(b). Although the Hanning window does not remove the spectral leakage it re-distributes it to nearby frequencies only. Thus better estimates are obtained at the expense of lower frequency resolution. An example is shown in Figure 3.11.

The transformation to the frequency domain preserves *all* the information contained in $y[n]$. The original signal over the window $k+1 \leq n \leq k+N$ can be unambiguously re-constructed using the *inverse Fourier transform*

$$y[n+k] = \frac{1}{N} \sum_{\omega=0}^{2\pi(N-1)/N} \text{FFT}[\omega, k] \mathrm{e}^{i\omega n}, \quad n = 1, 2, \cdots, N. \tag{3.28}$$

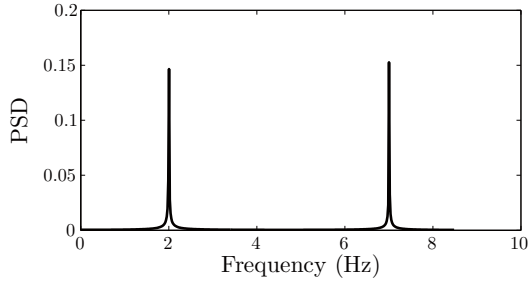


FIGURE 3.12: The effects of aliasing are shown. A test signal with 10 and 15Hz components, but sampled at $F_s = 17$Hz (1 second equals $N = 17$ samples) is used. Because the Nyquist criterion is not observed the frequency content of the signal is not reflected in the normalized PSD. $F_s = 17$Hz and the PSD is defined for frequencies up to 8.5Hz.

When computing Equation 3.27 the longer $y[n]$ is, the finer the *resolution* of the FFT. For example, if $F_s = 512$Hz and $N = 512$ samples are used, then the resulting frequency resolution is 1Hz because there are 512 values evenly distributed between 0 and 512Hz. If $N = 1024$ samples, then the resolution is 0.5Hz.

The second half of $\text{FFT}[\omega, k]$ can be discarded because frequencies beyond the Nyquist frequency $F_s/2$ are irrelevant, as explained in Section 3.2. If frequencies between $F_s/2 < f < F_s$ are not removed from $y[n]$ prior to sampling then *aliasing* (a form of computational noise) occurs, the effects of which are shown in Figure 3.12. This is measurement error that cannot be removed after sampling.

A related statistic is how much power each single frequency component contributes to the overall signal $y[n]$. This *power spectral density* (PSD) is defined as a function of ω as

$$\text{PSD}[\omega, k] = |\text{FFT}[\omega, k]|^2 , \tag{3.29}$$

where $|\cdot|$ denotes the absolute value. The notion of power is preserved between time and frequency domain: $\text{CORR}_y[0, k]$ in Equation 3.14 is equivalent to $\text{PSD}[\omega, k]$.

Equation 3.29 is not normalized and the comparison between signals of different amplitude and power can be misleading. A relative statistic can be derived by normalizing to the total signal power, that is divide Equation 3.29 by

$$\text{Normalizing Factor} = \sum_{\omega} \text{PSD}[\omega, k]. \tag{3.30}$$

With this type of normalization the PSD does not show how much power

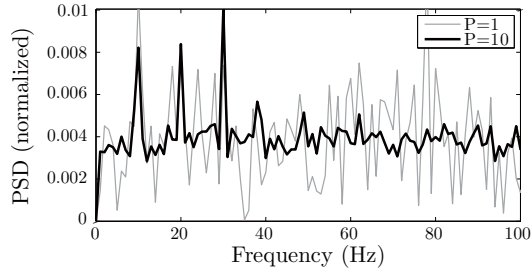


FIGURE 3.13: Effects of averaging to correctly represent the frequency content of noisy signals. One average ($P = 1$ in Equation 3.31) does not show the 10, 20 and 30Hz components. Using $P = 10$ does. The test signals are sampled at $F_s = 512$Hz and thus the PSD is defined up to 256Hz. The above plot only shows frequencies between 0-100Hz. The averaged PSD is then normalized using normalizing factor in Equation 3.30.

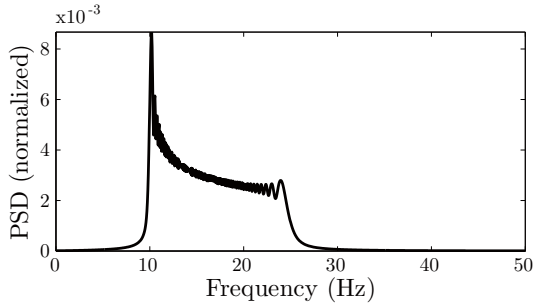
is in the signal, only the *relative* contribution of each frequency component. If it is important to compare the power of two signals then, so long as the two have been correctly normalized to the same scale, a complementary measure such as those presented in Section 3.3.1 should be used.

The statistical fluctuations in time domain signals make a single instance of a computed PSD, normalized or otherwise, unrepresentative of the frequency content of $y[n]$. Looking at averages over many trials is again more indicative of a general trend. If the N samples in the analysis window of $y[n]$ are divided into P segments of equal length, then an average PSD can be computed as

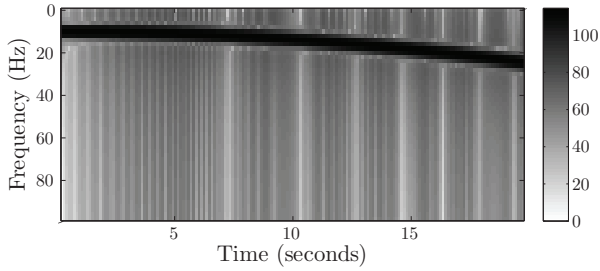
$$\mathrm{PSD}[\omega, k] = \frac{1}{P} \sum_{p=0}^{P-1} \left| \mathrm{FFT}_p[\omega, k] \right|^2 , \tag{3.31}$$

where $\mathrm{FFT}_p[\omega, k]$ is computed by Equation 3.27 applied to each N/P length segment. To understand the importance of this averaging consider a sinusoid mixture of 10, 20 and 30Hz embedded in a significant level of noise. The process of computing the PSD for $P = 1$ and $P = 10$ with $N = PF_s$ is shown in Figure 3.13. Using $P = 1$ provides a PSD estimate that has larger frequency resolution but that cannot isolate the important frequencies because the estimate is too noisy. With $P = 10$ the dominant frequencies 10, 20 and 30Hz show peaks well above background noise – this noise is averaged out – at the expense of reduced frequency resolution. The trade-off for a signal with N samples is then between better resolution (i.e., P is smaller) versus better characterization through a greater number of averages (P is larger). Too small a P and the PSD may look noisy. Too large a P and not enough frequencies are resolved.

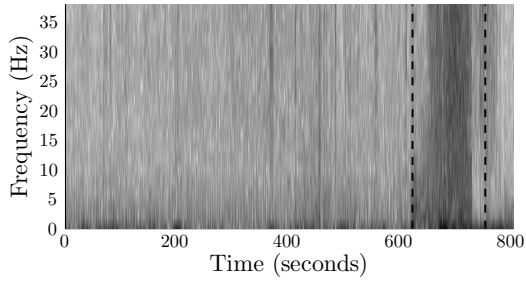
A final problem is once again that of stationarity. The PSD computed from



(a) PSD of a chirp



(b) Spectrogram of a chirp



(c) Spectrogram of EEG

FIGURE 3.14: The PSD of non-stationary signals. (a) shows the PSD of a *chirp*, described in Figure 3.8(a). It is obvious that the frequency content is between 10-25Hz, but it is not possible to say at what times this content occurs. (b) is a *spectrogram* that can localize both frequency and temporal aspects of the chirp. Spectrograms compute the PSD of short-time windows. (c) shows the spectrogram of an EEG sequence. A seizure can be identified at 620 seconds (between the dashed lines). Tracking in (b) and (c) is done using 0.5 second ($N = 256$ samples with sampling frequency $F_s = 512$Hz) non-overlapping analysis windows. The PSD at each 0.5 second window is defined for frequencies up to 128Hz, of which only the 0-40Hz range is shown.

a signal $y[n]$ that is non-stationary over the analysis window $n = k + 1, k + 2 \cdots k+N$ is only able to generally reflect its frequency content. Consider again the chirp signal shown in Figure 3.8(a). This signal's frequency content evolves from 10Hz to 25Hz over an analysis window, assumed to be 20 seconds for the computations here ($N = 10240$ samples when $F_s = 512$Hz). Its normalized PSD is shown in Figure 3.14(a), where it is evident that there is activity over the entire 10-25Hz range.

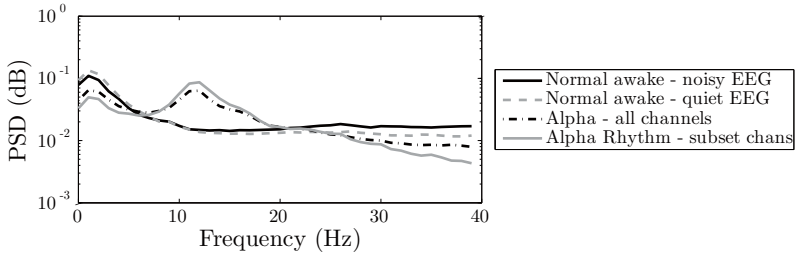
However *the PSD is by design incapable of isolating the time at which these frequencies occur*. One way to see this is by again inspecting the basis functions used to decompose $y[n]$. The time domain basis function in Equation 3.25 can define $y[n]$ for every single sample, regardless of the sampling rate. The sinusoidal basis functions used in the FFT computation require many more data samples to identify frequency content. Thus frequency resolution is gained at the expense of temporal resolution, and changes in frequency over the analysis window are always lumped together.

To track the temporal evolution of the frequency content we can instead draw a *spectrogram*. This is the sequential computation of the PSD of a signal for smaller analysis windows over time. In Figure 3.14(b) the analysis window used is now 0.5 seconds ($N = 256$ samples) long, using non-overlapping segments. In this figure the evolution from 10 to 25Hz, and the time at which these changes happen, is visible. The windows must be long enough to allow appropriate computation of Equation 3.31 (with sufficient averaging and high enough frequency resolution), yet short enough to correctly characterize the evolution of a non-stationary signal.

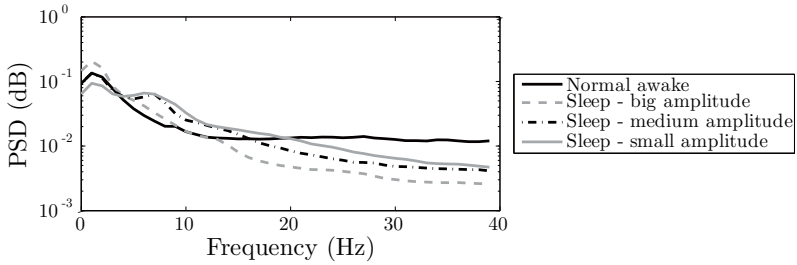
PSD analysis is an important tool used to understand the static and dynamic properties of the EEG, where static properties refer to locally stationary behavior, and dynamic properties aim to capture the time-evolving nature of the EEG. Figure 3.14 describes dynamic properties, and an example in which significant differences exist between the spectrogram of normal and seizure EEG is shown in (c).

The static properties of the EEG are highlighted in Figure 3.15, where the PSDs of typical EEG sequences described in Chapter 1 are computed. All data used are sampled at $F_s = 512$Hz and PSDs are computed on detrended data to which a Hanning window was applied prior to analysis. PSDs were computed to a resolution of 1Hz, meaning that $N = 512$ samples (1 second) were used for each computation. Given this, as many averages were taken as there was data available. For example, in (a) 30 seconds of data were used, thus there were $P = 30$ averages used to compute the PSD. In all cases at least 10 seconds of data were available and thus $P \geq 10$ averages in Equation 3.31. Important observations from this figure include:

- *PSDs of typical EEG sequences*: In Figure 3.15(a-d) the differences in PSD estimates between normal awake, normal asleep, pathological and artifactual activity reflected in common scalp recorded EEG are shown. Plots are computed with data taken from the same patient,

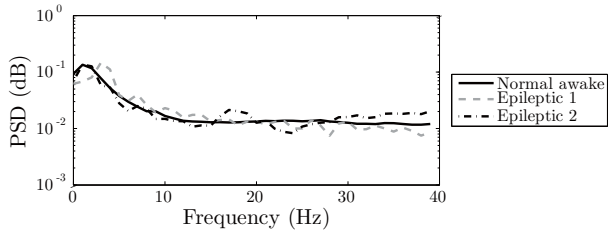


(a) PSD of normal EEG

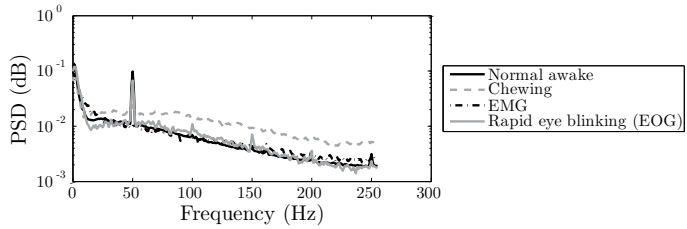


(b) PSD of sleep EEG

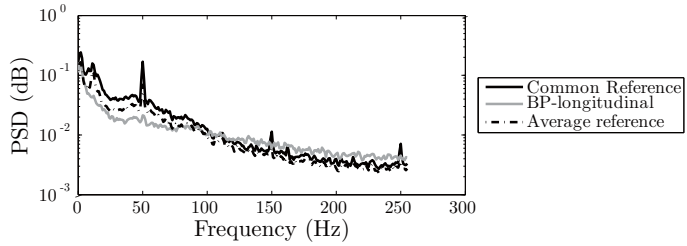
FIGURE 3.15: Example PSDs for EEG sequences. (a-f) show examples of the PSD of typical EEG sequences. Notice that even in normal states the EEG can look significantly different in (a), the lower frequency peaks of sleep EEG in (b), and the difficulty in differentiating epileptic EEG from normal sequences in (c). Artifact is easier to identify, particularly using the higher frequency range, as shown in (d). In (e) bipolar referencing is shown to de-emphasize low frequencies whilst emphasizing high frequencies. Finally the higher energy intra-cranial signals and low pass filtering resulting from spatial averaging of the skull and scalp are shown in (f) with simultaneous intra-cranial and scalp EEG records. All PSDs except in (f) are normalized to total PSD energy. All PSDs are computed on 1 second segments ($N = 512$ samples with $F_s = 512$Hz) for as much data as are available. An average of all segments is performed and the resulting PSD is normalized. The PSD is defined for frequencies up to 256Hz in all cases, even when a smaller range is shown. (Continued)



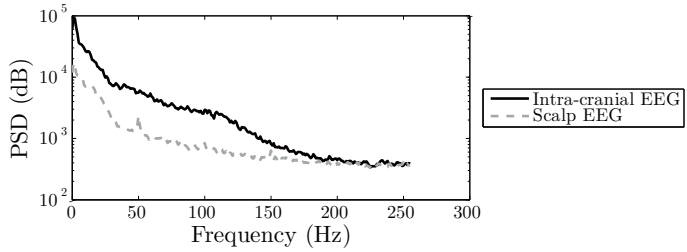
(c) PSD of seizure EEG



(d) PSD of artifactual EEG



(e) PSD of different referencing systems



(f) PSD of intra-cranial and scalp EEG

FIGURE 3.15: (Continued)

unless otherwise stated. The normal and sleep data used to compute the PSD of (a) were as in Figure 1.9 and epileptic data in (c) are taken from Figure 1.10.

(a) shows that the EEG is non-stationary because even the normal awake state in the same patient can vary significantly. One trace shows the pronounced alpha peak observed during the alpha rhythm, described in Chapter 1, while two other traces that seem very different in the time domain (shown in Figure 1.9) have almost identical PSDs once they have been normalized. The observed differences occur largely at higher frequencies, in itself an indication of non-stationarities perhaps driven by varying inputs. Selecting a subset of channels can emphasize certain features of the EEG. For example it is the occipital electrodes that are most affected by the alpha rhythm. When the PSDs of this subset of electrodes is calculated a 10Hz peak is more pronounced.

(b) shows that the frequency content in the sleep EEG varies from the normal awake EEG. Sleep EEG is known to contain slower and larger amplitudes than awake EEG. Although the amplitude difference cannot be seen here because the PSD energies have been normalized by Equation 3.30, in general low frequencies contribute more than high frequencies during sleep EEG. The alpha peak is not at all present during sleep, although different peaks due to spindles and REM may appear at times.

(c) shows that there are differences between epileptic and normal EEG. A segment from two different seizures, one with fundamental frequency 4Hz and another one with peaks at 1 and 15Hz, is shown. The latter can be hard to distinguish from normal EEG using PSD alone, hence the difficulties involved in seizure detection. Other measures such as synchronization, periodicity and energy must supplement frequency estimates.

Finally, (d) gives an idea of how common artifact may be differentiated from normal EEG. In general artifact that is muscular in nature (EMG or chewing) involves higher contributions from frequencies above roughly 40Hz, although they look similar to normal activity in the 1-40Hz range. EMG is not always separable from brain activity. Other common artifacts such as EOG (ocular/blinking artifact) have relatively higher contributions from frequencies < 2Hz and less contribution from frequencies between 10-25Hz. Shown also in one of these traces is a very large peak at 50Hz and its harmonics (100Hz,150Hz,200Hz...). These exist because of interference from electrical equipment that is powered by a 50Hz AC supply[7]. If analysis overlaps the 50Hz frequency range, a notch filter that removes this contribution should be applied to the data.

[7]This peak occurs at 60Hz in some countries, including the USA and Japan.

As an aside it is important to notice that there is an impressive difference in the way that the time and frequency domain describe the information in a signal, and yet because it is an invertible operation both time and frequency domain *contain the same information.*

- *Effects of referencing system*: Figure 3.15(e) shows that the referencing system used can affect the PSD distributions. These effects have been described in detail in Chapter 2; here we show that experiment coincides with theory. In Figure 2.17 it was predicted that bipolar referencing emphasizes high frequencies and attenuates low frequencies. This is supported in Figure 3.15(e). Also shown is the localizing nature of bipolar referencing that *de-emphasizes* the alpha peak because frontal electrodes are used. Some of these features are also observed when an average reference (taken over all available electrodes) is used, although the effects are not as pronounced because the spatial sampling frequency is too low in 21-channel scalp recorded EEG. Furthermore, the alpha peak is emphasized rather than de-emphasized seeing as an average montage considers both frontal and occipital electrodes.

- *Effects of skull*: Figure 3.15(f) shows the effects that the skull has on EEG records when the PSDs are computed from simultaneous intracranial and scalp records during a normal awake state. These PSD estimates have not been normalized to show firstly the larger amplitude/energy of intra-cranial versus scalp EEG at low frequencies; second that high frequency content is roughly the same; and third that a low pass filter for frequencies higher than about 20Hz occurs between cortex and scalp. These effects were all predicted in Chapter 2.

In the time domain a signal $y[n]$ can be constructed as the linear combination of *identity basis functions* described by Equation 3.25. In the *frequency domain* $y[n]$ is expressed in terms of *sinusoidal basis functions* capable of isolating activity at different frequencies, as in Equation 3.27. *Time domain and frequency domain representations contain exactly the same information*, but the features accentuated in each domain differ.

In particular, examples are shown where the *power spectral density* (PSD) of a time evolving signal can differentiate between many typical states of EEG. The PSD is a statistic that estimates the amount of signal power caused by events of each frequency. It is computed as in Equation 3.29. Comparing the relative energies between two signals is possible when the normalizing term in Equation 3.30 is used.

3.3.3 Time-Frequency Analysis

The time domain statistics in Section 3.3.1 fail to provide sufficient frequency content information, an aspect crucial to EEG classification, and the frequency domain techniques Section 3.3.2 designed for stationary processes can provide temporal information only by windowing $y[n]$. The difficulty is in selecting appropriate window sizes so that results are optimal and stationarity assumptions are not violated. In this section statistics *designed* to resolve both temporal and frequency content for non-stationary signals are presented. Several approaches exist, including Gabor atoms and Wigner-Ville distributions, but by far the one that has received the most attention in EEG is *wavelet analysis*.

A *wavelet* $\psi[n]$ is a function that [143]

1. Integrates to zero: $\sum_{n=-\infty}^{\infty} \psi[n] = 0$, and
2. Has finite power: $\sum_{n=-\infty}^{\infty} |\psi[n]|^2 < \infty$.

By themselves these properties do not make wavelets very special. With appropriate use, however, wavelets can be used as basis functions to provide a combination of both temporal and frequency information. Observe that time domain (identity) basis functions in Equation 3.25 violate the first condition, whilst frequency domain (sinusoidal) basis functions in Equation 3.27 violate the second.

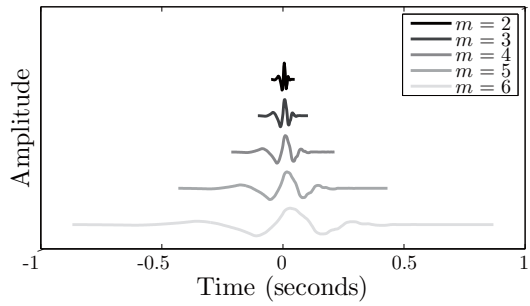
Some definitions are needed before this can be understood. Let $\psi_{ab}[n]$ represent a wavelet function that is shifted (or translated) by b samples and scaled by a, that is,

$$\psi_{ab}[n] = \frac{1}{\sqrt{a}} \psi[\frac{n-b}{a}]. \tag{3.32}$$

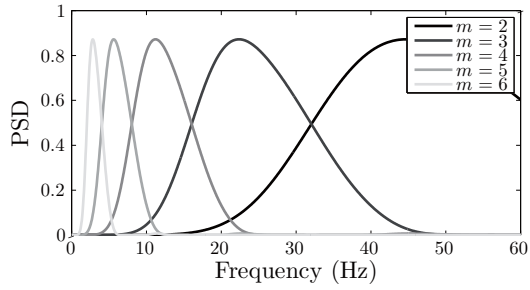
When $a = 1$ and $b = 0$ then $\psi_{10}[n]$ is known as the *mother wavelet*. With $0 < a < 1$ the mother wavelet contracts in time, and when $a > 1$ then $\psi_{ab}[n]$ stretches in time. The larger the a the longer the wavelet function is. Each value of a represents different *scales* at which temporal and frequency content can be extracted with different *resolutions*.

It is this scaling property that makes wavelets so powerful. A wavelet that is more compact in time can obtain new information in fewer samples than a wavelet that is temporally longer. Therefore smaller values of a give finer or more detailed temporal information than larger values of a. This is known as the *temporal scaling*. Temporal scaling for the Daubechies-4 (D4) wavelet proposed by Ingrid Daubechies in 1987 [143] is presented in Figure 3.16(a). Scales $a = 2^2, 2^3..2^6$, all with $b = 0$, are shown.

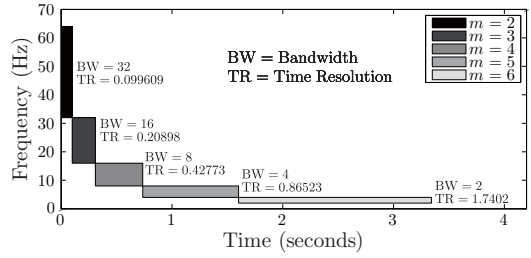
Now consider the frequency content of wavelets for different values of a. This is shown in Figure 3.16(b), where the PSD as computed by Equation 3.29 is applied to the wavelet at each scale. It is shown that different scales occupy different frequency regions. Larger values of a occupy lower regions of the frequency spectrum, and smaller values of a occupy higher frequencies.



(a) Time scale of wavelets



(b) Wavelet frequency bands



(c) Time-frequency resolution of wavelets

FIGURE 3.16: Wavelet Fundamentals. (a) and (b) respectively show the time and frequency scaling properties of the Daubechies-4 (D4) family for scales $m = 2 - 6$ and $b = 0$. Smaller scales give better temporal resolution but less frequency resolution, whereas larger scales do the opposite. In the frequency domain, larger scales occupy lower frequency regions, hence their ability to isolate behavior at different frequencies. The wavelets in (a) have been normalized to uniform amplitude. (c) shows the *uncertainty principle* of time-frequency analysis. Resolution can be high for either time or frequency, but not both. Intermediate scales show intermediate resolutions for both. Signals in this figure are sampled at $F_s = 512$Hz, thus 1 second is $N = 512$ samples.

Furthermore the *bandwidth* of each wavelet, that is, the range of frequency spectrum most relevant for each wavelet, becomes *smaller* for larger a – finer frequency details can be distinguished with larger values of a. This is known as the *frequency scaling* of wavelets. The $1/\sqrt{a}$ in Equation 3.32 ensures that the magnitude of the FFT at all scales is uniform.

$\psi_{ab}[n]$, described by Equation 3.32, is a function known as a *wavelet* that for small values of a has fine temporal resolution and coarse frequency resolution. Fine frequency resolution but coarse temporal resolution can be obtained with larger values of a. *Resolution can be high for either temporal or frequency information, but not both.* This is known as the *uncertainty principle* of time-frequency analysis, shown graphically in Figure 3.16(c). Depending on the application a suitable compromise must be reached with intermediate values of a.

Temporal evolution using wavelets can be obtained by *shifting* the wavelets with different values of b. This allows them to characterize the non-stationary nature of signals.

So far discussion has focused on a wavelet and its properties. To use wavelets to obtain temporal and frequency information from a signal $y[n]$ we must introduce the *wavelet transform* $\mathrm{W}_{ab}[n]$. For each value of a and b it is defined over the analysis window as

$$\mathrm{W}_{ab}[n,k] = \sum_{\tau=k+1}^{k+N} y[k]\psi_{ab}[n-\tau], \quad k+1 \leq n \leq k+N. \tag{3.33}$$

When applied to $y[n]$ the wavelet transform retains only the frequency content that is within the same bandwidth as that of $\psi_{ab}[n]$. This operation is known as *convolution* and its selective attenuation of the frequencies in $y[n]$ is known as *filtering*. More detailed information on these concepts including their broad applicability can be found in [138].

Orthogonal or *non-redundant* wavelets are those for which the frequency content at different scales do not overlap, so that when the wavelet transform is applied the resulting information is different at each scale. Orthogonal wavelets are achieved by *dyadic sampling*, that is, by restricting the dilations and translations of the mother wavelet to $a = 2^m$ and $b = 2^l$ with m and l both integers. If only orthogonal wavelets are allowed then a signal $y[n]$ is unambiguously represented by its decomposition into the different bases, so that [65]

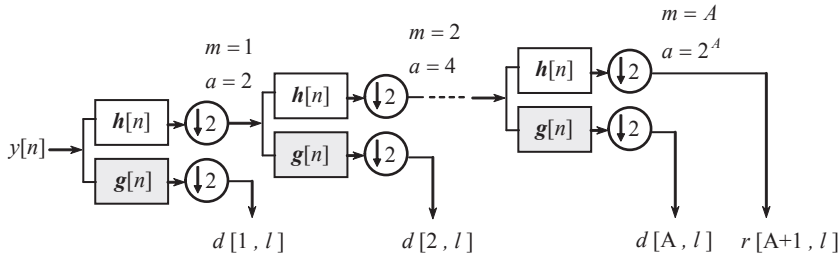


FIGURE 3.17: Wavelet Decimation. The figure shows the decimation procedure for wavelet decomposition into components shown in Equation 3.34. The coefficients $d[m,l]$ and $r[A+1,l]$ are computed for each value of m by iteratively filtering the previous stage result using filters $g[n]$ and $h[n]$ derived from the coefficients in Table 3.1, and then down sampling the result by a factor of 2.

Daubechies-4 (D4) Filter Coefficients							
0.1629	0.5055	0.4461	-0.0198	-0.1323	0.0218	0.0233	-0.0075

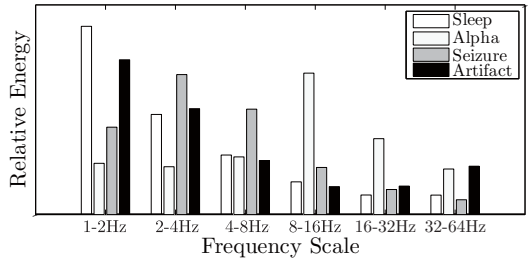
TABLE 3.1: Example wavelet filter coefficients.

$$y[n] = \sum_{m=-\infty}^{\infty} \left(\left(\sum_{m=1}^{A} (d[m,l]\psi_{ml}[n] \right) + r[A+1,l] \right). \tag{3.34}$$

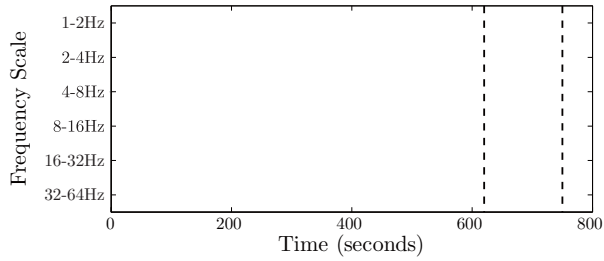
$d[m,l]$ are known as the *wavelet coefficients* for scales $a = 1, 2, ..2^A$, and $r[A+1,l]$ is the remainder of the signal representative of all scales larger than $a = 2^A$. The collection of $d[m,l]$ and $r[A+1,l]$ unambiguously describe $y[n]$, and $y[n]$ can be quickly and efficiently reconstructed exactly if all these coefficients are known [152, 151]. For the windowed time domain signals of length N the first summation reduces to $k+1 < l < k+N$. The D4 wavelet family shown in Figure 3.16(a) is orthogonal and can be used for the above analysis.

The values of $d[m,l]$ and $r[A+1,l]$ can be iteratively computed using cascaded filters. The process is shown and explained in Figure 3.17. This is simpler and faster than applying Equation 3.33 directly because the entire process, including calculation of wavelets and filters at each scale, can be performed by storing a few coefficients that describe the wavelet family. For the D4 family these coefficients are listed in Table 3.1. The interested reader can find detailed explanations of why this system works in [102] and [143]. Research into faster decimation also exists (see, for example, [199]).

Wavelet filters isolate spectral information in different frequency ranges



(a) Wavelet energy of EEG sequences



(b) Wavelet energy evolution

FIGURE 3.18: Example applications of wavelet filters. (a) shows the relative energy found in each frequency scale $k = 2 - 7$ for different states of EEG. The trends shown are the same as in Figure 3.15, demonstrating the power of wavelets to display similar information as PSD but more compactly. (b) shows wavelet decimation from Figure 3.16(d) applied to an EEG sequence. The seizure occurring at 620 seconds (between the dashed lines) is marked by an obvious change, signifying the applicability of wavelet theory to seizure recognition. Signals are sampled at $F_s = 512$Hz, thus 1 second is $N = 512$ samples. Tracking in (b) is performed using 2 second ($N = 1024$ samples) non-overlapping windows.

and can describe the EEG in similar ways as the PSD. A histogram of the relative average power at scales $m = 2$ to $m = 7$ for typical EEG sequences is shown in Figure 3.18(a). The frequency ranges for each scale correspond to those in Figure 3.16(b). A normalized measure is once more obtained by dividing by the total energy at all scales. The trends shown in this figure are the same as those observed in Section 3.3.2, but the representation is more compact. An example of relative energy of the different wavelet scales used to track changes over time is shown in Figure 3.18(b). Clear differences exist during the seizure.

3.3.4 Non-Linear Analysis

The statistics discussed so far are useful for extracting information from a signal $y[n]$ assumed to be either (1) generated by a linear system with state $\mathbf{z}[n]$ or (2) generated by non-linear dynamics that can be approximated by a very large dimensional linear system. Many signals that are generated by *non-linear* systems are treated as linear simply because the analysis is simpler and the theory is more developed. However even simple non-linear systems can lead to very large dimensional linear approximations, and non-linear analysis can be used to reveal information that may be missed. The hope is that by treating the system as non-linear a lower-dimensional (or simpler) representation of $y[n]$ exists.

Tools to analyze signals generated by a non-linear system with state $\mathbf{z}[n]$ are more useful now that fast computers are more readily available. Non-linear signal processing in the past few decades has developed significantly as a direct result of this advance in computer technology. Interest in the characterization of EEG from the perspective of non-linear statistics followed suit [70]: *If the EEG signal is generated by a non-linear system [60], can the non-linearity be used to the advantage in classification tasks?*

The problem is that non-linear statistics require significantly longer data sequences to be estimated correctly. This has to be traded against the hope that the dimension of the non-linear phase space is much smaller than the linear one. In an environment where a signal $y[n]$ is riddled with properties such as non-stationarity and noise non-linear analysis is difficult.

Nevertheless, non-linear dynamics are able to describe a much larger and richer range of behavior than linear ones. Attempting to understand an incredibly complex system such as the human brain requires analysis that reveals complex behavior, even if this analysis is limited.

Wherever possible linear analysis should be used to extract features from EEG data. The linear signal processing tools available are far more mature and better understood than non-linear ones. They generally require less data, making the assumed stationarity more likely.

However, linear analysis has been tried and has largely failed to reveal aspects of EEG dynamics including the identification of a pre-seizure state. In these cases non-linear analysis *may* be useful, *but results must be interpreted with care and limitations well understood.*

Assuming stationarity and the availability of sufficient data a challenge in estimating non-linear statistics is the presence of noise, discussed at the

beginning of Section 3.3. Noise can have considerable impact on the quality of computations. In some cases noise reduction techniques must be applied to the time series prior to computation [34].

This section presents some of the more commonly used methods in non-linear signal processing. These methods involve a two stage process in which the dynamics of a multidimensional $\mathbf{z}[n]$ must first be *reconstructed* from a single dimensional recording $y[n]$. This is known as *embedding* in the dynamical systems theory or *observation theory* in systems engineering.

Once the properties of $\mathbf{z}[n]$ have been reconstructed its non-linear features can be extracted. Three major categories of features exist: those that give an idea of how complicated a system is, known as *dimension* $\hat{N}$; those that give an idea of how predictable a system is, known as *Lyapunov exponents* λ_i; and those that give an idea of how random the system is, known as *entropy* H. Each of these statistics are discussed next. Further detail that that presented for both theoretical and practical aspects of this theory can be found in [2],[34],[82],[131],[159],[173] and [183].

3.3.4.1 Embedding Theory

When the equations that describe the dynamics of the system $\mathbf{z}[n]$ are known then it is possible to compute its features analytically. Often these equations are not available and statistics must be estimated using experimentally collected data. This signal, $y[n]$, is a single (or low) dimensional record that can be used to estimate statistics from a multidimensional system $\mathbf{z}[n]$. The properties of $\mathbf{z}[n]$ must be *reconstructed* from $y[n]$ to be able to capture the same dynamics of the underlying $\hat{N}$ dimensional system.

A fundamental idea in non-linear analysis (Takens/Aeyels's theorem, [4, 179]) is that this reconstruction can be performed by taking time delayed versions of $y[n]$, that is [179],

$$\hat{\mathbf{z}}[n] = (y[n], y[n+\tau], \cdots, y[n+(\hat{n}-1)\tau]). \tag{3.35}$$

τ is the time delay, typically chosen by minimum mutual information tests developed in [2], $\hat{n}$ is the *embedding dimension* and $\hat{\mathbf{z}}[n]$ is known as the *phase portrait* of $y[n]$ that spans an $\hat{n}$ dimensional *phase space*. This reconstruction can unambiguously capture the *co-ordinate invariant* dynamics of the original $\hat{N}$-dimensional system (given noise free data) if the embedding dimension $\hat{n} > 2\hat{N}+1$. In practice $\hat{n}$ can usually be much smaller [2], because typically the system state does not explore all of the phase space but only a (much) lower dimensional subset of this state space. The presence of noise in $y[n]$ limits but does not invalidate this method [179]. Note also that, as always, $y[n]$ should be normalized as in Equation 3.13 prior to embedding.

Formally, the reconstructed system $\hat{\mathbf{z}}[n]$ is related to the original system $\mathbf{z}[n]$ as

$$\mathbf{z}[n] = \Lambda(\hat{\mathbf{z}}[n]), \tag{3.36}$$

where Λ is some invertible map relating the original system $\mathbf{z}[n]$ to the reconstructed system $\hat{\mathbf{z}}[n]$. Equation 3.2 can be re-written as[8]

$$\begin{aligned}
\mathbf{z}[n+1] &= P(\mathbf{z}[n], \kappa[n], u[n], n) \\
\Lambda(\hat{\mathbf{z}}[n+1]) &= P\Lambda(\hat{\mathbf{z}}[n], \kappa[n], u[n], n) \\
\hat{\mathbf{z}}[n+1]) &= \Lambda^{-1}P\Lambda(\hat{\mathbf{z}}[n], \kappa[n], u[n], n) \\
\hat{\mathbf{z}}[n+1]) &= \hat{P}(\hat{\mathbf{z}}[n], \kappa[n], u[n], n),
\end{aligned}$$

where $\hat{P} = \Lambda^{-1}P\Lambda$ is the new map in embedded space replacing the original P.

The reconstructed system $\hat{\mathbf{z}}[n]$ does not look the same as the real $\mathbf{z}[n]$, but certain properties such as Lyapunov exponents, dimension and entropy (discussed later) remain the same because these do not depend on the coordinate system used. This reconstruction can and is used to estimate these parameters from a recorded signal.

The choice of $\hat{n}$ is not trivial. $\hat{n} \approx 5$ or 6 are typical values chosen for EEG analysis, usually justified by "good" results obtained for a particular application. These values are used because it is convenient for non-linear signal processing tools that $\mathbf{z}[n]$ be low dimensional (small $\hat{N}$). The dynamics of highly complex non-linear systems are unlikely to be captured by the EEG due to its low spatial resolution. Small $\hat{N}$ indicates that the EEG is capable of reproducing more of these dynamics.

However, *desiring low dimensionality does not make it so.* Proving the existence of low dimensional non-linearities is not easy, as shown in [145]. Whilst many papers ([10], [94], [145]) have concluded that there is sufficient indication of non-linear deterministic structure (as opposed to stochasticity) in the EEG, there has not been much evidence to support low dimensionality. In fact in Chapter 6 an incredibly simplified model is presented that in itself contains approximately 15 dimensions. This is quite something given that a simple 3-dimensional model of a neuron would suggest a brain with dimension $\hat{N} \approx 3 \times 10^9$, and hence embedding dimensions potentially in the order of 6×10^9! Tests for the presence of low dimensional dynamics on both intra-cranial and scalp EEG were performed in [10] and [26] for segments that were weakly stationary. The study showed that for intra-cranial EEG clear evidence of non-linearities exist, but that *strong indications of low dimensionality exist only during an epileptic seizure.* Scalp recordings indicated non-linearity but no evidence of underlying low dimensionality either in the seizure or non-seizure state. This suggests that either the dynamics captured at the scalp are of great complexity (more neurons contribute to the evoked potentials on the scalp) or that the characteristics of non-linearities are filtered by the skull and scalp and hidden by noise.

[8] The map Λ relates systems of very different dimensions, and as such is not as simple as implied by the mathematics presented here. However for illustration purposes only the argument can be simplified in this way.

Thus the choice of $\hat{n} \approx 5$ or 6 in EEG analysis is driven by necessity rather than correctness. Consider a moderately low resolution EEG that digitizes data to 10 bits. With 5 dimensions, phase space covers $(2^{10})^5 \approx 10^{15}$ possibilities. A typical weakly stationary 30 second EEG segment often used for analysis, sampled at 512Hz, produces data in the order of 10^5 samples – 10 orders of magnitude less than the size of phase space! One would hope that the amount of *relevant* phase space is sufficiently small to be describable by such small volumes of data, although these assumptions are only justifiable for special cases like during epileptic seizures in which the waveforms are simpler.

Even so a significant volume of literature utilizes such values with little or no admission to its shortcomings. *The implicit assumption is that embedding dimensions of such low magnitude are able to capture the* correct *dynamics rather than all the dynamics of* $\hat{\mathbf{z}}[n]$, *and consequently* $\boldsymbol{z}[n]$. This is not necessarily true, and the reader is encouraged to question the validity of results obtained through the use of inadequate parametrization.

Another concern is that of computation time. For an EEG signal $x_c[n]$ where $c = 1, 2..C_{TOT}$ channels are available and therefore C_{TOT} delay reconstructions are possible, the effective dimension of the overall phase space is $\hat{n} \times C_{TOT}$, even though some redundancy exists. For computational efficiency, there is a trade-off between C_{TOT} and $\hat{n}$. In some works, such as that in [37], 40 standard PCs were necessary to allow real time computation of most of the algorithms presented here for $C_{TOT} = 128$. This is not always (or usually!) feasible, so a subset of channels ($C_{TOT} = 4$ or 5) at appropriate sites are selected for analysis.

The following sections assume that the choice of $\hat{n}$ and τ and C_{TOT} in the analysis of an arbitrary signal $y[n]$ used to reconstruct a system $\hat{\mathbf{z}}[n]$ is adequate for the particular application.

3.3.4.2 Dimension – How Complex is a System?

The *dimension* of a system, described in Section 3.1, is a measure of the complexity in the dynamics of $\mathbf{z}[n]$. Dimension in Euclidean space is familiar and intuitive to most – it is the minimum number of coordinates required to unambiguously describe the location of a point in phase space $\mathbb{R}^{\hat{N}}$. For example the geographical location of a ball bouncing on a hill anywhere on the earth's surface can be defined by 3 variables – latitude, longitude, and height above sea level. In this case $\hat{N} = 3$. The dimension of a dynamical system can also be the number of variables necessary to describe its behavior [131], which in the case of our bouncing ball will be greater than 3 because it is necessary to know the velocity of the ball (in each of the 3 spatial dimensions) and aspects of the ball that define its movement and bounce (such as elasticity, volume, air pressure and temperature of the gas inside it, etc). Other definitions of dimension exist but in all cases an estimate of the dimension of a system provides a feel for its complexity. The dimension of a non-linear system, regardless of definition, shall from now on be referred to as $\hat{N}$.

The dimension $\hat{N}$ of non-linear systems can be a non-integer. The idea is that the actual dimension of a system is always an integer, but the effective phase space where the orbits lie most of the time may be a small volume of the entire space. To capture this non-integer *fractal dimensions* have been introduced. Several different types of fractal dimensions have been defined, including information dimension, capacity dimension and correlation dimension [131] of which only the latter is discussed here.

The correlation dimension computed from the re-constructed state variable $\hat{\mathbf{z}}[n]$ derived from a time domain signal $y[n]$ of length $n = 1, 2..N$, using embedding dimension $\hat{n}$, is given by:

$$\hat{N} = D_2(\hat{n}, \epsilon) = \lim_{n\to\infty} \lim_{r\to 0} \frac{\partial \ln C(\hat{n}, \epsilon)}{\partial \ln \epsilon}, \tag{3.37}$$

where $C(\hat{n}, \epsilon)$ is the correlation sum [10]

$$C(\hat{n}, \epsilon) = \frac{2}{N(N-1)} \sum_{n_i<n_j} \Theta(\epsilon - |\hat{\mathbf{z}}[n_i] - \hat{\mathbf{z}}[n_j]|). \tag{3.38}$$

$\Theta(y)$ is a sign function with a value of 1 if $y > 0$ and zero otherwise. The correlation sum gives the empirical probability that any two instances of $\hat{\mathbf{z}}[n]$ lie within a distance ϵ of each other [106]. For the limits in Equation 3.37 to exist ergodicity is required, that is, it is necessary that a long enough observation of the signal will allow us to infer the system's properties.

Ideally the calculation of Equation 3.37 requires an infinite amount of data that must be stationary. To track changes over time as well as to deal with the problems presented by the non-stationarities in $\hat{\mathbf{z}}[n]$, only finite (and not very long) lengths of data can be used. In such cases an *effective correlation dimension* $D_2^{eff}(\hat{n}, \epsilon)$ can be estimated for a limited range of ϵ where $D_2(\hat{n}, \epsilon)$ is found. That is,

$$D_2^{eff}(\hat{n}, \epsilon) = \frac{1}{N_\epsilon} \sum_{\epsilon=\epsilon_{lower}}^{\epsilon_{upper}} \frac{\partial \ln C(\hat{n}, \epsilon)}{\partial \ln \epsilon}, \tag{3.39}$$

where N_ϵ is the number of ϵ values in the interval $[\epsilon_{lower}, \epsilon_{upper}]$. Details on appropriate selection of ϵ_{lower} and ϵ_{upper} can be found in both [10] and [90].

Figure 3.19 shows the correlation sum calculated for awake, sleep and epileptic EEG, using embedding delay $\tau = 30$ samples (about 0.06 seconds at 512Hz sampling) and embedding dimension $\hat{n} = 4$. At least 30 seconds of data is used for each computation, meaning that at worst there are in the order of 15,000 samples available for computation. The effective correlation dimension $D_2^{eff}(\hat{n}, \epsilon)$ is calculated as the gradient of $C(\hat{n}, \epsilon)$ over which it follows a rough straight line. Here both the awake and asleep states show roughly equal complexity whilst the epileptic seizure is lower dimensional, as expected, because its slope is shallower.

Notice also the saturation effects, best observed for the 'normal awake'

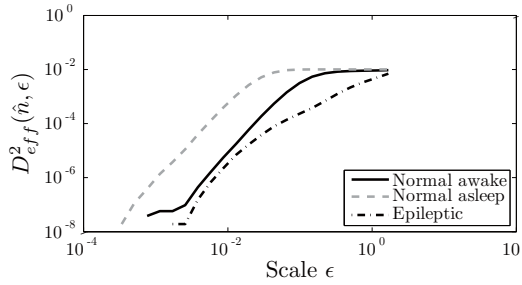


FIGURE 3.19: Estimated correlation integral for normal, sleep and seizure EEG. The computed correlation integral is a relative measure that should not be interpreted as an absolute value. The gradient of seizure curve is smaller than for sleep and awake EEG, indicating a smaller $D_2^{eff}(\hat{n}, \epsilon)$, and thus a smaller dimension. The dimensions of sleep and awake EEG are roughly similar to one another. $\tau = 0.06$ seconds is used, equivalent to $\tau = 30$ samples when $F_s = 512$Hz, with $\hat{n} = 4$. For each estimate as much data was used as is available, with at least 30 seconds of data for each sequence.

state. For small values of ϵ there is insufficient data and no points can be found that are closer to each other than this distance. The correlation integral approaches zero. For large ϵ all data can be found to lie within this distance of each other, and the correlation integral saturates.

3.3.4.3 Lyapunov Exponents – How Predictable is a System?

The predictability (or deterministic structure) of a system $\mathbf{z}[n]$ can be mathematically described by its *Lyapunov exponents*. These are statistics that give an idea of how much the system changes when a small perturbation or change is introduced [131, 159, 183]. For example, return to the bouncing ball example but now assume that it has come to a stop on a smooth flat surface. Pushing the ball softly moves the ball, but it will not go far from its resting point and it is easy to tell where it will end up. Its behavior is *predictable* and the Lyapunov exponents of this system are small. However, if the ball is resting on top of a narrow ledge any small force in the wrong direction can result in the ball falling off the edge. It could end up anywhere depending on the topography of the descent. This behavior is *less predictable* and its Lyapunov exponents are larger.

An $\hat{N}$ dimensional system has $\hat{N}$ Lyapunov exponents, denoted λ_i with $i = 1, 2..\hat{N}$. Each λ_i gives an idea of the amount of contraction ($\lambda_i < 0$) or expansion ($\lambda_i > 0$) of the system in the direction of the corresponding dimension [34, 82, 131, 183]. However since random perturbations are expected to cause the most change in direction of the largest Lyapunov exponent, $\lambda_{\max} = \max(\lambda_i)$ is sufficient to reliably and reproducibly characterize

the *predictability* in the dynamics of a system [72]. *It is the most un-predictable behavior that limits the predictability of a system.*

Again due to the limitations imposed by finite time signals and non-stationarities a short term estimate of $\lambda_{\max}$ is used [72, 71]. One example algorithm that estimates the largest Lyapunov exponent from a reconstructed system $\hat{\mathbf{z}}[n]$ assumes that $\lambda_{\max}$ can be computed by testing the exponential divergence of trajectories. Suppose we select an arbitrary time point n_0. We look at trajectories of length Δn, and compare the divergence between a trajectory starting at $\hat{\mathbf{z}}[n_0]$ and one starting at $\hat{\mathbf{z}}[n]$. Averaging over all such trajectories where $\hat{\mathbf{z}}[n]$ is ϵ close to $\hat{\mathbf{z}}[n_0]$ gives an estimate of the Lyapunov exponent. This is accomplished through the following computation [81]

$$\lambda_{\max}(\hat{n}, \epsilon, \Delta n) = \frac{1}{T} \sum_{n_0=1}^{T} \ln \left(\frac{1}{N_{n_0}} \sum_{|\hat{\mathbf{z}}[n]-\hat{\mathbf{z}}[n_0]| \leq \epsilon)} |\hat{\mathbf{z}}[n_0 + \Delta n] - \hat{\mathbf{z}}[n + \Delta n]| \right). \tag{3.40}$$

In the above, the difference between where $\hat{\mathbf{z}}[n_0]$ ends up at time $n_0 + \Delta n$ and where $\hat{\mathbf{z}}[n]$ ends up at the same time is measured. The average is taken over N_{n_0} points for values of n in which $\hat{\mathbf{z}}[n]$ is within ϵ distance of $\hat{\mathbf{z}}[n_0]$. The process is iterated over a large number of trials T, where a different starting point n_0 is selected for each.

$\lambda_{\max}(\hat{n}, \epsilon)$ is computable if there is a range of Δn over which $\lambda_{\max}(\hat{n}, \epsilon, \Delta n)$ remains constant. In theory this constant should remain the same for all Δn, but this does not occur in practice due to data record limitations. If all nearby trajectories end up in roughly similar positions, then $\lambda_{\max}(\hat{n}, \epsilon)$ is small and the system is predictable. Otherwise $\lambda_{\max}(\hat{n}, \epsilon)$ is large and the system is less predictable.

The two averages in Equation 3.40 normalize results by the number of trajectories N_{n_0} and the number of trials T. Averages result in better estimates in noisy data. Amplitude normalization for different systems must be applied to $y[n]$ prior to computation.

Lyapunov exponents computed for awake, sleep and epileptic EEG are shown in Figure 3.20, using $\tau = 30$ samples (about 0.06 seconds), embedding dimension $\hat{n} = 4$ and $\epsilon \approx 1\%$ of the signal amplitude. The estimations are roughly equal over the different Δn, although as is common in this type of analysis this convergence is very rough and averages must be taken over Δn. Relative to each other the seizure EEG is more predictable than normal or asleep EEG because its maximum Lyapunov exponent is closer to zero. Sleep and awake EEG seem to display similar predictability.

3.3.4.4 Entropy – How Random is the System?

Entropy is a measure first used in thermodynamics to give an idea (at the macro-scopic level) of how disorderly a system is. It is one of the most abused technical terms in the world, largely because so many interpretations exist.

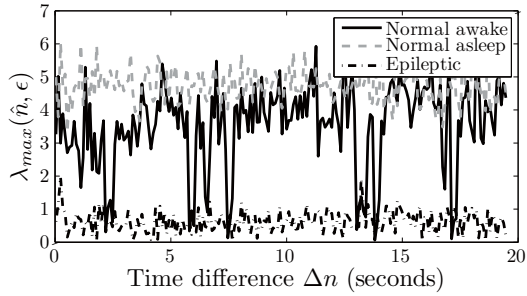
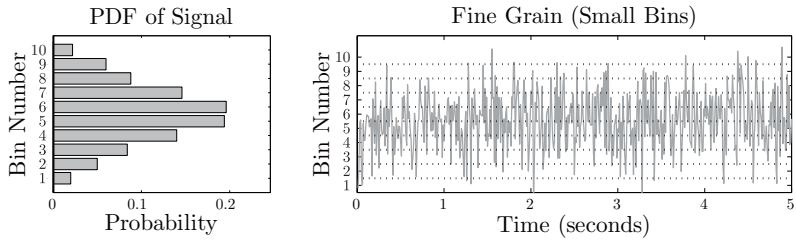


FIGURE 3.20: Estimated maximum Lyapunov exponent for normal, sleep and seizure EEG. The computed statistic is a relative measure that should not be interpreted as an absolute value. Notice the high level of variability observed in each case – this is typical in the estimation of Lyapunov exponents. However on average a flat line can be drawn for each. This line is much closer to zero for the seizure EEG, indicating that it is more predictable than sleep or awake EEG. $\tau = 0.06$ seconds is used, equivalent to $\tau = 30$ samples when $F_s = 512$Hz, with $\hat{n} = 4$ and ϵ of 1% of the signal amplitude, normalised within each window. For each estimate as much data were used as is available, with at least 30 seconds of data for each sequence.

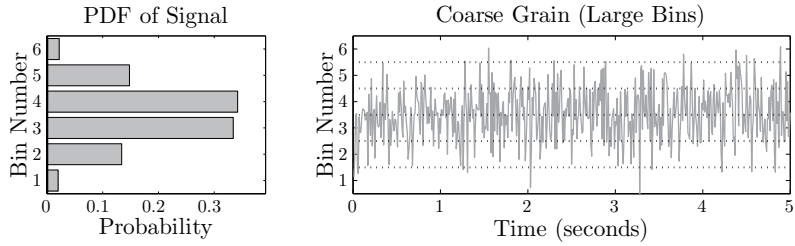
Entropy is a measure of *randomness* because it quantifies the amount of disorder in a system. It is also a measure of *information* because the level of randomness indicates how much information can be transferred in a single measurement taken from the system. It is a measure of *compressibility* because the less information carried in a signal, the more efficient that information can be communicated.

In general all entropies give an idea of, at least qualitatively, the same concept – the amount of order or disorder in a system. The more random something is, the more information we need to capture its essence. Let us return to the bouncing ball example, again at a stop. Unless something unforeseen happens the entire future of the ball is known: it will stay in the same place. It is a completely deterministic system and only one measurement (its position) is necessary to know everything about the future. Its entropy is *low* (zero). Now imagine that the ball is bouncing, but the direction that it bounces in each time is random and it is impossible to tell where it will go next. Measurements must be taken at all times because nothing can be said about the future of the ball. This random system has *high* entropy.

Let us first define a general entropy of a discrete probability distribution $p(k) \in [0, 1]$, where k counts the number of different possibilities, and $p(k)$ denotes the probability of event k occurring:



(a) Fine grain (small bins



(b) Coarse grain (large bins)

FIGURE 3.21: Coarse-graining in the computation of entropy. In (a) $K = 10$ bins are used to divide $y[n]$. This is *fine*-graining relative to (b), where only $K = 6$ bins are defined. The distance (in amplitude) between bins is ϵ. On the left of each graph is the PDF of the signal, that is, the probability $p(k)$ for each bin is shown. Notice that the PDF depends on the length of the window used. If only 0.5 seconds are used to compute it (as opposed to the full 5 seconds) then the PDF between 0-0.5 seconds will look different from that between 1-1.5 seconds. Long enough time intervals must be used so that the computed PDF remains stationary. This is especially important when fine graining is used.

$$H_q(\hat{n}, \epsilon) = \begin{cases} -\sum_{k=1}^{K} p(k) \log_2 p(k) & \text{if } q = 1 \\ \\ \frac{1}{1-q} \log_2 \sum_{k=1}^{K} (p(k))^q & \text{if } q > 1 \end{cases} . \tag{3.41}$$

A logarithm base 2 is used so that the entropy is defined in bits.

This entropy can be used in dynamical systems as follows. An arbitrary signal $y[n]$ is *quantized* into K symbols (that is, K possible measurements or outcome events), where the number of bins $k = 1..K$ is known as the *alphabet* and represents all the possible symbols in the quantized signal. For example if $y[n]$ is always between 0 and 1, $K = 5$ if the $[0 - 1]$ range is divided into 5 regions. This is also known as *coarse-graining* of $y[n]$ into ϵ-sized bins, and

in this case $\epsilon = 0.2$. The probability $p(k)$ of each symbol k occurring may be defined as

$$p(k) = \frac{\text{Time that } y[n] \text{ is in bin } k}{\text{Total length of } y[n] \ (N)}. \tag{3.42}$$

An example of coarse-graining for $K = 6$ and $K = 10$ is shown in Figure 3.21. The probability distribution $p(k)$ is also shown. For very fine graining it is important that sufficiently long signals be used to compute these probabilities. For example, if only 0.5 seconds are used to compute $p(k)$ then the PDF would look different at different times on the 5 seconds of data shown.

Coarse graining has the advantage of making the computation of entropy less susceptible to noise than estimates of Lyapunov exponents or correlation dimension. This comes at a cost – the difficulties that are avoided are instead shifted to the selection of an appropriate ϵ and alphabet scheme. Computation of $H_q(\hat{n}, \epsilon)$ in theory requires the supremum to be taken over all possible quantization schemes because quantization implies that entropy is *not* a coordinate independent system. The limit of $\epsilon \to 0$ is necessary, but not feasible for finite $y[n]$.

Several interpretations of Equation 3.41 exist. When embedding dimension $\hat{n} = 1$ (i.e., no embedding) and $q > 1$ it is known as the *block entropy*. Each integer q defines a different family of entropy. The special case $H_1(\hat{n} = 1, \epsilon)$ when $q = 1$ is known as the *Shannon Entropy* and it is the definition classically used in information theory to determine capacity of communication channels. For example, suppose a signal $y[n] \in (0, 1)$ is divided into $K = 8$ bins. Assuming all symbols in the alphabet are equally likely, that is, that Equation 3.42 yields $p(k) = \frac{1}{8}$ for $k = 1, 2 \cdots 8$, then

$$\begin{aligned} H_1(\hat{n} = 1, \frac{1}{8}) &= -\sum_{k=1}^{K=8} \frac{1}{8} \log_2 \frac{1}{8} && (3.43) \\ &= \frac{1}{8} \cdot 3 \cdot 8 && (3.44) \\ &= 3 \text{ bits.} && (3.45) \end{aligned}$$

This result tells that that in general we need 3 bits to represent the quantized $y[n]$ accurately.

The *Komolgorov-Sinai* entropy is an interpretation of Equation 3.41 that explores *transition probabilities* between measurements, computing the block entropy for a *sequence* of symbols. It is analogous to using embedding as in Equation 3.35 and corresponds to values of $\hat{n} > 1$ because $\hat{\mathbf{z}}[n]$ is an $\hat{n}$ dimensional object looking at a block of $\hat{n}$ past observations. The longer the sequences used to compute Komolgorov-Sinai, the higher the $\hat{n}$ that is used for reconstructing $\hat{\mathbf{z}}[n]$.

All the usual problems resulting from insufficient data for high dimensions and the selection of appropriate time delays apply. For a reconstructed system

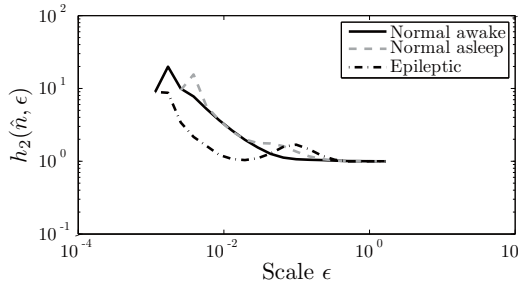


FIGURE 3.22: Komolgorov-Sinai entropy for normal, sleep and seizure EEG. The computed entropy is a relative measure that should not be interpreted as an absolute value. It is lower for seizure than for sleep or awake EEG over the range for which the correlation integral is valid, implying that seizures are more ordered than the other two. Sleep and awake EEG have a similar level of order. $\tau = 0.06$ seconds is used, equivalent to $\tau = 30$ samples when $F_s = 512$Hz, with $\hat{n} = 4$. For each estimate as much data was used as is available, with at least 30 seconds of data for each sequence.

with $\hat{n}$ dimensions and K symbols in the alphabet of each dimension, the total number of bins is $K^{\hat{n}}$. *As the dimension increases the required number of data points to estimate $H_q(\epsilon)$ to any degree of accuracy grows exponentially.*

For $q = 2$ the Komolgorov-Sinai entropy $h_2(\hat{n}, \epsilon)$ is related to the correlation sum $C(\hat{n}, \epsilon)$ defined in Equation 3.38 by

$$h_2(\hat{n}, \epsilon) = H_2(\hat{n} + 1, \epsilon) - H_2(\hat{n}, \epsilon) = \ln\left(\frac{C(\hat{n}, \epsilon)}{C(\hat{n} + 1, \epsilon)}\right). \tag{3.46}$$

$h_2(\hat{n}, \epsilon)$ is only valid for values of ϵ for which the gradient of $C(\hat{n}, \epsilon)$ is constant. Because $h_2(\hat{n}, \epsilon)$ is only valid for limiting behavior of $\hat{n}$ then this type of entropy requires even more data than Lyapunov exponents or correlation dimensions to compute [81].

