To my parents in Taiwan.
ACKNOWLEDGMENTS

I would like to thank Dr. Panagote M. Pardalos, Dr. Basim M. Uthman, Dr. J. Chris Sackellares, Dr. Deng-Shan Shiau, and Dr. W.A. Chaovalitwongse for their guidance and support. During the past five years, their energy and enthusiasm have motivated me to become a better researcher. I would also like to thank Dr. Van Oostrom, Dr. Paul Carney and Dr. Steven Roper for serving on my supervisory committee and for their insightful comments and suggestions.

I have been very lucky to have the chance to work with many excellent people in the Brain Dynamics Laboratory (B.D.L) and Center for Applied Optimization (C.A.O). I would like to thank Linda Dance, Wichai Suharitdamrong, Dr. Sandeep Nair and Dr. Michael Bewernitz from B.D.L and Dr. Altannar Chinchuluun, Dr. Onur Seref, Ashwin Arulselvan, Nikita Boyko, Dr. Alla Kammerdiner, Oleg Shylo, Dr. Vitaliy Yatsenko and Petros Xanthopoulos in C.A.O for their valuable friendship and kindness assistance.

I thank the staff in the Department of Neurology at the Gainesville VA Medical Center including David Juras and Scott Bearden. I thank Joy Mitchell from North Florida Foundation for Research and Education, Inc. for providing financial support during later part of my Ph.D. study. I also like to acknowledge the financial support provided by National Institutes of Health (NIH), Optima Neuroscience Inc. and North Florida Foundation for Research and Education, Inc. (NFFRE). There are many friends I have made at the University of Florida who gave me their help and support but were not mention here, I would like to take this opportunity to thank all of them for being with me throughout this wonderful journey.

Lastly, I wish to thank my lovely family in my home country Taiwan my father Ming-Chih Liu, my mother Ching-Wei Lin, my elder sister Tsai-Lin, Liu and my younger brother Chang-Hao, Liu for their selfless sacrifice and constant love.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>4</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>8</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>9</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>12</td>
</tr>
<tr>
<td><strong>CHAPTER</strong></td>
<td></td>
</tr>
<tr>
<td>1 OVERVIEW OF EPILEPSY</td>
<td>14</td>
</tr>
<tr>
<td>1.1 Introduction</td>
<td>14</td>
</tr>
<tr>
<td>1.2 Causes of Seizure</td>
<td>15</td>
</tr>
<tr>
<td>1.3 Classification of Epileptic Seizure (ICES 1981 revision)</td>
<td>16</td>
</tr>
<tr>
<td>1.3.1 Partial Onset Seizures</td>
<td>16</td>
</tr>
<tr>
<td>1.3.2 Generalized Seizures</td>
<td>19</td>
</tr>
<tr>
<td>1.3.3 Unclassified Seizures</td>
<td>20</td>
</tr>
<tr>
<td>1.4 Treatment for Epilepsy</td>
<td>20</td>
</tr>
<tr>
<td>1.4.1 Pharmacological Treatment</td>
<td>20</td>
</tr>
<tr>
<td>1.4.2 Surgical Section</td>
<td>21</td>
</tr>
<tr>
<td>1.4.3 Neurostimulator Implant</td>
<td>21</td>
</tr>
<tr>
<td>1.5 Neuron States and Membrane Potentials</td>
<td>23</td>
</tr>
<tr>
<td>1.5.1 Neuron States</td>
<td>23</td>
</tr>
<tr>
<td>1.5.2 Membrane Potentials</td>
<td>25</td>
</tr>
<tr>
<td>1.6 Recording Electric Brain Activity</td>
<td>27</td>
</tr>
<tr>
<td>1.6.1 Scalp EEG Recording</td>
<td>29</td>
</tr>
<tr>
<td>1.6.2 Subdural EEG Recording</td>
<td>29</td>
</tr>
<tr>
<td>1.6.3 Depth EEG Recording</td>
<td>30</td>
</tr>
<tr>
<td>1.7 Conclusions and Remarks</td>
<td>31</td>
</tr>
<tr>
<td>2 EPILEPSY AND NONLINEAR DYNAMICS</td>
<td>33</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>33</td>
</tr>
<tr>
<td>2.2 Dynamical Systems and State Space</td>
<td>34</td>
</tr>
<tr>
<td>2.3 Fractal Dimension</td>
<td>45</td>
</tr>
<tr>
<td>2.3.1 Correlation Dimension</td>
<td>46</td>
</tr>
<tr>
<td>2.3.2 Capacity Dimension</td>
<td>46</td>
</tr>
<tr>
<td>2.3.3 Information Dimension</td>
<td>47</td>
</tr>
<tr>
<td>2.4 State Space Reconstruction</td>
<td>47</td>
</tr>
<tr>
<td>2.5 Lyapunov Exponents</td>
<td>48</td>
</tr>
<tr>
<td>2.6 Phase/Angular Frequency</td>
<td>50</td>
</tr>
<tr>
<td>2.7 Approximate Entropy</td>
<td>51</td>
</tr>
<tr>
<td>2.8 Dynamical Support Vector Machine (D-SVM)</td>
<td>53</td>
</tr>
</tbody>
</table>
3 QUANTITATIVE COMPLEXITY ANALYSIS IN MULTI-CHANNEL INTRACRANIAL EEG RECORDINGS FROM EPILEPSY BRAIN ............................................. 60

3.1 Introduction ............................................................. 60
3.2 Patient and EEG Data Information ................................. 61
3.3 Proper Time Delay ...................................................... 63
3.4 The Minimum Embedding Dimension ............................... 65
3.5 Data Analysis ............................................................ 66
3.6 Conclusions .............................................................. 67

4 DISTINGUISHING INDEPENDENT BI-TEMPORAL FROM UNILATERAL ONSET IN EPILEPTIC PATIENTS BY THE ANALYSIS OF NONLINEAR CHARACTERISTICS OF EEG SIGNALS ........................................ 72

4.1 Introduction ............................................................. 72
4.2 Materials and Methods ............................................... 75
4.2.1 EEG Description .................................................... 75
4.2.2 Non-Stationarity ...................................................... 76
4.2.3 Surrogate Data Technique ........................................ 77
4.2.4 Estimation of Maximum Lyapunov Exponent ................. 78
4.2.5 Paired t-Test ........................................................ 79
4.3 Results ................................................................. 80
4.4 Discussion .............................................................. 84

5 OPTIMIZATION AND DATA MINING TECHNIQUES FOR THE SCREENING OF EPILEPTIC PATIENTS ................................................. 91

5.1 Introduction ............................................................. 91
5.2 EEG Data Information ............................................... 93
5.3 Independent Component Analysis ................................... 94
5.4 Dynamical Features Extraction ....................................... 95
5.4.1 Estimation of Maximum Lyapunov Exponent ................. 95
5.4.2 Phase/Angular Frequency ......................................... 96
5.4.3 Approximate Entropy .............................................. 97
5.5 Dynamical Support Vector Machine ................................. 98
5.6 Connectivity Support Vector Machine ............................... 100
5.7 Training and Testing: Cross Validation ............................ 102
5.8 Results and Discussions .............................................. 102
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>59</td>
</tr>
<tr>
<td>2-2</td>
<td>59</td>
</tr>
<tr>
<td>3-1</td>
<td>62</td>
</tr>
<tr>
<td>4-1</td>
<td>76</td>
</tr>
<tr>
<td>5-1</td>
<td>94</td>
</tr>
<tr>
<td>5-2</td>
<td>103</td>
</tr>
<tr>
<td>7-1</td>
<td>121</td>
</tr>
<tr>
<td>8-1</td>
<td>140</td>
</tr>
<tr>
<td>8-2</td>
<td>146</td>
</tr>
<tr>
<td>8-3</td>
<td>150</td>
</tr>
<tr>
<td>8-4</td>
<td>151</td>
</tr>
<tr>
<td>8-5</td>
<td>152</td>
</tr>
<tr>
<td>8-6</td>
<td>153</td>
</tr>
<tr>
<td>8-7</td>
<td>154</td>
</tr>
<tr>
<td>8-8</td>
<td>155</td>
</tr>
<tr>
<td>8-9</td>
<td>156</td>
</tr>
<tr>
<td>8-10</td>
<td>157</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1-1</td>
<td>Vagus nerve stimulation pulse generator</td>
</tr>
<tr>
<td>1-2</td>
<td>Vagus nerve stimulation electrode</td>
</tr>
<tr>
<td>1-3</td>
<td>Cortical nerve cell and structure of connections</td>
</tr>
<tr>
<td>1-4</td>
<td>Membrane potentials</td>
</tr>
<tr>
<td>1-5</td>
<td>Electrical potentials</td>
</tr>
<tr>
<td>1-6</td>
<td>EEG recording acquired by Hans Berger in 1929</td>
</tr>
<tr>
<td>1-7</td>
<td>Basic EEG patterns</td>
</tr>
<tr>
<td>1-8</td>
<td>International 10-20 electrode placement</td>
</tr>
<tr>
<td>1-9</td>
<td>Subdural electrode placement</td>
</tr>
<tr>
<td>1-10</td>
<td>Depth electrode placement</td>
</tr>
<tr>
<td>2-1</td>
<td>Rössler attractor</td>
</tr>
<tr>
<td>2-2</td>
<td>X component of Rössler system</td>
</tr>
<tr>
<td>2-3</td>
<td>Y component of Rössler system</td>
</tr>
<tr>
<td>2-4</td>
<td>Z component of Rössler system</td>
</tr>
<tr>
<td>2-5</td>
<td>Lorenz system</td>
</tr>
<tr>
<td>2-6</td>
<td>X component of Lorenz system</td>
</tr>
<tr>
<td>2-7</td>
<td>Y component of Lorenz system</td>
</tr>
<tr>
<td>2-8</td>
<td>Z component of Lorenz system</td>
</tr>
<tr>
<td>2-9</td>
<td>Estimation of Lyapunov exponent ($L_{max}$)</td>
</tr>
<tr>
<td>2-10</td>
<td>Temporal evolution of $STL_{max}$</td>
</tr>
<tr>
<td>2-11</td>
<td>Phase/Angular frequency of Lyapunov exponent ($\Omega_{max}$)</td>
</tr>
<tr>
<td>2-12</td>
<td>Three dimension plot for entropy, angular frequency and $L_{max}$ during interictal state</td>
</tr>
<tr>
<td>2-13</td>
<td>Three dimension plot for entropy, angular frequency and $L_{max}$ during preictal state</td>
</tr>
<tr>
<td>3-1</td>
<td>Electrode placement</td>
</tr>
</tbody>
</table>
7-6 Maximum clique algorithm ........................................... 131
7-7 Electrode selection using the maximum clique algorithm for Case 1 ........ 132
7-8 Electrode selection using the maximum clique algorithm for Case 2 ........ 133
7-9 Electrode selection using the maximum clique algorithm for Case 3 ........ 133
8-1 Nonlinear interdependences for electrode FP1 ........................... 147
8-2 Pairwise mutual information between for all electrodes- before v.s. after treatment 149
The basic mechanisms of epileptogenesis remain unclear and investigators agree that no single mechanism underlies the epileptiform activity. Different forms of epilepsy are probably initiated by different mechanisms. The quantification for preictal dynamic changes among different brain cortical regions have been shown to yield important information in understanding the spatio-temporal epileptogenic phenomena in both humans and animal models.

In the first part of this study, methods developed from nonlinear dynamics are used for detecting the preictal transitions. Dynamical changes of the brain, from complex to less complex spatio-temporal states, during preictal transitions were detected in intracranial electroencephalogram (EEG) recordings acquired from patients with intractable mesial temporal lobe epilepsy (MTLE). The detection performance was further enhanced by the dynamics support vector machine (D-SVM) and a maximum clique clustering framework. These methods were developed from optimization theory and data mining techniques by utilizing dynamic features of EEG. The quantitative complexity analysis in multi-channel intracranial EEG recordings is also presented. The findings suggest that it is possible to distinguish epilepsy patients with independent bi-temporal seizure onset zones (BTSOZ) from those with unilateral seizure onset zone (ULSOZ). Furthermore, for the ULSOZ patients, it is also possible to identify the location of the seizure onset zone in the brain. Improving clinician’s certainty in identifying the
epileptogenic focus will increase the chances for better outcome of epilepsy surgery in patient with intractable MTLE.

Recent advances in nonlinear dynamics performed on EEG recordings have shown the ability to characterize changes in synchronization structure and nonlinear interdependence among different brain cortical regions. Although these changes in cortical networks are rapid and often subtle, they may convey new and valuable information that are related to the state of the brain and the effect of therapeutic interventions. Traditionally, clinical observations evaluating the number of seizures during a given period of time have been gold standard for estimating the efficacy of medical treatment in epilepsy. EEG recordings are only used as a supplemental tool in clinical evaluations. In the later part of this study, a connectivity support vector machine (C-SVM) is developed for differentiating patients with epilepsy that are seizure free from those that are not. To that end, a quantitative outcome measure using EEG recordings acquired before and after anti-epileptic drug treatment is introduced. Our results indicate that connectivity and synchronization between different cortical regions at higher order EEG properties change with drug therapy. These changes could provide a new insight for developing a novel surrogate outcome measure for patients with epilepsy when clinical observations could potentially fail to detect a significant difference.
CHAPTER 1
OVERVIEW OF EPILEPSY

1.1 Introduction

Epilepsy was conceptualized during the Vedic period and described as “apasmara” (i.e., loss of consciousness) in the Ayurvedic literature of Charaka Samhita. Later in time, another ancient culture Babylon documented the causes of epilepsy were associated with spiritual matters and related to evil. Furthermore, this viewpoint was taken by Greek physician Hippocrates and described epilepsy as “the sacred disease”. About 1% (i.e., about 50 million of people) of the population is afflicted, making epilepsy the most common neurological disorder after stroke throughout the worldwide. The prevalence of epilepsy is generally higher in developing countries than in industrialized nations. Studies in industrialized countries suggest an annual incidence of epilepsy of approximately 100 per 100,000 of the general population [1]. Epilepsy occurs in all ages, the highest incidences occur in infants and in elderly [2; 3]. The cause of epilepsy resulted from a large number of factors, including head injury, stroke, brain tumor, central nervous system infections, developmental anomalies and hypoxia–ischemia.

The hallmark of epilepsy is recurrent seizures. Seizures are mediated by abrupt development of rhythmic firing of large groups of neurons in the cerebral cortex. These rhythmic discharges may begin locally in certain region of one cerebral hemisphere (partial seizures) or begin simultaneously in both cerebral hemispheres (generalized seizure). Partial onset seizures may remain confined within a particular region of the brain and cause no change in consciousness and relatively mild cognitive, sensory, motor or autonomic symptoms (simple partial seizures) or may spread to cause impaired consciousness during the seizure (complex partial seizure) along with a variety of motor symptoms, such as sudden and brief localized body jerks to generalized tonic-clonic activity (secondary generalized seizures). Primarily generalized seizures involve both hemispheres of the brain, cause altered in loss of consciousness during the occurrence
of the seizure with or without convulsion. In partial onset of epilepsy, close to 30% of patients remain refractory to medical treatment. In primary generalized epilepsy the great majority of patients responded to medical treatment. In either type of epilepsy uncontrolled seizures result in impacts on patients’ social, educational and vocational aspects of life. These restrictions have substantial effects on their quality of life, as well as their families and loved ones.

1.2 Causes of Seizure

Many abnormalities of the nervous system can result in seizure activities. Seizures can also occur in the normal nervous system when the metabolic balance is disturbed. The cause (etiology) of epilepsy may be not clearly known (idiopathic) or related to a particular disease. About 35% of all cases of seizure have no clearly definable causes, the followings are cause factors that might cause seizures:

1. **Genetic factors**: Some persons may have a genetic predisposition to the development of seizures. There is also an increased incidence of epilepsy in relatives of those with a seizure disorder.

2. **Head injury**: Seizures may develop at or around the time of injury.

3. **Stroke**: Seizures can occur at the time of a stroke or many years later. They may occur with strokes that result in lack of blood flow to the brain or with those that involve bleeding (hemorrhage) into or around the brain.

4. **Metabolic disturbances**: There are various metabolic substances in the body sometimes result in seizures. Some possible instances are given below:
   
   (a) Altered levels of sodium, calcium, or magnesium (electrolyte imbalance)
   (b) Kidney failure with increased urea in the blood (uremia) or changes that occur with kidney dialysis
   (c) Low blood sugar (hypoglycemia) or elevated blood sugar (hyperglycemia)
   (d) Lowered oxygen level in the brain (hypoxia)
   (e) Severe liver disease (hepatic failure) and elevation of associated toxins
5. **Infections**: Infections of the nervous system may result in seizure activity. These include infection of the covering of the brain and the spinal fluid (meningitis), infection of the brain (encephalitis), and human immunodeficiency virus (HIV) and related infections.

6. **Tumors**: Cancerous (malignant) and benign brain tumors may be associated with seizures. The location of the lesion influences the risk.

7. **Cerebral palsy**: Epilepsy is often a symptom of cerebral palsy, which results from lack of oxygen, infection, or trauma during birth or infancy.

8. **Febrile seizures**: Febrile seizures occur in small children and are caused by high fever.

1.3 **Classification of Epileptic Seizure (ICES 1981 revision)**

The ICES (Commission on Classification and Terminology of the ILAE, 1981), The ICES recognizes 18 subclasses of simple focal seizures belonging to four groups, four modalities of complex focal seizures, and three types of secondarily generalized seizures.

1.3.1 **Partial Onset Seizures**

Partial onset seizures are those in which, in general, the first clinical and EEG changes indicate initial activation of a limited group of neurons. A partial seizure is classified primarily on the basis of whether or not consciousness is impaired during the attack. When consciousness is not impaired, the seizure is classified as a simple focal seizure. When consciousness is impaired, the seizure is classified as complex focal seizures. In patients with impaired consciousness, aberrations of behavior (automatisms) may occur. A partial onset seizure may not terminate, but instead progress to a generalized motor seizure. Impaired consciousness is defined as the inability to respond normally to exogenous stimuli by virtue of altered awareness and/or responsiveness. There is considerable evidence that partial onset seizures usually have unilateral hemispheric involvement and only rarely have bilateral hemispheric involvement; complex partial onset
seizures, however, frequently have bilateral hemispheric involvement. Partial onset seizure can be classified into one of the following three groups:

1. **Simple partial onset seizures (consciousness not impaired)**

   The EEG features for simple partial onset seizures are local contralateral discharge starting over the corresponding area of cortical representation and the significant feature for the background EEG is local contralateral discharge. Simple partial onset seizures may have the following clinical features:

   (a) With motor signs
      
      i. Focal motor without march
      
      ii. Focal motor with march (Jacksonian)
      
      iii. Versive
      
      iv. Postural
      
      v. Phonatory (Vocalization or arrest of speech)

   (b) With somatosensory or special sensory symptoms (simple hallucinations, e.g., tingling, light flashes, buzzing)
      
      i. Somatosensory
      
      ii. Visual
      
      iii. Auditory
      
      iv. Olfactory
      
      v. Gustatory
      
      vi. Vertiginous

   (c) With autonomic symptoms or signs (including epigastric sensation, pallor, sweating, flushing, piloerection and pupillary dilatation)

   (d) With psychic symptoms (disturbance of higher cerebral function). These symptoms rarely occur without impairment of consciousness and are much more commonly experienced as complex focal seizures
      
      i. Dysphasic
2. Complex partial seizures (with impairment of consciousness)

The EEG for complex partial seizures usually show unilateral or, frequently, bilateral discharge, diffuse or focal in temporal or frontotemporal regions as for the background activity are unilateral or bilateral generally asynchronous focus; usually in the temporal or frontal regions.

(a) With impairment of consciousness at onset
   i. With impairment of consciousness only
   ii. With automatisms

(b) Simple partial onset followed by impairment consciousness
   i. With simple focal features followed by impaired consciousness
   ii. With automatisms

3. Partial seizures evolving to generalized tonic-clonic convulsions (GTC)

Seizures

These seizure may be generalized tonic-clonic, tonic, or clonic. The EEG show unilateral or, frequently, bilateral discharge become secondarily and rapidly generalized.

(a) Simple partial seizure evolving to generalized seizures

(b) Complex partial seizure evolving to generalized seizures

(c) Simple partial seizure evolving to complex focal seizures evolving to generalized seizures
1.3.2 Generalized Seizures

Generalized seizures are those in which the first clinical changes indicate involvement of both hemispheres. Consciousness may be impaired and this impairment may be the initial manifestation. Generalized seizures can be classified into the following groups:

1. **Absence seizures (Petit mal seizures)**

   The hallmark of the absence seizure is sudden onset, interruption of ongoing activities, the attack lasts from a few seconds to half minute and evaporates as rapidly as it commenced.

   (a) **Typical Absence seizures**: A typical absence seizure is characterized by generalized and bilaterally synchronous 3/sec spike and wave discharges (SWD) few seconds long to half a minute in duration. Like in most generalized epilepsies, SWD in absence seizures is maximal over the fronto central midline brain region and may start at a rate of around 4/sec, quickly slow down to 3-3.5/sec, and during the final phase of the absence, slow to about 2.5/sec. The EEG background activity is usually normal although paroxysmal activity (spike or spike-and-slow-wave complexes) may occur.

   (b) **Atypical absences**: The EEG configuration for atypical absence is more heterogeneous; may include irregular spike-and-slow-wave complexes, fast activity or other paroxysmal activity. The EEG abnormalities are bilateral but often irregular and asymmetrical. The EEG background of is usually abnormal paroxysmal activity (such as spike or spike-and-slow-wave complexes) frequently irregular and asymmetrical.

2. **Myoclonic Seizures**

   Myoclonic seizures usually appear as sudden jerks of arms and legs. Myoclonic seizures may last only a short period of time from less than a second for single jerks to a few seconds for repeated jerks.

3. **Clonic Seizures**
Clonic seizures consist of rhythmic jerking movements of the arms and legs, sometimes on both sides of the body.

4. **Tonic Seizures**

Tonic seizures are characterized by facial and truncal muscle spasms, flexion or extension of the upper and lower extremities, and impaired consciousness.

5. **Atonic Seizures**

Atonic seizures is also known as drop attacks, these seizures cause suddenly collapse or fall down during attack. After a few seconds, the consciousness would regain and are able to stand and walk.

6. **Generalized Tonic-clonic (Grand mal seizures)**

Generalized tonic-clonic is the most intense of all types of seizures, these are characterized by a loss of consciousness, body stiffening and shaking, and sometimes tongue biting or loss of bladder control. After the shaking subsides, a period of confusion or sleepiness usually occurs, lasting for a few minutes to a few hours.

1.3.3 **Unclassified Seizures**

The unclassified seizures are seizures which can not be classified into above categories mainly due to lack of information.

1.4 **Treatment for Epilepsy**

1.4.1 **Pharmacological Treatment**

For many years anti-epileptic drugs (AEDs) have been the most common therapeutic intervention for patients with epilepsy. The differences in chemical, pharmacological or clinical profiles among AEDs are considerably large. There exists no clear picture of the mechanism(s) of action for various AEDs. Moreover, AEDs are used to treat seizures that also cannot be defined as homogeneous in its clinical representation, but rather identified as several different syndromes whose fundamental mechanisms are not well defined. The severity of epilepsy assessed by the frequency of seizures during the duration of epilepsy prior to onset of therapeutic intervention. AEDs are taken periodically in fixed doses,
titrated to reach a steady state concentration in the body. The type and prescribed amount of AEDs are appropriately chosen to accomplish the most reduction in seizure frequency with the least accompanied side effects. With advances in pharmaceutical sciences, there are approximately 40% of individuals with epilepsy have seizures that are still refractory to AEDs therapy [4]. The evidence from both experimental and clinical studies which suggests that loss of efficacy of AEDs may develop during their long-term use in a minority of patients [5]. The most common mechanism is an increase in the rate of metabolism of the drug. For instance, many ”first generation” AEDs, including phenobarbital (PB), phenytoin (PHT), and carbamazepine (CBZ), stimulate the production of higher levels of hepatic microsomal enzymes, causing more rapid removal and breakdown of these AEDs from the circulation [6].

1.4.2 Surgical Section

For the patient who does not react to AEDs treatment, individuals may be benefited from epilepsy surgery. Epilepsy surgery has been performed on both children and adults for over 50 years. Seizure-free rate following temporal lobectomy are consistently 65% to 70% in adults [7]. However, studies have also shown that only individual with unilateral onset zone produces significant reductions in seizure frequency with current surgical routines [8].

1.4.3 Neurostimulator Implant

Various modalities of neurostimulation are currently being investigated for adjunctive treatment of intractable epilepsy. Vagus nerve stimulation (VNS) has been approved for the adjunctive treatment of partial onset epilepsy with or without secondary generalization in patients twelve years of age or older. The first animal studies demonstrating the anticonvulsant effect of VNS [9] based on previous observations that VNS blocked interictal EEG spiking [10] or desynchronized EEG in the thalamus or cortex [11; 12] in cats. Zabara extended this idea to demonstrate anticonvulsant activity of VNS in dogs [9]. Later on, McLachlan demonstrated suppression of inter ictal spikes and seizures by
VNS in rats [13]. Furthermore, chronic intermittent VNS has been shown to prevent recurrence of epileptic seizures in monkeys [14]. These observations lead to clinical trials to investigate the feasibility, safety, and efficacy of VNS in human patients. The first human implant was performed in 1988 and the first randomized active control study was performed in 1992. Encouraging results from the first two pilot studies lead to randomized double blind clinical trials that resulted in FDA approval of VNS as adjunctive therapy for intractable epilepsy in July, 1997. Implantation of the NCP system (VNS pulse generator and electrical leads) is normally performed under general anesthesia as an outpatient basis. The procedure usually takes two hours and requires two skin incisions; one in the left upper chest for housing the pulse generator in a pocket under the skin and the other is in left neck area above the collar bone to gain access to the vagus nerve and place two semi-circular electrodes and a neutral tether around the left vagus nerve (Figure 1-1).

Figure 1-1. Vagus nerve stimulation pulse generator
Figure 1-2. Vagus nerve stimulation electrode

Long term follow-up studies showed that prevention of recurrent seizures was maintained and adverse events decreased significantly over time [15; 16]. Positron emission tomography and functional MRI studies showed that VNS activates or increases blood flow in certain areas of the brain such as the thalamus [17; 18]. Cerebrospinal fluid (CSF) was analyzed in 16 subjects before, 3 months after, and 9 months after VNS treatment GABA (total and free) increased in low or high stimulation groups, aspartate marginally decreased and ethanolamine increased in the high stimulation group suggesting an increased inhibitory effect [19]. Krahl et al., suggested that seizure suppression induced by VNS may depend on the release of norepinephrine and they observed that acute or chronic lesions of the ”Locus coeruleus” attenuated VNS-induced seizure suppression [20].

1.5 Neuron States and Membrane Potentials

1.5.1 Neuron States

The birth of Electroencephalogram (EEG) has inspired the attempts to extract the subtle alternations in brain activity. It is known the fluctuations on EEG frequency and voltage arise from spontaneous interactions between excitatory and inhibitory neurons in circuit loops. If a neuron is stimulated, the membrane potential will be altered and this alternation can be classified into two different states Figure 1-3 is taken from Malmivu and Plonse (1993)[21].
1. **Excited state**: A neuron state with a less negative intraneural membrane potential compare to resting state of a neuron. The positive increase in voltage above the normal resting neuronal potential is called the excitatory postsynaptic potential (EPSP), if this potential rise high enough in the positive direction, it will elicit an action potential in the neuron.

2. **Inhibited state**: A neuron state with a more negative intraneural membrane potential compare to resting state of a neuron. An increase in negative beyond the normal resting member potential level is call an inhibitory postsynaptic potential (IPSP).

![Cortical nerve cell and structure of connections](image)

Figure 1-3. Cortical nerve cell and structure of connections

The excitatory neurons excite the target neurons. Excitatory neurons in the central nervous system are often glutamatergic neurons. Neurons in the peripheral nervous system, such as spinal motor neurons that synapse onto muscle cells, often use acetylcholine as their excitatory neurotransmitter. However, this is just a general tendency that may not always be true. It is not the neurotransmitter that decides excitatory or inhibitory action, but rather it is the postsynaptic receptor that is responsible for the
action of the neurotransmitter. Inhibitory neurons inhibit their target neurons. Inhibitory neurons are often interneurons. The output of some brain structures (neostriatum, globus pallidus, cerebellum) are inhibitory. The primary inhibitory neurotransmitters are GABA and glycine. Modulatory neurons evoke more complex effects termed neuron modulation. These neurons use neurotransmitters such as dopamine, acetylcholine, serotonin and others [22]. The effect of summing simultaneous postsynaptic potentials by activating multiple terminals on widely spaced areas of the membrane is called spatial summation and successive presynaptic discharges from a single presynaptic terminal, this type of summation is call temporal summation.

1.5.2 Membrane Potentials

The purposes of sending electrical signals are exchanging information between neurons, it is essential to understand how these signals arise. Figure 1-4 and figure 1-5 is taken from Malmivu and Plonse (1993) [21]. The sources of electrical potential can be categorized in to the following four different categorizes:
1. Resting membrane potential (-40mv—90 mv)
   A single neuron cell has a negative potential, implying that neurons have a means of generating a constant voltage across their membranes when at rest. This voltage is called the resting membrane potential, different type of neuron will have different resting membrane potential ranging from -40mv—90mv.

2. Receptor potentials
   Receptor potentials can be observed when neurons response to external stimuli, such as audio, light, etc.

3. Synaptic potentials
   The mutual interactions between neurons at synapses generates synaptic potentials.

4. Action potentials
   The action potentials have special properties that are different other potentials. First the action potential has all-or-none property, they occur only the when potential reach certain threshold value. Second, the amplitude of an action potential is independent of magnitude of the current used to evoke it; that is, larger current
do not elicit larger the amplitude of the action potential. Therefore, the intensity of a stimulus is encoded in the frequency of action potential rather than in their amplitude. Third, the action potential "travels" along the axon without fading out because the signal is regenerated at each adjacent membrane.