Shannon entropy $H_1(\hat{n}, \epsilon)$ can also be computed in $\hat{n} > 1$ dimensions. The advantage with $\hat{n} = 1$ is that computation is fast and easy. In higher dimensions the data requirements again become prohibitive but computation remains simpler than Equation 3.46. Shannon entropy is usually expressed with a logarithm base 2 so that results can be interpreted in bits per second, a property that is relevant to information and compression theory for which it is most often used.

The Shannon entropy is computed using a simple box-counting method, in which phase space is partitioned into ϵ sized bins or boxes and the probabilities are calculated by counting the number of data points in each bin. This is the same as the computation of $p(k)$ in Figure 3.21 but in multiple dimensions. To avoid the dependence of quantization on scale and magnitude a signal is

first normalized to the unit variance as in Equation 3.13. In practice the bins are often allowed to overlap so that edge effects can be reduced.

Figure 3.22 computes $h_2(\hat{n}, \epsilon)$ for the typical states, using $\tau = 30$ samples (about 0.06 seconds with 512Hz sampling) and $\hat{n} = 4$. Entropy of epileptic seizures (for the values of ϵ that are valid) is lower than either asleep or awake EEG, supporting the notion that there is less randomness in the system. A more detailed exploration of Shannon entropy can be found in Chapter 5.

The *non-linear* statistics discussed in this section provide information about the non-linear system $\mathbf{z}[n]$ that generates the signal $y[n]$. The co-ordinate invariant features of $\mathbf{z}[n]$ can be approximated by using the Takens/Aeyels theorem to *reconstruct* the system $\hat{\mathbf{z}}[n]$ from $y[n]$, as described by Equation 3.35. The co-ordinate invariant non-linear statistics can quantify the

- *Complexity*: The *dimension* $\hat{N}$ is a measure of how many degrees of freedom there are in a system. The correlation dimension is estimated as described by Equation 3.39.
- *Predictability*: The maximum *Lyapunov exponents* λ_{max} describe how predictable the system is by measuring the divergence of trajectories. It is estimated as described by Equation 3.40.
- *Randomness*: The *entropy* H quantifies the randomness of a system. Several entropy definitions exist. A generic definition can be found in Equation 3.41.

All methods require long data sequences and stationarity (as well as a much stronger form of ergodicity) in $y[n]$ to be estimated correctly. Their importance and their limitations in detecting and predicting epileptic seizures are involved, and are discussed separately next.

3.3.4.5 Non-Linear Dynamics and Analysis of the Epileptic EEG

The intrinsic non-linear properties of a system – dimension $\hat{N}$, Lyapunov exponents λ_i and entropy H – all give ideas of very similar properties. For example the choice of ϵ in each case is the partitioning of phase space with dimension $\hat{n}$ into boxes of size ϵ.

Using the box counting analogy, the correlation dimension measures how many points are found in nearby boxes; Lyapunov

exponents measure how far away from each other points starting in the same box end up; and entropy measures how likely it is that a point may be found in any one box. All give ideas of *predictability* or *randomness* of a system. *Correlation dimension, Lyapunov exponents and entropy interpret complexity by quantifying different ways in which a dynamical system evolves.*

Predictability and randomness are in principle very well suited to characterize epileptic seizures because its EEG is best described as a more ordered signal than that of normal activity. During a seizure a single channel becomes more oscillatory and therefore simpler. Many channels viewed together become less complex because they tend to resemble each other more. Hence both temporally and spatially there is reduced complexity, less randomness and more predictability. It is not surprising then that non-linear analysis has been widely used to analyze the epileptic EEG[9].

Dimension $\hat{N}$, Lyapunov exponents λ_i and entropy H must be estimated using an intermediate system $\hat{\mathbf{z}}[n]$ *reconstructed* from the measured signal $y[n]$. In EEG analysis $y[n] = x_c[n]$ is measured over a finite, non-stationary and noisy time-window. At the risk of repeating an important fact too often, weak stationarity may only be assumed in 20-30 seconds of recorded EEG, and this is typically insufficient to reliably compute any of these statistics[10]. Procedures by which an approximation rather than exact values of these statistics can be calculated in this environment were outlined in previous sections – effective correlation dimension $D_2^{eff}(\hat{n}, \epsilon)$, maximum Lyapunov exponent $\lambda_{\max}(\hat{n}, \epsilon)$ and entropy $H_q(\hat{n}, \epsilon)$.

The computation of both $D_2(\hat{n}, \epsilon)$ and $\lambda_{\max}(\hat{n}, \epsilon)$ is highly sensitive to noise, showing no regions of convergence when even 2-3% errors are present in the data [81]. Analysis on noisy data can sometimes be improved by determining the validity of the estimate through the rejection of a null hypothesis. These tests can in general only be applied off-line. For this reason $D_2^{eff}(\hat{n}, \epsilon)$ and $\lambda_{\max}(\hat{n}, \epsilon)$ are often estimated only from intra-cranial EEG that is less noisy than scalp EEG. Entropy and its variants are less susceptible to noise and have been applied more extensively.

The computational cost of non-linear analysis is high. Inefficiency in algorithms is an expense afforded only to problems in which no other solution is

[9]Unfortunately as is often the case when new methods become popular, much of the literature seems more a case of applying them to whatever data are there rather than to show an understanding of a complex problem. Readers should be warned that a certain degree of skepticism is needed when digging through the vast literature printed on the topic - often research is tested on very little data, and neither justification nor insight is provided.

[10]30 seconds of data at 512Hz sampling gives $15,360$ samples. With $\hat{n} = 4$ and $K = 10$ there are $10^4 = 10,000$ bins. With $\hat{n} = 6$ there are $10^6 = 1,000,000$ bins. Only if embedding dimension is low does using 30 second windows make any sense. However a low $\hat{n}$ will be an inadequate representation of the underlying system (unless there is a real collapse of dimension, as may be the case during epilepsy).

obvious. For seizure detection that requires fine temporal detail on-line algorithms based on non-linear measures are impractical. Non-linear analysis for detection of seizures has been explorative rather than conclusive. For example, efforts are made in [33] to determine how suitable H is for discriminating two states, and [129] is a comparison of the relative merits of linear versus non-linear methods for EEG analysis. Neither is specific to epilepsy.

3.4 Detection and Prediction of Seizures in Literature

We have now seen a good sample of the different methods available to analyze EEG data, and we have shown that at least in some cases these are able to characterize the difference between seizure and non-seizure activity. This section then looks, briefly, at how the different feature extraction methods have been used by seizure detectors and predictors as presented in the literature. It is assumed here that $y[n] = x_c[n]$.

Let us first look at epileptic seizure detection in the time domain. The statistics that deal with magnitudes and variance ($\mu_y[k], \sigma_y[k], c_y[k]$ and $v_y[k]$ presented in Section 3.3.1.1) are often only used as a second stage statistic because by themselves they are not sufficient to reliably characterize seizure activity. First a feature is extracted from $x_c[n]$ and its mean, variance, COV and total variation are computed over time. Some examples of detectors that use these methods include [55], [50], [165], [170]. There are many many more. In [186] and [187] it was proposed that variance and total variation can be better discriminators of seizure activity than mean because these can detect how regularity changes over time. This is discussed in more detail in Chapter 5.

Changes in periodicity and synchronization, Section 3.3.1.2 and Section 3.3.1.3 respectively, have been described in Chapter 1 as a key observable during many types of seizures. As such they are suitable features to extract for the detection of epilepsy. However although figures such as Figure 3.8(c) and Figure 3.9(f) show promise, perfect periodic activity in a single channel or perfect synchronization between channels is rarely observed, making the computation of $\text{CORR}_y[\tau, k]$ and $\text{XCORR}_{y_1,y_2}[\tau, k]$ unreliable. Again as stand alone statistics they have failed as seizure detectors. Furthermore these measures of synchrony are often not useful in intra-cranial analysis because EEG electrodes are close together and display high levels of synchrony even outside a seizure [75]. Nevertheless measures of periodicity have been used in detection strategies in texts such as [97] and [117], and measures of synchronization have been incorporated in [74] and [75].

Spectral characteristics of an EEG are important for state identification and it is not surprising that FFT based methods and its variants, described in Section 3.3.2, form part of many epileptic seizure detection systems. Recall

that for the most part exactly the same information is contained in the time domain as in the frequency domain, but representations in the frequency domain allow us to more easily identify features key to the detection of epileptic seizures.

For example a series of papers ([8], [45], [135] and [178]) explore the frequency content of seizures estimated using methods described here as well as auto-regressive (AR) based models. They reason that exploiting a combination of frequency and amplitude features to detect seizures is not new and virtually all methods capitalize on this notion. It is therefore important to know the features that have the most discriminating power.

Several attempts in which *spectral signatures* of epileptic seizures are extracted in order to detect seizures exist [163]. These involve the identification of stereotyped patterns in the evolution of frequency content that consistently present themselves before or during a seizure. However these signatures are patient dependent, cannot always be identified and their applicability on a larger scale is questionable.

Time-frequency methods such as wavelet filters and its variants are designed to cope with non-stationarities and they are today the method of choice to extract "frequency" information from the EEG. Classification using wavelets has been applied to recognize seizures as well as other pathologies including schizophrenia and obsessive compulsive behavior (OCD) [65]. DB4 is the most common choice of wavelet because it provides a smooth enough frequency filtering to appropriately characterize the EEG, whilst remaining computationally efficient [176].

Wavelet based feature extraction specific to seizure detection goes as far back as 1994 (in [162]) where their potential for the classification of epilepsy was first explored. Much of the early focus lay on identifying the frequencies that are important. It was found that in general higher frequencies (> 50Hz) corresponding to lower scales a can be ignored because of measurement noise, intermediate values of a can be used to detect the presence of artifact, whilst the ratios of higher scales can be used to distinguish between the normal and the seizure states.

Other time-frequency methods applied for the detection of epileptic seizures exist. In particular, the *Gabor atom density* derived from the original *matching pursuit algorithm* described in [41], [77], and [193] has claimed success in the detection of unspecific epileptic seizures. Gabor atoms are able to resolve different time and frequency scales in much the same way as wavelets, where a mother function is now translated and modulated at different frequencies. Gabor atoms are particularly tuned to pick out rhythmic activity, thus their ability to isolate the resonant frequencies of epileptic seizures.

Non-linear tools are thought to be too data intensive and computationally taxing to be used in the problem of seizure detection. However some work exists, for example in [198] where non-linear tools that require less data are developed for detection of seizures by first projecting signals into a linear space. This destroys the non-linear characteristics of EEG, but emphasizes regular-

ity. However the computational disadvantages of non-linear analysis remain, whilst providing no more information than linear measures could extract.

Seizure prediction has been attempted with much less success than detection. Prediction using accumulated energy time-domain measures (more sophisticated versions of the theory presented in Section 3.3.1.1) failed to identify the pre-seizure state because sleep and post-seizure EEG, which are high in energy, interfere with results. Energy based measures of this nature cannot be used to predict seizures in a stand-alone manner, but may be used as complementary features to other algorithms (see [39] for details).

Some believe that synchrony is key in prediction. Mormann et al [112, 113] and Chavez et al [28] rely on capturing the synchronicity between different channels of the EEG to predict the onset of seizures. Their work is based on calculations that use both Equations 3.19 and 3.24. The authors have demonstrated marked *decreases* in synchronicity long before an epileptic seizure, regardless of the type of test [112, 113]. This de-synchronization is explained by the slow entrainment of neurons into a seizure, where at the start of a seizure only the focus is involved and thereby its activity varies from surrounding sites. Under very limiting conditions it was found that correct prediction occurred for 80% of cases, and more if restricted to temporal lobe epilepsies, with de-synchronization occurring from 4 to 221 minutes before a seizure. In accordance with previous observations maximum synchronization is observed at the onset of a seizure. The examination of the spatial patterns of entrainment to aid optimisation of channel selection as well as to characterise the route to an epileptic seizure was introduced in [28] by applying the theory to narrowband signals.

Nevertheless little independent evidence exists that this algorithm is applicable to more than specific cases. Furthermore, it is implied in the calculation that the entrainment into seizure is slow – this is not true for many patients. However as a start to the difficult task of seizure prediction the ideas presented in this work merit reflection.

Spectral features, like in detection, form at least part of many published predictors of seizures. For example spectral signatures are used in [117]. The idea is that for many people the route to epilepsy is consistent and pattern recognition methods can be used to detect the precursors. However as in detection these signatures are patient dependent, cannot always be identified and their applicability on a larger scale is questionable. For those patients for whom this works it should be exploited as much as possible, but recognizing patterns does not infer anything about underlying dynamics, leading to missed events in cases where the evolution into a seizure is not typical.

The importance for non-linear tools described in Section 3.3.4 is found in prediction rather than detection of seizures. These methods assume the gradual entrainment of neurons into a seizure, slowly resulting in an EEG that is more ordered and less random. $D_2^{eff}(\hat{n}, \epsilon)$, $\lambda_{\max}(\hat{n}, \epsilon)$ and $H_q(\hat{n}, \epsilon)$ are expected to slowly decrease prior to a seizure, reaching their minimum immediately prior to the event. This assumption has been extensively explored.

$D_2^{eff}(\hat{n}, \epsilon)$ is used as a predictive measure in some of the literature, such as [90], where it is shown that the expected transitions to low dimensionality occur prior to a seizure. This was achieved by treating $D_2^{eff}(\hat{n}, \epsilon)$ as a non-absolute and rather informal definition of dimension. Predictive times ranged between 4 and 25 minutes. Higher temporal resolution can be achieved by the use of a more computationally efficient and adaptive algorithm proposed in [119].

Maximum Lyapunov exponents are explored in [71]. They notice a marked decrease in $\lambda_{\max}(\hat{n}, \epsilon)$ beginning over 2 hours prior to seizure onset, and a sharp post-seizure increase. Their algorithm adaptively selects the best channels to use for computation. Physiological significance has been implied in [72], where it was shown that pre-seizure entrainment of critical sites lasts significantly longer than post-seizure disentrainment. This suggests that seizures may be mechanisms by which the brain is able to reset itself after a gradual transition toward pathological behavior. If this is so, pharmacological (or other forms of) intervention with the goal of resetting the brain once entrainment begins may alleviate the need for a seizure to occur.

These types of statistics showed some promise at first but recent tests on simulated and real EEG data sets indicate that both *Lyapunov exponents and correlation dimension are unsuitable to predict epileptic seizures* [62, 88].

In some literature ([3, 152, 151]) the computation of entropy, Lyapunov exponents and correlation dimension is estimated from different wavelet frequency bands rather than different embedding dimensions, on the assumption that the dynamics could be disjoint at different scales. Insufficient testing makes results inconclusive and patterns common to many patients are not found.

In all cases of published predictors the validity of the algorithms is largely questionable. Data sets are often small, results are reproducible only for certain patients and the effects of state of vigilance and artifact are not explored.

Virtually all published detectors and many predictors of epileptic seizures use differences in frequency content as part of their feature set. Time-domain linear methods have typically been used as a supplement to other features. To capitalize on the advantage of both, time-frequency analysis such as wavelet analysis is becoming more popular.

In the case of prediction the focus is to use non-linear rather than linear tools. However after much work little success has ensued and it appears that prediction is a problem that will need patient specific methods, and most likely will have to involve other measurements in addition to the EEG.

Prediction is revisited in Chapter 7.

3.5 Conclusions

Designing a classifier for a complex system such as the epileptic EEG involves several stages of signal processing. This chapter focuses on the preprocessing and feature extraction aspects of the detection process shown in Figure 4.1.

> The selection of a feature set that can discriminate between different states is imperative to the performance of a classification system. For the epileptic EEG the features must differentiate between non-seizure, pre-seizure and seizure. The extraction method must target those aspects that make seizures different from normal brain activity. Each feature must add at least *some* new information because features that target exactly the same information do not add any discriminating power, and make the classifier more expensive computationally.
>
> The signal processing tools employed to extract these features must cope with *noise*, *non-stationarity* and possible *non-linearities* in the EEG.
>
> It is necessary that the signal be *normalized* so that comparisons of the extracted features can be made between different times, different electrodes and different people. Normalizing a signal entails removing its mean and making its features scale-invariant.

In this chapter we surveyed signal processing tools that can extract features from any signal, but that are particularly well suited to the study of the EEG. The collection of methods presented and the references to their literature is not exhaustive, but aims to be representative of what is available and has been tried over the last few decades.

However apt the features, a detection or prediction system cannot work without a suitable classifier or the combination of the information in an expert way. These are discussed next.

4

Classifying the EEG

In Figure 4.1 a complete detection system is divided into four tasks: preprocessing, feature extraction, decision making and expert system. The first two have been discussed in detail in Chapter 3. Here we focus on the rest, again with specific interest on detection and prediction of epileptic seizures. We assume that we have extracted a set of S features $(\zeta_1,\ \zeta_2, \cdots, \zeta_S)$ and we assume that these are suitable to distinguish between non-seizure, seizure and in the case of prediction pre-seizure EEG. We must now make good use of this information to *decide* in an *expert* manner to which of these classes the features belong.

A *classifier* makes decisions based on information taken from single channels, multiple channels or combinations of these. The classifier can be as simple as imposing a threshold on features, or can use more sophisticated machine learning algorithms. It must first be *trained* on how to make decisions, and only then can it be applied to previously unseen data. Several types of classifiers are discussed in Section 4.1.

The *expert system* is the overall *strategy* that is developed. Knowledge learned from preprocessing, feature extraction and classification is combined to make a final decision. It takes into account *contextual* differences such as those found between the EEG of different patients or different times of the day. The expert system is an entity that cannot be completely separated from either the classifier or the feature extractor because it encapsulates the entire process: a global strategy that determines what features to select, how to combine them and to what problem they are most applicable, so that

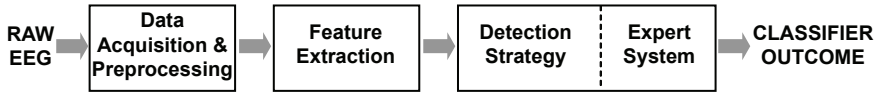


FIGURE 4.1: General structure of seizure detection algorithms. Data preprocessing and feature extraction describe ways that the original EEG signals are manipulated to extract information that can differentiate between the non-seizure, pre-seizure and seizure states. These two are discussed in Chapter 3. Classification and expert system, discussed in Chapter 4, are the way that these features are then used to make a decision as to which class they belong to.

performance is optimized. For example, extracting features from a signal is part of the feature extractor, but the selection of the *correct* features is part of the expert system because it requires knowledge of the problem. More examples are presented in Section 4.2.

4.1 Types of Classifiers

Classifiers are automated mathematical constructs that separate between predetermined classes, under the assumption that the feature set presented to it belongs to one of these classes. Formally,

A classifier maps an input space, in this case the feature set, to a decision space O each time a new feature set is presented

$$O = F(\zeta_1,\ \zeta_2, \cdots, \zeta_S). \tag{4.1}$$

The map $F(\cdot)$ may be a linear or non-linear combination of the features $(\zeta_1,\ \zeta_2, \cdots, \zeta_S)$.

Classifiers can be *deterministic* or *probabilistic*. The first draws strict boundaries between different classes and return a 'belong to' answer. Probabilistic classifiers determine the *likelihood* that a feature set belongs to a particular class. It returns a vector of the probability of belonging to each class.

Most classifiers discussed here are deterministic, although stochastic counterparts exist and are discussed where appropriate.

The quality of the decisions is only as good as the design of the classification system. For example, let us design a classifier to determine the owner of the bouncing ball. Let the extracted feature of the ball be its color, and we train the classifier that red balls belong to Mary and blue balls belong to Terry. What does the classifier do when it comes across a green ball? It was not trained to deal with this and its decision will be unpredictable. *It is important to select an appropriate* training set *– the data used to train the classifier to make decisions. The training set must be representative of all types of data likely to be presented to the classifier.* In most cases all the possibilities are not known, and it is wise for the classifier to contain an 'I do not know' state.

Sometimes, as is the case with detection of epilepsy, it is not possible to know all the types of data that will be presented. Consider the development of a classifier used to detect seizures that are very infrequent and of

rare morphology. EEG databases are unlikely to have examples of this rare phenomenon, and adding these seizures to the database may not be possible because typical EEG monitoring lasts days rather than months. The classifier cannot be trained to react to these events.

Once a classifier is trained its performance must be evaluated using a *testing set*. This is part of the validation process that ensures the classifier is behaving as desired. *The training and the testing data should where possible be mutually exclusive*. The development process should avoid using data in the test set, and the validation process should not contain any of the training data.

Three approaches to classification are presented in this section – *association rules*, *artificial neural networks* (ANN) and *support vector machines* (SVM). All methods attempt to find an optimal map $F(\cdot)$ between feature space and output space, the differences being the way in which decision boundaries are obtained.

4.1.1 Association Rules

Association rules use relationships between the features in the feature set and the correct output to determine an appropriate $F(\cdot)$. The process is usually manual and no formal training procedure exists. The feature set is inspected, simple relationships using features are selected and thresholds are used to make the decision. The relationships do not need to be simple but in practice often are because automated methods are better at finding the more complex ones.

A threshold is a *number* that separates between classes in a single dimension, e.g., when applied to a single feature ζ_s. A *hyperplane* is an analogous decision in multiple dimensions, e.g., when a combination of several features are thresholded together. The first is simpler to compute but allows for redundancy because features are treated independently. The second process is more powerful, but again automated methods discussed later are more suitable to find these. In all cases it is not always easy or even possible to draw a clear distinction between classes, as seen for example in Figure 4.3(b) and (c) where a simple line or hyperplane cannot unambiguously identify each class.

In epileptic seizure detection all early algorithms developed the association rules manually because no better way existed. The focus was on finding features that showed marked differences between classes rather than on finding complex maps $F(\cdot)$ [175, 177].

Recently more involved methods have been used to try reduce the number of false classifications in seizure detection [188]. Decision trees trained to represent $F(\cdot)$ are used in [135], and stochastic methods are presented in [156] where Bayesian analysis is used to determine the *conditional probability* that the partial EEG record belongs to a seizure. Relaxing of strict boundaries allows some flexibility in $F(\cdot)$, particularly when seizure probabilities can be adjusted given the presence of commonly observed artifact.

Most seizure prediction algorithms today use thresholding because the computational expense focuses on feature extraction rather than the training of more sophisticated classifiers.

4.1.2 Artificial Neural Networks

Artificial neural networks (ANN) are a mathematical, albeit rough, analogy to biological networks of neurons found in the brain. Each functional unit in an ANN resembles the behavior of a biological neuron, and when connected together *artificial neurons* or *nodes* are well suited for tasks such as pattern recognition and classification. In this chapter ANNs are described in the context of classification of the EEG where a feature set can belong only to a limited number of predefined classes. An ANN is *trained* to estimate the map $F(\cdot)$ in Equation 4.1. More information about a broader applicability of ANNs can be found in [64] and [100].

The idea for ANN began as early as 1943 when a mathematician and a neuroanatomist (Pitts and McCulloch) decided to collaborate and concluded that a sufficiently large network of simple "all-or-none" functional units could be used to approximate any realistic function [64]. Since then the field has grown tremendously, ranging in applications such as system identification, signal detection and classification, as well as modeling of real biological neural networks [100].

An artificial network contains $k = 1, 2..K$ interconnected nodes[1]. Each node is composed of three basic elements shown in Figure 4.2(a):

- *Inputs*: a set of M inputs to the node, $s_{k1}, s_{k2}..s_{kM}$, each with its own weight, $w_{k1}, w_{k2}..w_{kM}$, that characterizes the influence that each input has on that particular node. This structure is analogous to the dendrites of biological neurons, although any "synapse" in an ANN has adjustable weights that can take a positive or negative value. The strength and polarity of a real synapse is more or less fixed.

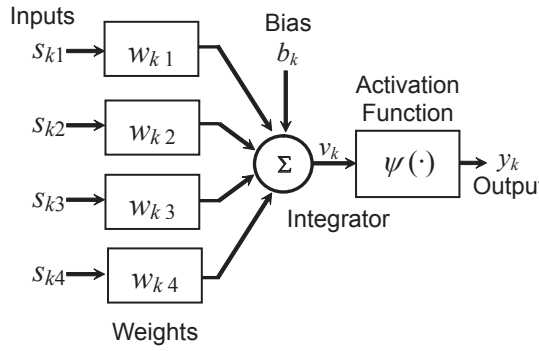
- *Integrator*: a summation mechanism that combines the inputs as

$$v_k = \sum_{m=1}^{M} w_{km} s_{km} + b_k. \tag{4.2}$$

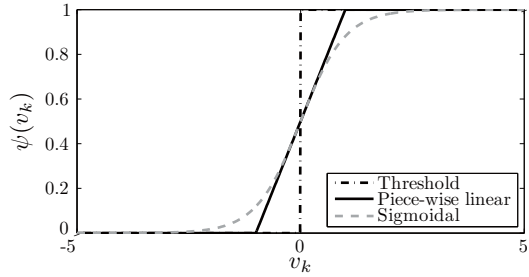
  This is analogous to the membrane potential observed at the soma of a real neuron. b_k is used to change the DC offset of v_k.

- *Activation Function*: a function $\psi(\cdot)$ that transforms v_k to the output of the node so that $y_k = \psi(v_k)$. In a biological neuron this is the function that determines the firing of an action potential. The activation function is used to limit the output range at each node.

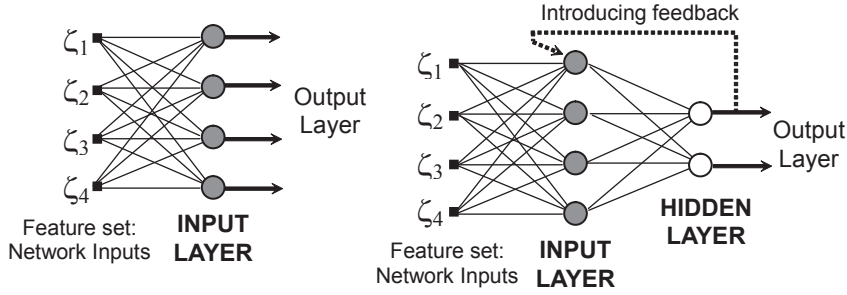
[1]The term *node* is used from now on, as opposed to neuron, so as to differentiate these artificial units from their biological counterparts.



(a)



(b) Example activation functions



(c) (d)

FIGURE 4.2: Artificial neural networks. (a) shows the components of a single artificial neuron, (b) shows typical activation functions, (c) shows an example of a single layer feedforward network with $M = 4$ inputs, $K = 4$ nodes, and with 4 outputs and (d) shows an example of a multi-layer feedforward network, with one input and one hidden layer. The dotted line shows how feedback may be added to this feedforward network. Networks in (c) and (d) are both fully interconnected.

It is with $\psi(\cdot)$ that non-linearity can be introduced into ANN. Figure 4.2(b) shows examples of typical activation functions. The simplest form is a threshold – when v_k is above a certain value, y_k is 1, otherwise it is zero. A larger range of outputs can be introduced by using a piecewise linear or a sigmoidal $\psi(\cdot)$; both are in Figure 4.2. The sigmoidal function is by far the most commonly used because it is differentiable [64]. An example sigmoidal function is

$$\psi(v_k) = \frac{1}{1 + \exp(-av_k)}, \tag{4.3}$$

where a is the parameter that determines its slope. Note the similarities between artificial neural networks and the modeling of biological neurons described in Chapter 6. A threshold activation function is similar to that used in the integrate and fire biological neural models, whereas the sigmoidal activation function is similar to that used in continuum models (see Equation 6.8).

Networks are created when multiple nodes are interconnected. Typically the nodes are arranged in *layers*, although in theory any configuration is possible. The layer to which the inputs of the classifier project is known as the *input layer*. Here the input layer is that to which the feature set $(\zeta_1, \zeta_2, \cdots, \zeta_S)$ is connected. The collection of nodes that connect directly to the output of the classifier is known as the *output layer*, corresponding to O in Equation 4.1. Any layers between the input and output layers are referred to as *hidden layers*.

Networks in this configuration can be loosely classified dependent on the number of layers and the types of interconnections found between them. They may be [64, 100]:

1. *Single Layer Feedforward*: The network consists of only one layer of nodes and the input layer projects directly onto the output layer. An example with $M = 4$ inputs can be seen in Figure 4.2(c). The outputs are not connected back to the input layers – it is a *feedforward* network. This type of network can approximate linear maps $F(\cdot)$.

2. *Multi-layer Feedforward*: The network consists of at least one hidden layer. The outputs of each layer project onto the next layer, and so forth, until the output layer. The output of each layer never projects to earlier layers – again it is feedforward. An example of an ANN with one hidden layer can be seen in Figure 4.2(d). Through the addition of at least one hidden layer non-linear maps $F(\cdot)$ can be estimated [178].

3. *Recurrent/Feedback*: A network in the configuration of (1) and (2) above, but where at least one output of a later layer projects back to the input of an earlier layer is called a *recurrent network* because *feedback* is added to the system. Adding recurrency allows the output of a node to be directly or indirectly affected by its previous outputs. Figure 4.2(d)

shows how the multi-layer feedforward network can be transformed to include feedback connections.

The number of connections in a network is also important. A network where every node in a layer is connected to every node in the next layer is *fully interconnected*. These are more flexible but require a greater number of weights and thus more processing power than *partially interconnected* networks. All networks shown in Figure 4.2 are fully interconnected.

Once a network configuration is chosen (dependent on the task) the ANN must be *trained* to estimate the map $F(\cdot)$. During training the *weights* in the ANN are selectively and recursively adapted until the desired behavior is observed. This is similar to how the brain learns tasks by adapting the way in which its neurons interact. The manner in which the weights are adjusted is called the *learning algorithm*.

The aim of the algorithm is then to, given an input and a desired output, adapt the weights in the ANN so that its output matches as closely as possible the desired output. This is done through the introduction of a *cost function* that penalizes divergence from the desired output. It may be as simple as the Euclidean distance between the ANN output and the desired output, in which case the cost function is a measure of the error. It is the aim of the learning algorithm to select the set of weights that minimize this cost function.

A plethora of such algorithms exist, each designed to work best with a particular network configuration. The *back-propagation* learning algorithm designed for multilayer feedforward ANN, most commonly used with one hidden layer, is an important one for classification problems. It is a form of steepest descent optimization where the gradient direction is calculated iteratively and in a computationally efficient manner. The weights of the nodes are adjusted so that the overall error between desired and actual output are minimized. Like all gradient based methods it suffers the problem that only local minima are found, and the final weights are normally a function of the initial randomly selected ones.

Another typical network configuration is a *self organizing map* (SOM) that utilizes *competitive learning* during the training process. Self organizing maps are single layer feedforward networks that are fully interconnected. The idea behind competitive learning is to emulate some behavior of a biological brain. Subgroups of neurons in the brain learn to react to a particular stimulus; thus sub-groups of nodes in an SOM learn to respond to certain input patterns. During the learning process nodes *compete* to be activated by the input, but it is only the node that minimizes the error that "wins" and has its weights adjusted. In this way nodes in an SOM learn to specialize, and are best suited to pattern recognition. Because nearby nodes usually react to similar patterns the SOM may be arranged in a 2D or 3D structure for visualization and interpretation of this distance. The initialization of weights during the learning algorithm is random so that initial reaction to input stimulus is different at each node.

In EEG classification ANN is more flexible than association rules in estimating the map $F(\cdot)$. The cost of implementing ANN is relatively small. ANN is a non-parametric approach that assumes nothing about the underlying model and as such no selection of parameters outside those already discussed is necessary. The process can in many cases be largely automated, whereas the development of association rules is cumbersome. Nevertheless there are disadvantages and in EEG classification these include:

- ANN is *adaptive* in that it can be retrained when the environment changes, making it suitable for non-stationary signals or for problems in which new data become available. However, this re-training process is not possible on-line and may need supervision by an ANN expert. In contrast association rules allow any user not familiar with the theory to change common thresholds so as to adjust performance [191]. Furthermore, adaptation does not guarantee that ANN performs better – in some cases it may in fact degrade performance.

- ANN can be organized so as to reflect a chosen confidence level for its decision. This is important.

- ANN can be trained to detect events which are *not* epileptic as well as those that are. This allows inclusion of artifactual data known to trigger false classifications so that the network can learn to reject them. It is more difficult to implement this using association rules.

- ANN, like any other classifier, can only behave as well as its teacher. If incorrect or insufficient data are used during training the reaction of the network can be unpredictable when presented with previously unseen data.

Error correcting for multi-layer feedforward networks and competitive learning for SOM are the most commonly used for classification of EEG seizures. Their relative merits are as follows:

- Backpropagation is guaranteed to converge to a minimum of the cost function, although convergence may be to a local rather than a global minimum. Less is known about competitive learning for SOM networks, and convergence even to a local minimum cannot be guaranteed.

- Backpropagation requires training data that is both epileptic and non-epileptic so that the ANN can learn to distinguish the two classes. SOM only requires examples of epileptic periods (and perhaps examples of typical artifact) so that neurons may learn these patterns. This avoids the problem of *over-representation* – the selection of the correct amount of data so that the relatively infrequent epileptic events are not under-represented during the training process.

- SOM can learn as many types of patterns as required, so long as there are enough nodes in the network. The selection of an appropriate number of nodes is a problem for both SOM and multi-layer networks. If too few nodes are used then the network is incapable of differentiating between complex patterns. If too many nodes are used then the network can potentially learn to react to noise in the data. This problem is known as *over-parametrization* [178, 190].

- Learning algorithms in ANN are computationally taxing, and convergence to a minimum may be particularly slow for backpropagating algorithms.

ANN have been incorporated into classification systems designed to detect epileptic seizures as early as 1994 in [162]. Given their popularity and their abilities in pattern recognition it is not surprising that their use remains strong today. Early work including that in [137] bypassed the need for feature extraction, using the raw EEG as an input to the ANN. However extracting features is a way to compress the input space so that training of ANN is faster.

SOM has been used by [45],[44] and [137] but it is multi-layer feedforward, typically with one hidden layer, that remain the most popular. These networks have been used for the classification of many EEG phenomena other than seizure activity. For example in [188] the system is used to recognize common EEG states such as sleep, alpha rhythms and artifact.

A series of publications [8, 178, 175, 177, 176] have recently emerged comparing the performance of different types of ANN, used for seizure detection, to each other as well as to common association rule methods. Spectral features were used for training. They conclude that feedforward ANN is the most suitable choice, although the results indicate that differences in performance are minimal and may be a consequence of the development process. SOM is not included in this analysis.

Early versions of one of the more sophisticated algorithms in the literature, `REVEAL`, suggest the use of a *probabilistic* ANN (or PNN). The major advantage is that it supports incremental learning – a new type of seizure can be added to the training set without the need to retrain the entire network. This comes at a computational cost that is nevertheless acceptable given the alternative (complete re-training of multi-layer networks). A series of papers that describe this process are [192, 194, 191].

4.1.3 Support Vector Machines

Support Vector Machines (SVM) are another type of machine learning suitable to estimate classification map $F(\cdot)$. SVMs are linear classifiers in that their aim is to find a hyperplane that separates classes of objects. The hyperplane that simultaneously minimizes error and maximizes the distance between the two classes forms the *classification boundary*.

An example of two classes that are linearly separable is shown in Figure

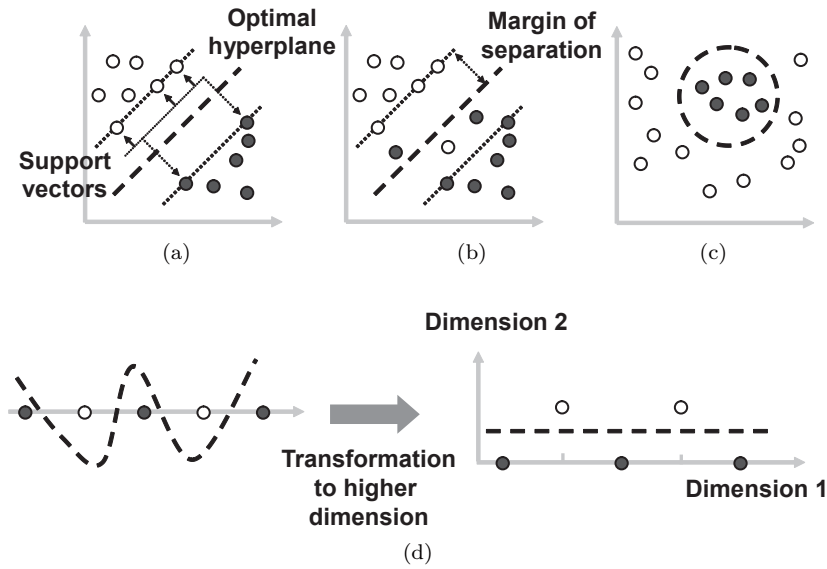


FIGURE 4.3: SVM classification problems. (a) is a linearly separable example, (b) is an non-linearly separable problem which may be solved by allowing *soft* margins and (c) is a completely non-linearly separable case in which linear separation may be achieved only through projection to higher dimensional space. (d) illustrates this, by showing how projecting a one dimensional space to two dimensions may make the classification linearly separable.

4.3(a). Here the hyperplane of separation in two dimensions is a straight line between the black and gray circles. The *optimal hyperplane* is the line that maximizes the *margin of separation* – the distance between classes, also marked in Figure 4.3(a).

Only a subset of the training data is used to compute the 'optimal' separation hyperplane. These data are known as *support vectors*, and they are the points that maximize the distance between the known classes. The SVM training algorithm determines which data to use as support vectors, and then uses these to find the optimal hyperplane. Finding the support vectors consumes most of the processing time. The interested reader may find the descriptions in [64] useful. An implementation of the algorithm can be found in [173].

Sometimes it is not possible to linearly separate two classes, as shown in Figure 4.3(b) and (c). In (b) new training data that lie on the wrong side of the decision boundary are available, and the two classes are no longer linearly separable. In such cases the margin of separation is made *soft*, and the aim of the SVM algorithm is then to compute the hyperplane that allows some errors to occur whilst still maximizing separation between support vectors. This is

achieved by the introduction of a *slack variable* that measures the deviation from ideal separation.

In Figure 4.3(c) slack variables do not help – the classes are in no way linearly separable. In such cases the SVMs can still be used, thanks to Cover's theorem that states that non-linearly separable data may become linearly separable with high probability if the input feature space is projected to (usually much) higher dimensions [64]. A simplistic and extremely unrealistic example that nonetheless illustrates this point can be seen in Figure 4.3(d). The left panel shows that in a single dimension no linear function can separate the two classes. However, imagine a transformation to two dimensions where a second dimension is introduced by assigning data the value of 0 and 1 in an alternating manner. This is shown in the right hand side of Figure 4.3(d). A line is now capable of separating the two classes. Cover's theorem does not say what an appropriate transformation should be, but several options exist, as discussed in [64].

SVM training algorithms are more generic than ANN in that their applicability is wider. For example, the back-propagation algorithm was designed to perform best for a particular type of network, whereas SVMs work for a very wide class of problems [64]. Furthermore, SVMs have several advantages over ANN for the purpose of seizure recognition:

- Both ANN and SVMs aim to find an 'optimal' decision boundary between classes, but whilst the former is computed by minimizing the classification error of the training set, SVMs maximize the *distance* between the boundary and both classes. This distinction implies that SVMs perform better than ANN for examples unseen in the training set – a property well suited to EEG data that is noisy and varied in nature [170].

- The selection of a subset of data to use as support vectors means that the performance of SVMs is not affected by the over-representation of data in the training process, as is the case with ANN [170]. Given that so much more non-epileptic data exists the danger of over-characterization is strong and this is a particularly useful property for seizure recognition. Less care is needed in the selection of a training set.

- Given identical training data SVMs *always* converge to the same answer, unlike ANN that may get stuck on local minima depending on (often random) initialization parameters [49].

- SVMs reduce the risk of over-parametrization or *over-fitting* through the use of the slack variable. This is not the case with ANN [49].

Nevertheless SVMs are not as popular as ANN in EEG seizure classification. Examples in which they have been used include the classification of PSD based features by [49] as early as 1999, and recently in [171] where the spectral features are now extracted by wavelet based methods. Testing on

larger EEG databases is necessary before the applicability of SVM to seizure detection can be determined.

Whether through the use of association rules, ANN or SVM all classifiers are designed to estimate an optimal mapping function $F(\cdot)$ between feature space and class boundaries. The differences between methods lie in how these decision boundaries are calculated, what assumptions are made about the underlying system and the computational complexity that is tolerated.

Association rules are non-formalized and non-automated methods of finding $F(\cdot)$. This makes them highly flexible at the expense of shifting all complications back to the designer – appropriate models and assumptions must be derived. ANN and SVM are more structured methods in that these underlying assumptions already exist for a generic class of problems. This makes them less flexible in general, the advantage being that both ANN and SVM are non-parametric and require the choice of only a few parameters by the designer. Thus, unlike association rules, they can be largely automated to estimate a suitable $F(\cdot)$ that does not rely on assumptions about the nature of the data.

4.2 Expert System

This section discusses in more general terms what may be incorporated into a complete detection or prediction system to account for differences observed between and within EEG traces.

Choosing the correct number and type of features to use in a classifier is part of the expert system. The process cannot be approached naively with a "more is better" approach because this is simply not true. Overparametrization does not only make a problem computationally harder but also adds unwanted redundancy. *If a feature is to be added to the feature set it must provide new or at least partially independent information.* Most features extracted from EEG, even if they employ different feature extraction methods, display some degree of correlation between them. In such cases a multivariate approach, where features are combined prior to any decision making, is likely to provide better results [75].

An EEG record is typically composed of many signals located on different

parts of the head. Features can be extracted from each signal $x_c[n]$ separately, where $c = 1, 2..C_{TOT}$ is the signal number, or from groups of signals combined together. Signals can be classified independently, resulting in C_{TOT} separate decisions, or features can be combined prior to the classification process. In any case the location of a signal has impact on the decision and decisions should make use of this *spatial context*.

There is also *temporal context*. Background activity can affect a decision, and expert systems should capitalize on this information. Temporal context is implicitly included in the feature extraction process through the selection of window sizes and their overlap, but questions such as how a classifier reacts to recent detections or how it reacts when a person is asleep or awake is information that can be added by the expert system.

Finally performance can depend on *environmental context*, where algorithm parameters can be tuned to a particular case. For detection and prediction of epileptic seizures this is most often done by making the classification *patient specific*. Features that are particular to a person are targeted. Almost all published algorithms claim a certain degree of user-tuneability, although ANN and SVM based classifiers less so.

It is a reasonable assumption that complicated problems require sophisticated solutions [98]. The long history of EEG seizure detection and prediction has shown that naive methods cannot be successful and creativity in the expert system is required. For example signals other than EEG such as heart rate could provide additional useful information.

The remainder of this section expands on comments made thus far, giving examples of appropriate combination of features in the decision making process (Section 4.2.1), the different types of spatial and temporal context (Section 4.2.2) and how to make algorithms patient specific (Section 4.2.3). The reader is advised that any combination of these is possible as a complete strategy for a classification system.

4.2.1 Processing Decisions

Decisions can be made for each EEG channel separately, at different points in time, with different degrees of certainty, etc. At the end of the day, however, a detector must give a yes/no/don't know answer. This is what is meant by processing of decisions. How is the information of all channels combined? Is this done at the feature extraction stage or at the classification stage?

Complete strategies can involve the combination of many different classifiers. Association rules can be combined with ANN or SVM by applying them to features before or after the classification process. Figure 4.4 shows some possible configurations applied to features extracted independently from four EEG channels. The structure of (a) is as in [188], the structure of (b) is as in [171] (but with SVM), and the structure of (c) is as in [194]. Sometimes rules can be developed as a two stage detection process, one example being in [98] where features are first classified using association rules. As a first stage 75%

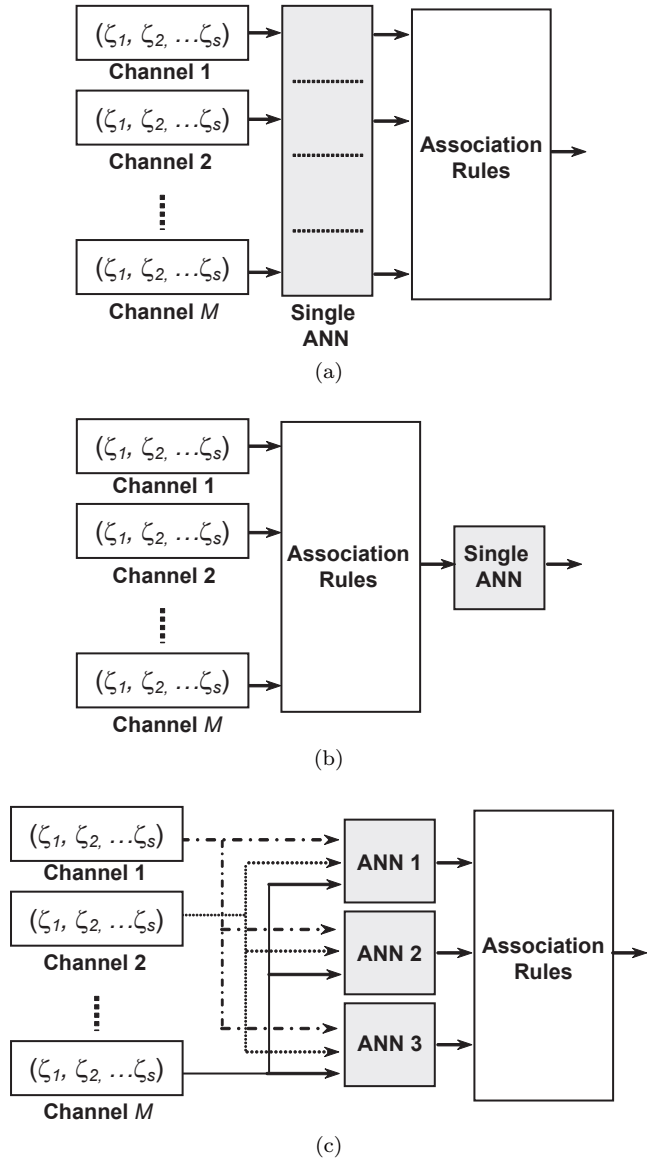


FIGURE 4.4: Example expert classifier combinations. In (a) each feature is passed, independently, to the same ANN. Association rules between these decisions are applied to obtain a final outcome. In (b) association rules are applied to the feature set first, the outcome of which is passed to a single ANN to make a decision. In (c) different ANN systems have been trained to recognize particular features. Decisions are made using all features at once for each "if/else" ANN rule, and association rules combine these as a final stage to obtain a decision.

of the data are efficiently rejected in this way. In the second stage detections are fine-tuned using more sophisticated classification, in this case ANN.

A final example of how an expert system may differ from a simple feature and threshold scheme is the one in [176], where the expertise is now put into the training process. Many identical ANNs are trained simultaneously, but because they are initialized randomly they converge to different local minima. The training system then selects the neural network that performs the best through a form of competitive learning. An article that explains how this works is [19].

4.2.2 Spatio-Temporal Context

An expert system can add context by deciding how a classification is made given (1) the patient, (2) the time and (3) the location. For a patient unspecific detector only (2) and (3), the *spatio-temporal* context, is relevant. Both temporal and spatial information are implicitly included in any classification system through the choice of windowing and through the combination of EEG channels. Here, more explicit methods to include spatio-temporal context are discussed.

The simplest form of temporal context is to report a true detection only if consecutive detections over many windows are made. This is useful because it eliminates the mis-classification of short bursts. It is incorporated in most systems in one form or another, an example being in [50] that requires at least two detections in a row.

Comparing to a *background* is another way by which temporal context can be added, for example a classifier can adapt its behavior dependent on whether the patient is asleep or awake, can ignore short burst by smoothing features over time or can use relative rather than absolute values wherever a seizure is expected to differ significantly from the background. Features can be made relative to short term background (a few seconds) or long term background (minutes to hours). Some detectors ([128, 127]) use median filters instead of averages to quantify background activity so as to avoid the interference of short, large amplitude bursts. *Forgetting factors* and variable window lengths can be used to place more importance on recent events by either omitting or attenuating data that are less recent [194]. The work in [55] was the first to incorporate background updating, using 20 seconds at a time and including a gap of 12 seconds between background and present so that slow onset seizures are not missed.

Most detectors incorporate spatial context at the end of the classification process. A decision is first made for each channel independently and a true detection is reported only when more than a few channels agree. The trade-off is in the number of channels to use to avoid spurious detections versus missing focal seizures that only manifest on a few signals. This last is also the reason why channels are treated separately in the first place. How some contextual information may be included by selective grouping is discussed in Chapter 5.

In prediction of seizures spatial context is often a much more involved process, requiring the careful selection of subsets of channels, for example the synchronization work done in [112, 113] where the predictor measures differences in synchronization at different sites. In the extraction of maximal Lyapunov exponents [71] there is selective inclusion of channels based on how similarly they behave. The role of such multivariate methods remains largely unexplored for detection of seizures [75].

4.2.3 Patient Specificity

Patient specific algorithms in EEG are suitable only when sufficient data are available. For the detection or prediction of seizures this means that a sufficient number of epileptic events must be recorded so that parameters can be tuned to a particular person. Often this is possible only for long-term monitoring where days rather than hours of EEG are collected [156].

Once sufficient seizure data are available, the expert system can be trained to detect (or predict) seizures only if they are *stereotyped*, that is, when the extracted features are consistently similar before or during the event. Fortunately for epilepsy the number of different types of seizures particular to a person is often small, and so long as an instance of each is present in the training process the classifier can be tuned. The algorithm can also be trained to recognize typical events that trigger false detections, often patient specific manifestations themselves [74].

Algorithms developed for patient un-specific cases can perform better when tuned to the patient specific case [45], but methods developed for patient specific cases cannot usually be generalized to be patient un-specific. Their performance is poor, as demonstrated in the work of [171], [139] and [140].

Patient specificity can be as simple as selecting channels known to be involved in seizures. For example focal seizures that never spread to the entire brain can be analyzed using only a subset of channels. This is particularly important for intra-cranial onset detectors when seizures begin in a very local and specific region. More sophisticated methods utilize spectral ([163, 191, 192]) and temporal signatures ([139] and [140]) – patterns that are stereotyped and can be used to recognize events. All methods that employ the extraction of signatures are by necessity patient specific.

Although the search for the perfect patient un-specific algorithm continues for the detection of seizures, in prediction it is more or less accepted that classifiers *must* be customized to the individual [45]. Prediction of seizures is much more susceptible to the balance between the correct selection of features and spatio-temporal context than is detection. All methods in seizure prediction that have shown promise are patient specific.

4.3 Conclusions

Designing a classifier for a complex system such as the epileptic EEG involves the stages shown in Figure 4.1. The features extracted using methods described in Chapter 3 must be combined, analyzed and classified in an expert manner.

Given a feature set as an input, a classifier (association rules, ANN, SVM or otherwise) determines which class this feature set most likely belongs to. Machine learning algorithms such as those used for ANN or SVM help in standardizing and automating the task of determining the decision rule.

An expert system addresses the more difficult question: how do we make our classifier work best?

A deep understanding of the problem is required to know what expert knowledge to incorporate into the overall detection system. For epileptic seizure detection and prediction it involves adding contextual information that exploits the spatial and temporal aspects of the EEG.

Detection, classifiers and expert system are not based on a true understanding of the relationship between the epileptic brain behavior and the EEG. Instead these algorithms try to emulate human decision making using the same EEG records. This chapter together with Chapter 3 surveyed signal processing, classifiers and expert systems applied to the detection and prediction of seizures using EEG. The analysis presented and the references to the literature are not exhaustive – they point to the type of research conducted in the past few decades. Although hundreds of papers have been published in this area, it is fair to say that at this stage most are explorative. This is a testament to the difficulties in the classification of the epileptic EEG, discussed next.

5

Seizure Detection

Automated detection of epileptic seizures from EEG records is an old problem, and work continues because detectors are driven by technology. However to date little or no standardization exists in what 'good' performance or 'good' detection means.

To our knowledge no comprehensive review of seizure detection algorithms where state of the art detection strategies are compared *on the same dataset* has been published. The aim of this chapter is to explore the difficulties of seizure detection as well as to provide a norm under which current and future seizure detectors can be validated. We work toward developing a systematic and consistent framework that defines standard performance metrics.

Of most importance is the standardization of the data that is being used. An algorithm is developed on a *training* dataset and evaluated on a *testing* dataset.

To validate an algorithm as a successful detector guidelines are proposed in [53]. The *testing* dataset must:

1. Be different from the *training* dataset (to some extent).
2. Contain many different types of seizures.
3. Not be pre-selected in any way (e.g., based on seizure type) unless specifically required by the detector.

A 'sufficiently large' database minimizes the risk that performance comparisons between algorithms are biased, but is nevertheless an arbitrary and subjective requirement. It is not too difficult to get hold of thousands of gigabytes of data, but who controls how representative this data is? How is the bias of the professional marking the epileptic events accounted for? *Comparisons between algorithms tested on different 'sufficiently large' datasets may be unrepresentative.* For this reason an additional guideline is proposed.

The comparison of performance between detectors should be standardized. This requires an *independent evaluation* and/or a *common dataset* used to test *all* algorithms.

Whilst there are a few independent studies that evaluate the performance of algorithms (see [132, 194]), to date no mass testing of seizure detection algorithms on the same dataset exists. We attempt to remedy this in this chapter: several algorithms are tested and compared on the same dataset. The reported performance is of course relative rather than absolute. It is the hope that with appropriate ethics clearance this database may be made publicly available so that results here become benchmarks for future research. This initiative would also make it possible for researchers to test their work.

The *metrics* used to compare algorithms must also be made standard. The number of true detections (*true positive rate* (TPR)) may be as important as the number of false detections (*false positive rate* (FPR)). These, along with others, are discussed in more detail later.

Clear and pre-defined metrics that compare the performance of an algorithm must be reported. These metrics must wherever possible be objective rather than subjective.

The problem statement of seizure detection is discussed and formalized in Section 5.1. It includes details of the data and the metrics used throughout this chapter, both of which depend on the application.

Once a standard dataset and a way in which to evaluate performance is defined, it is then possible to explore the difficulties in designing a successful detector. In this chapter three key aspects of a detector such as that shown in Figure 4.1 are identified and tested separately:

1. *Evaluation of Classification Methods*: A detector bases its decision on a classification method, as presented in Section 4.1. The aim in this test is to determine the relative ability of ANN and SVM to classify epileptic EEG data using the same features as input to all classifiers. This is discussed in Section 5.2.

2. *Evaluation of Patient Un-specific Seizure Detectors*: Current patient un-specific seizure detection algorithms are available by the hundreds but are seldom tested on comparable datasets. The aim of this test is to determine the relative performance of leading seizure detectors, as

well as to provide a benchmark for future tests. Performance is evaluated based on the ability of a detector to identify any part of a seizure measured from scalp EEG records. This is discussed in Section 5.3.

3. *Evaluation of Onset Detectors*: The ability to detect the beginning of a seizure opens the door for new technology aimed at aborting a seizure once it has begun. The aim of this test is to determine the relative performance of different features as onset detectors on the intra-cranial EEG. This is discussed in Section 5.4.

The details and rationale behind the selected tests are all outlined in the relevant sections.

5.1 The Problem of Seizure Detection

There are many reasons why neurophysiologists need seizures to be detected from EEG. Seizure detection is useful for diagnosis: long hours of EEG must be reviewed and classified. Finding seizures is an important aspect of this. Sometimes it is necessary that seizures be detected on-line so that the spread of the seizure can be monitored, sometimes by injecting a radiation dye (radionuclide) at the time the seizure starts. It is important that the beginning of a seizure be identified, particularly for patients whose seizures are rare. Seizure detection is also useful for treatment: timely drug delivery, electrical stimulation and even predictors require seizures or their precursors to be detected reliably.

In all cases the automation of the detection process saves many hours and in some cases makes treatment viable. It is not surprising that after decades of research seizure detection remains an important focus of research. The motivation to improve detectors as well as develop new technology that target specific applications remains strong.

Why is seizure detection so difficult? Part of the problem is indeed that seizures can manifest themselves in many ways. The features most commonly observed in the EEG, reiterated from Chapter 1, are

- *Oscillations*: Each channel often becomes more oscillatory, in contrast to the examples of normal EEG shown in Figure 1.9.
- *Synchronization*: The EEG channels behave more like each other and activity is global.
- *Large amplitude*: After seizure onset the EEG is often much larger in amplitude than prior to a seizure.

However it is important to remember that the task is made difficult because

- Not all seizures display all or any of these features.
- Seizures can vary significantly between patients, within the same patient, within the same seizure at different times and within the same seizure on different channels.
- Many non-epileptic phenomena including artifact often have features similar to seizures.

So, coding a detector to target activity that 'is different from background and not artifact' seems like a good idea. However, extracting any one feature, or a set of features, capable of discriminating between seizure and non-seizure EEG (let alone artifact) is difficult.

This leads to yet another aspect that makes seizure detection difficult – *even human experts often cannot agree on what is a seizure and what is not.* It was shown that when four experts were asked to review the same EEG record, marked independently before by another expert, only 92% of the events were also identified by one of the new experts, and just below 80% were identified by two or more experts [193]. Even less agreement was evident as to the time of onset and termination of seizures.

If human experts cannot agree it is difficult to interpret the results presented by any published detector. For comparison to be fair these must be tested on the same database that has been marked by the same experts. This database may contain bias, after all they are marked by human beings, but by making this bias uniform to all algorithms then the relative result can be evaluated.

In the remainder of this chapter several tests are conducted whereby the same database, marked by the same human experts, is tested with the same performance criteria. This EEG database is described next, followed by the definition of standard metrics used for evaluation.

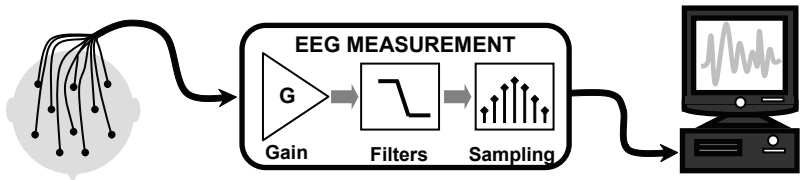
5.1.1 The EEG Database

Four groups of EEG data are available for testing. All datasets are acquired using the Compumedics^{TM} E-series EEG whose specifications are summarized in Table 5.1. The impedance of each channel is tested prior to recording – it must be high so that minimal current leaks through the electrodes. Each sample is recorded with noise levels guaranteed to below 2μV. Hardware filters are applied so that recorded signal lies between $0.15 - 105$Hz prior to sampling at 256 or 512Hz. Additional filters can be applied post-acquisition. Figure 5.1(a) shows a typical scalp EEG recording setup, and (b) shows the dimensions of typical scalp, subdural and depth electrodes. More detailed specifications and usage instructions can be found in [30].

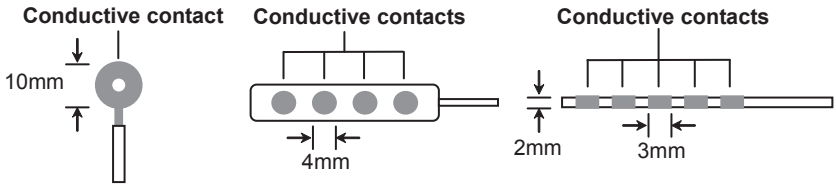
All scalp data was acquired using the international 10-20 electrode configuration system shown in Figure 1.8. Twenty-one channels are recorded with reference to an electrode located near Cz and re-referenced post-acquisition.

No. of Channels	up to 64
ADC Resolution	14 bits ±1 LSB bit
Sampling Rate	256 or 512Hz (Software Configurable)
Sampling Type	Sample and hold on all channels
Input Ranges	1, 2, 4, 8 and 16mVp-p
Noise	2μVp-p maximum Guaranteed 0.16Hz-105Hz
Electrode Offset	Up to ±350mV allowable DC offset
Electrode Impedance	100kΩ range Software allowable testing
Filters	High Pass 0.15Hz (Hardware) Low Pass 105Hz (Hardware) Notch 50Hz, 60Hz (Software) Additional software filters available
Amplifier Accuracy	±1%
CMMR	>110dB with 10Ω imbalance

TABLE 5.1: Compumedics E-series EEG specifications [30].



(a) Scalp EEG recording setup



(b)

FIGURE 5.1: A typical scalp EEG recording setup is shown in (a). Electrodes attached to the scalp feed to a computer for storage after amplification, filtering and sampling. In (b) are example dimensions of typical scalp, subdural and depth electrodes.

The electrode configuration of intra-cranial data was specific to the subject's condition.

All scalp records were marked by professional electroencephalographers from St Vincent's Hospital, Melbourne, Australia. To reduce bias, bursts shorter than 20 seconds are not used for evaluation. To reduce bias from patients with frequent seizures only the first 10 are included in testing. All EEG data after the first 10 seizures is discarded. There is no other form of pre-selection prior to testing other than to ensure that seizures were electrographically identifiable.

The details of each dataset are outlined next and summarized in Table 5.2. Which datasets and how they are used for each evaluation is discussed in more detail in the relevant section. Wherever possible all data is normalized using Equation 3.13.

5.1.1.1 Group 1 – Scalp EEG Data (< 6 Seizures per Patient)

This database contains scalp-recorded EEG from 15 patients. Between 1-5 seizures are recorded for each patient. A total of 41 seizures are identified in 361 hours of EEG recording.

The data in this group are used in the evaluation of patient un-specific seizure detectors (Section 5.3). When acquiring scalp EEG data it is not unusual to have some corrupted recordings due to, for example, improper attachment of electrodes. Where appropriate, performance is also evaluated when channels with excessive amounts of artifact are removed from analysis. These results are labeled as 'good channels only'.

5.1.1.2 Group 2 – Scalp EEG Data ($6 - 10$ Seizures per Patient)

This database contains scalp-recorded EEG from 6 epileptic patients. Between 6-10 seizures are recorded for each patient. A total of 51 seizures and 165 hours of EEG are available.

The data in this group are also used in the evaluation of patient un-specific seizure detectors but is differentiated from Group 1 because of the large number of seizures that each patient experiences. This makes it ideal for patient specific applications that require a separable training and testing dataset. It is used in the evaluation of classification methods (Section 5.2). As with Group 1 (and where appropriate), results are compared to when only good channels are used.

5.1.1.3 Group 3 – Scalp EEG Data, Non-Epileptic Patients

This database contains scalp-recorded EEG from 8 patients that are *not* epileptic. Only 11.5 hours of EEG records exist because routine monitoring of non-epileptic subjects is more rare in hospitals.

The data in this group are useful to complement the performance evaluation of patient un-specific seizure detectors (Section 5.3) because they provide

Patient #	**Seizure #**	**Mean Sz Length** (secs)	**Record Time** (hours)	**No. EEG Channels**	**Sampling Rate** (Hz)
Group 1 – Scalp EEG data, < 6 seizures per patient.					
Number of patients = 15, Seizures = 41, Total testing time = 361.9 hours					
1	3	185	6.7	32	512
2	4	349	32.3	32	512
3	2	268	25.2	32	512
4	3	140	24.1	32	512
5	1	66	6.9	32	512
6	4	58	15.8	32	512
7	3	149	65.5	32	512
8	3	80	31.3	32	512
9	5	74	40.8	32	512
10	3	109	41.9	32	512
11	1	109	17	32	512
12	1	99	4.5	32	512
13	5	48	37.5	32	512
14	1	92	7.3	32	512
15	2	74	5	32	512
Group 2 – Scalp EEG data, 6 – 10 seizures per patient.					
Number of patients = 6, Seizures = 51, Total testing time = 165.6 hours					
1	7	65	20.6	32	512
2	10	529	16.6	32	512
3	10	135	14.4	32	512
4	6	22	30.3	32	512
5	8	24	58.4	32	512
6	10	52	25.4	32	512
Group 3 – Scalp EEG data, Non-epileptic Patients.					
Number of patients = 8, Total testing time = 11.4 hours					
1	n/a	n/a	0.7	32	256
2	n/a	n/a	1.6	32	256
3	n/a	n/a	1.8	32	256
4	n/a	n/a	2.2	32	256
5	n/a	n/a	1.1	32	256
6	n/a	n/a	0.3	32	256
7	n/a	n/a	0.4	32	256
8	n/a	n/a	3.1	32	256
Group 4 – Intracranial EEG data					
Number of patients = 3, Seizures = 16, Total testing time = 96.5 hours					
1	6	123	23.5	6	512
2	5	82	23.8	6	512
3	5	199	49.2	6	512

TABLE 5.2: Summary of available EEG data divided into testing groups.

a comparison of how performance is affected by the presence of epileptic bursts between seizures. Whilst 11.5 hours is small in comparison to other databases it is used only as complementary information until more data can be gathered.

5.1.1.4 Group 4 – Intra-Cranial EEG Data

This database consists of intra-cranially recorded EEG from 3 epileptic patients being monitored for surgical resection. From this database a total of 16 clinical seizures and 96.5 hours of EEG are used here. Sub-clinical seizures are also marked but not used in the analysis. The database was gathered, marked and made available by the Freiburg Seizure Prediction Contest (For details, see `https://epilepsy.uni-freiburg.de/freiburg-seizure-prediction-project/`).

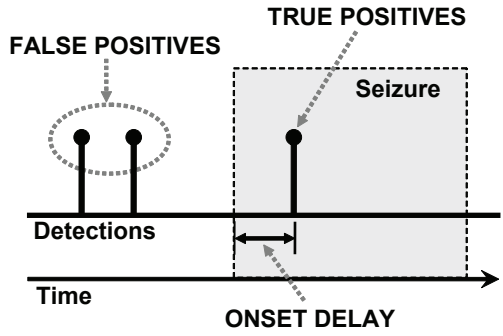


FIGURE 5.2: Above is a representation of what a true positive (TP) and a false positive (FP) are. If a detection occurs during the seizure, it is a true positive. If a detection is made outside this time, then it is a false positive. Also marked is the detection onset delay.

The configuration and number of electrodes for each patient differed. However in all cases only six signals, bipolar referenced, that are most obviously involved in seizure activity as judged by visual inspection are used.

Both the electrographic and clinical onsets are marked, making this database useful for testing of onset detectors described in Section 5.4. The length of time between electrographic and clinical manifestation was, on average, 27 seconds for Patient 1, 2 seconds for Patient 2 and 23 seconds for Patient 3.

5.1.2 Performance Evaluation Metrics

Ideally a metric used to evaluate the performance of a detector should be objective. In reality subjectivity cannot be avoided because the EEG records must be marked by a human expert with his/her own individual bias. This can be minimized when *relative* rather than absolute performance is reported, as is the case when a common dataset is used. The manner in which seizure detections are reported is formalized prior to testing so as to increase comparability between tests.

A *positive detection* is reported when flagged by the algorithm. Detections within 60 seconds of each other are grouped so that continuous bursts of positive detections are not over-represented

A *true positive* (TP) is reported when a positive detection occurs within the time marked as a seizure by a human expert. Due to ambiguous onset/offset times, a TP is also reported when

a positive detection occurs within 60 seconds of onset and offset of a marked seizure. Only one true positive is reported per seizure. A seizure that goes undetected is a *false negative* (FN).

A *false positive* (FP) is reported when a positive detection occurs outside 3 minutes before and after seizure. This 3 minute period is omitted so that epileptic activities frequently observed to lead up to a seizure are not counted as FPs. Only one FP is reported for every 60 seconds of continuous FPs.

Combining all these, it is only possible to make a maximum of one (true or false) detection per minute. Example true and false detections can be found in Figure 5.2.

Once TPs and FPs have been identified, the metrics that describe the performance of an algorithm can be formalized. These include[1]:

- *True Positive Rate* (TPR) *and Sensitivity*: This metric gives the probability that a seizure is correctly identified. Formally it is defined as

$$\text{TPR} = \frac{\text{Total number of TPs}}{\text{Total number of seizures}}. \tag{5.1}$$

  Sensitivity and TPR are equivalent measures.

- *False Positive Rate* (FPR) *and Specificity* (S): The FPR is the number of expected false positives that occur per hour of non-epileptic EEG. Specificity is a similar measure that describes the proportion of record that is marked. Many definitions exist, but here it is defined as the fraction of non-seizure time that is incorrectly marked as a seizure.

$$\text{S} = \frac{\text{Total FP time}}{\text{Total non-seizure time}}. \tag{5.2}$$

  S and FPR provide similar information but are not the same.

- *Onset Delay*: The onset delay is the time it takes for a detector to identify a seizure after its electrographic onset. Electrographic onset, in turn, is defined as the first time that at least 2 channels are visibly involved in the seizure activity. This is more subjective than other metrics because it depends heavily on how a human expert selects an often unclear point in time.

[1]Many different definitions of the metrics specified here exist. Although they are usually similar the differences are sufficient to make comparing results difficult. In this chapter the definitions important to the problem of seizure detection are formalized, again so that the problem can be standardized.

	High TPR	**Low FPR/S**	**Short Onset Delay**
Clinical diagnosis	Moderate	Moderate	Very low
Radiation injection	Moderate	Moderate	Very high
Seizure predictor	High	Moderate	Very high

TABLE 5.3: Relative importance of performance based on application.

TPR, FPR and S each give important information about the performance of an algorithm. Whilst TPR and FPR are sufficient to give an electroencephalographer a measure of performance, specificity S gives an additional idea of the ability of an algorithm to discriminate by estimating the proportion of time that a detector is flagging seizures. For example a detector that marks 90% of the record as seizure can have a misleading TPR of 100%, but looking at an FPR of 54, although very high, does not intuitively provide this information. There is redundancy in reporting both FPR and S but for review purposes one may be more instinctive to interpret than the other. Since they are not difficult to calculate it does not hurt to have both.

The computed TPR, FPR and S must not be not over-influenced by any one patient. True detections have been limited by allowing a maximum of 10 seizures per patient (see description of EEG data) to reduce bias. For example, it is not uncommon for a single EEG record to contain 20-30 seizures. In a database of 100 seizures this accounts for 25% of occurrences. If these seizures are particularly easy (or hard) to detect, results are not truly representative on how well seizures are detected across patients, particularly if the application of interest is patient un-specific. In a similar way some patients exhibit an excessive number of false detections. In all cases the FPR is capped at 15 per hour. A report of how many patients exceeded this FPR is given where appropriate.

Another consideration is that many epileptic bursts occur throughout an EEG record even though these do not develop as seizures. Whilst detection of these are not TPs, they are also not FPs. The issue is dealt with by including results using the Group 3 database, where all records are taken from non-epileptic patients. Since no part of these records are seizure related they give a better estimate of what the true FPR is. This eliminates the need to specify whether a false detection is 'interesting' or not as was done in some studies including [51] and [156]. This strategy is only possible for scalp records for ethical reasons.

The *true positive rate* (TPR), *false positive rate* (FPR), *specificity* (S) and where appropriate the *onset delay* are reported as metrics that can be used to compare the performance of each algorithm.

To avoid over-representation of any one patient its TPR and FPR is limited to 10 seizures per person and 15 false detections per hour. This is particularly important for patient un-specific tests.

FPRs are reported for both the epileptic records as well as for non-epileptic records so that bursts of epileptic activity that are neither TPs or FPs do not bias results. Specificity is only listed for epileptic records to give an idea of the proportion of EEG that can be discarded by the detector.

The interpretation of the performance depends on the application because of the relative importance of each metric. For example a short onset delay is important for a system designed to stop a seizure once it has started, but not relevant for a clinical review of EEG. Table 5.3 shows the relative importance of performance for the applications listed at the beginning of this section.

In [192] an example combined metric P that weighs the relative performance of a patient un-specific seizure detector is presented:

$$P = \frac{\max[0, \mathrm{S} - \mathrm{S}_{max}]}{\sqrt{2}} \sqrt{\mathrm{TPR} \times (1 - \log_{10}(\ \min[\mathrm{FPR}_{max}, \max[\mathrm{FPR}_{min}, \mathrm{FPR}]]\ )}, \tag{5.3}$$

where FPR_{min}, FPR_{max}, S_{max} are the minimum or maximum values that are tolerated from each metric, defined by the user and depending on the application. This looks like a complicated formula but is only an example of how to weigh the importance of each metric, for instance if P is zero when the FPR or S exceed a maximum allowable level. P can be increased proportional to an increasing TPR.

Other formulae for P can be conceived. This may require inclusion of onset delay, or weigh the metrics in a different manner. Only the metrics are reported here.

As a summary of the performance of a detector a *receiver operating characteristic* (ROC) curve is reported, where possible. The TPR is plotted as a function of FPR when a detection parameter is altered. An example is in Figure 5.3. It shows the trade-off between true detections and false detections because an increase in true detections is usually followed by an increase in false detections. In general the closer the ROC is to the top left corner (TPR=100%, FPR=0) and the further away it is from a random detection rate (dotted line) the better the performance of the detector.

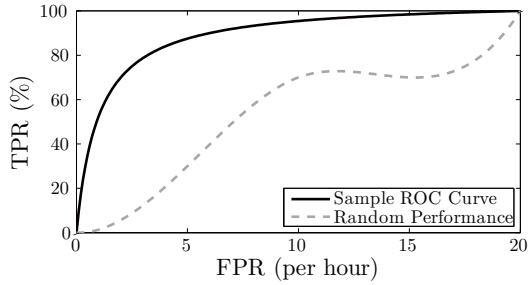


FIGURE 5.3: A sample *receiver operating characteristic* (ROC) curve. An ROC is a graphical representation of performance where the TPR is plotted versus the FPR when a detection parameter is varied. The closer the ROC curve is to the top left corner (TPR=100%, FPR=0) and the further it is to the random detection rate the better the performance.

5.2 Evaluation of Classification Methods

The job of a *classifier* in a seizure detector is to label features extracted from data as belonging to either seizure or non-seizure activity. Three different classes have been described in Section 4.1: *association rules*, *artificial neural networks* (ANN) and *support vector machines* (SVM). ANN and SVM have been introduced as an alternative to association rules because they largely automate the identification of a decision boundary, in particular simplifying the task when multi-dimensional feature vectors are involved.

This section is dedicated to the evaluation of the performance of ANN versus SVM for the classification of the epileptic EEG. In the literature many detection algorithms employ either ANN or SVM with no justification as to the choice, other than to propose that they are a better alternative to association rules. There are few references (e.g., [188, 190, 156]) providing numerical evidence as to whether SVM or ANN is suited to pattern recognition in the EEG, and even fewer try to distinguish between them.

The aim of this evaluation is to determine which type of classifier, when trained and tested on the same data, is best capable of correctly identifying characteristic behavior in the EEG. A patient specific test ensures that the high variability of observables between patients does not skew results. At the same time the (arguably smaller) variability found within the same patient determines which of the classifiers can generalize their results to similar but previously unobserved examples.

Only Group 2 data are used for this evaluation because each record contains many seizures, making patient specific training possible. For consistency between patients only the first 6 seizures in each are considered. Scalp rather

than intra-cranial EEG is used because the noisier environment makes separability between seizure and non-seizure activity much more difficult.

5.2.1 Feature Extraction

This evaluation is concerned with relative rather than absolute performance of each classifier. Whilst the feature extraction process is key in the performance of a detector here it is only necessary that they be capable of distinguishing between typical seizure and non-seizure activity – it is not essential that they be optimal. Frequency analysis is used because of its well demonstrated ability to differentiate between seizure and non-seizure EEG.

Specifically, 2 second non-overlapping windows of single channel EEG data are filtered into frequency bands of 2-4Hz, 4-8Hz, 8-16Hz and 16-32Hz to keep in line with dyadic-style distributions used in wavelet analysis. These frequency ranges include all frequencies most commonly associated with epileptic seizures whilst excluding higher frequencies often associated with artifact.

The mean signal power (squared value) after filtering is calculated for each frequency band. Each value is normalized by the total power, calculated as the sum over all bands. The resultant feature vector consists of four normalized energy values for each of the C_{TOT} channels. One epoch is characterized by the features in *all* channels and the classifiers are trained and tested on this $4 \times C_{TOT}$ dimensional feature vector.

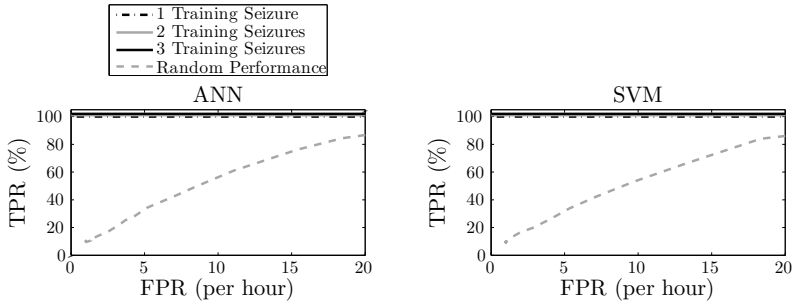
5.2.2 ANN Training and Testing

The performance of many types of ANN can degrade when too much data of one type is presented to it during training. This is not ideal because in seizure detection at least 95% of recorded data is non-seizure. It follows that performance depends on the data selection [52]. Extensive testing revealed that too often the ANN training did not converge when both seizure and non-seizure data are used for training, and thus could not be compared to that of SVM.

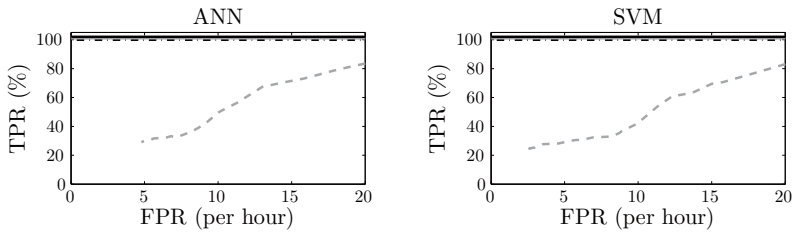
For this reason this evaluation utilizes a self organizing map (SOM). This is a type of ANN that maps itself to the topology of a given training set, thus can be trained using only seizure data. See Section 4.1.2 for a more detailed description of SOM, and Figure 4.2 for some examples of ANN node configuration. All data within a seizures in the training set is used to train the SOM.

It is important that the number of nodes in the network is large enough to describe the topology of the feature vector whilst small enough to avoid over-fitting. Many network configurations with different number of nodes are tried here for each patient, and the one that performs best is kept. Typically this optimal performance was obtained with 5 to 6 nodes.

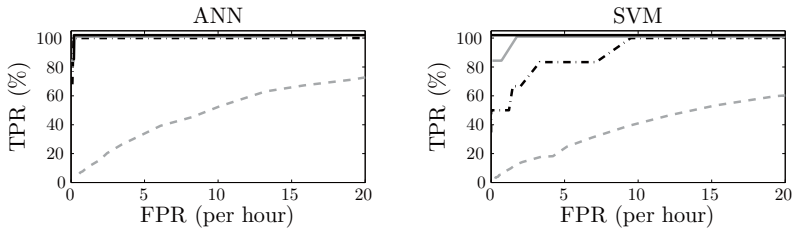
Testing consists of presenting an input feature vector to each node in the network and measuring the distance between them. In theory if the input is



(a) Patient 1

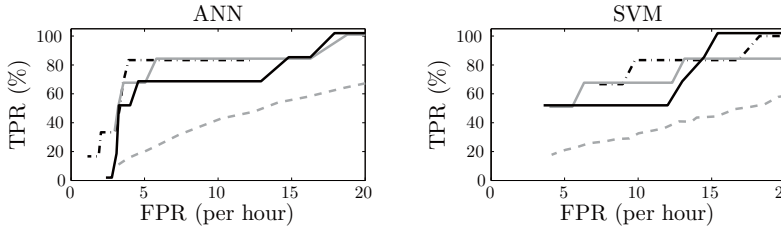


(b) Patient 2

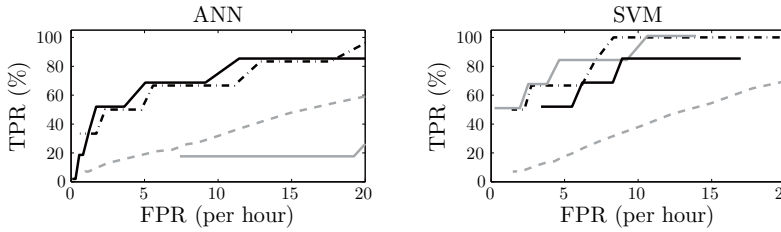


(c) Patient 3

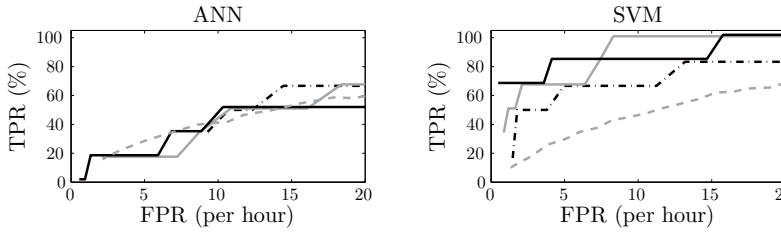
FIGURE 5.4: Individual patient performance on Group 2 data. In each plot results are shown when 1, 2 and 3 seizures are used for training. These are compared to random detection rates, calculated by performance rates when seizures are placed at random times along the EEG, averaged over 1000 trials. In patients 1, 2 and 3 both ANN and SVM are capable of 100% TPR with near zero FPR. Patients 4, 5 and 6 have less stereotyped seizure activity. Detection rates are comparable in Patient 4 for both ANN and SVM, but performance is not far from random. SVM is capable of higher TPRs than ANN in both Patient 5 and 6. ANN performs no better than random guessing in Patient 6. (Continued)



(d) Patient 4



(e) Patient 5



(f) Patient 6

FIGURE 5.4: (Continued)

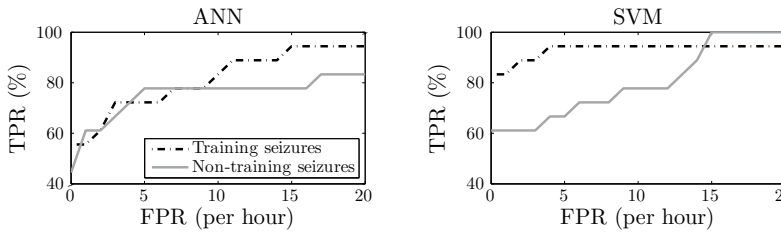


FIGURE 5.5: Average performance on Group 2 data, where the first 3 seizures are used for training and the next 3 used for testing. SVM is shown to outperform ANN when detecting seizures it has been trained on, whilst achieving similar performance to ANN when presented with unseen EEG data.

close to a particular node then its features are very similar to some subset of the training sequence. The distance between the input feature vector and the closest node is stored for thresholding.

5.2.3 SVM Training and Testing

The SVM can be trained on both seizure and non-seizure data because, unlike ANN, the process is not overly affected by over-representation of non-seizure data. Again all seizure data in the training set is used to train the SVM. However to reduce the size of the training set only one third of non-seizure EEG is used. This did not affect results but did speed up the training time.

Because of the complexity of the EEG it is unlikely that seizure and non-seizure data are linearly separable. A non-linear projection, as described in Section 4.1.3, is used. Several projection functions were tested but ultimately a Gaussian radial basis proved best. This is of course not exhaustive testing of SVM but proved sufficient for this investigation, and the question as to which projection function works best is not considered here.

The training of the SVM used the same features as the ANN described earlier. That is, a $4 \times C_{TOT}$ dimensional feature vector consisting of frequency information at different scales for each of the C_{TOT} channels is presented to the SVM.

Testing consists of presenting an input feature vector of the same dimension to the SV, and thresholding the output directly.

5.2.4 Results and Comparisons

Both SVM and ANN are tested using the first 6 seizures in Group 2 data. In both cases all reasonable efforts are made to optimize performance within each patient. In the case of ANN several different configurations, including

alterations to the number, type and layout of nodes, have been studied. Only the best results are reported here.

ROC curves (described in Section 5.1.2) summarizing results for each patient are shown in Figure 5.4 for both ANN and SVM. Each plot shows the performance of the classifier when 1, 2 and 3 seizures are used during training. Also shown are the expected detection rates when the seizures are placed in random locations along the EEG record, averaged over 1000 trials. These are used to give an indication as to how much better than random each of the classifiers perform, given the distribution of the particular EEG record. Recall that the closer the ROC curves are to the top left corner, and the further they are from random, the better the performance.

Patients 1, 2 and 3 are examples in which 100% TPR with near 0 false positives are achievable, both with ANN and SVM. These patients have seizures whose features are both stereotyped and differ significantly from the non-seizure EEG. Both algorithms are capable of differentiating the two. In Patient 3 SVM struggles a little more than ANN when only 1 seizure is used for training, but catches up in performance to ANN once more seizures are used.

Separating seizure and non-seizure activity in Patient 4 is more difficult. The ROC curves for both ANN and SVM is much closer to random than in Patients 1-3. The performance of ANN and SVM is comparable.

SVM outperforms ANN in both Patient 5 and 6. In Patient 5 SVM is capable of reaching a TPR of 100% as compared to 80% with ANN, albeit this occurs at high FPRs[2]. In Patient 6 ANN performs no better than random, whilst SVM is again capable of 100% detection rate when more than 1 seizure is used for training.

In Figure 5.5 a summary of the performance averaged over all patients is presented. In this figure 3 seizures have been used for training and 3 for testing. SVM is very good at detecting seizures it has been trained on, with $> 90\%$ TPR achievable for FPRs as low as 4. This TPR is only achievable for ANN with FPRs as high as 15. For the testing set, both ANN and SVM perform comparably for reasonably low FPRs. This suggests that SVM is superior to ANN at recognizing features it has been trained on, whilst achieving comparable results to ANN with previously unseen data.

Arguably although SVM does better in more of the patients both ANN and SVM perform reasonably similarly in most cases. However from a more subjective point of view the implementation and training of SVM requires many less design considerations, is much faster to train and is less dependent on the training examples than is ANN. The small quantitative advantage that SVM has on ANN is magnified by these qualitative observations.

[2]In all cases high FPRs are expected because little effort has been made to optimize the features used for evaluation. Again, this is because relative rather than best performance is under review.

In summary, the relative ability of ANN and SVM to discriminate seizure from non-seizure EEG in a noisy environment has been evaluated.

Quantitatively SVM outperforms ANN, albeit in many cases the performance is comparable. SVM seems to be particularly apt at identifying features it has been trained on, whilst still maintaining a reasonably good ability to identify unseen data. This is likely so because SVM is trained on both seizure and non-seizure data, whereas only seizure EEG is used to train the SOM. Quantitative results can be seen in Figure 5.4 and Figure 5.5.

Qualitatively SVM is both easier to implement and less prone to the idiosyncrasies of the data. This admittedly could be the reason that SVM outperforms ANN – the best configuration for ANN was not found – but does not change the fact that good results with SVM are easier to achieve.

5.3 Evaluation of Patient Un-Specific Seizure Detectors

A typical EEG machine is capable of recording around 64 channels concurrently, generating gigabytes of information a day. Given that 95-99% of this data for epileptic patients appears useless for diagnosis [52] the manual review process is not only tedious, time consuming and expensive, but also prone to human error.

In this section a review of current published software-based automated patient un-specific seizure detectors that use scalp EEG data is conducted[3]. It is an evaluation of the detection algorithms used for review of EEG data, where the requirement is that the detector identifies a seizure at least once during its progress, be this at the beginning, middle or end.

Since standardization is the purpose of this study, although many hundreds of algorithms exist only a subset of the most developed were implemented. The selection was based on several criteria, including

- *Non-specific epilepsies*: This is an evaluation of a broad range of seizure

[3]This review does not look at spike detection even though it is also an important aspect in the diagnosis of epilepsy. Spike detection is a problem in its own with very different considerations than those of seizure detection.

types. It specifically excludes work for which performance was only reported on absence seizures, such as that in [178], [175], [177] and [8].

- *Volume and type of EEG database*: Only work validated through testing on a large database is used. Furthermore preference is given to algorithms that make a clear distinction between the training and testing datasets. Preference is also given to those whose database is recorded in the international 10-20 electrode configuration system.

- *Adult EEG*: The current database described in Section 5.1.1 does not include neonate data which differs significantly from adult EEG available here [52].

- *Normalized feature extraction*: Preference is given to algorithms whose features are normalized. This avoids implementation of algorithms that depend on the specifics of the EEG measurement system.

Algorithms that are not implemented because they do not comply to these criteria may be compared to the benchmarks presented here once the database becomes publicly available.

Of the algorithms selected only the latest published implementation is used, under the assumption that these improve on previous results. All efforts are made so that our work here remains true to the published material, including re-sampling and filtering of the database as well as selection of EEG channels that coincide with those used. Unlike Section 5.2 this specifically excludes the optimization of results.

A total of 4 detection systems are implemented, three of which are commercially available solutions. Each algorithm is tested on Group 1, Group 2 and Group 3 databases. Each algorithm and its results is summarized individually next, in chronological order of publishing.

5.3.1 Algorithm 1: `Monitor`

5.3.1.1 Algorithm Description

`Monitor` is a commercially available detector whose origins begin in the early 1980s. Its performance has been shown inferior to many modern detectors, but due to its historical role in the problem of seizure detection it is included in this analysis as reference.

The algorithm is described in detail in [50] and [51] and summarized in Figure 5.6. Readers may also find it useful to refer to relevant material in [55], [52], [54], [139], [140] and [141]. `Monitor` is a multi-feature rule based classifier. The three normalized features are extracted from each EEG channel after it is de-composed into segments (see [54]). These features include the relative amplitude of a current epoch to past averages (RA), a coefficient of variation that measures the variability in a signal (COV) and the average duration of these segments relative to the background. The expert system incorporates

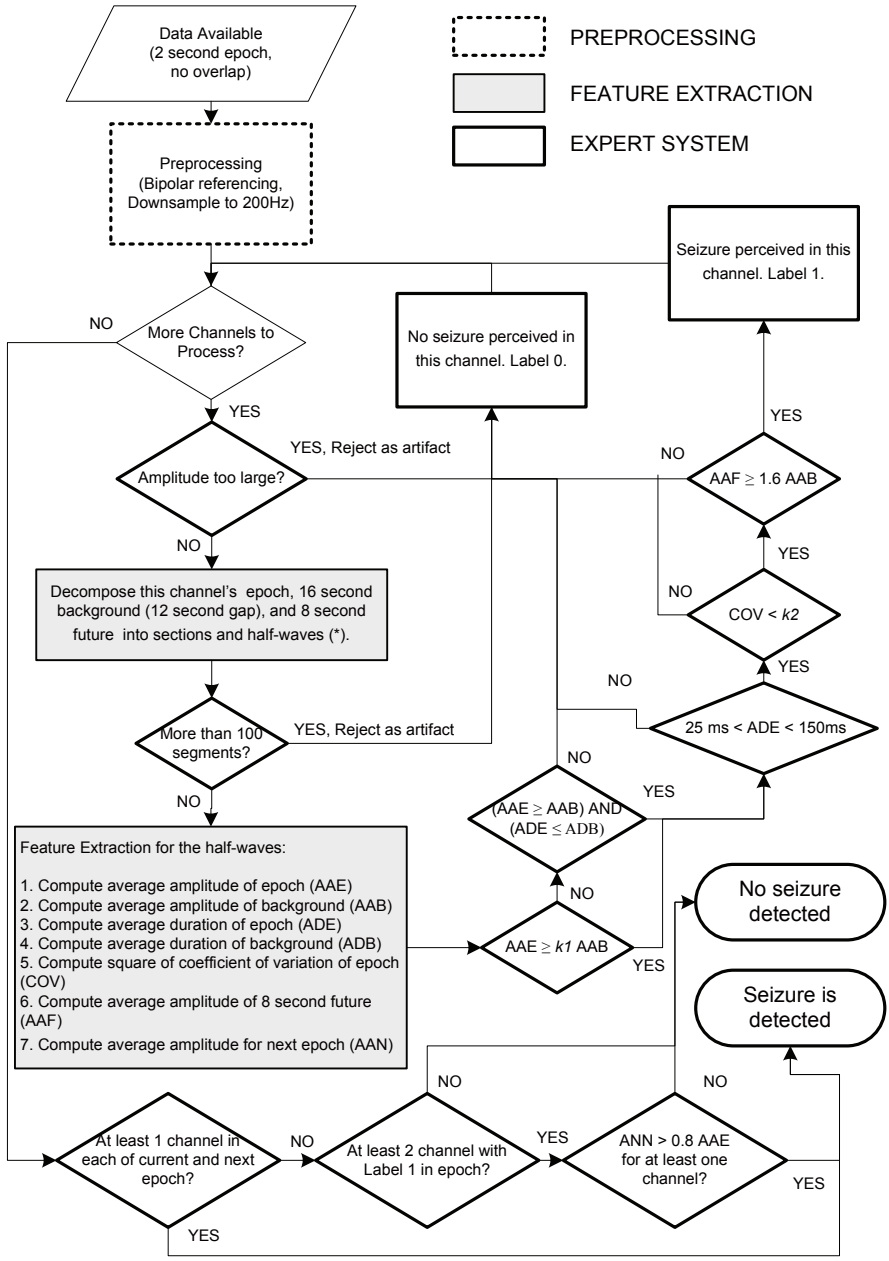


FIGURE 5.6: `Monitor` detection algorithm summary. In the latest published original implementation $k1 = 3$ and $k2 = 0.36$. The ROC curves in Figure 5.7 are calculated by varying $k1$ (the value of $k2$ did not alter results significantly).

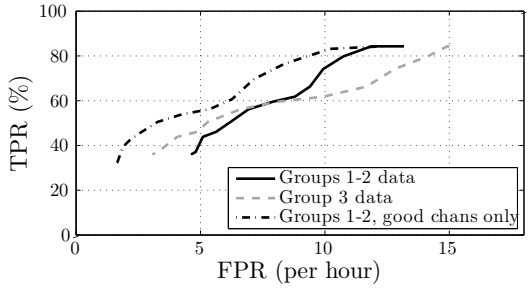


FIGURE 5.7: `Monitor` ROC, generated by varying $k1$ in Figure 5.6. A maximum TPR of 84.3%is achievable only when the FPR is as high as 11.9per hour. A more moderate FPR of 8reduces TPR to 60%. However these calculations included 2 patients for whom FPR was capped at a maximum of 15 per hour, and results are expected to be affected by this. The FPR in Group 3 is somewhat higher but overall not inconsistent with the FPR in Groups 1 and 2.

limited temporal and spatial contextual information by requiring detections to occur over more than one channel or more than one epoch.

5.3.1.2 Results

The algorithm was implemented verbatim from [50] and [51], using MATLAB version 7.1 and run on a standard 1400Hz Intel processor running Microsoft Windows XP operating system. Code can be made available upon request. Since this is a pure thresholding technique no training was required and the algorithm was applied to all available EEG data. Prior to analysis data were pre-filtered and downsampled to 200Hz to coincide with the specifics of the dataset used in the publications.

The results are summarized in Figure 5.7. The ROC curves are generated by varying $k1$ shown in Figure 5.6, responsible for thresholding RA $= \frac{\text{AAE}}{\text{AAB}}$. $k2$ was also varied but results are less significantly affected by this parameter.

`Monitor` has been evaluated many times with results that vary wildly dependent on the database. Originally a TPR of roughly 73% with an FPR of ≈ 1 per hour is reported in [51]. Since then, TPR of 74.4% with FPR of 3.02 per hour in [44], TPR of 31% with FPR of 0.1 per hour in [194] and TPR of 43% with FPR of 0.78 per hour in [156] are reported. Here a maximum TPR of 84.3%is achievable only when the FPR is as high as 11.9 per hour. A more moderate FPR of 8reduces TPR to 60%. It is difficult to know which results to believe because little context about each implementation is known. Results here can be used as reference since all algorithms are tested on a common dataset.

The discrepancies are most likely explained by the differences in recording

equipment. The data used in [51] are recorded with an EEG machine whose sampling rate is much lower. Although data here are down-sampled to coincide with the published results the waveforms could look significantly different because of other differences in the hardware (e.g., anti-aliasing filters, noise levels, etc). Alternatively, the recording environment for data included here may contain significantly higher levels of noise.

When Group 3 data are used the FPR on average increases, although not by much. Group 3 database is not large enough to be conclusive, but results indicate that initial calculations of FPR using databases taken from epileptic subjects are reasonably accurate.

5.3.2 Algorithm 2: `CNet`

5.3.2.1 Algorithm Description

`CNet` is one of the first algorithms to use ANN. The detection algorithm is described in detail in [44] and [45] and summarized in Figure 5.8. It is a two-stage approach in which an initial evaluation is applied to groups of channels averaged together. If a preliminary detection occurs in any one group of channels further evidence from individual channels is gathered before a true detection is reported.

Spectral information from both averaged as well as single EEG channels is extracted by computing spectrograms of 4-second, zero overlap EEG epochs. Prior to computation each epoch is first filtered using a matched filter that selectively attenuates frequencies least prevalent in seizure train data. A 2DFFT[4] is applied to the spectrogram so that spectral peaks are emphasized. The distance of this feature vector to each node is presented to a trained SOM (see Section 4.1.2 for discussion on ANN and SOM). This distance (LDE) is normalized by a 30 minute moving background (LDB) before thresholding.

The SOM is trained on pre-selected epileptic examples (in the order of 100 4-second epochs) chosen to demonstrate a wide variety of phenomena. More than one epoch from a single seizure may be used. An advantage of SOM is that no non-seizure data are necessary and thus over-representation of non-seizure data is not a problem.

Other than ANN training the expert system consists of epoch rejection based on relative EEG amplitudes between current, past and future epochs. Temporal information is incorporated by requiring detections to occur at least twice in 15 seconds, but the only form of spatial contextual information is in the manner that groups are formed in the preliminary detection stage.

[4] A 2DFFT (2 dimensional FFT) is computed by taking the FFT of each row of a matrix, then the FFT of each column of the result. More details can be found in [44] and [45].

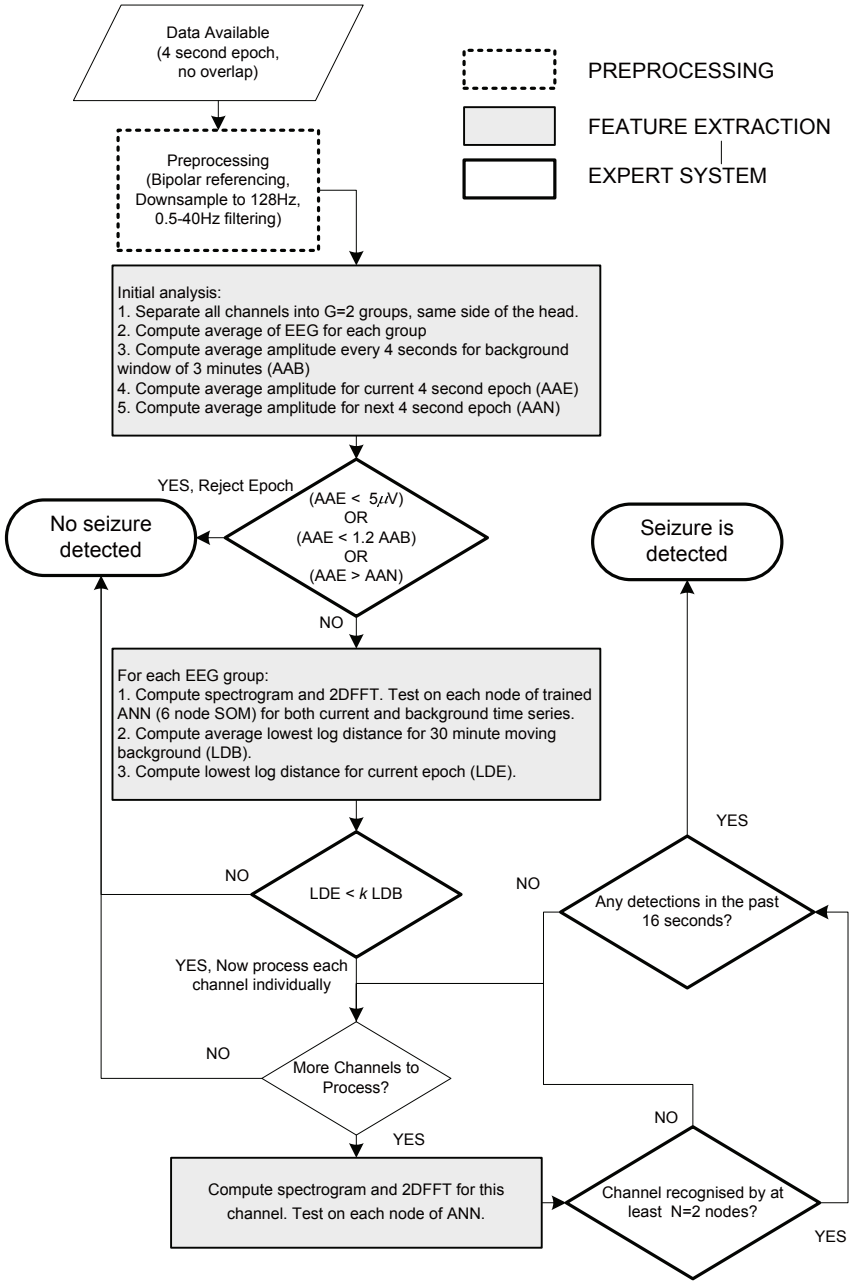


FIGURE 5.8: `CNet` detection algorithm summary. The ROC curves in Figure 5.9 are calculated by varying k. In the latest published original implementation $k = 0.8$.

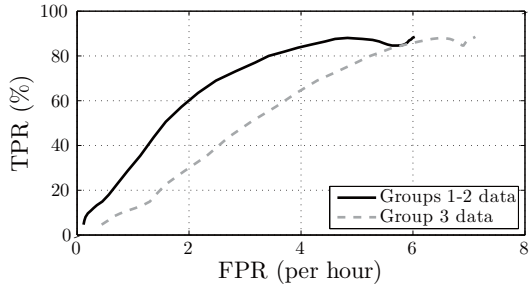


FIGURE 5.9: `CNet` ROC, generated by varying k in Figure 5.8. For Group 1 and 2 data a maximum TPR of 85.4%is achieved coinciding with an FPR of 4.4per hour. No patients exceeded an FPR of 15 for these calculations. The FPR in Group 3 are again slightly higher.

5.3.2.2 Results

The algorithm was implemented using MATLAB version 7.1 and run on a standard 1400Hz Intel processor running Microsoft Windows XP operating system. Code can be made available upon request. Since the original trained ANNs were not available, performance was estimated by averaging 10 trials. Each trial was trained using seizure data extracted from half the patients randomly selected from Group 1 and 2 and tested on the remaining data. *Under no circumstances is a patient whose data are included in the training set used during testing.*

The average results are presented in Figure 5.9. These curves are generated by varying k in Figure 5.8. This parameter determines how different the current epoch must be from the 30 minute moving background to be detected as a potential seizure.

Original reported performance is an average TPR of 90-93% and an FPR between 1.25-1.39 per hour. Although the TPR is only slightly lower in Figure 5.9 – a maximum TPR of 85.4%– this occurs at a significantly higher FPR of 4.4per hour. This could again be a result of the differences in the recording environment, seeing as their databases are recorded at much lower sampling rates and their evaluation of `Monitor` resulted in many lower FPRs than here. `CNet` is evaluated independently in [194] and although only one value is reported the performance as well as their database is more consistent with those here.

The FPRs of the non-epileptic Group 3 data are slightly higher.

5.3.3 Algorithm 3: `Reveal`

5.3.3.1 Algorithm Description

`Reveal` is a commercially available seizure detector designed to target rhythmic activity. The detection algorithm is described in detail in [194] and summarized in Figure 5.10. Other relevant material includes [191] and [192].

The *matching pursuit algorithm* is applied to 2 second windows (1 second overlap) to extract features known as *Gabor atoms*, described in Section 3.4, used to describe the two most prominent rhythmic components in the epoch (see Section 3.4). The amplitude, duration and frequency of each atom is then stored for analysis.

Unlike `Monitor` and `CNet` in which relative values are computed using fixed background windows, `Reveal` adaptively selects the best separation between background and current epochs by analyzing the similarities in the time series. A seizure is reported only when the difference between background and current epochs is large enough. To determine this difference two rule-based tests and 4 different ANN tests are performed. Rather than lumping all rules in a single ANN, as with `CNet`, each ANN targets smaller rules and is trained separately. In this way all temporal and spatial context is lumped within the 2 rule-based and 4 ANN tests.

5.3.3.2 Results

The algorithm was tested using a trial evaluation of `Reveal` version 2007.12.07, downloadable from `http://www.eeg-persyst.com/`. This was desired over self implementation because of the complex and lengthy training periods, prone to error if repeated. The software is designed to support data recorded using the Compumedics equipment used here, and no transformations prior to testing are necessary. Since training is provided by the suppliers of the software all EEG data is used for evaluation.

The average results are presented in Figure 5.11. These curves are generated by varying seizure perception likelihood k shown in Figure 5.10.

The reported performance in [194] is a maximum TPR of about 84% at a very low FPR of 0.6 per hour. The maximum TPR in Figure 5.11 is 77.5%with a corresponding FPR of 2.8per hour. This FPR can be reduced to roughly 1per hour when visually unreliable channels are manually removed, as shown in green. However to be fair if this leniency is afforded here it should also be applied to other algorithms. For example, removing bad channels in `Monitor` reduces its FPR, as shown in Figure 5.7. `CNet` remains relatively unaffected and this curve is not shown. Since these unreliable channels were often the result of loose or disconnected channels – an artifact that is not difficult to identify – `Reveal` should consider automated removal.

Again, the FPRs of the non-epileptic Group 3 data are slightly higher.

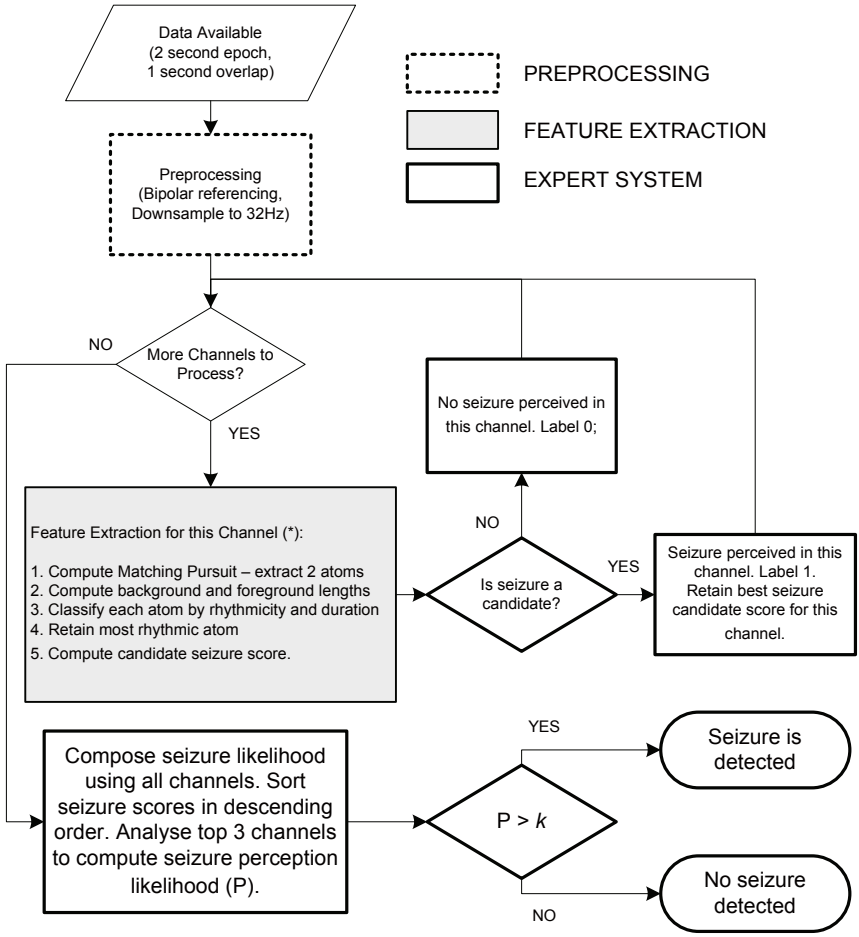


FIGURE 5.10: `Reveal` detection algorithm summary. Much of the complexity in this algorithm has been shifted to the training of 6 individual rules, 4 of which are ANNs, that target different aspects of seizure detection and include the spatial and temporal context. The seizure perception threshold k is a user defined variable ranging from $0 - 1$.

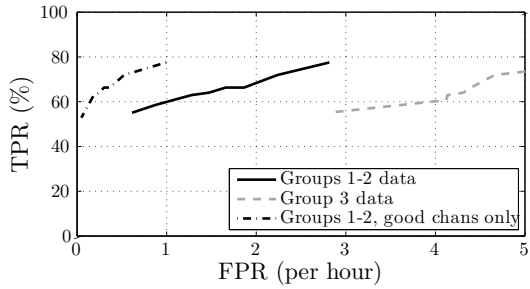


FIGURE 5.11: `Reveal` ROC, generated by varying k in Figure 5.10. The maximum achievable TPR is 77.5%with a corresponding FPR of 2.8per hour, a poorer performance than the originally published results in [194] that give a maximum TPR of 84% at a very low FPR of 0.6 per hour. However by manually omitting unreliable channels the FPR can be reduced to about 1per hour, as shown.

5.3.4 Algorithm 4: `Saab`

5.3.4.1 Algorithm Description

The `Saab` algorithm, named here after the principal author, is designed by the same research group that also developed `Monitor`. The detection algorithm is described in detail in [156] and summarized in Figure 5.12. Readers may also be interested in an intra-cranial EEG implementation discussed in [57].

Similar features (namely RA and COV) are used, but are applied to the coefficients of wavelet decimation at different scales rather than the raw data. A third feature, RSE, computes the energy in each of the wavelet scales relative to the energy at all scales. The 5 scales used are representative of $50-100$Hz, $25-50$Hz, $12-25$Hz, $6-12$Hz and $3-6$Hz. Scales 2, 3 and 4 are used to characterize epileptic events because most seizures contain frequencies in these ranges. The remaining scales are used to identify artifact.

The innovation in `Saab` is the pure probabilistic approach to classification. For each channel the conditional probability that a seizure occurs given the features RA, COV and RSE ($P(seizure|features)$) is computed using Bayes' theorem

$$P(seizure|features) = \frac{P(features|seizure)P(seizure)}{P(features)}. \qquad (5.4)$$

The terms on the right are trained on observations from seizure data. RA, COV, RSE in each scale is collected from seizure and non-seizure training examples, the total range is divided into 5 distinct levels and probabilities of each are recorded. Detailed explanations of the training process are found in [156].

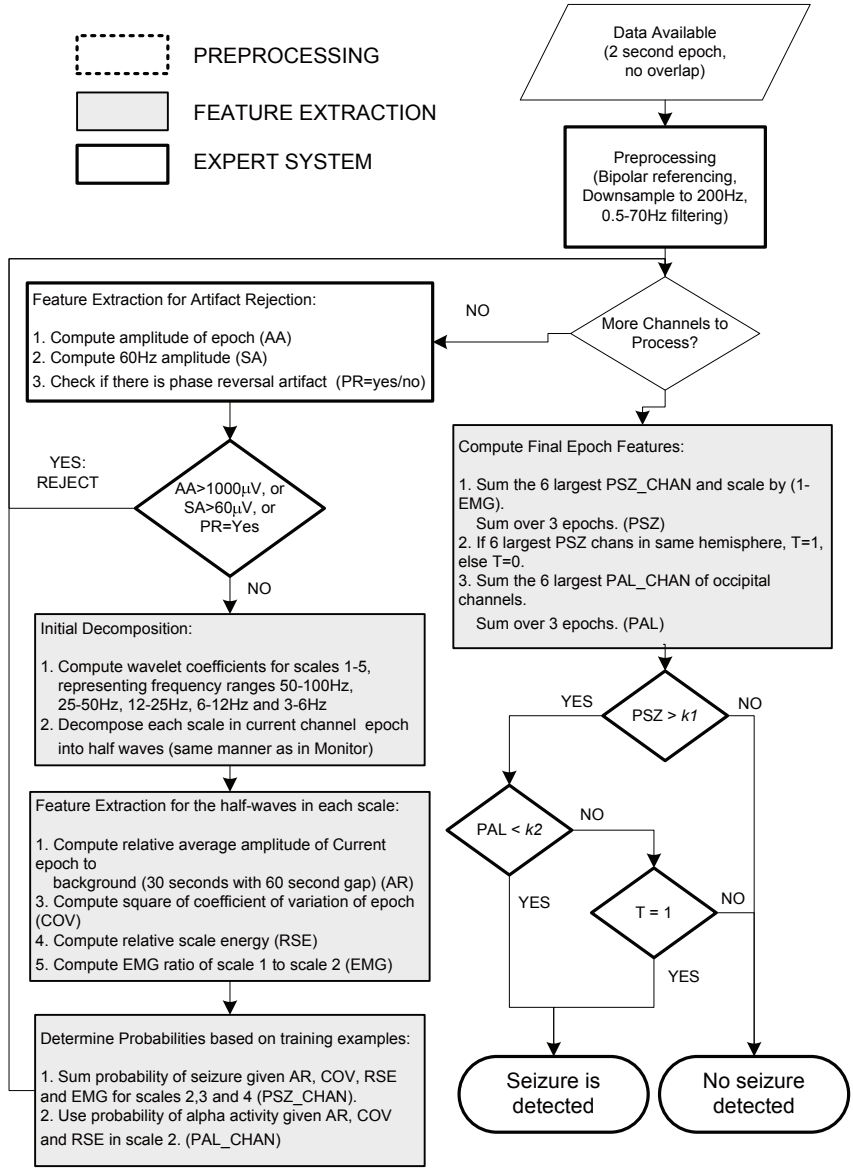


FIGURE 5.12: `Saab` detection algorithm summary. In the original published implementation $k1 = 4$ and $k2 = 2$. The ROC curve in Figure 5.13 is calculated by varying $k1$ (the value of $k2$ did not alter results significantly).

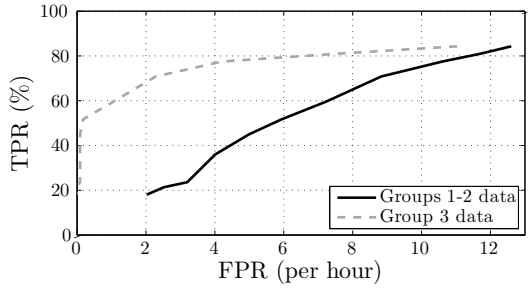


FIGURE 5.13: `Saab` ROC, generated by varying $k1$ in Figure 5.12. A maximum TPR of 84.3%is achievable only when the FPR is as high as 12.6per hour. A more moderate FPR of 7reduces TPR to 60%. No patients had their FPR higher than 15 per hour, but several were close. Surprisingly the FPR in Group 3 is significantly lower than that in Group 1 and 2, a result not observed with any other algorithm.

The expert system incorporates more sophisticated artifact rejection that ignores epochs with disconnected or loose electrodes as well as those that have abnormally large amplitudes. Alpha rhythm data are also rejected through the computation of the conditional probability $P(alpha|features)$ trained in a similar manner as in Equation 5.4. Spatial contextual information is incorporated by summing $P(seizure|features)$ over 6 channels with highest probabilities and placing importance on the location of these channels. In this way evidence is amassed over several channels, providing a likelihood rather than a probability that a seizure is present in any one epoch.

5.3.4.2 Results

The algorithm was implemented using MATLAB version 7.1 and run on a standard 1400Hz Intel processor running Microsoft Windows XP operating system. Like `CNet` a training phase is required – the same 10 trial combinations used in Section 5.3.2 are used for averaging. However unlike `CNet` the training data included non-seizure as well as seizure data.

The average results are presented in Figure 5.13, generated by varying the parameter $k1$ shown in Figure 5.12, responsible for thresholding overall seizure likelihood PSZ. Varying $k2$, the threshold for alpha activity likelihood PAL, was also attempted but results are less significantly affected by this variable. No real difference in results is observed when unreliable channels are manually removed, and thus this curve is not shown.

The results reported in literature are as high as TPR of 76% with an FPR of 0.34. These results vary significantly from those observed here, where for an equivalent TPR at least 10 false positives per hour are reported. Whilst an improvement on `Monitor` is made, as reported in [156], this improvement

does not compare with the superior performance of CNet or Reveal. Authors have been extended an invitation to submit their own code with their own training sequences, but no reply was received.

The discrepancies may, like in Monitor, be explained by the EEG acquisition which may be more prone to noise in our database. Another possible reason is the difficulty in selecting the data used to train the algorithm. Like CNet this algorithm required 10 different trials with 10 different training sequences. However in contrast to CNet in which performance over the 10 trials was relatively consistent, this was not so for Saab. Furthermore training relies on both epileptic and non-epileptic data and suffers from over-representation problems such as those of ANNs, avoided by CNet through the use of SOM. Saab depends heavily on the data used in the training process. CNet, on the other hand, is more robust to these inclusions.

This is the only algorithm for which results vary significantly for non-epileptic database; in this case a *lower* FPR is achieved for Group 3.

5.3.5 Comparisons and Conclusions

Good performance for a patient un-specific seizure detector, where detection anytime during a seizure is sufficient, is a moderately high TPR with a moderately low FPR (see Table 5.3). What moderate means is left up to the individual. Here it is proposed that an FPR of 6 per hour is reasonable since this rejects at least 90% of the record that must be reviewed. This is significantly better than, say, an FPR of 15 per hour that rejects only 75% of the record. On the matter of sensitivity if experts can only agree with each other 80% of the time (see Section 5.1) then anything above this value outperforms human classification.

The relative performance of all algorithms can be compared from Figure 5.7, Figure 5.9, Figure 5.11 and Figure 5.13. From these it is concluded that CNet provides the highest TPR, although if a slightly lower TPR is acceptable FPR can be reduced significantly by Reveal. Monitor and Saab produce less acceptable FPRs for this database.

However the above observations do not provide a complete picture. These results are averages where each plotted value is calculated using the same threshold for all patients. Although the averages are indicative of general performance trends they may be un-representative when the inferior performance of any one patient may be improved dramatically simply by adjusting a user-defined threshold. This type of patient specificity can be applied to all algorithms in a fair manner, and is presented in Table 5.4 and Table 5.5.

Table 5.4 reports the best achievable TPR when a maximum FPR of 6 per hour is allowed. When an FPR of 6 is not possible the lowest FPR above this threshold and its respective TPR is reported. This gives an idea of how good performance is for a reasonable number of false detections. The results for CNet and Reveal remain relatively unchanged because all patients in both have average FPRs lower than this limit. However the results for Monitor and

`Saab` change significantly because tuning to a particular patient can improve results significantly. For example, prior to tuning `Monitor` was shown to have an average TPR of 60%with an FPR of 8. After tuning, `Monitor` can achieve a TPR of 68.5%with an FPR of 5.87. Some patients, namely 1, 2 and 5 in Group 1, exhibit very high FPRs, close to and exceeding 15. These patients skew results and the algorithm performs adequately for the majority of patients. Similar observations hold for `Saab`.