1.6 Recording Electric Brain Activity

The first brain electrical scalp recordings of human was performed by Hans Berger in 1929 [23]. Since then the EEG recordings has been the most common diagnosis tool for epilepsy. EEG measures the electrical activity of the brain. EEG studies are particularly important when neurologic disorders are not accompanied by detectable alteration in brain structure. It is accepted that the neurons in the thalamus play an important role in generating the EEG signals. The synchronicity of the cortical synaptic activity reflects the degree of synchronous firing of the thalamic neurons that are generating the electrical activities. However, the purposes of these electrical activities and EEG oscillations are largely unknown.

![Figure 1-6. EEG recording acquired by Hans Berger in 1929](image)

The configurations of EEG recordings play an important role in determining the normal brain function from abnormal. The most obvious EEG frequencies of an awake, relaxed adult whose eyes are closed is 8-13 Hz also known as the alpha rhythm. The alpha rhythm is recorded best over the parietal and occipital lobes and is known to be associated with decreased levels of attention. When alpha rhythm are presented, subjects commonly report that they feel relaxed and happy. However, people who normally experience more alpha rhythm than usual have not been shown to be psychologically different from those with less. Another important EEG frequencies is the beta rhythm, people are attentive to
When an external stimulus or are thinking hard about something, the alpha rhythm is replaced by lower-amplitude, high-frequency (> 13 Hz). This is beta rhythm oscillations. This transformation is also known as EEG arousal and is associated with the act of paying attention to stimulus even in a dark room with no visual inputs. A transient is an event which clearly stands out against the background. A sharp wave is a transient with 70ms–200ms in duration. A spike wave is a transition with less than 70ms in duration. A spike that follow by a slow wave is called a spike-and-wave complex, which can be seen in patients with typical absence seizure. In the cases having two or more spike occur in sequence forming multiple spike complex call polyspike complex, if they are followed by a slow wave, they are called polyspike-and-wave complex. Spike and sharp waves are
abnormal features in EEG recordings. Figure 1-7 is taken from Malmivu and Plonse (1993) [21].

1.6.1 Scalp EEG Recording

Scalp EEG recording is a most common recording method for monitoring the electrical activity of the brain. See figure 1-8. The recording electrodes are placed on the scalp of the head and record electrical potential differences between the recording electrodes. However, recordings acquired from scalp are usually contaminated by multiple sources of artifacts such as movement artifacts, chewing artifacts, eye movement, vertex waves and sleep spindles, etc. The international 10-20 electrode placement system is commonly used for routine scalp EEG recording [24]. Figure 1-8 is taken from Malmivu and Plonse (1993) [21].

![Diagram of scalp electrode placement]

Figure 1-8. International 10-20 electrode placement

1.6.2 Subdural EEG Recording

The subdural recordings provide less unwanted information in the signals by placing the electrodes under the scalp. See figure 1-9. It requires surgical procedure to place the subdural recording electrodes and the risks of infection increase with the amount...
of time that subdural electrodes stay in subdural regions. The subdural electrodes are usually placed unilaterally to locate the epileptic focus hemisphere. While the scalp EEG recordings provide good resolution for regional localization, it does not localize the focus as precise as subdural EEG recording.

Figure 1-9. Subdural electrode placement

1.6.3 Depth EEG Recording

The depth EEG recording is an invasive recording method. It offers the best spacial resolution among all the EEG recording methods. See figure 1-10. The depth recording
electrodes are usually placed in hippocampus, amygdala and certain neocortical regions. Subdural and depth recording is usually used when individual’s brain fails to response for AEDs and considering epilepsy surgery. The main goal for the epilepsy surgery is to remove the tissue in the brain where seizure are initiated.

![Figure 1-10. Depth electrode placement](image)

### 1.7 Conclusions and Remarks

The search for a common mechanism or brain abnormality that initiates seizures has been the challenging task in epilepsy research. Different forms of epilepsy can casus by different mechanisms; however, they has one important factor in common, during a seizure the abnormal rhythm firing from a large group of neurons. A seizure event can be divided into three states. The state for an individual experiencing a seizure is called as ictus state and preictal state is the aura period or warning prior to a seizure; for example patients may experience unusual smell during preictal states. The state for an individual after enduring a seizure attack is called postictal state. Electrographically, during interictal states, there are evidences of abnormal neuron activities exist in the brain. For example, interictal epileptic discharges, they are brief epileptiform activities which provide the straight forward information for investigating some basic mechanisms for epilepsy. Furthermore, properties of the interictal discharges appear to be the general factors that determine cortical susceptibility to epilepsy [25]. Like many phenomena occur in nature, there exist certain build up period prior to some major events.
Clinical findings in supporting the existence of the preictal state including an increase in cerebral blood flow, oxygen supply related factors in the brain and changes in heart rate [26–31]. From above findings, it is reasonable to hypothesize that a seizure is starting from small abnormally discharging from neurons that recruit and entrain the neighboring neurons into a larger or full seizure. This hypothesis is particularly clear for the focal onset seizures. These recruitment, entrainment and transition phenomena take place in a brain state, the preictal state. Recently, for patients with MTLE several authors have shown it is possible to detect the preictal transitions using quantitative EEG analysis [32–39].

In chapter 2, we investigated the existence of preictal states. We quantified and detected the changes in EEG dynamics that are associated with preictal state in EEG recordings. Three different dynamical measures were used: 1. Largest Lyapunov exponent, a measure is known for measuring the chaoticity of the steady state of a dynamical system, 2. Phase information of the largest Lyapunov spectrum, based on theory from topology and information theory and 3. Approximate entropy, a method for measuring the regularity or predictability of time series. The dynamics support vector machine (D-SVM) was subsequently introduced for improving the performance of the preictal detection. In chapter 3, we analyzed the complexity of EEG recordings using methods developed from nonlinear dynamics and showed the EEG complexity changes prior to the seizure onsets. In the chapter 4, we investigated the differences in nonlinear characteristics between patients with independent bi-temporal seizure onset zone (BTSOZ) and patients with unilateral seizure onset zone (ULSOZ). In chapter 5, we introduced the support vector machines to classify EEG recordings acquired from seizure free and none seizure free patients. In chapter 6, we discussed different methods for spatiotemporal EEG time series analysis. In chapter 7, we proposed a maximum clique framework to study the brain cortical networks. In the chapter 8, we studied the medical treatment effects on the structure of brain cortical networks.
CHAPTER 2
EPILEPSY AND NONLINEAR DYNAMICS

2.1 Introduction

The beginning and termination of epileptic seizures reflect intrinsic, but poorly understood properties of the epileptic brain. One of the most challenging tasks in the field of epilepsy research has been remained the search for basic mechanism that underlies seizures. Traditional studies into the seizure activity have focus upon neuronal apparatuses such as neurotransmitters, receptors or specific ionophores. However, a seizure involves large portions of the cerebral cortex, therefore, it is likely that investigation into the epileptic brain as a system will elucidate important greater information than traditional approaches.

The development of preictal transitions can be considered as a sudden increase of synchronous neuronal firing in the cerebral cortex that may begin locally in a portion of one cerebral hemisphere or begin simultaneously in both cerebral hemisphere. By observing the occurrence of epileptic seizures, it is reasonable to believe that there are multiple states exist in a epileptic brain and the sequences of the states are not deterministic. The preictal transitions are detectable EEG dynamical changes by applying methods developed from nonlinear dynamics. Several groups have reported that seizures are not sudden transitions in and out of the abnormal ictal state; instead, seizures follow a certain dynamical transition that develops over time \[32–39\] see \[40; 41\] for review. In an study of Pijn et al. in 1991, authors were able to demonstrate decrease in the value of correlation dimension at seizure onset in the rat model. In early 1990s, Iasemidis et al., first estimated the largest Lyapunov exponent and reported seizure was initiated detectable transition period by analyzing spatiotemporal dynamics of the EEG recordings; this transition process is characterized by: (1) progressive convergence of dynamical measures among specific anatomical areas “dynamical entrainment ” and (2) following the overshot brain resetting mechanism during post ictal state. Martinerie et al., (1998)
reported decrease in complexity quantify by the correlation density. Le Van Quyen et al. (1999) showed drop in dynamical similarity before seizures using a measure called “dynamical similarity index”; Lehnertz and Elger (1998) demonstrated seizure prediction by “dynamical complexity” time series analysis. Litt et al., (2001) showed increase in accumulated “signal energy” prior to seizure onset. See also [42]. Mormann et al., (2003) detected the preictal state based on decrease in “synchronization” measures [43].

The basic text of nonlinear dynamics and nonlinear dynamical models are presented in the following sections. Nonlinear dynamical measures namely (the largest Lyapunov exponent ($L_{\text{max}}$), Phase/ Angular frequency ($\Phi$), Approximate entropy ($ApEn$)) were used for detecting the preictal transitions in intracranial EEG recordings acquired from patients with intractable MTLE. Since the underlying dynamics of preictal transitions is changing from case to case, this demands sophisticated analytical tools which have the ability for identifying the changes of brain dynamics when preictal transitions occur. The preictal detection performance was further improved by proposed dynamics support vector machine (D-SVM), a classification method developed from optimization theory and data mining techniques. The detection performances were summarized in the later part of this chapter.

### 2.2 Dynamical Systems and State Space

In this section, the basic theory about nonlinear dynamical systems will be given. A dynamical system consists of a set of $d$ state variables, such that each state of the system map to a point $\varphi \in M$. Thus $M$ is $d$-dimensional manifold. A system is said to be a dynamical system if state of the system changes with time. Let us denote $\varphi(t)$ be the state of a system at time $T$, as time evolve (e.g., $t = 0 \rightarrow t = 10,000$), the evolution of the state of the system through state space will form a path. This is path is call “trajectory”. If the current state $\varphi(t)$ uniquely determines all the future state in time, the system is said to be a deterministic dynamical system. If the mapping is not unique, the system is called a stochastic dynamical system. As $\varphi(t)$ evolve for a sufficient amount of time (e.g.,
$t = 0 \rightarrow t = \infty$, the structure of the trajectory (path) will shrink for a dissipative system. For a dissipative system, after a sufficient long time, the number of variables $d$ used to describe state space reduced to a small set of $A$. This set of state variable is called an “attractor”. An attractor can be classified into one of the following four different categories:

1. **Saddle point**: For any given initial conditions, after a sufficient long time, the solution may converge to the same final state (fixed point). An example for this attractor is a constant series.

$$x(t) = x(0), \ t \rightarrow \infty \quad (2-1)$$

2. **Limit cycle**: Instead of converging to a fixed point the dynamical system may converge to a set of states, which are visited periodically. A limited cycle attractor is a closed trajectory through state space.

$$x(t) = x(t + T), \quad (2-2)$$

where T denotes the period of this cycle.

3. **Limit tori**: A limit tori attractor is a limit cycle attractor with multiple period. This attractor will no longer be closed and limited cycle becomes a limit torus.

4. **Strange attractor**: The existence of this type of attractor was unknown until the development of nonlinear dynamics. A strange attractor is defined as an attractor that shows sensitivity to initial conditions (exponential divergence of neighboring trajectories), it may appear to be stochastic in time domain. A strange attractor exhibits regular structure in the phase space (See Figure 2-1, 2-5 for Rössler and Lorenz attractor).

Recall an attractor is a set of state variables; geometrically an attractor can be a point, a curve, a manifold, or even a complicated set with a fractal structure known as the “strange attractor”. Describing these attractors has been one of the achievements of chaos
In the following sections, several methods for qualifying the dynamical attractors are described; results on application for real world EEG recording are also included.

Rössler attractor [44]

\[
\begin{align*}
\dot{x} &= -(y + z), \quad (2-3) \\
\dot{y} &= x + ay, \quad (2-4) \\
\dot{z} &= b + xz - cz. \quad (2-5)
\end{align*}
\]

\((x, y, z) \in \mathbb{R}^3\)

System parameters: \([a=0.45, b=2.0, c=4.0]\);

Initial conditions: \([0 \ 0.01 \ -0.01]\).
Figure 2-2. X component of Rössler system
Figure 2-3. Y component of Rössler system
Figure 2-4. Z component of Rössler system
The Lorenz attractor is introduced by Edward Lorenz in 1963, it is three-dimensional dynamical system and deterministic [45]. The strange attractor in this case is a fractal of Hausdorff dimension between 2 and 3. Grassberger (1983) has estimated the Hausdorff dimension to be 2.06 ± 0.01 and the correlation dimension to be 2.05 ± 0.01 [46].

**Lorenz attractor**

\[
\begin{align*}
\dot{x} &= -\sigma(y - x) \quad (2-7)
\end{align*}
\]

\[
\begin{align*}
\dot{y} &= x(\rho - z) - y \quad (2-8)
\end{align*}
\]

\[
\begin{align*}
\dot{z} &= xy - \beta z \quad (2-9)
\end{align*}
\]

\[
(x, y, z) \in \mathbb{R}^3
\]

System parameters: [\sigma = -10, \rho = 28, \beta = -0.01];

Initial conditions: [0 0.01 -0.01];

Length: 40 seconds;

Sample rate: 50 Hz.
Figure 2-5. Lorenz system
Figure 2-6. X component of Lorenz system
Figure 2-7. Y component of Lorenz system
Figure 2-8. Z component of Lorenz system
2.3 Fractal Dimension

The term "fractal" was first introduced by Mandelbrot in 1983. Roughly speaking, a fractal is a set of points that when looked at smaller scales, resembles the whole set. The concept of fractal dimension refers to a non-integer or fractional dimension originates from fractal geometry. Strange attractors often have a structure that is not simple; they are often not manifolds and actually have a highly fractured character. The dimension that is most useful takes on values that are typically not integers. These non-integer dimensions are called fractal dimensions. For any attractor, the dimension can be estimated by looking at the way in which the number of points within a sphere of radius \( r \) scales as the radius shrinks to zero. The geometric relevance of this observation is that the volume occupied by a sphere of radius \( r \) in the dimension \( d \) behaves as

For regular attractors, irrespective to the origin of the sphere, the dimension would be the dimension of the attractor. But for a chaotic attractor, the dimension varies depending on the point at which the estimation is performed. If the dimension is invariant under the dynamics of the process, we will have to average the point densities of the attractor around it. For the purpose of identifying the dimension in this fashion, we find the number of points \( y(k) \) within a sphere around some phase space location \( x \). This is defined by:

\[
n(x, r) = \frac{1}{N} \sum_{k=0}^{N} \Theta r - | y(k) - x |, \tag{2–11}
\]

where \( \Theta \) is the Heaviside step function such that \( \Theta(n) = 0 \) for \( n < 0 \), \( \Theta(n) = 1 \) for \( n \geq 0 \).

This counts all the points on the orbit \( y(k) \) within a radius \( r \) from the point \( x \) and normalizes this quantity by the total number of points \( N \) in the data. Also, we know that the point density, \( \rho(x) \), on an attractor does not need to be uniform for a strange attractor. Choosing the function as \( n(x; r)^{q-1} \) and defining the function \( C(q; r) \) of two variables \( q \) and \( r \) by the mean of \( n(x; r)^{q-1} \) over the attractor weighted with the natural
density $\rho(x)$ yield:

$$C(q, r) = \int d^x \cdot x\rho(x) \cdot (x, r)^{q-1} = \frac{1}{M} \sum_{k=1}^{M} \left[ \frac{1}{K} \sum_{n=1, n\neq k}^{M} \Theta(r - |y(n) - y(k)|) \right]^{q-1}. \quad (2-12)$$

This $C(q, r)$ is well known correlation integral and the fractal dimension of the system is estimated:

$$C(q, r) \approx r^{(q-1)D_q}, \quad (2-13)$$

for small $r$ that the function $\log[C(q,r)]$ behave linearly with $\log[r]$ for true dimension.

### 2.3.1 Correlation Dimension

When $q$ takes on the value 2, the definition of the fractal dimension, $D_q$, assumes a simple form that lends it to reliable computation. The resulting dimension, $D_2$, is called the correlation dimension of the attractor, and is estimated as the slope of the log-log plot given by:

$$D_2 = \lim_{r \to 0} \frac{\log[C(r)]}{\log[r]}, \quad (2-14)$$

Correlation Dimension has been shown to have the ability in capturing the preictal transition in many studies. For example, A. Babloyantz and A. Destexhe, (1986) shown the existence of chaotic attractor in phase space from EEG acquired from an a patient with absence seizure [47]. Pijn et al., (1997) showed in temporal lobe epilepsy patients that epileptic seizure activity often, but not always, emerges as a low-dimensional oscillation [48]. It was also found that correlation dimension decreases in deep sleep stages, thus reflecting a synchronization of the EEG. A decrease in correlation dimension has been related to the abnormal synchronization behaviors on EEG recordings in epilepsy and other pathologies such as Alzheimers, dementia, Parkinson [48–50].

### 2.3.2 Capacity Dimension

The capacity dimension ($D_0$) is estimated as the number of spheres of radius $\epsilon$ or the number open sets required to cover the metric space. Let us define the number $N(\epsilon)$ as a
function of $\epsilon$ as $\epsilon \to 0$:

$$D_0 = \lim_{\epsilon \to 0} \frac{\log[N(\epsilon)]}{\log[\epsilon]}.$$

(2–15)

### 2.3.3 Information Dimension

The information dimension $D_1$ is a generalization of the capacity that are relative probability of cubes used to cover the attractor. Let $I$ denote the information function:

$$I = - \sum_{i=1}^{N} P_i(r) \log P_i(r),$$

(2–16)

$P_i(r)$ is the normalized probability of an element $i$ is covered such that $\sum_{i=1}^{N} P_i(r) = 1$.

Information dimension is defined as:

$$D_1 = \sum_{i=1}^{N} \frac{P_i(r) \log P_i(r)}{\log(r)},$$

(2–17)

$D_2 \leq D_1 \leq D_0$ if elements of the fractal is equally likely to be visited in the state space.

### 2.4 State Space Reconstruction

Most dynamical properties are contained within almost any variable and its time lags. It is not necessary to reconstruct to entire state space from the measured variable since the attractor dimension will often evolved in a much smaller dimension. The method called state-space reconstruction was proposed by Takens (1981) for reconstructing the state space for a dynamical system. For a series of observations acquired from a dynamical system, the state space reconstruction transforms the observations into state space using an embedding coordinate map $\vartheta : M \rightarrow \mathbb{R}^m$,

$$x(t) = \vartheta(\varphi(t)),$$

(2–18)

where $m$ is the embedding dimension. The transform function $\vartheta$ must be unique (i.e., has no self intersection). Whitney (1936) proved a theorem which can also be used for finding the embedding dimension [51]. $\vartheta : M \rightarrow \mathbb{R}^{2d+1}$; $\vartheta$ embedding is an open and dense set in the space of smooth
For example two planes of dimension \( d_1 \) and \( d_2 \) embedded in \( m \) dimensional space will intersect if \( m \leq d_1 + d_2 \), it is clear that if \( d_1 = d_2 = d \) the embedding dimension need at least \( 2d + 1 \) to avoid the intersections in the state space. However, if only \( s \) subset of the degrees of freedom of \( M \) is represented in our measurement \( x(t) = \vartheta(\varphi(t)) \), it is impossible to obtain additional information. A technique called method of delay is employed to retreat the information from previous times with a embedding window \( \tau \) and form a set of reconstructed delay vector \( x(t) \),

\[
x(t) = (x(t), x(t - \tau), x(t - 2\tau), ..., x(t - (m - 1)\tau)),
\]

and the duration of each embedding vector is

\[
\bar{\tau} = (m - 1) \cdot \tau.
\]

A much more general situation for time-lagged variables constitute an adequate embedding provided the measured variable is smooth and coupled to all the other variables is proved by Takens, and the number of time lag is at least \( 2d + 1 \) [52].

\( \vartheta : M \to \mathbb{R}^{2d+1} \) is an open and dense set in the space of pairs of smooth maps \((f,h)\), where \( f \) is the dynamical system measure by function \( h \).

### 2.5 Lyapunov Exponents

The concept of Lyapunov exponents was first introduced in by A.M. Lyapunov. Lyapunov developed “Lyapunov Stability” concepts to measure the stability of a dynamical system. It quantifies the rate of separation of nearby trajectories in the state space. In this section I describe the method for estimating the Lyapunov exponents. For a dynamical system, sensitivity to initial conditions is quantified by the Lyapunov exponents. For example, consider two trajectories with nearby initial conditions on an attracting manifold. Eckmann and Ruelle (1985) pointed out that when the attractor is chaotic, the trajectories diverge, on average, at an exponential rate characterized by the largest Lyapunov exponent [53]. For a dynamical system as time evolves the sphere
evolves the principal axes expend or contract at rate quantify by Lyapunov exponents. The number of Lyapunov exponent equals to the number of equation (e.g. number of state variable) used to describe the system. The stability is therefore quantified by the average over all the Lyapunov exponents. Wolf et al., (1985) illustrate the Lyapunov spectrum using geometrical interpretation [54]. For a n dimensional system, as time evolves, the order Lyapunov exponent is corresponding to the most expanded to the most contracted principal axes.

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq ... \geq \lambda_n,$$  \hspace{1cm} (2–21)

Iasemidis et al. first used the maximum Lyapunov exponent to show the EEG recordings exhibit abrupt transient drops in chaoticity before seizure onset [55–61]. The maximum Lyapunov exponent is defined by:

$$L \triangleq \frac{1}{N_a} \cdot \sum_{i=1}^{N_a} \log_2 \frac{|\delta X_{i,j}(\Delta t)|}{|\delta X_{i,j}(0)|},$$  \hspace{1cm} (2–22)

$$\delta X_{i,j}(\hat{t}) \triangleq X(t_i + \hat{t}) - X(t_j + \hat{t});$$  \hspace{1cm} (2–23)

where \( t_i = t_0 + (j - 1) \cdot \Delta t \), with and \( i \in [1, N_a] \) and \( \hat{t} \in [0, \Delta t] \), \( \Delta t \) is maximum evolution time for \( \Delta X_{i,j} \).

$$\Delta X_{i,j}(0) = X(t_i) - X(t_j),$$  \hspace{1cm} (2–24)

is perturbation of the fiducial orbit at \( t_i \), and

$$\Delta X_{i,j}(\Delta t) = X(t_i + \Delta t) - X(t_j + \Delta t),$$  \hspace{1cm} (2–25)

is the evolution of \( \Delta X_{i,j}(0) \) after \( \Delta t \).

The \( X(t_i) \) is vector of the fiducial trajectory \( \phi_t(X(t_0)) \), where \( t = t_0 + (i - 1) \cdot \Delta t \), \( X(t_0) = x(t_0), ..., X(t_0 + (p - 1)\tau)^T \), and \( X(t_j) \) is a properly chosen vector adjacent to \( X(t_i) \) in the state space. \( N_a \) is necessary number of iterations for the search through reconstructed state space (with embedding dimension \( p \) and delay \( \tau \)), from \( N \) data points...
and duration $T$. If $D_t$ is the sampling period, then

$$T = (N - 1) \cdot D_t = N_a \cdot \Delta t - (p - 1) \cdot \tau.$$  \hspace{1cm} (2–26)

If the evolution time $\Delta t$ is given in second, then the unit of $L$ is $\text{bit/sec}$. The selection of $p$ is based from Takens’ embedding theorem and was estimated from epoches during ictal EEG recordings. Takens’ embedding theorem is defined:

$$f_x(t) \triangleq (x(t), x(t + \tau), ... x(t + 2n \cdot \tau))^T,$$ \hspace{1cm} (2–27)

using the above defined $f_x$, even if one only observes one variable $x(t)$ for $t \to \infty$, one can construct an embedding of the system into a $p = 2m + 1$ dimensional state space.

The dimension of the ictal EEG attractor is found between 2 to 3 in the state space. Therefore according to Takens’ the embedding dimension would be at least $p = 2 \cdot 3 + 1 = 7$. The selection of $\tau$ is chosen as small as possible to capture the highest frequency component in the data.

2.6 Phase/Angular Frequency

Phase/ angular frequency $\Omega_{max}$ estimates the rate of change of the stability of a dynamical system. Thus, it complements the Lyapunov exponent, which measures the local stability of the system. The difference in phase between two evolved states $X(t_i)$ and
Figure 2-10. Temporal evolution of $STL_{\text{max}}$

$X(t_i + \Delta t)$ is defined as $\Delta \Phi_i$. The average of the local phase differences $\Delta \Phi_i$ between two states in the phase space.

$$\Delta \Phi = \frac{1}{N_\alpha} \cdot \sum_{i=1}^{N_\alpha} \Delta \Phi_i,$$

(2–28)

where $N_\alpha$ is the total number of phase differences estimated from the evolution of $X(t_i)$ to $X(t_i + \Delta t)$ in the state space, and

$$\Delta \Phi_i = |\arccos\left(\frac{X(t_i) \cdot X(t_i + \Delta t)}{\|X(t_i)\| \cdot \|X(t_i + \Delta t)\|}\right)|.$$

(2–29)

2.7 Approximate Entropy

Approximate Entropy ($ApEn$) is a “regularity statistic”, it determines the complexity of a system. It is introduced by Pincus (1991) [62]. It can differentiate between regular and irregular data in instances where moment statistics (e.g. mean and variance) approaches fail to show a significant difference. Applications include heart rate analysis.
in the human neonate and in epileptic activity in electrocardiograms (Diambra, 1999) [63]. Mathematically, as part of a general theoretical framework, ApEn has been shown to be the rate of approximating a Markov chain process [62]. Most importantly, compared ApEn with Kolmogrov-Sinai (K-S) Entropy (Kolmogrov, 1958), ApEn is generally finite and has been shown to classify the complexity of a system via fewer data points via theoretical analysis of both stochastic and deterministic chaotic processes and clinical applications [62; 64–66]. Here I give brief description about ApEn calculation for a time series measured equally in time with length \( n \). Suppose \( S = s_1, s_2, ..., s_n \) is given and use the method of delay we obtain the delay vector \( x_1, x_2, ..., x_{n-m+1} \) in \( R^m \):

\[
x_i = s_i, s_{i+1}, ..., s_{i+m-1},
\]

\[
C_m^i(r) = \frac{\text{number of } x_j \text{ such that } d(x_i, x_j) \leq r_f}{N - m + 1},
\]

\[
(2-30)
\]

\[
(2-31)
\]
where \( m \) is given as an integer and \( r_f \) is a positive real number. The value of \( N \) is the length of compared subsequences in \( S \), and \( r_f \) specifies a tolerance level.

\[
d(x_i, x_j) = \max_{0 \leq k \leq m-1} | s_{i+k} - s_{j+k}|, \tag{2-32}
\]

\( d(x_i, x_j) \) represents the maximum distance between vectors \( x_i \) and \( x_j \) in their respective scalar components.

\[
\Phi^m(r_f) = \sum_{i=1}^{n-m+1} \ln \frac{C^m_i(r_f)}{n-m-1}, \tag{2-33}
\]

Finally the approximate entropy is given by:

\[
ApEn(m, r_f, N) = \Phi^m(r_f) - \Phi^{m+1}(r_f). \tag{2-34}
\]

The parameter \( r_f \) corresponds to an a priori fixed distance between neighboring trajectory and \( r_f \) is chosen according to the standard deviation estimated from data. Hence, \( r_f \) can be viewed as a filtering level and the parameter \( m \) is the embedding dimension determining the dimension of the phase space. Heuristically, \( ApEn \) quantifies the likelihood that subsequences in \( S \) of patterns that are close and will remain close on the next increment. The lower \( ApEn \) value indicates that the given time series is more regular and correlated, and larger \( ApEn \) value means that it is more complex and independent.

### 2.8 Dynamical Support Vector Machine (D-SVM)

The underlying dynamics of preictal transitions is changing from case to case, this requires analytical tools which is capable for identifying the changes in brain dynamics when preictal transitions take place. The detection performance is further improved by the dynamics support vector machine (D-SVM), a method developed from optimization theory and data mining techniques by utilizing dynamic features of EEG.

D-SVM performs classification by constructing an \( N \)-dimensional hyper plane that separates the data into two different classes. The maximal margin classifier rule is used to construct the D-SVM. The objective of maximal margin D-SVM is to minimize the bond
on the generalization error by maximizing the margin with respect to the training data sets. Consider a problem with two class, where a classifier is sought to separate two class of points. The general framework of D-SVM formulation can be written as follows: Let us define two data points in the training set, each belonging to one class, $x^-$ and $x^+$.

\[ \langle w \cdot x^+ \rangle = 1 \quad \text{1 denote one class in the study data sets}, \quad (2-35) \]

\[ \langle w \cdot x^- \rangle = -1 \quad \text{-1 denotes another class in the study data sets}, \quad (2-36) \]

A hyperplane $(w, b)$ is called a canonical hyperplane such that

\[
\min : \frac{1}{2} \| w \| ^2 + \frac{C}{2} \sum_{i=1}^{n} \epsilon_i^2, \quad (2-37)
\]

subject to

\[
y_i (\langle w \cdot x_i \rangle + b) \geq 1 - \epsilon_i, \quad (2-38)
\]

where $C$ is a parameter to be chosen by the user, $w$ is the vector perpendicular to the separating hyperplane, $b$ is the offset and $\epsilon$ are referring to the slack variables for possible infeasibility of the constraints. With this formulation, ones wants to maximize the margin between two classes by minimizing $\| w \| ^2$. The second term of the objective function is used to minimize the misclassification errors that are described by the slack variables $\epsilon_i$. Introducing positive Lagrange multipliers $\alpha_i$ to the inequality constraints in DSVM model, we obtain the following dual formulation:

\[
\min_{\alpha} \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i y_j \alpha_i \alpha_j x_i x_j - \sum_{i=1}^{n} \alpha_i \quad (2-39)
\]

s.t.

\[
\sum_{i=1}^{n} y_i \alpha_i = 0; \quad (2-40)
\]

\[
0 \leq \alpha_i \leq C, i = 1, \ldots n. \quad (2-41)
\]

The solution of the primal problem is given by $w = \sum_i \alpha_i y_i x_i$, where $w$ is the vector that is perpendicular to the separating hyperplane. The free coefficient $b$ can be found
from $\alpha_i(y_i(w \cdot x_i + b) - 1) = 0$, for any $i$ such that $\alpha_i$ is not zero. D-SVM maps a given EEG data set of binary labeled training data into a high dimensional feature space and separates the two classes of data linearly with a maximum margin hyperplane in the dynamical feature space. In the case of nonlinear separability, each data point $x$ in the input space is mapped into a different dynamical feature space using some nonlinear mapping function $\varphi$. Figure 2-12 and 2-13 show the 3D plot for entropy, angular frequency, and $L_{\text{max}}$ during interictal (100 data points dynamical features 2 hours prior to seizure onset) and preictal state (100 data points dynamical features sampled 2 mins prior to seizure onset).