Table 5.5 reports the best achievable TPR when a maximum FPR of 15 per hour is allowed. This gives an idea of the best possible TPR of the algorithm, but also how FPR is affected by this performance. Again `CNet` and `Reveal` remain relatively unchanged because their average FPRs are well below the maximum allowable 15 per hour. However in both `Monitor` and `Saab` significant improvements can be made, at the expense of a much increased FPR.

In addition Table 5.6 shows a patient specific summary of FPR in Group 3. These FPR correspond to an average TPR of 70% or above[5]. With a similar performance metric for all patients the FPR may be compared on a relatively even playing field. Again, it is evident that very high FPRs occur in `Monitor` to achieve equivalent results. This is not the case for `Saab` where Group 3 yields much lower FPRs than Group 1 and 2. The Group 3 database must be expanded before definitive conclusions can be made.

The specificities are provided in these tables for the reader that finds these numbers more intuitive. In any case some new information is provided. Whereas to calculate FPR 60 second blocks are used, specificity S does not do this. For example in Table 5.4 `CNet` achieves an average FPR of about double that of `Reveal` (4.03versus 1.84), whilst the specificity of `CNet` is half that of `Reveal` (0.84 %versus 2.26%). From this information we can deduce that `CNet` generates false positives that are short in duration and far apart, whereas those in `Reveal` are generally lumped together.

In summary, Figures 5.7, 5.9, 5.11, 5.13, Table 5.4, Table 5.5 and Table 5.6 demonstrate that

- The overall best TPR is achieved by `CNet`.
- The overall best FPR, with a sufficiently high TPR, is achieved by `Reveal`.
- The performance of `Monitor` and `Saab` can be significantly improved on a patient specific basis, but overall performance remains inferior to `CNet` and `Reveal`.

[5]A TPR of 70% was selected because this is a reasonable performance that is achieved for most patients by most algorithms.

The advantages of using a common dataset for evaluation become clear – had the reported performances been believed unconditionally these observations would not be evident. A fairer evaluation would include data recorded with different equipment, so that biases observed by algorithms such as `Monitor` and `Saab` that are sampled at much lower frequencies may be removed.

5.4 Evaluation of Onset Seizure Detectors

For many applications it is important to detect the beginning of the seizure. In this section we evaluate detection strategies based on how quickly a seizure is marked once it has begun. To narrow scope we concentrate on the class of problems where detection of a seizure must occur before clinical onset, applicable to situations such as the delivery of anti-epileptic drugs at critical moments. Only group 4 data are used because onset detectors of this nature are most often applied to intra-cranial data, where a much longer lead-up to seizures is frequently visible.

The *onset delay* metric described in Section 5.1.2 is defined here relative to the *clinical onset* of a seizure. The electrographic onset does not necessarily coincide with the clinical onset and may occur seconds or minutes earlier. A positive onset delay indicates a detection before the clinical onset, whilst a negative one indicates detection after clinical onset. We re-name this metric the *warning time*.

5.4.1 Feature Extraction

A few complete algorithms for the detection of intra-cranially measured seizures have been reported. See, for example, the work in [128], [127] and [59]. However the focus of the study here is to the evaluate how well the different types of statistics presented in Chapter 3 can detect seizure onsets[6]. We select one time domain feature (cross correlation), one frequency domain feature (PSD), one time-frequency based method (wavelets) and one non-linear statistic (correlation dimension).

[6] A study of this nature exists in [74] and [75]. However the methods of these studies are not available, and furthermore TPR and FPR metrics are not included. Here we aim to address this.

Patient	MONITOR TPR (%)	MONITOR FPR (per hour)	MONITOR S (%)	CNET TPR (%)	CNET FPR (per hour)	CNET S (%)	REVEAL TPR (%)	REVEAL FPR (per hour)	REVEAL S (%)	SAAB TPR (%)	SAAB FPR (per hour)	SAAB S (%)
Group 1 – Scalp EEG data, < 6 seizures per patient.												
1	100	15	10.91	100	4.87	1.23	100	0	0	86.67	5.41	1.35
2	25	15	23.09	100	4.18	1.06	100	3.39	3.76	97.5	4.04	1.41
3	50	5.56	1.16	100	4.45	1	100	0.04	0.02	20	7.25	5.74
4	66.67	4.47	1.15	96.67	4.59	0.91	100	0	0	100	5.05	1.72
5	100	1.41	0.18	100	1.92	0.3	100	0.31	0.26	20	4.61	1.45
6	75	4.95	1.24	67.5	5.64	1.13	100	0.07	0.03	100	2.51	0.39
7	100	5.56	4.53	33.33	1.92	0.34	100	0.14	0.1	73.33	5.46	3.87
8	33.33	5.75	3.27	66.67	2.02	0.33	33.33	3.03	2.12	6.67	7.32	4.41
9	40	14.12	11.09	96	5.81	1.31	80	0.85	0.41	38	5.62	4.14
10	66.67	5.06	1.07	66.67	1.73	0.32	66.67	0.05	0.04	0	0	0
11	100	2.15	0.28	100	1.77	0.26	100	0	0	100	3.87	1
12	100	3.22	1.65	10	5.44	1.02	100	13.81	19.35	90	3.13	0.63
13	100	4.79	1.55	88	2.89	0.61	60	0.36	0.14	8	3.17	0.83
14	100	5.57	1.54	100	4.79	0.84	100	0	0	100	3.63	0.81
15	50	4.07	1.61	100	4.03	0.95	100	0	0	0	0	0
Group 2 – Scalp EEG data, 6 – 10 seizures per patient.												
1	100	5.15	1.17	67.14	5.99	1.2	100	0	0	35.71	4.8	3.52
2	100	5.56	2.54	80	3.64	0.64	80	0.54	0.16	72	5.01	1.57
3	40	1.4	0.39	99	5.96	1.48	100	15	20.43	58	5.53	2.83
4	75	4.3	0.96	75	5.49	1.21	75	1.03	0.55	80	5.41	1.98
5	71.43	5.97	4.5	100	3.34	0.66	85.71	0.12	0.07	85.71	5.33	1.87
6	40	4.19	0.9	100	5.53	1.15	0	0	0	21	4.29	1.56
Average Performance Group 1 and 2												
	68.5	**5.87**	**3.56**	**86**	**4.1**	**0.85**	**77.5**	**1.84**	**2.26**	**54.5**	**4.35**	**1.96**

TABLE 5.4: Patient specific results for the evaluation of general seizure detectors, described in Section 5.3. Here the best TPR when the FPR is limited to a maximum of 6 per hour is reported for each patient. If an FPR of 6 cannot be achieved, the lowest FPR is used. Unreliable channels have not been removed. These results give an idea of how TPR is affected when the FPR is at a reasonably low level. It also highlights how patient specificity can improve the results shown in ROC curves taken from averages. For example, `Monitor` with an average TPR and FPR of 60%and 8are improved to 68.5%and 5.87when patient specific thresholds are defined. Furthermore some patients are responsible for many FPR, and if excluded `Monitor` may exhibit results comparable to `CNet`.

Patient	MONITOR TPR (%)	MONITOR FPR (per hour)	MONITOR S (%)	CNET TPR (%)	CNET FPR (per hour)	CNET S (%)	REVEAL TPR (%)	REVEAL FPR (per hour)	REVEAL S (%)	SAAB TPR (%)	SAAB FPR (per hour)	SAAB S (%)
Group 1 – Scalp EEG data, < 6 seizures per patient.												
1	100	15	10.91	100	4.87	1.23	100	0	0	96.67	11.6	3.26
2	25	15	23.09	100	4.18	1.06	100	3.39	3.76	100	12.89	6.36
3	100	14.92	5	100	4.45	1	100	0.04	0.02	95	14.03	12.38
4	100	13.34	11.79	100	6.83	1.59	100	0	0	100	14.86	6.12
5	100	1.41	0.18	100	1.92	0.3	100	0.31	0.26	90	13.69	6.37
6	75	11.65	5.12	100	7.87	1.71	100	0.07	0.03	100	2.51	0.39
7	100	11.05	8.12	33.33	1.92	0.34	100	0.14	0.1	100	13.96	9.21
8	66.67	13.06	7.95	66.67	2.02	0.33	33.33	3.03	2.12	53.33	12.78	8.02
9	40	14.12	11.09	100	7.08	1.75	80	0.85	0.41	82	14.48	11.92
10	100	13.46	4.12	66.67	1.73	0.32	66.67	0.05	0.04	73.33	12.96	11.1
11	100	6.71	1.18	100	1.77	0.26	100	0	0	100	13.84	6.55
12	100	3.22	1.65	100	9.79	2.54	100	13.81	19.35	90	3.13	0.63
13	100	7.94	4.17	88	2.89	0.61	60	0.36	0.14	8	3.17	0.83
14	100	5.57	1.54	100	4.79	0.84	100	0	0	100	13.41	3.79
15	50	4.07	1.61	100	4.03	0.95	100	0	0	60	13.51	13.03
Group 2 – Scalp EEG data, 6 – 10 seizures per patient.												
1	100	9.24	2.19	72.86	6.76	1.36	100	0	0	87.14	11.2	9.52
2	100	7.02	3.67	81	6.86	1.59	80	0.54	0.16	89	13.28	5.09
3	40	1.4	0.39	100	6.58	1.72	100	15	20.43	79	12.68	8.82
4	75	13.17	16.41	75	5.49	1.21	75	1.03	0.55	90	11.97	5.8
5	71.43	13.32	16.18	100	3.34	0.66	85.71	0.12	0.07	100	11.19	4.64
6	60	13.65	4.32	100	5.53	1.15	0	0	0	31	7.32	3.29
Average Performance Group 1 and 2												
	75.3	**9.92**	**6.7**	**89.4**	**4.8**	**1.07**	**77.5**	**1.84**	**2.26**	**77.2**	**11.36**	**6.53**

TABLE 5.5: Patient specific results for the evaluation of general seizure detectors, described in Section 5.3. Here the best TPR when the FPR is limited to a maximum of 15 per hour is reported for each patient. Unreliable channels have not been removed. These results give an idea of how high the TPR can rise to, irrespective of FPR. Alternatively, they give an idea of how high FPR can go. `CNet` and `Reveal` are shown to be consistent in keeping their FPR low, whilst keeping TPR high. On the other hand `Monitor` and `Saab` show that high TPRs can only be achieved with very high FPR.

	MONITOR		**CNET**		**REVEAL**		**SAAB**	
Av. TPR	73%		76%		72%		71%	
Av. FPR	9.65 p/h		2.83 p/h		2.24 p/h		8.83 p/h	
Pat.	**FPR** (p/hour)	**S** (%)	**FPR** (p/hour)	**S** (%)	**FPR** (p/hour)	**S** (%)	**FPR** (p/hour)	**S** (%)
Group 3 – Scalp EEG data, Non-epileptic Patients.								
1	15	5.22	4.53	0.76	0	0	3.54	0.42
2	7.09	1.26	7.96	2.53	4.25	3.15	2.98	0.59
3	6.17	1.3	3.77	1.02	5.55	3.54	2.53	0.91
4	15	3.4	6.25	1.31	1.11	0.74	1.28	0.1
5	12.1	2.44	10.09	1.9	0.9	0.36	1.66	0.32
6	6.99	1.07	6.03	1.04	9.6	7.35	4.02	0.37
7	15	5.94	0	0	0	0	15	9.79
8	15	16.12	4.7	0.78	4.68	2.47	10.77	2.42
Average Performance Group 3								
	11.54	**4.59**	**5.42**	**1.17**	**3.26**	**2.2**	**5.22**	**1.87**

TABLE 5.6: Patient specific results for the evaluation of general seizure detectors, described in Section 5.3. Here the FPR and specificity for group 3 data is reported when the average TPR of each algorithm is higher than (but as close to as possible to) 70%. This value is chosen for comparison because it is a reasonable (although low) performance that is nevertheless achievable by all algorithms. Particularly in the case of `Monitor` some patients exhibit much higher FPR than others, capped at a maximum of 15 per hour.

All methods use 10 second windows with an overlap of 9 seconds. A much longer window than that in Section 5.3 is used because it is expected that a much slower onset of seizures is observed intra-cranially. This makes the computed statistics more robust but can result in shorter warning times. A 3-40Hz bandpass filter is applied, and data are normalized to unit variance (see Equation 3.13) prior to analysis. All features are extracted as described next, and then thresholded in a patient specific manner.

5.4.1.1 Cross Correlation (XCORR)

The methods used to compute the cross correlation between two signals are found in Section 3.3.1. Specifically, the normalized cross correlation defined in Equation 3.19 is used. This statistic has a maximum value of 1 when two signals are identical.

The normalized cross correlation is computed for all 30 possible combinations of the 6 channels in each patient. Only the maximum cross correlation, observed across all channels and all time delays, is recorded for thresholding in each 10 second window.

5.4.1.2 Power Spectral Density (PSD)

The methods used to compute the power spectral density (PSD) of a signal can be found in Section 3.3.2. Specifically the PSD is computed as given by Equation 3.31. $P = 10$ averages are used over a 10 second window, resulting in a 1Hz resolution of the computed PSD. A Hanning window of 1 second long

(512 samples at sampling rate $F_s = 512$Hz) is applied to each segment prior to computation of its PSD.

To obtain a normalized statistic the ratio of low to high frequency energy content is computed. This is done by summing the PSD between frequency range 3-8Hz and dividing by the sum over the 9-30Hz range. It is expected that during a seizure there is a characteristic shift, typically to lower frequencies, in the relative energy observed [74]. Note that because we are working with ratios it is not necessary to normalize the computed PSD as per Equation 3.30.

In each 10 second window the ratio is computed for each of the 6 available channels. The largest ratio is retained for thresholding.

5.4.1.3 Wavelet Analysis (WAV)

The methods used to compute wavelet coefficients of a signal at different scales are exactly those used in Figure 3.18(a) and described in Section 3.3.3. To get an analogous statistic to the PSD ratio of low to high frequencies, scales $m = 5 - 6$ (corresponding to frequencies 2-8Hz) are summed and divided by the sum of scales $m = 3 - 4$ (corresponding to 8-32Hz). Thus this statistic targets (qualitatively) the same phenomena as the PSD statistic, but in a different way.

Again, in each 10 second window the ratio is computed for each of the 6 available channels. The largest ratio is retained for thresholding.

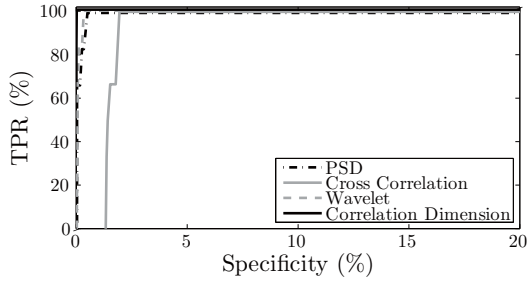
5.4.1.4 Correlation Dimension (CD)

The methods use to compute the correlation dimension of a signal can be found in Section 3.3.4. First a time-delay reconstruction as in Equation 3.35 is performed using $\tau = 30$ samples (0.06 seconds at 512Hz sampling) and dimension $\hat{n} = 4$. The correlation sum is then computed (Equation 3.38) and the gradient in Equation 3.39 is used as an estimate of correlation dimension. $\epsilon_{lower} = 0.1$ and $\epsilon_{upper} = 2$ are used for computation, consistent across all data because it has been normalized to unit variance prior to analysis.

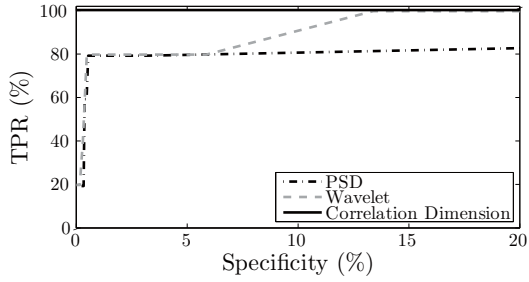
Unlike other statistics the correlation dimension is computed on all 6 channels at once, i.e., all channels are lumped into the one calculation. The rationale behind this is that the correlation dimension requires more data to be robustly computed; moreover the complexity of a system should be viewed as a whole and not on individual channels. This is done for each 10 second window.

5.4.2 Results and Comparisons

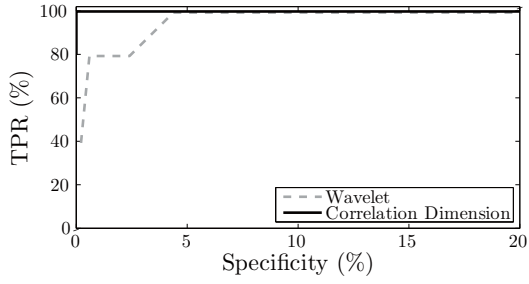
The features were extracted from the 3 datasets in group 4 EEG records and visually inspected for characteristics during a seizure. With the exception of the cross correlation all statistics are capable of detecting at least part of the majority of seizures. In fact these results are quite good compared to those



(a) Patient 1

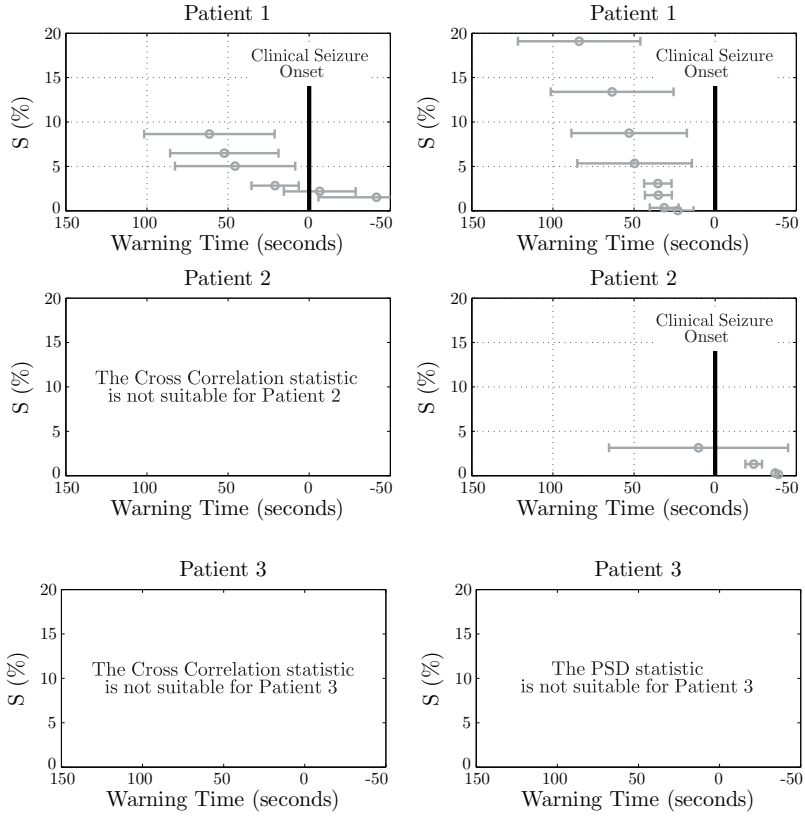


(b) Patient 2



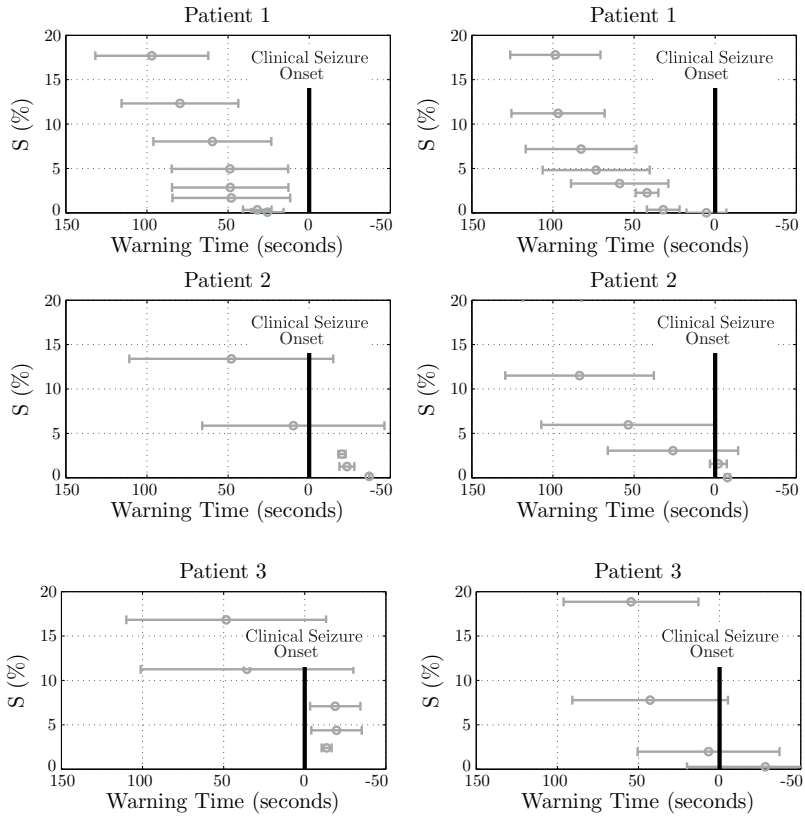
(c) Patient 3

FIGURE 5.14: ROC curves for each of the 3 patients in Group 4 data, using the metrics defined in Section 5.4.1. Not all tests are appropriate for all patients, and only the ones in which characteristic changes are visible during the beginning of a seizure are shown. Overall, correlation dimension performs the best when compared to all other linear statistics because it has 100% TPR with zero specificity in all 3 patients. Cross correlation performs the worst because it is not applicable to 2 out of 3 patients, and has the highest specificity for 100% TPR for Patient 1.



(a) Cross Correlation (b) PSD

FIGURE 5.15: Onset delay evaluation. Each graph shows the computed onset warning for one patient, averaged over all detected seizures. The variance of each computed mean is shown as horizontal lines on this mean, and the specificity for different thresholds shown in the vertical axis. The clinical seizure occurs at time zero, so that positive values indicate a detection is made *before* the clinical onset. Here it is seen that the cross correlation in column 1 is only applicable to Patient 1 and, coupled with results in Figure 5.14(a), provides a warning time of about 40 seconds with 100% TPR and 5% specificity. (Continued)



(c) Wavelet (d) Correlation Dimension

FIGURE 5.15: (Continued) In columns 2 and 3 the results of the PSD and wavelet statistics are shown. Both perform similarly, not surprising given that they are designed to target the same information. Again Patient 1 performs well with clear, consistent warning times. Patients 2 and 3 results are less favorable, with no consistent positive warning times possible with either of these statistics. In column 4 the correlation dimension shows a much better performance for all patients; coupled with 100% TPR in Figure 5.14 it is capable of, in most cases, giving positive warning times. However, higher specificities must be tolerated for Patients 2 and 3, and this improved performance comes at the expense of a much greater computational cost.

reported in Section 5.3, highlighting the importance of both patient specificity as well as the differences between scalp and intra-cranial EEG. However since the focus of this study is on onset/warning time, greater importance was placed on identifying characteristic behavior at the beginning of the seizure, and in particular those that occur before the clinical onset. For example, the slower oscillations of seizures lead to a larger value of the PSD statistic during a seizure. However in patient 1 there is a period of high frequency activity in the lead up to generalization, and the onset detector is based on a *decrease* of the PSD statistic. The result is earlier detection in general, at the expense of some undetected events. The test derived for each of the statistics is patient specific, as is appropriate for the class of problems under investigation.

First it is important to see, on a patient by patient manner, how well the features are able to detect seizures. The results are shown in Figure 5.14 when thresholds are varied. Only the statistics for which characteristic changes were identified are shown in each patient's ROC curves. Specificities rather than FPRs are reported because in this application it is more useful to know the percentage of time that a system such as automated drug delivery will be wrongly active.

For Patient 1 all four statistics were capable of 100% TPR with specificity less than 2%, and less than 1% if we exclude cross correlation. Results are not as good for Patient 2 where only the correlation dimension can achieve 100% TPR and the cross correlation cannot be used at all. Similar difficulties exist for Patient 3. From discussions in Chapter 3 it is not surprising that cross correlation is a less suitable detector of intra-cranially measured seizures because electrodes are close to one another and greater synchronicity is observed at all times, not just during a seizure. The correlation dimension statistic is the best at detecting all seizures in all patients, whilst making no false detections. This comes at the expense of much greater computational complexity than for the linear statistics.

Let us now turn our attention to the warning times these statistics provide, shown in Figure 5.15. Each graph shows the computed onset warning time for one patient, averaged over all detected seizures (i.e., undetected seizures are not included). Results must be coupled with those in Figure 5.14 to be complete. The variance of each computed mean is shown as horizontal lines on this mean, and the specificity for different thresholds shown in the vertical axis. The clinical seizure occurs at time zero, so that positive values indicate a detection is made *before* the clinical onset.

Each column in this figure shows the performance of one statistic applied to all patients, and each row shows the results for all statistics applied to one patient. Looking first at cross correlation in column 1, only Patient 1 benefits from its use. Average warning times of more than 20 seconds occur with 100% TPR and a specificity of 2% or greater. This is a fairly good performance overall.

The nature of the PSD and wavelet statistics is very similar, and the results in columns 2 and 3 reflect this. Patient 1 exhibits 30-40 seconds warning, 100%

TPR and a specificity of about 0.5%, overall better performance than the cross correlation. Prospects are not so good for Patient 2 where 100% TPR is only possible using the wavelet statistic, and with high specificity. The variability in the computed onset times indicate that detections are rarely made before clinical seizure onset. The PSD statistic is not applicable to Patient 3, and whilst it is possible to use the wavelet statistic for this patient the onset warning times have specificities that are too high.

The story is much more positive for correlation dimension in the last column, and the added computational complexity is justified in Patients 2 and 3. Whilst not perfect, warning times for most detections occur before clinical onset (at 100% TPR) with 6% specificity in Patient 2 and 8-10% for Patient 3. Average warning times are close to 50 seconds in both cases.

When designing a *patient specific* onset detector it is not necessary to use the same statistic for all patients – it suffices to choose the best one. For example Patient 1 benefits from any of the four statistics discussed here, and weighing up the computational expense versus the warning time, it is probably best to use PSD or wavelets over correlation dimension. Recall that Patient 1 has an average of 27 seconds between electrographic and clinical onset, thus the window to react is relatively long. This is not so for Patient 2 who has an average of 2 seconds to react. In this case it may be necessary to live with higher specificity as well as the computationally more cumbersome correlation dimension. Alternatively in such cases it may be necessary to shift to an even smaller scale measurement that may reveal longer lead-up times to clinical onsets.

When results such as those here are considered it may be worthwhile asking the question: *is there a need for seizure prediction?* Given that so little progress has been made in this regard in the last decade (see Chapter 7) it may be necessary to move to a detection based regime. If research focus shifts to development of technology that can intervene quickly then Patient 1 and to a smaller extent Patients 2 and 3 can be helped immediately and with very simple signal processing. Of course the results are still patient specific and may not be applicable to all, but this is no different to the expectations placed on current seizure predictors.

In summary, Figures 5.14 and 5.15 reveal that, for intra-cranial EEG recordings, it is sometimes possible to consistently detect a seizure well before its clinical onset. This occurs with more warning when a non-linear statistic such as correlation dimension is used, but it comes at a computational expense that is not necessary in some cases because linear statistics perform comparably.

5.5 Conclusions

A detector must be designed on an application specific basis. Similarly its evaluation is only possible when appropriate data and performance metrics are selected. In this chapter we evaluated several aspects of epileptic seizure detection using the EEG.

In summary, this chapter shows that:

- *Evaluation of Classifiers*: On average SVM outperforms ANN and is simpler to implement.
- *Evaluation of Patient-Unspecific Detectors*: `Reveal` achieved the best FPR with sufficiently high TPR, whilst the best TPR is achieved by `CNet`.
- *Evaluation of Onset Detectors*: Correlation dimension proved the most robust at the expense of a much greater computational burden.

This list by no means implies all applications and all implementations of seizure detectors have been evaluated. It is only *the beginning* to how such an evaluation may be performed in a standard manner.

Algorithms should be tested in a standard way, using *standard data* and *standard performance metrics*. Results reported in any other way cannot be compared.

The data used in this chapter will be made publicly available with the publication of this book, so that this standard can be maintained. Visit this book's website for updates.

6

Modeling for Epilepsy

"The best model of a cat is a cat. Preferably the same cat."

- "Philosophy of Science", 1945, Arturo Rosenblueth (1900-1970)

A *dynamic model* is a set of mathematical equations capable of simulating the behavior of a *system*, at least in part. Its purpose is to *explain* the system behavior: What is it capable of? What is it *not* capable of? If the model is supported by physical observations, can these be explained by the model? The set of equations must be complex enough to describe the dynamics of interest but preferably simple enough so that mathematical tools exist to analyze them. At a bare minimum the model must be computable.

To construct a model of an epileptic brain one must look for cases where epilepsy exists – if the best model of a cat is a cat, why not simply look at a cat? Or, better yet, the same cat? The problem with this approach is that mice can also have epilepsy (so can humans for that matter). A model that can only explain the epilepsy of the one cat is unlikely to be useful in developing an understanding of epilepsy in general. If instead we start with 'epilepsy is common to all mammals' we may model characteristics that are present in all forms.

In this chapter models of the brain capable of describing epileptic seizures are presented. These are limited to models that are *dynamic* rather than static, to reflect the time-evolving nature of both the seizures as well as the measured EEG used to corroborate them. They are also limited to those that reflect the activity of *networks* rather than single neurons. Recall that in the order of thousands of neurons must collaborate for seizure-like activity to occur [73].

Limited resources, computation technology, or data, coupled with the possible presence of chaos, makes explaining complex systems difficult. Even very well known systems such as the weather are too complex and cannot be modeled accurately because of computational limitations and lack of data. A human brain is of similar, if not greater, complexity than the weather: if the activity of each of the 100 billion (10^{11}) neurons in the brain is to be known then in the weather example this is analogous to making a measurement of temperature, pressure, humidity, etc., every 50-100 meters of the earth's surface. This requirement is not feasible neither economically nor computationally. The same is true for the brain. We cannot measure the activity of every

single neuron, and even if we could it would still be very difficult to draw useful interpretations from such vast amounts of data.

Simplifications allow us to manage our problem. Some examples include:

- *Model average activity*: Assume that changes in weather conditions over distances in the order of, say, 50km are insignificant. In this way the order of the *global* problem can be reduced by a factor of a million. A reduction of similar order can be achieved in the brain – the 50,000 or so neurons in a *cortical column* discussed in Chapter 1 are believed to behave roughly in synchrony, thus reducing the complexity of the problem significantly. The assumed synchrony is an approximation because the time scales required to cause a column to act 'as one' are much smaller than those of interest in the study of epilepsy.

- *Reduce parameter space*: In a complex system the fewer variables or parameters used the better. The benefits are exponential. Think of a model of N compartments put together, where the equations of each compartment have M parameters each with K possible values. Reducing the number of parameters M by one results in a *multiplicative* reduction in complexity of $K^{M\times N}$! It is important to have sufficient detail to be representative of the system of interest, but it is more important to reduce the number of variables to the absolute minimum. In the weather example, the model can be simplified by omitting information such as, say, elevation above sea level, if this does not significantly affect the temperature on a flat landscape. In the brain this means omitting information such as the differences in conductivity between gray and white matter, dendritic shape and size, or concentrating on the firing patterns rather than the physiology of the neurons that produce them if these details do not, on average, affect macro-scopic EEG.

- *Tackle a smaller problem*: For the epileptic brain the sub-system most often studied is the cortical column. Its activity is formulated independent of the remainder of the brain, allowing for the understanding of the behavior that is not a result of global network interactions.

The models presented in this chapter employ all these simplifying strategies.

We seek a model of the entire brain, but this is too difficult a task at present. Models of neuronal networks presented in this chapter attempt to understand the *uniform* activity in *sub-systems* of the brain. They focus on how the behavior exhibited by these sub-systems is affected by *parameters* representative of the underlying biology. These parameters represent *average* activity in a meso-scopic network of neurons. One such sub-system

is the *cortical column*, although the theory is not restricted to this.

Simplifying complex models is necessary but it comes at a cost. It relies on assumptions that do not always hold and hence conclusions must be checked carefully.

A model can be *deterministic* or contain *stochastic* components. It is rare to find a completely deterministic system where a set of mathematical equations determines dynamics exactly. Most models are a combination of deterministic and stochastic elements and it is important to understand the different types of random contributions:

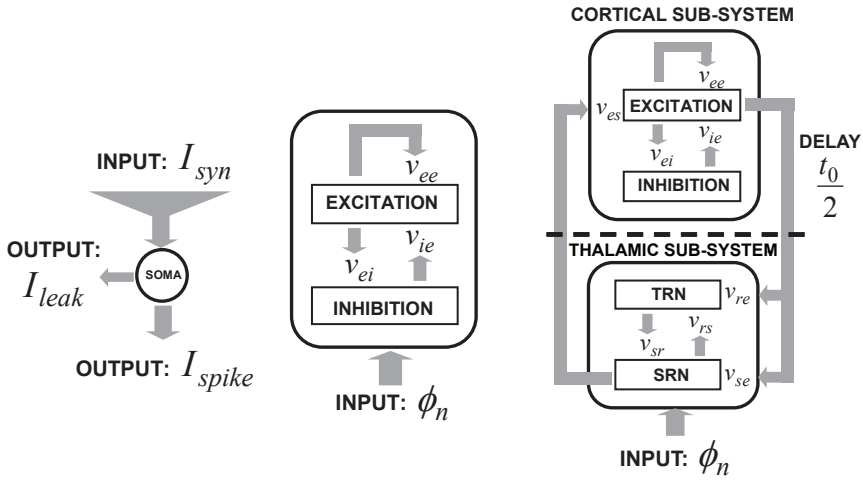
- *Model simplification*: Using a simplified representation results in discrepancies between true and model behavior. The simplifications are necessary to make the problem tractable, but comparisons between model and real data reveal fluctuations that cannot be explained by the model. These may be accounted for by a stochastic component. If the simplifying assumptions are valid then these errors can be kept small.

- *Stochastic activity*: The mechanics of certain aspects of a system are sometimes not understood, either because suitable experimental data do not exist or because its complexity cannot, at present, be captured by analytical methods. Even though the underlying activity may be deterministic these components appear random and can be approximated as such. These random elements are *part of the system itself* and have particular properties (e.g., mean, variance, distribution) gathered from experimental data or assumed based on realistic constraints. An example is presented in Section 6.2 where a model based on the stochastic nature of action potential firing rates is presented.

- *Stochastic inputs*: The system equations can be deterministic but *driven by* a stochastic process, that is, the *input* to the system is somehow random. One example is presented in Section 6.3 where the model equations are deterministic (even though they are derived from stochastic observations) but the input to this model is stochastic. Again, the real input to this system could in actuality be deterministic, but is so complex that it is approximated as a random process.

Any combination of the above is possible.

In the remainder of this chapter the emphasis is on the *creation* of models suitable to describe the EEG. Real EEG data are used to *validate* the model. Whereas in Chapter 2 the emphasis was on the passive mechanisms that affect the EEG, the focus here is on the *active* generators of electrical activity within the brain.



(a) Single neuron (b) Micro-scopic (c) Meso-scopic

FIGURE 6.1: Neural models at different scales. In (a) a single neuron model is shown. The variables of choice are the input and output currents emergent from the cell body (soma). In (b) the activity in a network of neurons is modeled by the interactions between neurons of different types – in this case the excitatory and inhibitory neurons in the cortex. v_{ab} denotes the strength of the response that neuron of type b has on neuron of type a. Larger scale models can be generated by joining neural populations together, as in (c), where the primary pathways between cortical and subcortical networks are shown. In (b) and (c) the parameters and equations must be derived from *averages* over neural populations. TRN refers to the *Thalamic Reticular Nucleus*, and SRN to the *Sensory Relay Nucleus*.

In this chapter, Section 6.1 explains the parameters that are relevant in the meso-scopic and macro-scopic modeling of brain dynamics, expanding on the information presented in Section 1.1.

In Sections 6.2-6.4 we present three different classes of models relating to three different scales in the brain: those derived from *micro-scopic* activity (Section 6.2); those that take averages over neural populations so as to describe the phenomena of the EEG (Section 6.3); and those that target large-scale dynamics (Section 6.4). In all cases it is *networks* of *cortical* neurons that are of interest because these are believed to play an integral role in generating and sustaining seizures.

Finally in Section 6.5 we see how we can make use of these models. We try to *understand* aspects of epileptic seizures inferred from these models, and present practical ways in which these models are used in practice.

6.1 Physiological Parameters of Neural Models

In this section we introduce the *parameters* and *variables* that are important in modeling the dynamics of the brain. The models themselves will be described in detail later. Let us go back to the definition of a system in Equation 3.2, and in particular the state transition formula, rewritten in continuous time as

$$\frac{d\mathbf{z}(t)}{dt} = P(\mathbf{z}(t), \kappa(t), u(t), t) \; ; \; \mathbf{z}(t_0). \tag{6.1}$$

This equation describes the evolution of the system state $\mathbf{z}(t)$, given the current state $\mathbf{z}(t_0)$, a parameter set $\kappa(t)$ and inputs $u(t)$ which both depend on time t. The behavior is determined by the map P.

All dynamic models of the brain can be written in this general form, and can be made as detailed or as coarse as required by the application by changing each of its components, possibly including infinite dimensional state $\mathbf{z}(t)$. Dynamic models have *variables* incorporated into the state $\mathbf{z}(t)$, and it is the dynamics (or changes) of these variables that we want to observe when we study them. They also have a set of *parameters* $\kappa(t)$ that parametrize the family of maps P. Parameters are quantities that are altered to cause changes in the dynamics of the variables, but are typically considered constant over the time scale of interest. That is,

$$\frac{d\kappa(t)}{dt} = 0 \tag{6.2}$$

is assumed. However in reality both parameters and variables are capable of change, thus the explicit dependence of $\kappa(t)$ on time t. But whereas the changes in parameters are caused through means external to the model, the changes in the variables *must* occur through the dynamics represented by the mathematical equations. In any case the distinction between parameter and variable is somewhat artificial from a mathematical point of view because an equation in the form of Equation 6.2 can always be incorporated into Equation 6.1.

There is no set of rules as to which quantities should be parameters and which should be variables. General guidelines include:

- *Time scale of change*: Parameters are those that change slowly relative to the time scale of interest. In most mathematical analysis parameters are assumed static. Sometimes this is true, for example the length of an axon remains relatively unchanged throughout time. Others are capable of slow changes. For example the number of neurons in a network is relatively static in the time scale of 1 hour, but when looking at time scales as long as decades then cell death can be important in dynamics that depend on neural populations. Another example is the plasticity

of the brain, constantly learning and adapting to change, but at a much longer time scale than the EEG.

- *Availability of experimental evidence*: In general, parameters are those that are measurable and for which experiments have been or will have to be done. Realistic constraints on these parameters exist. Variables are more explorative in that a general understanding of realistic values exist but it is more difficult to determine how small changes affect their behavior.

An example of a parameter versus a variable is that of neurotransmitters and neuromodulators, explained in Chapter 1. The modeling of EEG dynamics is concerned with time scales of milliseconds. At these time scales the *effects* that any one neurotransmitters has on the system (as opposed to the changes in neurotransmitter concentration) are assumed roughly uniform and for simplicity these effects are represented as static parameters. However over time the neuromodulators modulate the reaction to neurotransmitters, resulting in changes in EEG dynamics. When it is the dynamics of neurotransmitters that are of interest then equations where these are variables and the neuromodulators are parameters must be derived. A parameter can become a variable at different time scales.

A *variable* in a model is a part of the state, or can be derived from the state, parameters, and inputs to the model. The *parameters* $\kappa(t)$ are properties of the system, which when modeling real systems are in principle experimentally measurable.

For all intents and purposes parameters are considered *static*, although in a system such as the brain some parameters are capable of changing over time. This change occurs at a much slower rate than the changes exhibited by the variables. If this is not the case then these parameters must themselves be included as variables in the model.

In the remainder of this section we outline the parameters $\kappa(t)$ of most importance for single neuron and network models. Under the assumption that in the time scale of interest these parameters are constant, the explicit dependence on time t for parameters described here is omitted for better legibility. Again, effort is made to restrict discussions to those parameters believed relevant in the modeling of macro-scopic dynamics of the epileptic brain.

The reader can reference a list of the parameters described here, along with their notation, in Appendix 6.A. The equations in the form of Equation

6.1 that explain the behavior of the brain using the parameters defined here are outlined in later sections.

6.1.1 Parameters in Single Neurons

The parameters of interest at the single neuron level vary with the model. The well known Hodgkin-Huxley model of the action potential ([66]) describes the ion concentrations inside and outside the neural membrane. These concentrations are the variables that we want to observe, and the constant quantities that define how easily charge is transmitted through the membrane are the parameters. This model provides a very realistic description of currents generated in the firing process. However the set of equations, although computable for a single neuron, become too unwieldy even with a moderate number of inter-connected neurons. Focusing on parameters at this scale (both spatial and temporal) is not useful for networks of neurons because they predominantly are concerned with larger spatial and longer temporal scales.

A simpler and more analytically tractable alternative is to model neurons as *integrate-and-fire* units, that is, model the currents on the dendritic tree as inputs that are integrated (summed) at the cell body to determine if an action potential is fired as an output. Unlike the Hodgkin-Huxley approach the morphology of the inputs and outputs is not important – both currents are modeled as *spikes*. Changes in parameter values can alter the observed firing pattern of the modeled neuron, and consequently can be used to represent the different types of neurons, if such detail is necessary. Thus the parameters that are important in this case are the properties of the neuron that generate action potentials, including the resting membrane potential V_P that determines the voltage of the membrane with no input; the passive membrane time constant τ_p and membrane capacitance c that determine how fast changes in membrane potentials are possible; and voltages V_{TH} and V_R that determine when an action potential is fired and the potential of the membrane after this occurs. The full parameter set listed in Appendix 6.A can be used to model the variables for membrane potential $V(t)$ and the currents $I_{leak}(t)$, $I_{spike}(t)$ and $I_{syn}(t)$ shown in Figure 6.1(a)) that describe $V(t)$.

6.1.2 Parameters in Networks of Neurons

A mathematical formulation of a *collection* of interacting neurons may or may not make use of the characteristics of a single neuron. Examples of both are presented in Sections 6.2-6.4. In either case the network parameters must be *derived* or *hypothesized* from evidence provided by ensembles of single neurons because a solution that is consistent at all spatial scales is preferable. The mean firing rate $Q_a(t)$ of neurons of type a is one such example, where the parameters in this function describe averages over neuron populations.

One important aspect in a network of neurons is their expected histology, that is, the type and number of neurons in this network. In a slice of cortex

there are excitatory (e) and inhibitory (i) neurons and their numbers $N_{a=e,i}$ must be estimated from experimental data[1]. Each neuron receives input from a specific number of synapses C_{ab} formed from neuron of type b onto neuron of type a. Whereas a can take the value only of the types of neurons *within* the network ($a = e, i$), b can take an additional value of $b = n$ where n represents input from neurons *outside* the network. External inputs are usually modeled as stochastic elements, as is discussed in more detail later. An illustration of this is shown in Figure 6.1(b).

These parameters are measurable and can be used directly in the model or indirectly to deduce other parameters. For example if v_{ab} denotes the magnitude of the response that neurons of type b induce on neurons of type a then v_{ab} is proportional to the number of synapses C_{ab} and the current strength s_b caused by a synapse with neuron type a. These latter are also experimentally measurable quantities.

Recall that dividing a system into sub-systems is necessary to simplify the development of macro-scopic models. A *patch* of cortex, as opposed to the entire cortex, is one example of a sub-system. It can be modeled independently, as in Figure 6.1(b), or in conjunction with other sub-systems, as in Figure 6.1(c) where the cortex interacts with thalamic networks. In such cases it is important to incorporate information about each sub-system (N_a, C_{ab} and s_b) as well as the strength of interactions *between* sub-systems, also experimentally observable quantities. Additional information such as propagation time $t_0/2$ of signals between sub-systems in Figure 6.1(c) is also necessary. To distinguish between sub-systems different subscripts are used. Excitatory and inhibitory neurons in the sub-cortex are henceforth denoted s and r respectively to differentiate them from excitatory and inhibitory neurons in the cortex.

In summary, a neural population with N neurons behaves according to the number N_a of each neuron type a, the number of synapses C_{ab} resulting from other neurons of type b, and the strength of the synapse s_b. All these parameters are used directly or indirectly to model neural networks.

Connecting *sub-systems* of neural networks together requires other similar parameters that describe their interactions, including propagation delays. The manner in which the sub-systems are connected affects dynamics.

[1]Of course, for an accurate model it may also be important to model the different types of excitatory and/or inhibitory neurons found in the brain, along with many other details omitted in this discussion. However for the purposes of this book, as well as for the very real necessity of a *tractable* mathematical model, many of these important details are not included here. In theory these models can be extended to include any number of different types of neurons.

In all cases the parameters must reflect the physiology of real neuronal networks, either by restricting them to realistic ranges determined from experimental data or by using knowledge of the systems to infer their behavior.

Using the physiology in Section 1.1 and the parametrization discussed here we are now in a position to write *mathematical equations* of neural models at the different scales. We want to use these equations to describe the *active* generators of electrical activity, largely ignoring the passive mechanisms such as volume conduction explained in Chapter 2. Although these are important in describing the EEG we must wait until future developments allow the mathematics and the computing to describe large regions of the cortex.

To avoid lengthy discussions of all available models and their variants, a subset that highlight the key issues are selected. These are believed to be the most general, whilst still relevant to describing epileptic seizures in both scale and parametrization.

6.2 Micro-Scopic (Statistical) Models

We start the description of neuronal dynamics at the micro-scopic scale. A class of models that capture the statistical behavior of ensembles of neurons based on physiology at the micro-scopic scale is called an *integrate-and-fire* (IF) neuron model. They provide a biologically inspired representation of *background* network activity in a *local* and *homogeneous* region of the cortex. They are based on the simpler characteristics of neural firing (nowhere near the complexity of the Hodgkin-Huxley[2]model) but are nevertheless capable of replicating experimentally observed behavior of networks of interconnected cortical neurons in vivo.

Integrate-and-fire (IF) models describe the statistics of the activity in a network of cortical neurons. The model may be *stochastic* if it includes a stochastic description of the input. The model may be deterministic if it only describes the distribution

[2]The Hodgkin-Huxley model describes the generation and propagation of an action potential based on trans-membrane cellular currents and concentrations. In this book we have focused on phenomena at a larger scale than this, and have never described the biochemical and gating mechanisms that make this possible. Thus a model of this nature is not included here, but the interested reader may consult [66].

of the statistics. Typical distribution quantities used as part of the state $\mathbf{z}(t)$ are:

- $\nu(t)$: The mean spiking rate of action potentials in the network.
- $CV(t)$: The coefficient of variation of the inter-spike interval.
- $\mu(t)$: The mean membrane potential in the network.
- $\sigma(t)$: The standard deviation of the membrane potential.
- $\tau(t)$: The effective membrane time constant.

The most important advantage of this class of models is that experimental data of the above variables exists, which can be used to *validate* the model.

Following is a brief description of the mathematics along with important results. A comprehensive review on IF models can be found in [24] and [25]. Other related literature includes [23] and [108]; this is by no means a complete list.

6.2.1 Model Summary

Biologically inspired models first look at a single neuron. For those interested in a reference of the model only, a summary of all parameters, variables and equations can be found in Appendix 6.B, complemented by parameters in Appendix 6.A.

A single neuron fires an action potential depending on the balance of incoming and outgoing currents. The membrane potential $V(t)$ is the difference in potential between the inside and the outside of the cell body (it is the proper definition of a voltage). Changes in $V(t)$ occur according to three dominant currents shown in Figure 6.1(a): the passive currents in the membrane used to restore ion concentrations ($I_{leak}(t)$), the currents formed by the synaptic inputs ($I_{syn}(t)$) and the output current of the action potential ($I_{spike}(t)$). Each of these currents groups together the behavior generated by many different classes of ion channels.

The dynamics of $V(t)$ are described by

$$c\frac{dV(t)}{dt} = -I_{leak}(t) - I_{spike}(t) + I_{syn}(t). \tag{6.3}$$

The capacitance c scales the rate of change to the membrane potential.

When there are no input and output currents the membrane voltage relaxes towards the resting membrane potential V_P, which is assumed constant. When currents are present, $I_{leak}(t)$ works to restore the ion concentrations to this

resting state. Thus $I_{leak}(t)$ is proportional to the difference $(V(t) - V_P)$. Mathematically[3],

$$I_{leak}(t) = \frac{c}{\tau_p}(V(t) - V_P), \tag{6.4}$$

where the membrane time constant τ_p is used to capture the speed at which the leak current restores $V(t)$ to V_P.

$I_{spike}(t)$ is the current caused by an action potential spiking on the axon. It can be described by

$$I_{spike}(t) = c(V_{TH} - V_R)\delta(V(t) - V_{TH}). \tag{6.5}$$

This says that the membrane potential reaches threshold (V_{TH}), fires an action potential and resets the membrane to the reset potential (V_R). Here $V_R < V_{TH}$. There is a refractory period τ_r before the membrane potential begins to evolve again.

Finally, $I_{syn}(t)$ represents the effects of the synaptic inputs to the cell body. $I_{syn}(t)$ can be described in terms of currents or conductances[4]. Both have been analyzed in literature (see, for example, [24]) but only the results for conductance based models are presented here for reasons discussed later. The equation for synaptic current in conductance-based synaptic input is given by

$$I_{syn}(t) = \sum_{a=e,i,n} c\alpha_a(V(t) - V_a)\wp_a(t), \tag{6.6}$$

where the subscripts e, i, n represent excitatory, inhibitory and external (excitatory) synaptic activity respectively, as discussed in Section 6.1. The inputs can often be recurrent because past outputs can affect present behavior. This is represented by $\wp_a(t)$, a function drawn from a suitable distribution that integrates firing times of pre-synaptic neurons as well as transmission delays of post-synaptic potentials. The parameters α_a are a unitless measure of conductance of synapses formed from neuron a (more details can be found in [108]).

The above is a description of how each neuron works, modeling pre- and post-synaptic currents, their integration and consequent firing. A neural network is constructed by putting many of these neurons together. For analysis, a network of N neurons, with N_e excitatory and N_i inhibitory is used.

[3]This equation is effectively Ohm's law. Because voltages are defined only up to a constant we could have re-written a simplified equation as $I_{leak}(t) = \frac{c}{\tau_p}V^{\#}(t)$, where $V^{\#}(t) = (V(t) - V_P)$. However because the resting membrane potential V_P plays an important role in some aspects of the neuron's dynamics it has become convention to explicitly include it.

[4]Current based models try to emulate the currents passing through the dendritic tree, whereas conductance based ones emulate how conductivity is changed because of these synaptic currents (see Chapter 2).

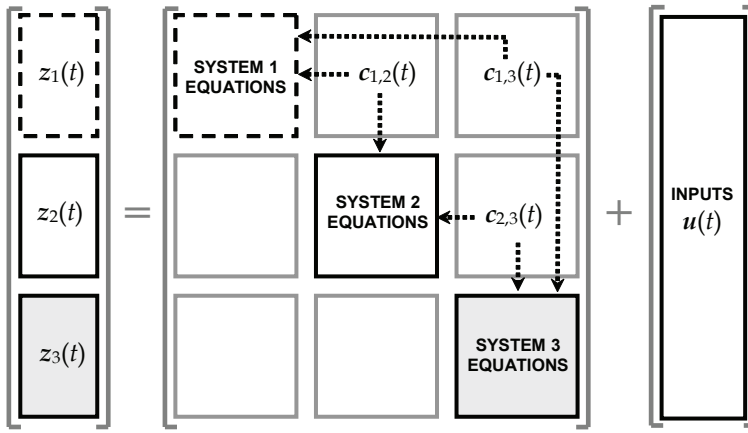


FIGURE 6.2: Pseudo-matrix representation of aggregate models. Here the equations of one large model are formed by joining three smaller (identical) models together. The sub-systems are defined by $\mathbf{z}_1(t)$, $\mathbf{z}_2(t)$ and $\mathbf{z}_3(t)$ and their interactions are modeled via coupling parameters $c_{1,2}(t)$, $c_{1,3}(t)$ and $c_{2,3}(t)$. The dimensions of aggregate models grow very quickly with the number of sub-systems, and with the added complexity of coupling parameters this naive implementation quickly becomes unmanageable. To make the manipulation of larger models more tractable simplifications that remove unnecessary detail should be applied.

Each neuron of type $a = e, i$ has a number of synaptic inputs (C_{an} external, C_{ae} excitatory and C_{ai} inhibitory). Connections are assumed sparse ($1 << [C_{ae}, C_{ai}] << N$) and chosen at random.

To model network activity we need to make the jump from Figure 6.1(a) to Figure 6.1(b). There are several ways to go about this. The naive approach is that if we want to model N neurons we need to have N copies of the single neuron model repeated, along with their interactions. This is because the behavior of the single neuron is a subset of the behavior of an ensemble of neurons. Pictorially (in pseudo matrix notation) this scenario can be seen in Figure 6.2 for $N = 3$. Each neuron is modeled with its own set of equations, with the interactions as well as the inputs affecting each unit.

However having N sets of Equations 6.3, 6.4, 6.5 and 6.6 is analytically difficult, perhaps even impossible, with a large enough N. The dimensions of the matrix in Figure 6.2 grow linearly with the number of neurons.

As explained earlier when modeling complex systems it is necessary to make *simplifying assumptions*. For example, instead of modeling each neuron individually, we can model the *average* behavior of the ensemble, as done for

ensemble IF models. This reduces the complexity of the model and makes the problem tractable, under the assumption that it is a 'good enough' approximation to the real activity. Further simplifying assumptions can be made:

1. Because at any one time each neuron receives many small inputs from a network that is sparsely but heavily interconnected (that is, there are lots of connections but there could have been many many more), the excitatory and inhibitory inputs can be modeled as a stochastic process with a Poisson distribution[5]. The mean firing rate in the network is equal for both excitatory and inhibitory inputs under the generalizing (and perhaps unrealistic) assumption that excitatory and inhibitory neurons have the same characteristics. The strength of the synaptic inputs are given by $\gamma_{ae,ai}(t) = \nu(t)C_{ae,ai}$, where $\nu(t)$ is the average firing rate of neurons in the network.

2. The external input is modeled as a temporally homogeneous Poisson process, that is, the input properties do not change over time and has constant intensity (mean firing rate ν_n, and $\gamma_{an} = \nu_n C_{an}$).

3. Because connections are assumed sparse, correlations between firing times of different neurons are negligible. Neurons in such an ensemble can be modeled as firing independently[6].

4. Correlations between firing times of the same neuron can be ignored because experimental observations show that spike times follow a Poisson distribution. This is only valid when the neuron is involved in background activity.

5. On average, firing times of a single neuron can be regarded as independent.

Putting all this together and the fact that inputs are many and small implies that the system can be described by a probability density function of membrane potentials $P(V(t))$. This probability density function obeys the well known Fokker-Planck equation. The network dynamics can be studied by analyzing $P(V(t))$, and with the appropriate use of boundary conditions the stationary case ($P^*(V) = P(V(t \to \infty))$) can be solved *analytically*

$$\tau^* \frac{\partial P^*(V)}{\partial t} = 0 = \Theta^2 \frac{\partial^2 P^*(V)}{\partial V^2} + \frac{\partial}{\partial V}\left[(V - M)P^*(V)\right], \tag{6.7}$$

[5]The Poisson distribution is a stochastic process often used to describe arrival times, where the probability of an event (in this case an action potential) only depends on the time since the last event occurred. See, for example, [93]

[6]This does not imply that neurons cannot fire at the same time - the dynamics of the system can bring the model to synchrony. The assumption of uncorrelated firing times implies that, when talking about average behavior, the existence of one spike cannot be directly attributed to another spike. Like all assumptions this one may not hold at all under all circumstances.

with M and Θ corresponding to the mean and variance of $P^*(V)$ respectively. This stationary distribution also has analytic solutions for steady state solutions of the variables of interest: $\nu^* = \nu(t \to \infty)$, $CV^* = CV(t \to \infty)$, $\mu^* = \mu(t \to \infty)$, $\sigma^* = \sigma(t \to \infty)$ and $\tau^* = \tau(t \to \infty)$, all of which can be found in [108].

6.2.2 Validation and Limitations

The most important result is that conductance-based IF models are able to replicate (at least qualitatively) the behavior of more realistic models based on the Hodgkin-Huxley equations. Furthermore, the observed behavior produces quantities for ν^*, CV^*, μ^*, σ^* and τ^* consistent with what has been observed experimentally in vivo (shown in Figure 6.3(a)), something that current-based IF models cannot do [108]. The advantage of this is that a relatively simple model for which analytic solutions exist can be used to study network dynamics. The Hodgkin-Huxley model does not *scale* to network models because it is too complex.

IF models represent the activity in large-scale networks, *but can still make a valid analytical connection to micro-scopic models.*

Aspects that have been ignored in the model include a synaptic time constant, non-sparse connectivity (both long and short range connections), correlations in the times of spikes, adaptation of neural firing, synaptic adaptation and failure, and dendritic morphology [24]. Keeping in mind that we are interested only in the modeling for the recognition of epileptic seizures (and not more complex tasks such as learning, for which adaptation is important), *the behavior that we are interested in is captured by a relatively simple model.* This suggests that the elements that are important in network dynamics are already part of this model, and the simplifications are somewhat validated.

Several activity levels have been identified through simulations. Figure 6.3(b) shows some examples of these: a *high-rate regular* firing and a *quiescent* state both thought to be unrealistic or abnormal cases, and a *low-rate irregular* firing state believed to describe the background neural activity. This last is in agreement with both observed experimental data and Hodgkin-Huxley models.

What is not clear from published material is how IF models relate to epileptic seizures. Analytic solutions have identified bifurcations and oscillatory regions in parameter space, but are these seizures? If so, what is the physiological interpretation that leads to epileptic seizures? Does the activity in the epileptic focus resemble the quiescent or the high rate activity shown in Figure 6.3(b)? As far as the validity of the model goes, experimentally available data mostly belongs to low-rate, irregular firing, so it is difficult to infer what part

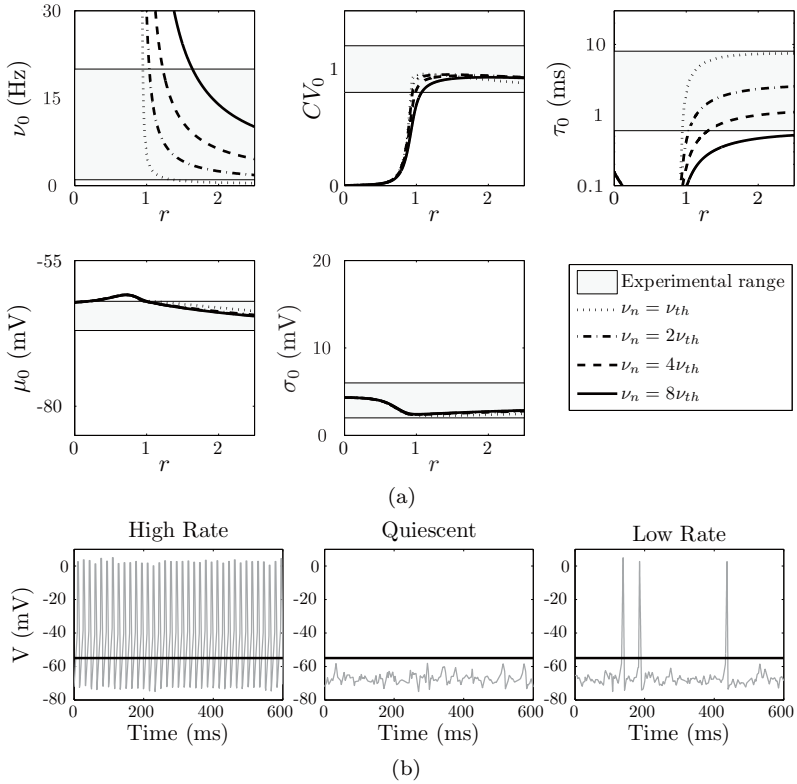


FIGURE 6.3: The advantage of conductance based IF models is that complexity is reduced and solutions are consistent with experimentally observed ranges, as shown in (a) where the IF model outputs are compared to real data. This is true for a wide range of external input mean spiking rate ν_n, shown as a function of $\nu_{th} = 1.4$Hz – the minimum input rate required to bring a neuron to threshold. r is the ratio of inhibition to recurrent excitation in the network [108]. In (b) are examples of some excitation levels identified in the model: *high-rate regular firing* (left), *quiescent* (middle) and *low-rate irregular* (right). Recall that firing of action potentials is achieved when the threshold voltage ($V_{TH} = -55$mV, as marked with a horizontal line) is crossed. The first two cases are believed unrealistic or abnormal behavior, whereas the low-rate irregular firing coincides with experimentally observed background activity. Graphics were reproduced with permission from its original publication in [108] using data provided by the authors.

of parameter space should/can represent epileptic seizures. Work specific to the disorder has to our knowledge not yet been published.

Another unanswered question is: *how does the IF network activity relate to the EEG?* Few biologically feasible structural constraints have been put on network connectivity. What is the scale for which the model is valid? An arbitrarily sized patch of cortex, a cortical column, a minicolumn? The role of interactions between sub-systems within the brain have also been ignored, most importantly the thalamus, responsible for regulating flow of information and believed to be involved in the generation or spread of at least some types of seizures. How far can a simple model that has the advantage that analytical solutions exist be pushed to understand network dynamics? Some of these issues are addressed next.

6.3 Meso-Scopic (Phenomenological) Models

Meso-scopic models operate at a larger scale than the micro-scopic models because they include interactions between sub-systems of the brain, as in Figure 6.1(c). Again, the move to a larger scale necessitates a change in the simplifying assumptions because while models in the form of Figure 6.1(b) describe a subset of the behavior in (c) we also want to ensure that the computations do not become unmanageable. This concept is analogous to that from moving from a single/few neurons to networks, as is explained in Figure 6.2.

As a consequence, the class of models described here are based on biology that is further removed from the dynamics of single neurons. They are *phenomenological* because the variables are related to the EEG – they describe macro-scopic phenomena. In the process some of the micro-scopic biology inherent in IF models is lost and many of the parameters important in the meso-scopic models are no longer directly measurable. Although these meso-scopic models are capable of qualitatively describing some EEG waveforms, it is difficult to infer what the smaller-scale mechanisms that lead to a particular waveform are.

The mathematical framework for these models is not specifically constrained to local activity. However insightful results only apply to *local background* activity of a *homogeneous* network of neurons. In this sense they are not too different from IF models. The difference lies in the modeling approach, in particular the inclusion of inter-connections between neural networks.

Phenomenological models describe the *potential fields* generated by cortical and sub-cortical neurons. The equations describing the system are, on the whole, deterministic. The model itself is *stochastic* only when a random external input is as-

sumed. The variables of interest relate to the fields generated by the different types of neurons involved and describe features of their distribution. For neurons of type a in a network these are

- $Q_a(t)$: The mean spiking rate of action potentials.
- $V_a(t)$: The mean membrane potential.
- $\phi_a(t)$: The field generated by the dendritic currents.

In all cases the neurons involved are the excitatory and inhibitory connections in the cortex ($a = e, i$ respectively) as well as the excitatory and inhibitory neurons in the thalamus ($a = s, r$).

The main advantage of these models is that they are readily related to the EEG waveforms under the assumption that these recordings are proportional to the dendritic fields of the excitatory neurons in the cortex ($\phi_e(t)$), as discussed in Chapter 2.

A large volume of work has been published about these models, pioneered by Walter Freeman [43] who in the early 1990s worked with the olfactory bulb to produce models at the level of a cortical column. Improvements and modifications have been made over time but the original concepts remain. The group of researchers involved in [148], [144], [146], [147], [154] and [22] advanced the model by involving interactions with sub-cortical networks, thus explaining many of the common EEG rhythms. Work in [95] is also relevant. A summary of the salient results of this model follows.

6.3.1 Model Summary

The meso-scopic models describe *average* activity of a population of neurons and the mechanisms of a single neuron are no longer important. The model is explained here, but those interested in a quick reference only can find a summary in Appendix 6.C, complemented by parameters in Appendix 6.A.

The neurons of type a in an interconnected network such as that in Figure 6.1(b) fire action potentials distributed according to a sigmoidal function of the membrane voltage. The firing rate is an average of the population and at any one time depends on the mean membrane potential $V_a(t)$, the mean taken over the ensemble of neurons considered. Although the slope of the sigmoid varies for neurons of different types the average firing rate $Q_a(t)$ can be described by a single sigmoidal function

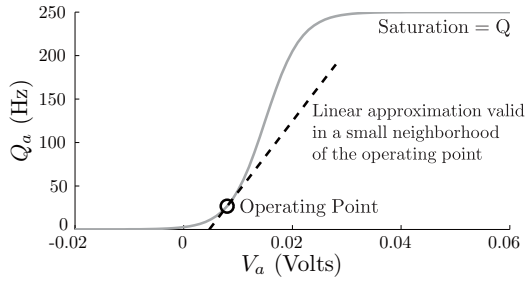


FIGURE 6.4: A typical relationship between average firing rate Q_a versus mean membrane potential V_a for ensembles of neurons of type a. Here $\mathbf{Q} = 250\text{s}^{-1}$, $\sigma = 3.3\text{mV}$ and $\bar{V}_{th} = 15\text{mV}$. Q_a is a non-linear function that at an operating point can be approximated by a linear function as shown. This *linearization* is an approximation that is only valid for values close to the operating point.

$$Q_a(t) = \frac{\mathbf{Q}}{1 + e^{-[V_a(t) - \bar{V}_{TH}]/\sigma}}, \tag{6.8}$$

where $\bar{V}_{TH}, \sigma$ and $\mathbf{Q}$ are experimentally measurable parameters that describe the mean firing threshold, standard deviation and maximum firing rate considered over the ensemble of neurons, respectively. A typical Q_a is shown in Figure 6.4. $Q_a(t)$ (and similar forms of this equation) is widely used in these models as the mean firing rate of a population of neurons in response to the mean membrane potential, even though experimental justification remains vague[7].

Thus action potentials are fired at a rate indicated by $Q_a(t)$ resulting in an electric field $\phi_a(t)$, proportional to this firing rate if the axons are short, as is the case with inhibitory neurons of type $a = i$. Excitatory cortical neurons ($a = e$) have longer projections and the field they generate is related to the firing rate via a low pass filter, denoted by the operator D_{γ_e}, so that $D_{\gamma_e}\phi_e(t) = Q_e(t)$. The operator is defined as

$$D_{\gamma e} = \frac{1}{\gamma_e^2}\frac{d^2}{dt^2} + \frac{2}{\gamma_e}\frac{d}{dt} + 1, \tag{6.9}$$

where γ_e is a measure inversely proportional to the range of axons of cortical neurons, and thus also inversely proportional to the time it takes to travel along an axon, on average.

[7]The input-output functions of most cortical neurons are well known, and they are not really like the sigmoidal function shown in this figure. This function is used to represent the average of these input-output functions.

Action potentials feed into dendrites of neurons within the same population. In some models, a second sigmoidal function is used to describe the firing of post-synaptic potentials (PSPs) but this is ignored here for simplicity, thus the assumption that $\phi_a(t)$ is proportional to $Q_a(t)$. The observation that the membrane potential requires co-incidence of many dendritic inputs leads to the postulation of a further filter relating the dendritic field to the mean membrane potential as

$$D_{\alpha\beta}V_a(t) = \sum_b v_{ab}\phi_b(t), \tag{6.10}$$

where $D_{\alpha\beta}$ is a second order differential operator defined as

$$D_{\alpha\beta} = \frac{1}{\alpha\beta}\frac{d^2}{dt^2} + (\frac{1}{\alpha} + \frac{1}{\beta})\frac{d}{dt} + 1. \tag{6.11}$$

The parameters $1/\alpha$ and $1/\beta$ are the time constants of the filters, which physiologically relate to the rise and decay times of the cell-body potential produced by an impulse at a dendritic synapse. Together they define the shape of the filter. v_{ab} is the strength of interaction between neurons of type b with a, proportional to the number of synapses from b to a. In summary Equation 6.10 describes the mean membrane potential of neuron of type a resulting from the the summed dendritic input from all other neurons, filtered by the dendritic tree in a manner described by $D_{\alpha\beta}$.

Equations 6.8-6.10 describe the 'mean' behavior of an arbitrary neuron type within a local network. If the network is representative of a cortical area, for example a cortical column as in the IF models, then $a = e, i$ and $b = e, i, n$, where n is the input from other cortical or sub-cortical populations. In this averaged form all neurons of a given type behave alike. The EEG is proportional to $\phi_e(t)$ – the field produced by excitatory (pyramidal) neurons that align in the cortex (see Chapter 2).

Whilst capable of reproducing much of the observed clinical phenomena, the cortical column alone is incapable of explaining the generation of certain common EEG rhythms (e.g., alpha) and oscillations (e.g., epileptic seizures). The addition of connections with another similar but sub-cortical thalamic sub-system introduces delays and feedback that are believed by some to be responsible for these phenomena. The modification is physiologically reasonable because the thalamus is responsible for relaying and regulating sensory information to and from the cortex.

Cortico-thalamic and thalamo-cortical projections occur through the *Sensory Relay Network* (SRN), responsible for relaying external sensory information to the cortex, and the *Thalamic Reticular Nucleus* (TRN), an inhibitory interface to regulate flow of information between cortex and thalamus. Transmission delays ($t_0/2$) are introduced between the cortex and the thalamus. Thus t_0 is the round-trip time for a signal passing through the thalamus and back to the cortex. All other delays are assumed negligible at the frequencies of the EEG.

The behavior of both the thalamus and the cortex follows Equations 6.8-6.10, with the addition of neurons s and r, representing the neurons in the SRN and TRN respectively. The equations are also modified to include the feedback and relay that occurs between the two neuron populations. The complete setup along with all neuron types in each sub-system is as shown in Figure 6.1(c). The arrows indicate all the physiologically possible flows of information.

Despite the many simplifications involved even this model remains relatively complex to analyze. A full behavior analysis (the collection of all possible solutions) is prohibitively complex. The main ingredient that complicates the matter is the non-linear characteristics of $Q_a(t)$ in Equation 6.8.

A further complication is the inherent delay in the model. Despite its apparent simplicity this model is still an infinite dimensional model, and its full bifurcation analysis remains an open question and is beyond the scope of the present text. The interested reader can find information on bifurcation analysis in texts such as [83] and [159].

Nevertheless some analysis is possible and is discussed next.

6.3.2 Analysis: Linearization, Stability and Instability

A typical analysis proceeds as follows. Assuming a constant input of a certain magnitude, one can compute a *steady state solution* (when all variables are constant). Note that there may be many steady states, not all biologically relevant. Mathematically we want to find a solution for the system $\mathbf{z}(t)$

$$P(\mathbf{z}, \kappa, u, t) = 0. \tag{6.12}$$

Dependence on time t has been removed because we want constant solutions, that is, we want $\mathbf{z}(t) = \mathbf{z}$, $\kappa(t) = \kappa$ and $u(t) = u$.

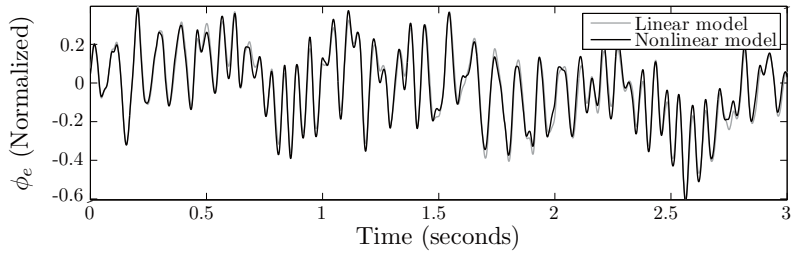
Next we can check which of the solutions will be observed in practice by analyzing the local stability properties of this solution. A common way to do this is through its *linearization*.

Linearization is a way for the *non-linear* model from Section 6.3.1 to be made *linear*. Linear systems are desired because they are much better understood and more tools are available to analyze them.

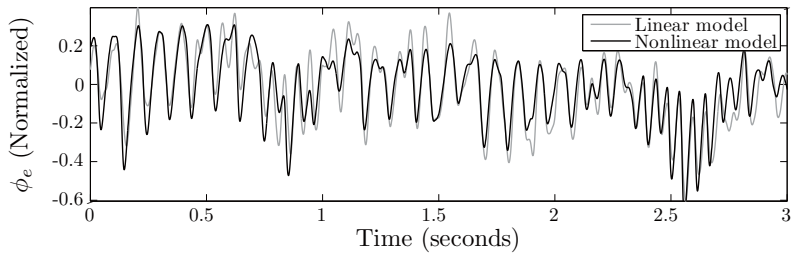
Linearization is achieved by *approximating* an arbitrary non-linear curve $f(x)$ by a linear one. This is done at an *operating point* x_0, as shown in Figure 6.4 for $f(x) = Q_a$ and $x = V_a$, the only non-linear component in the meso-scopic model. The linear approximation $\hat{f}(x)$ is given by

$$\hat{f}(x) \approx \left.\frac{\partial f(x)}{\partial x}\right|_{x=x_0} (x - x_0) + f(x_0), \tag{6.13}$$

where the derivative is the gradient of $f(x)$ at $x = x_0$ and $f(x_0)$ is a constant that is zero when equations are shifted so that the operating point



(a) ϕ_n with variance of 0.001



(b) ϕ_n with variance of 20

FIGURE 6.5: In (a) *small* changes in the input ϕ_n show that simulations of the non-linear and the linearized system are almost identical. In (b) the changes in ϕ_n are large and simulations no longer agree because the linearized system is an approximation valid only for small changes around an operating point. ϕ_n is modeled as zero mean stochastic white noise with variance 0.001 in (a) and 20 in (b). Such large changes are necessary because the system operating point is relatively unaffected by ϕ_n.

is at the origin. This linear approximation is very similar to the original non-linear function ($\hat{f}(x) \approx f(x)$) only around the chosen operating point and results are valid only around x_0. Simulations of the linearized and non-linear system should be almost identical if the changes made to the system are small. This is shown to be true in Figure 6.5 when linearization is applied to Q_a in the meso-scopic model.

In the model the operating point represents a mode of behavior that is not expected to change too quickly because parameters remain relatively stable over time. If large changes occur the linearization is no longer valid. This is a major limitation in analysis of this type: time-evolutions can only be explored with relatively small changes around operating points. Thus often analysis of this kind involves understanding many different operating points, and even then only a small subset of all possible non-linear behavior is explored. Nevertheless it often suffices as the behavior around operating points is generally the most relevant.

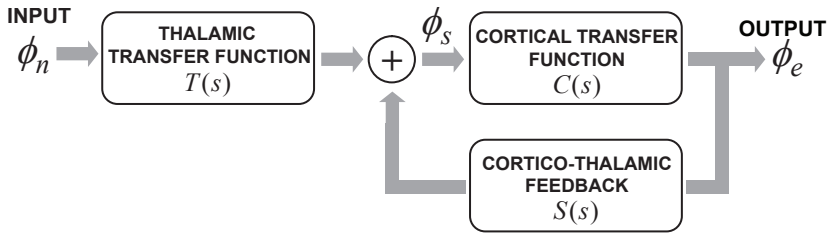


FIGURE 6.6: The equations describing the model shown in Figure 6.1(c) can be re-written in terms of its steady-state linear behavior. The transfer function is a frequency domain representation of the interactions between cortical ($C(s)$), sub-cortical ($S(s)$) and thalamo-cortical ($T(s)$) sub-systems. The *feedback* loop formed can be analyzed for steady-state stability. Equations for $C(s)$, $S(s)$ and $T(s)$ can be found in Appendix 6.C.

With a little bit of work Figure 6.6 shows the resulting linearized mesoscopic model, divided into sub-systems $C(s)$, $S(s)$ and $T(s)$ representative of cortical, subcortical and thalamo-cortical interactions. These are spectral domain representations with $s = 2\pi fj$ for steady-state solutions and f is frequency in Hz. For the EEG the relevant range is $0 < f < 100$Hz. $C(s)$, $S(s)$ and $T(s)$ are known as the *transfer functions* of the system and they explain the relationship between the input and the output of the system. They are

$$\begin{aligned}
C(s) &= L(s)G_{es}\,\frac{A(s)}{1-L(s)G_{ei}}\,\frac{1}{1-L(s)G_{ee}\frac{A(s)}{1-L(s)G_{ei}}} \\
S(s) &= \frac{L(s)G_{se}+L(s)G_{sr}L(s)G_{re}}{1-L(s)G_{sr}L(s)G_{rs}}e^{-st_0} \\
T(s) &= \frac{L(s)G_{sn}}{1-L(s)G_{sr}L(s)G_{rs}}e^{-st_0/2} \\
A(s) &= \frac{1}{(1+s/\gamma_e)^2} \\
L(s) &= \frac{1}{(1+s/\alpha)(1+s/\beta)}.
\end{aligned} \tag{6.14}$$

G_{ab} are the *gains* of the linearized system that result from the linearization process. They are calculated as

$$G_{ab} = \left.\frac{\partial Q_b}{\partial V_b}\right|_{V_b=V_b^*} v_{ab} = \frac{\phi_b^*}{\sigma}\left(1-\frac{\phi_b^*}{Q}\right)v_{ab} \tag{6.15}$$

where ϕ_b^* and V_b^* are the solutions of $\phi_b(t)$ and $V_b(t)$ at equilibrium.

An overall system transfer function between input $\phi_n(s)$ and output $\phi_e(s)$ is given by

$$H(s) = \frac{\phi_e(s)}{\phi_n(s)} = T(s)\frac{C(s)}{1 - C(s)S(s)}. \tag{6.16}$$

Now that we have a simpler (linear) model we can use linear tools to analyze its stability. A solution is stable provided small changes in the present do not have a major effect on the solutions of the indefinite future. We want a small change in the starting condition to lead to responses that remain close forever.

We can use the transfer function of a system directly to infer stability. The denominator of this function has *roots* or solutions, also known as *poles*. *If all real parts are negative the system is stable.* The imaginary axis forms the *instability boundary*. Poles that are purely imaginary indicate oscillations or resonances. In this case the value of the pole denotes the frequency of the resonance, in Hz times 2π. Poles that are very close to the imaginary axis also form resonances (although stable if negative) because small perturbations of the system can lead the system closer to the imaginary axis. The dominant frequency is determined by the pole that is closest to the instability boundary[8]. In our model the input ϕ_n is stochastic and thus a wide range of frequencies are excited. If there are any poles close to the imaginary axis the simulations will show these resonances.

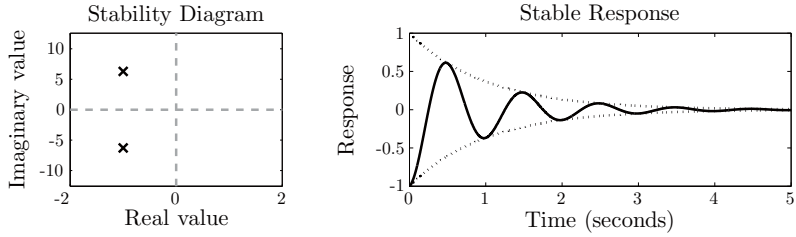
It is useful to visualize stability by plotting the real parts of all poles on the horizontal axis and imaginary parts on the vertical axis. A pictorial representation for a system with only 1 pair of complex poles is shown in Figure 6.7.

A *stable* linear system is incapable of describing oscillations. Transitions into oscillatory behavior in the non-linear system are transitions to instability in a linearized system. Distinguishing between stability and instability is important in the study of epilepsy. Epileptic seizures are identified with oscillatory waveforms; thus in the linearized system transitions to epileptic seizures are represented by transitions into instability.

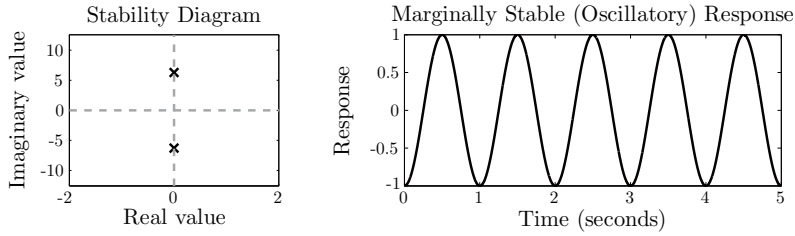
The stability of a linear system is often studied through its transfer functions: *A linear system is stable if the real part of the roots of the denominator, also known as poles, are negative.* Graphical representation of the poles is an easy way to determine if the system is stable or not.

---

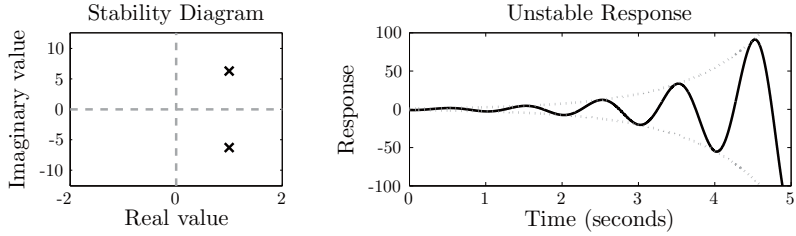
[8]The theory and proof behind the following information is an important result in electrical engineering, not presented here only for brevity. Information can be found in most introductory signals and systems books, such as [174], or [6] for simpler and more intuitive explanations.



(a) Stable system



(b) Marginally stable system



(c) Unstable system

FIGURE 6.7: Stability of a *linear* system can be inferred by looking at the *poles* (zeros of the denominator) of the relevant transfer functions. The real part of the poles is plotted on the horizontal axis and the imaginary on the vertical axis. In (a) we see a *stable* system because its two poles both have negative real parts. The simulated time domain signal (step response) oscillates but eventually settles to its *steady-state* value. In (b) is a *marginally stable* system because its two poles lie on the imaginary axis. The simulated time series is oscillatory. Finally in (c) we see an *unstable* system because its two poles have positive real parts. The simulated time series continues to grow indefinitely. The imaginary axis is called the stability boundary because it marks the transition from stable to unstable behavior. This stability analysis can be applied to non-linear systems that have been linearized, although an unstable system in the linearized system may be stable in the non-linear one. Biologically, the unstable linear response will be limited in the non-linear system and often give rise to a periodic bounded response.

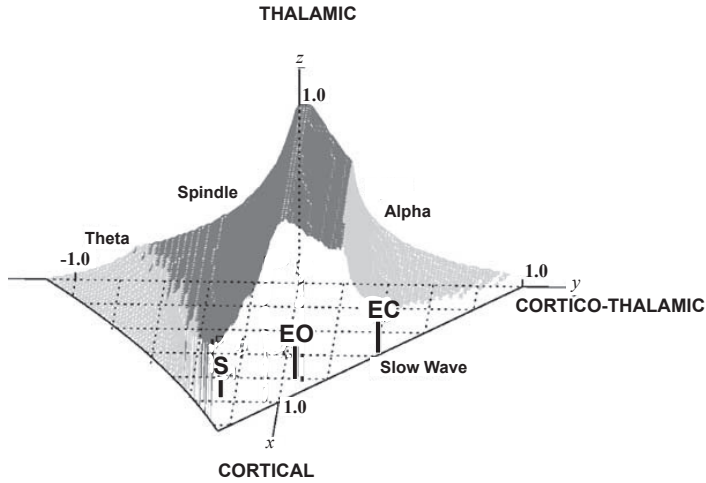
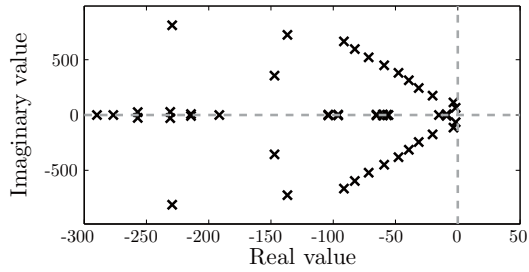


FIGURE 6.8: Stability boundaries of the meso-scopic model, shown in a reduced co-ordinate system that separates the stability of $C(s)$ in the x axis, $S(s)$ in the y axis and $T(s)$ in the z axis. Specifically the axes map to $x = \frac{G_{ee}}{1-G_{ei}}$, $y = \frac{S_d+S_i}{[(1-S_r)(1-G_{ei})]}$ and $z = -S_r\frac{\alpha\beta}{(\alpha+\beta)^2}$ with $S_r = G_{sr}G_{rs}$, $S_d = G_{es}G_{se}$ and $S_i = G_{es}G_{sr}G_{rs}$. Each shaded region shows a different type of instability that can be used to describe the EEG oscillations of different seizures. Stable regions can explain the non-epileptic EEG. Also shown are the location in parameter space of different states of awareness: eyes open (EO), eyes closed (EC) and one type of sleep (S). This figure is reproduced from [22], page 3, with permission from Oxford University Press.

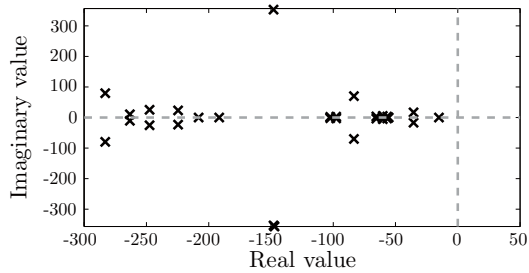
Using this theory and a reduced co-ordinate system defined by

$$\begin{aligned} x &= \frac{G_{ee}}{1-G_{ei}} \\ y &= \frac{S_d+S_i}{[(1-S_r)(1-G_{ei})]} \\ z &= -S_r\frac{\alpha\beta}{(\alpha+\beta)^2}, \end{aligned} \tag{6.17}$$

with $S_r = G_{sr}G_{rs}$, $S_d = G_{es}G_{se}$ and $S_i = G_{es}G_{sr}G_{rs}$. The authors of [146] have identified the stability boundaries of Equation 6.16 that are shown in Figure 6.8. Here a particular set of parameter values (κ) correspond to a behavior represented by a location in this three-dimensional plot. If the behavior of the non-linear system is oscillatory then the linearized system is unstable and its poles have crossed the stability boundaries shown. If instead



(a) $t_0 = 80$ms



(b) $t_0 = 0$ms

FIGURE 6.9: The importance of the delay t_0 between cortex and sub-cortex is demonstrated in the above example. The 10Hz alpha rhythm in (a) is only possible because poles can get close to the instability boundaries. When $t_0 = 0$ in (b) the poles are driven further away from the imaginary axis – these instabilities are more difficult to generate under normal conditions. A non-zero t_0 also introduces an infinite number of poles (not shown) allowing for greater variability in the behavior of the model.

the behavior is stable in the linearized system then these boundaries have not been breached.

The significance of this work is that *instabilities of different frequencies have been identified*. One parameter (or several) can be changed so that the behavior crosses one of these stability boundaries. This is representative of transitions between oscillatory and stable behavior, theoretically linked to transitions between seizure and non-seizure activity as well as common rhythms. The oscillations can take different frequencies and are able to describe the various types of seizures that are observed in the real EEG.

Next we describe how to interpret these stability boundaries in the context of EEG and epileptic seizures in particular.

6.3.3 Validation and Limitations: Rhythms in the EEG

The behavior of the meso-scopic model can be analyzed by studying how parameters affect the spectrum of the simulated EEG defined by Equation 6.16. While some parameters affect only the power of the signal (e.g., G_{sn}) others affect the range of the frequency spectrum (e.g., γ_e) or the position of the spectral peaks or *resonances* (e.g., α, β). Whether these resonances occur at all is largely determined by the inclusion of the delay $t_0/2$ between cortex and sub-cortex.

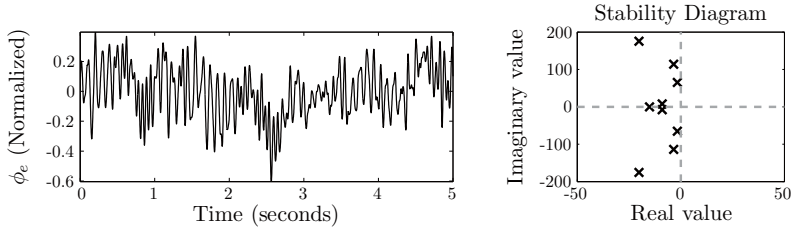
As an example Figure 6.9 shows the plot of the poles (stability analysis) of the model with identical parameters except $t_0 = 80$ms in (a) and $t_0 = 0$ms in (b). The delay is not only responsible for introducing more poles into the system (in fact, an infinite number) but also pushes the entire plot closer to the instability boundary, thus making rhythms more likely. *Without delays resonances occur very rarely.* This is true generally.

At this point in time results are such that models of similar nature to this meso-scopic model are the most suitable to describe epileptic seizures at the scale of the EEG, whilst maintaining relatively strong links to the biology beneath. Next we show that its dynamics are consistent with both normal and epileptic waveforms. The solutions presented are *numerical simulations* because, unlike the IF models, no analytic solution for the phenomenological model exists.

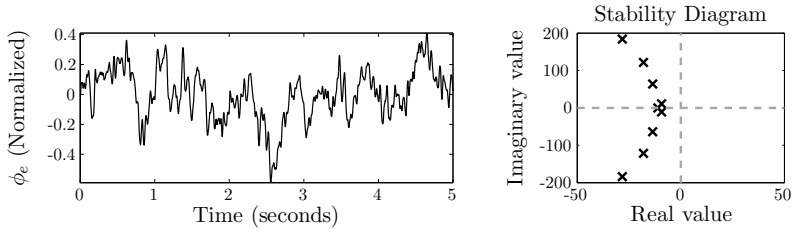
6.3.3.1 Simulating the Normal EEG

The meso-scopic model has been used to fit real single channel EEG data acquired from intra-cranial and scalp electrodes and was shown to produce waveforms and spectra similar to typical observed phenomena. Examples of these waveforms and their spectra can be found in Figure 6.10. (a) and (b) has been associated to the alpha rhythm with eyes closed and eyes open respectively, whereas (c) is similar to sleep-like waveforms. Their spectra, found in Figure 6.12(a), show their accordance with results shown in Figure 3.15 – a strong 10Hz resonance for the alpha rhythm, also present but to a much lesser extent in the eyes open example where the alpha activity is dampened by visual stimulus. During sleep it is the lower frequencies that dominate.

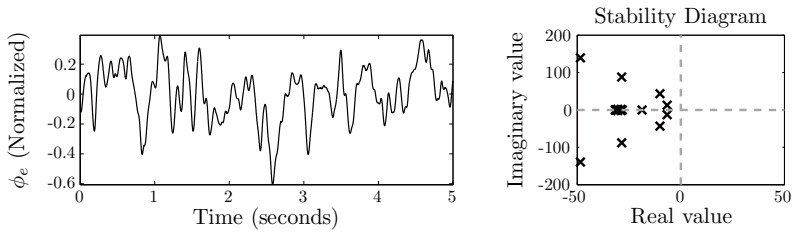
A plot of the poles is also provided. In all cases the system is stable because the instability boundary is not breached, but in the case of the alpha rhythm some poles are very close to the stability boundary. The magnitude of these poles indicates that the frequency of the dominant resonance is $f = \frac{w}{2\pi} \approx 10$Hz – i.e. consistent with the alpha rhythm. With the eyes open in (b) poles with similar imaginary magnitude also exist but these are no longer the ones closest to the instability boundary thus the 10Hz resonance does not dominate. During sleep in (c) the dominant poles are of much lower frequencies – the poles of larger frequencies are too distant to contribute significantly to the signal.



(a) Alpha rhythm – Eyes closed



(b) Alpha rhythm – Eyes open



(c) Sleep

FIGURE 6.10: Example simulated waveforms that have been associated, by visual inspection, to typical EEG signals, as labeled. These are stable examples because all poles have negative real parts – the waveform is driven by the stochastic input ϕ_n, modeled as white noise. (a) is more oscillatory than (b) or (c) because its poles are very close to the imaginary axis. The dominant frequency of oscillation corresponds to the closest pole – roughly 10Hz. The simulated PSDs of these waveforms can be found in Figure 6.12(a).

The model is validated by these relationships to real experimental data and also by the location of the waveforms in the stability diagram presented in Figure 6.8. For example, changes between eyes open and eyes closed is predominantly a shift in the y axis representative of the cortico-thalamic stability. Closing the eyes removes the influence of external stimulus relayed by the thalamus. Sleep stages occur close to instability boundaries, which is supported by the fact that many people experience seizures in the transition in and out of sleep. This figure also postulates that the alpha rhythm is itself an 'instability', and a region of parameter space that corresponds to an oscillation of roughly 10Hz has been identified.

6.3.3.2 Simulating the Seizure EEG

Recall that epileptic seizures can be identified in the linearized system from their instabilities. Linearization is an *approximation* to non-linear behavior, but if the linearized system is unstable it can no longer predict the behavior of the non-linear one (by and large). Thus an unstable linearized system can be purely oscillatory in the non-linear domain because 'new' solutions exist in the non-linear behavior.

In Figure 6.11 we simulate three types of seizure-like activity possible with the meso-scopic model. In all cases the linearized system is unstable because at least one of the poles have positive real parts. In (a) and (b) the oscillations are caused by instabilities at 20Hz. In contrast the waveform in Figure 6.11(c) shows an oscillation of much lower frequency – roughly 3Hz.

Notice that in (a) and (b) the poles are capable of estimating the frequency content observed in the simulated signal, as shown in Figure 6.12(b). This is because in both these there is only one pair of unstable poles. However in (c) three pairs of poles have had time to cross the imaginary axis. In this case many other bifurcations may have taken place, and the linear system is no longer adequate to describe what is observed in the simulated time series. For example, the 3Hz activity is not predicted by the magnitude of any of these poles. Analysis of multiple bifurcations is beyond the scope of this book.

Each oscillation is the result of a different parameter κ that corresponds to a different location (breaching the instability boundaries) in Figure 6.8. The different types of instabilities have been identified in [146] to be representative of the different types of epilepsies, as labeled in Figure 6.8. Whereas the *spindle* instabilities shown in Figure 6.11(a) and (b) are reminiscent of the focal (i.e., secondarily generalized) seizures, the 3Hz waveform in (c) is related to absence epilepsy. That the different instabilities shaded in Figure 6.8 cover distinct regions supports the theory that different mechanisms are responsible for focal and non-focal epilepsies. Variations in waveforms between epilepsies of the same type are explained by slight changes in the parameter space.

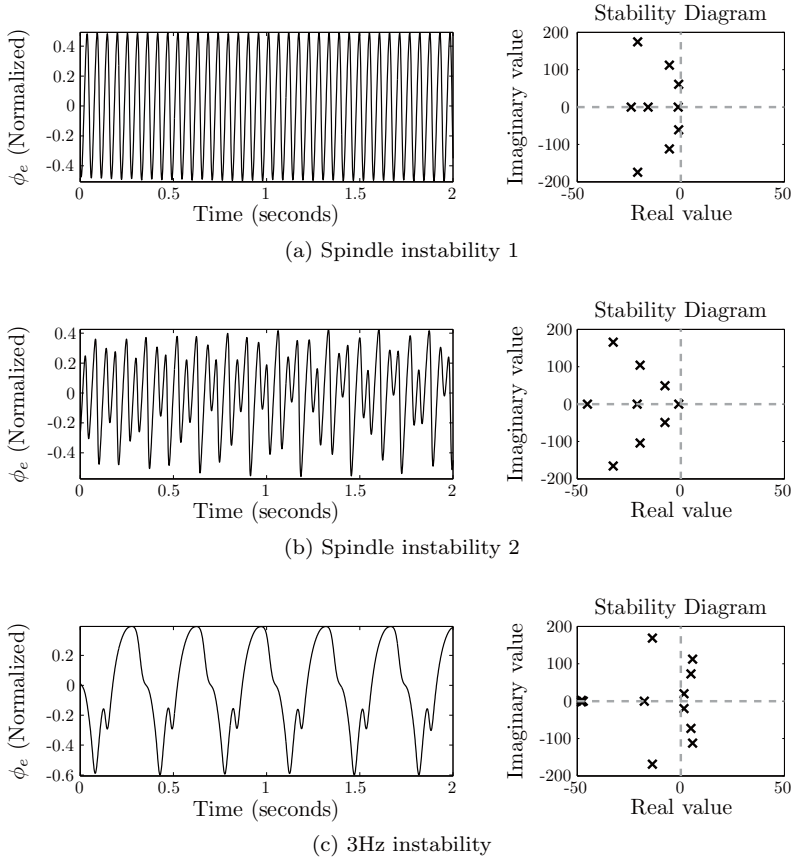
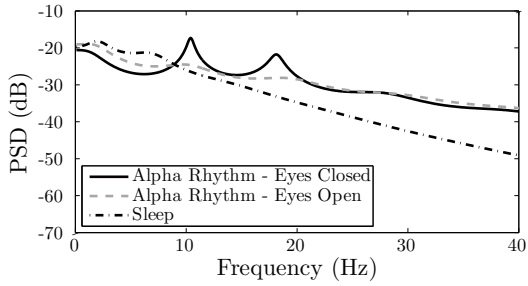
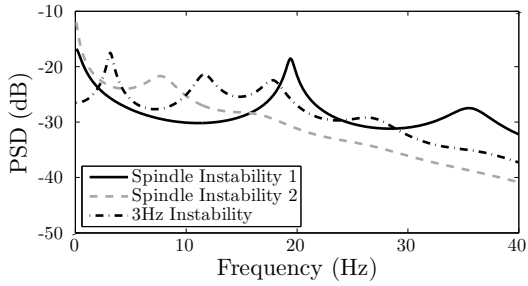


FIGURE 6.11: Example simulated waveforms corresponding to different types of instabilities. The input ϕ_n has hardly any effect on the output. (a) and (b) show *spindle* instabilities dominated by a 20Hz signal. These look like and thus have been associated with stages in seizures caused by focal epilepsies. (c) shows a 3Hz waveform often observed in absence epilepsies. The spectra show that in (a) and (b) the dominant frequencies correspond to the unstable poles that are closest to the imaginary axis. This is not so in (c) where there are multiple unstable poles, meaning that the system may have undergone further bifurcations. The linearized system in this case is incapable of predicting the behavior of the simulated signal. The simulated PSDs of these waveforms can be found in Figure 6.12(b).



(a) Normalized PSD of normal states



(b) Normalized PSD of epileptic states

FIGURE 6.12: Spectra of model simulated waveforms. In (a) are the normal states simulated in Figure 6.10, and in (b) are the epileptic states simulated in Figure 6.11. The spectra in both figures show that the dominant frequencies are those corresponding to the pole that is closest to the instability boundary. Also note that the spectra in (a) closely correspond to the real EEG examples provided in Chapter 3.

6.3.3.3 Caution

Even with increasing volume of support for this type of meso-scopic models it is important to remember that their validity remains largely speculative. For example this model assumes that the generation of the alpha rhythm is possible because of the cortico-thalamic loop. Model simulations and analysis support this theory. However other models can provide similar support for the theory that the alpha rhythm is due to the cortico-cortical interactions (see [95]). Without any experimental evidence to refute either paradigm each research group continues seeking and finding support for their own theory. The researchers of the meso-scopic model here have gone to great lengths to link these models to normal EEG, epileptic EEG, sleep disorders, etc. However *the models are still local and they cannot explain everything.*

Even though close relationships between data and model are demonstrated

in this section it is only the *phenomena* that is explained by the model. The more complex effects (e.g., cognition) are beyond the scope of this level of models. Furthermore although these models are used to explain global phenomena their analysis is performed at a local scale – the inference to global dynamics remains speculative.

In any case the phenomenological models are the closest and most developed models capable of describing EEG dynamics in terms of meso- if not micro-scopic biological detail.

6.3.4 Relationship to Micro-Scopic Models

Let us now look at the connection between the micro-scopic (IF) models presented in Section 6.2 and the meso-scopic models presented here. Even though the meso-scopic models incorporate sub-cortical interactions both models are not inconsistent with one another. Both begin by describing the behavior of ensembles of neurons such as those in Figure 6.1(b), and although not specifically in the same proportions, there is no reason why these ensembles cannot be the same.

Assuming they are both descriptions of similar networks, then the difference lies in the experimental data used to validate them. The IF models are easily linked to and supported by experimental evidence measured at the micro-scopic scale, but little is known about how they relate to macro-scopic measurements. The meso-scopic models, on the other hand, are easily linked to macro-scopic (EEG) measurements but the link to the micro-scopic scale is not clear.

Ideally we would like a unified model that is easily related to both micro-scopic and macro-scopic activity. We don't need *direct* measurables, but if we were to understand the link between measurements at different scales then disorders such as epilepsy that transcend spatial and temporal scales could be better understood. For example, epileptic seizures are a macro-scopic phenomena because large regions of the brain must be involved to sustain such activity. However a seizure often starts at a meso-scopic scale (e.g., at an abnormal tissue focus) and the *reason* that it starts is likely explained by physiology and biochemistry of the micro-scopic tissue.

In any case the mathematical links between such models remain largely unexplored, even though definitive links do exist. Granted some differences are irreconcilable because the fundamental strategy to derive the equations is not the same. Most notable is the differentiation between types of neurons in the meso-scopic model, where a different $Q_a(t)$, $V_a(t)$ and $\phi_a(t)$ exists for each of $a = e, i, r, s$[9]. This differentiation is important in demonstrating the co-incidence between the model and the EEG: It is only the cortical excitatory field ($\phi_e(t)$) that affects this measurement, but this variable does not exist

[9]Interestingly this differentiation is not uniform throughout, for example the parameters describing mean threshold potentials $\bar{V}_{th}$ and its variance σ are assumed the same for all neurons.

in the micro-scopic model. A considerable amount of work is necessary to re-write these equations in this form.

Other differences are not so large. For example, many links can be made between the parameters and variables in both models, the most obvious being mean spiking rate ($\nu(t)$, $Q_a(t)$) and mean membrane potential ($\mu(t)$, $V_a(t)$). These are the same quantities in each model, the only difference being that in the meso-scopic case they are dependent on the neuron type a and in the micro-scopic they are common to the whole population. Would an average taken over $a = e, i$ equate the two directly? (Should it?) Alternatively it may be relatively simple to incorporate different populations of neurons into the IF models.

The links between the IF and meso-scopic model must be made *explicit* so that the meso-scopic model can be validated – or alternatively in-validated – physiologically as well as phenomenologically. Even though solutions will likely remain numerical rather than analytic they would be corroborated by biological evidence, an important step toward the creation of a single unified model in which the inter-scale interactions are considered (see Figure 1.11).

Alternatively a new model can be created in which each of the sub-populations of the meso-scopic models in Figure 6.1(c) is replaced by the IF equations and linked in a similar way. However it seems wasteful to disregard the large volume of work on the meso-scopic models. The smaller task is in identifying these relationships.

6.4 Macro-Scopic Models (Future Outlook)

No computationally tractable macro-scopic model of brain activity based on the active sources in the brain exists, even for a few centimeters of cortex. Recall that the behavior of the single neuron is a subset of the behavior at a (meso-scopic) network, and these are themselves a subset of the behavior at the macro-scopic level. Joining such sub-systems together quickly becomes (mathematically and computationally) problematic.

However until mathematical analysis and computational simulations can catch up it is still worth discussing the *framework*, consistent at all spatial and temporal scales, that macro-scopic models may take. In Chapter 1 (see Figure 1.11) it was postulated that larger scale models must account for the dynamics

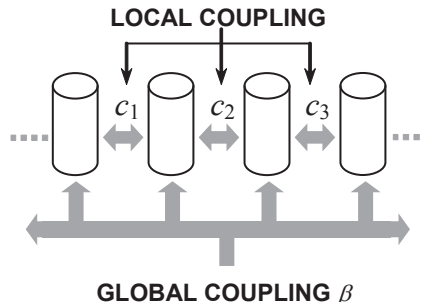


FIGURE 6.13: A macro-scopic model can be derived by joining sub-systems such as those in Figure 6.1(b) and (c). Systems are coupled with strength $c_{1,2,3,...}$ as shown. If all sub-systems are coupled to each other the complexity of the macro-scopic model becomes unmanageable. A simpler approach is to have a single coupling parameter β that describes the average strength of interaction between sub-systems.

at smaller scales, while at the same time large-scale dynamics influence the behavior of the smaller scales. While this is true, models at both (or all) scales must exist for such a unified solution to be formulated. The focus is then on developing a framework consistent at all scales. The basic idea of an approach to macro-scopic modeling is presented here. The reader interested in more detail can refer to [122] and [120].

A meso-scopic model is derived by connecting micro-scopic sub-systems together. Similarly it follows that a macro-scopic model can be constructed by connecting these meso-scopic sub-systems together, as shown in Figure 6.13, so as to be consistent with the organization of the brain as a whole. Connections such as cortico-cortical projections, which by far outnumber thalamo-cortical ones but have been ignored in both IF and phenomenological models, can be included.

However the problem quickly becomes intractable because not only must each sub-system be modeled by the described equations, but more parameters indicating the strength of interactions between them must be included. In addition the macro-scopic models cannot ignore the effects of volume conduction discussed in Chapter 2. The spatial scales involved are larger and the conducting properties of the materials must be added to complement active sources. The complexity of the problem grows very rapidly.

A gross simplification is to use a single extra parameter, β, that represents how much of the dynamics are due to local versus global activity. β is referred to as the *background excitability control parameter* that determines the amount of coupling between sub-systems. This is a simplistic approach that is nevertheless capable of explaining much of the *global* EEG dynamics. This idea is elaborated in some detail in [120] and [122]. They postulate that in

states of low cognition (e.g., anesthesia, sleep, resting alpha activity) β is (arbitrarily) higher and it is the global modes that dominate. Slower frequencies are more evident, particularly in scalp EEG which due to spatial integration and filtering are automatically biased toward cortical global activity. More complex/alert states of cognition lead to a reduction of low frequency components and an increase in high frequencies in the EEG, suggesting greater independence in activity between different sites in the cortex. Here it is the local models that dominate, and β decreases accordingly.

In their discussions the parameter β is dependent on non-specific chemicals (e.g., neuromodulators), but can be related to the effects of slow acting neuromodulators that are responsible (in part) for determining the excitability of regions of cortex to local versus global contributions. The thalamus is also thought responsible in the regulation of the amount of coupling between sub-systems.

Although the results are abstracted from much of the micro-scopic biology, the framework is very well developed and experimentally substantiated. Their work concentrates on global aspects such as the effects of volume conduction, spatial filtering and the shape and size of the head, but it does not deny the importance of local dynamics. *This framework does not exclude the possibility of incorporating smaller scale models to represent the active mechanisms of the brain, their point being simply that the local models should not underestimate the importance of the global structure.*

Thus all models presented in this chapter are consistent with one another, or rather they are not inconsistent. However the links between meso- and macro-scopic models remain more tenuous than those to IF models. At the smaller scales gross simplifications that ignore biology as well as the characteristics of the organization of the brain have been made for the sake of a tractable solution. In particular the cortico-cortical connectivity, which by far outnumber the sub-cortical projections in the human brain, have been ignored by both micro- and meso-scopic models.

So many problems remain at the smaller scales that it is unlikely that links to more global models will be considered in the near future. However the importance of cortico-cortical connectivity is undeniable, and the framework presented here allows for their inclusion. It is abstract enough to be consistent at all scales, provided suitable progress is made in both mathematics and computational power.

6.5 Practical Use of Models

In this section we explore two ways in which the models presented in this chapter can be used to make practical progress in the field of epilepsy. In Section 6.5.1 we use the models to help us *understand*, *investigate* and *infer*

how seizures are generated and how they can subsequently be terminated. In Section 6.5.2 we see how we can directly *apply* these models to tell us about the EEG, even when the models themselves are very simple.

6.5.1 Epileptic Seizure Generation

Since relatively little is known about epileptic seizures (given the amount of research undertaken) it is not surprising that the dynamics of these models have been used to try to explain it. The focus of research at the moment lies on understanding the generation and subsequent termination of seizures. To a large extent the high-level causes remain unexplored, for example hyperventilation, sleep deprivation and stress, that increase the likelihood of seizures. Future research may reveal their impact upon the generation of seizures.

Mathematical model-based epilepsy research studies how *parameters* in the model can be altered to create seizure-like activity. The link to the underlying physiology depends upon how closely the parameters describe the underlying physiology.

While these models are not specifically designed for studying epileptic seizures they can be used because normal brains are also capable of seizing. If a seizure is a state to which the activity of this brain is altered, the question is then (a) how does it start and (b) how does it spread, as discussed in Chapter 1? How it stops is important as well. Because the meso-scopic models remain local, the question of spread cannot be addressed unless coupled with more cortical columns. The material presented here concentrates on *initiation* and in part on *termination* of epileptic seizures.

6.5.1.1 Seizure Initiation

Identifying the region of parameter space that corresponds to seizure-like activity can at most explain the state that the brain is in when seizing – by itself this information is insufficient to explain how the system is driven to this state. *Transient* analysis of the system, in the form of altering the system parameters, can be used instead. In [22] a single parameter is chosen and changed over time to explain generation and termination of absence and focal seizures. The mechanisms behind each of these types of epilepsies are explained in the context of system stability. One example of seizure transitions is presented in Figure 6.14, where the parameter v_{se} that describes the strength of interaction between cortex and sub-cortex is varied as shown to drive a brain behaving normally into seizure.

It is important to understand that the examples shown in Figure 6.11

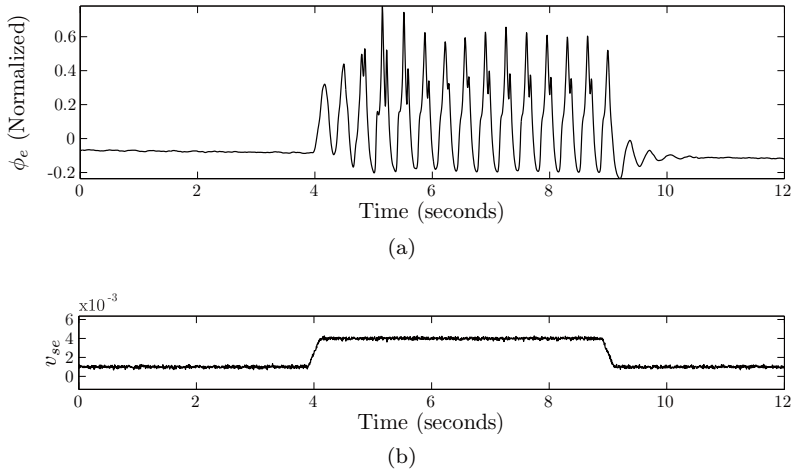


FIGURE 6.14: Sample transition in and out of system stability to replicate transitions in and out of seizure. v_{se} is varied as in (b), with additive stochastic white noise. In [22] such methods have been used to infer the importance of v_{se} on the generation of these types of seizures but its significance remains uncertain seeing as many parameters can be used to generate similar waveforms.

and Figure 6.14 are *instances* rather than explanations. Virtually all parameters can be varied sufficiently so as to drive the system to instability/seizure. Blindly searching through parameter space does not lead to a better understanding, especially since it is known that not all parameter changes are physiologically realistic. The research should be corroborated by experimental data or animal models.

Finally, a distinction must be made between the model used here and a model of an epileptic focus. The meso-scopic model was designed for a normal region of the brain. Focal epilepsies often involve abnormal composition of neural tissue and its behavior cannot necessarily be explained by these models even if the range in the parameters is large. Thus the information provided here could simply be how normal tissue is driven to seizure, in which case it is an explanation of spread rather than initiation. Nevertheless this type of information should not be dismissed because many epilepsies, even focal ones, begin with changes in the anatomy that are not obvious. In order to understand the behavior of abnormal focal neurons it may be necessary to employ the IF models presented in Section 6.2.

6.5.1.2 Seizure Termination by Electrical Stimulation

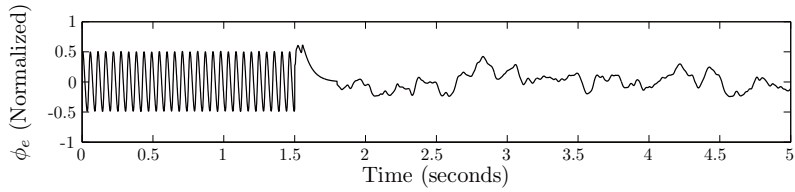
A seizure most often terminates naturally by unknown changes in parameter space. The mechanisms remain speculative even though links to model parameters can be made at times. Intervention to abate a seizure is desired because continuing seizures can cause brain damage. While drugs are a drastic means of doing so when appropriate medical supervision is available[10]an ideal solution is full automation of the process. That is, we want an automated method that can abort a seizure on-demand once it has already begun. Recently a lot of research has been dedicated to implantable *electrical stimulators* used to stop a seizure once it has started. This is in contrast to stimulators such as VNS that deliver continuous stimulus so as to control seizures, described in Section 1.1.5

Experimentally there is no clear evidence that this class of stimulators works. Here we briefly look at the possibility that electrical stimulators to terminate seizures are at least *theoretically* viable. Since the models are based on physiology, does the introduction of external current or voltage sources change the dynamics? Can stimulation be used to change the seizing brain to a stable state?

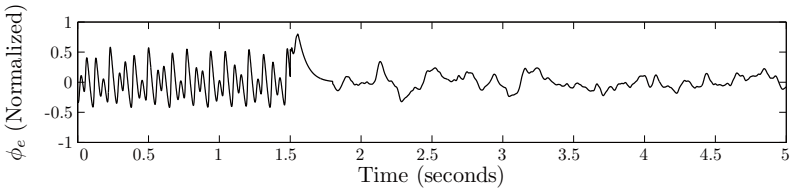
Figure 6.15(a), (b) and (c) shows the same seizure-like waveforms as in Figure 6.11 but with an external stimulus applied at time $t = 1.5$ seconds. The stimulus is shown in (d) – a *positive* field that sums with local fields. In all cases the seizure-like activity is stopped, and the system is driven to stability. In the spindle instabilities representative of the focal epilepsies ((a) and (b)) the termination occurs when the stimulus is applied to either cortex or subcortex. For the 3Hz instability in (c) representative of non-focal epilepsies the termination is only successful when the stimulus is applied to the thalamus. This is consistent with the view that the non-focal epilepsies and in particular absence seizures are thalamic in origin. Whether the state that the brain is driven to with stimulation is physiologically safe and whether the seizure returns once stimulus ceases is not addressed by these simulation.

While a lot of work must be done to translate the indications contained in these figures into experimental or clinical trials, the analysis is useful because it implies that at least theoretically electrical stimulation can be used to terminate seizures. Before experimental results follow, closer links to the biology and greater mathematical formalism are necessary so that the correct stimulation parameters (e.g., duration, amplitude) can be inferred from the models. Future work may also be able to facilitate understanding of *why* electrical stimulation should work, if indeed it does.

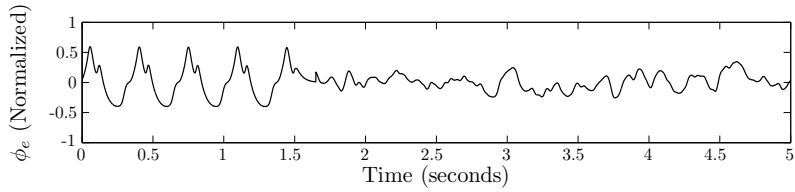
[10]Here we are talking about drugs used to abort seizures, that is, they are administered on demand. This is in contrast to drugs used continuously to control seizures, described in Chapter 1.



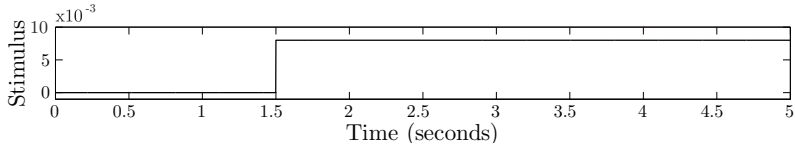
(a) Spindle instability 1



(b) Spindle instability 2



(c) 3Hz instability



(d) Applied stimulus

FIGURE 6.15: Example transitions out of seizure with the application of an external electrical stimulus, shown in (d). The spindle instabilities are stopped by cortical stimulation, whilst the 3Hz instability can only be stopped with thalamic stimulus. This is consistent with the belief that non-focal seizures are thalamic in origin. The simulations show that at least theoretically electrical stimulation may be used to stop seizures once they have started. In the figures above signals after stimulus have been amplified so that the state of 'normal' versus 'seizing' activity is more obvious.

6.5.2 Limitations of the EEG

So far in this chapter we have described models that attempt to mimic the physiology of the brain to as much detail as possible. At the same time many simplifications were necessary so that the mathematical problem is tractable. Whilst making the problem easier to compute, removing detail inevitably led to models that were difficult to link to either physiology or to macro-scopic measurements such as the EEG.

It is important to understand the limitations of what simplified models can tell us about a system, but having limitations does not make a model unusable. In this section we highlight this by discussing the limitations of the EEG using a simple linear model of the brain for which the model details itself are not known.

Consider the EEG as the output of a dynamical system that represents the brain. Suppose for a moment that the actual brain dynamics can be represented by a discrete time linear system

$$z[n+1] = Fz[n] + \eta[n]. \tag{6.18}$$

This is an instance of 3.2, where $\eta[n]$ is a disturbance or noise signal, presumably representative of external input into the cortex. $\eta[n]$ may be modeled as a white noise, but it suffices to assume that it is independent of the state $z[n]$. Also, assume that the EEG signal $y[n]$ is simply a linear map from the state,

$$y[n] = Cz[n]. \tag{6.19}$$

Let the dimension of $z[n]$ be N and the dimension of $y[n]$ be M, with a typical M much smaller than N.

It is reasonable to postulate that the state is a very high dimensional object (several billion variables at least). Over a limited period of time, the period over which the noise characteristics may be considered stationary, it is plausible to think of F as representing a collection of interconnected oscillators. This means that all the eigenvalues of the matrix F have modulus 1.

In a generic context it is possible to reconstruct the state $z[n]$ from observations of the EEG signal, the system output $y[n]$, over a sufficiently long window of observation (say using the celebrated Kalman filter). In principle this means that at least N/M time samples are required, that is, in order to find $z[n]$ it is at least necessary to observe $y[n]$, $y[n+1]$, $y[n+2]$ ... till $y[n+N/M]$. *With N being of the order of hundred billion and M of the order of hundred, this means that a few billion time samples are required!* Unfortunately, such a long observation window (longer than 500 hours, using 512Hz sampling) is not at all compatible with the horizon over which the dynamics can be considered stationary. Neither are such observation windows practical.

The problem is actually worse. With F being a matrix with eigenvalues of modulus 1, the uncertainty in $z[n]$ grows linearly over time (because the variance of $z[n]$ is linear in time, driven by the variance of the noise input $\eta[n]$).

Even if it were possible to reconstruct $z[n]$ after having observed $y[n]$, $y[n+1]$, till $y[n+N/M]$, this knowledge is virtually meaningless as the difference between $z[n]$ and $z[n+N/M]$ is equally of the order of N/M. The Kalman filter is essentially meaningless in this context, as the variance of the error between the estimate of the state and the state variable itself will grow indefinitely.

The same idea can be gleaned from an information theoretic point of view. The amount of information available at each measurement point is captured by the number of bits obtained, $M \times B = MB$, where B is the resolution of the recorded EEG in bits per sample. Over a horizon time T the total number of bits gathered is $T \times M \times B = TMB$. On the other hand, $z[n]$ contains N states, and in order to know $z[n]$ to within b bits precision requires $N \times b = Nb$ bits. The brain map, being a collection of interconnected oscillators, is known to be measure preserving. Furthermore the uncertainty in $z[n]$ grows due to the noise, linearly in time, say at a rate g. So, in order to learn the state $z[n]$ it is necessary that TMB exceeds $Nb + gT$, at some time T in the future. Again it follows that the observation window has to be large enough, and in order for this to be true then

$$T > Nb/(MB - g). \tag{6.20}$$

Furthermore for this to be meaningful, it has been assumed that at any instant in time the amount of information gathered MB is actually larger than the amount of information destroyed by the noise, i.e., $MB > g$. This is a generous assumption that is impossible to be certain of, but it is used here because otherwise the reconstruction of $z[n]$ quite simply cannot be done. The conclusion is again that it will take in the order of a few thousand hours of EEG observations in order to learn the state of a "stationary" brain.