![3D plot for entropy, angular frequency, and Lmax during interictal state](image)

Figure 2-12. Three dimension plot for entropy, angular frequency and Lmax during interictal state
2.9 Statistical Distance

In this section, we introduced $T_{ij}$ index, a statistical measure to estimate the difference of EEG recordings in the dynamical measures. The $T_{ij}$ index at time $t$ between the dynamical profiles of EEG recording at $i$ and $j$ is defined as:

$$T_{ij}(t) = \left| \overline{D_{ij}^t} \right| × \frac{\hat{\sigma}_{ij}^t}{\sqrt{N}},$$  \hspace{1cm} (2-42)

where $\left| \overline{D_{ij}^t} \right|$ denotes the absolute value of the average of all paired differences

$$\left| \overline{D_{ij}^t} \right| = (L_i^t - L_j^t) \mid t \in w(t)), \hspace{1cm} (2-43)$$
over a moving window $w(t)$ defined as

$$w(t) = \left[ \frac{t}{T} - N + 1, \frac{t}{T} \right],$$

(2–44)

where $N$ denotes number of $L_{max}$ in the moving window and $\hat{\sigma}_{ij}^t$ denotes the standard deviation of the sample $D_{ij}$ within $w(t)$. Asymptotically, $T_{ij}(t)$ follows the t-distribution with $N - 1$ degree of freedom. We used $N = 30$ (i.e. averages of 30 paired differences of values from dynamical profiles per moving window).

2.10 Cross-Validation

The leave-one-out rule is used for the statistical cross-validation, cross-validation is a method for estimating the generalization errors based on the re-sampling approach. The decision models are trained $k$ times, in which one of the subsets from training is left out each time, by using only the omitted subset to produce the error criterion of interests. If $k$ equals the sample size, this is called leave-one-out cross-validation. Since we consider the statistical distance between EEG epochs as our decision rule, we call this technique a statistical cross-validation.

2.11 Performance Evaluation of D-SVM

To evaluate the classifier, we categorize the classification into two classes, positive and negative. Four subsets of classification is considered:

- True positive (TP): True positive answers of a classifier denoting correct classification of positive cases;
  A classification result is considered to be true positive if the D-SVM classify a preictal EEG epoch as a preictal EEG sample.

- True negative (TN): True negative answers denoting correct classifications of negative cases;
  A classification result is considered to be true negative if the D-SVM classify a interictal EEG epoch as a interictal EEG sample.
• False positive (FP): False positive answers denoting incorrect classifications of negative cases into the positive cases;
   
   A classification result is considered to be true positive if the D-SVM classify a interictal EEG epoch as a preictal EEG sample.

• False negative (FN): False negative answers denoting incorrect classifications of positive cases into the negative cases;
   
   A classification result is considered to be true positive if the D-SVM classify a preictal EEG epoch as a interictal EEG sample.

The performance of the D-SVM is evaluate using sensitivity and specificity:

\[
\text{Sensitivity} = \frac{TP}{(TP + FN)}
\]

\[
\text{Specificity} = \frac{TN}{(TN + FP)}
\]

The sensitivity can be interpreted as the probability of accurately classifying EEG epochs in the positive case. Specificity can be consider as the probability of accurately classifying EEG epochs in the negative class. In general, one always wants to increase the sensitivity of classifiers by attempting to increase the correct classifications of positive cases (TP). On the other hand, false positive rate can be considered as \((1 - \text{specificity})\) which one wants to minimize.

### 2.12 Patient Information and EEG Description

The information of the patients and the EEG recording are summarized in the table below. For each patient, we randomly selected 200 epochs from interictal state, each epoch is 10.24 seconds long in duration as input to D-SVM classification scheme. The interictal and ictal state is defined as: 1. interictal state: 1 hour away from ictal state 2. preictal state: 5 minutes data length prior to ictal state

### 2.13 Results

The results of this study indicate that D-SVM can correctly the detect preictal state with high sensitivity and specificity. For the patients with bi-lateral seizure onset zone the performance of D-SVM is better than those with uni-lateral seizure onset zone. The
Table 2-1. Patient information and EEG description

<table>
<thead>
<tr>
<th>Patient no.</th>
<th>Gender</th>
<th>Age</th>
<th>Focus region(s)</th>
<th>Number of seizure</th>
<th>Length of recording</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F</td>
<td>45</td>
<td>R.H</td>
<td>12</td>
<td>8 days</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>42</td>
<td>R.H/L.H</td>
<td>18</td>
<td>12 days</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>30</td>
<td>R.H</td>
<td>6</td>
<td>5 days</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>39</td>
<td>R.H/L/H</td>
<td>4</td>
<td>3 days</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>52</td>
<td>R.H</td>
<td>9</td>
<td>5 days</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>65</td>
<td>R.H</td>
<td>7</td>
<td>6 days</td>
</tr>
</tbody>
</table>

Table 2-2. Performance for D-SVM

<table>
<thead>
<tr>
<th>Patient no.</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>91.3 %</td>
<td>91.4 %</td>
</tr>
<tr>
<td>2</td>
<td>93.2 %</td>
<td>95.2 %</td>
</tr>
<tr>
<td>3</td>
<td>90.9 %</td>
<td>91.7 %</td>
</tr>
<tr>
<td>4</td>
<td>94.1 %</td>
<td>96.4 %</td>
</tr>
<tr>
<td>5</td>
<td>84.7 %</td>
<td>92.0 %</td>
</tr>
<tr>
<td>6</td>
<td>89.3 %</td>
<td>94.5 %</td>
</tr>
</tbody>
</table>

Results also confirm that the preictal and interictal brain state are differentiable using approaches developed from nonlinear dynamics.

### 2.14 Conclusions

Although evidence for existing the preictal state and preictal transitions were reported by several studies, the performance for pervious preictal detection is limited due to the rapid changes dynamics of the brain. We proposed D-SVM to embed the EEG dynamics to a higher multidimensional feature space and use the statistical distance to improve the performance for the preictal detection. With the above findings and advantages of nonlinear dynamics, the significance of EEG recordings has been raised and information that are beyond the second moment of the EEG has been further revealed for practical and prognostic uses. These findings appear to be well-founded now, but may become irrelevant later on with the new revelations on the dynamics of the epileptic brain. In future, more cases will needed to be included to validate our findings in this study.
CHAPTER 3
QUANTITATIVE COMPLEXITY ANALYSIS IN MULTI-CHANNEL INTRACRANIAL EEG RECORDINGS FROM EPILEPSY BRAIN

3.1 Introduction

Epilepsy is a brain disorder characterized clinically by temporary but recurrent disturbances of brain function that may or may not be associated with destruction or loss of consciousness and abnormal behavior. Human brain is composed of more than 10 to the power 10 neurons, each of which receives electrical impulses (known as action potentials) from others neurons via synapses and sends electrical impulses via a single output line to a similar (the axon) number of neurons (Shatz 1981). When neuronal networks are active, they produce a change in voltage potential, which can be captured by an electroencephalogram (EEG).

The EEG recordings represent the time series that match up to neurological activity as a function of time. The structure of EEG recordings represent the inter activities among the groups of neurons. Many investigators have applied nonlinear dynamical methods to a broad range of medical applications. Recent developments in nonlinear dynamics have shown the abilities to explain some underlining mechanisms of brain behavior [67–71]. It is known that a dynamical system with d degree of freedom may evolve on a manifold with a lower dimension, so that only portions of the total number of degree of freedom are actually active. For a simple system with limit cycles, it is obvious that time-delay embedding produce an equivalent reconstruction of the true state. According to embedding theorem from Whitney (1936), an arbitrary D-dimension curved space can be mapped into a Cartesian (rectangular) space of $2d + 1$ dimensions without having any self intersections, hence satisfying the uniqueness condition for an embedding [51]. Sauer et al. (1991) generalized Whitneys and Takens’ theorem to fractural attractors with dimension $D_f$ and showed the embedding space only need to have a dimension greater than $2D_f$ [72]. Although it is possible for a fractal to be embedded in another fractal, we only consider the integer embedding. Takens delay embedding theorem
(Takens, 1981) also provided that the time lagged variables constitute an adequate embedding provided the measured variables is smooth and couples to all the variables, and number of time lags is at least $2D + 1$ [52]. For the above reasons, we employed a method proposed by Cao (1997) to estimate the minimum embedding dimension of EEG time series [73]. Like some other exiting methods, Caos method is also under the concepts of false-nearest-neighbors. The false-nearest-neighbors utilized on the fact that if the reconstruction space has not enough dimensions, the reconstruction will perform a projection, and hence will not be an embedding of the desired system [74]. As a result of giving a too low embedding dimension while processing the embedding procedure, two points which is far away in the true state space will be mapped into close neighbor in the reconstruction space. These are then the false neighbors. Caos method does not require large amount of data points, is not subjective and it is not time-consuming find the proper minimum embedding dimension. The EEG recordings was divided into non-overlapping single electrode segments of 10.24 s duration, each of which was estimated for the minimum embedding dimension. Under the assumption the EEG recordings within each 10.24 s duration was approximately stationary [56], we evaluated the underlining dynamical behavior by looking at the minimum embedding dimension over time.

The remaining of this chapter is organized as follows. In Sections. 2 and 3, we describe the data information and explain the algorithm for estimating the minimum embedding dimension estimation. The results from two patients with a total number of six temporal lobe epilepsy (TLE) are given in Section 4. In Section 5, we discuss the results of our findings with respect the use of this algorithm and the function of nonlinear dynamical measurements in the area of seizure control.

3.2 Patient and EEG Data Information

Electrocardiogram (EEG) recordings from bilaterally placed depth and subdural electrodes (Roper and Glimore, 1995) in patients with medically refractory partial seizures of mesial temporal origin were analyzed in this study. Electrode placement. A Inferior
Table 3-1. Patients and EEG data statistics for complexity analysis

<table>
<thead>
<tr>
<th>Patient #</th>
<th>Gender</th>
<th>Age</th>
<th>Focus (RH/LH)</th>
<th>Length of EEG (hr.)</th>
<th>Number of seizure</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>M</td>
<td>19</td>
<td>RH</td>
<td>20h 37m 05s</td>
<td>3</td>
</tr>
<tr>
<td>P2</td>
<td>M</td>
<td>33</td>
<td>LH</td>
<td>09h 43m 57s</td>
<td>3</td>
</tr>
</tbody>
</table>

transverse and B lateral views of the brain, illustrating the depth and subdural electrode placement for EEG recordings are depicted. Subdural electrode strips are placed over the left orbitfrontal (LOF), right orbitofrontal (ROF), left subtemporal (LST), and right subtemporal (RST) cortex see Figure 3-1. The EEG recording data for epilepsy patients were obtained as part of pre-surgical clinical evaluation. They had been obtained using a Nicolet BMSI 4000 and 5000 recording system, using a 0.1 Hz high-pass and a 70 hz low pass filter. Each recording contains a total number of 28 to 32 intracranial electrodes (8 subdural and 6 hippocampal depth electrodes for each cerebral hemisphere). Prior to storage, the signals were sampled at 200Hz using an analog to digital converter with 10 bits quantization. The recordings were stored digitally onto high fidelity video type. Two epilepsy subjects (see Table 3-1) were included in this study.

Figure 3-1. Electrode placement
3.3 Proper Time Delay

We used the mutual information function to estimate the proper time delay between successive components in delay vectors. In theory, the time delay used for time delay vector reconstruction is not the subject of the embedding theorem. Since the data are assumed to have infinite precision, from mathematical point of view delay time can be chosen arbitrary. On the other hand, it is essential to have a good estimation for proper time delay when dealing with none artificial data. For none artificial data, the time delay parameter can affect the dynamical properties under studying, if time delay is very large, the different coordinates may be almost uncorrelated. In this case, the attractor may become very complicated, even if the underlying true attractor is simple. If delay is too small, there is almost no difference between the different components between delay vectors, such that all points are accumulated around the bisectrix in the embedding space. Therefore, it is suggested to look for the first minimum of the time delay mutual information [75]. The concept of mutual information is given as below

Mutual information is originated from information theory and it has been used for measuring interdependence between two series of variables. Let us denote the time series of two observable variables as \( X = \{x_i\}_{i=1}^N \) and \( Y = \{y_j\}_{j=1}^N \), where \( N \) is the length of the series and the time between consecutive observations (i.e. sampling rate) is fixed. The mutual information between observations \( x_i \) and \( y_j \) is defined as:

\[
I_{x,y}(x_i, y_j) = \log \left( \frac{P_{x,y}(s_i, q_j)}{P_x(s_i)P_y(q_j)} \right),
\]

where \( P_{x,y}(x_i, y_j) \) is the joint probability density of \( x \) and \( y \) evaluated at \( (x_i, y_j) \) and \( P_x(x_i), P_y(y_j) \) are the marginal probability densities of \( x \) and \( y \) evaluated at \( x_i \) and \( y_j \) respectively. The unit of mutual information is in bit, when based 2 logarithm is taken.

If \( x \) and \( y \) are completely independent, the joint probability density \( P_{x,y}(x_i, y_j) \) equals to the product of its two marginal probabilities and the mutual information between
\( x_i \) and \( y_j \) vanished thus \( I_{x,y}(x_i, y_j) \) equals to zero. The mutual information measure is symmetric, i.e., \( I_{x,y}(x_i, y_j) = I_{y,x}(y_i, x_j) \).

Fraser and Swinney (1986) showed one could obtain the average mutual information between two time series \( S \) and \( Q \) with length \( n \) \( s_1, s_2, ..., s_n \) and \( q_1, q_2, ..., q_n \) using the entropies \( H(S) \), \( H(Q) \) and \( H(S, Q) \). Suppose we have observed the first series of interest with a set of \( n \) outcomes \( s_1, s_2, ..., s_n \), and each outcome is associated with probabilities \( P_s(s_1), P_s(s_2), ..., P_s(s_n) \) and same as \( Q \).

\[
H(S) = -\sum_i p_s(s_i) \log[p_s(s_i)]. \tag{3–2}
\]

\[
H(Q) = -\sum_i p_q(q_i) \log[p_q(q_i)]. \tag{3–3}
\]

The average amount of uncertainty that a measurement of \( s \) reduces the uncertainty of \( q \) is given

\[
I(Q, S) = H(Q) - H(Q \mid S) = I(S, Q). \tag{3–4}
\]

In other words, “By knowing a measurement of \( s \), how many bits on average can be predicted about \( q \)?”

\[
I(S, Q) = \int P_{sq}(s, q) \log \left[ P_{sq}(s, q) / P_s(s)P_q(q) \right] ds \cdot dq. \tag{3–5}
\]

Suppose a variable \( v \) is investigated by being sampled with sampling interval \( T_s \). Let such process be the context of system \( S \) and system \( Q \), let \( s \) be the measurement of \( v \) at time \( t \), and let \( q \) be the measurement at time \( t + T_s \). Using these measurement to define systems \( S \) and \( Q \), mutual information \( I(Q, S) \) can be calculated. Thus, mutual information becomes a function of \( T_s \). For this problem, mutual information will be the number of bits of \( v(t + T_s) \) that can be predicated, on average, when \( v(t) \) is known. One wants to pick \( T_s \) should be chosen so that \( v(t + T_s) \) is as unpredictable as possible. Maximum unpredictability occurs at minimum of predictability; that is, at the minimum in the mutual information. Because of the exponential divergence of chaotic trajectories,
the first local minimum of $I(Q, S)$, rather than some subsequent minimum, should probably be chosen for the sampling interval $T_s$.

### 3.4 The Minimum Embedding Dimension

Dynamical systems processing $d$ degree of freedom which may choose to proceed on a manifold of much lower dimension, so that only small portions of the degrees of freedom are actually active. In such case it is useful to estimate the behaviors of degrees of freedom over a period of time, and it is obvious that this information can be obtained from that dimension of attractor from the corresponding system. If one chooses the embedding dimension too low this results in points that are far apart in the original phase space being moved closer together in the reconstruction space. Takens delay embedding theorem states that a pseudo-state space can be reconstructed from infinite noiseless time series (when one choose $d > 2d_A$) is often been used when reconstructing the delay vector [52].

There are several classical algorithms used to obtain the minimum embedding dimension [76; 74; 77]. The classical approaches usually require huge computation power and vast among of data. Another limitation of these algorithms is that they usually subjective to different types of data. We evaluated the minimum embedding dimension of the attractors from the EEG by using Caos method. The notions here followed “Practical method for determining the minimum embedding dimension of a scalar time series”. Suppose that we have a time series $(x_1, x_2, x_3, ..., x_N)$. Applying the method of delay we obtain the time delay vector as follows:

$$y_i(d) = (x_i, x_{i+\tau}, ..., x_{i+(d+1)\tau}), i = 1, 2, ..., N - (d - 1)\tau; \quad (3-6)$$

where $d$ is the embedding dimension and $\tau$ is the time-delay and $y_i(d)$ means the $i$th reconstructed vector with embedding dimension $d$. Similar to the idea of the false nearest neighbor method, defining

$$a(i, d) = \frac{\|y_i(d + 1) - y_{n(i,d)}(d + 1)\|}{\|y_i(d) - y_{n(i,d)}(d)\|}, i = 1, 2, ..., N - d\tau \quad (3-7)$$
where \( \| \cdot \| \) is some Euclidian distance and is given in this paper by maximum norm. Define the mean value of all \( a(i, d) \) as

\[
E(d) = \frac{1}{N - d\tau} \sum_{i=1}^{N-d\tau} a(i, d).
\]  

(3–8)

\( E(d) \) is depend only on the dimension \( d \) and the time delay \( \tau \). The minimum embedding dimension is founded when \( E_1(d) = E(d+1)/E(d) \) saturated when \( d \) is larger than some value \( d_0 \) if the time series comes from an attractor. The value \( d_0 + 1 \) is the estimated minimum embedding dimension.

### 3.5 Data Analysis

The first step in the data analysis was to divide the EEG data into non-overlapping windows of 10.24 seconds in duration for nonstationarity purposes. This procedure was to ensure that the underlining dynamical properties were approximately stationary. For each divided window, the first step of estimating the minimum embedding dimension is to construct the delay coordinates using method of delay proposed by. The time delay \( \tau \) was obtained from the first local minimum of the mutual information function. We used these time delay vectors as inputs to Cao’s method for the minimum embedding dimension estimation. The minimum embedding dimension was calculated over time for EEG recordings with 29 electrodes at six brain regions (RTD, RST, ROF, LTD, LST, and LOF) from epilepsy patients. Each brain region contains 46 electrodes; the average of the minimum embedding dimension \( d_0 \) is taken as representation to the underlining brain dynamics. We shall study the minimum embedding dimension in the following three different time periods: \textit{interictal}, \textit{ictal} and \textit{postictal}. These three different time period are defined as follows: 1. \textit{interictal} state: 1 hour away prior to ictal state 2. \textit{preictal} state: 2 minutes data length prior to ictal state 3. \textit{postictal} state: 1 hour after the ictal state Figures 3-2,3-3,3-4,3-5 and 3-6 show typical the minimum embedding dimension over time for six seizures. One can observe the behavior of the average minimum embedding dimension over time for six brain cortical regions. The minimum embedding dimension
Figure 3-2. Average minimum embedding dimension profiles for Patient 1 (seizure 1) showed stable during the interictal state. In other words, the underlying degree of freedom is uniformly distributed over the interictal state in the EEG recordings. The results indicated the lowest minimum embedding dimension were found within the epileptic zone during interictal state (the RST electrodes in Figs. 2, 3, and 4; the LTD electrodes in Figs. 5 and 6). The complexity of the EEG recordings from the epileptic region is lower than that from the brain regions. The values of the minimum embedding dimension from all brain regions start decrease and converge to a lower value as the patient proceed from interictal to ictal state. The underlining dynamical changes before entering ictal period were consistently detected by the algorithm.

3.6 Conclusions

In this chapter, we investigate the degree of complexity for EEG recordings by estimating the minimum embedding dimension. The algorithm we use for the minimum embedding estimation is faster and requires less data points to obtain accurate results.
It is computationally efficient and certainly less time consuming compared to some classical procedures for estimating embedding dimension estimation. The results of this study confirm that it is possible to predict an seizure based on nonlinear dynamics of multichannel intracranial EEG recordings. For majority of seizures the spatiotemporal dynamical features of the preictal transition are similar to that of the preceding seizure. This similarity makes it possible to apply optimization techniques to identify electrode sites that will participate in the next preictal transition, based on their behavior during the previous preictal transition. At present the electrode selection problems were solved efficiently and the solutions were optimally attained. However, future technology may allow physicians to implant thousands of electrode sites in the brain. This procedure will help us to obtain more information and allow to have a better understanding about the epileptic brain. Therefore, in order to solve problems with larger number of recording electrodes, there is a need to develop computationally fast approaches for solving
Figure 3-4. Average minimum embedding dimension profiles for Patient 1 (seizure 3) large-scale multi-quadratic 0–1 programming problems. Our results also are compatible with the findings about the nature of transitions to ictal state in invasive EEG recordings from patients with seizures of mesial temporal origin. The development of multi-quadratic 0–1 programming modeling is in progress.
Figure 3-5. Average minimum embedding dimension profiles for Patient 2 (seizure 4)
Figure 3-6. Average minimum embedding dimension profiles for Patient 2 (seizure 5, 6)
CHAPTER 4
DISTINGUISHING INDEPENDENT BI-TEMPORAL FROM UNILATERAL ONSET IN EPILEPTIC PATIENTS BY THE ANALYSIS OF NONLINEAR CHARACTERISTICS OF EEG SIGNALS

4.1 Introduction

In this present study, we investigate the difference in nonlinear characteristics of electroencephalographic (EEG) recordings between epilepsy patients with independent bi-temporal seizure onset zone (BTSOZ) and those with unilateral seizure onset zone (ULSOZ). Eight adult patients with temporal lobe epilepsy were included in the study, five patients with ULSOZ and three patients with BTSOZ. The approach was based on the test of nonlinear characteristics, defined as the distinction from a Gaussian linear process, in intracranial EEG recordings. Nonlinear characteristics were tested by the statistical difference of short-term maximum Lyapunov exponent $STL_{\text{max}}$, a discriminating nonlinear measure, between the original EEG recordings and its surrogates. Distributions of EEG nonlinearity over different recording brain areas were investigated and were compared between two groups of patients. The results from the five ULSOZ patients showed that the nonlinear characteristics of EEG recordings are significantly inconsistent ($p < 0.01$) over six different recording brain cortical regions (left and right temporal depth, sub temporal and orbitofrontal). Further, the EEG recordings acquire from focal regions of the brain exhibit higher degree of nonlinearities than the homologous contralateral regions and the nonlinear characteristics of EEG are uniformly distributed over the recording areas in all three patients with BTSOZ. These results suggest that it is possible to efficiently and quantitatively determine whether an epileptic patient has ULSOZ based on the proposed nonlinear characteristics analysis. For the ULSOZ patients, it is also possible to identify the focal area. However, these results will have to be validated in a larger sample of patients. Success of this study can provide more essential information to patients and epileptologists and lead to successful epilepsy surgery.
According to the estimates from Epilepsy Foundation, approximately 1% of the US population has been diagnosed with epilepsy [1]. Epilepsy may be treated with drugs, surgery, a special diet, or an implanted device programmed to stimulate the vagus nerve (VNS therapy). For the patient who does not react to Anti-epileptic drugs (AEDs) or other types of the tremens, epilepsy surgery is often considered because it offers the potential for cure of seizures and successful psychosocial rehabilitation. The usefulness of resective surgery for the treatment of carefully selected patients with medically intractable, localization-related epilepsy is clear. Seizure-free rate following temporal lobectomy are consistently 65% to 70% in adults. Epilepsy surgery lies on carefully evaluation of the candidates for surgery, surgical intervention may be carried out with a high probability of success if the area of seizure onset is consistently and repeatedly from the same portion of the brain. The focus localization procedure becomes one of the most important tasks in the pre-surgical examination. Some seizures are resulted from cortical damage. Neuroimaging can help in identifying and localizing the damage regions in the brain and therefore, the focus. Currently, brain magnetic resonance imaging (MRI) provides the best structural imaging study. However, the most common tool in epilepsy diagnosis is electroencephalogram (EEG). The EEG recordings reflect interactions between neurons in the brain. For routine per-surgical evaluation, the patient has to stay in the epilepsy monitoring unit (EMU) for 4 to 5 days and is expected to obtain 4 seizures in the recording. However, in some cases the hospital stay may be longer in order to have enough seizures to identify the focus area. Therefore, a efficient method for determining the suitability of a patient based on the analysis of the shorter EEG recordings would not only improve the outcome of the seizure control but also reduce the patients’ financial burden for the pre-surgical evaluation procedures. In general, EEG captures the spatiotemporal information of the underlying neurons activities nearby the recording electrodes. It is accepted that EEG recording contains non-linear mechanisms at microscopic level. Many studies have shown the presences of nonlinearity in the EEG recordings from both humans
and animal models. Casdagli et al. (1996) reported nonlinearity found using correlation integral and surrogate data technique in intracranial EEG recordings from two patients with temporal lobe epilepsy (TLE) [78]. Andrzejak et al. (2006) showed by focusing on nonlinearity and a combination of nonlinear measures with surrogates appears as the key to a successful characterization of spatiotemporal distribution of epileptic process [71]. B. Weber et al. (1998) also shown evidences for the usefulness of nonlinear time series analysis for the characterization of the spatio-temporal dynamics of the primary epileptogenic area in patients with TLE [79]. Using correlation dimension K. Lehnertz et al. (1995) reported the variance of the EEG dimension during interictal allowed the primary epileptogenic area to be characterized in exact agreement with the results of the presurgical work-up [49]. Many researchers also have shown that the intracranial EEG recordings exhibit certain characteristics that are similar to chaotic systems. For example, Sackellares, Iasemidis et al. first used the maximum Lyapunov exponent, a measure of chaoticity, to show the EEG recordings exhibit abrupt transient drops in chaoticity before seizure onset [32; 60; 80]. By following the concept of spatiotemporal dynamical entrainment (i.e., similar degree of chaoticity between two EEG signals), this group further reported that, during the interictal state, the number of recording sites entrained to the epileptogenic mesial temporal focus was significantly less than that of the homologous contralateral electrode sites [57]. These result suggested that it is possible to identify the epileptogenic focus by examining the dynamical characteristics of the interictal EEG signals. The first part of this study, a phase randomization surrogate data technique was used to generate surrogate EEG signals. By randomizing the phase of the Fourier amplitudes, all information which is not contained in the power spectrum is lost. The surrogate EEG will have the same linear properties thus the same power spectrum and equal coefficients of a linear autoregressive (AR) model. If the original distribution of $L_{max}$ is significantly different from its surrogate, it will be the evidence for the nonlinearity. In the second part of this study, eight adult patients with temporal lobe epilepsy were
included to extend this study. We test the hypothesis that, in patients with unilateral seizure onset zone, the degree of nonlinearity are different over the brain regions, and the focus areas exhibit higher degree of nonlinearity during interictal, preictal, and postictal periods. Further, for patients with independent bi-temporal seizure onset zones, the distribution of nonlinearity would be uniform over brain regions.

4.2 Materials and Methods

4.2.1 EEG Description

EEG recordings were obtained from bilaterally placed depth and subdural electrodes (Roper and Gilmore, 1995), multi-electrode 28–32 common reference channels were used in this study. Figure 4-1 a inferior transverse views of the brain, illustrating approximate depth and subdural electrode placement for EEG recordings are depicted. Subdural electrode strips are placed over the left orbitofrontal (LOF), right orbitofrontal (ROF), left subtemporal (LST), and right subtemporal (RST) cortex. Depth electrodes are placed in the left temporal depth (LTD) and right temporal depth (RTD) to record hippocampal activity.

EEG recordings obtained from eight patients with temporal lobe epilepsy were included in this study. See Table 4-1. Five patients were clinically determined to have unilateral seizure onset zone (ULSOZ) and the remaining three patients were determined to have independent bi-temporal seizure onset zone (BTSOZ). For each patient, three seizures were included in the EEG recordings. Segments from interictal (at least one hour before the seizure), preictal (immediately before the seizure onset) and postictal (immediately after the seizure offset) time intervals corresponding to each seizure were sampled for testing the hypothesis. Two electrodes from each brain area were included, a total of 12 electrodes were analyzed for each patient. The EEG recordings were sampled using amplifiers with input range of 0.6 $mV$, and a frequency range of 0.5–70Hz. The recordings were stored digitally on videotapes with a sampling rate of 200 Hz, using an
Table 4-1. Patients and EEG data statistics

<table>
<thead>
<tr>
<th>Patient #</th>
<th>Gender</th>
<th>Age</th>
<th>Focus (RH/LH)</th>
<th>Length of EEG (hr.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>M</td>
<td>19</td>
<td>RH</td>
<td>6.1</td>
</tr>
<tr>
<td>P2</td>
<td>M</td>
<td>45</td>
<td>RH</td>
<td>5.4</td>
</tr>
<tr>
<td>P3</td>
<td>M</td>
<td>41</td>
<td>RH</td>
<td>5.8</td>
</tr>
<tr>
<td>P4</td>
<td>F</td>
<td>33</td>
<td>RH</td>
<td>5.3</td>
</tr>
<tr>
<td>P5</td>
<td>F</td>
<td>38</td>
<td>RH</td>
<td>6.3</td>
</tr>
<tr>
<td>P6</td>
<td>M</td>
<td>44</td>
<td>RH/LH</td>
<td>5.5</td>
</tr>
<tr>
<td>P7</td>
<td>F</td>
<td>37</td>
<td>RH/LH</td>
<td>4.6</td>
</tr>
<tr>
<td>P8</td>
<td>M</td>
<td>39</td>
<td>RH/LH</td>
<td>5.4</td>
</tr>
</tbody>
</table>

analog to digital (A/D) converter with 10 bit quantization. In this study, all the EEG recordings were viewed by two independent board certified electroencephalographers.

Figure 4-1. 32-channel depth electrode placement

4.2.2 Non-Stationarity

Non–stationarity is an fundamental difficulty for time series analysis. The existing of non–stationarity in a measured time series will result in no invariant measures. Stationarity will cause errors for many algorithms when one is trying interpret the results of an invariant measure. In most cases, one can try to remove the stationarity by using variety of filters or divided the time series into a number of shorter epochs and assume the underlying dynamics to be approximately stationary within each divided epochs. In this
study, prior to calculate $L_{max}$ the EEG recordings were first divided into non-overlapping window with 10.24 second in duration. The same segmentation procedure was also used by Iasemidis et al., (1991) such segmentation technique is often applied especially for medical time series [55].

4.2.3 Surrogate Data Technique

The degree of nonlinearity of a signal can be examined by testing the null hypothesis: “The signal results from a Gaussian linear stochastic process ”. One way to test this hypothesis is to estimate the difference in a discriminating statistic between the original EEG and its surrogate [81]. There are three different procedures for surrogate data.

1. Surrogates are realizations of independent identically distributed ($iid$) random variables with the same mean, variance, and probability density function as the original data. The $iid$ surrogates were generated by randomly permuting in temporal order the samples of the original series. This shuffling process will destroy the temporal information and thus generated surrogates are mainly random observation drawn (without replacement) from the same probability distribution as original data.

2. Fourier transform (FT) surrogates are constrained realizations of linear stochastic processes with the same power spectra as the original data. FT surrogate series were constructed by computing the FT of the original series, by substituting the phase of the Fourier coefficients with random numbers in the range while keeping unchanged their modulus, and by applying the inverse FT to return to the time domain. To render completely uncoupled the surrogate pairs, two independent white noises where used to randomize the Fourier phases.

3. Auto regressive (AR) surrogates are typical realizations of linear stochastic processes with the same power spectra as the original series. By generating a Gaussian time series with the same length as the data, and reordered it to have the same rank distribution. Take the Fourier transform of this and randomize the phases (FT). Finally, the surrogate is obtained by reordering the original data to have the same
rank distribution as the inverse Fourier transform. These surrogates were obtained by fitting an AR model to each of the two original series, using pairs of independent white noises as model inputs to produce completely uncorrelated surrogate series.

In this presented study, we employed the second algorithm to generate the surrogate data. By shuffling the phases but keeping the amplitude of the complex conjugate pairs at the same time the surrogates will have the same power spectrum (autocorrelation) as the data, but will have no nonlinear determinism. For each EEG epoch, ten surrogates will be produced to insured Fourier phases are completely randomized. This surrogate algorithm has been applied and combining with correlation integral, a measure sensitive to a wide variety of non-linearities, was used for detection for nonlinearity by Casdigali et al. (1996) [78].

4.2.4 Estimation of Maximum Lyapunov Exponent

The method for estimation of the Short Term Maximum Lyapunov Exponent ($STL_{\text{max}}$) for nonstationary EEG recordings has been demonstrated in [55; 54]. In this section, we will only give a short description and basic notation of our mathematical models used to estimate $STL_{\text{max}}$. First, let us define the following notation.

- $N_a$ is the number of local $STL_{\text{max}}$’s that will be estimated within a duration $T$ data segment. Therefore, if $D_t$ is the sampling period of the time domain data, $T = (N - 1)D_t = N_a\Delta t + (p - 1)\tau$.
- $X(t_i)$ is the point of the fiducial trajectory $\phi_t(X(t_0))$ with $t = t_i$, $X(t_0) = (x(t_0), \ldots, x(t_0 + (p - 1) * \tau))$, and $X(t_j)$ is a properly chosen vector adjacent to $X(t_i)$ in the phase space.
- $\delta X_{i,j}(0) = X(t_i) - X(t_j)$ is the displacement vector at $t_i$, that is, a perturbation of the fiducial orbit at $t_i$, and $\delta X_{i,j}(\Delta t) = X(t_i + \Delta t) - X(t_j + \Delta t)$ is the evolution of this perturbation after time $\Delta t$.
- $t_i = t_0 + (i - 1) * \Delta t$ and $t_j = t_0 + (j - 1) * \Delta t$, where $i \in [1, N_a]$ and $j \in [1, N]$ with $j \neq i$. 


• $\Delta t$ is the evolution time for $\delta X_{i,j}$, that is, the time one allows $\delta X_{i,j}$ to evolve in the phase space. If the evolution time $\Delta t$ is given in second, then $L$ is in bits per second.

• $t_0$ is the initial time point of the fiducial trajectory and coincides with the time point of the first data in the data segment of analysis. In the estimation of $L$, for a complete scan of the attractor, $t_0$ should move within $[0, \Delta t]$.

Let $L$ be an estimate of the short term maximum Lyapunov exponent, defined as the average of local Lyapunov exponents in the state space. $L$ can be calculated by the following equation:

$$L = \frac{1}{N_a \Delta t} \sum_{i=1}^{N_a} \log_2 \frac{|\delta X_{i,j}(\Delta t)|}{|\delta X_{i,j}(0)|}.$$  \hspace{1cm} (4-1)

4.2.5 Paired t-Test

The comparison between the $L_{\text{max}}$ value estimated from original EEG recordings and the mean of its 10 surrogates in three different time periods: 1. interictal (normal) 2. per-ictal (per-seizure) 3. postictal (post-seizure). The mean of Lmax value from 10 surrogates are taken to construct 95% confident interval. This hypothesis testing will exam whether there exits any significant difference in Lmax values between the original and surrogated EEG recordings. This 95% confident interval follows the t-distribution with degree of freedom equal to 9; the critical value of $t$-distribution is 2.26. The null hypothesis will be rejected if the Lmax values estimated from the original EEG recordings falls outside the 95% confident interval. Further if rejected, the original EEG recordings is considered lost information after performing the surrogate technique. The mean of T-index value for each 10.24 s epochs is served as an indicator of how far away (dissimilar) the original EEG recordings are from the surrogated EEG (null hypothesis). This statistical testing is derived from the well-known paired-t test for comparisons for each paired-dependent observation, it was employed here as a measure of statistical distance between pairs of $L_{\text{max}}$ profiles over a time window. The $T_{ij}$ index at time $t$ between the Lmax profiles of $i$ and $j$ is defined as:
\[ T_{ij}(t) = | \bar{D}_{ij}^t | \times \frac{\hat{\sigma}_{ij}^t}{\sqrt{N}} \]  

(4-2)

where \( | \bar{D}_{ij}^t | \) denotes the absolute value of the average of all paired differences

\[ | \bar{D}_{ij}^t | = (L_i^t - L_j^t) | t \in w(t) \]  

(4-3)

over a moving window \( w(t) \) defined as

\[ w(t) = \left[ \frac{t}{T} - N + 1, \frac{t}{T} \right] \]  

(4-4)

where \( N \) denotes number of \( L_{max} \) in the moving window and \( \hat{\sigma}_{ij}^t \) denotes the standard deviation of the sample \( D_{ij} \) within \( w(t) \). Asymptotically, \( T_{ij}(t) \) follows the t-distribution with \( N - 1 \) degree of freedom. We used \( N = 30 \) (i.e. averages of 30 paired differences of \( L_{max} \) values per moving window). Since each \( L_{max} \) value was derived from a 10.24 second of EEG epoch, the length of our moving window is approximately about 5 minutes. A critical value \( T_{\frac{\alpha}{2}} \) from t-distribution with N-1 (=29) degrees of freedom at significance level \( \alpha = 0.05 \) was used to test the null hypothesis \( H_0 \). For the T-index to reject \( H_0 \), \( T_{ij}(t) \) should be greater than 2.045 \( (T_{0.005,29}) \).

### 4.3 Results

Our study results were plotted in different figures for different state. Figure 4-2, 4-3 show the \( L_{max} \) profile from both original and its surrogates respectively, for analysis purposes, here we define interictal state is the time duration at least 60 mins prior to seizure onset, preictal is the time duration 10 mins prior to seizure onset, finally postictal is 5 mins after seizure onset. The results clearly showed the nonlinearity was detected in three different states, this further clarified \( L_{max} \) was capable to capture information which is not contained in linear AR model. The differences between original \( L_{max} \) and its surrogates was found largest in the EEG recorded from epileptic focus regions.

The \( STL_{max} \) values (the discriminating statistic) estimated from original EEG and its surrogates, and T-index profiles during interictal (Fig. 4-4), pre-ictal (Fig. 4-5)
Figure 4-2. Degree of nonlinearity during preictal state and post-ictal (Fig. 4-6) state, respectively, in a ULSOZ patient. Each figure contains two areas, one from the focal area (top two panels) and another from the homologous contralateral hippocampus area (bottom two panels). From these figures, it is clear that the EEG recorded from the focus area exhibits higher distinction from Gaussian linear processes than those recorded from the homologous contralateral hippocampus area in all three states. Further, the differences of $STL_{max}$ values in the focus area increased from interictal to preictal, and reached to the maximum in the postictal state, but the difference remained the stable in the homologous contralateral area.
Figures 4-7(A), 4-8(A) and 4-9(A) show the degree of nonlinearity (quantified mean T-index values) in EEG for ULSOZ patients, during interictal, preictal and postictal states, respectively. Figures 4-7(B), 4-8(B) and 4-9(B) show Multiple comparisons of nonlinearities in each pair of recording areas (A = LTD, B = RTD, C = LST, D = RST, E = LOF, F = ROF).

The results demonstrated that the nonlinearities were inconsistent across recording areas for all five patients. The results from ANOVA showed that there exist significantly recording area effects on the degree of nonlinearity in all three states (p-values = 0.0019, 0.0012, 0.0015 for interictal, preictal and postictal, respectively). Further, multiple comparisons (shown in Figures 4-7b, 4-8b and 4-9b) revealed that significantly differences
Figure 4-4. \( STL_{\text{max}} \) and T-index profiles during interictal state

\((p\text{-value} < 0.05)\) in the degree of nonlinearity exists between the focus area (Right Hippocampus, RTD) and its homologous contralateral brain area (Left Hippocampus, LTD) in three states, with higher degree of nonlinearity in focal area. The degree of nonlinearity across recorded areas for three BTSOZ patients was shown in Figures 4-10, 4-11 and 4-12 for interictal, preictal and postictal states, respectively. It is observed that the degree of nonlinearity was uniformly distributed over recorded areas. The results from ANOVA revealed that the recording area effects on the degree of nonlinearity in states were not significant \((p\text{-values} = 0.9955, 0.9945, 0.9975\) for interictal, preictal and postictal, respectively).
4.4 Discussion

In this study, we demonstrated the usefulness of nonlinear dynamics measures and investigated the degree of nonlinearity for EEG signals in different brain area for epilepsy patients with and without unilateral seizure onset zone. The degree of nonlinearity was defined as the distinction of the signal from Gaussian linear processes. The method combined the estimation of Short-Term Maximum Lyapunov Exponents ($STL_{max}$), a nonlinear discriminating statistic, and surrogate time series techniques. The hypotheses tested were that (1) there exists difference in signal nonlinearities across recording brain regions for patients with unilateral seizure onset zone, (2) EEG nonlinearities are distributed uniformly across recording brain regions for patients with bi-temporal seizure onset zone, and (3) in patients with unilateral seizure onset zone, the focal area can be identified by comparisons of EEG nonlinearities among recording brain regions. The
results of this study suggest that the distribution of EEG signal nonlinearities across recording brain areas in ULSOZ patients is different from BTSOZ patients. In each of the five test patients with ULSOZ, the EEG nonlinearities were significantly inconsistent among recording areas. On the other hand, the EEG nonlinearities were uniformly distributed across brain areas in each of the three test patients with BTSOZ. These results were consistent during the interictal, preictal and postictal periods. Thus it may be possible to efficiently and quantitatively, with a short duration of EEG recording, determine whether an epileptic patient has unilateral focal area that he/she could be a candidate for epilepsy surgery treatment. If these results can be validated in a large sample patient, the duration of EEG monitoring procedure for a BTSOZ epileptic patient could be greatly shortened. This will not only reduce the cost of the EEG monitoring
Figure 4-7. Nonlinearities across recording areas during interictal state for ULSOZ patients

Figure 4-8. Nonlinearities across recording areas during perictal state for ULSOZ patients
procedure, but also will decrease the risk of infection caused by the implanted recording electrodes. Large sample of patients with ULSOZ and BTSSOZ will be required for reliable estimation of sensitivity and specificity of this method. Correct identification of the focal area in ULSOZ patients is a challenging task. An obvious question could be whether any of brain areas where the EEG signals recorded from is close enough to the actual focal area. If not, it would be very difficult to identify the focal area by an analysis on these EEG signals. Other issues such as the number of recording areas and number of recording electrodes in each area could also affect the results of the analysis. In each of the five ULSOZ patients studied here, EEG signals were recorded from six different brain areas: left and right hippocampus, subtemporal, and orbitofrontal regions. All five patients were clinically determined to have focal area in the right hippocampus. During the interictal state, focal area (right temporal depth) consistently exhibited higher degree of nonlinearity than in the contralateral temporal depth and subtemporal areas (significant observations in 4 out of 5 patients). Similar findings were also observed during preictal and postictal states. These results suggest that it is possible to identify the focal area.
Figure 4-10. Nonlinearities across recording areas during interictal state for BTSOZ patients in patients with ULSOZ. Further studies on a larger sample of patients to validate these results are warranted. Success of this study will provide more much-needed information to guide electroencephalographer and clinician to improve the likelihood of successful surgery.
Figure 4-11. Nonlinearities across recording areas during preictal state for BTSOZ patients
Figure 4-12. Nonlinearities across recording areas during postictal state for BTSOZ patients

p-value = 0.9975
CHAPTER 5
OPTIMIZATION AND DATA MINING TECHNIQUES FOR THE SCREENING OF EPILEPTIC PATIENTS

5.1 Introduction

Detecting and identifying the important abnormal electroencephalogram (EEG) complex by visual examination is not only a time consuming task but also requires fully attentions form the electroencephalographer. In this study, we investigate the possibility for classifying EEG recordings between seizure free patients and patients still suffering from seizure attack using the support vector machine (SVM). Two multi-dimensional SVMs, connectivity SVM (C-SVM) and dynamics SVM (D-SVM), were proposed to identify the EEG recordings acquired from epileptic patients. The C-SVM uses connectivity feature that extracted from EEG recording through mutual information and D-SVM uses three dynamical measures (1. Angular frequency 2. Approximate entropy 3. Short-term largest lyapunov exponent) input for the EEG classification. One hour scalp EEG recording was acquired from each subject (5 class 1, 5 class 2) in this study. Prior to C-SVM classification, the independent component analysis (ICA) methods were applied to remove the noise in the EEG recording for improving the performance od the SVM. D-SVM achieved 94.7% accuracy when identifying class 2 subjects compared to 69.4% accuracy with C-SVM.

Epilepsy is the most common disorders of nervous systems. Preliminary findings on the costs of epilepsy show the total cost to the nation for 2.3 million people with epilepsy was approximately $12.5 billion. The high incidence of epilepsy originates from the fact that it occurs as a result of a large number of factors, including , febrile disturbance, genetic abnormal mutation, developmental deviation as well as brain insults such as central nervous system (CNS) infections, hypoxia, ischemia, and tumors.

Neuron or groups of neurons generate electrical signals when interacting or transmitting information between each other. The EEG recordings capture the local field potential around electrodes that generate from neuron in the brain. Through visual inspection,
the electroencephalographers can search for specific EEG configurations and link it to particular physiological states or neurological disorders. However, performing the visual inspection on long term EEG recordings is time consuming and requires continuously cautions from examiner. Inaccurate diagnosis could lead to severe consequences, especially in life-threatening conditions such as in emergency room (ER) or intensive care unit (ICU). There is currently no reliable tool for rapid EEG screening that can quickly detect and identified the abnormal configurations in EEG recordings. There is a need for developing a reliable technique which would serve as an initial medical diagnosis and prognosis tool.

SVM has been successfully implemented for biomedical research on analyzing very large data sets. Moreover SVM has been recently applied for the use of epileptic seizure prediction and it has been shown to achieved 76% sensitivity and 78% specificity for EEG recordings from 3 patients [82]. Nurettin Acir and Cuneyt Guzelis introduced a two-stage procedure SVM for the automatic epileptic spikes detection in a multi-channel EEG recordings [83]. Bruno Gonzalez-Velldnet et al., reported it is possible to detect the epileptic seizures using three features of the electroencephalogram (EEG), namely, energy, decay (damping) of the dominant frequency, and cyclostationarity of the signals [84]. Along with this directions, the abnormal EEG identification problem can be modeled as binary classification problem – “normal or abnormal ”. Embedded with neuron network and connectivity concepts we first proposed and described an application of connectivity support vector machine C-SVM, C-SVM is based on network modeling concepts and connectivity measures to compare the EEG signals recorded from different brain regions. A detail flow chart of the proposed C-SVM framework is given in Figure ??. We also uses three dynamical features of EEG 1. Angular frequency 2. Approximate entropy 3. Short-term largest lyapunov exponent to conduct the dynamical SVM in the second part of this study.
5.2 EEG Data Information

In this study, the dataset consists of continuous short-term (about 60 minutes) multi-channel scalp EEG recordings of 10 epileptic patients, 5 with medically intractable epilepsy (TLE) and 5 were seizure after treatment. The 19-32 channels scalp EEG recordings were obtained using standard 10-20 system, Nicolet BMSI 6000. Figure 5-1 shows the location of the electrodes on the scalp.

Table 8-1 shows the EEG description from 10 subjects, EEG signals were recorded at sampling rate 250Hz. For consistency, we analyze and investigate EEG time series using bipolar electrodes only from 18 standard channels for every patient. EEG recordings from each subjects were inspected by certificated electroencephalographers. We randomly and uniformly sample two 30-second EEG epochs from each subject. Since EEG recordings were digitized at the sampling rate of 250 Hz, the length of each EEG epoch 7,500 points.
Table 5-1. EEG data description

<table>
<thead>
<tr>
<th>Patient</th>
<th>Duration (minutes)</th>
<th>Length (points)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 1</td>
<td>28.71</td>
<td>430,650</td>
</tr>
<tr>
<td>A 2</td>
<td>29.87</td>
<td>448,025</td>
</tr>
<tr>
<td>A 3</td>
<td>20.89</td>
<td>313,375</td>
</tr>
<tr>
<td>A 4</td>
<td>30.19</td>
<td>452,875</td>
</tr>
<tr>
<td>A 5</td>
<td>29.94</td>
<td>449,150</td>
</tr>
<tr>
<td>N 1</td>
<td>31.50</td>
<td>472,464</td>
</tr>
<tr>
<td>N 2</td>
<td>32.90</td>
<td>493,464</td>
</tr>
<tr>
<td>N 3</td>
<td>28.06</td>
<td>420,964</td>
</tr>
<tr>
<td>N 4</td>
<td>21.90</td>
<td>328,464</td>
</tr>
<tr>
<td>N 5</td>
<td>33.33</td>
<td>499,464</td>
</tr>
</tbody>
</table>
| Total   | 287.29             | 4,309,395       

5.3 Independent Component Analysis

ICA algorithms are a family of related methods for unmixing linearly mixed signals using only recorded time course information, e.g., blind to detailed models of the signal sources as required by earlier signal processing approaches [85]. The ICA algorithms consider the higher-order statistics of the separate data maps recorded at different time points, with no regard for the time order in which the maps occur.[86–88]. Assuming a linear statistical model, we have

\[ \mathbf{x} = \mathbf{A} \mathbf{y}, \]  

(5–1)

where \( \mathbf{x} \) and \( \mathbf{y} \) are random vectors with zero mean and finite covariance; \( \mathbf{A} \) is a rectangular matrix with at most as many columns as rows.

The Elements of vector \( \mathbf{x} \) are the independent components, \( x_1, \ldots, x_n \), which are \( n \) linear mixtures observed. Elements of vector \( \mathbf{y} \) are the independent components, \( y_1, \ldots, y_m \). Without loss of generality, zero mean assumption can always be made. If the observable variables \( x_i \) do not have zero mean, it can always be centered by subtracting the sample mean that produces the zero-mean model. \( \mathbf{A} \) has elements \( a_{ij} \) for \( i = 1, \ldots, n \) and \( j = 1, \ldots, m \). Values of \( n \) and \( m \) may be different.
The ICA model describes how the observed data are generated by a process of mixing the components \( y_i \). All we observe is the random vector \( x \), and we have to estimate both \( A \) and \( y \). The starting point of \( y \) for ICA is simply generated based on the assumption that the components \( y_i \) are statically independent. After estimating the matrix \( A \), and computing its inverse, called it \( B \), we then have obtained the independent components:

\[
y = Bx. \tag{5-2}
\]

In this study, we used the Gaussian Kernel to improve the performance of C-SVM. The Gaussian is defined as

\[
K(x_i, x_j) = \exp\left(-\frac{||x_i - x_j||^2}{2\sigma^2}\right).
\]

The output of Gaussian kernel depends on the Euclidean distance of \( x_i \) from \( x_j \), where \( i, j \) are indices of samples. The support vectors are at the center of Gaussian kernel and determine the area influenced support vectors. It maps data into a different space, and thus improves the linear separability of data.

### 5.4 Dynamical Features Extraction

#### 5.4.1 Estimation of Maximum Lyapunov Exponent

The method for estimation of the Short Term Maximum Lyapunov Exponent \((STL_{max})\) for nonstationary data (e.g., EEG time series) has been demonstrated in [55]. In this section, we will only give a short description and basic notation of our mathematical models used to estimate \( STL_{max} \). First, let us define the following notation.

- \( N_a \) is the number of local \( STL_{max} \)'s that will be estimated within a duration \( T \) data segment. Therefore, if \( D_t \) is the sampling period of the time domain data, \( T = (N - 1)D_t = N_a\Delta t + (p - 1)\tau \).

- \( X(t_i) \) is the point of the fiducial trajectory \( \phi_t(X(t_0)) \) with \( t = t_i \), \( X(t_0) = (x(t_0), \ldots, x(t_0 + (p - 1) \ast \tau)) \), and \( X(t_j) \) is a properly chosen vector adjacent to \( X(t_i) \) in the phase space.
• $\delta X_{i,j}(0) = X(t_i) - X(t_j)$ is the displacement vector at $t_i$, that is, a perturbation of the fiducial orbit at $t_i$, and $\delta X_{i,j}(\Delta t) = X(t_i + \Delta t) - X(t_j + \Delta t)$ is the evolution of this perturbation after time $\Delta t$.
• $t_i = t_0 + (i - 1) * \Delta t$ and $t_j = t_0 + (j - 1) * \Delta t$, where $i \in [1, N_a]$ and $j \in [1, N]$ with $j \neq i$.
• $\Delta t$ is the evolution time for $\delta X_{i,j}$, that is, the time one allows $\delta X_{i,j}$ to evolve in the phase space. If the evolution time $\Delta t$ is given in second, then $L$ is in bits per second.
• $t_0$ is the initial time point of the fiducial trajectory and coincides with the time point of the first data in the data segment of analysis. In the estimation of $L$, for a complete scan of the attractor, $t_0$ should move within $[0, \Delta t]$.

Let $L$ be an estimate of the short term maximum Lyapunov exponent, defined as the average of local Lyapunov exponents in the state space. $L$ can be calculated by the following equation:

$$L = \frac{1}{N_a \Delta t} \sum_{i=1}^{N_a} \log_2 \left| \frac{\delta X_{i,j}(\Delta t)}{\delta X_{i,j}(0)} \right|. \quad (5-3)$$

### 5.4.2 Phase/Angular Frequency

Phase/ angular frequency $\Omega_{\text{max}}$ estimates the rate of change of the stability of a dynamical system. Thus, it complements the Lyapunov exponent, which measures the local stability of the system. The difference in phase between two evolved states $X(t_i)$ and $X(t_i + \Delta t)$ is defined as $\Delta \Phi_i$. The average of the local phase differences $\Delta \Phi_i$ between two states in the phase space.

$$\Delta \Phi = \frac{1}{N_{\alpha}} \cdot \sum_{i=1}^{N_{\alpha}} \Delta \Phi_i, \quad (5-4)$$

where $N_{\alpha}$ is the total number of phase differences estimated from the evolution of $X(t_i)$ to $X(t_i + \Delta(t))$ in the state space, and

$$\Delta \Phi_i = \left| \arccos\left( \frac{X(t_i) \cdot X(t_i + \Delta t)}{\|X(t_i)\| \cdot \|X(t_i + \Delta t)\|} \right) \right|. \quad (5-5)$$
5.4.3 Approximate Entropy

Approximate Entropy (ApEn) is a “regularity statistic” that determines change in complexity of a system. It was introduced by Pincus (1991) [62]. It can differentiate between regular and irregular data in instances where moment statistics (e.g., mean and variance) approaches fail to show a significant difference. Applications include heart rate analysis in the human neonate and in epileptic activity in electrocardiograms (Diambra, 1999) [63]. Mathematically, as part of a general theoretical framework, ApEn has been shown to be the rate of approximating a Markov chain process [62]. Most importantly, compared ApEn with Kolmogorov-Sinai (K-S) Entropy (Kolmogrov, 1958), ApEn is generally finite and has been shown to classify the complexity of a system via fewer data points via theoretical analysis of both stochastic and deterministic chaotic processes and clinical applications [62; 64–66]. Here I give brief description of ApEn calculation for a time series measured equally in time with length \( n \), \( s = s_1, s_2, ..., s_n \) is given by first form a sequence of vector \( x_1, x_2, ..., x_{n-m+1} \) in \( R^m \) using:

\[
x_i = u_i, u_{i+1}, ..., u_{i+m-1},
\]

\[
C^m_i(r) = \frac{\text{number of } x_j \text{ such that } d(x_i, x_j) \leq r_f}{N - m + 1},
\]

where \( m \) is given as an integer and \( r_f \) is a positive real number. The value of \( l \) is the length of compared subsequences in \( S \), and \( r_f \) specifies a tolerance level.

\[
d(x_i, x_j) = \max_{0 \leq k \leq m-1} | u_{i+k} - u_{j+k} |,
\]

\( d(x_i, x_j) \) represents the maximum distance between vectors \( x_i \) and \( x_j \) in their respective scalar components.

\[
\Phi^m(r_f) = \sum_{i=1}^{n-m+1} \ln \frac{C^m_i(r_f)}{n - m - 1}.
\]

Finally the approximate entropy is given by:

\[
\text{ApEn}(m, r_f, N) = \Phi^m(r_f) - \Phi^{m+1}(r_f).
\]
The parameter $r_f$ corresponds to an a priori fixed distance between neighboring trajectory points and frequently, $r_f$ is chosen according to the standard deviation estimated from data. Hence, $r_f$ can be viewed as a filtering level and the parameter $m$ is the embedding dimension determining the dimension of the phase space. Heuristically, $ApEn$ quantifies the likelihood that subsequences in $S$ of patterns that are close and will remain close on the next increment. The lower $ApEn$ value indicates that the given time series is more regular and correlated, and larger $ApEn$ value means that it is more complex and independent.

5.5 Dynamical Support Vector Machine

Dynamical Support Vector Machine (D-SVM) performs classification by constructing an $N$-dimensional hyperplane that separates the data into two different classes. The maximal margin classifier rule is used to construct the D-SVM. The objective of maximal margin D-SVM is to minimize the bond on the generalization error by maximizing the margin with respect to the training data sets. Consider a problem with two class, where a classifier is sought to separate two class of points. The D-SVM formulation can be written as follows: Let us define two data points in the training set, each belonging to one class, $x^-$ and $x^+$. 

\[ \langle w \cdot x^+ \rangle = 1 \]  
\[ \langle w \cdot x^- \rangle = -1 \]

\[ \langle w \cdot x^+ \rangle = 1 \]  
\[ \langle w \cdot x^- \rangle = -1 \]

\[ \langle w \cdot x^- \rangle = -1 \]  
\[ \langle w \cdot x^+ \rangle = 1 \]

A hyperplane $(w, b)$ is called a canonical hyperplane such that

\[
\min : \frac{1}{2} \|w\|^2 + \frac{C}{2} \sum_{i=1}^{n} \epsilon_i ^2
\]

subject to

\[ y_i (\langle w \cdot x_i \rangle + b) \geq 1 - \epsilon_i \]

where $C$ is a parameter to be chosen by the user, $w$ is the vector perpendicular to the separating hyperplane, $b$ is the offset and $\epsilon$ are referring to the slack variables for possible
infeasibilities of the constraints. With this formulation, ones wants to maximize the margin between two classes by minimizing $\| w \|^2$. The second term of the objective function is used to minimize the misclassification errors that are described by the slack variables $\epsilon_i$. Introducing positive Lagrange multipliers $\alpha_i$ to the inequality constraints in D-SVM model, we obtain the following dual formulation:

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i y_j \alpha_i \alpha_j x_i x_j - \sum_{i=1}^{n} \alpha_i$$  \hspace{1cm} (5–15)
s.t.
\[ \sum_{i=1}^{n} y_i \alpha_i = 0 \quad (5-16) \]
\[ 0 \leq \alpha_i \leq C, i = 1, \ldots, n \quad (5-17) \]

The solution of the primal problem is given by \( w = \sum_{i} \alpha_i y_i x_i \), where \( w \) is the vector that is perpendicular to the separating hyperplane. The free coefficient \( b \) can be found from \( \alpha_i(y_i(w \cdot x_i + b) - 1) = 0 \), for any \( i \) such that \( \alpha_i \) is not zero. D-SVM map a given EEG data set of binary labeled training data into a high dimensional feature space and separate the two classes of data linearly with a maximum margin hyperplane in the dynamical feature space. In the case of nonlinear separability, each data point \( x \) in the input space is mapped into a different space using some nonlinear mapping function \( \varphi \). A nonlinear kernel function, \( k(x, \hat{x}) \), can be used to substitute the dot product \( \langle \varphi(x), \varphi(\hat{x}) \rangle \). This kernel function allows the D-SVM to operate efficiently in a nonlinear high-dimensional feature space without being adversely affected by dimensionality of that space.