The above argument certainly undermines any attempt of using embedding ideas presented in Section 3.3.4 in order to recover brain dynamics. Indeed apart from the fact that in practice it is not possible to know the map F, which was tacitly assumed knowledge in the above argument, the observation window required to achieve this "in principle" possible state reconstruction is not realistic at all.

Does this imply that it is not possible to detect or predict an epileptic event from an EEG record? No, because knowledge of the state is not called for to make this decision. It is of course feasible to consider the condition of an epileptic seizure as being described by some binary valued decision map from the state: $E(z) = 1$ indicating the presence of an epileptic seizure and $E(z) = 0$ indicating no epileptic seizure active. That such a map E exists is a consequence of the definition of state itself. Nevertheless, it is also clear that deciding on the presence of an epileptic seizure can be regarded as a decision from the history of EEG observations as well, after all that is the way clinicians make this decision. The issue is one of how to learn this decision map from a collection of examples, which in this book has been approached using machine learning ideas, through the construction of an expert system.

6.6 Conclusions

The formulation of mathematical models of neural dynamics depends largely on the scale of interest, both spatial and temporal. This in turn depends on the application. The models presented here are geared toward the understanding of epileptic seizures, although they are in no way restricted to this application alone.

In the understanding of neural dynamics of epileptic seizures it is important that a model be able to explain normal as well as seizure behavior because an epileptic brain has seizures only part of the time. Without this, transitions in and out of seizures cannot be studied. The differences between the epileptic brain and the normal brain must be encoded in the parameters of the system.

Having said this, one of the major limitations that current models have is that their analysis is largely numerical and limited to stationary behavior. Little work has been done toward understanding the time evolution of neural dynamics; in particular the transition in and out of seizures remains inconclusive.

At this stage neural models at the meso-scopic scale are still being validated, that is, they are in the process of being compared to real data to see if they have been formulated correctly. At the same time these models are being used to gain insight into the nature and source of epileptic seizures.

Even with this limited understanding it is still possible to utilize these models in practical applications, for example to infer the limitations of the EEG as is done in Section 6.5.2. This casts further doubt on the use of measures that utilize embedding theory for the prediction of epileptic seizures. The next chapter discusses, in more general terms, the question of how predictable seizures really are.

6.A Physiological Parameters and Notation

Notation

Subscript e	Refers to excitatory neurons in the cortex.
Subscript i	Refers to inhibitory neurons in the cortex.
Subscript r	Refers to neurons in the TRN.
Subscript s	Refers to neurons in SRN.
Subscript n	Refers to input external to the network, most often excitatory.

Single Neuron Model Equation Symbols

c	Membrane capacitance.
τ_p	Membrane time constant.
τ_r	Absolute refractory period.
V_P	Resting potential.
V_{TH}	Firing threshold.
V_R	Reset potential (after firing).
$V_{re,ri}$	Constant reversal potentials of excitatory/inhibitory neurons.
α_a	Conductance of synapse from neuron type a.

Network Architecture

N	Number of neurons in network.
N_a	Number of neurons of type $a = e, i$ in the network.
C_{ab}	Number of synapses from neuron type b to *each* neuron of type a.
ν_a	Firing rates of neurons of type a.
γ_{ab}	Strength of synaptic inputs from neurons of type b onto neurons a, with $\gamma_{ab} = \nu_b C_{ab}$.

Average Network Behavior

Q	The maximum mean firing rate (of action potentials) of a neuron population.
$\bar{V}_{TH}$	The mean threshold for a neuron to fire an action potential.
σ	The variance of $\bar{V}_{TH}$, assumed to have Gaussian distribution.
γ_e	A measure inversely proportional to the range of axons of cortical neurons. $\gamma_e = \frac{\upsilon_e}{r_e}$ where υ_e is the propagation speed of the action potential.

6.B Summary of IF Model

Variables and Steady State Solutions (subscript '0')

$\nu(t),\ \nu_0$	Mean spiking rate.
$CV(t),\ CV_0$	Coefficient of variation of inter-spike interval.
$\mu(t),\ \mu_0$	Mean membrane potential.
$\sigma(t),\ \sigma_0$	Standard deviation of membrane potential.
$\tau(t),\ \tau_0$	Effective membrane time constant.

Model Specific Functions

$V(t)$	Membrane potential, defined between the inside and outside of a neuron's membrane.
$I_{leak}(t)$	Current due to passive leak of membrane.
$I_{spike}(t)$	Current describing the spiking mechanism of the neuron.
$I_{syn}(t)$	Current describing the effect of synaptic input to the neuron.
$\wp(t)$	Recurrent input function to each neuron.
$P(V,t)$	Probability distribution of membrane potential $V(t)$.
$P^*(V)$	Stationary solution to $P(V(t))$.

Changes in the membrane potential $V(t)$ of a neuron cell body are dictated by incoming current $I_{syn}(t)$ and outgoing currents $I_{leak}(t)$ and $I_{spike}(t)$

$$c\frac{dV(t)}{dt} = -I_{leak}(t) - I_{spike}(t) + I_{syn}(t). \tag{6.21}$$

Spiking currents occur because the membrane voltage reaches threshold V_{TH}

$$I_{spike}(t) = c(V_{TH} - V_R)\delta(V(t) - V_{TH}). \tag{6.22}$$

Passive leak currents exist to return the membrane potential to the resting state V_P

$$I_{leak}(t) = \frac{c}{\tau_p}(V(t) - V_P). \tag{6.23}$$

The synaptic inputs here are *conductance*-based currents modeled as

$$I_{syn}(t) = \sum_{a=e,i,n} c\alpha_a(V(t) - V_a)\wp_a(t), \tag{6.24}$$

where α_a is the trans-membrane conductance of neurons of type a, and $\wp_a(t)$ is a function that defines the recurrent input to neurons of type a. This

recurrent input is determined by the number of pre-synaptic spikes, summed temporally over a neural population, accounting also for a transmission delay.

Connecting N sets of such equations to form networks of neurons quickly becomes an intractable problem. Instead simplifying assumptions can be made over a group of neurons, as listed in Section 6.2, leading to a model that describes the probability distribution of $V(t)$. This distribution $P(V(t))$ follows the Fokker-Planck relationship, which for the stationary case $P^*(V) = P(V(t \to \infty))$ is given by

$$\tau^* \frac{\partial P^*(V)}{\partial t} = 0 = \Theta^2 \frac{\partial^2 P^*(V)}{\partial V^2} + \frac{\partial}{\partial V}\left[(V - M)P^*(V)\right], \tag{6.25}$$

with M and Θ corresponding to the mean and variance of $P^*(V)$ respectively. The advantage of this model is that analytical solutions exist for ν^*, CV^*, μ^*, σ^* and τ^*. These along with derivations can be found in [108]. Their simulated values comply with in vivo observations of a group of neurons.

6.C Summary of Phenomenological Model

Model Specific Functions and Operators

$Q_a(t)$	Mean spiking rate of a group of neurons of type $a = e, r, s$. Also known as pulse density. This is related to the cell body potential V_a.
$V_a(t)$	Mean membrane potential after inputs from the dendrites have been summed and filtered. This is relative to a resting potential.
$\phi_a(t)$	Potential fields induced by a group of neurons of type $a = e, i, r, s$.
$D_{\gamma e}$	Characterizes the low pass filtering effects that cortical damping has on the potential in the cortex.
$D_{\alpha\beta}$	Characterizes the low pass filtering effects that dendrites have on incoming signals.

Parameters κ (Steady State)

v_{ab}	Parameters that characterize the effects that neurons of type b have on neurons of type a.
G_{ab}	Gain of each stage in the linearized system, proportional to v_{ab}.
β, α	Parameters that characterize low pass filtering of dendrites. In biology these parameters represent the inverse rise and fall times of the potential produced by an impulse at a synapse.
$t_0/2$	Propagation delay between cortical and subcortical systems.
ϕ_n	The external input to the brain, assumed constant.

The mean spiking rate (see Figure 6.4) is described by equation

$$Q_a(t) = \frac{\mathbf{Q}}{1 + e^{-[V_a(t) - \bar{V}_{TH}]/\sigma}}, \quad a = e, s, r, \tag{6.26}$$

where $\mathbf{Q}$ is the maximum firing rate of a neural population and σ is the variance of $\bar{V}_{TH}$, all quantities defined in Appendix 6.A.

The potential fields $\phi_a(t)$ generated in the dendrites are proportional to the firing rate, as explained in Chapter 2. Excitatory neurons in the cortex experience low pass filtering ($D_{\gamma e}$) because their axons have non-negligible lengths.

$$\begin{aligned}
D_{\gamma e}\phi_e(t) &= Q_e(t) \\
\phi_i(t) &= Q_e(t) \\
\phi_s(t) &= Q_s(t) \\
\phi_r(t) &= Q_r(t) \\
D_{\gamma e} &= \frac{1}{\gamma_e^2}\frac{d^2}{dt^2} + \frac{2}{\gamma_e}\frac{d}{dt} + 1 = \left(\frac{1}{\gamma_e}\frac{d}{dt} + 1\right)\left(\frac{1}{\gamma_e}\frac{d}{dt} + 1\right).
\end{aligned}$$

In general fields $\phi_a(t)$ are assumed proportional to firing rates, with the exception of $\phi_e(t)$ that experiences damping by the dendritic tree because excitatory neurons have longer projections. Notice that $\phi_i(t)$ is proportional to $Q_e(t)$, another simplification that makes the model more tractable by removing a variable.

The mean membrane voltage $V_a(t)$ of each neuron type is determined by the strengths of interactions with the fields $\phi_a(t)$. Dendrites filter the incoming dendritic input ($D_{\alpha\beta}$).

$$\begin{aligned}
D_{\alpha\beta}V_e(t) &= v_{ee}\phi_e(t) + v_{ei}\phi_i(t) + v_{es}\phi_s(t - \frac{t_0}{2}) \\
D_{\alpha\beta}V_r(t) &= v_{re}\phi_e(t - \frac{t_0}{2}) + v_{rs}\phi_s(t) \\
D_{\alpha\beta}V_s(t) &= v_{se}\phi_e(t - \frac{t_0}{2}) + v_{sr}\phi_r(t) + v_{sn}\phi_n(t) \\
D_{\alpha\beta} &= \frac{1}{\alpha\beta}\frac{d^2}{dt^2} + (\frac{1}{\alpha} + \frac{1}{\beta})\frac{d}{dt} + 1 = \left(\frac{1}{\alpha}\frac{d}{dt} + 1\right)\left(\frac{1}{\beta}\frac{d}{dt} + 1\right).
\end{aligned}$$

The non-linear components $Q_a(t)$ in the model are approximated by linear functions around operating point – defined as an equilibrium when $V_a(t)$, $a = e, r, s$, are stationary. The resulting linearized system is described by the transfer functions in Figure 6.6. Formally we replace $\frac{d}{dt}$ by s, giving a simple algebra to quickly compute relationships between variables.

$$
\begin{aligned}
T(s) &= \frac{L(s)G_{sn}}{1 - L(s)G_{sr}L(s)G_{rs}} e^{-st_0/2} \\
S(s) &= \frac{L(s)G_{se} + L(s)G_{sr}L(s)G_{re}}{1 - L(s)G_{sr}L(s)G_{rs}} e^{-st_0} \\
C(s) &= L(s)G_{es} \, \frac{A(s)}{1 - L(s)G_{ei}} \, \frac{1}{1 - L(s)G_{ee}\frac{A(s)}{1-L(s)G_{ei}}}, \\
A(s) &= \frac{1}{(1 + s/\gamma_e)^2} \\
L(s) &= \frac{1}{(1 + s/\alpha)(1 + s/\beta)} \\
s &= j2\pi f \quad \text{for sinusoidal steady state, or } s \leftrightarrow \frac{d}{dt} \text{more generally.}
\end{aligned}
$$

The gains G_{ab} can be calculated from the steady-state behavior at the equilibria V_b^* and ϕ_a^*.

$$G_{ab} = \left.\frac{\partial Q_b}{\partial V_b}\right|_{V_b = V_b^*} v_{ab} = \frac{\phi_b^*}{\sigma}\left(1 - \frac{\phi_b^*}{Q}\right) v_{ab}.$$

Note that there may be many equilibrium solutions.

The overall transfer function $H(s)$ between input $\phi_n(s)$ and output $\phi_e(s)$ is

$$H(s) = \frac{\phi_e(s)}{\phi_n(s)} = T(s)\frac{C(s)}{1 - C(s)S(s)}.$$

7

On the Predictability of Seizures

> "The farther backward you can look, the farther forward you are likely to see."
>
> *- Sir Winston Churchill, (1874-1965)*

Prediction is a difficult task. To be able to forecast the future an in-depth knowledge of both past and present is required. The more complex a system is, the more information that is necessary to predict the same distance into the future. Although Winston Churchill's words were probably spoken in a more philosophical sense they hold true also for numerical analysis – *how far back in time must we look in order to be able to tell anything about the future?*

Let us first define what we mean by prediction and how it differs from detection. In a real system measurements are only available of the past. A detector uses these measurements to make a decision about this past and the present. A predictor, on the other hand, uses them to make a statement about the future. All predictors are forms of detectors, the difference being in the time at which the measurement is taken relative to the time for which the detection is made. For example, a seizure detector tells us that a seizure has begun, whilst a predictor tells us if a seizure will begin.

The weather example is once again useful in pointing out the difficulties of forecasting the future. Its prediction has become a daily convenience to many of us: we turn our TVs on and expect that there is some truth in what is said the weather will be like tomorrow, the day after, or even next week. But how are these predictions made? The most comprehensive and versatile solution is to create a *physical model* of the earth's atmosphere that accounts for all factors that affect the weather at any point on this earth. Humidity, topography, air pressure and even how many people decide to drive to work all need to be considered. This is an incredibly complex system for which even our most modern modeling tools may not be sophisticated enough.

> "When the number of factors coming into play in a phenomenological complex is too large, scientific method in most cases fails us. One need only think of the weather, in which case prediction even for a few days ahead is impossible. Nevertheless no one doubts that

> we are confronted with a causal connection whose causal components are in the main known to us. Occurrences in this domain are beyond the reach of exact prediction because of the variety of factors in operation, not because of any lack of order in nature".
>
> *- Science, Philosophy and Religion, Albert Einstein, (1879-1955)*

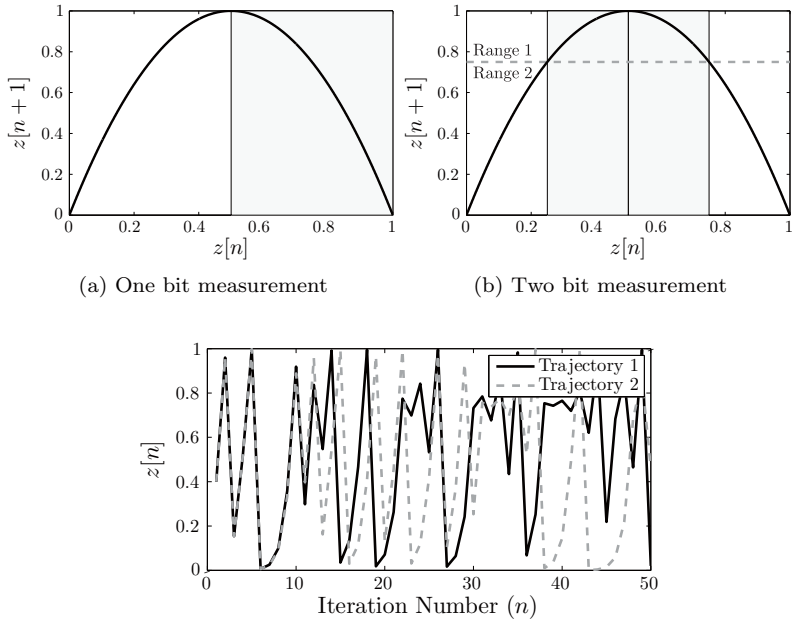
Thus although a physical model exists it may be too complex to realize, even with appropriate simplifications. In such cases an alternate approach is to base predictions on the assumption of a *stochastic* model. These models forecast future events by restricting the *range* of possibilities, unlike a purely deterministic model where an exact outcome is predicted[1]. Statistics are gathered over time and future events predicted based only on past observations. We limit our prediction to allow for variability, so we may predict that it is *likely* to rain, rather than it *will* rain. The more information that is collected, so long as this information is still valid, the more accurate the forecasted range is likely to be because we are predicting a statistic drawn from a probability distribution.

The statistics of a purely stochastic model do not themselves explain anything about the physical processes involved (even though physical processes may be inferred through the study of these statistics). These models are known as *black box* or *data driven* methods.

A model does not have to be purely physical or purely black box. A *gray box model* is when some aspects of a system are known and modeled by a physical process, and others are unknown and modeled by black box components. Weather prediction is an example of a gray box stochastic model because physical models with built-in allowance for the variability in measurements are used. They are effective because monitoring of weather patterns has been in place for many decades, and the statistics can be used by experts to interpret the outputs of the physical models. Although a full physical model is less limited in what it can tell us than black or gray box models, for such a complex system a gray box stochastic model is today the only practical solution.

The same observations hold true for seizure prediction – epilepsy is undoubtedly a behavior of one of the most complex systems that humans are trying to understand. Current "predictors" of seizures rely on the assumption of stochasticity simply because our models of the brain, presented in the

[1] In Chapter 6 physical models with stochastic components are described. This is unlike purely stochastic models where none of the physical signaling are considered in the mathematics.



(a) One bit measurement

(b) Two bit measurement

(c) Trajectory differences with 0.001 deviation at the start

FIGURE 7.1: The predictability of a map is affected by the measurement precision. In (a) and (b) is the map described by Equation 7.1 where $z[n]$ is measured using 2 values (one bit precision) and 4 values (2 bits precision) respectively. In (a) one bit cannot be used to calculate $z[n+1]$ because both ranges of $z[n]$ span the entire range of $z[n+1]$. In (b) two bits in $z[n]$ divides $z[n+1]$ into two ranges, thus $z[n+1]$ can be predicted to an accuracy of 1 bit. In (c) it is shown how even this simple map is very unpredictable under the presence of small errors. A difference of 0.001 in $z[0]$ of each trajectory results in very different behavior after only 10 iterations.

previous chapter, are not developed enough to be usable for this task. Black box statistical analysis applied on collected data such as the EEG are used to attempt a prediction.

The predictability of a system depends on how the collected data are *measured* and thus is a function of the technology used to make it. This applies to both physical and stochastic models. Think for example of a simple single dimensional system whose purely deterministic map is described by

$$z[n+1] = 4z[n](1 - z[n]), \quad 0 < z[0] < 1. \tag{7.1}$$

This says that if a measurement of the current sample $z[n]$ exists then the value of the next sample $z[n+1]$ can be predicted. This map is restricted to the range $0 < z[n] < 1$ shown in Figure 7.1. Now consider a measurement of $z[n]$ described by only 1 bit, that is, the only information available is whether $z[n] < 0.5$ or $z[n] \geq 0.5$ as in Figure 7.1(a). Then this measurement cannot be used to predict $z[n+1]$ because the measurement error results in a prediction range that spans the entire range of $z[n+1]$. However if 2 bits are used (the range of $z[n]$ is divided into 4, as in Figure 7.1(b)) then $z[n+1]$ can be predicted to an accuracy of 1 bit. In this system $z[n+k]$ can be predicted to 1 bit accuracy if $z[n]$ is known to k bits accuracy.

Successful prediction requires the extraction of information derived from the measurement of the past and the knowledge of the dynamics.

In Chapter 2 the many limitations of EEG measurements are described in detail. Whilst the recording equipment itself is reasonably precise the volume conducting properties of the head make EEG a macro-scopic representation of millions of neurons at a time. Even if these volume conducting properties did not average out the activity of large regions of the brain, recall that in the Preface to this book we estimated that the EEG reveals 1 bit of information per second for every cortical column (10^5 neurons). Predictors that rely purely on the EEG are wanting to forecast into the future with very limited information, which given the complexity of the brain seems improbable. In any case whether EEG data are sufficiently accurate or localized to make successful predictions remains dubious.

Finally prediction is difficult because a system may not be very predictable in the first place. Let's examine Equation 7.1 again. If $z[n]$ is known exactly and to infinite precision then the future of z can be predicted forever. Seeing as $z[n]$ cannot be known exactly any error multiplies roughly by a factor of 2 every iteration of the map. With an error of 0.01% $z[n]$ can be predicted only 13 steps ahead to 1 bit accuracy, only 9 steps with an error of 0.1% and only 6 steps when the error must be contained to less than 1%. This is where the essence of *chaos theory*, colloquially referred to as the butterfly effect, stems from. Small errors can have large consequences, as shown in Figure 7.1(c). A system may have inherent limitations as to how predictable it is because of (1) measurement error, (2) stochastic fluctuations in the system or the model of the system or (3) the model used to make predictions does not approximate system behavior to sufficient accuracy.

Irrespective of whether physical, gray or black box models are used, and whether the EEG as a measurement is appropriate, there is a more fundamental question that has rarely been asked: *"are seizures predictable?"*, and if so

"how predictable are they?" They have been assumed predictable because up to 50% of people suffering from epilepsy are able to predict their own seizures, usually with warning symptoms such as headaches or mood alterations that appear well before the clinical onset [63, 142]. In some cases people are able to learn how to use these warning symptoms (also known as *auras*) to prevent their incidence [161]. However, short term auras are accepted as the point at which a seizure has already begun without yet impairing consciousness. Long term auras have been linked to the build-up of epileptic activity. Auras are therefore *detectors* rather than predictors of epilepsy[2]. A systematic study that determines if seizures are indeed predictable does not to the best of our knowledge exist.

Given that auras often occur well before clinical onset the question *"is there even a need for seizure prediction?"* is also valid. The purpose of seizure prediction algorithms is to provide sufficient warning to allow for some form of intervention. This is predominantly important for patients that cannot be surgically or pharmacologically treated – by knowing when a seizure is imminent they can retain some level of control over their own lives. It is also useful for the more timely delivery of fast acting drugs or electrical stimulation, in this way reducing the side effects associated with traditional pre-emptive treatments. But if the development of seizures is slow enough so that the electrographic onset (sometimes manifest as auras) is present some time before impairing consciousness then the same can be achieved through the *detection* rather than prediction of these symptoms. It was shown in Section 5.4 that this is already (sometimes) feasible for intra-cranial records with detection algorithms available today. Some research groups (e.g., [13]) decided that this direction is the most logical, but although preliminary results are promising whether the seizure can be aborted once it has started has not yet been adequately answered [111] (refer to Section 6.5.1).

Detection based treatment has not made prediction of seizures redundant. Preventing seizures from occurring could be more beneficial than stopping them once they have begun. For example, there are cases in which epileptic people feel the build-up of an oncoming seizure over hours, days or even months in which they crave for a seizure to occur so that they may feel better. Although the existence of this build-up is contentious, it supports the hypothesis of seizures as reset mechanisms – the solution to a problem rather than the cause[3] [72]. If prediction were possible then this reset could be performed externally, thereby abating the need for a seizure – a preventative rather than reactive measure.

[2]Another argument for the existence of a pre-seizure state is the ability of some dogs to "predict" seizures. However these dogs also predict pseudo-seizures and are once again examples of detectors rather than predictors, most likely through motor-based symptoms [35, 85].

[3]Although it is also possible that this desire for the seizure is a psychological need by the patient to associate some positivity to the situation.

On a more practical note prediction rather than detection provides a longer warning time in which action can be taken. This is important seeing as breaching the blood-brain barrier in the delivery of fast acting drugs as well as the use of electrical stimulation are both novel methods for which much research is still necessary. If longer horizons are provided then traditional drugs that take minutes rather than seconds could be used today with a successful predictor.

Predictors today are far from clinically applicable. Most prediction algorithms are based on many assumptions: they assume the existence of a dynamic and deterministic model – even though this model is not known; they assume that the measured EEG is appropriate in the representation of this model – this may not be true (see Chapter 2); they assume that relatively short amounts of data can be used to infer statistical invariants such as Lyapunov exponents and synchronicity – this has been disproved in Chapter 3. Although it is not impossible that dynamical characteristics can be tracked in this way, it seems unlikely that these methods are reliable considering the volatility observed in EEG records.

The performance of these statistics as predictors creates further doubt as to what they are really detecting in the first place. They work only in very specific cases and have for the most part been invalidated in recent publications, when tested on larger and independent datasets [11, 61, 62, 88, 101]. In some cases the methods were shown to perform no better than random (see [111] for details). Furthermore nothing is known about the nuances of the data being tested: What do the EEG records leading up to the seizure look like? What do the inter-seizure events look like? Are they stereotyped? Is it possible that these methods are effectively a form of signature recognition, and are therefore yet another example of a detector rather than predictor? Without knowing exact details of the data under test (beyond the typical information such as sampling rate, number of channels, etc) then these questions are unanswerable.

Methods based on signature recognition exploit only phenomenological information. They do not attempt to understand anything about the generation of the seizures. They cannot, themselves, be used to infer anything generally about the dynamics of epilepsy, and seizures that are not typical are likely missed. It has been argued, and it could be true, that these statistics are indeed measuring a slow change in the EEG that is representative of the dynamics. Patients for whom this does not work can be explained by a recruitment that is too fast to notice – it all makes sense under this paradigm. Just as valid, however, is the argument that what is being detected is a stereotyped pattern leading to a seizure. We must accept this possibility. Until continuous, standardized data sets that include a blind validation phase are available prediction (as opposed to detection) of epilepsy cannot be verified.

Fortunately measures have been put in place making a more unified effort possible. The International Workshop of Seizure Prediction is a regular meeting scheduled every two years promoting collaboration and providing common datasets for testing. The third such meeting was held in Freiburg, Germany,

in 2007. The standardization of datasets as well as the introduction of validation tests that compare algorithms against random predictors [160, 195, 196] provide a more rigorous platform if not a solution to the problem of seizure prediction.

Successful or otherwise; predictors or detectors; dynamical indicators or stochastic fluctuations, current prediction has not addressed the fundamental question: *how predictable are seizures?* The work in this chapter is an attempt to answer this, in part, with particular interest placed in the identification of *memory* in the epileptic brain. If no memory exists and seizures are caused by a sudden, abrupt transition then they are unpredictable and therapy should concentrate on their detection. If, on the other hand, memory exists then arguably seizures can be predicted.

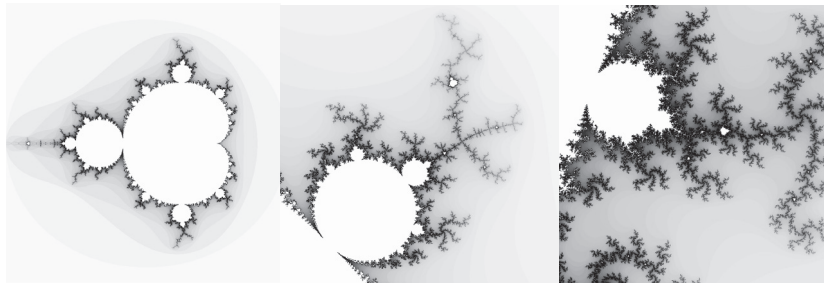
How long this memory is determines what avenues of treatment are possible as well as the type of data that can be used to make these predictions. In this chapter data sequences consist of only the times of epileptic events. These times are analyzed to determine if memory exists in the generation of epileptic seizures. These datasets span from 3 hours to 25 years, and at first glance look like random events that cannot be predicted. The aim here is to determine if some structure can be identified in this seemingly stochastic process, in which case information about future seizure times can be obtained from past seizure events.

The methods used to identify the existence of memory involve the detection of *scale-invariance* in these datasets. This is a phenomenon that indicates that behavior at different time scales is similar to each other. Scale-invariance does not itself indicate memory unless it is of a specific type. Because the words *scale-invariance*, *power-laws*, *long-range memory* and many other terms that are all relevant to this study are frequently used, confused and abused in literature, the next section is dedicated to defining and discussing what this terminology means. A description of the datasets is then given in Section 7.3, and the analysis and discussion of results are presented in Section 7.4 and Section 7.5 respectively.

7.1 Predictability – Terminology Made Clear

In the context of predictability, the notions of *self-similarity*, *scale-invariance*, *self-organization*, *criticality*, *power-laws* and *long-range dependency* are often used liberally and, to confuse things, interchangeably, without much explanation. This section is here to clarify some of the concepts.

We start with the most general concept, *self-similarity*. An arbitrary entity (e.g., a picture, an object, a time-series) is said to be self-similar if the properties of the entity are the same when it is looked at as a whole or in parts. A common example is that of *fractals*, the most famous of which is



(a) (b) (c)

FIGURE 7.2: Mandelbrot set. (a) is the complete Mandelbrot set, (b) zooms into part of (a), and (c) zooms into part of (b). Self-similarity is evident because the same types of structures are present regardless of the zoom scale. The diagrams were generated using the software `Fractal Explorer`, available for download on Matlab Central File Exchange (`http://www.mathworks.com/matlabcentral/fileexchange/`).

the Mandelbrot set shown in Figure 7.2. (a) shows the Mandelbrot set in its entirety. Enlarging a part of the set, as in (b), results in an image that shows patterns that are very similar to the whole. This continues indefinitely at finer scales. The properties at different scales do not look *exactly* the same as each other, but similar types of structures exist. Self-similarity exists in the real world – coastlines, snowflakes and even the flows and eddies in Leonardo da Vinci's drawings are considered self-similar (within certain spatial scales) [18].

Scale-invariance is an instance of self-similarity. The Mandelbrot set in Figure 7.2 is spatially scale invariant because similar properties are present at all spatial scales of the picture. More relevant to this chapter is temporal scale-invariance, which for a stochastic process implies that the statistical properties at different time scales (e.g., days versus hours versus minutes) effectively remain the same. Scale-invariance has been studied for earthquake frequency, internet traffic and economic data [130], and has also been identified in many biological systems including the timing of ion-channel opening in neurons, auditory nerve fiber action potentials and human heartbeats [7].

Self-organization is the ability of a system (of interest here – the brain) to organize itself into a state of increased complexity without the need for external interference. The system can change thanks to properties of the network and its dynamics and not because of external input. *Self-organized criticality* (SOC) is then the ability of this system to evolve from a state that is not self-similar or scale-invariant state to one that is. These concepts are explained very well in [15]. The identification of SOC in brain function is an active area of research because it is believed that understanding why it

exists will reveal important mechanisms in brain function. For example in [17] a study of spontaneous activity in in vitro brain slices revealed that SOC exists in the 'avalanches' of activity that occur, a trend that may be important for optimal information transfer and stability in cortical networks. In [153] evidence of self-organized criticality is found in the focus of some types of epilepsies, where the seizure itself is the self-similar state. These observations are important in the development of models because they must be capable of replicating this behavior. Conversely these observations can help validate models. For example the integrate-fire models described in Chapter 6 have been shown to exhibit SOC-like behavior [197].

So, regardless of the presence of SOC, systems can be scale-invariant (and therefore self-similar) in both the spatial and temporal domain. *How can the scale-invariance be identified?* From here on we will concentrate on temporal scale invariance because of its relevance to later sections. The concepts presented here are applicable also to statistics that can be extracted from spatial information.

Implied by temporal scale-invariance is that short/small events occur frequently and long/large events occur infrequently, but with non-negligible probability. If an arbitrary function $f(t)$ describes a temporal scale-invariant process, then the log-log plot of $f(t)$ versus time t is a straight line. That is,

$$\log(f(t)) = -\beta \log(t) + c, \tag{7.2}$$

or alternatively

$$f(t) = \exp(c)t^{-\beta} = Ct^{-\beta}. \tag{7.3}$$

This is known as a *power-law* with positive scaling exponent[4] β and positive $c = \log C$. β is the negative of the gradient of the straight line formed in the log-log domain. $f(t)$ displays scale invariance because dilating or contracting time t by a constant k (that is, changing the temporal scale) simply multiplies $f(t)$ by k^{β}, also a constant. The power-law becomes

$$\log(f(kt)) = -\beta \log(kt) + c = -\beta \log(t) - \beta \log(k) + c, \tag{7.4}$$

and the dilation results in an additive factor of $-\beta \log(k)$ in the log-log domain. *The gradient/exponent of the power-law remains the same upon dilation.*

The word "power-law" thus describes a particular relationship that may be present in an arbitrary function. By itself it does not mean much – events with random distributions that obey a power-law are easily generated. However under certain conditions if power-laws are observed in the statistics of a system then information about the underlying process, such as its predictability, may be inferred.

[4]Here the negative sign in front of β is there to demonstrate a decay or negative gradient in the relationship. This is convenient for the analysis in this chapter but not necessary in the general sense of a power-law.

Scale-invariance is an example of *self-similarity*, in which the statistical properties at large scales are the same as those at small scales. Scale-invariance may be temporal or spatial. *Self-organized criticality* is the ability of a system to, without external input, make its dynamics scale-invariant when a critical stage is reached.

If the log-log plot of a function $f(t)$ versus time t is a straight line with gradient β, then the process is said to obey a *power-law*, and it is a special cases of a scale-invariant process.

We now introduce the concept of *short-range* versus *long-range dependence*. The autocorrelation of $f(t)$ (defined in Equation 3.14 in Chapter 3) describes how much events at different times depend on one another. If the auto-correlation decreases very fast then current events can say very little about future events. This is known as *short-range dependence* (SRD). The most trivial example of an SRD process is an independently drawn random variable where the auto-correlation for all values other than $t = 0$ is negligible. This is the shortest type of dependence, i.e., none at all.

However, if the auto-correlation decreases more slowly then there is some information about the future in the present event. If the decay is particularly slow then the dependence spans far into the future. This is known as *long-range dependence* (LRD). Processes that exhibit LRD are thus said to have long memory. Formally, the area under the auto-correlation of an SRD process is finite, and the decay near zero is very fast, whereas the area under the auto-correlation of an LRD process is infinite.

Determining whether a process is SRD or LRD can be done by examining the power-laws that exist in the statistics of collected data, so long as it is stationary. The amount of dependence in a process can be characterized by an exponent α, with values $0 \leq \alpha < 1$. This exponent is calculated from the gradient β extracted from a power-law relationship[5]. For the trivial (random) SRD process, $\alpha = 0$. The closer α gets to 1 the longer the dependence and therefore the longer the memory in the process. Higher values of α represent smoother trends and less volatility.

The values of α have been restricted to the range $0 \leq \alpha < 1$ so that they only test for LRD. Furthermore, this restriction implies that if a process is scale-invariant then it is also LRD, and vice versa [130]. This is not true in the more general sense. Values outside this range from observed data are indicative of non-stationarity, noise or the existence of more complex dynamics not described by the presence of LRD in the data.

[5]The relationship between α and β depends on the method of extracting the power-law. More on this in Section 7.2

The exponent α is often re-expressed in terms of the *Hurst exponent* H, the "index of dependence", which can in turn be directly related to fractal dimensions such as those described in Chapter 3. So as not to confuse matters, the relationship between H and α is assumed to be $\alpha = 2H - 1$, so that $H = 0.5$ indicates a random distribution (i.e., no memory), and values of $0.5 < H < 1$ are indicative of LRD. Since α and H can in any case always be expressed in terms of each other only α is used from now on.

Is LRD a desirable property? There are two ways in which the existence of LRD may be interpreted:

1. *LRD is bad news*: The presence of long memory creates many problems in analysis, especially for short data sequences. The structure of the underlying system is complex and difficult to analyze because statistics computed from data whose probability distribution obeys a power law require (an often prohibitively) large number of datapoints to be accurate. The observation time must be long enough so that at least some of the infrequent events occur.

2. *LRD is good news*: *At least there is some structure in the system.* Consider measurements that look random, but are in fact LRD. For a completely random system the acquisition of more data does not make the system more predictable because no additional information is gained about the future. However, because memory exists in an LRD process the longer the period of observation the more predictable that it becomes [130]. In theory it is possible to reduce the prediction error to an arbitrarily low level by increasing the observation time accordingly. In practice, of course, this is often not feasible.

The information in this section can be used to determine the predictability of seizures. Data that consist of only the times at which epileptic events start were collected from different patients spanning different intervals of time. These events *look* random, but may in fact be correlated. Determining whether this system has memory, that is, *identifying* the presence of long-range dependence, could mean that these events are predictable. If no memory exists then this type of data cannot be used to predict seizures, although other forms of data may still be usable.

In summary, if a power-law exists in the statistics of a function $f(t)$ then it displays *scale-invariance*. The gradient of the power law, β, can be used to determine if there is memory in the system or not by computing the *scaling exponent* α. The relationship between α and β depends on the method of computing the power law.

If $\alpha = 0$, then there is no memory in the system. If $0 <$

$\alpha < 1$, then long-range dependency or memory exists in the system. Memory or LRD can mean that a system becomes more predictable with longer observation time. This information is important in discerning how predictable epileptic seizures are.

Robust methods to estimate α are described next.

7.2 How to Estimate LRD

The clinical data used for analysis (described in detail in Section 7.3) are in the form of discrete time points at which epileptic events start. The strength or duration of the events is often not known. This is called a *point process*, that can be equally expressed as either discrete times at which events occur or as the inter-event times. The inter-event times of these datasets look random in nature.

This section illustrates the process of robust estimation of long range dependence and/or scale-invariance through the use of well-known generators of both LRD and random (trivial SRD) processes. These methods can then be applied to the clinical data. To generate point processes samples are drawn from well-known probability distributions, with and without memory.

7.2.1 Example Distributions

For a random process, the most widely used distribution is the *Gaussian* or *Normal* distribution, with probability density function (PDF)

$$P_{GAUSS}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right). \tag{7.5}$$

The probability of x lying in the range (m, n) is given by

$$P_{GAUSS}(x \in (m,n)) = \int_m^n P_{GAUSS}(x)dx. \tag{7.6}$$

Equation 7.5 is a model of a probability distribution. If a data sequence of length N obeys this distribution then empirically, and for a sufficiently large number of events, the sample mean approaches μ and the variance approaches σ^2.

To simulate a Gaussian process we draw or generate a sequence of i.i.d. (independent and identically distributed) random samples $X[n]$ using this distribution, for $n = 1, 2, \cdots, N$ trials. Point processes such as the ones in the clinical datasets can only take positive values, so negative values are

rejected. Results are not affected by these transformations[6]. These positive values are then representative of the inter-event times. An example time series drawn from $P_{GAUSS}(x)$ with mean $\mu = 3.5$ and variance $\sigma^2 = 1$ (both in arbitrary units) is shown in Figure 7.3(a). Provided that the samples are drawn independently then there is no correlation and thus no memory in this time series. This is the example used for the trivial SRD process with expected $\alpha = 0$.

A stochastic LRD process is more difficult to simulate. A well accepted example is *Fractional Gaussian Noise* (fGN), derived from *Fractional Brownian Motion* (fBM, also known as a random walk process). fBM is a random discrete time series $B_\alpha[n]$ with $n = 1, 2, \cdots, N$. If the path is constructed as

$$B_\alpha[n] = \frac{1}{n} \sum_{i=0}^{n} X[i], \quad n = 1, 2, \cdots, N, \tag{7.7}$$

where $X[n]$ are randomly generated samples, and the distribution follows

$$B_\alpha[N] \sim N^{\frac{\alpha+1}{2}} X[1], \tag{7.8}$$

where $\sim$ denotes an equality in distribution, then $B_\alpha[N]$ is Brownian motion with scaling parameter α. The properties at different scales (i.e., different N) scale according to a scaling exponent α. If $X[n]$ are i.i.d. variables drawn from Gaussian distribution described by Equation 7.5 then fGN is defined as

$$F_\alpha[n] = B_\alpha[n] - B_\alpha[n-1], \quad n = 2, 3, 4, \cdots, N. \tag{7.9}$$

The distribution of F_α follows that of a Gaussian, with variance proportional to the delay between samples (in this case 1) to the power of $\alpha + 1$. When $\alpha = 0$ then the elements in $F_{\alpha=0}$ are independent, and the random walk $B_{\alpha=0}$ is truly random. With $0 < \alpha < 1$ the system experiences memory in that if an increment in a particular direction occurs, then it is likely that the motion in B_α continues in this direction. The larger the α, the greater memory this process exhibits.

fGN with only positive values can be interpreted as inter-event times. F_α is used here as the example of an LRD process. An example of fGN and fBM with $\alpha = 0.8$ and $N = 1000$ is shown in Figure 7.3(b) and (c) respectively. Note that fGN looks random, even though slightly different than the Gaussian process in (a). fBM, on the other hand, displays more structure.

The probability density function (PDF) of our test signals is shown next to each simulated time-series in Figure 7.3, and is compared to the expected distribution drawn in a solid line. They are derived as explained by Equation 3.42. Notice the much greater variability in the case of fBM relative to

[6]Better probability distributions that generate discrete and positive point processes exist (e.g., the Poisson distribution), but no analogous distribution exists for an LRD process. This makes comparisons between findings more tedious. In any case, all observations made about the random case in this section are the same regardless of whether samples are drawn from a Gaussian or a Poisson distribution.

the Gaussian and fGN PDFs. These last two converge very quickly to their expected distributions. In contrast the PDF of fBM is very volatile. This variability is a consequence of the LRD present in this data: many more data points are needed for convergence to the expected distribution. LRD processes are notoriously difficult to simulate for this reason.

The existence of scale-invariance is not evident at face value in the simulated fBM and further processing is necessary.

7.2.2 Computing α

The remainder of this section shows how α can be estimated from sequences simulated according to the above two distributions (Gaussian and fGN). The random Gaussian distribution or trivial SRD case is referred to from now on to have probability density function $P_{tSRD}(x)$. The fGN is referred to have PDF $P_{LRD}(x)$ with $\alpha = 0.8$. We perform analysis on data sequences of three lengths: $N = 500$, $N = 1000$ and $N = 10000$ events.

The first step in determining LRD is to check whether scale-invariance is likely. The easiest way to do this is to derive *inter-event probability histogram* (IPH). They are calculated as the probability that an event of a particular length, in this case the inter-event times, has of occurring[7] [7]. For data from an unknown distribution this requires knowledge about the maximum resolution. In the case of the processes drawn from $P_{tSRD}(x)$ and $P_{LRD}(x)$ the maximum resolution is known and the IPH can be drawn.

The IPH is drawn for the simulated point processes of different lengths in Figure 7.4(a) and (b), where the latter is now the PDF of fBM. The histograms are presented in a log-log plot to show that the probability of large events for $P_{tSRD}(x)$ is very low. It is shown next to the expected distribution and a very fast (*exponential*) decay occurs at large events. In contrast (b) shows that large events occur at non-trivial probability for $P_{LRD}(x)$, independent of the number of samples N used. The histograms are drawn next to a reference exponential decay that has best fit at small event durations. Long event durations do not decay exponentially as in (a). This slow decay is sometimes referred to as a *heavy tail* (although more formal definitions exist) for which long/large events occur at non-negligible probabilities [130].

From the purely mathematical point of view there is no reason why a heavy tailed distribution has to be LRD, but in processes derived from dynamical systems heavy tails generally lend themselves to the existence of long memory [130]. *Although a heavy tail does not necessarily indicate the existence of LRD, LRD cannot exist without one.* This first step is a simple way to determine whether it is worth continuing with analysis. For example, the process generated by $P_{tSRD}(x)$ can be rejected, whereas the process generated by $P_{LRD}(x)$ requires further analysis and validation.

Traditional methods of estimating the fractal exponent α include the iden-

[7] Again, probability histograms are calculated as in Equation 3.42.

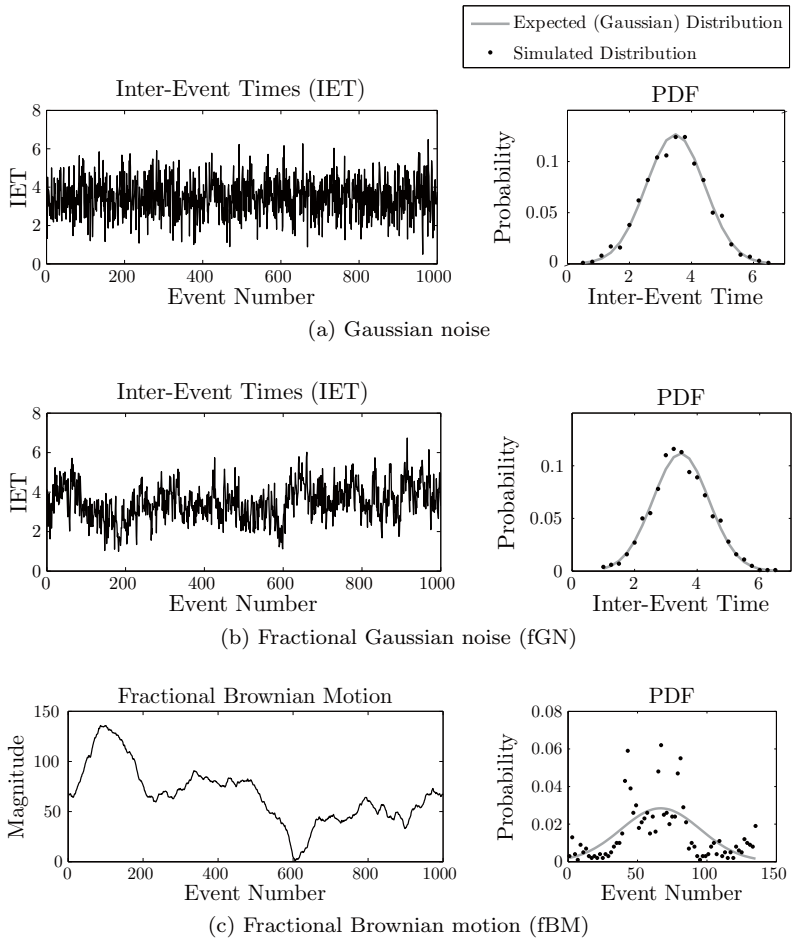
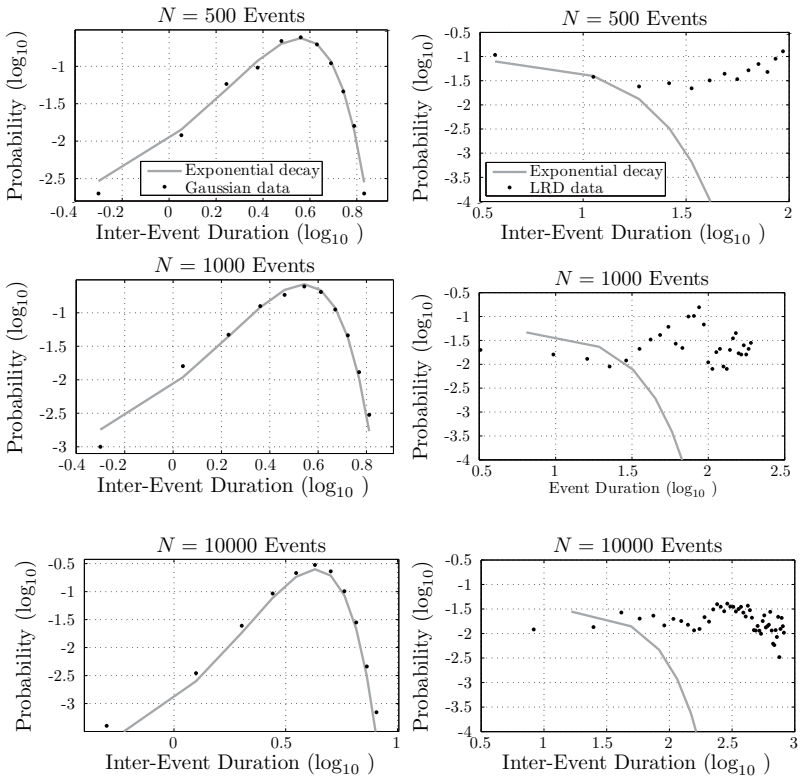


FIGURE 7.3: Simulated time series for (a) Random (Gaussian) process, (b) Fractional Gaussian Noise (fGN) with $\alpha = 0.8$ and (c) the consequent Fractional Brownian Motion (fBM). Upon visual inspection both the (a) and (b) look predominantly random even though some structure exists in fGN. Next to each figure is the probability density function (PDF) of the simulated time-series. (a) and (b) have similar PDFs that converge to their expected distributions quickly. The PDF of fBM, on the other hand, is more volatile and does not converge to the expected function. This is because the existence of long memory, brought about by the introduction of the parameter α, makes the convergence time of the simulated time-series much longer. This is the problem when long memory exists: many data points are necessary to provide conclusive results.



(a) Random (Gaussian) Process (b) LRD Process

FIGURE 7.4: Inter-event probability histograms (IPH) drawn on log-log plots for both the random ($\alpha = 0$) and the LRD ($\alpha = 0.8$) simulated time series. (a) shows that the IPH of random events falls exponentially so that the probability of large events is negligible. (b) shows that the IPH for an LRD process has large events that occur at non-trivial probability. This is known as a *heavy tail*, a property that is necessary for memory to exist.

tification of power-laws constructed in a variety of ways. These methods are described in great detail in references such as [18], [36] and [130], and typically involve higher order statistics (e.g., variance and higher) because first order statistics (e.g., mean) do not reveal the differences between random and LRD processes [130]. Variance is the most often used second order statistic. Other typical methods include:

- *Auto-correlation*: Defined in Section 3.3.1 in Chapter 3. The auto-correlation not only gives an idea of how far in time there is non-trivial coupling, but can also be used directly to estimate α. If a power-law is

evident in the log-log plot of absolute auto-correlation magnitude versus delay then its gradient $\beta = 1 - \alpha$ [7].

- *Power Spectral Density (PSD)*: Defined in Section 3.3.2 in Chapter 3. Since the auto-correlation can be related to the PSD, it follows that α can also be estimated by analyzing the frequency content of the signal. Long memory relates to low frequencies (and slow time scales). The power-law generated at *low frequencies* in the PSD scales as $\beta = \alpha$.

- *Fano Factor*: This is an estimation of the variance in the number of events observed in a time period T. Each T is representative of a different time scale. The gradient of the Fano factor scales as $\beta = \alpha + 1$ for increasing T. In the random case, the Fano factor does not depend on the duration of the observation, resulting in an $\beta = 1$ ($\alpha = 0$) as expected. Any deviation from $\beta = 1$ indicates a complexity or richness of information not present in a process with no memory [7].

The above examples are all simple ways of estimating the scaling exponent α, related to β in different ways depending on the way that the power-law is drawn[8]. However although these methods work well under optimal conditions they are each wrought with their own idiosyncrasies for data series that are short, noisy and not stationary – results are difficult to interpret.

Wavelet decomposition is used as an alternative. Wavelets work well because they are themselves scale-invariant processes, a property that is not true for other analysis tools [130]. They have proved effective in detecting the existence of many different types of long memory. Under ideal conditions they are not better at estimating α than previously mentioned methods [7], but they have the advantage of coping well with non-stationarity, provided that the changes are smooth enough [7, 36, 130].

Wavelets have been described in detail in Section 3.3.3. They isolate activity at different frequencies and time scales, and, as such, they are intuitively related to the estimation of α through PSD based methods, where larger scales correspond to lower frequencies[9]. If an orthogonal mother wavelet is used and only dyadic sampling is allowed (that is, let dilation factor $a_m = 2^m$ for $m = 1, 2, 3...$) then the wavelet coefficients between scales are only weakly correlated. Thus the analysis at each scale is largely de-coupled from all other scales [130].

The variance s_m of the wavelet coefficients $d[m, l]$ at scale m is defined as

[8]The estimated value of α may not be the same as the true scaling exponent. It is the aim of this work to find out to what confidence the value of α may be representative of the true value.

[9]In fact, if the DB-2 (also known as the Haar wavelet) is used then the analysis can be *equivalent* to the Fano factor calculation because they both measure variance at different time scales in similar ways. The difference is that the Fano factor does not handle the correlation present between scales very well, whereas the wavelet methods do, provided that sampling is dyadic or exponential ($a_m = i^m$ for any number i) [130].

$$s_m = \frac{1}{L_m} \sum_{l=1}^{L_m} |d[m,l]|^2 , \tag{7.10}$$

where there are L_m coefficients at each scale m. This is a second order statistic that can be used to estimate α. Because each value of m represents an exponential increase in bandwidth as well as center frequency, plotting $\log_2(s_m)$ versus m is the equivalent of a log-log plot (also known as a *scalogram*) used to identify the power-law. In reality a small bias g_m needs to be introduced because of the non-linear nature of the logarithm, so that the plot versus m becomes

$$y_m = log_2(s_m) - g_m. \tag{7.11}$$

A good approximation of g_m when L_m is large enough and under the simplifying assumption that coefficients at different m are completely decorrelated is given by [36]

$$g_m \approx -\frac{1}{L_m \ln 2}, \tag{7.12}$$

where $\approx$ denotes an approximate estimate.

To detect scaling using y_m versus m a *region of alignment* must be identified in this plot. Not all scales need to be involved in the long memory process, but if a sufficient number of scales line up then a power-law, which may then be tested for LRD, is evident. In practice the region of alignment must consist of at least 4 dyadic scales. This number is somewhat arbitrary, seeing as memory could exist in as little as 2 scales, but it is entirely too easy for up to 3 scales to line up randomly. A minimum of 4 scales is necessary before results are taken seriously [130]. To test for LRD the gradient at the region of alignment can be equated to the scaling exponent, that is, $\beta = \alpha$.

In reality the points are unlikely to line up perfectly. However the quantity y_m is only accurate to a certain degree of confidence dependent on the number of available coefficients. The variance s_m is a statistic and more data make its estimate more accurate. Since L_m halves at each iteration m the errors expected at each scale increase monotonically with larger scales. An estimate of the expected variance σ_m^2 in y_m can also be used to compute error bars, an approximation of which is given by

$$\sigma_m^2 \approx \frac{2}{L_m \ln^2 2}. \tag{7.13}$$

Error bars are be drawn on y_m as multiples of σ_m for different confidence levels. More data increase the number of coefficients L_m at each scale m. Thus there is a decrease in the variance σ_m and the error bars get smaller. A larger L_m also makes more scales available for analysis and results are generally more conclusive. A small dataset can lead to problems that imply that (1) the observation time may not be sufficiently long for LRD to be detectable,

or (2) error bars become large enough so that LRD cannot be distinguished from random. In (2) the method is incapable of rejecting either possibility. Longer time series do not remove but do reduce the chance of these limitations becoming significant.

7.2.3 Simulations

To apply the above theory on our distributions $P_{tSRD}(x)$ and $P_{LRD}(x)$, the following steps are taken:

1. *Generate inter-event sequences*: This is done using Equations 7.5 and 7.9 respectively for a random (trivial SRD) and LRD process, ensuring that all values are positive. Three sequences of length $N = 500$, $N = 1000$ and $N = 10000$ are generated for each case.

2. *Create time-series*: A continuous-time process is approximated by generating a time-series in which a zero indicates no event, and a 1 indicates an event [10]. In this case, values were arbitrarily assigned a resolution of the minimum inter-event time observed in generated sequences. Using a finer resolution does not affect the estimates (see later analysis).

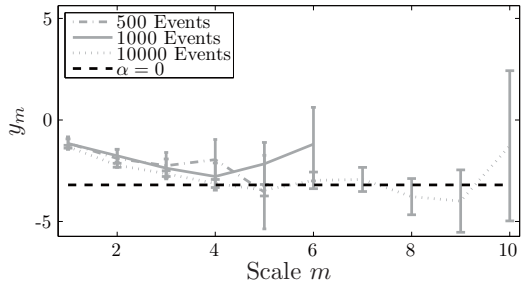
3. *Use wavelet tools*: Calculate y_m and σ_m for as many scales as are available[11]. Although the number of scales depends on the resolution of the time series created in step 2, the number of dyadic scales available for $N = 500$ is roughly 1 less than for $N = 1000$ and 4-5 less than for $N = 10000$.

4. *Calculate* α: Select the interval over which, within confidence limits, points align in the scalogram of y_m versus m. Compute α using a line of best fit. A χ^2 goodness of fit test is computed to determine how well the data fit this line. A confidence level Q is provided to evaluate goodness-of-fit. A value of $Q > 0.05$ is deemed an acceptable fit. For the readers not familiar with χ^2 goodness-of-fit tests, this is explained in [86].

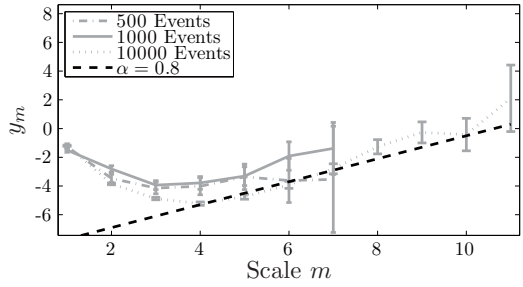
5. *Distinguish LRD*: Determine if the value of α falls within values indicative of LRD. If not, reject as noise.

[10] This step is performed because inter-event times are inherently discrete-time processes, whereas the theory presented is relevant to continuous time only. Note that a way to filter the discrete-time process so that methods can be applied directly to this data exists. Details on this can be found in [36] and [130], but results in this chapter were not significantly affected by the choice of method.

[11] The basis for the code used in this analysis is freely available for download on `http://www.cubinlab.ee.unimelb.edu.au/~darryl/secondorder_code.html` for non-commercial purposes. The authors, Patrice Abry and Darryl Veitch, retain copyright of this code.



(a) Random Process



(b) LRD Process

FIGURE 7.5: Wavelet estimation tools are shown to reliably estimate α at large scales for both the random and the LRD processes, even for simulations involving relatively low number of events. The figures show the power-law relationships formed between y_m and m. The expected value of $\alpha = 0$ for the random case and $\alpha = 0.8$ in the LRD case is included in each plot for comparison.

7.2.4 Results

Figure 7.5(a) and (b) show estimates for the trivial SRD and LRD process respectively. The expected $\alpha = 0$ for SRD and $\alpha = 0.8$ for LRD are also plotted for comparison. In both cases even for relatively low number of events the limiting value of the gradient of y_m versus m approaches the expected α within error bars that increase with m. All analysis here is performed using a DB-3 wavelet (see Chapter 3).

How robust are the estimations of LRD? Several factors can affect the calculations: choice of resolution in step 2, noise in the data, and choice of wavelet type and order in step 3. We explore how each of these affect the estimates.