### 5.6 Connectivity Support Vector Machine

In this subsection, we describe the framework of C-SVM. Instead of modeling the deterministic evolution of the physiological state from time, we now model the evolution of an ensemble of possible states by implementing EEG representation as a “information path way” or “Brain Connectivity”. The brain connectivity can be formulated as follows.

Let \( G \) be an undirected graph with vertices \( V_1, \ldots, V_n \), where \( V_i \) represents electrode \( i \). There is an edge (link) with the weight \( w_{ij} \) for every pair of nodes \( V_i \) and \( V_j \) corresponding to the connectivity of the brain dynamics between these two electrodes. The connectivity or synchronization can be viewed as been activated by interactions between neurons in the local circuitry underlying the recording electrodes. Figure 5-4 represents a hypothetical brain graph in which each connected path denotes the underlying connectivity. With this graph model, the attributes of C-SVM inputs are the pair-wised relation between two time series profiles rather than time stamps of a time series profile. In this context, the
Figure 5-3. Support vector machines

input of C-SVM is the degree of connectivity between different brain regions. Given \( n \) time series data points, each with \( m \) time stamps, the proposed framework will decrease the number attributes by \( 2(n-1)/m \) times. Let \( l \) be the total number of data points, the dimensionality can be reduced from \( A \in \mathbb{R}^{l \times n \times m} \) to \( \tilde{A} \in \mathbb{R}^{l \times n \times \frac{(n-1)}{2}} \).

The connectivity among the 18 EEG channels is calculated for each sample as shown in Figure 5-4. The connectivity are Euclidean Distance-based between channel \( i \) and channel \( j, i \neq j \), for \( i, j = 1, \ldots, 18 \). For example: let \( C_i \) and \( C_j \) denote EEG time series from channel \( i \) and channel \( j \), respectively. Each epoch of time series has length 30 seconds, which is equal to 7,500 points. So the size of vector \( C_i \) (or \( C_j \)) is 7,500. The connectivity between \( C_i \) and \( C_j \) using Euclidean distance we obtain:

\[
EU_{ij} = \frac{\sum_{k=1}^{7500} (C_i^k - C_j^k)^2}{7500}.
\]

and 18 \( \times \) 17 connectivity profiles for each sample. Thus, the C-SVM transforms each EEG time series sample into this with (18 \( \times \) 17) number of attributes. C-SVM largely reduces number of attributes from \( m = (18 \times 7500) \) to \( m = (18 \times 17) \) and also saves memory resources and computational time.
Figure 5-4. Connectivity support vector machine

5.7 Training and Testing: Cross Validation

To reduce the bias of training and test data, we proposed to implement the cross validation technique. Cross validation is extensively used as a method to estimate the generalization error on “re-sampling”. We proposed to implement 5-fold cross-validation to estimate the accuracy of the classification. Generally in $n$-fold cross validation, EEGs data will be divided into $n$ subsets of (approximately) equal size. C-SVM and D-SVM will be trained and tested $n$ times, in which one of the subsets from training is left out each time, and tested on the omitted subset. After $n$ iterations, we averaged the classification performance for different iteration. We performed this cross validation procedure in 10 repetitions to ensure that the training and testing EEG data was unbiased.

5.8 Results and Discussions

The results of classification performance using D-SVM and C-SVM with gaussian kernels are summarized in Table 5-2. The proposed C-SVM without Gaussian kernel produced average accuracy of 69.4% and D-SVM using dynamical features obtained from EEG recordings produced 94.7% in classification.
Table 5-2. Results for D-SVM using 5-fold cross validation

<table>
<thead>
<tr>
<th>D-SVM/C-SVM</th>
<th>Dynamical features</th>
<th>UNICA</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-fold CV</td>
<td>1</td>
<td>46%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>92%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>94%</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>98%</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>96%</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>92%</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>96%</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>95%</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>94%</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>95%</td>
</tr>
</tbody>
</table>

Average % of correctness

<table>
<thead>
<tr>
<th></th>
<th>D-SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>94.7%</td>
</tr>
<tr>
<td></td>
<td>69.4%</td>
</tr>
</tbody>
</table>

The SVM has a very long statistical foundation and assure the optimal feasible solution for a set of training data, given a set of features and the operation of the SVM. In this study, we attempted to study the separability between abnormal EEG and normal EEG using different EEG features. We tested the performance on scalp EEG recordings from normal individuals and abnormal patients. The EEG data was filtered using ICA algorithm. ICA filters the noise in EEG scalp data, keeps essential structure and makes better representable EEG data sets. The Euclidean distance based C-SVM was proposed to evaluate the connectivity among different brain regions. The dynamical features were generated as input for D-SVM, the classification results of the proposed D-SVM are very encouraging. The results indicated that D-SVM improves classification accuracy compare to C-SVM. It gives an average accuracy of 94.7%. The dynamical features provide a subset in the feature space and improve classification accuracy.
CHAPTER 6
SPATIO-TEMPORAL EEG TIME SERIES ANALYSIS

6.1 Introduction

The degree of synchronization is an important indication of how information is processing in the brain. The quantification of neuronal synchronization has been investigated using different approaches, from linear cross-correlation to phase synchronization or advanced dynamical interdependence analysis. These synchronization tools have also been applied to EEG recordings and have been shown to be able to detect increased in synchronization measures prior to the seizures [89]. Synchronization can be quantified in both space and time domain. For a multi-variate system, understanding the interactions among its various variables, whose behavior can be represented along time as time-sequences, presents many challenges. One of the key aspects of highly synchronized systems with spatial extent is their ability to interact both across space and time, which complicates the analysis greatly. In biological systems such as the central nervous system, this difficulty is compounded by the fact that the components of interest have nonlinear complicated dynamics that can dictate overall changes in the system behavior. The exact figure of how to quantify the information exchanges in a system remains ambiguous. Studies on multi-variate time series analysis have resulted in development of a wide range of signal-processing tools for quantification of synchronization in systems. However, the general consensus on how to quantify this phenomenon is largely uncertain. In the literature, synchronization between variables can be categorized as identical synchronization, phase synchronization and generalized synchronization. In the following chapters, I undertake an in-depth analysis of preictal and interictal synchronization behavior, focusing on EEG recordings from patients with temporal lobe and generalized seizures.
6.2 Second Order Synchronization Measures

6.2.1 Cross Correlation Function

Cross correlation is the most common linear techniques for quantifying the relationship between two variables. It measures the linear relationship between two variables and it is a symmetrical measure. Consider an identically distributed stationary stochastic process \( X \). Assuming ergodicity, represent a time series acquired at equal spaced time point \( x(n), n = 1, 2, \ldots, p \), where \( p \) is the number of total number of observations in \( x(n) \); same as \( y(n), n = 1, 2, \ldots, q \), where \( q \) is the number of total number of observations in \( y(n) \). Cross-correlation between two discrete time series \( x(n) \) and \( y(n) \) can be defined as a function of lags, \( \tau \) as follows:

\[
R_{xy}(\tau) = \frac{1}{N - \tau} \sum_{i=1}^{N-\tau} x_i y_{i+\tau}
\]

where \( \tau = -N - 1, -N, \ldots, 0, \ldots, N - 1 \). This cross correlation is symmetric \( R_{xy}(\tau) = R_{xy}(-\tau) \). It can also be shown that

\[
C_{xy} = R_{xy} - \mu_x \mu_y,
\]

where \( \mu_x, \mu_y \) are the estimated mean of \( x(n) \) and \( y(n) \).

Cross-correlation coefficient between \( x(n) \) and \( y(n) \) is defined as cross covariance normalized by the product of the square root of the variances of two observed series, as follows:

\[
\rho_{xy}(\tau) = \frac{C_{xy}(\tau)}{\mu_x \mu_y},
\]

where \( \mu_x \mu_y \) are the standard deviation of \( x(n) \) and \( y(n) \). Cross-correlation coefficient bounded between -1 and +1.

The as the frequency domain, define \( x(\omega) = F(x(n)) \) and \( Y(\omega) = F(y(n)) \) as the Fourier transform equivalents of \( x(n) \) and \( y(n) \). If the cross-spectrum \( C_{xy}(\omega) \) and auto-spectrums \( C_{xx}(\omega) \) and \( C_{yy}(\omega) \) are defined by normalized cross-coherence and can be
represented as
\[ \sigma_{xy}(\omega) = \frac{C_{xy}(\omega)}{(C_{xx}(\omega)C_{yy}(\omega))^{1/2}}. \]  
(6-4)

The cross-coherence quantifies the degree of coupling between \( X \) and \( Y \) at given frequency \( \omega \) and it is also bounded between -1 and +1.

### 6.2.2 Partial Directed Coherence

Partial Directed Coherence a frequency domain based Granger-causality technique which says that an observed time series \( x_j(n) \) causes another series \( x_i(n) \), if knowledge of \( x_j(n) \)'s past significantly improves prediction of \( x_i(n) \) [90]. However, the reverse case may or may not be true. To make a quantitative assessment of the amount of linear interaction and the direction of interaction among multiple time-series, the concept of Granger-causality can be used and into the development multi autoregressive model (MVAR). The partial directed coherence from \( j \) to \( i \) at a frequency \( \omega \) is given by:

\[ \pi_{ij}(\omega) = \frac{A_{ij}(\omega)}{a_j^H(\omega)}, \]  
(6-5)

where for \( i = j \)

\[ A_{ij}(\omega) = 1 - \sum_{r=1}^{p} a_{ij}(r)e^{-j2\pi r}; \]  
(6-6)

and for \( i \neq j \)

\[ A_{ij}(\omega) = -\sum_{r=1}^{p} a_{ij}(r)e^{-j2\pi r}; \]  
(6-7)

\( a_{ij} \) are the multivariate auto-regressive (MAR) coefficient at lag \( r \), obtained by least-square solution of MAR model

\[ x = \sum_{r=1}^{p-1} A^T x(p - r) + \epsilon, \]  
(6-8)

here
\[ A^T = \begin{pmatrix}
  a_{r11}^r & a_{r12}^r & \cdots & a_{r1N}^r \\
  \vdots & \vdots & \ddots & \vdots \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{rN1}^r & a_{rN2}^r & \cdots & a_{rNN}^r
\end{pmatrix} \] (6–9)

and

\[ x(p - r) = \begin{pmatrix}
  x_1(p - r) \\
  x_2(p - r) \\
  x_3(p - r) \\
  \vdots \\
  x_N(p - r)
\end{pmatrix} \] (6–10)

p denotes the depth of the AR model, r denotes the delay and n is the prediction error or the white noise.

Note that \( \pi_\omega \) quantifies the relative strength of the interaction of a given signal source j with regard to signal i as compared to all of js interactions to other signals. It turns out that the PDC is normalized between 0 and 1 at all frequencies. If i=j, the Partial Directed Coherence represents the casual influence from the earlier state to its current state.

The MVAR approaches have been used to determine the propagation of epileptic intracranial EEG activity in temporal lobe and mesial seizures [2-3, 10, 19-20]. However, these models strictly require that the measurements be made from all the nodes, or the directional relationships could be ambiguous. In addition, there remains no clear evidence of causality relationships among the cortical regions as suggested “the nature of synchronization is mostly instantaneous or without any detectable delay”[91].

The general, nonlinearity are commonly inherent within neuronal recordings, the above linear measures are typically restricted to measure statistical dependencies up to the second order [92]. If observations are Gaussian distributed, the 2nd order statistics are sufficient to capture all the information in the data. However, in practice, EEG data
sets are either non- Gaussian or quasi-Gaussian rendering the linear statistical measures inadequate [78].

### 6.3 Phase Synchronization

The notion of synchronization was introduced to physics by Huygens for he was trying to describe two coupled frictionless harmonic oscillators in 1673 [93]. Conceptually, if the rhythms of one signal are in harmony with that of the other, the two signals are known to be phase locked. Phase synchronization can therefore be defined as the degree to which two signals are phase locked. For this Huygens’ classical case, phase synchronization is usually defined as locking of the phases of two oscillators.

\[
\phi_{n,m} = n\varphi_a(t) - m\varphi_b(t) \leq \text{constant,} \tag{6–11}
\]

where \( n \) and \( m \) are integers, \( \varphi_a(t) \) and \( \varphi_b(t) \) denote the phases of the oscillators, and \( \phi_{n,m} \) is defined as their relative phase.

Rosenblum et al, (1996) generalized the above phase locking formula by the weaker condition of phase entrainment [94]:

\[
|\phi_{n,m}| = |n\varphi_a(t) - m\varphi_b(t)| < \text{constant}, \tag{6–12}
\]

by even weaker condition of frequency locking:

\[
\tilde{\eta}_{n,m} = n\tilde{\eta}_a - m\tilde{\eta}_a = n\frac{d\tilde{\varphi}_a}{dt} - m\frac{d\tilde{\varphi}_b}{dt} = 0, \tag{6–13}
\]

\( \tilde{\cdot} \) denotes averaging over time, and \( \varphi_{n,m} \) the relative frequency of the system.

Most of real world signals have broad spectra. For example, EEG signal recordings are usually in the range of 0.1 to 1000 Hz even though they are usually band pass filtered between 0.1 and 70 Hz since a major portion of the energy is contained in that spectrum. The EEG can be classified roughly into five (5) different frequency bands, namely the delta (0-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), Beta (12-16 Hz) and the Gamma (16-80 Hz) frequency bands. Freeman demonstrated evidence of phase locking between EEG
frequency bands across different regions of the brain, leading to certain clinical events such as evoked potentials [95; 96]. Similarly, it is also believed that phase synchronization across narrow frequency EEG bands, pre-seizure and at the onset of seizure may provide useful hints of the spatio-temporal interactions in epileptic brain [33; 97; 35]. Hilbert transform is used to compute the instantaneous parameters $\varphi_a(t)$ and $\varphi_b(t)$ of a time-signal. Consider a real-valued narrow-band signal $x(t)$ concentrated around frequency $f_c$. Define $x(t)$ as

$$\bar{x} = x(t) \times \frac{1}{\tau t},$$

(6–14)

where $\bar{x}(t)$ can be regarded as the output of the filter with impulse respond

$$h(t) = \frac{1}{\tau t}, -\infty < t < \infty,$$

(6–15)

excited by an input signal $x(t)$. This filter is called a Hilbert transformer. Hilbert transforms are accurate only when the signals have narrow-band spectrum, which is often unrealistic for most real-world signals. Pre-processing of the signal such as decomposing it into narrow frequency bands is needed before we apply Hilbert transformation to compute the instantaneous parameters. Certain conditions need to be checked to define a meaningful instantaneous frequency on a narrow-band signal. It has been reported that the distinct differences in the degree of synchronization between recordings from seizure-free intervals and those before an impending seizure, indicating an altered state of brain dynamics prior to seizure activity [89].

6.4 Mutual Information

The concept of mutual information dates back to the work of Shannon in 1948 [98]. Generally, mutual information measures the information obtained from observations of one random event for the other. It is known that mutual information has the capability to capture both linear and nonlinear relationships between two random variables since both linear and nonlinear relationships can be described through probabilistic theories. Here in our model, the mutual information measures how much information of EEG time
series acquired from electrode $x$ is presented by electrode $y$ and vice versa. Let $X$ be the set of data points where its possible realizations are $x_1, x_2, x_3, \ldots, x_n$ with probabilities $P(x_1), P(x_2), P(x_3), \ldots P(x_n)$. The Shannon entropy $H(X)$ of $X$ is defined as Mutual information has been applied for measuring the interdependency between two time series. Many previous studies have shown its superior performance over the traditional linear measures [99–104].

Kraskov et al., 2004 introduced two classes of improved estimators for mutual information $M(X, Y)$ from samples of random points distributed according to some joint probability density $\mu(x, y)$. In contrast to conventional estimators based on histogram approach, they are based on entropy estimates from $k$–nearest neighbour distances. Let us denote the time series of two observable variables as $X = \{x_i\}_{i=1}^N$ and $Y = \{y_j\}_{j=1}^N$, where $N$ is the length of the series and the time between consecutive observations (i.e., sampling period) is fixed. Then the mutual information is given by:

$$I(X, Y) \approx I_{\text{binned}}(X, Y) = \sum_i \sum_j P_{x,y}(x_i, y_j) \log \left( \frac{P_{x,y}(x_i, y_j)}{P_x(x_i)P_y(y_j)} \right).$$

(6–16)

where $p_x(i) = \int_i \cdot dx$, $p_y(i) = \int_i \cdot dy$ and

$$p(i, j) = \int_i \int_j p(x, y) dx dy$$

(6–17)

“$\int_i$ ”denotes the integral over bin $i$. If $n_x(i)$ and $n_y(j)$ are the number of data points in the $i$th bin of $X$ and $j$th bin of $Y$; $n(i, j)$ is the number of data points in the intersection bin $(i, j)$. The probabilities are estimated as $p_x(i) \approx n_x(i)/N$, $p_y(j) \approx n_y(j)/N$ and $p(i, j) \approx n_x(j)/N$. Rather then bin approach the mutual information can be estimated from $k$-nearest neighbor statistics.

We first estimate $\hat{H}(X)$ from $X$ by

$$\hat{H}(X) = -\frac{1}{N} \sum_{i=1}^N P(X = x_i).$$

(6–18)
For $X$ and $Y$ time series we define $d_{ij}^{(x)} = \|x_i - x_j\|, d_{ij}^{(y)} = \|y_i - y_j\|$ as the distances for $x_i$ and $y_i$ between every other point in matrix spaces $X$ and $Y$. One can rank these distances and find the $knn$ for every $x_i$ and $y_i$. In the space spanned by $X, Y$, similar distance rank method can be applied for $Z = (X, Y)$ and for every $z_i = (x_i, y_i)$ one can also compute the distances $d_{ij}^{(z)} = \|z_i - z_j\|$ and determine the $knn$ according to some distance measure. The maximum norm is used in this study:

$$d_{ij}^{(z)} = \max\{\|x_i - x_j\|, \|y_i - y_j\|\}, \quad d_{ij}^{(x)} = |x_i - x_j|.$$  

(6–19)

Next let $\frac{\epsilon(i)}{2}$ be the distance between $z_i$ and its $k$th neighbor. In order to estimate the joint probability density function ($p.d.f.$), we consider the probability $P_k(\epsilon)$ which is the probability that for each $z_i$ the $k$th nearest neighbor has distance $\frac{\epsilon(i)}{2} \pm d\epsilon$ from $z_i$. This probability means that $k - 1$ points have distance less than the $k$th nearest neighbor and $N - k - 1$ points have distance greater than $\frac{\epsilon(i)}{2}$ and $k - 1$ points have distance less than $\frac{\epsilon(i)}{2}$. $P_k(\epsilon)$ is obtained using the multinomial distribution:

$$P_k(\epsilon) = k \binom{N - 1}{k} \frac{dp_i(\epsilon)}{d\epsilon} p_i^{k-1}(1 - p_i)^{N-k-1},$$  

(6–20)

where $p_i$ is the mass of the $\epsilon$-ball. Then the expected value of $\log p_i$ will be:

$$E(\log p_i) = \psi(k) - \psi(N),$$  

(6–21)

where $\psi(\cdot)$ is the digamma function:

$$\psi(t) = \frac{\Gamma(t)}{\Gamma(t)} - \frac{d\Gamma(t)}{dt},$$  

(6–22)

where $\Gamma(\cdot)$ is the gamma function. It holds that $\psi(1) = C$ where $C$ is the Euler - Mascheroni constant ($C \approx 0.57721$). The mass of the $\epsilon$-ball can be approximated (if we consider the probability density function inside the ball is the same) by:

$$p_i(\epsilon) \approx c_i e^{-d} P(X = x_i),$$  

(6–23)

111
where $c_{d_x}$ is the mass of the unit ball in the $d_x$-dimensional space. From Eq.\((6–23)\) we can find an estimator for $P(X = x_i)$:

$$\log[P(X = x_i)] \approx \psi(k) - \psi(N) - dE(\log \epsilon(i)) - \log c_{d_x}, \quad (6–24)$$

finally with Eq\((6–24)\) and Eq\((6–25)\) we obtain the Kozachenko-Leonenko entropy estimator for $X$ [105]:

$$\hat{H}(X) = \psi(N) - \psi(k) + \log c_{d_x} + \frac{d_x}{N} \sum_{i=1}^{N} \log \epsilon(i), \quad (6–25)$$

where $\epsilon(i)$ is twice the distance from $x_i$ to its $k$-th neighbor in the $d_x$ dimensional space.

For the joint entropy we have:

$$\hat{H}(X, Y) = \psi(N) - \psi(k) + \log(c_{d_x}c_{d_y}) + \frac{d_x + d_y}{N} \sum_{i=1}^{N} \log(\epsilon(i)). \quad (6–26)$$

The $I(X; Y)$ is now can be obtained by Eq.\((6–16)\). The problem with this method is that a fixed $k$ is used in all estimators but the distance metric in different scaled spaces (marginal and joint) are not comparable. To avoid such problem, instead of using a fixed $k$, $n_x(i) + 1$ and $n_y(i) + 1$ are used in obtaining the distances (where $n_x(i)$ and $n_y(i)$ are the number of samples contained the bin $[x(i) - \frac{\epsilon(i)}{2}, x(i) + \frac{\epsilon(i)}{2}]$ and $[y(i) - \frac{\epsilon(i)}{2}, y(i) + \frac{\epsilon(i)}{2}]$ respectively) in the $x$–$y$ scatter diagram. The Eq.\((6–26)\) becomes:

$$\hat{H}(X) = \psi(N) - \psi(n_x(i) + 1) + \log c_{d_x} + \frac{d_x}{N} \sum_{i=1}^{N} \log(\epsilon(i)). \quad (6–27)$$

Finally the Eq.\((6–16)\) is rewritten as:

$$I_{knnr}(X; Y) = \psi(k) + \psi(N) - \frac{1}{N} \sum_{i=1}^{N} [\psi(n_x(i) + 1) + \psi(n_y(i) + 1)]. \quad (6–28)$$

### 6.5 Nonlinear Interdependencies

Arnhold et al., (1999) introduced nonlinear the interdependence measures for characterizing directional relationships (i.e. driver & response) between two time sequences [106]. In this section, we investigate directional relationships using the nonlinear
interdependencies measures among different brain regions during before and after add-on AEDs treatment for the patient with ULD.

Given two time series \( x \) and \( y \), using method of delay to obtain delay vectors \( x_n = (x_n, \ldots, x_{n-(m-1)\tau}) \) and \( y_n = (x_n, \ldots, x_{n-(m-1)\tau}) \), where \( n = 1, \ldots, N \), \( m \) is the embedding dimension and \( \tau \) denotes the time delay [52]. Let \( r_{n,j} \) and \( s_{n,j} \), \( j = 1, \ldots, k \) denote the time indices of the \( k \) nearest neighbors of \( x_n \) and \( y_n \). For each \( x_n \), the mean Euclidean distance to its \( k \) neighbors is defined as

\[
R_n^k(X) = \frac{1}{k} \sum_{j=1}^{k} (x_n - x_{r_{n,j}})^2,
\]
and the \( Y \)-conditioned mean squared Euclidean distance is defined by replacing the nearest neighbors by the equal time partners of the closest neighbors of \( y_n \)

\[
R_n^{(k)}(X|Y) = \frac{1}{k} \sum_{j=1}^{k} (x_n - x_{s_{n,j}})^2.
\]

The delay \( \tau = 5 \) is estimated by auto mutual information function, the embedding dimension \( m = 10 \) is obtained using Cao’s method using 10 sec EEG selected during interictal state and a Theiler correction is set to \( T = 50 \) [73; 107].

If \( x_n \) has an average squared radius \( R(X) = (1/N) \sum_{n=1}^{N} R_n^{(N-1)}(X) \), then \( R_n^{(k)} \approx R_n^{(k)}(X) < R(X) \) if the system are strongly correlated, while \( R_n^{(k)}(X|Y) \approx R(X) > R_n^{(k)}(X) \) if they are independent. Accordingly, it can be define and interdependence measure \( S^{(k)}(X|Y) \) as

\[
S^{(k)}(X|Y) = \frac{1}{N} \sum_{n=1}^{N} \frac{R_n^{(k)}(X)}{R_n^{(k)}(X|Y)}.
\]

Since \( R_n^{(k)}(X|Y) \geq R_n^{(k)}(X) \) by construction,

\[
0 < S^{(k)}(X|Y) \leq 1
\]

Low values of \( S^k(X|Y) \) indicate independence between \( X \) and \( Y \), while high values indicate synchronization.
Arnhold et al., (1999) introduced another nonlinear interdependence measure
\[ H^{(k)}(X|Y) \]
as
\[ H^{(k)}(X|Y) = \frac{1}{N} \sum_{n=1}^{N} \log \frac{R_n(X)}{R_n^{(k)}(X|Y)} . \] (6–33)

\( H^{(k)}(X|Y) = 0 \) if \( X \) and \( Y \) are completely independent, while it is possible if closest in \( Y \) implies also closest in \( X \) for equal time indexes. \( H^{(k)}(X|Y) \) would be negative if close pairs in \( Y \) would correspond mainly to distant pairs in \( X \). \( H^{(k)}(X|Y) \) is linear measures thus is more sensitive to weak dependencies compare to mutual information. Arnhold et al., (1999) also showed \( H \) was more robust against noise and easier to interpret than \( S \).

Since \( H \) is not normalized Quiroga et al., (2002) introduced another \( N(X|Y) \):
\[ N^{(k)}(X|Y) = \frac{1}{N} \sum_{n=1}^{N} \frac{R_n(X) - R_n^{(k)}(X|Y)}{R_n(X)} , \] (6–34)
which is normalized between 0 and 1. The opposite interdependencies \( S(Y|X), H(Y|X) \), and \( N(Y|X) \) are defined in complete analogy and they are in general not equal to \( S(X|Y), H(X|Y) \), and \( N(X|Y) \), respectively. Using nonlinear interdependencies on several chaotic model (Lorenz, Roessler, and Heénon models) Quiroga et al., (2000) showed the measure \( H \) is more robust than \( S \).

The asymmetry of above nonlinear interdependencies is the main advantage over other nonlinear measures such as the mutual information or the phase synchronization. This asymmetry can give information about “driver-response” relationships but can also reflect different properties of dynamical systems when it is importance to detect causal relationships. It should be clear that the above nonlinear interdependencies measures are bivariate measures. Although it quantified the “driver-response” for given input-the whole input space under study might be driven by other unobserved system(s).

6.6 Discussions

It is believed that synchronization occurs due to both local and global discharges of the neurons. From the epilepsy perspective, quantifying the changes in spatiotemporal interactions could potentially lead to the development of seizure-warning systems and
detect the effect of therapeutic interventions. This quantification would also help us identify the regions that actively participate during epileptic seizures. In the later chapter, I investigate possibilities for identify the regions in the brain that are actively participate prior to epileptic seizures and the effect of therapeutic interventions on EEG recordings. Although it is known that the fluctuations in EEG frequency and voltage arise from spontaneous interactions between excitatory and inhibitory neurons in circuit loops, the cause of neuron discharge remains unclear. The synchronization of neuron activity is considered to be important for information processing in the developing brain. There is now clear evidence that there are distinct differences between the immature and mature brain in the pathophysiology and consequences of seizures, abnormal synchronization activity in the developing brain can result in irreversible alterations in neuronal connectivity [108]. In cognitive task studies, In 1980, Freeman found “more regular” spatiotemporal activities in EEG for a brief period of time when the animal inhaled a familiar odor until the animal exhaled [109; 96]. Some earlier studies also indicated the significant role of synchronization for physiological systems in humans; the detectable alters in synchronization phenomena have been associated to a number of chronic, acute diseases or the normality of brain. [110; 111]. In the field of epilepsy research, several authors have suggested direct relationship between alters in synchronization phenomena and onset of the epileptic seizures using EEG recordings. For example, Iasemidis et al., (1996) reported from intracranial EEG that the entrainment in the largest Lyapunov exponents from critical cortical regions is a necessary condition for onset of seizures for patients with temporal lobe epilepsy [112; 61]; Le Van Quyen et al., showed epileptic seizure can be anticipated by nonlinear analysis of dynamical similarity between recordings [35]. Mormann et al., showed the preictal state can be detected based on a decrease in synchronization on intracranial EEG recordings [89; 43]. The highly complex behavior on the EEG recordings is considered to normality of brain state, while transitions into a lower complexity brain state are regarded as a pathological
normality losses. Synchronization patterns were also found to differ somewhat depending on the epileptic syndrome, with primary generalized absence seizures displaying more long-range synchrony in all frequency bands studied (355 Hz) than generalized tonic motor seizures of secondary (symptomatic) generalized epilepsy or frontal lobe epilepsy [113]. In next chapters, we will use the concepts and mathematical frameworks introduced in this chapter to cluster the EEG between normal and abnormal epoches and to detect the effect in EEG resulted from AEDs treatments.
CHAPTER 7
CLUSTERING ELECTROENCEPHALOGRAM (EEG) SIGNALS TO STUDY MESIAL TEMPORAL LOBE EPILEPSY (MTLE)

7.1 Introduction

In this study, we propose a network analysis framework to study the evolution of epileptic seizures. We apply a signal processing approach, derived from information theory, to investigate the synchronization of neuronal activities, which can be captured by electroencephalogram (EEG) recordings. Two network-theoretic approaches are proposed to globally model the synchronization the brain network. We observe some unique patterns related to the development of epileptic seizures, which can be used to illuminate the brain function governed by the epileptogenic process during the period before a seizure. The proposed framework can provide a global structural patterns in the brain network and may be used in the simulation study of dynamical systems (like the brain) to predict oncoming events (like seizures). To analyze long-term EEG recordings in the future, we discuss how the Markov-Chain Monte Carlo (MCMC) methodology can be applied to estimate the clique parameters. This MCMC framework fits very well with this work as the epileptic evolution can be considered to be a system with unobservable state variables and nonlinearities.

Neural activity is manifested by electrical signals known as graded and action potentials. Berger’s demonstration in 1929 has shown that it is possible to record the electrical activity from the human brain, particularly the neurons located near the surface of the brain [23]. While we often think of electrical activity in neurons in terms of action potentials, the action potentials do not usually contribute directly to the electroencephalogram (EEG) recordings. In fact, for scalp EEG recordings, the EEG patterns are mainly the graded potentials accumulated from hundreds of thousands of neurons. The EEG patterns vary greatly in both amplitude and frequency. The amplitude of the EEG reflects the degree of synchronous firing of the neurons located around the recording electrodes. In general, the high EEG amplitude indicates that neurons are
activated simultaneously. Low EEG frequency indicates less responses of the brain, such as sleep, whereas higher EEG frequency implies the increased alertness. Given the above descriptions, an acquired EEG time series can be defined as a record of the fluctuating brain activity measured at different times and spaces. The high degree of synchronicity for two different brain regions implies strong connectivity among them and vice versa. We will interchangeably use the terms synchronicity and connectivity for rest of the chapter.

Although the brain may have originally emerged as an organ with functionally dedicated regions, recent evidence suggests that the brain evolved by preserving, extending, and re-combining existing network components, rather than by generating complex structures de novo [114; 115]. This is significant because it suggests: (1) the brain network is arranged such that the functional neural complexes supporting different cognitive functions share many low-level neural components, and (2) the specific connection topology of the brain network may play a significant role in seizure development. This line of thinking is also supported by [70], which demonstrates that specific connected structures are either significantly abundant or rare in cortical networks. If seizures evolve in this fashion, then we should be able to make some specific empirical hypotheses regarding the evolution of seizures, that might be borne out by investigating the synchronization between the activity in different brain areas, as revealed by quantitative analysis of EEG recordings. The goal of this study is to test the following two hypotheses. First, we should expect the brain activity in the orbitofrontal areas are highly correlated while the activity in the temporal lobe and subtemporal lobe areas are highly correlated with their own side (left only or right only) during the pre-seizure period. The high correlation can be viewed as a recruitment operation initiated by an epileptogenic area through a regular communication channel in the brain. Note that the connection of these brain areas has been a long-standing principle in normal brain functions and we believe that the same principle should hold in the case of epilepsy as well. Second, we should expect some brain regions to be consistently active, manifested
by a higher degree of synchronization among EEG electrodes within the same region, during the pre-seizure state. We postulate that the active connection may be driven by seizure evolution, regulating abnormal communications in the epileptogenic brain areas or vulnerable areas in the brain network. To test these hypotheses, we herein propose network-theoretical methods through a multivariate statistical analysis of EEGs to study the seizure development by investigating the topological structure of the brain connectivity network. Epileptic seizures involve the synchronization of large populations of neurons [116]. Measuring the connectivity and synchronicity among different brain regions through EEG recordings has been well documented [99; 117; 69]. The structures and the behaviors of the brain connectivity have been shown to contain rich information related to the functionality of the brain [118; 67; 68]. More recently, the mathematical principles derived from information theory and nonlinear dynamical systems have allowed us to investigate the synchronization phenomena in highly non-stationary EEG recordings. For example, a number of synchronization measures were used for analyzing the epileptic EEG recordings to reach the goals of localizing the epileptogenic zones and predicting the impending epileptic seizures [99; 106; 119; 38; 120]. These studies also suggest that epilepsy is a dynamical brain disorder in which the interactions among neuron or groups of neurons in the brain alter abruptly. Moreover, the characteristic changes in the EEG recordings have been shown to have clear associations with the synchronization phenomena among epileptogenic and other brain regions. When the conductivities between two or among multiple brain regions are simultaneously considered, the univariate analysis alone will not be able to carry out such a task. Therefore it is appropriate to utilize multivariate analysis. Multivariate analysis has been widely used in the field of neuroscience to study the relationships among sources obtained simultaneously. In this study, the cross mutual information (CMI) approach is applied to measure the connectivity among brain regions [75]. The CMI approach is a bivariate measure and has been shown to have ability for
quantifying the connectivity of the EEG signals [101; 97]. The brain connectivity graph is then constructed where vertices in the graph represent the EEG electrodes.

Every distinct pair of vertices is connected by an arc with the length equal to the connectivity quantified by CMI. After constructing a brain connectivity graph, which is a complete graph, we then remove arcs of connectivity below a specified threshold value to preserve only strong couplings of electrode pairs. Finally, we employ a maximum clique algorithm to find a maximum clique in which the brain regions are strongly connected. The maximum clique size can be, in turn, used to represent the amount of largest connected regions in the brain. The maximum clique algorithm reduces the computational effort for searching in the constructed brain connectivity graph. The proposed graph-theoretic approach offers an easy protocol for inspecting the structures of the brain connectivity over time and possibly identifying the brain regions where seizures are initiated.

7.2 Epilepsy as a Dynamical Brain Disorder

Epilepsy can be caused by multiple factors. Some people may even begin having seizures from their childhood. Epilepsy in children can result from almost everything related to the brain development or function. Lack of oxygen supply can cause cerebral palsy and seizure for example new born infants that suffer a lack of oxygen supply to the brain before or during birth have higher risks for developing epilepsy in their lives. Epilepsy can also occur in adult subjects that have bleeding in the brain as a result of prematurity or defective blood vessels in the brain. Some studies have also reported epilepsy can be induced by genetic changes. Some patients are born with genes related to epilepsy that can cause them to develop epilepsy for example the Unverricht-Lundborg disease. The majority of patients who have seizures are first treated with anti-epileptic drugs (AEDs). About 70% to 80% of these patients will become seizure free after the AEDs treatment. The choice of the AEDs depends on several factors, including the type of seizures, the age of the subject, and the potential side effects of the medicine. Some
patients, however, do not respond to the usual pharmacological treatments and will then consider undergoing the epilepsy surgery.

7.3 Data Information

In this study, the EEG recordings were obtained from bilaterally, surgically implanted macro electrodes in the hippocampus, temporal and frontal lobe cortexes of 2 patients who underwent pre-surgical clinical evaluation for possible surgical treatment of intractable temporal lobe epilepsy. The recordings were obtained using a Nicolet BMSI 4000 recording system with amplifiers of an input range of 0.6 mV, sampling rate of 200 Hz and filters with a frequency range of a 0.5–70 Hz. Each recording included a total of 26 to 32 intracranial electrodes (8 subdural and 6 hippocampal depth electrodes for each cerebral hemisphere, and a strip of 4 additional electrodes if deemed necessary by the neurologist). The recorded EEG signals were digitized and stored on magnetic media for subsequent off-line analysis.

Table 7-1. Patient information for clustering analysis

<table>
<thead>
<tr>
<th>Patient</th>
<th>Gender</th>
<th>Age</th>
<th>Number of electrodes</th>
<th>Seizure onset zone</th>
<th>Duration of EEG recordings (days)</th>
<th>Number of seizures</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Male</td>
<td>29</td>
<td>26</td>
<td>R. Hippocampus</td>
<td>6.07</td>
<td>19</td>
</tr>
<tr>
<td>2</td>
<td>Male</td>
<td>37</td>
<td>30</td>
<td>L./R. Hippocampus</td>
<td>9.88</td>
<td>11</td>
</tr>
</tbody>
</table>

7.3.1 Multivariate Analysis on EEG Signals

In any study of the brain connectivity network, signal processing and data mining techniques are required to extract useful information buried in the raw EEG data. We can categorize signal processing techniques into two types based on the number of sources, univariate and multivariate measures. Univariate measures process the information from a single data source, such as a single electrode. There are several signal properties one can extract using univariate measures such as power spectrum, autocorrelation, entropy, and divergent rates. Multivariate measures (also referred as spatio-temporal measures) allow one to determine the synchronization that commonly occurs among a group of sources. This is very important because multivariate measures can uniquely determine specific
types of connection between two or more sources, and quantifying the synchronization between different brain areas (measured by different electrodes) is crucial to a greater understanding of the brain connectivity network. The synchronization may be attributable to the brain’s anatomical, functional, or dynamical connectivity. In this study, the synchronization patterns are postulated to reflect the seizure evolution (epileptogenic process), and we shall use electrode synchronization as a similarity measure of EEG signals from different brain areas. This is fine in theory, however there are a few complexity issues in calculation of multivariate measures. First, in spite of the theoretical capability of multi-variate methods to calculate common patterns from several sources simultaneously, the calculation complexity increases exponentially with the number of sources.

Therefore, we use multivariate measures for quantifying the synchronization from only 2 electrodes at a time. Specifically, a simple signal processing used to calculate the synchronization between electrode pairs is employed in this study. Then we apply a data mining technique based on network-theoretical methods to the multivariate analysis of EEG data.

7.3.2 Brain Synchronization

In general, statistical similarity measures can be categorized into two groups: linear and nonlinear dependence measures. The linear measure is mainly used for measuring a linear relationship between two or more time series. For example, the most commonly used measure is cross-correlation function, which is a standard method of estimating the degree of correlation in time domain between two time series. The result of a cross correlation function can be calculated at different time lags of two time series to show the level of redundancy at different time points. Frequency coherence is another linear similarity measure, which calculates the synchrony of activities at each frequency [121]. Although the information from cross-correlation function and frequency coherence has been shown to be identical [122], the similarity between two EEG signals in different frequency bands such as delta, theta, beta, alpha and gamma, is still commonly used to investigate EEG
similarity patterns [123; 121]. For example, [124] used frequency coherence measures to investigate the interactions between medial limbic structures and the neocortex during ictal periods (seizure onsets). In another study by [125], the coherence pattern of cortical areas from epileptic brain was investigated to identify a cortical epileptic system during interictal (normal) and ictal (seizure) periods.

Although linear measures are very useful and commonly used, they are insensitive to nonlinear coupling between signals, and non-linearities are quite common in neural contexts. To be able to investigate more of the interdependence between EEG electrodes, nonlinear measures should be applied. Nonlinear measures have been widely used to determine the interdependence among EEG signals from different brain areas. For example, [106] and [35] studied the similarity between EEG signals using nonlinear dynamical system approaches. They applied a time-delay embedding technique to reconstruct a trajectory of EEG in phase space and used the idea of generalized synchronization proposed by [126] to calculate the interdependence and causal relationships of EEG signals.

We propose an approach to investigate and quantify the synchronization of the brain network, specifically tailored to study the propagation of epileptogenic processes. [127] investigated this propagation, where the average amount of mutual information during the ictal period (seizure onset) was used to identify the focal site and study the spread of epileptic seizure activity. Subsequently, [128] applied the information-theoretic approach to measure synchronization and identify causal relationships between areas in the brain to localize an epileptogenic region. Here, we apply an information-theoretic approach, called cross-mutual information, which can capture both linear and nonlinear dependence between EEG signals, to quantify the synchronization between nodes in the brain network. In order to globally model the brain network, we represent the brain synchronization network as a graph.
7.4 Graph-Theoretic Modeling for Brain Connectivity

Modeling the brain network as a graph is not new. In the past decade, several studies have attempted to use network-theoretic methods to study the brain network. For example, the topological relationships between brain networks and social networks were proposed by [70]. In an earlier study, [115] demonstrated that the brain evolved a highly efficient network architecture whose structural connectivity (or motif) is capable of generating a large repertoire of functional states. In another recent study, the brain network graph was investigated to verify that the re-use of existing neural components played a significant role in the evolutionary development of cognition [114].

Applying network/graph-theoretic methods to EEG signals, we can model the brain connectivity/synchronization network as a complete graph $G(V, E)$, where $V$ is a set of vertices and $E$ is a set of edges. Vertices (also called nodes) are represented by EEG electrodes (also referred as channels). Edges (also called arcs) are represented by the synchronization/similarity between 2 EEG electrodes whose degrees correspond to the edge weights. In short, a brain connectivity network can then be constructed as a graph whose vertices are EEG electrodes and the weighted edges are the coupling strength of electrode pairs. Every pair of vertices is connected by a weighted edge. In this study, we focus on the structural changes in the brain connectivity network that may be related to the seizure evolution. The structural changes could be represented by connectivity fractions/partitions through aggregation and segregation of the brain network. In this study, we propose two network-theoretic approaches, spectral partitioning and maximum clique, to identify independent/segregated and clustered brain areas.

7.4.1 Cross–Mutual Information (CMI)

The concept of CMI dates back to the work of Shannon in 1948[98]. Generally, CMI measures the information obtained from observations of one random event for the other. It is known that CMI has the capability to capture both linear and nonlinear relationships between two random variables since both linear and nonlinear relationships
can be described through probabilistic theories. Here in our model, the CMI measures how much information of EEG time series acquired from electrode \( x \) is presented by electrode \( y \) and vice versa. Let \( X \) be the set of data points where its possible realizations are \( x_1, x_2, x_3, ..., x_n \) with probabilities \( P(x_1), P(x_2), P(x_3), ... \). The Shannon entropy \( H(X) \) of \( X \) is defined as

\[
H(X) = -\sum_{i=1}^{n} p_i \ln p_i. \tag{7-1}
\]

Shannon entropy measures the uncertainty content of \( X \). It is always positive and measured in bits, if the logarithm is taken with base 2. Now let us consider another set of data points \( Y \), where all possible realizations of \( Y \) are \( y_1, y_2, y_3, ..., y_n \) with probabilities \( P(y_1), P(y_2), P(y_3), ... \). The degree of synchronicity and connectivity between \( X \) and \( Y \) can be measured by the joint entropy of \( X \) and \( Y \), defined as

\[
H(X, Y) = -\sum_{i,j} p_{ij}^{XY} \ln p_{ij}^{XY}. \tag{7-2}
\]

where \( p_{ij}^{XY} \) which is the joint probability of \( X = X_i \) and \( Y = Y_j \). The cross information between \( X \) and \( Y \), \( CMI(X, Y) \), is then given by

\[
CMI(X, Y) = H(Y) - H(X|Y) = H(X) - H(Y|X) \tag{7-3}
\]

\[
= H(X) + H(Y) - H(X, Y) \tag{7-4}
\]

\[
= \int \int p_{XY}(x, y) \log_2 \frac{p_{XY}(x, y)}{p_X(x)p_Y(y)} dxdy. \tag{7-5}
\]

The cross mutual information is nonnegative. If these two random variables \( X, Y \) are independent, \( f_{XY}(x, y) = f_X(x)f_Y(y) \), then \( CMI(X, Y) = 0 \), which implies that there is no correlation between \( X \) and \( Y \). The probabilities are estimated using the histogram based box counting method. The random variables representing the observed number of pairs of point measurements in histogram cell \( (i, j) \), row \( i \) and column \( j \), are respectively \( k_{ij}, k_i \) and \( k_j \). Here, we assume the probability of a pair of point measurements outside the area covered by histogram is negligible, therefore \( \sum_{i,j} P_{ij} = 1 \) [75; 129].
Figure 7-1. EEG epochs for $RTD_2$, $RTD_4$ and $RTD_6$ (10 seconds)

Figure 7-2. Scatter plot for EEG epoch (10 seconds) of $RTD_2$ vs. $RTD_4$ and $RTD_4$ vs. $RTD_6$

Figure 7-3 shows the CMI values measured from right mesial temporal depth ($RT(D)$) regions. Figure 7-2 displays the scatter plots for the EEG recorded from the same ($RT(D)$) brain region. From the scatter plot, it is clear that EEG recordings between $RT(D)2$ and $RT(D)4$ have weak linear correlation which have also yielded lower CMI values in Figure 7-3. The stronger linear relationship is discovered between $RT(D)4$ and $RT(D)6$ and this linear correlation pattern has resulted in higher CMI values. Prior to measuring the CMI, we first divided EEG recordings into smaller non-overlapping EEG epochs. The segmentation procedure is widely utilized to subdue the non-stationary nature of the EEG recordings. The changes of EEG pattern tend to appear very briefly,
examples include sharp wave transients, spikes, spike-wave complexes, and spindles. Working on shorter EEG epochs will insure the stationarity for the underlying processes and thus any change in the connectivity can be detected. Therefore, a proper length of EEG epochs has to be determined for measuring the connectivity among EEG recordings. We chose the length of the EEG epochs equal to 10.24 seconds (2048 points), which has also been utilized in many pervious EEG research studies [58; 130]. The brain connectivity measured using CMI form the complete graph, in which each node has an arc to every other adjacent vertex. In the procedure for removing the insignificant arcs (weak connection between brain regions), we first estimated an appropriate threshold value by utilizing the statistical tests. We determined this threshold by observing the statistical significance over the complete connectivity graph, this threshold value was set to be a value where the small noise is eliminated, but yet the real signal is not deleted [131].
7.4.2 Spectral Partitioning

In the spectral partitioning procedure, we applied the normalized cut to partition graph into two natural partitions (clusters), in which two groups of electrodes can be separated into two highly synchronized groups. For visualization purposes, a 10-second segment of the similarity matrix ($W$) can be represented by a two-dimensional bitmap shown in Figures 7-5(a) and 7-5(b).

Note that this matrix is symmetric because the mutual information measure has no coupling direction. In each row and column of this bitmap, the color represents the synchronization level. Note that we will ignore the diagonal of the matrix because we can always find a very high level of self-synchronization.

After applying the normalized cut, we calculated an eigenvector corresponding to the second smallest eigenvalue. Subsequently, we separated electrodes into two groups with the minimum cut or separation with minimum cost by applying the threshold value at 0. Using the eigenvector in Figure 7-5c, electrodes were separated into two clusters. It is easy to observe a clear separation of these two clusters through the value of eigenvector, in which a sharp transient from R(S)T4 to L(O)F1 is used as a separating point. The first group of synchronized electrodes is from L(S)T, L(T)D, R(S)T and R(T)D areas. The second group is from L(O)F and R(O)F areas. After the
first iteration, it is clear that the synchronization in the LD-LT-RD-RT cluster is not uniform throughout all electrodes in the cluster. Therefore, we consequently performed another iteration of spectral partitioning on the LD-LT-RD-RT cluster to find highly synchronized groups of electrodes within the cluster. This procedure can be viewed as a hierarchical clustering. After rearranging the electrodes based on the synchronization level, we found two sub-clusters of electrodes in the bitmap shown in Figure 7-5b. As shown in Figure 7-5d, the value of eigenvector indicates that there are two sub-clusters within the LD-LT-RD-RT cluster by applying the threshold of 0. The LD-LT-RD-RT cluster was separated into two sub-clusters: LD-LT and RD-RT. This observation suggests that there exists a highly synchronized pattern in the same side of temporal lobe as well as in the entire orbitofrontal area. This finding can be considered as a proof of concept that the seizure evolution also follows a regular communication pattern in the brain network.
7.4.3 Maximum Clique Algorithm

In this section, we discuss the results from analyzing the structural property of the brain network using the maximum clique approach. As mentioned earlier, the idea of applying the maximum clique technique is different from the one using the spectral partitioning approach as we are only interested in the most highly synchronized group of electrodes in the brain network. We adopted the algorithm to find a maximum clique in the brain connectivity graph after deleting the insignificant arcs in the original complete graph as follows: Let $G = G(V, E)$ be a simple, undirected graph where $V = \{1, \ldots, n\}$ is the set of vertices (nodes), and $E$ denotes the set of arcs. Assume that there is no parallel arcs (and no self-loops joining the same vertex) in $G$. Denote an arc joining vertex $i$ and $j$ by $(i, j)$. We define a clique of $G$ as a subset $C$ of vertices with the property that every pair of vertices in $C$ is connected by an arc; that is, $C$ is a clique if the subgraph $G(C)$ induced by $C$ is complete. Then, the maximum clique problem is to find a clique $C$ with maximum cardinality (size) $|C|$. The maximum clique problem can be represented in many equivalent formulations (e.g., an integer programming problem, a continuous global optimization problem, and an indefinite quadratic programming). In this paper, we represent it in a simple integer programming form given by

We analyzed 3 epochs of 3-hour EEG recordings, 2 hours before and 1 hour after a seizure, from Patient 2 who had the epileptogenic areas on both right and left mesial temporal lobes. Figures 7-7 and 7-8 demonstrate the electrode selection of the maximum clique group during two hours before and one hour after the seizure onset. During the period before the seizure onset, both figures manifested a pattern where all the LD electrodes were consistently selected to be in the maximum clique. During the seizure onset, the size of the maximum clique increases drastically. This is very intuitive because, in temporal lobe epilepsy, all of the brain areas are highly synchronized. We visually inspected the raw EEG recordings before and during the seizure onsets and found a similar semiological pattern of the seizure onset - electrodes from the L(T)D areas.
algorithm: maximum clique
begin
    sort all nodes based on vertex ordering
    LIST = ordered nodes
    cbc = 0 current best clique size
    depth = 0 current depth level
    enter-next-depth(LIST, depth)
end

procedure: enter-next-depth(LIST, depth)
begin
    1 m = the number of nodes in the LIST
    2 depth = depth + 1
    3 for a node in position i in the LIST
    4 if depth + (m - i) ≤ cbc then
        5 return prune the search
    6 else
        7 mark node i
        8 if no adjacent node then
            9 cbc = depth (maximum clique found)
        10 else
            11 enter to next depth (adjacent node of i, depth)
        12 end
    13 end
    14 unmark node i
    15 if depth = 1
    16 delete node i from LIST
    17 end
    18 end
end

Figure 7-6. Maximum clique algorithm

initiated a highly organized rhythmic patterns and the patterns started to propagate throughout all the brain areas. We initially speculated that the epileptogenic areas could be the ones that are highly synchronized long before a seizure onset. In the previous case, we observed that the L(T)D electrodes are the one that started the seizure evolution. However, in a further investigation of EEG recordings from the same patient, we found some contrast results. In Figure 7-9, the electrode selection pattern of the maximum clique demonstrates a very highly synchronized group of electrodes in both left and right orbitofrontal areas during the 2-hour period preceding the seizure. After visual inspection on the raw EEG recordings, this seizure was initiated by the R(T)D area. Generally, it is
known that the orbitofrontal areas communicate with each other more than other parts of the brain. This has led us to the conclusion that the brain areas that are selected to be in the maximum clique are the vulnerable brain areas, rather than the epileptogenic areas. In other words, the brain area(s) that are highly synchronized could be governed or manipulated by the epileptogenic areas so that they continuously show strong neuronal communication through the synchronization of EEG signals (measured by cross-mutual information).

Figure 7-7. Electrode selection using the maximum clique algorithm for Case 1

7.4.4 Implications of the Results

In normal brain functions, the orbitofrontal areas (both left and right) of the brain are highly synchronized active most of the time as it is considered to be the brain’s executive function, and the temporal lobe areas are separated into left and right cortical hemispheres that work independently from each other. We hypothesized that this operation in the brain should be applied to the epileptic brains, even in the pre-seizure period. As we predicted, from the spectral partitioning results, both left and right orbitofrontal areas were also highly synchronized and active as well as right and left temporal lobe areas during the pre-seizure state. This suggests that the epileptogenic
processes slowly develop themselves through a regular communication channel in the brain network, rather than abruptly disrupt, collapse, or change the way brains communicate. From this observation, we postulate that this phenomenon may be a reflection of neuronal recruitment in seizure evolution. This observation confirms our first hypothesis. In
addition, we have found that nodes in the brain network are clustered during the seizure evolution. Most brain areas seem to be communicating with their physiological neighbors during the process. The key process of seizure evolution could be the step where the epileptogenic area(s) govern or manipulate the other vulnerable, or easily synchronized, brain areas to communicate with their neighbors. This can be viewed as a recruitment of other brain areas done by the epileptogenic area(s). In most cases, the recruitment of seizure development should start with a weaker group, which in our case is represented by a vulnerable brain area. After enough neurons have been recruited, the disorders of epileptic brains spread out abnormal functions from than localized areas of cortex or other vulnerable areas throughout the cortical networks and the entire brain network. This phenomenon was shown by the results of our maximum clique approach, which confirms our second hypothesis. In addition, a different type of maximum clique patterns may be useful in the identification of incoming seizures. This study suggests that, in the future, this framework may be used as a tool to provide practical seizure interventions. For example, one can locate and stimulate the brain areas that seem to be vulnerable to the seizure evolution by electrical pulses through the monitoring process of the maximum clique. This will drastically reduce the risk of seizure to epilepsy patients.

7.5 Discussion and Future Work

In this study, we attempted to study seizure evolution by investigating some neuronal interactions among different brain areas. Analyzing multidimensional time series data like multichannel EEG recordings is a very complex process. The study of the brain network needs to involve the neuronal activities from not only a single source or a small group of sources, but also the entire brain network. Here we applied the cross-mutual information technique, a measure widely used in the information theory, to capture the neuronal interactions through the brain’s synchronization patterns. Then we modeled the global interactions using network/graph-theoretic approaches, spectral partitioning and maximum clique. These approaches are used to generalize the brain network investigation
to capture synchronization patterns among different sources (brain areas). The idea of analyzing EEG recordings from several sources (multiple electrodes) is very crucial since the knowledge from local information (i.e., single electrode) is very limited. In our future study, we plan to incorporate the knowledge of general brain communication in the brain network. For example, [114] demonstrated the evolution of cognitive function through quantitative analysis of fMRI data.

The proposed framework can provide a global structural patterns in the brain network and may be used in the simulation study of dynamical systems (like the brain) to predict oncoming events (like seizures). For example, an ON-OFF pattern of electrode selection in the maximum clique over one period of time can be modeled as a binary observation in a discrete state in a Markov model, which can be used to simulate the seizure evolution in the brain. In addition, the number of electrodes in the maximum clique can be used to estimate the minimum number of features and explain dynamical models or the parameters in time series regression. Note that the proposed network model represents an epileptic brain as a graph, where there exist several efficient algorithms (e.g., maximum clique, shortest path) for finding special structure of the graph. This idea has enabled us, computationally and empirically, to study the evolution of the brain as a whole. The Monte-Carlo Markov Chain (MCMC) framework may be applicable in our future study on long term EEG analysis. The MCMC framework has been shown very effective in data mining research [132]. It can be used to estimate the graph or clique parameters in epileptic processes from EEG recordings. Since long term EEG recordings are very massive, most simulation techniques are not scalable enough to investigate large-scale multivariate time series like EEGs. The use of MCMC makes it possible to approximate the brain structure parameters over time. More importantly, the MCMC framework can also be extended to the analysis of multi-channel EEGs by generating new EEG data points while exploring the data sequences using a Markov chain mechanism. In addition, we can integrate the MCMC framework with a Baysian approach. This can be
implemented in on-line simulation-based brain clique estimation scheme, which employs sequential sampling, electrode selection (maximum clique), and MCMC moves. Although the implementation of this MCMC framework remains to be further investigated, we expect that this framework will be very fast and efficient.
Assessing the severity of myoclonus and evaluating the efficacy of antiepileptic drugs (AEDs) treatment for patients with Unverricht-Lundborg Disease (ULD) have traditionally utilized the Unified Myoclonus Rating Scale (UMRS). EEG recordings are only used as a supplemental tool for the diagnosis of epilepsy disorders. In this study, mutual information and nonlinear interdependence measures were applied on EEG recordings to identify the effect of treatment on the coupling strength and directionality of information transport between different brain cortical regions. Two 1-hour EEG recording were acquired from four ULD subjects during the time period of before and after treatment. All subjects in this study are from the same family with similar age (48± 3 years) and ULD history (~37.75 years). Our results indicate that the coupling strength was low between different brain cortical regions in the patients with less severity of ULD. The effects of the treatment was associated with significant decrease of the coupling strength. The information transport between different cortex regions were reduced after treatment. These findings could provide a new insight for developing a novel surrogate outcome measure for patients with epilepsy when clinical observations could potentially fail to detect a significant difference.

EEG recording system has been the most used apparatus for the diagnosis of epilepsy and other neurological disorders. It is known that changes in EEG frequency and amplitude arise from spontaneous interactions between excitatory and inhibitory neurons in the brain. Studies into the underlying mechanism of brain function have suggested the importance of the EEG coupling strength between different cortical regions. For example, the synchronization of EEG activity has been shown in relation to memory process [133; 134] and learning process of the brain [135]. In a pathophysiological study, different brain synchronization/desynchronization EEG patterns are shown to be induced
by hippocampal atrophy in subjects with mild cognitive impairment, therefore EEG recordings could represent a tool for differential diagnosis [136].

In epilepsy research, several authors have suggested direct relationship between change in synchronization phenomena and onset of epileptic seizures using EEG recordings. For example, Iasemidis et al., reported from intracranial EEG that the nonlinear dynamical entrainment from critical cortical regions is a necessary condition for onset of seizures for patients with temporal lobe epilepsy [112; 37; 60]; Le Van Quyen et al., showed epileptic seizure can be anticipated by nonlinear analysis of dynamical similarity between recordings [35]; Mormann et al., showed the preictal state can be detected based on a decrease in synchronization on intracranial EEG recordings [89; 43]. The highly complex behavior on EEG recordings is considered as normality of brain state, while transitions into a lower complexity brain state are regarded as a pathological normality losses. Synchronization patterns were also found to differ somewhat depending on epileptic syndromes, with primary generalized absence seizures displaying more long-range synchrony in frequency bands (3-55 Hz) than generalized tonic motor seizures of secondary (symptomatic) generalized epilepsy or frontal lobe epilepsy [113].

In this study, mutual information and nonlinear interdependence measures were applied on the EEG recordings to identify the effect of treatment on the coupling strength and directionality of information transport between different brain cortical regions [137; 100–102; 104]. The EEG recordings were obtained from patients with ULD.