First, in Figure 7.6 step 3 is applied to the LRD process for different resolutions, ranging from low (smaller than the smallest inter-event time) to

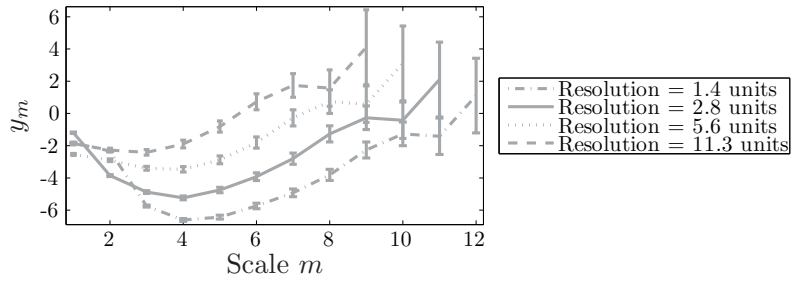
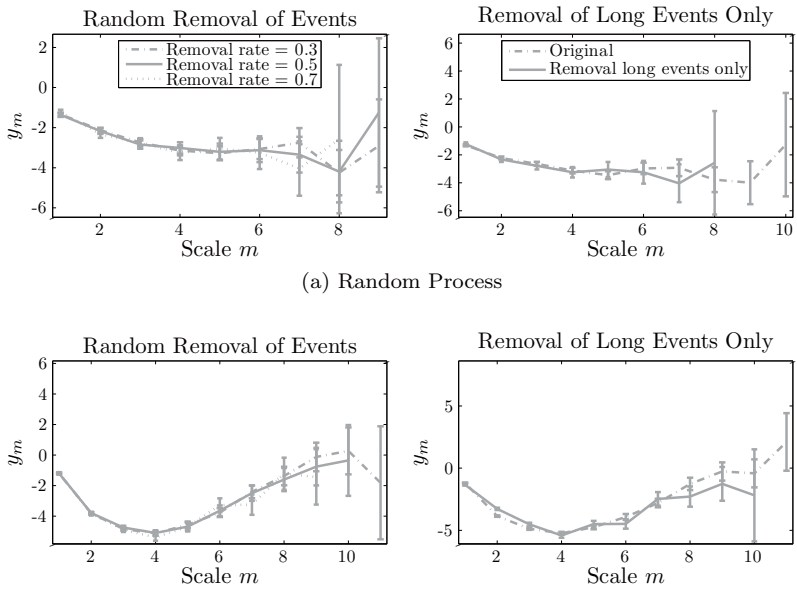


FIGURE 7.6: LRD calculation for changes in resolution. Changing the resolution at which α is estimated does not affect the gradient at large scales, even though the raw values of y_m do change. Thus although it is not valid to use a resolution higher than that used in the recording process, so long as the selected resolution is reasonable analysis can continue with little thought to its impact on estimates. This is in stark contrast to traditional estimators of α that are very volatile under such design choices.

high (larger than the smallest inter-event time) in arbitrary units. The coarser the resolution (higher value) the fewer scales that are available for analysis. The values of y_m themselves *are* affected by the choice of resolution, but the gradient α at large scales is not. This again is a result of the independence of wavelet coefficients at different scales. Using a resolution greater than that of the recorded data, although in general meaningless from an analysis point of view, does not affect the results significantly. Furthermore using coarser resolutions also does not affect results, and since coarser resolutions may be interpreted as greater tolerance to noise the wavelet based estimates are fairly robust to noisy conditions. *As long as the resolution used is within the ballpark figure of the true resolution, wavelet estimation will not be adversely affected by this choice.*

How are estimates affected by missing data? Or rather *how is the analysis affected by removal of events?* To answer this, step 3 was computed for both random and LRD processes with events randomly removed at different rates [67]. The results are in Figure 7.7, and they show that even for large removal rate of 0.7 (that is, 70% of the data missing) the trends at the larger scales remain relatively unaffected. Given a fixed recording epoch the uncertainty of results grows as more data go missing, i.e., there are fewer scales available for analysis. Of course if you increase the observation window this effect would disappear.

However if we are selective in the types of events that are removed by targeting only the long events, the values of y_m at larger scales m, are affected. This is also shown in Figure 7.7 for both LRD and random processes, and it



(a) Random Process

(b) LRD Process

FIGURE 7.7: The estimate of α in an LRD process is robust when events are missing from datasets. However, the targeted removal of large events destroys this structure because LRD is dependent on their inclusion. This is demonstrated for both the random and the LRD process above. Neither the random nor the selective removal of long events changes the estimate of $\alpha = 0$ for the random process, as seen in (a). On the other hand the structure at large scales is shown to suffer for the LRD process in (b) when long events are not present.

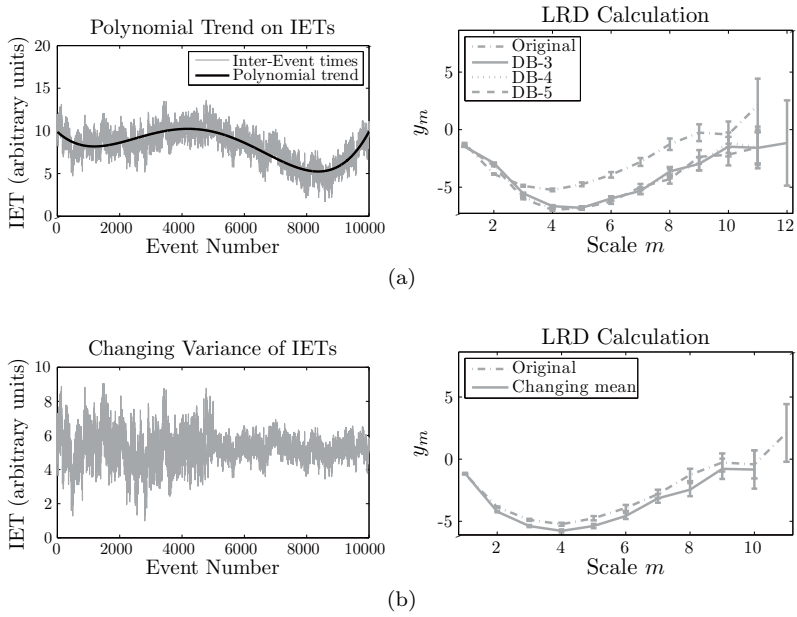


FIGURE 7.8: Wavelet based tools for the computation of α are robust under smooth changes. In (a) the mean of an LRD process with $\alpha = 0.8$ is modulated with a smooth polynomial. Though the actual values of y_m change when compared to the original, the estimated gradient α at large scales does not, even for relatively low wavelet orders. Wavelets are also shown to be robust under changes in the variance of data, as shown in (b) where a sharp transition is observed half way through the sequence. The ability of wavelets to cope with such changes make them more suitable for data such as the one described in Section 7.3 because it is expected to contain such transitions.

occurs because the properties of scale-invariance for larger m are dependent on the longer events even though they occur at lower probabilities.

Non-stationarities in the data may also affect the estimation of α. If there are reasons to believe that the data are stationary, as is the case for our simulated sequences, then the estimations of α should be taken seriously [130]. However real data can be affected by non-stationarities. Traditional methods of estimation of LRD could not be used to analyze these cases because they cannot cope with changes in stationarity. Wavelet tools *can*, provided these changes are smooth enough. This is shown in Figure 7.8. In (a) the mean of the events is changed (smoothly) by modulating the original time series by a polynomial trend. It is seen that although the values of y_m are affected (as compared to the original data), the α estimated from the gradient at large m is not. Changes in variance of events are abruptly introduced in the second half of the sequence in (b), and again the estimate of α is unchanged.

These observations are supported by theory. The order O of the Daubechies wavelet is capable of removing polynomial trends of order up to $O-1$ [7, 130]. For example the DB-3 ($O = 3$) wavelet, which has been used in all analysis so far, is capable of removing quadratic trends. Increasing the order increases the capabilities of coping with less stationarity. The choice of wavelet family in this type of analysis is not as important as the choice of wavelet order, so long as the wavelets used are orthogonal. Since wavelets are naturally equipped with the ability to change order quickly and easily once again they are superior to traditional estimation methods.

In practice the selection of the wavelet order should depend on the data because the smoothness of the non-stationarities is unknown. The order of the wavelet is systematically increased until stable results are observed, and the lowest order wavelet is used. Low order wavelets are desired because less data are necessary. The tradeoff is between the number of scales available for analysis versus higher order wavelets that can cope with more complicated forms of non-stationarity. Figure 7.8(a) shows that the introduced trend is smooth enough so that selection of higher order wavelets does not change results.

Finally, because for real data it is not known whether the non-stationarity is smooth enough changes over time should be checked for less obvious volatility. An accepted method is to break up the data into segments to determine if the values of α are consistent over time. This should be done for any estimation method used, not just wavelet based tools. This step is not shown here for the simulated sequences as it is known that these are stationary.

Wavelet tools described in Section 3.3.3 in Chapter 3 are effective in estimating α by using the variance of the wavelet coefficients at each scale m. Wavelet methods are robust with respect to:

1. *Resolution/Noise*: The choice of resolution, and thus the allowed tolerance for noise, does not affect the estimate of α, only the number of scales available for analysis.

2. *Missing Data*: Even with a relatively large number of missing events (i.e., events that have gone unrecorded) estimates of α using wavelet methods are not affected.

3. *Non-Stationarity*: A wavelet of order O is not affected by non-stationarity of polynomial trends of order less than or equal to $O - 1$.

The ability for wavelets to cope with non-stationary data as well as the flexibility in the parameter choice O that allows different levels of non-stationarity makes wavelets much more powerful than traditional methods.

Now that practical aspects of the detection of the presence of long memory have been discussed they can be applied to real data. This is done in Section 7.4 and discussed in Section 7.5, but first a detailed description of the clinical data is provided in the next section.

7.3 Seizure Frequency Dataset

A summary of the epilepsy data available for analysis is in Table 7.1. There are 6 datasets of the times at which epileptic events occur. Each represents a different subject (person or rat). Inter-event times can be extracted and time-series generated with a maximum resolution shown in Table 7.1. This value refers to the best resolution that the events were reliably recorded at by the patient or at which the events were extracted from an EEG time-series. The resolution at which analysis was performed may differ from this value. Specific numbers are given in Section 7.4 but should never be better than that shown here. The various time series are plotted in Figure 7.9.

There are two types of data. Datasets 1-4 are records of epileptic events maintained by the patient him-herself over a period of 1-25 years. These records are prone to much noise – faulty memory may lead to erroneous event times; lapses in discipline may mean some events are missing; only clinical epileptic events are present since sub-clinical seizures are not detectable by the patient; drug dosages usually change over many years, perhaps affecting the stationarity of event frequency. Nevertheless this is the best data of this nature that we are ever likely to have, seeing that EEG monitoring over long time scales is not feasible. Analysis of such data may be helped by using coarser resolution than that shown in Table 7.1, as well as by removing time periods in which lapses in the record-keeping occur (these usually have been

Frequency of Epileptic Events – Dataset Summary					
Dataset	Data Source	Duration	Sz. #	Resolution	Comments
1	Human (diary entries)	25 years (1982-2006)	911	1 day	Many drug dosage changes
2	Human (diary entries)	10 years (1997-2006)	1050	1 day	Seizures occur in clusters (665 clusters)
3	Human (diary entries)	5 years (2001-2006)	412	0.5 days	Drug changes unavailable
4	Human (diary entries)	1 year (2005-2006)	397	15 mins	Drug changes unavailable
5	Human (EEG)	4 days	2722	1 secs	Subject experienced an unusually large number of epileptiform discharges
6	Rat (EEG)	3 hours	1429	0.5 secs	Spontaneous epilepsy

TABLE 7.1: Seizure frequency data.

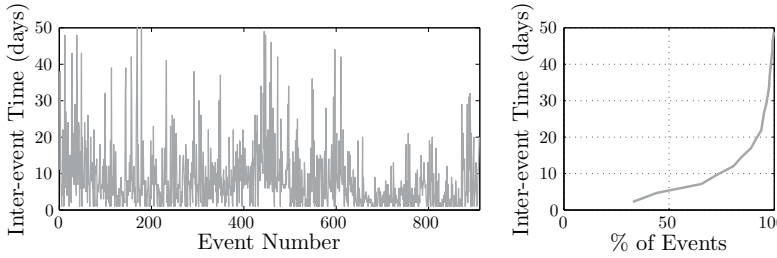
annotated by the patient). Furthermore, one can make the perhaps unjustified assumption that drug dosages change the mean frequency or variance of events but not the intrinsic relationship between events. The use of wavelet analysis that can remove the effect of such changes if they are smooth enough is then justified.

Datasets 5-6 are shorter-term sequences extracted from EEG recordings. Dataset 5 belongs to a patient who experienced an unusually large number of epileptiform discharges in the EEG. These events are not necessarily clinical. That there are many events makes this 4-day EEG monitoring suitable for analysis of this type. Finally, Dataset 6 belongs to a 3 hour EEG record taken from a rat recorded in the Shanghai Institute of Brain Functional Genomics (East China Normal University, Shanghai, China). During a surgical procedure in preparation for a different experiment the rat developed frequent spontaneous epilepsy. The records for the rat were taken in 1 hour periods. No drug changes occurred throughout the entire recording time. In Datasets 5-6 it was also possible to extract the duration of the events as well as their times, so that the 'strength' of the events is known.

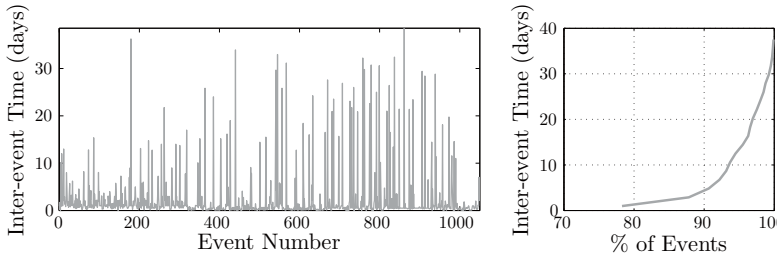
The short duration of both these studies, as well as the extraction of events directly from EEG records, makes their integrity greater than Datasets 1-4 because events are known to at most be very rarely missed. The stationarity in the epileptic discharges is also more likely. Non-stationarity may be experienced in Dataset 5 because of the day-night changes, although this is shown not to affect results in the next section.

One could argue that a typical epileptic patient is unlikely to experience this number of epileptiform discharges, or that the analysis on rat EEG data may not translate to human epilepsy. However the purpose of this chapter is to detect the presence of memory in the epilepsy, thus although the analysis may be relevant only to specific cases it could have greater implications in our understanding of epilepsy as a whole.

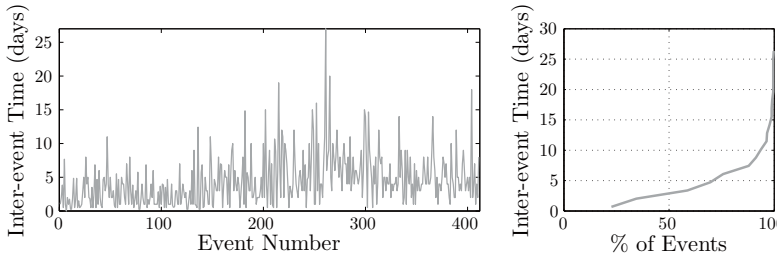
In any case, all datasets are relatively short in the number of observed



(a) Dataset 1

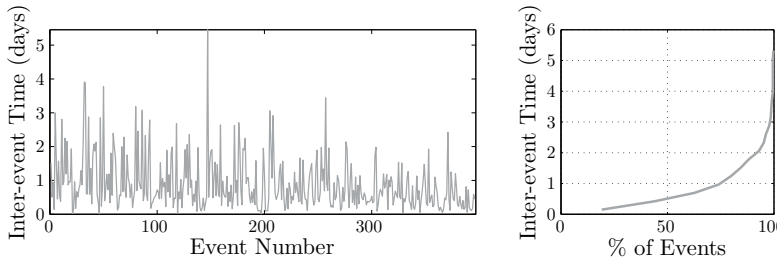


(b) Dataset 2

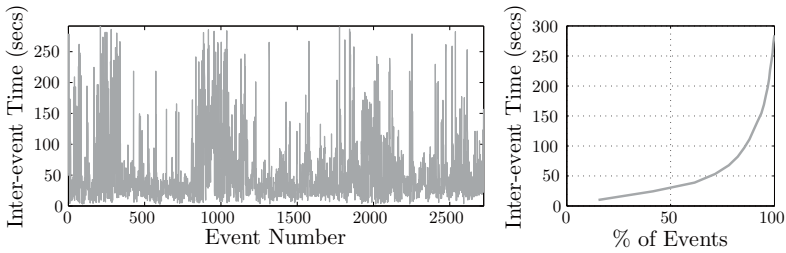


(c) Dataset 3

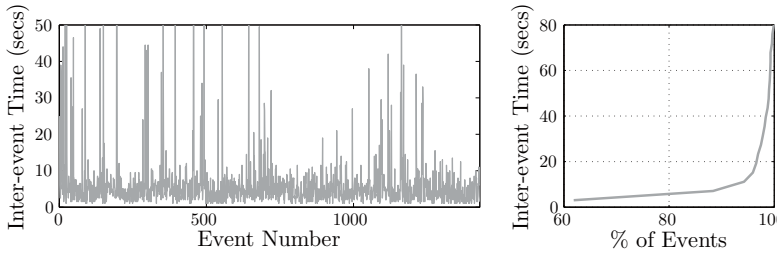
FIGURE 7.9: The inter-event times for Datasets 1-6. The type of data for each plot is summarized in Table 7.1. Each of these events look random and non-stationary. The aim of the analysis here is to find some structure in the form of memory in this seemingly stochastic data. (Continued)



(d) Dataset 4



(e) Dataset 5



(f) Dataset 6

FIGURE 7.9: (Continued)

events. The computation of memory is limited by fairly large error bars, and care in the interpretation of these results is necessary. The analysis of this data, and their relevance to predictability, is discussed next.

The available data consists of records of times at which epileptic events occur. This data can

- Have a substantial number of missing events.
- Be non-stationary due to changes in drugs, evolution of the epilepsy or changes in sleep/awake cycles.

Missed events are not expected to be selective of length (i.e., both short and long events are excluded) and short-term non-stationarities due to day and night changes are expected to be smooth changes in variance and mean. *Wavelet analysis is suitable to detect the presence of LRD under these conditions.*

7.4 Analysis – Estimation of α

Visual inspection of the data in Figure 7.9 does not reveal any obvious pattern in the epileptic inter-event times. The data looks random, even if the distributions are not clear and not necessarily the same for all datasets. The data also looks non-stationary – obvious changes in mean can be seen in Datasets 1, 3 and 5, whilst obvious changes in variance can be identified in datasets 2, 4 and 5. These changes may be due to mechanisms of the generators of epileptic activity in each case, external factors such as drug dosage changes, or in the case of dataset 5 may simply be the difference between day and night time. In any case the presence of these changes indicates that we should use wavelet analysis tools.

Let us first see if long memory *may* exist in the data by drawing inter-event probability histograms (IPH) for each dataset. These can be seen in Figure 7.10, in log-log plots, along with comparative exponential (fast) decay. In all cases the probability of large events decays slower than exponential, and in all cases it is also possible to identify a straight line that can be fitted to the IPH. Thus a power-law exists, and this heavy tail indicates that long memory may be present. Further analysis is validated for all datasets[12].

[12] In principle it is possible to use this power-law directly to estimate the scaling exponent α, but like other traditional estimation methods the results can vary greatly with choice of histogram intervals, and much more care needs to be taken in the interpretation of results.

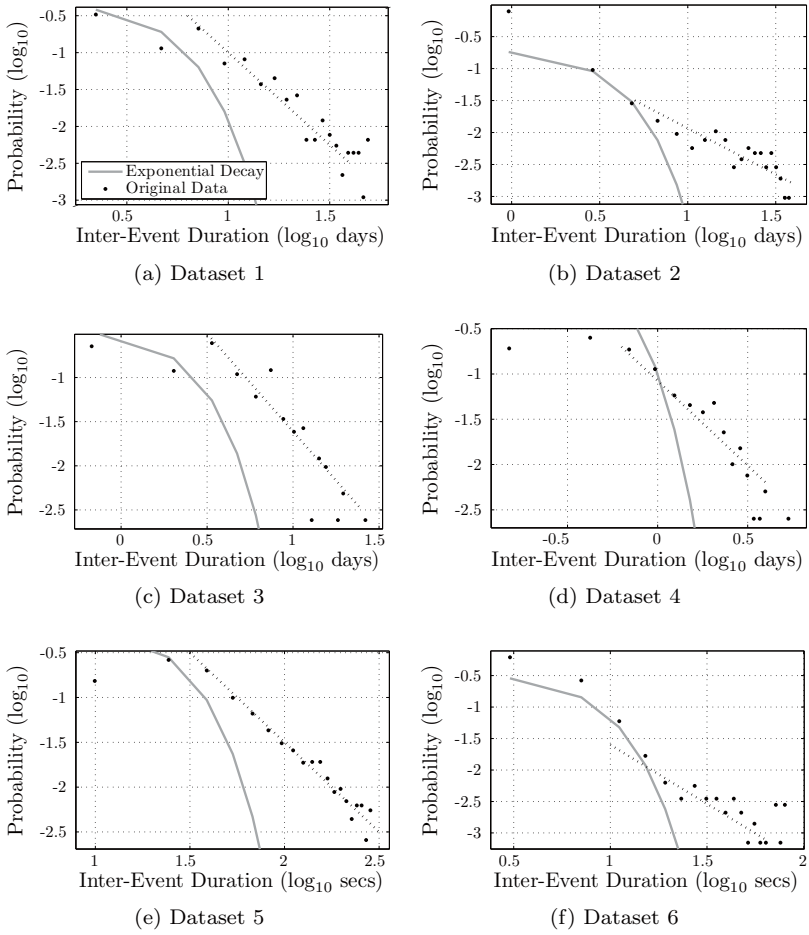


FIGURE 7.10: Inter-event probability histograms (IPH) for each of the 6 datasets in a log-log plot. The probability of long inter-event times in all cases is non-negligible. Moreover at least visually it seems that all these distributions obey a power-law because the data follow a linear trend (shown as a dotted line for reference). Because this power-law exists, further analysis in search for the existence of memory is justified. Exponential decay is plotted in each graph to emphasize this linear trend.

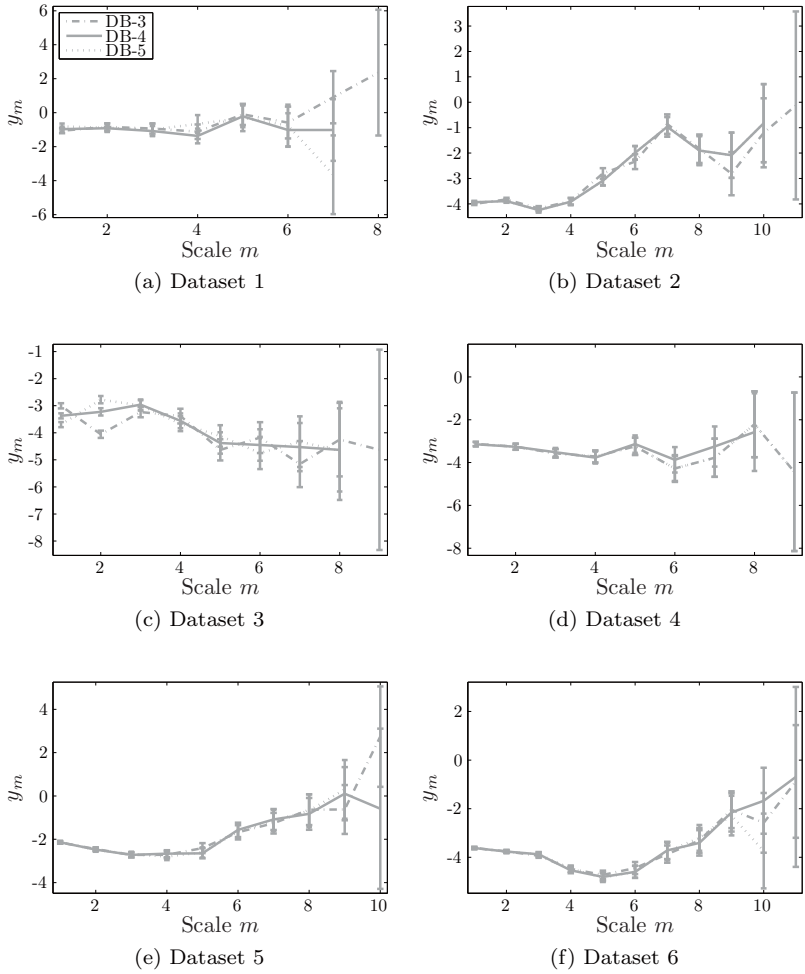


FIGURE 7.11: Calculations of y_m using Daubechies wavelets DB-3, DB-4 and DB-5 on Datasets 1-6. The above graphs validate the use of the DB-3 wavelet for further analysis because the results are stable under changes in the wavelet order. Using wavelet orders higher than necessary would reduce number of scales available for analysis, which is not desirable when the datasets are already so short.

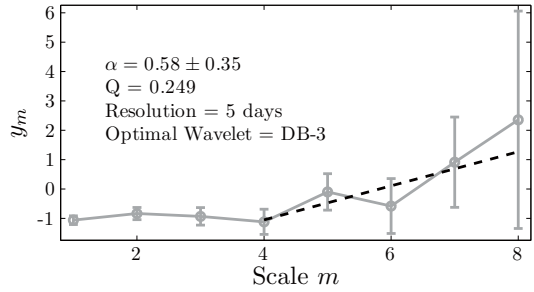
Before estimating α, the analysis is quickly performed for different order O of the Daubechies wavelets. This is shown in Figure 7.11 for $O = 3, 4$ and 5. In all cases the results are stable under the choice of O; thus the smallest order ($O = 3$) is chosen and the DB-3 wavelet is used throughout the remainder of the data analysis.

Figure 7.12 shows the estimated α for all datasets. Each graph shows y_m versus m and the line of regression used to estimate α, as well as the range of m for which it was calculated. Also shown are the resolutions used, the estimated α including error bounds and the goodness of fit parameter Q (greater than 0.05 in all cases, indicating a reasonable fit).

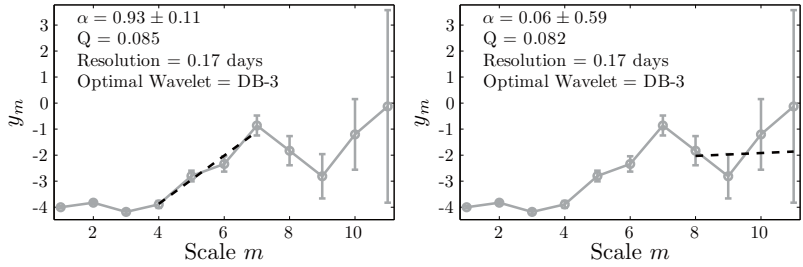
Recall that a value of $0 < \alpha < 1$ at large scales indicates LRD, whilst a value of $\alpha = 0$ indicates no memory. The problem is sometimes in selecting the correct region of alignment for which α should be estimated. Consider Datasets 4. Two regions of alignment are shown, one in which $\alpha \approx 0$ and one in which $\alpha = 0.65$. Should the larger range be chosen, because the result is more likely to be correct? What if the LRD is only present at larger scales, and it is only the un-availability of data that means that memory goes undetected? In both cases the previously mentioned minimum of 4 scales are used, thus both are valid observations. However when one starts examining the error bars it becomes clear that the $\alpha = 0.65$, although suggestive of long memory, has error bars (± 0.62) that span the entire range of $0 < \alpha < 1$. This not only means that if there is long memory its effect cannot be estimated, but also that randomness cannot be rejected because a result very close to $\alpha = 0$ is possible. *The error bars are important in determining the confidence to which results can be interpreted.*

Similar observations can be made about Dataset 2, in which several regions of alignment are possible. Initially it is tempting to say that there is evidence for scale-invariance between scales $4 \leq m \leq 7$. However this dataset is known to have seizures that occur in clusters and the time scales at which these clusters occur correspond to these scales. It could be that within the clusters memory exists, and this is indicated by this region of alignment. However this cannot be used to infer long-range memory beyond these clusters because an $\alpha \approx 0$ is estimated for larger scales. *It is important to use knowledge of the nature of the data or the data collection techniques when interpreting results.* Although it seems that a region of alignment with non-zero gradient may develop at larger scales, this is only observed for 3 values of m, an insufficient number to be taken seriously. More data would be necessary to infer memory at longer scales.

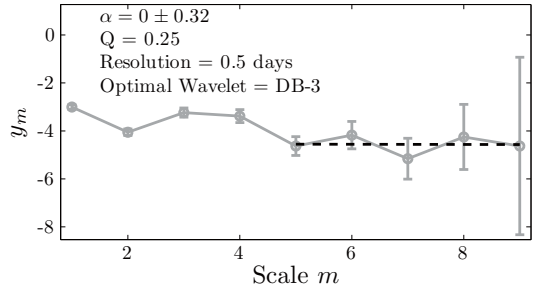
The remainder of the datasets seem a little more straightforward. An $\alpha = 0$ is a logical conclusion for Datasets 2, 3 and 4, although in all cases neither the existence or lack of memory can be rejected seeing as the error bars resulting from insufficiently long data series span both the $\alpha = 0$ and significant proportions of the $0 < \alpha < 1$ range. It does tell us that further analysis would be fruitless, and these data are rejected. In any case these graphs are important in pointing out the follies of using histograms such as



(a) Dataset 1

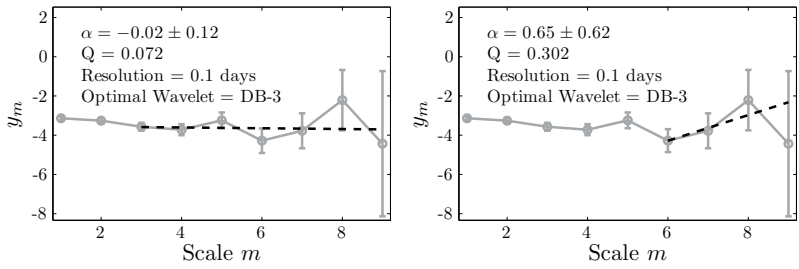


(b) Dataset 2

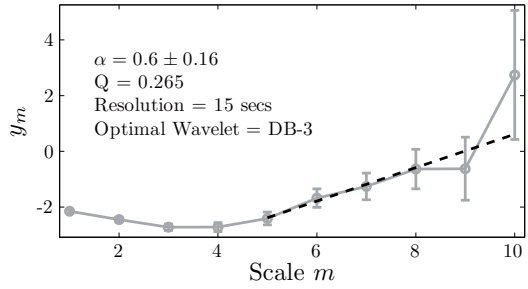


(c) Dataset 3

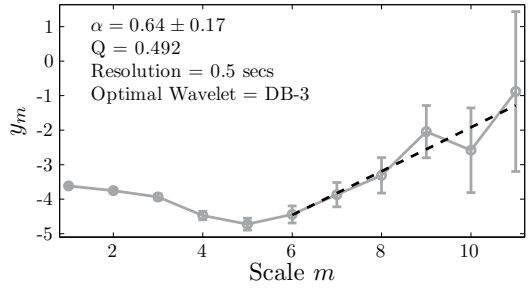
FIGURE 7.12: Calculations of α for Datasets 1-6. A clear trend is identified for Datasets 1, 3, 5 and 6, but the selection of an appropriate range to calculate α from is less clear in Datasets 2 and 4, where different ranges yield different gradients. In Datasets 2-4 the more likely conclusion is that $\alpha = 0$, that is, no memory exists. However the small number of data points used in this analysis means that error bars are too large to conclusively reject the existence of LRD. Further analysis using this data is futile. Datasets 1, 5 and 6 on the other hand indicate that memory exists. Further validation tests are necessary before conclusions can be drawn. (Continued)



(d) Dataset 4



(e) Dataset 5



(f) Dataset 6

FIGURE 7.12: (Continued)

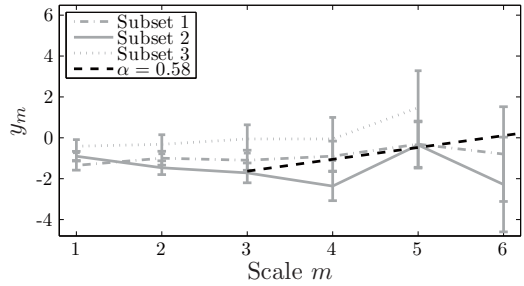
those in Figure 7.10 to infer LRD, as is done in much of the literature. In this figure a power law is clearly identified in Datasets 2, 3 and 4 and it is tempting to assume that if the computed exponent is correct, that this would be LRD. It was necessary to use the better defined wavelet tools to determine that error bars are too large to draw any conclusions.

Datasets 1, 5 and 6 are a different story because the estimated α in all of them indicates LRD with errors that exclude the possibility of $\alpha = 0$. It would be tempting to suggest that long memory exists in all of these, but first there is need to evaluate how robust these results are. A stationarity test is performed for each of these 3 datasets in Figure 7.13 by breaking down the time series into 3 consecutive segments and comparing the results of each segment to the originally computed α. This is particularly important for Dataset 1 which is known to be of a 25 year duration. Again by using knowledge of where the data come from we can see that a stationarity test is important because the data are likely to be affected by countless factors occurring over such a very long time. Figure 7.13(a) reveals that the skepticism is justified because there are large fluctuations occurring between each of the 3 tests. Even though the error bars are quite large and in most cases they overlap each other, this overlap is very small. Thus Dataset 1 is not very robust in informing us whether LRD is present. Pursuing analysis on this data is not likely to validate or in-validate the existence of LRD, even though the non-zero gradients in the scalograms shown in Figure 7.13(a) do indicate that LRD may exist over long periods in spite of the non-stationarity.

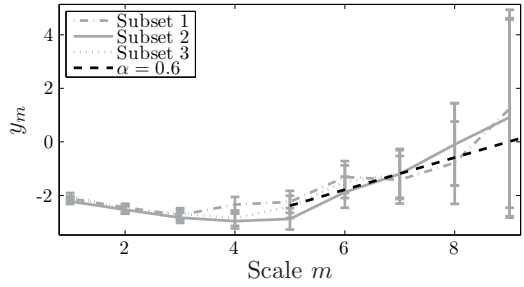
This is not so for Datasets 5 and 6. The stationarity test shows that similar results are observable in all segments, with error bars spanning largely the same space in all cases. This suggests that the assumption of stationarity is valid and results can be taken seriously. Further robustness tests are seen in Figure 7.14, where events were removed at random at different removal rates. No change is observed in the estimate of α. This test is performed because even though both these datasets are less likely to contain missing events than Datasets 1-4, EEG records were marked manually and errors are possible. This gives further confidence in the existence of long memory. For completeness, a comparison is made in this same figure to the case in which only long events are removed. Visible changes in the characteristics of the calculated y_m, as well as the corresponding gradient, are observable in both cases. Long range memory is destroyed by the removal of these events.

Furthermore we can infer the length of the memory by relating the scales m for which LRD is identified to its temporal scale equivalent. We know the resolution and we know the number of events, thus we can say that Dataset 5 identifies memory in the (conservative) range of 16 minutes to 8 hours, and Dataset 6 in the range of 1 to 16 minutes. Longer memory may be present but we do not have access to more data.

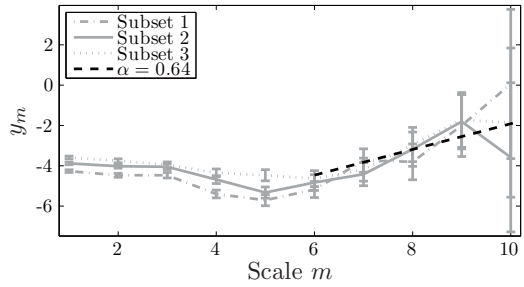
It is important to note that the presence of LRD cannot be rejected for Datasets 1-4, but neither can a pure noise model. The errors caused by insufficient data (because events are rare) are large and make this type of



(a) Dataset 1

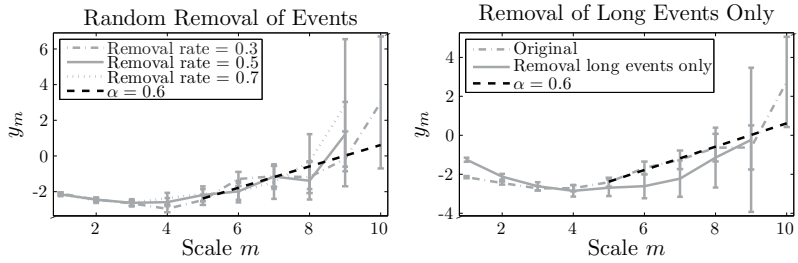


(b) Dataset 5

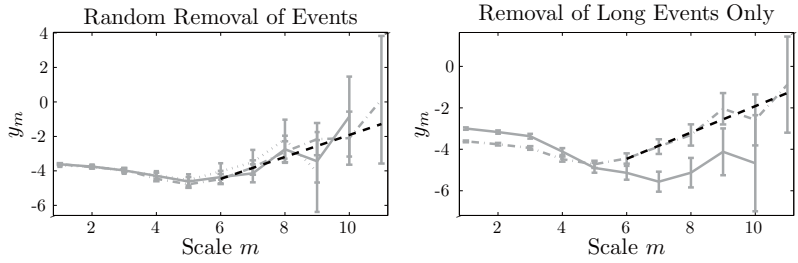


(c) Dataset 6

FIGURE 7.13: Stationarity analysis for Datasets 1, 5 and 6. Datasets were partitioned into 3 consecutive blocks and compared to the initial estimate of α as shown in Figure 7.12. The above shows that Datasets 5 and 6 are likely stationary because each partition gives an α similar to the original calculation. Dataset 1, on the other hand, is more volatile and stationarity is unlikely. This is expected because Dataset 5 and 6 span only hours and days, whereas Dataset 1 spans 25 years over which many changes in the characteristics of the epilepsy likely occur. Because of this lack of stationarity further analysis is unlikely to yield conclusive results, and thus Dataset 1 must be rejected as a candidate for determining LRD.



(a) Dataset 5



(b) Dataset 6

FIGURE 7.14: An estimated α for an LRD process is robust under random removal of events. This is shown to be the case for both Datasets 5 and 6, even for relatively large removal rates. This is not the case when removal of large events is targeted, where the structure at large scales is partly destroyed. The above figures support the existence of LRD in Datasets 5 and 6. The original estimated α is shown in all plots for comparison.

data unsuitable. The analysis in this section can validate an α indicative of LRD for Datasets 5 and 6 only. It is interesting that in both these datasets, which were extracted from EEG, the length of the events is known and that these lengths also follow a power-law indicative of long memory. This can be seen in Figure 7.15. Although possibly useful for modeling purposes in other studies this information does not add any value to the determination of how predictable seizures are, and is only included here for interest's sake.

An α value consistent with long memory up to hours long has been found in 2 out of the 6 datasets. Memory may exist in the processes described by the other 4 datasets, but the data are unable to confirm or deny this, either because the nature of recordings is inadequate, because clinical events are not frequent enough to provide sufficient amounts of data or because memory simply does not exist. It was wavelet tools that allowed these

observations to be made; other tools such as IPH can (more easily) lead to the wrong conclusion.

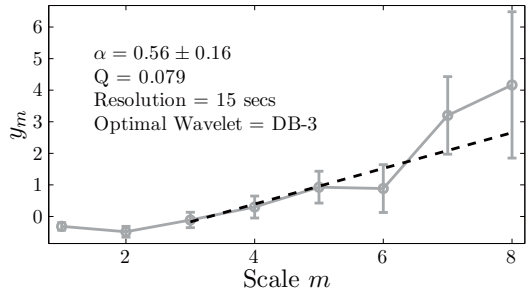
The implications of the discoveries listed here are discussed next.

7.5 Memory and Predictability of Seizures

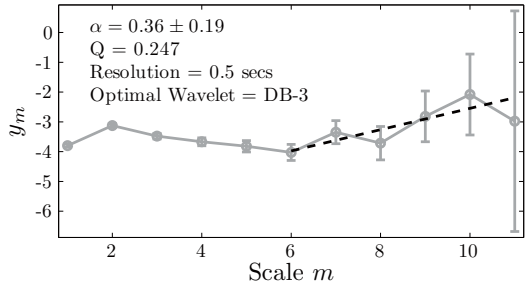
What do these findings tell us about how predictable seizures are? The fact that α was indicative of long memory does not make the underlying process a long memory one. The results simply tell us that *the data are consistent with a model in which LRD is present*, or rather *a short-range memory model is excluded*. It is always possible that the structure discussed in Section 7.4 is due to something else entirely. However, since results have proved stationary over time and because of the nature of the data they do tell us that a purely random model of the generation of epileptic activity is not suitable. A richness of structure exists that may suitably be modeled by stochastic processes with long memory, but the existence of another model equally capable of describing this phenomena is not ruled out [130].

Furthermore, even if the generation of events is due to long memory, the results do not in any way tell us what physical mechanisms are responsible for the existence of this structure. Is it a phenomenon brought about from the neural network architecture in the brain? Perhaps it is due to sensory input? More likely it is to do with both of these, and more. Thus the existence of the scale-invariance has implications for the types of models chosen to simulate brain behavior – they need to be capable of reproducing this type of behavior. It should be kept in mind that the models presented in Chapter 6 do not contain any mechanisms in which memory longer than a few milliseconds is possible. If these local models are accepted as suitable for small scales then the development of global models needs to somehow incorporate it.

If data from this chapter were used to develop these physical models then it is only the short-term studies acquired from clinical records (e.g., datasets 5 and 6) that are usable. In long term qualitative studies the number of events is small and records are unreliable under stationarity tests. This is both good news and bad: short-term records are more easily maintained, but most epileptic patients do not experience a suitably large number of epileptic events in these time frames. Thus if prediction models are derived from this type of data there is a limit in the type of patients that can be helped, once again emphasizing the patient specific nature of a likely predictor. In any case even if these short-term datasets cannot be used directly to predict seizures they can be used to provide insight and ensure that physical models (of a larger scale than those in Chapter 6) are consistent with LRD. They can then



(a) Dataset 5



(b) Dataset 6

FIGURE 7.15: The length of the epileptic events, which are available because data were extracted from EEG records, is shown to also follow a power-law indicative of memory in this relationship as well. The above is not directly applicable to the determination of whether epileptic events are predictable or not, but is of interest in the development of models that must be able to replicate this behavior.

be applied to the problem of prediction, an improvement on the black box predictors available today. The relationships in the length of events briefly introduced in the previous section may also be used to this end.

Another implication on the development of suitable large-scale physical models is that the dimension $\hat{N}$ of the state space in Equation 3.2 must be incredibly large, in fact infinite as a worst case scenario. While this is consistent with the architecture of the brain, known to be very complex, it implies (as does LRD) that prohibitively long data records are needed and this forebodes bad news for the ability to predict seizures at all.

In any case the aim of this chapter is not to develop these models; it is simply to identify the existence of scaling to try and understand if seizures are predictable. The results show evidence of long range memory in the system. This is apparent for studies that span hours and days, but not months or years. Perhaps when one thinks of the capabilities of the brain it is not surprising

that such memory is possible in the brain – what is surprising is that this memory exists between epileptic events, and that it is not short.

For prediction, is this type of memory a good thing or a bad thing? The fact that memory exists at all in epilepsy, beyond very short times, is a positive discovery: at least seizures *could* be predictable! If no memory was observed then no clear evidence exists that seizures are at all predictable. Because event times are not random information about the timing of a current and past seizure does contain information about future seizures. The more information of this type that is available, the more predictable seizures may be.

This is a simplistic argument that holds only for cases in which many epileptiform bursts are observed in short enough time period so that data are roughly stationary. Perhaps this study only tells us what type of data are unsuitable for prediction. Global events measurable by scalp or intracranial EEG are rare and using such measurements is not wise. Predictors should focus on scales at which epileptic events occur more frequently. For example, if this type of memory also exists in local discharges from focal lesions, where events occur much more frequently and more data are available, then a measurement at this scale could be more suitable for prediction. Intracranial records using micro-electrodes that detect activity at these scales exist but have to date largely been ignored in the problem of seizure prediction. An exception to this is the detection/intervention approach taken in [14].

That this memory appears to be so long (up to 8 hours!) casts further doubt on how current predictors of epilepsy work because they typically depend on short-term EEG, often restricting themselves to inter-seizure EEG. Of course the type of data they are looking at is different than the one here and perhaps it is sufficient to achieve success, but given their limited performance it still seems unlikely that this information can be ignored. Why don't current techniques look at a few days rather than a few hours? In theory better baselines can be established in this way, although in practice such long sequences of data cannot be processed easily and the types of features that are extracted are already too computationally demanding.

The identification of scaling regions in data spanning time scales of up to 8 hours, possibly more, implies that the generation of the next epileptic event relies in part on events that occurred a long time ago. This means that to predict seizures it may be necessary to include data representative of these time scales, and it indicates a need to use observations *other* than the EEG. Current predictors of epilepsy do not utilize such long memory.

However, the fact that memory is likely gives hope that seizures are predictable. Future efforts should move toward the development of physical rather than black box models consistent with the existence of LRD discovered here.

7.6 Conclusions

If prediction algorithms are likely to ever evolve beyond black box methods then we must look past the traditional use of EEG records, which are limited in the amount of information they can reveal about the fine-scale behavior of the brain. The work presented in this chapter supports this argument, numerically, by inspecting the point-processes of epileptic events at different time scales.

> Epileptic event times, although seemingly random, are shown here to contain some structure with memory found between events, up to a scale of 8 hours. The type of structure found indicates that long range memory exists in the generators of epileptic seizures.

These findings suggest that current prediction algorithms may be using insufficient amounts of data that do not account for the presence of such long memory. Future efforts must consider this, perhaps by integrating observations other than the EEG into the process.

Although the revelations are useful, the methods used in this chapter are not themselves practical for the implementation of a seizure prediction algorithm – the number of data points required is often prohibitive.

> For this work to be usable directly in prediction algorithms, research must shift to a different type of data in which epileptic events are very frequent. This could simply mean a reduction in the scale at which the EEG is recorded, perhaps using single microelectrodes as opposed to more global electrode arrays.

Although it is unlikely that this data by itself will ever be used as a predictor it is feasible that it could be used instead as a support mechanism to make current predictors more robust. This support can be as simple as the introduction of expected prediction errors.

The identification of scaling in epilepsy could lead to an alternate route of research in which this information is used for the validation or the development of models capable of replicating this behavior. Interestingly LRD was identified in both rat and human EEG data. Is it possible that we have a network only explanation for epilepsy? If so, the brain *structure* may be more

important than the actual dynamics. How much do chemical processes affect the generation of epileptic activity? We should use whatever clues we have access to so as to build an appropriate model of epilepsy.

These models could subsequently be used in the prediction process, and although it is not the point of this text to suggest how this may be done, a predictor and even a detector based on the properties of network architecture, neuron properties and other such factors has a clear advantage over traditional black box time-series analysis. Until the generators of epileptic activity are better understood it is our belief that it is unlikely that implantable devices will ever move beyond the detection regime. This is of course still a workable solution for some – preventing seizures from occurring is the goal after all – but is unlikely to serve the wide variety of epilepsies that exist.

8

Concluding Remarks

The EEG provides us with a window into the brain, and was used in this book to study detection and prediction of epileptic seizures. From a physical principles point of view the EEG measures voltages that are due to electric charges distributed across the volume contained by the skull. There is interference from other electromagnetic field sources outside the skull, but in the general environment of the recording equipment it can be minimized (although not completely eliminated). The EEG measurement reflects the instantaneous charge distribution across the skull volume. The temporal changes in the EEG signal are due to the motion of these charges as a consequence of brain activity as well as artifacts such as muscular/ocular movement. The primary purpose of the EEG is to consider brain activity and ignore the artifact.

Further interpretation of the typical current distribution in the brain, making use of the physiology of the neuronal structure, leads to the model of current dipoles as the primary source for the EEG signal. This in turn reveals that the extracranial or scalp EEG is indeed a very blunt instrument which only records a measurable voltage deviation when large groups of neurons are acting in a coordinated manner. This coordination can be a consequence of the synchronized activity of neurons, or an event of "chance" in a much larger population of neurons. From the brain geometry it follows that the scalp EEG measures cortex activity, as the cortex effectively shields the rest of the brain from the EEG instrumentation. An electrode pair's EEG recording is typically influenced by the cortex activity in a radius of about 2.5cm around each electrode. The scalp EEG is an observation of brain activity at a truly macroscopic scale.

When we analyze EEG records it is important to realize that voltages are only defined up to a constant, and hence only make sense as a voltage difference between two points in space. Moreover, the recording equipment (the location of the electrodes in particular) and the skull geometry play a significant role in determining the magnitude of the measurement signal. Given a recording over a period of time it therefore makes sense to normalize the signal by eliminating the mean (set the mean to zero), and scaling the signal so as to make its standard deviation one. Neither feature, the mean or the variance, have meaning.

From a complexity point of view it is equally clear that the typical scalp EEG is a blunt instrument. The cortex contains in the order of 100 billion neurons, whereas a typical EEG record may contain in the order of a 100

parallel voltage recordings. There is simply not enough information in the EEG to be able to reconstruct the intricate behavior of the brain. Depth electrodes that record the activity of a single neural cell may indeed assist to unravel at this microscopic level the behavior of a single neural cell. But even the meso-scopic scale behavior of a few thousand neurons in a cortical column remains elusive, regardless of which EEG method is used. There remains much more work to be done from a modeling point of view.

Nevertheless, the EEG provides a very useful window into brain activity. It is relatively straightforward to extract features that distinguish clearly between sleep and awake conditions, regardless of the individual being measured. No doubt, the sleep and awake conditions imply some form of difference in behavior at the macro-scopic scale, which can and does indeed show up clearly in any EEG recording.

However, identifying the EEG features that allow one to detect epileptic episodes and differentiate it from normal brain behavior is more intricate. Epilepsy is a collective for a very large group of brain misbehaviors that manifest themselves in a variety of ways in the EEG. There are not only many epilepsy variations between different people, but also in the same patient one may observe differences between epileptic episodes even in the same EEG record.

Using machine learning technology, an expert system that mimics and indeed rivals the performance of a clinician in deciding which epochs of an EEG recording correspond to epileptic behavior can be constructed. The basic features of any such expert system consist of three phases, first signal pre-conditioning, next feature extraction, typically using windowed data, and finally detection. Which features are relevant depend on the application, and typically are learned from annotated EEG recordings. Patient independent and patient specific detection can be catered for in much the same way. Detector performance can be expressed using metrics such as true positive vs. false negative detection rates. Excellent detector results can be obtained with features that track synchronization across a number of recordings. This necessarily implies that the detected seizure affects a significant area of the cortex (occupying a large proportion of the spatial neighborhood of the relevant electrodes).

Intracranial EEG show greater potential in detecting epileptic epochs, at the cost of a rather unwelcome major surgical intervention. In the context of focal epilepsy, the proximity of the intracranial EEG electrodes allows the acquisition of cleaner signals, which makes it easier to learn and to interpret signal features.

Detection of an epileptic episode is in general feasible with only a minor delay from the epileptic onset (in the EEG record). In the case of intracranial EEG, detection may therefore occur even before behavioral clinical symptoms become observable. This is unlikely the case for scalp EEG because of its lack of area specificity, so that onset and detection typically only occur when the epilepsy has progressed significantly and affects a large portion of the cortex.

Early detection opens the possibility for intervention. This is an important topic for further research, with great clinical promise.

Prediction of epilepsy is of course even better. However, is there a clearly identifiable pre-seizure state in the EEG? Research to date is inconclusive. From a dynamical system's point of view there are two very different ways to interpret epilepsy with. First, one can consider epilepsy as the brain's behavior after a bifurcation. A particular condition in the brain changes over time, and the corresponding brain dynamics switch dramatically, from a normal to an epileptic state of behavior once this condition exceeds a certain threshold. Alternatively, it may be that epilepsy is just the behavior of the brain in another part of the state space, a part that a normal functioning brain does not visit. The former could well be a predictable phenomenon, provided the particular brain condition is identifiable from the EEG signal. However the latter model of epilepsy almost defies prediction from a purely EEG point of view, and would best be approached using EEG together with other measurements. Despite these dynamic differences, both modes of operation are entirely compatible with detection of epileptic activity using the EEG.

In order to elucidate the question of predictability of epilepsy the notion of functional memory between epileptic events is useful. If successive epileptic events were not functionally related, predictability is futile. Early evidence points to a long-range dependency in the time series of epileptic onsets. This is positive in that predictability is not outright falsified. It is however somewhat of a Phyrrus victory because long range dependency requires long observation times, and moreover it is indicative of the enormous difficulty one faces in predicting long intervals between successive epileptic events. Further work is called for. It appears essential to analyze (long) EEG time series for long range dependencies and associated power laws in epileptic onset as well as seizure severity to shed more light on this.

Despite its near 100 year history, our collective understanding of the EEG is very much elementary, not in the least because of the enormous complexity of the brain it deals with. There are a great many questions that have been answered, but also there remain many interesting open problems.

Glossary

action potential The response of a neuron to the integration of information incoming from the *dendrites*. The action potential propagates through the cell's axon.

AED See *anti-epileptic drug*.

aliasing A type of *measurement error* introduced when the *sampling rate* is lower than twice the highest frequency found in a signal.

animal model A model of a disorder or pathology that uses a living but non-human subject to replicate the symptoms.

ANN See *artificial neural network*.

anti-epileptic drug (AED) A drug prescribed to stop or minimize the number of seizures.

artifact Activity measured by the EEG that does not originate in the brain. For example, the movements of the eye generate an artifact that interferes with the EEG.

artificial neural network (ANN) A type of *classifier* based on a network of mathematical units designed to mimic the activity of a biological neuron.

association rules A type of *classifier* based on simple relationships between the outputs of a *feature extractor*.

auto-correlation A statistical quantity that estimates how much a signal repeats itself over time.

axon A neural fiber that conducts *action potentials*, the *outputs* generated by a neuron's *cell body*.

bifurcation An abrupt change or split of activity into a different mode of behavior. In mathematics it is a very rigorously defined concept.

bioelectromagnetism The interaction of electric and magnetic fields in biological tissue.

black box dynamic model A mathematical model of a *system* based only on measured data and not on its underlying physical construct.

capacitance A quantity that describes the ability of a body to store charge.

cell body The part of a *neuron* that integrates the *inputs* coming from the *dendrites* to generate the *outputs* projected by the *axon*.

central nervous system (CNS) The organ formed by all the *neurons* in the human body, which together determine behavior. The CNS spans the brain as well as the spinal cord.

cerebral cortex A 1-2mm thick sheet of highly interconnected *neurons* that form the outer layer of the brain. See also *cortical fold*.

cerebrospinal fluid (CSF) The clear fluid found inside and surrounding the brain.

charge A fundamental electric property of sub-atomic particles that defines the *electric field* generated by its interaction with other charges. Charge can be positive or negative.

chemical current See *current*.

classifier A method that identifies the class that a set of extracted features belong to. See for example *ANN*, *SVM* and *association rules*.

CNS See *central nervous system*.

computational noise An error introduced into the estimate of a statistic because of insufficient data or algorithmic issues.

conductivity A quantity that describes the ability of a material to conduct or transmit charge, defined per unit length of the material.

correlation dimension A non-linear statistical quantity that estimates the non-integer dimension of a system. It gives an idea of the complexity of the system generating the signal.

cortical column A small region of brain containing roughly 50,000 neurons that behave similarly to one another. It is believed to be a functional unit of the brain. A cortical column is sometimes referred to as a *macro-column* because a smaller *mini-column* consisting of approximately 100 neurons has also been proposed.

cortical fold The natural folds of the cortex. The exposed curved part of a fold is known as the *gyrus*, while the walls found deeper are known as the *sulci*.

cortico-cortical projection Neural connectivity between different parts of the cortex.

cross-correlation A statistical quantity that estimates how similar two signals are to one another.

CSF See *cerebrospinal fluid*.

current The movement of charged particles. When the particles are electrons the current is known as an *electric current*. When the particles are charged ions the current is known as a *chemical current*.

current dipole The *electric field* generated by the combination of a (theoretical) current source with a current sink. Current dipoles are most commonly used to approximate average electric activity in a small volume of brain.

DBS See *deep-brain stimulation*.

deep-brain stimulation (DBS) The application of electric charge to regions of the brain beneath the cortex.

dendrite Branched fibers that form the *inputs* to a neuron's *cell body*.

detection The correct identification of activity after it occurs.

dimension In mathematics dimension typically refers to the minimum number of co-ordinates required to describe a system. Different definitions exist (see, for example, *correlation dimension*) but all give an idea of the complexity of the system.

dipole The combination of two *charges*, typically used in reference to its generated *electric field*. If the two charges are electric then it is known as an *electric dipole*.

dynamic model A mathematical representation of the behavior of a *system*.

EEG See *electroencephalogram*.

electric current See *current*.

electric dipole See *dipole*.

electric field The *vector* representation of the electric force generated by a particular charge distribution.

electric potential Also known as *voltage*, it is the *scalar* potential energy necessary to move a unit charge between two points. It is related to the *vector electric field*.

electrical stimulation The application of electric *charge* to stimulate activity.

electricity The movement of electric *charge*.

electrode A conductive material that enables recordings of electrical activity. In the *EEG* the electrodes can be placed on the scalp, on the cortex, or in deeper brain structures.

electroencephalogram (EEG) A *measurement system* that records the time-evolving *voltages* generated by the brain.

entropy A non-linear statistical quantity that estimates how much redundancy there is in an observed signal. It gives an idea of the complexity of the system generating the signal.

epilepsy A neurological disorder characterized by repeated occurrences of *seizures*.

EPSP See *post-synaptic potential*.

expert system A sophisticated method of selecting and combining the outputs of both the *feature extraction* and *classifier* stages to optimize the performance of a detector.

extra-cellular fluid The fluid found outside a cell.

false positive When an incorrect detection is made by a classifier.

false positive rate (FPR) See *specificity*.

fast Fourier transform (FFT) An efficient way of computing the *Fourier transform* from a digital *signal*.

feature extraction The estimation of *statistics* from a *signal*, typically after *windowing*. In a detector these features are passed to the *classifier*.

FFT See *fast Fourier transform*.

filter A mathematical transform that selectively removes components of a signal. For example, a low pass filter removes all high frequencies from a signal.

Fourier transform A mathematical transform that converts a signal from *time domain* to *frequency domain*.

FPR See *false positive rate*.

frequency domain The representation of *signals* by their frequency components.

gray matter Nerve tissue primarily consisting of neurons.

gyrus See *cortical fold*.

histology A description of the micro-scopic cellular physiology of tissue.

homogeneous medium A material whose properties (e.g., electrical) are uniform in all (infinite) directions.

in vitro A biological experiment conducted in dead (but preserved) tissue.

in vivo A biological experiment conducted in live tissue.

inhomogeneous medium A material whose properties (e.g., electrical) change over space.

input The *signals* that are inserted into a *system*, used to determine its *outputs*.

integrate-and-fire unit A mathematical model of a single neuron based on statistics of observed events rather than complex anatomical details.

intra-cellular fluid The fluid found inside a cell.

intra-cranial EEG EEG measured by electrodes placed beneath the skull, either on the cortex (cortical EEG) or in deeper structures (depth EEG).

ion gate Channels that allow the transmission of ions between the inside and outside of *neurons*. The gates are typically triggered open by chemical, electrical or physical processes.

IPSP See *post-synaptic potential*.

linear system A *system* whose mathematical descriptors are linear, that is, they are additive and homogeneous. Their responses obey the principle of superposition.

linearization The approximation of a non-linear function by a linear one, valid only around the point at which the linearization is performed.

long range dependence (LRD) A relationship where the current value of a signal/object depends on values that occurred far away in time or space. Also known as *long range memory*.

long range memory See *long range dependence*.

LRD See *long range dependence*.

Lyapunov exponent A non-linear statistical quantity that estimates how fast the datapoints in a *signal* diverge from each other. It gives an idea of the complexity of the system generating the signal.

macro-column See *cortical column*.

macro-scopic model A mathematical model that describes activity at the *macro-scopic scales*, which in the brain refers to modeling the activity in centimeters of brain tissue.

macro-scopic scale Refers to activity of large regions of the brain, spanning scales of cm rather than mm.

map (mathematical) The mathematical relationship between two quantities.

mean A statistical quantity that estimates the average value of a set of data.

measurement A record of the activity of a *system*. For example, see *electroencephalogram*.

measurement error A *noise* or error introduced to a recorded *signal* because of the inability of measurement equipment to measure activity to infinite precision.

membrane potential The *electric potential* between the *intra-cellular* and *extra-cellular* fluid.

meso-scopic scale Refers to the activity of a small network of around 50,000 neurons found in the mm scale (see *cortical column*).

micro-scopic scale Refers to the activity of a few neurons found in the μm scale.

mini-column See *cortical column*.

monopole A single *charge*, typically used in reference to its generated *electric field*.

neural network The complex connectivity between (up to) billions of neurons. Neural networks are necessary for the *central nervous system* to function.

neuromodulator A type of *neurotransmitter* that acts more diffusely and affects large populations of neurons simultaneously. Like neurotransmitters, there are many different types of neuromodulators.

neuron A cell that is the basic functional unit of the *central nervous system*.

neurotransmitter A chemical used to facilitate communication between neurons (at the synapses) in response to an incoming *action potential*. There are many different types of neurotransmitters.

non-linear system A *system* whose mathematical descriptors are not additive, homogeneous, or both. Their responses does not obey the principle of superposition. (See *linear system*.)

normalization The removal of the mean and scale dependence of a signal so that it may be compared to all other signals recorded using different methods.

onset delay A performance metric that represents the delay between the onset of an event and its detection by a classifier.

output The activity projected outside of a *system* in response to its *inputs* as well as its internal processes.

parameter A relatively *stationary* quantity used in a mathematical expression to describe properties of a *system*.