ULD is one type of Progressive Myoclonic Epilepsy (PME); a rare epilepsy disorder with complex inheritance. ULD was first described by Unverricht in 1891 and Lundborg in 1903 [138; 139]. AEDs is mainstay for the treatment of ULD with overall unsatisfactory efficacy. Due to the progression of the severity of myoclonus, the efficacy of AEDs treatment is difficult to be determined clinically especially in the later stages of the disease. The EEG recordings for ULD subjects usually demonstrate abnormal slow background, generalized high-amplitude 3–5Hz spike waves or poly spike and wave
complexes. Normal background EEG can be observed between generalized spike and wave discharges. Studies have shown increasing background slowing of EEG or no change in patients with more advanced disease stages for patients with ULD [140; 141]. Furthermore, AEDs associated generalized slowing of EEG background rhythms has been reported with highly variable from patient to patient [142]. However, it was difficult to determine how strong the correlation between EEG slowing and disease progression was since the intensification of drug treatment during the later stages of illness might have contributed to the EEG slowing [140].

The clinical observations have been the most common method for evaluating the influence and effectiveness of AEDs interventions in patients with epilepsy and other neurological disorders. More specifically, efficacy of treatment is usually measured by comparing the seizure frequency during treatment to a finite baseline period. EEG recordings are mainly used as supplemental diagnostic tools in medical treatment evaluations. Other than counting the number of seizures as a measure for treatment effect there is currently no reliable tool for evaluating treatment effects in patients with seizure disorders. A quantitative surrogate outcome measure using EEG recordings for patients with epilepsy is desired.

This rest of this chapter is organized as follows. The background of the patients and the parameters of the EEG recordings are given in section II. In section III. The methods for identify the coupling strength and directionality of information transport between different brain cortical regions are described. The quantitative analysis, statistical tests and results are presented in Section IV. The conclusion and discussion are given in Section V.

8.2 Data Information

The EEG recordings were acquired using a Nicolet BMSI recording system and the international 10-20 electrode placement system (Fp1, Fp2, F3, F4, C3, C4, A1, A2, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, Fz, Cz, and Pz). The acquired EEG recordings
was digitized using a 12-bit analog-digital converter. The EEG recordings were band pass filtered at 0.1-70Hz. The sample frequency was set to 250Hz. An additional 60Hz notch filter was applied to reduce the artifact induced by electronic devices in the EEG recordings. In this study, 1-hour EEG recordings were obtained before and after a new AEDs treatment was added. All the EEG recordings were reviewed by a board certified electroencephalographer and artifact-free baseline EEG segments were selected for the quantitative analysis. The information of the ULD subjects in this study are summarized in Table 1. The EEG recordings were recorded approximately at the same time of a day with subjects in a relaxed state; one set of EEG recordings was acquired before treatment started and the other set was acquired after at least four weeks after a new AEDs treatment was added. In this study, the severity of the ULD patients were evaluated by performing the Unified Myoclonus Rating Scale (UMRS), a statistically validated clinical rating instrument for evaluating individuals with myoclonus.

Table 8-1. ULD patient information

<table>
<thead>
<tr>
<th>Patient</th>
<th>Gender</th>
<th>Age</th>
<th>ULD Onset age</th>
<th>UMRS score (Before)</th>
<th>UMRS score (After)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Female</td>
<td>47</td>
<td>9</td>
<td>98</td>
<td>48</td>
</tr>
<tr>
<td>2</td>
<td>Male</td>
<td>45</td>
<td>10</td>
<td>80</td>
<td>65</td>
</tr>
<tr>
<td>3</td>
<td>Male</td>
<td>50</td>
<td>12</td>
<td>50</td>
<td>66</td>
</tr>
<tr>
<td>4</td>
<td>Male</td>
<td>51</td>
<td>11</td>
<td>68</td>
<td>54</td>
</tr>
</tbody>
</table>

8.3 Synchronization Measures

8.3.1 Mutual Information

In this section, we describe the estimation of mutual information[103]. Let us denote the time series of two observable variables as $X = \{x_i\}_{i=1}^N$ and $Y = \{y_j\}_{j=1}^N$, where $N$ is the length of the series and the time between consecutive observations (i.e., sampling period) is fixed. Then the mutual information is given by:

$$I(X;Y) = \sum_i \sum_j P_{x,y}(x_i, y_j) \log \left( \frac{P_{x,y}(x_i, y_j)}{P_x(x_i)P_y(y_j)} \right).$$  \hspace{1cm} (8-1)
One can obtain the mutual information between $X$ and $Y$ using the following equation [143]:

$$I(X; Y) = H(X) + H(Y) - H(X, Y), \quad (8-2)$$

where $H(X), H(Y)$ are the entropies of $X, Y$ and $H(X, Y)$ is the joint entropy of $X$ and $Y$. Entropy for $X$ is defined by:

$$H(X) = -\sum_i p(x_i) \log p(x_i). \quad (8-3)$$

The units of the mutual information depends on the choice on the base of logarithm. The natural logarithm is used in the study therefore the unit of the mutual information is $\text{nat}$. We first estimate $\hat{H}(X)$ from $X$ by

$$\hat{H}(X) = -\frac{1}{N} \sum_{i=1}^{N} P(X = x_i). \quad (8-4)$$

For $X$ and $Y$ time series we define $d_{ij}^{(x)} = \|x_i - x_j\|, d_{ij}^{(y)} = \|y_i - y_j\|$ as the distances for $x_i$ and $y_i$ between every other point in matrix spaces $X$ and $Y$. One can rank these distances and find the $knn$ for every $x_i$ and $y_i$. In the space spanned by $X, Y$, similar distance rank method can be applied for $Z = (X, Y)$ and for every $z_i = (x_i, y_i)$ one can also compute the distances $d_{ij}^{(z)} = \|z_i - z_j\|$ and determine the $knn$ according to some distance measure. The maximum norm is used in this study:

$$d_{ij}^{(z)} = \max\{\|x_i - x_j\|, \|y_i - y_j\|\}, \quad d_{ij}^{(x)} = |x_i - x_j|. \quad (8-5)$$

Next let $\frac{\epsilon(i)}{2}$ be the distance between $z_i$ and its $k^{th}$ neighbor. In order to estimate the joint probability density function ($p.d.f.$), we consider the probability $P_k(\epsilon)$ which is the probability that for each $z_i$ the $k^{th}$ nearest neighbor has distance $\frac{\epsilon(i)}{2} \pm d\epsilon$ from $z_i$. This $P_k(\epsilon)$ represents the probability for $k - 1$ points have distance less than the $k^{th}$ nearest neighbor and $N - k - 1$ points have distance greater than $\frac{\epsilon(i)}{2}$ and $k - 1$ points have
distance less than $\varepsilon_i$. $P_k(\varepsilon)$ is obtained using the multinomial distribution:

$$P_k(\varepsilon) = k \binom{N-1}{k} \frac{dp_i(\varepsilon)}{d\varepsilon} \varepsilon_i^{k-1}(1 - p_i)^{N-k-1}, \quad (8-6)$$

where $p_i$ is the mass of the $\varepsilon$-ball. Then the expected value of $\log p_i$ is

$$E(\log p_i) = \psi(k) - \psi(N), \quad (8-7)$$

where $\psi(\cdot)$ is the digamma function:

$$\psi(t) = \Gamma(t)^{-1} \frac{d\Gamma(t)}{dt}, \quad (8-8)$$

where $\Gamma(\cdot)$ is the gamma function. It holds when $\psi(1) = C$ where $C$ is the Euler-Mascheroni constant ($C \approx 0.57721$). The mass of the $\varepsilon$-ball can be approximated (if considering the p.d.f inside the $\varepsilon$-ball is uniform) as

$$p_i(\varepsilon) \approx c_d\varepsilon_i^d P(X = x_i), \quad (8-9)$$

where $c_{d_x}$ is the mass of the unit ball in the $d_x$-dimensional space. From Eq.(8-9) we can find an estimator for $P(X = x_i)$

$$\log[P(X = x_i)] \approx \psi(k) - \psi(N) - dE(\log \varepsilon(i)) - \log c_{d_x}, \quad (8-10)$$

finally with Eq(8-10) and Eq(8-4) we obtain the Kozachenko-Leonenko entropy estimator for $X$ [105]

$$\hat{H}(X) = \psi(N) - \psi(k) + \log c_{d_x} + \frac{d_x}{N} \sum_{i=1}^{N} \log \varepsilon(i), \quad (8-11)$$

where $\varepsilon(i)$ is twice the distance from $x_i$ to its $k$-th neighbor in the $d_x$ dimensional space.

For the joint entropy we have

$$\hat{H}(X, Y) = \psi(N) - \psi(k) + \log(c_{d_x}c_{d_y}) + \frac{d_x + d_y}{N} \sum_{i=1}^{N} \log(\varepsilon(i)). \quad (8-12)$$

The $I(X; Y)$ is now readily to be estimated by Eq.(8-2). The problem with this estimation is that a fixed number $k$ is used in all estimators but the distance metric in
different scaled spaces (marginal and joint) are not comparable. To avoid this problem, instead of using a fixed \( k, n_x(i) + 1 \) and \( n_y(i) + 1 \) are used in obtaining the distances (where \( n_x(i) \) and \( n_y(i) \) are the number of samples contained the bin \([x(i) - \frac{\epsilon(i)}{2}, x(i) + \frac{\epsilon(i)}{2}]\) and \([y(i) - \frac{\epsilon(i)}{2}, y(i) + \frac{\epsilon(i)}{2}]\) respectively) in the \( x-y \) scatter diagram. The Eq.(8–12) becomes:

\[
\hat{H}(X) = \psi(N) - \psi(n_x(i) + 1) + \log c_d + \frac{d_x}{N} \sum_{i=1}^{N} \log \epsilon(i). \quad (8–13)
\]

Finally the Eq.(8–2) is rewritten as:

\[
I_{knnr}(X; Y) = \psi(k) + \psi(N) - \frac{1}{N} \sum_{i=1}^{N} [\psi(n_x(i) + 1) + \psi(n_y(i) + 1)]. \quad (8–14)
\]

### 8.3.2 Nonlinear Interdependencies

Arnhold et al., (1999) introduced the nonlinear interdependence measures for characterizing directional relationships (i.e. driver & response) between two time sequences [106]. Given two time series \( x \) and \( y \), using method of delay we obtain the delay vectors:

\[ x_n = (x_n, ..., x_{n-(m-1)\tau}) \] and \( y_n = (x_n, ..., x_{n-(m-1)\tau}) \), where \( n = 1, ... N, m \) is the embedding dimension and \( \tau \) denotes the time delay [52]. Let \( r_{n,j} \) and \( s_{n,j} \), \( j = 1, ..., k \) denote the time indices of the \( k \) nearest neighbors of \( x_n \) and \( y_n \). For each \( x_n \), the mean Euclidean distance to its \( k \) neighbors is defined as

\[
R_{kn}^k(X) = \frac{1}{k} \sum_{j=1}^{k} (x_n - x_{r_{n,j}})^2, \quad (8–15)
\]

and the \( Y \)-conditioned mean squared Euclidean distance is defined by replacing the nearest neighbors by the equal time partners of the closest neighbors of \( y_n \)

\[
R_{kn}^{(k)}(X|Y) = \frac{1}{k} \sum_{j=1}^{k} (x_n - x_{s_{n,j}})^2. \quad (8–16)
\]

The delay \( \tau = 5 \) is estimated using auto mutual information function, the embedding dimension \( m = 10 \) is obtained using Cao’s method and the Theiler correction is set to \( T = 50 \) [73; 107].
If $x_n$ has an average squared radius $R(X) = (1/N) \sum_{n=1}^{N} R^{(n-1)}(X)$, then $R^{(k)}_n(X) \approx R^{(k)}_n(X) < R(X)$ if the system are strongly correlated, while $R^{(k)}_n(X|Y) \approx R(X) > R^{(k)}(X)$ if they are independent. Accordingly, it can be define and interdependence measure $S^{(k)}(X|Y)$ as

$$S^{(k)}(X|Y) = \frac{1}{N} \sum_{n=1}^{N} \frac{R^{(k)}_n(X)}{R^{(k)}_n(X|Y)}.$$  \hspace{1cm} (8–17)

Since $R^{(k)}_n(X|Y) \geq R^{(k)}_n(X) \text{ by construction}$,

$$0 < S^{(k)}(X|Y) \leq 1$$ \hspace{1cm} (8–18)

Low values of $S^k(X|Y)$ indicate independence between $X$ and $Y$, while high values indicate synchronization.

Arnhold et al., (1999) introduced another nonlinear interdependence measure $H^{(k)}(X|Y)$ as

$$H^{(k)}(X|Y) = \frac{1}{N} \sum_{n=1}^{N} log \frac{R_n(X)}{R^{(k)}_n(X|Y)}.$$ \hspace{1cm} (8–19)

$H^{(k)}(X|Y) = 0$ if $X$ and $Y$ are completely independent, while it is possible if closest in $Y$ implies also closest in $X$ for equal time indexes. $H^{(k)}(X|Y)$ would be negative if close pairs in $Y$ would correspond mainly to distant pairs in $X$. $H^{(k)}(X|Y)$ is linear measures thus is more sensitive to weak dependencies compare to mutual information. Arnhold et al., (1999) also showed $H$ was more robust against noise and easier to interpret than $S$.

Since $H$ is not normalized Quiroga et al., (2002) introduced another $N(X|Y)$:

$$N^{(k)}(X|Y) = \frac{1}{N} \sum_{n=1}^{N} \frac{R_n(X) - R^{(k)}_n(X|Y)}{R_n(X)},$$ \hspace{1cm} (8–20)

which is normalized between 0 and 1. The opposite interdependencies $S(Y|X)$, $H(Y|X)$, and $N(Y|X)$ are defined in complete analogy and they are in general not equal to $S(X|Y)$, $H(X|Y)$, and $N(X|Y)$, respectively. Using nonlinear interdependencies on several chaotic model (Lorenz, Roessler, and Heénon models) Quiroga et al., (2000) showed the measure $H$ is more robust than $S$. 

144
The asymmetry of above nonlinear interdependencies is the main advantage over other synchronization measures. This asymmetry property can give directionality of information transport between different cortical regions. Furthermore the “driver-response” relationships but can also reflect different properties of brain functions when it is important to detect causal relationships. It should be clear that the above nonlinear interdependencies measures are bivariate measures. Although it quantified the “driver-response” for given input-the whole input space under study might be driven by other unobserved sources.

8.4 Statistic Tests and Data Analysis

In this study, the mutual information and nonlinear interdependence measures were estimated for every 10 seconds (2500 EEG data points) of continuous EEG recordings. The bootstrap re-sampling approach was adapted for deriving estimates on the measures[144]. Ten seconds of continuous EEG epoch is randomly sampled from every channels and this sampling procedure was repeated with replacement for 30 times. The reference A1 and A2 channels (inactive regions) are excluded from the analysis. Two sample $t$-test ($N=30, \alpha = 0.05$) is used to test the statistical differences on mutual information and nonlinear interdependence during before and after treatment. Low mutual information and information transport between different brain cortical regions were observed in our subjects with less severity of ULD. Furthermore, for each patient both mutual information and information transport between different brain cortical regions decrease after AEDs treatment. $t$-test for mutual information are summarized in Table 8-2, the topographical distribution for mutual information is also plot in heatmaps shown in Figures 8-2.

The significant “driver-response” relationship is revealed by $t$-test. After $t$-test the significant information transport between Fp1 and other brain cortical regions is shown in Fig. 8-1. The edges with an arrow starting from Fp1 to other channel denote $N(X|Y)$ is significant larger then $N(Y|X)$, therefore Fp1 is the driver, and vice versa. The
above results suggest the existing the treatment effects on the coupling strength and directionality of information transport between different brain cortical regions.

Table 8-2. Topographical distribution for treatment decoupling effect (DE: Decouple Electrode (DE))

<table>
<thead>
<tr>
<th>Electrode</th>
<th>DE for P1</th>
<th>DE for P2</th>
<th>DE for P3</th>
<th>DE for P4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Fp1</td>
<td>F3, C4, P4, F7</td>
<td>Fp2, F3, F8, T5</td>
<td>F3, F7</td>
<td>F3, P3, Fz, T5</td>
</tr>
<tr>
<td></td>
<td>T4, T5, O1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Fp2</td>
<td>C3, C4, F8, T4</td>
<td>Fp1, F4, T6, O2</td>
<td>F8</td>
<td>C3, C4, T5, P4</td>
</tr>
<tr>
<td></td>
<td>T5, Pz</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) F3</td>
<td>Fp1, C4, P4, O2</td>
<td>Fp1, C3, P3, Pz</td>
<td>Fp1</td>
<td>Fp1, F7</td>
</tr>
<tr>
<td>(4) F4</td>
<td>C4, P4 O2</td>
<td>Fp2, P4, O2, Fz</td>
<td>Cz</td>
<td>C4, Fz</td>
</tr>
<tr>
<td>(5) C3</td>
<td>Fp2, C4, P3, O1</td>
<td>F3, P3, O1</td>
<td>O1 Fp2, P3, O1</td>
<td></td>
</tr>
<tr>
<td>(6) C4</td>
<td>Fp1, Fp2, F3, F4, C3</td>
<td>P4, T6, O2</td>
<td>P4, O2</td>
<td>Fp1, Fp2, F3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7) P3</td>
<td>C3, O2</td>
<td>F3, C3, T3, Pz</td>
<td>O1 Fp1, C3, Cz</td>
<td></td>
</tr>
<tr>
<td>(8) P4</td>
<td>Fp1, F3, F4</td>
<td>F4, C4</td>
<td>C4</td>
<td>Fp2</td>
</tr>
<tr>
<td>(9) O1</td>
<td>Fp1, C3</td>
<td>C3, T5, T3, F7</td>
<td>P3, Pz</td>
<td>C3</td>
</tr>
<tr>
<td>(10) O2</td>
<td>F3, F4, P3</td>
<td>Fp2, F4, C4</td>
<td>C4</td>
<td>Pz</td>
</tr>
<tr>
<td>(11) F7</td>
<td>Fp1, Fz, Pz</td>
<td>C3, T5, O1</td>
<td>Fp1</td>
<td>F3</td>
</tr>
<tr>
<td>(12) F8</td>
<td>Fp2</td>
<td>Fp1, C4, P4</td>
<td>Fp2</td>
<td>Fp2</td>
</tr>
<tr>
<td>(13) T3</td>
<td>NONE</td>
<td>P3, O1</td>
<td>NONE</td>
<td>NONE</td>
</tr>
<tr>
<td>(14) T4</td>
<td>Fp1, Fp2, Cz</td>
<td>Fp2, O1</td>
<td>NONE</td>
<td>Fp2</td>
</tr>
<tr>
<td>(15) T5</td>
<td>Fp1, Fp2</td>
<td>Fp2, C4</td>
<td>NONE</td>
<td>Fp2, Cz</td>
</tr>
<tr>
<td>(16) T6</td>
<td>NONE</td>
<td>C4</td>
<td>NONE</td>
<td>NONE</td>
</tr>
<tr>
<td>(17) Fz</td>
<td>F7</td>
<td>Fp1, F4</td>
<td>Cz Fp1, Fp2, F3</td>
<td></td>
</tr>
<tr>
<td>(18) Pz</td>
<td>Fp2, F7, Cz</td>
<td>F3, P3, Cz</td>
<td>O2</td>
<td>O2</td>
</tr>
<tr>
<td>(19) Cz</td>
<td>T4, Pz</td>
<td>Pz</td>
<td>F4, Fz</td>
<td>Fp2, T5, P3</td>
</tr>
</tbody>
</table>

8.5 Conclusion and Discussion

The effectiveness of the new AEDs is currently accessed using myoclonus severity with the UMRS. As mentioned above, it is not easy to perform such evaluation scheme precisely especially in the later stages of the disease. Furthermore, the UMRS is a skewed measure that may not detect functional changes in a patient when these changes may be clinically important. The outcome from UMRS in this study did not evaluate the severity of the patients accurately, as the P1 was with less severity of ULD determined by the clinical experienced neurologist. The present study measure the synchronization behaviors and nonlinear interdependences, in a straightforward manner, in the cortical network during
Figure 8-1. Nonlinear interdependences for electrode FP1 before and after treatment. Although the results indicate that the mutual information and nonlinear interdependencies measures could be useful in determining the treatment effects for patients with ULD. To prove the usefulness of the proposed study, a larger patient population is needed. The approaches in this study are a bivariate measures, since a multivariate measure is not easy to model and has not been resolved. The decoupling between frontal and occipital cortical regions may be caused by decreased driving force deep inside the brain. In other words, the effect of the treatment may reduce the couple strength between thalamus and cortex in ULD subjects. Nevertheless, the limitations must be mentioned, it has been reported that it is necessary to take into account the
interdependence between thalamus and cortex when analyzing genesis of generalized spike-and-wave discharges [145–147].
Figure 8-2. Pairwise mutual information between for all electrodes- before v.s. after treatment
Table 8-3. Patient 1 before treatment nonlinear interdependencies

| Electrode Pair | \( S(X|Y) \) | \( S(Y|X) \) | \( H(X|Y) \) | \( H(Y|X) \) | \( N(Y|X) \) | \( N(Y|X) \) | \( C_{xy} \) |
|---------------|---------|---------|---------|---------|---------|---------|---------|
| \( X-Y \)     | mean ± SD | mean ± SD | mean ± SD | mean ± SD | mean ± SD | mean ± SD | mean ± SD |
| Fp1-Fp2       | 0.19 ± 0.21 | 0.21 ± 0.22 | 0.69 ± 0.79 | 0.69 ± 0.80 | 0.34 ± 0.36 | 0.34 ± 0.36 | 0.43 ± 0.45 |
| Fp1-F3        | 0.37 ± 0.10 | 0.44 ± 0.08 | 1.40 ± 0.39 | 1.38 ± 0.32 | 0.69 ± 0.10 | 0.70 ± 0.09 | 0.88 ± 0.05 |
| Fp1-F4        | 0.20 ± 0.07 | 0.34 ± 0.27 | 0.77 ± 0.63 | 0.77 ± 0.36 | 0.41 ± 0.20 | 0.44 ± 0.18 | 0.59 ± 0.25 |
| Fp1-C3        | 0.23 ± 0.06 | 0.31 ± 0.04 | 0.88 ± 0.40 | 0.75 ± 0.24 | 0.49 ± 0.14 | 0.48 ± 0.11 | 0.71 ± 0.09 |
| Fp1-C4        | 0.13 ± 0.09 | 0.22 ± 0.13 | 0.46 ± 0.40 | 0.40 ± 0.28 | 0.28 ± 0.23 | 0.27 ± 0.18 | 0.42 ± 0.28 |
| Fp1-P3        | 0.17 ± 0.04 | 0.26 ± 0.06 | 0.82 ± 0.48 | 0.56 ± 0.28 | 0.40 ± 0.17 | 0.35 ± 0.11 | 0.58 ± 0.12 |
| Fp1-P4        | 0.15 ± 0.08 | 0.25 ± 0.06 | 0.62 ± 0.60 | 0.41 ± 0.25 | 0.32 ± 0.19 | 0.28 ± 0.14 | 0.50 ± 0.15 |
| Fp1-O1        | 0.14 ± 0.04 | 0.36 ± 0.06 | 0.40 ± 0.51 | 0.21 ± 0.20 | 0.17 ± 0.20 | 0.10 ± 0.10 | 0.15 ± 0.26 |
| Fp1-O2        | 0.13 ± 0.03 | 0.23 ± 0.02 | 0.39 ± 0.28 | 0.21 ± 0.14 | 0.22 ± 0.15 | 0.14 ± 0.09 | 0.18 ± 0.21 |
| Fp1-F7        | 0.39 ± 0.11 | 0.37 ± 0.11 | 1.46 ± 0.40 | 1.42 ± 0.40 | 0.68 ± 0.17 | 0.68 ± 0.17 | 0.84 ± 0.16 |
| Fp1-F8        | 0.20 ± 0.06 | 0.24 ± 0.06 | 0.65 ± 0.33 | 0.60 ± 0.23 | 0.37 ± 0.18 | 0.36 ± 0.14 | 0.40 ± 0.33 |
| Fp1-T3        | 0.26 ± 0.08 | 0.31 ± 0.05 | 1.19 ± 0.52 | 1.08 ± 0.47 | 0.59 ± 0.17 | 0.57 ± 0.17 | 0.76 ± 0.20 |
| Fp1-T4        | 0.15 ± 0.04 | 0.24 ± 0.13 | 0.59 ± 0.51 | 0.46 ± 0.32 | 0.30 ± 0.23 | 0.26 ± 0.17 | 0.33 ± 0.29 |
| Fp1-T5        | 0.18 ± 0.05 | 0.28 ± 0.12 | 0.76 ± 0.49 | 0.54 ± 0.33 | 0.38 ± 0.19 | 0.34 ± 0.17 | 0.56 ± 0.22 |
| Fp1-T6        | 0.12 ± 0.02 | 0.21 ± 0.02 | 0.32 ± 0.24 | 0.16 ± 0.11 | 0.18 ± 0.14 | 0.10 ± 0.09 | 0.18 ± 0.17 |
| Fp1-Fz        | 0.27 ± 0.10 | 0.67 ± 1.51 | 1.13 ± 0.65 | 1.18 ± 0.34 | 0.57 ± 0.17 | 0.64 ± 0.11 | 0.80 ± 0.09 |
| Fp1-Cz        | 0.19 ± 0.05 | 0.36 ± 0.38 | 0.74 ± 0.38 | 0.65 ± 0.23 | 0.42 ± 0.17 | 0.42 ± 0.14 | 0.63 ± 0.11 |
| Fp1-Pz        | 0.16 ± 0.04 | 0.27 ± 0.06 | 0.63 ± 0.49 | 0.41 ± 0.22 | 0.35 ± 0.19 | 0.29 ± 0.14 | 0.49 ± 0.17 |
Table 8-4. Patient 1 after treatment nonlinear interdependencies