PDF See *probability distribution function*.

permittivity A quantity that describes the ability of a material to store charge, defined per unit length of the material.

phenomenological model A mathematical model that describes the phenomena of the observed signal rather than the underlying physical components. In this book the phenomenological models are used to describe the macro- or meso-scopic potentials recorded by the EEG.

physical dynamic model A mathematical model of a system based on the underlying physical construct (e.g., physiology) of the system.

post-synaptic potential (PSP) The response of a cell to an incoming *action potential*. It can be excitatory (EPSP) or inhibitory (IPSP). EPSPs promote further transmission whilst IPSPs inhibit it.

power law The mathematical relationship formed when a straight line is observed in a log-log plot.

power spectral density (PSD) The absolute squared magnitude of the *Fourier transform* of a signal.

prediction The forecasting of activity in the future.

preprocessing The initial stages of a detector used to *normalize* and standardize a signal prior to feature extraction.

probability distribution function (PDF) A function that describes the probability of each value of a *stochastic variable* occurring.

PSD See *power spectral density*.

PSP See *post-synaptic potential*.

quantization The process of converting an analog signal into a digital signal by dividing the signal space into a finite number of bins.

receiver operating characteristic (ROC) curve A visualization of the performance of a detector by plotting the *false-positive rate* versus the true-positive rate.

resistivity The reciprocal of *conductivity*, that is, the ability of a material to impede the transmission of charge.

ROC See *receiver operating characteristic*.

sampling The conversion of a continuous signal into a sequence of discrete values, typically performed in the measurement process. The rate at which samples are acquired is known as the *sampling frequency*.

sampling frequency / sampling rate See *sampling*.

scalar A quantity that has only magnitude (as opposed to a *vector*).

scale-invariance When the properties of a signal/object are the same when analyzed at different temporal/spatial scales. Scale-invariance is an instance of *self-similarity*.

scalogram A plot of the wavelet scale number versus a logarithmic quantity. It can be used to identify *power laws*.

scalp EEG EEG measured by electrodes placed on the scalp.

seizure A temporary impairment to normal brain function where there is excessive and highly synchronous firing of action potentials. A seizure can affect the entire brain (*generalized seizure*) or parts of the brain (*partial/focal seizure*).

self-similarity When the properties of an arbitrary entity are the same when it is looked at as a whole or in parts.

sensitivity A performance metric that represents the rate of correct detections made by a classifier. Also known as *true positive rate*.

sensory reticular nucleus (SRN) A *sub-system* of the thalamus responsible for relaying sensory input onto the cortex.

short-range memory A relationship where the current value of a signal/object depends on recent previous values.

signal A collection of *measurements* taken from a *system* over time.

signal processing The analysis of *signals* used to extract statistical information.

spatial filter A *filter* that operates on the spatial (as opposed to temporal) domain.

specificity A performance metric that represents the rate of incorrect detections made by a classifier. It is closely related to *false positive rate*.

spectrogram The computation of the *power spectral density* over time using *windowed* signals.

spike A discrete representation of an *action potential*.

SRN See *sensory reticular nucleus*.

stability analysis The process of determining whether a system is *stable* or *unstable*.

stable system When the *steady state* of a system is finite – that is, when the system responses are close at a point in time, they remain close over indefinite periods of time.

standard deviation The square root of *variance*.

stationarity When a quantity (e.g., a parameter) does not change its values or properties over time.

statistic A quantity (e.g., a *variable*) whose average properties are understood as belonging to an ensemble of data.

steady state The resting state of the system in response to a constant input.

stochastic A random quantity or *variable* that may be described using a *probability distribution function*.

sub-system A smaller *system* that can be combined with other sub-systems to describe the entire *system*.

sulcus See *cortical fold*.

support vector machine (SVM) A type of *classifier* based on identifying the vectors in the observed feature space that best separates between classes.

SVM See *support vector machine*.

synapse The connecting region between *axons* and *dendrites* that allows transmission of information between neurons.

synchronization When events at different locations occur at the same time.

system A description of an entity, such as the brain, that by its physiological or physical make-up constrains the *signals* that are generated. The *outputs* of a system are determined by its internal structure as well as its *inputs*.

thalamic reticular nucleus (TRN) A *sub-system* of the thalamus responsible for regulating the amount of sensory input relayed onto the cortex.

thalamo-cortical projection Neural connectivity between thalamus and cortex.

thalamus A region of the brain found sub-cortically that is responsible, among other things, for regulation of sensory input into the cortex. It is also believed to be responsible for synchronizing activity between different parts of the cortex.

time domain The representation of *signals* over time.

TPR See *true positive rate*.

TRN See *thalamic reticular nucleus*.

true positive When a correct detection is made by a classifier.

true positive rate (TPR) See *sensitivity*.

unstable system When the system is not stable, i.e., there are system responses that are close at an instant in time, and then diverge arbitrarily far from each other in the future.

vagal nerve stimulation (VNS) The application of electric *charge* to the vagus nerve, typically used to reduce the number of epileptic seizures.

variable A quantity, typically time-evolving, used in a mathematical expression to describe properties of a *system*.

variance A statistical quantity that describes the divergence of data.

vector A quantity that has both magnitude and direction (as opposed to a *scalar*).

VNS See *vagal nerve stimulation*.

voltage See *electric potential*.

volume conductor A medium, typically modeled in 3 dimensions, that conducts electric charge according to its electric properties: *conductivity*, *resistivity* and *permittivity*.

wavelet analysis An analysis method that extracts both frequency and time domain information from a signal using specialized mathematical functions called wavelets.

white matter Nerve tissue primarily consisting of connecting fibers rather than neurons.

windowing The partitioning of a signal into smaller segments so that its temporal evolution can be analyzed. The applied windows may be rectangular or non-rectangular (e.g., Hanning window, Hamming window).

Bibliography

[1] Proposal for revise clinical and electroencephalographic classification of epileptic seizures. From the commission on classification and terminology of the international league against epilepsy. *Epilepsia 22*, 4 (1981), pp. 489–501.

[2] Abarbanel, H. *Analysis of observed chaotic data*. Springer, 1996.

[3] Adeli, H., Ghosh-Dastidar, S., and Dadmehr, N. A wavelet-chaos methodology for analysis of EEGs and EEG sub-bands to detect seizure and epilepsy. *IEEE Transactions on Biomedical Engineering 54*, 2 (2007), pp. 205–211.

[4] Aeyels, D. Generic observability of differentiable systems. *SIAM Journal on Control and Optimization 19*, 5 (1981), pp. 595–603.

[5] Akhtari, M., Bryant, H., Emin, D., Merrifield, W., Mamelak, A., Flynn, E., Shih, J., Mandelkern, M., Matiachov, A., Ranken, D., Best, E., DiMauro, M., Lee, R., and sutherling, W. A model for frequency dependence of conductivities of the live human skull. *Brain Topography 16*, 1 (2003), pp. 39–55.

[6] Albertos, P., and Mareels, I. *Feedback and Control for Everyone*, 1st ed. Springer, 2010.

[7] Aldroubi, A., and Unser, M. *Wavelets in medicine and biology*. CRC Press, 1996.

[8] Alkan, A., Koklukaya, E., and Subasi, A. Automatic seizure detection in EEG using logistic regression and artificial neural network. *Journal of Neuroscience Methods 148* (2005), pp. 167–176.

[9] Amarb, A., Levya, M., McComba, J., and Apuzzo, M. Vagus nerve stimulation for control of intractable seizures in childhood. *Pediatric Neurosurgery 34* (2001), pp. 218–223.

[10] Andrzejak, R., Lehnertz, K., Mormann, F., Rieke, C., David, P., and Elger, C. Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Physical Review E 64*, 6 (2001), pp. 61907.

[11] Aschenbrenner-Scheibe, R., Maiwald, T., Winterhalder, M., Voss, H., Timmer, J., and Schulze-Bonhage, A. How well can epileptic seizures be predicted? An evaluation of a nonlinear method. *Brain 126* (2003), pp. 2616–2626.

[12] Avoli, M., Louvel, J., Pumain, R., and Köhling, R. Cellular and molecular mechanisms of epilepsy in the human brain. *Progress in Neurobiology 77* (2005), pp. 166–200.

[13] Aziz, J., Karakiewicz, R., Genov, R., Bardakjian, B., Derchansky, M., and Carlen, P. Towards real-time in-implant epileptic seizure prediction. *Proceedings of the 28th Annual IEEE Engineering in Medicine and Biology Conference* (2006).

[14] Aziz, J., Karakiewicz, R., Genov, R., Bardakjian, B., Derchansky, M., and Carlen, P. Towards real-time in-implant epileptic seizure prediction. *Proceedings of the 28th IEEE EMBS Annual International Conference* (2006).

[15] Bak, P., and Chen, K. Self-organized criticality. *Scientific American January* (1991), pp. 26–33.

[16] Banks, G., Regan, K., and Beran, R. *ILAE Comission on Economic Aspects of Epilepsy. Cost of Epilepsy: Proceedings of the 20th International Epilepsy Congress.* Wehr-Baden:Ciba-Geigy Verlag, 1995, ch. The prevalence and direct costs of epilepsy in Australia, pp. 39–48.

[17] Beggs, J., and Plenz, D. Neuronal avalanches in neocortical circuits. *The Journal of Neuroscience 23*, 35 (2003), pp. 11167–11177.

[18] Beran, J. *Statistics for long-memory processes: Monographs on statistics and applied probability.* Chapman and Hall, 1994.

[19] Blackmore, K., Williamson, R., Mareels, I., and Sethares, W. Online learning via congregational gradient descent. *Mathematics of Control, Signals, and Systems 10*, 4 (1997), 331–363.

[20] Boon, P., Vonck, K., Vandekerckhove, T., D'have, M., Nieuwenhuis, L., Michielsen, G., Vanbelleghem, H., Goethals, I., Caemaert, J., Calliauw, L., and De Reuck, J. Vagus nerve stimulation for medically refractory epilepsy; efficacy and cost-benefit analysis. *Acta Neurochirurgica 141* (1999), pp. 447–453.

[21] Braitenberg, V., and Schuz, A. *Anatomy of the cortex: Statistics and geometry.* Springer-Verlag, 1991.

[22] Breakspear, M., Roberts, J., Terry, J., Rodrigues, S., Mahant, N., and Robinson, P. A unifying explanation of primary generalized seizures through nonlinear brain modeling and bifurcation analysis. *Cerebral Cortex 16*, 9 (2006), pp. 1296–1313.

[23] Brunel, N. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of Computational Neuroscience 8* (2000), pp. 183–208.

[24] Burkitt, A. A review of the integrate-and-fire neuron model: I. Homogeneous synaptic input. *Biological Cybernetics 95*, 1 (2006), pp. 1–19.

[25] Burkitt, A. A review of the integrate-and-fire neuron model: II. Inhomogeneous synaptic input and network properties. *Biological Cybernetics 95*, 2 (2006), pp. 97–112.

[26] Casdagli, M., Iasemidis, L., Savit, R., Gilmore, R., Roper, S., and Sackellares, J. Non-linearity in invasive EEG recordings from patients with temporal lobe epilepsy. *Electroencephalography and Clinical Neurophysiology 102* (1997), pp. 98–105.

[27] Chavas, J., and Marty, A. Coexistence of excitatory and inhibitory GABA synapses in the cerebellar interneuron network. *Journal of Neuroscience 23*, 6 (2003), pp. 2019.

[28] Chavez, M., Le Van Quyen, M., Navarro, V., Baulac, M., and Martinerie, J. Spatio-temporal dynamics prior to neocortical seizures: Amplitude versus phase coupling. *IEEE Transactions on Biomedical Engineering 50*, 5 (2003), pp. 571–583.

[29] Cheng, D. *Field and wave electromagnetics*, 2nd ed. Addison-Wesley Publishing Company Inc., 1992.

[30] Compumedics^{TM}. *Compumedics profusion EEG user guide*, 12th ed., 2002.

[31] Cooper, R., Winter, A., Crow, H., and Walter, W. Comparison of subcortical, cortical and scalp activity using chronically indwelling electrodes in man. *Electroencephalography and Clinical Neurophysiology 18* (1965), pp. 217–228.

[32] Da Silva, L., Kamphuis, W., Titulaer, M., Vreugdenhil, M., and Wadman, W. An experimental-model of progressive epilepsy - the development of kindling of the hippocampus of the rat. *Italian Journal of Neurological Sciences 16*, 1-2 (1995), pp. 45–57.

[33] Diambra, L., Bastos de Figueiredo, J., and Malta, C. Epileptic activity recognition in EEG recording. *Physica A 273* (1999), pp. 495–505.

[34] Diks, C. *Nonlinear time series analysis: Methods and applications*. World Scientific Publishing Co., 1999.

[35] Doherty, M., and Haltiner, A. Wag the dog: Skepticism on seizure alert canines. *Neurology 68* (2007), pp. 308.

[36] Doukhan, P., Oppenheim, G., and Taqqu, M. *Theory and applications of long-range dependence*. Birkhäuser, 2003.

[37] Elger, C., Mormann, F., Kreuz, T., Andrzejak, R., Rieke, C., Sowa, R., Florin, S., David, P., and Lehnertz, K. Characterizing the spatio-temporal dynamics of the epileptogenic process with nonlinear EEG analysis. *Proceedings of the 7th IEEE International Workshop in Cellular Neural Networks and their Applications* (2002), pp. 228–248.

[38] Elger, G., Hoppe, C., Falkai, P., Rush, A., and Elger, C. Vagus nerve stimulation is associated with mood improvements in epilepsy patients. *Epilepsy Research 42*, 2-3 (2000), pp. 203–210.

[39] Esteller, R., Echauz, J., D'Alessandro, M., Worrell, G., Cranstoun, S., Vachtsevanos, G., and Litt, B. Continuous energy variation during the seizure cycle: towards an on-line accumulated energy. *Clinical Neurophysiology 116*, 3 (2005), pp. 517–526.

[40] Ferree, T., Eriksen, K., and Tucker, D. Regional head tissue conductivity estimation for improved EEG analysis. *IEEE Transactions on Biomedical Engineering 47*, 12 (2000), pp. 1584–1592.

[41] Franaszczuk, P., Bergey, G., Durka, P., and Eisenberg, H. Time-frequency analysis using the matching pursuit algorithm applied to seizures originating from the mesial temporal lobe. *Electroencephalography and Clinical Neurophysiology 106* (1998), pp. 513–521.

[42] Freeman, W. Simulation of chaotic EEG patterns with a dynamic model of the olfactory system. *Biological Cybernetics 56*, 2-3 (1987), pp. 139–150.

[43] Freeman, W., Chang, H., Burke, B., Rose, P., and Badler, J. Taming chaos: stabilization of aperiodic attractors by noise. *IEEE Trans. on Circuits and Systems 44*, 10 (1997), pp. 989–996.

[44] Gabor, A. Seizure detection using a self organising neural network: Variation and comparison with other detection strategies. *Electroencephalography and Clinical Neurophysiology 107* (1998), pp. 27–32.

[45] Gabor, A., Leach, R., and Dowla, F. Automated seizure detection using a self-organised neural network. *Electroencephalography and Clinical Neurophysiology 99* (1996), pp. 257–266.

[46] Gabriel, C., Gabriel, S., and Corthout, E. The dielectric properties of biological tissues : i. literature survey. *Physics in Medicine and Biology 41* (1996), pp. 2231–2249.

[47] Gabriel, S., Lau, R., and Gabriel, C. The dielectric properties of biological tissues : ii. measurements in the frequency range 10Hz to 20GHz. *Physics in Medicine and Biology 41* (1996), pp. 2251–2269.

[48] Gabriel, S., Lau, R., and Gabriel, C. The dielectric properties of biological tissues : iii. parametric models for the dielectric spectrum of tissues. *Physics in Medicine and Biology 41* (1996), pp. 2271–2293.

[49] Gonzalez-Vellon, B., Sanei, S., and Chambers, J. Support vector machines for seizure detection. *Proceedings of the 3rd IEEE International Symposium on Signal Processing and Information Technology* (2003), pp. 126–129.

[50] Gotman, J. Automatic recognition of epileptic seizures in the EEG. *Electroencephalography and Clinical Neurophysiology 54* (1982), pp. 530–540.

[51] Gotman, J. Automatic seizure detection: improvements and evaluation. *Electroencephalography and Clinical Neurophysiology 76* (1990), pp. 317–324.

[52] Gotman, J. Automatic detection of seizures and spikes. *Journal of Clinical Neurophysiology 16*, 2 (1999), pp. 130–140.

[53] Gotman, J., Flanagan, D., Rosenblatt, B., Bye, A., and Mizrahi, E. Evaluation of an automatic seizure detection method for the newborn EEG. *Electroencephalography and Clinical Neurophysiology 103* (1997), pp. 363–369.

[54] Gotman, J., and Gloor, P. Automatic recognition and quantification of interictal epileptic activity in the human scalp EEG. *Electroencephalography and Clinical Neurophysiology 41* (1976), pp. 513–529.

[55] Gotman, J., Ives, J., and Gloor, P. Frequency content of EEG and EMG at seizure onset: possibility of removal of EMG artefact by digital filtering. *Electroencephalography and Clinical Neurophysiology 52* (1981), pp. 626–639.

[56] Gray, H. *Anatomy of the human body.* Lea & Febiger, 1918.

[57] Grewal, S., and Gotman, J. An automatic warning system for epileptic seizures recorded on intracerebral EEGs. *Clinical Neurophysiology 116* (2005), pp. 2460–2472.

[58] Grossman, S. *Elementary linear algebra*, 5th ed. Saunders College Publishing, 1994.

[59] Haas, S., Frei, M., and Osorio, I. Strategies for adapting automated seizure detection algorithms. *Medical Engineering and Physics 29*, 8 (2007), pp. 895909.

[60] Haken, H. *Brain dynamics: Synchronisation and activity patterns in pulse-coupled neural nets with delays and noise*. Springer-Verlag, 2002.

[61] Harrison, M., Frei, M., and Osorio, I. Accumulated energy revisited. *Clinical Neurophysiology 116* (2005), pp. 711–715.

[62] Harrison, M., Osorio, I., Frei, M., Asuri, S., and Lai, Y. Correlation dimension and integral do not predict epileptic seizures. *Chaos 15* (2005), pp. 33106.

[63] Haut, S., Hall, C., Le Valley, A., and Lipton, R. Can patients with epilepsy predict their own seizures? *Neurology 68* (2007), pp. 262–266.

[64] Haykin, S. *Neural networks: A comprehensive foundation*, 2nd ed. Delhi : Pearson Education, 1999.

[65] Hazarika, N., Chen, J., Tsoi, A., and Sergejew, A. Classification of EEG signals using the wavelet transform. *Signal Processing 59* (1997), pp. 61–72.

[66] Hodgkin, A., and Huxley, A. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology 117* (1952), pp. 500–544.

[67] Hohn, N., Veitch, D., and Abry, P. Does fractal scaling at the IP level depend on TCP flow arrival processes? *Proceedings of the ACM SIGCOMM Internet Measurement Workshop* (2002), pp. 6368.

[68] Hoppensteadt, F., and Izhikevich, E. Thalamo-cortical interactions modeled by weakly connected oscillators: Could the brain use FM radio principles? *BioSystems 48* (1998), pp. 85–94.

[69] Hyvärinen, A., Karhunen, J., and Oja, E. *Independent Component Analysis*. John Wiley & Sons, Inc., 2001.

[70] Iasemidis, L. Epileptic seizure prediction and control. *IEEE Transactions on Biomedical Engineering 50*, 5 (2003), pp. 549–558.

[71] Iasemidis, L., Shiau, D., Chaovalitwongse, W., Sackellares, J., Pardalos, P., Principe, J., Carney, P., Prasad, A., Veeramani, B., and Tsakalis, K. Adaptive epileptic seizure prediction system. *IEEE Transactions on Biomedical Engineering 50*, 5 (2003), pp. 616–627.

[72] Iasemidis, L., Shiau, D., Sackellares, J., Pardalos, P., and Prasad, A. Dynamical resetting of the human brain at epileptic seizures: Application of nonlinear dynamics and global optimization techniques. *IEEE Transactions on Biomedical Engineering 51*, 3 (2004), pp. 493–506.

[73] Jefferys, J. Models and mechanisms of experimental epilepsies. *Epilepsia 44*, Suppl. 12 (2003), pp. 44–50.

[74] Jerger, K., Netoff, T., Francis, J., Sauer, T., Pecora, L., Weinstein, S., and Schiff, S. Early seizure detection. *Journal of Clinical Neurophysiology 18* (2001), pp. 259–268.

[75] Jerger, K., Weinstein, S., Sauer, T., and Schiff, S. Multivariate linear discrimination of seizures. *Clinical Neurophysiology 116* (2005), pp. 545–551.

[76] Jones, D. *Methods in electromagnetic wave propagation.* Oxford University Press, 1979.

[77] Jouny, C., Franaszczuk, P., and Bergey, G. Characterization of epileptic seizure dynamics using gabor atom density. *Clinical Neurophysiology 114* (2003), pp. 426–437.

[78] Joyce, C., Gorodnitsky, I., and Kutas, M. Automatic removal of the eye movement and blink artifacts from EEG data using blind component separation. *Psychophysiology 41* (2004), pp. 313–325.

[79] Jung, T., Makeig, S., Humphries, C., Lee, T., McKeown, M., Iragui, V., and Sejnowski, T. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology 37* (2000), pp. 163–178.

[80] Kandel, E., Schwartz, J., and Jessell, T. *Principles of neural science*, 4th ed. McGraw-Hill Companies, 2000.

[81] Kantz, H., and Schreiber, T. *Nonlinear time series analysis*, 2nd ed. Cambridge University Press, 2003.

[82] Kapitaniak, T. *Chaos for engineers: Theory, application and control.* Springer-Verlag, 1998.

[83] Khalil, H. *Nonlinear systems*, 3rd ed. Prentice Hall, 2002.

[84] Kotsopoulus, I., Evers, S., Ament, A., and De Krom, M. Estimating the costs of epilepsy: An international comparison of epilepsy cost studies. *Epilepsia 42*, 5 (2001), pp. 634–640.

[85] Krauss, G. J., Choi, S., and Lesser, R. Pseudoseizure dogs. *Neurology 68* (2007), pp. 308–309.

[86] Kreyszig, E. *Advanced engineering mathematics*, 9th ed. John Wiley & Sons, Inc., 2005.

[87] Lai, Y., Drongelen, W., Ding, L., Hecox, K., Towle, V., Frim, D., and He, B. Estimation of in vivo human brain-to-skull conductivity ratio from simultaneous extra- and intra-cranial electrical potential recordings. *Clinical Neurophysiology 116* (2005), pp. 456–465.

[88] Lai, Y., Harrison, M., Frei, M., and Osorio, I. Controlled test for predictive power of lyapunov exponents: Their inability to predict epileptic seizures. *Chaos 14*, 3 (2004), pp. 630–642.

[89] Lamsa, K., and Taira, T. Use-dependent shift from inhibitory to excitatory GABAA receptor action in SP-O interneurons in the rat hippocampal CA3 area. *Journal of Neurophysiology 90*, 3 (2003), pp. 1983–95.

[90] Lehnertz, K., and Elger, C. Can epileptic seizures be predicted? Evidence from nonlinear time series analysis of brain electrical activity. *Physical Review Letters 80*, 22 (1998), pp. 5019–5022.

[91] Lehnertz, K., Mormann, F., Kreuz, T., Andrzejak, R., Rieke, C., David, P., and Elger, C. Seizure prediction by nonlinear EEG analysis. *IEEE Engineering in Medicine and Biology Jan-Feb* (2003), pp. 57–63.

[92] Lemieux, L., McBride, A., and Hand, J. Calculation of electrical potentials on the surface of a realistic head model by finite differences. *Physics in Medicine and Biology 41* (1996), pp. 1079–1091.

[93] Leon-Garcia, A. *Probability and random processes for electrical engineering.* Addison-Wesley, 1993.

[94] Li, D., Zhou, W., Drury, I., and Savit, R. Linear and nonlinear measures and seizure anticipation in temporal lobe epilepsy. *Journal of Computational Neuroscience 15* (2003), pp. 335–345.

[95] Liley, D., Cadusch, P., and Dafilis, M. A spatially continuous mean field theory of electrocortical activity. *Network: Computation in Neural Systems 13* (2002), pp. 67–113.

[96] Litt, B., and Lehnertz, K. Seizure prediction and the preseizure period. *Current Opinion in Neurology 15* (2002), pp. 173–177.

[97] Liu, A., Hahn, J., Heldt, G., and Coen, R. Detection of neonatal seizures through computerized EEG analysis. *Electroencephalography and Clinical Neurophysiology 1* (1982), pp. 30–37.

[98] Liu, H., Zhang, T., and Yang, F. A multistage, multimethod approach for automatic detection and classification of epileptiform EEG. *IEEE Transactions on Biomedical Engineering 49*, 12 (2002), pp. 1557–1566.

[99] Lowery, M., Stoykov, N., Dewald, J., and Kuiken, T. Volume conduction in an anatomycally based surface EMG model. *IEEE Transactions on Biomedical Engineering 51*, 12 (2004), pp. 2138–2147.

[100] Luo, F., and Unbehauen, R. *Applied neural networks for signal processing.* Cambridge University Press, 1997.

[101] Maiwald, T., Winterhalder, M., Aschenbrenner-Scheibe, R., adn A Schulze-Bonhage, H. V., and Timmer, J. Comparison of three nonlinear seizure prediction methods by means of the seizure prediction characteristic. *Physica D 194* (2004), pp. 357–368.

[102] Mallat, S. *A wavelet tour of signal processing*, 2nd ed. Academic Press, 1999.

[103] Malmivuo, J., and Plonsey, R. *Bioelectromagnetism: Principles and applications of bioelectric and biomagnetic fields.* Oxford University Press, 1995.

[104] Markovsky, I., Willems, J., Van Huffel, S., and De Moor, B. *Exact and approximate modeling of linear systems: A behavioral approach.* SIAM, 2006.

[105] Marsden, E., and Tromba, A. *Vector calculus*, 4th ed. W.H. Freeman and Company, 1996.

[106] Martinerie, J., Adam, C., Le Van Quyen, M., Baulac, M., Clemenceau, S., Renault, B., and Varela, F. Epileptic seizures can be anticipated by non-linear analysis. *Nature Medicine 4* (1998), pp. 1173–1176.

[107] McLachlan, R., Sadler, M., Pillay, N., Guberman, A., Jones, M., Wiebe, S., and Schneiderman, J. Quality of life after vagus nerve stimulation for intractable epilepsy: is seizure control the only contributing factor? *European Neurology 50*, 1 (2003), pp. 16–19.

[108] Meffin, H., Burkitt, A., and Grayden, D. An analytical model for the 'large, fluctuating synaptic conductance state' typical of neocortical neurons in vivo. *Journal of Computational Neuroscience 16* (2004), pp. 159–175.

[109] Misra, D. *Practical electromagnetics: From biomedical sciences to wireless communication.* John Wiley and Sons, 2007.

[110] Mitra, S. *Digital signal processing: A computer based approach*, 3rd ed. McGraw-Hill, 2006.

[111] Mormann, F., Andrzejak, R., Elger, C., and Lehnertz, K. Seizure prediction: the long and winding road. *Brain 130*, 3 (2007), pp. 314–33.

[112] Mormann, F., Andrzejak, R., Kreuz, T., Rieke, C., David, P., Elger, C., and Lehnertz, K. Automated detection of a preseizure state based on a decrease in synchronization in intracranial electroencephalogram recordings from epilepsy patients. *Physical Review E 67*, 2 (2003), pp. 21912.

[113] Mormann, F., Kreuz, T., Andrzejak, R., David, P., Lehnertz, K., and Elger, C. Epileptic seizures are preceded by a decrease in synchronization. *Epilepsy Research 53* (2003), pp. 173–185.

[114] Munck, J., Dijk, B., and Spekreijse, H. Mathematical dipoles are adequate to describe realistic generators of human brain activity. *IEEE Transactions on Biomedical Engineering 35*, 11 (1988), pp. 960–966.

[115] Munck, J., and Peters, M. A fast method to compute the potential in the multisphere model. *IEEE Transactions on Biomedical Engineering 40*, 11 (1993), pp. 1166–1174.

[116] Murphy, J., and Patil, A. Stimulation of the nervous system for the management of seizures: Current and future developments. *CNS Drugs 17*, 2 (2003), pp. 101–115.

[117] Niederhauser, J., Esteller, R., Echauz, J., Vachtsevanos, G., and Litt, B. Detection of seizure precursors from depth EEG using a sign periodogram transform. *IEEE Transactions on Biomedical Engineering 51*, 4 (2003), pp. 449–458.

[118] Niedermeyer, E., and Da Silva, F. *Electroencephalography: Basic principles, clinical applications, and related fields*, 4th ed. Williams and Wilkins, 1999.

[119] Notley, S., and Elliott, S. Efficient estimation of a time varying dimension parameter and its application to EEG analysis. *IEEE Transactions on Biomedical Engineering 50*, 5 (2003), pp. 594–602.

[120] Nunez, P. *Neocortical dynamics and human EEG rhythms*. Oxford University Press, 1995.

[121] Nunez, P., and Silberstein, R. On the relationship of synaptic activity to macroscopic measurements: Does co-registration of EEG with fMRI make sense? *Brain Topography 13*, 2 (2000), pp. 79–96.

[122] Nunez, P., and Srinivasan, R. *Electric fields of the brain: The neurophysics of EEG*, 2nd ed. Oxford University Press, 2006.

[123] Nunez, P., Srinivasan, R., Westdorp, A., Wijesinghe, R., Tucker, D., Silberstein, R., and Cadusch, P. EEG coherency i: Statistics, reference electrode, volume conduction, laplacians, cortical imaging, and interpretation at multiple scales. *Electroencephalography and Clinical Neurophysiology 103* (1997), pp. 499–515.

[124] OED-Online. *The Oxford English Dictionary.* Oxford University Press, 1989.

[125] O'Leary, J., and Goldring, S. *Science and epilepsy: Neuroscience gains in epilepsy research.* Raven Press Books, 1976.

[126] Oostendorp, T., Delbeke, J., and Stegeman, D. The conductivity of the human skull: Results of in vivo and in vitro measurements. *IEEE Transactions on Biomedical Engineering 47*, 11 (2000), pp. 1487–1492.

[127] Osorio, I., Frei, M., Giftakis, J., Peters, T., Ingram, J., Turnbull, M., Herzog, M., Rise, M., Schaffner, S., Wennberg, R., Walczak, T., Risinger, M., and Ajmone-Marsan, C. Performance reassessment of a real-time seizure-detection algorithm on long ECoG series. *Epilepsia 43*, 12 (2002), pp. 1522–1535.

[128] Osorio, I., Frei, M., and Wilkinson, S. Real-time automated detection and quantitative analysis of seizures and short-term prediction of clinical onset. *Epilepsia 39*, 6 (1998), pp. 615–627.

[129] Päivinen, N., Lammi, S., Pitkänen, A., Nissinen, J., Penttonen, M., and Grönfors, T. Epileptic seizure detection: a nonlinear viewpoint. *Computer Methods and Programs in Biomedicine 79* (2005), pp. 151–159.

[130] Park, K. *Self-similar network traffic and performance evaluation.* John Wiley & Sons, Inc., 2000.

[131] Parker, S., and Chua, L. *Practical numerical algorithms for chaotic systems.* Springer-Verlag, New York, Inc., 1989.

[132] Pauri, F., Pierelli, F., Chatrian, G., and Erdly, W. Long-term EEG-video-audio monitoring: computer detection of focal EEG seizure patterns. *Electroencephalography and Clinical Neurophysiology 82* (1992), pp. 1–9.

[133] Perucca, E. An introduction to antiepileptic drugs. *Epilepsia 46*, Suppl. 4 (2005), pp. 31–37.

[134] Pitkänen, A., Schwartzkroin, P., and Moshé, S. *Models of seizures and epilepsy*. Academic Press, 2005.

[135] Polat, K., and Günes, S. Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast fourier transform. *Applied Mathematics and Computation 187*, 2 (2007), pp. 1017–1026.

[136] Polderman, J., and Willems, J. *Introduction to mathematical systems theory: A behavioral approach*. Springer, New York, 1998.

[137] Pradhan, N., Sadasivan, P., and Arunodaya, G. Detection of seizure activity in EEG by an artificial neural network: a preliminary study. *Computers and Biomedical Research 29* (1996), pp. 303–313.

[138] Proakis, J., and Manolakis, D. *Digital signal processing: Principles, algorithms, and applications*, 3rd ed. Prentice Hall, 1996.

[139] Qu, H., and Gotman, J. Improvement in seizure detection performance by automatic adaptation to the EEG of each patient. *Electroencephalography and Clinical Neurophysiology 86* (1993), pp. 79–87.

[140] Qu, H., and Gotman, J. A seizure warning system for long-term epilepsy monitoring. *Neurology 45* (1995), pp. 2250–2254.

[141] Qu, H., and Gotman, J. A patient-specific algorithm for the detection of seizure onset in long-term EEG monitoring: possible use as a warning device. *IEEE Transactions on Biomedical Engineering 44*, 2 (1997), pp. 115–122.

[142] Rajna, P., Clemens, B., Csibri, E., Dobos, E., Geregely, A., Gottschal, M., György, I., Horvath, A., Horvath, F., Mezöfi, L., Velkey, I., Veres, J., and Wagner, E. Hungarian multicenter epidemiologic study of the warning and initial symptoms (prodome aura) of epileptic seizures. *Seizure 6* (1997), pp. 361–368.

[143] Rao, R., and Bopardikar, A. *Wavelet transforms: Introduction to theory and applications*. Addison-Wesley Longman, Inc., 1998.

[144] Rennie, C., Wright, J., and Robinson, P. Mechanisms of cortical electrical activity and emergence of gamma rhythm. *Journal of Theoretical Biology 205* (2000), pp. 17–35.

[145] Rieke, C., Mormann, F., Andrzejak, R., Kreuz, T., David, P., Elger, C., and Lehnertz, K. Discerning nonstationarity from nonlinearity in seizure-free and preseizure EEG recordings from epilepsy patients. *IEEE Transactions on Biomedical Engineering 50*, 5 (2003), pp. 634–639.

[146] Robinson, P., Rennie, C., and Rowe, D. Dynamics of large-scale brain activity in normal arousal states and epileptic seizures. *Physical Review E 65* (2002), pp. 41924.

[147] Robinson, P., Rennie, C., Rowe, D., O'Connor, S., Wright, J., Gordon, E., and Whitehouse, R. Neurophysical modeling of brain dynamics. *Neuropsychopharmacology 28* (2003), pp. 74–79.

[148] Robinson, P., Rennie, C., and Wright, J. Propagation and stability of waves of electrical activity in the cerebral cortex. *Physical Review E 56*, 1 (1997), pp. 826–840.

[149] Roland, P., and Zilles, K. Structural divisions and functional fields in the human cerebral cortex. *Brain Research Reviews 26* (1998), pp. 87–105.

[150] Romero, S., Mananas, M., and Barbanoj, M. Ocular reduction in EEG signals based on adaptive filtering, regression and blind source separation. *Annals of Biomedical Engineering 37*, 1 (2009), pp. 176–191.

[151] Rosso, O., and Mairal, M. Characterization of time dynamical evolution of electroencephalographic epileptic records. *Physica A 312* (2002), pp. 469–504.

[152] Rosso, O., Martin, M., and Plastino, A. Brain electrical activity analysis using wavelet-based informational tools. *Physica A 313* (2002), pp. 587–608.

[153] Rosso, O., Martin, M., and Plastino, A. Evidence of self-organization in brain electrical activity using wavelet-based information tools. *Physica A 347* (2005), pp. 444–464.

[154] Rowe, D., Robinson, P., and Rennie, C. Estimation of neurophysiological parameters from the waking EEG using a biophysical model of brain dynamics. *Journal of Theoretical Biology 231*, 2 (2004), pp. 413–433.

[155] Rush, A., and Driscoll, D. Current distribution in the brain from surface electrodes. *Anesthesia and Analgesia 47*, 6 (1968), pp, 717–723.

[156] Saab, M., and Gotman, J. A system to detect the onset of epileptic seizures in scalp EEG. *Clinical Neurophysiology 116* (2005), pp. 427–442.

[157] Saha, S., and Williams, P. Electric and dielectric properties of wet human cortical bone as a function of frequency. *IEEE Transactions on Biomedical Engineering 39*, 12 (1992), pp. 1298–1304.

[158] Sander, J., and Hart, Y. *Epilepsy: Questions and answers.* Merit Publishing International, 1997.

[159] Sastry, S. *Nonlinear systems: Analysis, stability and control.* Springer-Verlag, New York, Inc., 1999.

[160] Schelter, B., Winterhalder, M., Maiwald, T., Brandt, A., Schad, A., Schulze-Bonhage, A., and Timmer, J. Testing statistical significance of multivariate time series analysis techniques for epileptic seizure prediction. *Chaos 16* (2006), pp. 13108.

[161] Schid-Shönbein, C. Improvement of seizure control by psychological methods in patients with intractable epilepsies. *Seizure 7* (1998), pp. 261–270.

[162] Schiff, S. Wavelet transforms for epileptic spike and seizure detection. *IEEE 2* (1994), pp. 1214–1215.

[163] Schiff, S., Colella, D., Jacyna, G., Hughes, E., Creekmore, J., Marshall, A., Bozek-Kuzmicki, M., Benke, G., Gaillard, W., Conry, J., and Weinstein, S. Brain chirps: spectrographic signatures of epileptic seizures. *Clinical Neurophysiology 111* (2000), pp. 953–958.

[164] Schimpf, P., Ramon, C., and Haueisen, J. Dipole models for the mboxEEG and MEG. *IEEE Transactions on Biomedical Engineering 49*, 5 (2002), pp. 409–418.

[165] Schindler, K., Wiest, R., Kollar, M., and Donati, F. EEG analysis with simulated neuronal cell models helps to detect pre-seizure changes. *Clinical Neurophysiology 113* (2002), pp. 604.

[166] Schmidt, R. *Epilepsy.* BJ Wilder, 1968.

[167] Sekino, M., Inoue, Y., and Ueno, S. Magnetic resonance imaging of electrical conductivity in the human brain. *IEEE Transactions on Magnetics 41*, 10 (2005), pp. 4203–4205.

[168] Shepherd, G. *Synaptic organization of the brain*, 4th ed. Oxford University Press, 1998.

[169] Sherman, S., and Guillery, R. The role of the thalamus in the flow of information to the cortex. *Philosophical Transactions of the Royal Society of London B 357* (2002), pp. 1695–1708.

[170] Shoeb, A., Edwards, H., Connolly, J., Bourgeois, B., Treves, T., and Guttag, J. Patient-specific seizure onset detection. *Epilepsy and Behavior 5* (2004), pp. 483–498.

[171] Shoeb, A., Schachter, S., Schomer, D., Bourgeois, B., Treves, S., and Guttag, J. Detecting seizure onset in the ambulatory setting: demonstrating feasibility. *Proceedings 27th IEEE EMBS* (2005).

[172] Srinivasan, R., Nunez, P., Tucker, D., Silberstein, R., and Cadusch, P. Spatial sampling and filtering of EEG with spline laplacians to estimate cortical potentials. *Brain Topography 8*, 4 (1996), pp. 355–366.

[173] Steeb, W. *The nonlinear workbook*, 3rd ed. World Scientific Publishing Co. Pte. Ltd., 2005.

[174] Stuller, J. *An Introduction to Signals and Systems.* Thomson Learning, 2007.

[175] Subasi, A. Epileptic seizure detection using dynamic wavelet network. *Expert Systems With Applications 29* (2005), pp. 343–355.

[176] Subasi, A. EEG signal classification using wavelet feature extraction and a mixture of expert model. *Expert Systems With Applications 32*, 4 (2006), pp. 1084–1093.

[177] Subasi, A., Alkan, A., Koklukaya, E., and Kiymik, M. Wavelet neural network classification of EEG signals by using AR model with MLE pre-processing. *Neural Networks 18* (2005), pp. 985–997.

[178] Subasi, A., and Ercelebi, E. Classification of EEG signals using neural network and logistic regression. *Computer Methods and Programs in Biomedicine 78* (2005), pp. 87–99.

[179] Takens, F. Detecting strange attractors in turbulence. *Lecture Notes in Mathematics 898* (1981), pp. 366–381.

[180] Temkin, O. *Falling sickness: History of epilepsy from the Greeks to the beginnings of modern neurology*, 2nd ed. John Hopkins U.P., 1994.

[181] Thakor, N., and Tong, S. Advances in quantitative electroencephalogram analysis methods. *Annual Review of Biomedical Engineering 6* (2004), pp. 453–495.

[182] Thompson, S. *Elementary lessons in electricity and magnetism.* Macmillan Company, 1902.

[183] Tong, H. *Dimension estimation and models.* World Scientific Publishing Co., 2003.

[184] Van Burik, M., and Peters, M. Estimation of the electric conductivity from scalp measurements: Feasibility and application to source localization. *Clinical Neurophysiology 111* (2000), pp. 1514–1521.

[185] Van Putten, M., Kind, T., Visser, F., and Lagerburg, V. Detecting temporal lobe seizures from scalp EEG recordings: a comparison of various features. *Clinical Neurophysiology 116* (2005), pp. 2480–2489.

[186] Varsavsky, A., and Mareels, I. Patient un-specific detection of epileptic seizures through changes in variance. *Proceedings of the 28th Annual IEEE Engineering in Medicine and Biology Conference* (2006).

[187] Varsavsky, A., and Mareels, I. A complete strategy for patient un-specific detection of epileptic seizures using crude estimations of entropy. *Proceedings of the 29th Annual IEEE Engineering in Medicine and Biology Conference* (2007).

[188] Webber, W., Lesser, R., Richardson, R., and Wilson, K. An approach to seizure detection using artificial neural networks. *Electroencephalography and Clinical Neurophysiology 98* (1996), pp. 250–272.

[189] Weinstein, S. The anticonvulsant effect of electric fields. *Current Neurology and Neuroscience Reports 1* (2001), pp. 155–161.

[190] Weng, W., and Khorasani, K. An adaptive structure neural network with application to EEG automatic seizure detection. *Neural Networks 9*, 7 (1996), pp. 1223–1240.

[191] Wilson, S. A neural network method for automatic and incremental learning applied to patient-dependent seizure detection. *Clinical Neurophysiology 116* (2005), pp. 1785–1795.

[192] Wilson, S. Algorithm architectures for patient dependent seizure detection. *Clinical Neurophysiology 117* (2006), pp. 1204–1216.

[193] Wilson, S., Scheuer, M., Plummer, C., Young, B., and Parcia, S. Seizure detection: correlation by human experts. *Clinical Neurophysiology 114* (2003), pp. 2156–2164.

[194] Wilson, S., Scheuerb, M., Emersonc, R., and Gabord, A. Seizure detection: evaluation of the Reveal algorithm. *Clinical Neurophysiology 115*, 10 (2004), pp. 2280–2291.

[195] Winterhalder, M., Maiwald, T., Voss, H., Aschenbrenner-Scheibe, R., Timmer, J., and Schulze-Bonhage, A. The seizure prediction characteristic: A general framework to assess and compare seizure prediction methods. *Epilepsy and Behavior 4* (2003), pp. 318–325.

[196] Wong, S., Gardner, A. B., Krieger, A. M., and Litt, B. A stochastic framework for evaluating seizure prediction algorithms using hidden markov models. *Journal of Neurophysiology 97* (2007), pp. 2525–2532.

[197] Worrell, G., Cranstoun, S., Echauz, J., and Litt, B. Evidence for self-organised criticality in human epileptic hippocampus. *Neurophysiology, Basic and Clinical 13*, 16 (2002), pp. 2017–2021.

[198] Yaylali, I., Kocak, H., and Jayakar, P. Detection of seizures from small samples using nonlinear dynamic system theory. *IEEE Transactions on Biomedical Engineering 43*, 7 (1996), pp. 743–751.

[199] Zhang, Z., Kawabata, H., and Liu, Z. Electroencephalogram analysis using fast wavelet transform. *Computers in Biology and Medicine 16*, 2 (2001), pp. 429–440.

Index

A

Anti-epileptic drugs (AEDs), 18
Artificial neural networks (ANN)
  activation function, 158
  adaptiveness, 162
  backpropogation, 162
  description, 158
  error-correcting, 162
  inputs, 158
  integrators, 158
  learning algorithms, 161
  network configurations, 160–161
  non-linearity and, 160
  overview, 158
  self-organizing maps, 161, 162, 163, 185
  SVMs, *versus,* 164, 184, 188–189
  testing, 185, 188
  training, 162, 185

B

Black box models, 266
Brain. *See also* Head, human; *specific brain areas*
  behavior, stochastic, 35
  chemistry, complexity of, 12
  complexity of, 1
  conductivity, 66
  dynamic model, 29–30
  functions, 3
  mammalian, 2–3
  mathematical model of, 28–29
  models of, 216–217
  neural networks, 2
  signals from, 1
  spatial scales of (*see* Spatial scales)
  tissue capacitance, 63, 65

C

Caton, Richard, 20
Central nervous system, 2
Cerebral cortex, 2, 4*f*. *See also* *specific lobes*
Cerebrospinal fluid, 60–61
  conductivity of, 66
  spatial filter, acting as, 77
Classifiers, EEG
  artificial neural networks (ANN) (*see* Artificial neural networks (ANN))
  association rules, 157–158, 162, 167
  definition, 156
  evaluation, 174
  feature extraction, 185
  performance, relationship between, 157
  role of, 184
  support vector machines (SVM) (*see* Support vector machines (SVM))
CNet algorithm, 194, 196, 201, 203
Coefficient of variation (COV), 106–107
Competitive learning, 161
Conductivity, 84
Cortex
  behavior of, modeling, 234
  cortical folds, 56
  current dipole in, 55
  intra-cranial EEG, 72
  macro-columns, 39
  mesosources, 58–59
  mini-columns, 39
  neocortex, 54
  neurons in, 39

pyramidal cells in, 54–55
spatial filtering, 79
Cortico-cortical projections, 10–11
Criticality, 269
Current dipoles, 47, 50, 52
cortex, in, 55
synchronicity of, 58

D
Deep brain stimulation, 20
Dydadic sampling, 133
Dynamic models of epilepsy, 32–33
overview, 215–216

E
Electric current, 84
Electric dipoles, 45
Electric fields, 42, 44, 84
Electric potential, 44, 45, 84
expression of, 46
Electrical stimulation treatment
deep brain stimulation, 20
increasing attention toward, 19
termination of seizures via, 252
vagal nerve stimulation (*see* Vagal nerve stimulation)
Electroencephalogram (EEG)
amplitude, 175
artifacts, 24, 27–28, 95, 129
background state, 27
brain conductivity, 66
cerebrospinal fluid conductivity, 66
classifiers (*see* Classifiers, EEG)
cortical folds, modeling of, 56
cortical recordings, 72
dipole synchronicity, 58–59, 82
dynamic models, 83
dynamics, 37, 38, 68, 79–80
electric potential (*see* Electric potential)
epileptic (*see* Epileptic EEGs)
expert system (*see* Expert systems, EEG)
expert-knowledge rejection, 27–28
frequency domain analysis (*see* Frequency domain analysis)
inter-seizure epileptiform, 26, 27
intra-cranial EEG, 21, 39, 72, 82
limits of, 83, 254–255
low resolution, 139
macro-scopic models, 31–32, 35
measurement error, 38
measurement system of, 1, 20–21, 35, 37–38, 68, 71
micro-scopic models, 30–31, 35
noisy measurement, 105
normal, 22, 24, 241, 243
origins, 20
oscillations, 175, 239
pre-seizure epileptiform, 27
predictive, 27, 28
referencing electrodes (*see* Referencing electrodes)
relative voltage, 94
scalp conductivity, 67
scalp EEG, 21, 22, 39, 65, 82, 305–306
scalp recordings, 72, 73
seizure detection (*see* Seizure detection)
seizures, during, 24–26, 27, 243 (*see also* Epilepsy; Seizures)
signal processing (*see* Signal processing, EEG)
skull conductivity, 67
sleep, during, 24
spatial information, 27
spatial integration, 53–54
spatial resolution of signals, 83
stable regions, 239
sub-systems, rhythm of, 31
synchronization, 175
temporal filtering, 79
temporal integration, 58–59
time domain analysis (*see* Time domain analysis)

time-frequency analysis (*see* Time-frequency analysis)
typical readings, 305–306
usefulness, 27, 305–306
volume conductor, 38, 40, 41, 49–50, 63, 84–86
waveforms, 82
Entropy
defining, 142–143
dynamical systems, as part of, 144–145
general, of a discrete probability distribution, 143–144
Komolgorov-Sinai entropy, 145
Shannon entropy, 146–147
Epilepsy. *See also* Seizures
average activity model, 216
continuous, 17
deterministic models, 217
diagnosis, 17
dynamic models of (*see* Dynamic models of epilepsy)
EEGs of (*see* Epileptic EEGs)
electrical stimulation treatment (*see* Electrical stimulation treatment)
focal epilepsy (*see* Focal epilepsy)
generalization, 33
initiation, 33
large-scale phenomenon of, 32–33
mysteries of, 250
neurons, behavior of, 80–81
non-focal epilepsy (*see* Non-focal epilepsy)
stochastic models, 217
stress, applying, for diagnosis, 17
surgical resection for treatment (*see* Surgical resection)
symptoms, 32
Epileptic EEGs
asynchronous activities, 80
overview, 80
spatial scales, 80
synchronous activities, 80
Equivalent current dipoles, 47–48
Excitatory post-synpatic potential (EPSP), 8, 21
Expert systems, EEG, 155–156
environmental context, 167
overview, 166–167
patient-specific algorithms, 170
processing decisions, 167, 169
spatio-temporal context, 169–170

F

Focal epilepsy
complex seizures, 16
generalization, 15–16
histology, 15
inhibitory brain activity, 15
initiation, 14, 15
overview, 14–15
progressive nature of, 14–15
pyramidal neurons in, 32–33
simple seizures, 16
Fourier, Jean Baptiste Joseph, 121
Fractals, 269–270
Frequency domain analysis. *See also* Time-frequency analysis
basis functions, 120–122
Fourier transform, 120–122
inverse Fourier transform, 122–123
Nyquist frequency, 123
overview, 120
power spectral densities (PSD), 123–124, 126, 129
sleep EEGs, 129
Frontal lobe, 2
Functional magnetic resonance imaging (fMRI), 39

G

Gabor atom density, 150–152
Glial cells, 5
Gray box models, 264, 285

Gray matter, 3, 60
conductivity of, 66
isotropy of, 66
Green's function, 48, 50
Gyri, 3

H
Hanning windows, 103, 122
Head, human
four-sphere model, 61*f*
layers of, 60
model of, 60–61
skull (*see* Skull)
spherical harmonics, 61, 63
Hodgkin-Huxley model, 221, 223
Homogenous media
conductivity of, 48–49
overview, 48
permittivity of, 48
Hurst exponent, 273

I
Inhibitory post-synpatic potential (IPSP), 8, 21
Inhomogenous media
geometry of, 49–50
overview, 49
Inter-event probability histograms, 291, 294, 297, 299–300

K
Komolgorov-Sinai entropy, 145

L
Learning algorithms, 161
Linear systems, 237
Long-range dependency, 269, 272, 273
auto-correlation, 278–279
datasets, 274–275
estimating, 274
fano factor, 279
Gaussian distribution, 275–276
normal distribution, 274
power-spectral density, 279
probability distribution, 274
robustness, 282–283
scale invariance, 276, 278, 286
simulations, 281
wavelet order selection, 286
Lyapunov exponents, 141–142

M
Macro-scopic potentials, 48
Magnetic fields, 42
Magnetic resonance imaging (MRI), 39
Magnetoencephalogram (MEG), 39
Monitor algorithm, 191, 193–194, 203
Muscle artifacts (EMG), 95

N
Neocortex, 54
Neural models
biologically inspired, 224–227
conductance-based IF models, 228, 230
cortico-cortical connectivity, 249
experimental evidence, availability of, 220
guidelines, 219
Hodgkin-Huxley model (*see* Hodgkin-Huxley model)
integrate-and-fire models, 221, 223
limitations of, 254–255, 256
macro-scopic models, 247–249
meso-scopic phenomenological models, 230–231, 236, 245–246, 256
micro-scopic statistical models, 223–224, 246–247
networks of neurons, parameters, 221–222
parameters, 219
scale of change, 219–220
single neurons, parameters, 221
sub-systems, 222–223
unified models, 246
variables, 219

Neuromodulators, 12–13, 220. *See also* Neurotransmitters
Neurons, 2
  axons, 5
  chemical currents, 5–6
  dendrites, 5
  epilepsy, behavior in, 80–81
  firing rate, 231
  inputs, 227
  models of (*see* Neural models)
  pyramidal (*see* Pyramidal neurons)
  recordings of, on EEG, 50
  seizures, behavior during, 14
  soma, 5
  synaptic terminals, 5, 6, 8
Neurotransmitters, 6, 8
  concentration of, 220
  dynamics of, 220
  role of, 12
  speed of, 12–13
Non-focal epilepsy
  bifurcation point, 16
  generalization, 16, 17
  initiation, 16
  overview, 16
  spread, 16
Non-linear analysis
  artificial neural networks (ANN), in (*see* Artificial neural networks (ANN))
  complexity, 147
  data sequences, 136
  dimensions, system, 139–141
  embedding theory, 137–138
  Lyapunov exponents, 141–142, 148
  overview, 136
  predictability, 147, 148
  randomness, 147, 148

O

Occipital lobe, 2
Onset seizure detectors
  correlation dimension, 208
  cross-correlation (XCORR), 207
  feature extraction, 204, 207, 208, 212
  overview, 204
  power spectral density (PSD), 207–208, 212–213
  warning times, 212–213
  wavelet analysis (WAV), 208
Orthogonal wavelets, 133

P

Parietal lobe, 2
Patient un-specific seizure detectors
  CNet algorithm, 194, 196, 201, 203
  Monitor algorithm, 191, 193–194, 203
  overview, 190–191
  performance analyses, 202–203
  Reveal algorithm, 197, 199, 203
  Saab algorithm, 199, 201–202, 203
Permittivity, 84
Power-laws, 269, 271
Predicting seizures, 150
  black box models, 266
  data measurement, relationship between, 265–266
  difficulties of, 263, 307
  gray box models, 264
  long-range dependency (*see* Long-range dependency)
  memory, relationship between, 300–302
  overview, 263
  personal (patient) predictions, 266–267, 268
  physical models, 266
  signature recognition, 268–269
  simulations, 281
  stochastic models, 264
Pyramidal neurons, 5, 8–9
  cortex, in, 54
  focal-point epilepsy, in, 32–33

R
RC circuits, capacitance of, 86–87
Receiver operating characteristic (ROC) curve, 183, 193
Referencing electrodes, 130
  average reference, 74
  bipolar referencing, 74–75, 76–77
  linked ears/mastoids, 74
  noise levels, 74–75
  overview, 73–74
Resistance, 84
Resistivity capacitance, 84
Reveal algorithm, 197, 199, 203

S
Saab algorithm, 199, 201–202, 203
Scalar notation, 42–43
Scale-invariance, 269, 270, 271, 272, 276, 278, 286
Seizure detection, 89–90
  amplitude, difficulties with, 175
  database group 1, 178
  database group 2, 178
  database group 3, 178–179
  database group 4, 179–180
  database size, 173
  difficulties inherent in, 175–176
  false negative detections, 181
  false positive detections, 181, 182, 183, 203
  metrics for comparison of algorithms, 174
  onset delay, 181
  onset detectors (*see* Onset seizure detectors)
  oscillations, difficulties with, 175
  overview, 173
  patient un-specific seizure detectors (*see* Patient un-specific seizure detectors)
  positive detections, 180–181
  receiver operating characteristic (ROC) curve, 183, 193
  scalp data, 176, 178
  specificity, 182
  synchronization, difficulties with, 175
  true positive detections, 180–181, 182, 183
Seizure frequency dataset
  analysis, 291
  integrity of, 288
  inter-event probability histograms (*see* Inter-event probability histograms)
  overview, 287–288
  sequences, 288
Seizures. *See also* Epilepsy
  classifications of, 14, 89–90
  complex, 16
  continuous, 17
  detection of (*see* Seizure detection)
  EEG appearance (*see* Electroencephalogram (EEG))
  expert system classification (*see* Expert systems)
  frequency dataset (*see* Seizure frequency dataset)
  initiation, 250–251
  low dimensionality during, 138
  mysteries of, 250
  neuron behavior in, 14
  post-seizure (offset) period, 14
  predicting (*see* Predicting seizures)
  simple, 16
  termination via electrical stimulation, 252
Self-organized criticality, 270–271
Self-organizing maps, 161, 162, 163, 185, 269, 270–271
Self-similarity, 269
Sensory relay network, 233
Shannon entropy, 146–147
Signal processing, EEG, 27, 83
  auto-correlations, 111, 113, 115
  averages *versus* instances, 96–99

basis functions, 120–122, 130
classification, 90
determinism, 103
expert system, 90
feature extraction, 90, 96
frequency domain analysis (*see* Frequency domain analysis)
instantaneous amplitude, 106
linearity, 103
mathematical representation of, 91–93
noise, 99
non-linearity, 103–104
normalization, 94–95, 104, 105, 106–107, 123–124
overview, 89–90
periodicity, 108, 110–111
phase locking, 118
preprocessing, 90, 93–95
purpose of, 89
raw EEG signal, 107–108
sleep/awake morphologies, 90–91
stationarity, 100–101, 110
stochasticity, 104
synchonization, 115, 118–120
time domain analysis (*see* Time domain analysis)
variance, 106–107
windowing, 101–103
Skull, 61
conductivity of, 67
effect on EEGs, 130
spatial filtering of, 75–77, 78
spherical harmonics, 75–76
Slack variables, 164
Spatial scales, 3
macro-scopic, 3–4, 8–10
meso-scopic, 3, 8–10
micro-scopic, 3, 5–6, 8
Status epilepticus, 17
Stochastic models, 264
Stochastic processes, 33–34
Sucli, 3
Support vector machines (SVM)
ANNs, *versus,* 164, 184, 188–189
description, 163
overview, 163–165
slack variables, 164
testing, 188
training, 188
Surgical resection, 18–19
Synaptic transmission, 6, 8
excitatory post-synpatic potential (EPSP), 8
gating mechanisms, 52–53
inhibitory post-synpatic potential (IPSP), 8

T

Temporal lobe, 2
Temporal scales, 3
Thalamic reticular nucleus, 233
Thalamo-cortical projections, 11–12
Thalamus
behavior of, 234
projections, 11–12
regulation of information to cortex, 12
role of, 12
Thalmus, 2
Time domain analysis
overview, 105–106
signal amplitude, 106
standard deviation, 106
variance, 106–107
Time-frequency analysis
overview, 131
temporal evolution, 133
wavelet coefficients, 134–135

V

Vagal nerve stimulation
effectiveness, 19
prevalence as treatment choice, 19
side effects, 19–20
Vector notation, 42–43
Voxels, 50

W

White matter, 11, 60
    conductivity of, 66
    isotropy of, 66