| Electrode Pair | $S(X|Y)$ mean ± SD | $S(Y|X)$ mean ± SD | $H(X|Y)$ mean ± SD | $H(Y|X)$ mean ± SD | $N(Y|X)$ mean ± SD | $N(Y|X)$ mean ± SD | $C_{xy}$ mean ± SD |
|----------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| Fp1-Fp2        | 0.38 ± 0.18         | 0.37 ± 0.16         | 1.44 ± 0.93         | 1.46 ± 0.89         | 0.62 ± 0.26         | 0.64 ± 0.22         | 0.81 ± 0.18         |
| Fp1-F3         | 0.39 ± 0.11         | 0.48 ± 0.07         | 1.56 ± 0.49         | 1.37 ± 0.25         | 0.72 ± 0.10         | 0.70 ± 0.08         | 0.87 ± 0.05         |
| Fp1-F4         | 0.24 ± 0.08         | 0.32 ± 0.07         | 0.87 ± 0.58         | 0.83 ± 0.34         | 0.46 ± 0.20         | 0.50 ± 0.15         | 0.71 ± 0.14         |
| Fp1-C3         | 0.22 ± 0.07         | 0.30 ± 0.04         | 0.88 ± 0.29         | 0.81 ± 0.18         | 0.50 ± 0.12         | 0.50 ± 0.09         | 0.72 ± 0.08         |
| Fp1-C4         | 0.19 ± 0.06         | 0.25 ± 0.04         | 0.62 ± 0.32         | 0.66 ± 0.22         | 0.36 ± 0.18         | 0.42 ± 0.12         | 0.63 ± 0.16         |
| Fp1-P3         | 0.16 ± 0.05         | 0.27 ± 0.03         | 0.54 ± 0.25         | 0.49 ± 0.14         | 0.33 ± 0.13         | 0.33 ± 0.10         | 0.58 ± 0.09         |
| Fp1-P4         | 0.14 ± 0.04         | 0.24 ± 0.04         | 0.47 ± 0.32         | 0.42 ± 0.17         | 0.27 ± 0.17         | 0.29 ± 0.10         | 0.53 ± 0.14         |
| Fp1-O1         | 0.13 ± 0.03         | 0.23 ± 0.03         | 0.34 ± 0.30         | 0.27 ± 0.15         | 0.18 ± 0.15         | 0.18 ± 0.10         | 0.44 ± 0.12         |
| Fp1-O2         | 0.12 ± 0.03         | 0.19 ± 0.04         | 0.23 ± 0.21         | 0.21 ± 0.13         | 0.10 ± 0.13         | 0.11 ± 0.09         | 0.31 ± 0.16         |
| Fp1-F7         | 0.20 ± 0.06         | 0.17 ± 0.03         | 0.60 ± 0.77         | 0.62 ± 0.36         | 0.28 ± 0.36         | 0.29 ± 0.37         | 0.36 ± 0.44         |
| Fp1-F8         | 0.21 ± 0.05         | 0.20 ± 0.09         | 0.87 ± 0.63         | 0.43 ± 0.22         | 0.49 ± 0.22         | 0.534 ± 0.25        | 0.64 ± 0.47         |
| Fp1-T3         | 0.26 ± 0.12         | 0.31 ± 0.08         | 0.93 ± 0.44         | 0.95 ± 0.38         | 0.51 ± 0.24         | 0.54 ± 0.19         | 0.72 ± 0.23         |
| Fp1-T4         | 0.15 ± 0.04         | 0.21 ± 0.06         | 0.48 ± 0.38         | 0.38 ± 0.25         | 0.22 ± 0.22         | 0.22 ± 0.18         | 0.32 ± 0.35         |
| Fp1-T5         | 0.15 ± 0.04         | 0.25 ± 0.03         | 0.52 ± 0.25         | 0.46 ± 0.15         | 0.31 ± 0.13         | 0.30 ± 0.10         | 0.58 ± 0.08         |
| Fp1-T6         | 0.14 ± 0.04         | 0.23 ± 0.03         | 0.33 ± 0.26         | 0.26 ± 0.15         | 0.18 ± 0.10         | 0.18 ± 0.08         | 0.38 ± 0.08         |
| Fp1-Fz         | 0.27 ± 0.10         | 0.38 ± 0.06         | 1.03 ± 0.51         | 1.06 ± 0.32         | 0.55 ± 0.16         | 0.60 ± 0.11         | 0.79 ± 0.09         |
| Fp1-Cz         | 0.19 ± 0.07         | 0.28 ± 0.04         | 0.69 ± 0.37         | 0.68 ± 0.18         | 0.40 ± 0.17         | 0.44 ± 0.09         | 0.65 ± 0.10         |
| Fp1-Pz         | 0.16 ± 0.06         | 0.26 ± 0.03         | 0.50 ± 0.27         | 0.46 ± 0.14         | 0.30 ± 0.13         | 0.31 ± 0.09         | 0.57 ± 0.08         |
| Electrode Pair | $S(X|Y)$ | $S(Y|X)$ | $H(X|Y)$ | $H(Y|X)$ | $N(Y|X)$ | $N(Y|X)$ | $C_{xy}$ |
|---------------|---------|---------|---------|---------|---------|---------|---------|
| $X - Y$       | mean ± SD | mean ± SD | mean ± SD | mean ± SD | mean ± SD | mean ± SD | mean ± SD |
| Fp1-Fp2      | 0.68 ± 0.17 | 0.70 ± 0.18 | 2.28 ± 0.50 | 2.20 ± 0.45 | 0.84 ± 0.09 | 0.84 ± 0.08 | 0.94 ± 0.06 |
| Fp1-F3       | 0.53 ± 0.14 | 0.58 ± 0.13 | 1.90 ± 0.35 | 1.75 ± 0.35 | 0.79 ± 0.07 | 0.78 ± 0.08 | 0.86 ± 0.09 |
| Fp1-F4       | 0.40 ± 0.09 | 0.51 ± 0.08 | 1.58 ± 0.40 | 1.40 ± 0.45 | 0.71 ± 0.11 | 0.68 ± 0.13 | 0.79 ± 0.14 |
| Fp1-C3       | 0.34 ± 0.09 | 0.47 ± 0.07 | 1.37 ± 0.43 | 1.40 ± 0.46 | 0.54 ± 0.23 | 0.63 ± 0.12 | 0.74 ± 0.11 |
| Fp1-C4       | 0.25 ± 0.08 | 0.41 ± 0.13 | 1.03 ± 0.44 | 1.04 ± 0.56 | 0.43 ± 0.21 | 0.49 ± 0.19 | 0.64 ± 0.19 |
| Fp1-P3       | 0.28 ± 0.07 | 0.39 ± 0.05 | 1.16 ± 0.46 | 1.38 ± 0.54 | 0.46 ± 0.26 | 0.58 ± 0.16 | 0.69 ± 0.13 |
| Fp1-P4       | 0.23 ± 0.06 | 1.00 ± 3.44 | 0.95 ± 0.42 | 1.09 ± 0.55 | 0.43 ± 0.20 | 0.49 ± 0.18 | 0.62 ± 0.20 |
| Fp1-O1       | 0.22 ± 0.07 | 0.34 ± 0.07 | 0.87 ± 0.43 | 1.15 ± 0.57 | 0.33 ± 0.29 | 0.47 ± 0.20 | 0.57 ± 0.19 |
| Fp1-O2       | 0.22 ± 0.05 | 0.34 ± 0.06 | 0.92 ± 0.37 | 1.04 ± 0.49 | 0.42 ± 0.17 | 0.46 ± 0.17 | 0.54 ± 0.22 |
| Fp1-F7       | 0.43 ± 0.15 | 0.47 ± 0.15 | 1.68 ± 0.44 | 1.76 ± 0.56 | 0.71 ± 0.21 | 0.74 ± 0.19 | 0.85 ± 0.19 |
| Fp1-F8       | 0.30 ± 0.06 | 0.40 ± 0.10 | 1.30 ± 0.39 | 1.24 ± 0.41 | 0.61 ± 0.13 | 0.60 ± 0.11 | 0.68 ± 0.11 |
| Fp1-T3       | 0.37 ± 0.11 | 0.41 ± 0.12 | 1.45 ± 0.50 | 1.89 ± 0.54 | 0.59 ± 0.14 | 0.59 ± 0.15 | 0.75 ± 0.29 |
| Fp1-T4       | 0.25 ± 0.07 | 0.37 ± 0.08 | 0.96 ± 0.43 | 1.05 ± 0.58 | 0.41 ± 0.23 | 0.47 ± 0.20 | 0.58 ± 0.25 |
| Fp1-T5       | 0.25 ± 0.06 | 0.31 ± 0.05 | 1.09 ± 0.42 | 1.41 ± 0.62 | 0.44 ± 0.23 | 0.55 ± 0.21 | 0.63 ± 0.17 |
| Fp1-T6       | 0.21 ± 0.06 | 0.36 ± 0.14 | 0.95 ± 0.47 | 1.15 ± 0.57 | 0.39 ± 0.27 | 0.49 ± 0.19 | 0.55 ± 0.24 |
| Fp1-Fz       | 0.43 ± 0.11 | 10.63 ± 55.36 | 1.78 ± 0.28 | 1.55 ± 0.36 | 0.76 ± 0.07 | 0.72 ± 0.10 | 0.86 ± 0.07 |
| Fp1-Cz       | 0.32 ± 0.07 | 0.44 ± 0.04 | 1.39 ± 0.37 | 1.33 ± 0.47 | 0.61 ± 0.15 | 0.63 ± 0.15 | 0.75 ± 0.13 |
| Fp1-Pz       | 0.24 ± 0.06 | 0.35 ± 0.06 | 0.92 ± 0.40 | 1.04 ± 0.51 | 0.37 ± 0.24 | 0.49 ± 0.17 | 0.61 ± 0.17 |
| Electrode Pair | $S(X|Y)$ mean ± SD | $S(Y|X)$ mean ± SD | $H(X|Y)$ mean ± SD | $H(Y|X)$ mean ± SD | $N(Y|X)$ mean ± SD | $N(Y|X)$ mean ± SD | $C_{xy}$ mean ± SD |
|---------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| Fp1-Fp2       | 0.40 ± 0.11      | 0.44 ± 0.09      | 1.52 ± 0.33      | 1.32 ± 0.30      | 0.73 ± 0.09      | 0.68 ± 0.09      | 0.88 ± 0.04      |
| Fp1-F3        | 0.33 ± 0.05      | 0.46 ± 0.03      | 1.26 ± 0.21      | 1.30 ± 0.16      | 0.66 ± 0.06      | 0.69 ± 0.05      | 0.83 ± 0.03      |
| Fp1-F4        | 0.21 ± 0.05      | 0.31 ± 0.05      | 0.78 ± 0.23      | 0.72 ± 0.25      | 0.44 ± 0.11      | 0.44 ± 0.13      | 0.62 ± 0.11      |
| Fp1-C3        | 0.16 ± 0.03      | 0.28 ± 0.02      | 0.55 ± 0.12      | 0.49 ± 0.12      | 0.35 ± 0.07      | 0.34 ± 0.07      | 0.55 ± 0.06      |
| Fp1-C4        | 0.14 ± 0.03      | 0.24 ± 0.02      | 0.39 ± 0.08      | 0.39 ± 0.17      | 0.22 ± 0.06      | 0.25 ± 0.11      | 0.34 ± 0.10      |
| Fp1-P3        | 0.13 ± 0.03      | 0.24 ± 0.02      | 0.29 ± 0.12      | 0.23 ± 0.17      | 0.17 ± 0.06      | 0.16 ± 0.06      | 0.14 ± 0.10      |
| Fp1-P4        | 0.13 ± 0.03      | 0.23 ± 0.02      | 0.29 ± 0.11      | 0.27 ± 0.12      | 0.16 ± 0.08      | 0.18 ± 0.08      | 0.12 ± 0.11      |
| Fp1-O1        | 0.12 ± 0.03      | 0.22 ± 0.02      | 0.24 ± 0.12      | 0.26 ± 0.12      | 0.13 ± 0.09      | 0.16 ± 0.09      | 0.03 ± 0.11      |
| Fp1-O2        | 0.13 ± 0.03      | 0.23 ± 0.02      | 0.27 ± 0.10      | 0.30 ± 0.12      | 0.14 ± 0.08      | 0.20 ± 0.08      | 0.08 ± 0.14      |
| Fp1-F7        | 0.29 ± 0.18      | 0.33 ± 0.20      | 1.12 ± 0.65      | 1.09 ± 0.63      | 0.54 ± 0.31      | 0.53 ± 0.30      | 0.66 ± 0.37      |
| Fp1-F8        | 0.22 ± 0.05      | 0.23 ± 0.06      | 0.78 ± 0.28      | 0.51 ± 0.24      | 0.45 ± 0.12      | 0.30 ± 0.15      | 0.51 ± 0.25      |
| Fp1-T3        | 0.20 ± 0.03      | 0.28 ± 0.04      | 0.75 ± 0.13      | 0.72 ± 0.14      | 0.43 ± 0.07      | 0.43 ± 0.08      | 0.64 ± 0.07      |
| Fp1-T4        | 0.14 ± 0.03      | 0.19 ± 0.04      | 0.42 ± 0.12      | 0.29 ± 0.15      | 0.24 ± 0.08      | 0.15 ± 0.14      | 0.19 ± 0.25      |
| Fp1-T5        | 0.14 ± 0.02      | 0.24 ± 0.03      | 0.33 ± 0.11      | 0.18 ± 0.08      | 0.19 ± 0.06      | 0.11 ± 0.06      | 0.18 ± 0.14      |
| Fp1-T6        | 0.13 ± 0.02      | 0.12 ± 0.02      | 0.28 ± 0.10      | 0.28 ± 0.08      | 0.16 ± 0.08      | 0.18 ± 0.08      | 0.03 ± 0.09      |
| Fp1-Fz        | 0.31 ± 0.07      | 0.43 ± 0.05      | 1.21 ± 0.30      | 1.18 ± 0.24      | 0.62 ± 0.09      | 0.64 ± 0.08      | 0.79 ± 0.04      |
| Fp1-Cz        | 0.15 ± 0.04      | 0.23 ± 0.04      | 0.40 ± 0.15      | 0.43 ± 0.24      | 0.24 ± 0.06      | 0.29 ± 0.08      | 0.45 ± 0.07      |
| Fp1-Pz        | 0.13 ± 0.02      | 0.23 ± 0.02      | 0.26 ± 0.10      | 0.26 ± 0.06      | 0.14 ± 0.06      | 0.18 ± 0.07      | 0.12 ± 0.11      |
Table 8-7. Patient 3 before treatment nonlinear interdependencies

| Electrode Pair | $S(X|Y)$ | $S(Y|X)$ | $H(X|Y)$ | $H(Y|X)$ | $N(Y|X)$ | $N(Y|X)$ | $C_{xy}$ |
|----------------|---------|---------|---------|---------|---------|---------|--------|
| Fp1-Fp2        | 0.41    | 0.36    | 2.19    | 2.16    | 0.79    | 0.80    | 0.91    |
|                | ±1.74   | ±0.13   | ±1.03   | ±0.96   | ±0.20   | ±0.15   | ±0.14   |
| Fp1-F3         | 0.37    | 0.49    | 2.00    | 1.94    | 0.75    | 0.75    | 0.84    |
|                | ±0.11   | ±0.09   | ±0.94   | ±1.03   | ±0.16   | ±0.16   | ±0.13   |
| Fp1-F4         | 0.32    | 0.89    | 1.55    | 1.56    | 0.64    | 0.67    | 0.68    |
|                | ±0.093  | ±2.56   | ±0.84   | ±0.86   | ±0.18   | ±0.16   | ±0.21   |
| Fp1-F7         | 0.24    | 1.03    | 1.52    | 1.43    | 0.60    | 0.63    | 0.70    |
|                | ±0.08   | ±3.59   | ±0.73   | ±0.71   | ±0.20   | ±0.19   | ±0.23   |
| Fp1-C3         | 0.15    | 0.25    | 0.78    | 0.81    | 0.36    | 0.40    | 0.4    |
|                | ±0.11   | ±0.17   | ±0.72   | ±0.70   | ±0.28   | ±0.29   | ±0.34   |
| Fp1-C4         | 0.16    | 5.37    | 0.88    | 0.82    | 0.40    | 0.40    | 0.48    |
|                | ±0.04   | ±24.1   | ±0.67   | ±0.62   | ±0.24   | ±0.22   | ±0.25   |
| Fp1-P3         | 0.15    | 0.71    | 0.88    | 0.79    | 0.37    | 0.38    | 0.37    |
|                | ±0.06   | ±2.64   | ±0.70   | ±0.69   | ±0.22   | ±0.20   | ±0.28   |
| Fp1-P4         | 0.15    | 9.50    | 0.78    | 0.69    | 0.36    | 0.35    | 0.33    |
|                | ±0.05   | ±28.22  | ±0.62   | ±0.55   | ±0.22   | ±0.21   | ±0.32   |
| Fp1-O1         | 0.04    | 0.05    | 0.66    | 0.80    | 0.22    | 0.21    | 0.33    |
|                | ±0.04   | ±0.05   | ±0.66   | ±0.80   | ±0.22   | ±0.21   | ±0.33   |
| Fp1-O2         | 0.39    | 1.68    | 2.08    | 2.04    | 0.75    | 0.72    | 0.83    |
|                | ±0.13   | ±6.70   | ±1.03   | ±1.21   | ±0.18   | ±0.24   | ±0.21   |
| Fp1-F8         | 0.21    | 0.28    | 1.19    | 1.31    | 0.49    | 0.534   | 0.64    |
|                | ±0.09   | ±0.09   | ±0.92   | ±1.02   | ±0.28   | ±0.25   | ±0.26   |
| Fp1-T3         | 0.22    | 0.40    | 1.27    | 1.33    | 0.52    | 0.53    | 0.64    |
|                | ±0.10   | ±0.51   | ±1.00   | ±1.17   | ±0.25   | ±0.28   | ±0.23   |
| Fp1-T4         | 0.19    | 8.14    | 1.17    | 1.25    | 0.51    | 0.54    | 0.61    |
|                | ±0.08   | ±30.12  | ±0.76   | ±0.88   | ±0.22   | ±0.22   | ±0.29   |
| Fp1-T5         | 0.17    | 14.48   | 1.23    | 1.25    | 0.50    | 0.48    | 0.55    |
|                | ±0.07   | ±43.45  | ±0.96   | ±1.11   | ±0.24   | ±0.25   | ±0.26   |
| Fp1-T6         | 0.06    | 14.22   | 0.74    | 0.81    | 0.24    | 0.20    | 0.28    |
|                | ±0.16   | ±4.32   | ±0.90   | ±0.99   | ±0.38   | ±0.44   | ±0.38   |
| Fp1-Fz         | 0.37    | 2.66    | 1.81    | 1.73    | 0.73    | 0.72    | 0.79    |
|                | ±0.12   | ±6.68   | ±0.74   | ±0.84   | ±0.12   | ±0.18   |
| Fp1-Cz         | 0.17    | 27.75   | 1.10    | 1.05    | 0.45    | 0.49    | 0.53    |
|                | ±0.06   | ±88.79  | ±0.68   | ±0.73   | ±0.19   | ±0.19   | ±0.18   |
| Fp1-Pz         | 0.64    | 3.62    | 0.95    | 0.92    | 0.40    | 0.41    | 0.44    |
|                | ±2.72   | ±18.61  | ±0.71   | ±0.79   | ±0.24   | ±0.24   | ±0.27   |
Table 8-8. Patient 3 after treatment nonlinear interdependencies

| Electrode Pair | $S(X|Y)$ mean ± SD | $S(Y|X)$ mean ± SD | $H(X|Y)$ mean ± SD | $H(Y|X)$ mean ± SD | $N(Y|X)$ mean ± SD | $N(Y|X)$ mean ± SD | $C_{xy}$ mean ± SD |
|----------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| Fp1-Fp2        | 1.02 ± 1.74         | 0.56 ± 0.13         | 2.19 ± 1.03         | 2.16 ± 0.96         | 0.79 ± 0.20         | 0.80 ± 0.15         | 0.91 ± 0.14         |
| Fp1-F3         | 0.37 ± 0.11         | 0.49 ± 0.09         | 2.00 ± 0.94         | 1.94 ± 1.03         | 0.75 ± 0.16         | 0.75 ± 0.16         | 0.84 ± 0.13         |
| Fp1-F4         | 0.32 ± 0.09         | 0.89 ± 2.56         | 1.55 ± 0.84         | 1.56 ± 0.86         | 0.64 ± 0.18         | 0.67 ± 0.16         | 0.68 ± 0.21         |
| Fp1-C3         | 0.24 ± 0.08         | 1.03 ± 3.59         | 1.52 ± 0.73         | 1.43 ± 0.71         | 0.60 ± 0.20         | 0.63 ± 0.19         | 0.70 ± 0.23         |
| Fp1-C4         | 0.15 ± 0.11         | 0.26 ± 0.17         | 0.78 ± 0.72         | 0.81 ± 0.70         | 0.36 ± 0.28         | 0.40 ± 0.29         | 0.4 ± 0.34          |
| Fp1-P3         | 0.16 ± 0.04         | 5.37 ± 24.1         | 0.88 ± 0.67         | 0.82 ± 0.62         | 0.40 ± 0.24         | 0.40 ± 0.22         | 0.48 ± 0.25         |
| Fp1-P4         | 0.15 ± 0.06         | 0.71 ± 2.64         | 0.88 ± 0.70         | 0.79 ± 0.69         | 0.37 ± 0.22         | 0.38 ± 0.20         | 0.37 ± 0.28         |
| Fp1-O1         | 0.15 ± 0.05         | 9.50 ± 28.22        | 0.78 ± 0.62         | 0.69 ± 0.55         | 0.36 ± 0.22         | 0.35 ± 0.21         | 0.33 ± 0.32         |
| Fp1-O2         | 0.04 ± 0.04         | 0.05 ± 0.05         | 0.66 ± 0.66         | 0.80 ± 0.80         | 0.22 ± 0.22         | 0.21 ± 0.21         | 0.33 ± 0.33         |
| Fp1-F7         | 0.39 ± 0.13         | 1.68 ± 6.70         | 2.08 ± 1.03         | 2.04 ± 1.21         | 0.75 ± 0.18         | 0.72 ± 0.24         | 0.83 ± 0.21         |
| Fp1-F8         | 0.21 ± 0.09         | 0.28 ± 0.09         | 1.19 ± 0.92         | 1.31 ± 1.02         | 0.49 ± 0.28         | 0.53 ± 0.25         | 0.64 ± 0.26         |
| Fp1-T3         | 0.22 ± 0.10         | 0.40 ± 0.51         | 1.27 ± 1.00         | 1.33 ± 1.17         | 0.52 ± 0.25         | 0.53 ± 0.28         | 0.64 ± 0.23         |
| Fp1-T4         | 0.19 ± 0.08         | 8.14 ± 30.12        | 1.17 ± 0.76         | 1.25 ± 0.88         | 0.51 ± 0.22         | 0.54 ± 0.22         | 0.61 ± 0.29         |
| Fp1-T5         | 0.17 ± 0.07         | 14.48 ± 43.45       | 1.23 ± 0.96         | 1.25 ± 1.11         | 0.50 ± 0.24         | 0.48 ± 0.25         | 0.55 ± 0.26         |
| Fp1-T6         | 0.06 ± 0.16         | 14.22 ± 4.32        | 0.74 ± 0.90         | 0.81 ± 0.99         | 0.24 ± 0.38         | 0.20 ± 0.44         | 0.28 ± 0.38         |
| Fp1-Pz         | ± 0.37             | ± 2.66             | ± 1.81             | ± 1.73             | ± 0.73             | ± 0.72             | ± 0.79             |
| Fp1-Cz         | 0.17 ± 0.06         | 27.75 ± 88.79       | 1.10 ± 0.68         | 1.05 ± 0.73         | 0.45 ± 0.19         | 0.49 ± 0.19         | 0.53 ± 0.18         |
| Fp1-Pz         | 0.64 ± 2.72         | 3.62 ± 18.61        | 0.95 ± 0.71         | 0.92 ± 0.79         | 0.40 ± 0.24         | 0.41 ± 0.24         | 0.44 ± 0.27         |
Table 8-9. Patient 4 before treatment nonlinear interdependencies

| Electrode Pair | S(X|Y)  | S(Y|X)  | H(X|Y)  | H(Y|X)  | N(Y|X)  | N(Y|X)  | C_{xy} |
|----------------|-------|--------|--------|--------|--------|--------|--------|
|                | mean ± SD | mean ± SD | mean ± SD | mean ± SD | mean ± SD | mean ± SD | mean ± SD |
| Fp1-Fp2        | 0.51 ± 0.10 | 0.54 ± 0.08 | 1.92 ± 0.47 | 1.68 ± 0.50 | 0.79 ± 0.07 | 0.76 ± 0.07 | 0.91 ± 0.04 |
| Fp1-F3         | 0.45 ± 0.07 | 0.52 ± 0.05 | 1.63 ± 0.28 | 1.28 ± 0.23 | 0.73 ± 0.08 | 0.68 ± 0.06 | 0.78 ± 0.07 |
| Fp1-F4         | 0.25 ± 0.08 | 0.36 ± 0.06 | 0.96 ± 0.31 | 0.79 ± 0.22 | 0.49 ± 0.14 | 0.47 ± 0.11 | 0.59 ± 0.09 |
| Fp1-C3         | 0.21 ± 0.04 | 0.28 ± 0.02 | 0.66 ± 0.28 | 0.53 ± 0.28 | 0.36 ± 0.16 | 0.34 ± 0.13 | 0.47 ± 0.13 |
| Fp1-C4         | 0.14 ± 0.10 | 0.21 ± 0.12 | 0.41 ± 0.31 | 0.35 ± 0.29 | 0.23 ± 0.18 | 0.22 ± 0.17 | 0.28 ± 0.21 |
| Fp1-P3         | 0.17 ± 0.06 | 0.24 ± 0.03 | 0.47 ± 0.22 | 0.36 ± 0.18 | 0.24 ± 0.14 | 0.23 ± 0.11 | 0.21 ± 0.15 |
| Fp1-P4         | 0.16 ± 0.05 | 0.47 ± 1.29 | 0.43 ± 0.21 | 0.29 ± 0.17 | 0.21 ± 0.13 | 0.19 ± 0.10 | 0.18 ± 0.18 |
| Fp1-O1         | 0.17 ± 0.05 | 0.23 ± 0.03 | 0.45 ± 0.26 | 0.27 ± 0.20 | 0.21 ± 0.16 | 0.17 ± 0.12 | 0.03 ± 0.14 |
| Fp1-O2         | 0.16 ± 0.05 | 0.23 ± 0.04 | 0.46 ± 0.24 | 0.30 ± 0.20 | 0.22 ± 0.14 | 0.18 ± 0.13 | 0.06 ± 0.20 |
| Fp1-F7         | 0.45 ± 0.10 | 0.50 ± 0.11 | 1.76 ± 0.28 | 1.51 ± 0.26 | 0.77 ± 0.06 | 0.72 ± 0.05 | 0.87 ± 0.05 |
| Fp1-F8         | 0.36 ± 0.05 | 0.39 ± 0.07 | 1.27 ± 0.38 | 0.91 ± 0.29 | 0.65 ± 0.11 | 0.54 ± 0.13 | 0.74 ± 0.11 |
| Fp1-T3         | 0.21 ± 0.05 | 0.80 ± 1.98 | 0.81 ± 0.25 | 0.61 ± 0.26 | 0.44 ± 0.12 | 0.38 ± 0.13 | 0.55 ± 0.10 |
| Fp1-T4         | 0.18 ± 0.07 | 0.28 ± 0.06 | 0.51 ± 0.31 | 0.33 ± 0.25 | 0.27 ± 0.19 | 0.21 ± 0.16 | 0.35 ± 0.18 |
| Fp1-T5         | 0.17 ± 0.05 | 0.25 ± 0.03 | 0.45 ± 0.26 | 0.24 ± 0.19 | 0.20 ± 0.18 | 0.16 ± 0.12 | 0.15 ± 0.14 |
| Fp1-T6         | 0.16 ± 0.06 | 0.24 ± 0.04 | 0.37 ± 0.26 | 0.23 ± 0.15 | 0.17 ± 0.18 | 0.14 ± 0.10 | 0.10 ± 0.14 |
| Fp1-Fz         | 0.41 ± 0.06 | 0.48 ± 2.10 | 1.47 ± 0.22 | 1.19 ± 0.22 | 0.70 ± 0.06 | 0.65 ± 0.06 | 0.77 ± 0.06 |
| Fp1-Cz         | 0.18 ± 0.06 | 0.63 ± 2.10 | 0.61 ± 0.28 | 0.49 ± 0.25 | 0.28 ± 0.19 | 0.13 ± 0.13 | 0.40 ± 0.13 |
| Fp1-Pz         | 0.17 ± 0.06 | 0.24 ± 0.02 | 0.48 ± 0.23 | 0.29 ± 0.16 | 0.23 ± 0.16 | 0.20 ± 0.10 | 0.16 ± 0.15 |
Table 8-10. Patient 4 after treatment nonlinear interdependencies

| Electrode Pair | S[X|Y] mean ± SD | S[Y|X] mean ± SD | H[X|Y] mean ± SD | H[Y|X] mean ± SD | N[Y|X] mean ± SD | N[Y|X] mean ± SD | C_{xy} mean ± SD |
|---------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Fp1-Fp2       | 0.34 ± 0.07     | 0.41 ± 0.07     | 1.41 ± 0.30     | 1.22 ± 0.24     | 0.69 ± 0.08     | 0.64 ± 0.08     | 0.87 ± 0.04     |
| Fp1-F3        | 0.32 ± 0.04     | 0.46 ± 0.03     | 1.24 ± 0.27     | 1.25 ± 0.15     | 0.64 ± 0.08     | 0.67 ± 0.05     | 0.82 ± 0.03     |
| Fp1-F4        | 0.21 ± 0.06     | 0.31 ± 0.05     | 0.76 ± 0.24     | 0.68 ± 0.26     | 0.43 ± 0.11     | 0.42 ± 0.13     | 0.60 ± 0.11     |
| Fp1-C3        | 0.18 ± 0.03     | 0.29 ± 0.02     | 0.58 ± 0.12     | 0.52 ± 0.12     | 0.35 ± 0.07     | 0.36 ± 0.06     | 0.55 ± 0.07     |
| Fp1-C4        | 0.15 ± 0.03     | 0.24 ± 0.02     | 0.38 ± 0.08     | 0.40 ± 0.15     | 0.22 ± 0.06     | 0.26 ± 0.10     | 0.32 ± 0.10     |
| Fp1-P3        | 0.14 ± 0.03     | 0.25 ± 0.03     | 0.33 ± 0.09     | 0.22 ± 0.08     | 0.20 ± 0.06     | 0.16 ± 0.05     | 0.15 ± 0.15     |
| Fp1-P4        | 0.12 ± 0.02     | 0.25 ± 0.10     | 0.27 ± 0.09     | 0.27 ± 0.07     | 0.15 ± 0.07     | 0.17 ± 0.06     | 0.09 ± 0.10     |
| Fp1-O1        | 0.14 ± 0.03     | 0.22 ± 0.14     | 0.25 ± 0.11     | 0.22 ± 0.10     | 0.15 ± 0.10     | 0.14 ± 0.08     | -0.17 ± 0.18    |
| Fp1-O2        | 0.13 ± 0.03     | 0.24 ± 0.08     | 0.26 ± 0.14     | 0.28 ± 0.07     | 0.14 ± 0.09     | 0.17 ± 0.09     | -0.09 ± 0.15    |
| Fp1-F7        | 0.38 ± 0.10     | 0.43 ± 0.21     | 1.46 ± 0.19     | 1.44 ± 0.07     | 0.70 ± 0.07     | 0.69 ± 0.07     | 0.86 ± 0.05     |
| Fp1-F8        | 0.23 ± 0.05     | 0.24 ± 0.25     | 0.91 ± 0.25     | 0.59 ± 0.28     | 0.50 ± 0.10     | 0.32 ± 0.18     | 0.45 ± 0.31     |
| Fp1-T3        | 0.20 ± 0.04     | 0.29 ± 0.15     | 0.72 ± 0.14     | 0.69 ± 0.14     | 0.41 ± 0.11     | 0.41 ± 0.08     | 0.63 ± 0.08     |
| Fp1-T4        | 0.15 ± 0.02     | 0.20 ± 0.17     | 0.44 ± 0.14     | 0.28 ± 0.15     | 0.26 ± 0.10     | 0.15 ± 0.13     | 0.19 ± 0.27     |
| Fp1-T5        | 0.13 ± 0.02     | 0.25 ± 0.07     | 0.30 ± 0.06     | 0.17 ± 0.05     | 0.18 ± 0.05     | 0.11 ± 0.05     | 0.19 ± 0.13     |
| Fp1-T6        | 0.13 ± 0.02     | 0.23 ± 0.28     | 0.28 ± 0.25     | 0.25 ± 0.16     | 0.16 ± 0.08     | 0.15 ± 0.06     | 0.01 ± 0.10     |
| Fp1-Fz        | 0.30 ± 0.05     | 0.42 ± 0.24     | 1.19 ± 0.20     | 1.17 ± 0.08     | 0.61 ± 0.08     | 0.63 ± 0.06     | 0.78 ± 0.04     |
| Fp1-Cz        | 0.15 ± 0.04     | 0.25 ± 0.10     | 0.41 ± 0.14     | 0.42 ± 0.24     | 0.24 ± 0.08     | 0.28 ± 0.07     | 0.43 ± 0.10     |
| Fp1-Pz        | 0.13 ± 0.02     | 0.25 ± 0.08     | 0.28 ± 0.09     | 0.25 ± 0.06     | 0.15 ± 0.06     | 0.17 ± 0.06     | 0.10 ± 0.10     |
REFERENCES


H.B. Lundborg. *Die progressive Myoclonus-Epilepsie (Unverrichts Myoclonie)*. Almqvist and Wiksell, Uppsala, 1903.


BIOGRAPHICAL SKETCH

Chang-Chia Liu earned his B.S. degree in spring 2000 from the Department of Industrial Engineering, Da-Yeh University in Taiwan. He went to the United States in spring 2002 and joined the University of Florida in fall 2002. He received dual M.S. degrees from Departments of Industrial and Systems Engineering and J. Crayton Pruitt Family Biomedical Engineering in spring 2004 and fall 2007, respectively. His research interests include global optimization, time series analysis, chaos theory, and nonlinear dynamics with applications in biomedicine.