

## CHAPTER 6

# Epilepsy Detection and Monitoring

Nicholas Fisher, Sachin Talathi, Alex Cadotte, and Paul R. Carney

Epilepsy is one of the world's most common neurological diseases, affecting more than 40 million people worldwide. Epilepsy's hallmark symptom, seizures, can have a broad spectrum of debilitating medical and social consequences. Although antiepileptic drugs have helped treat millions of patients, roughly a third of all patients are unresponsive to pharmacological intervention.

As our understanding of this dynamic disease evolves, new possibilities for treatment are emerging. An area of great interest is the development of devices that incorporate algorithms capable of detecting early onset of seizures or even predicting them hours before they occur. This lead time will allow for new types of interventional treatment. In the near future a patient's seizure may be detected and aborted before physical manifestations begin. In this chapter we discuss the algorithms that will make these devices possible and how they have been implemented to date. We investigate how wavelets, synchronization, Lyapunov exponents, principal component analysis, and other techniques can help investigators extract information about impending seizures. We also compare and contrast these measures, and discuss their individual strengths and weaknesses. Finally, we illustrate how these techniques can be brought together in a closed-loop seizure prevention system.

### 6.1 Epilepsy: Seizures, Causes, Classification, and Treatment

Epilepsy is a common chronic neurological disorder characterized by recurrent, unprovoked seizures [1, 2]. Epilepsy is the most common neurological condition in children and the third most common in adults after Alzheimer's and stroke. The World Health Organization estimates that there are 40 to 50 million people with epilepsy worldwide [3]. Seizures are transient epochs due to abnormal, excessive, or synchronous neuronal activity in the brain [2]. *Epilepsy* is a generic term used to define a family of seizure disorders. A person with recurring seizures is said to have epilepsy. Currently there is no cure for epilepsy. Many patients' seizures can be controlled, but not cured, with medication. Those resistant to the medication may become candidates for surgical intervention. Not all epileptic syndromes are life-long conditions; some forms are confined to particular stages of childhood. Epilepsy

should not be understood as a single disorder, but rather as a group of syndromes with vastly divergent symptoms all involving episodic abnormal electrical activity in the brain.

Roughly 70% of cases present with no known cause. Of the remaining 30%, the following are the most frequent causes: brain tumor and/or stroke; head trauma, especially from automobile accidents, gunshot wounds, sports accidents, and falls and blows; poisoning, such as lead poisoning, and substance abuse; infection, such as meningitis, viral encephalitis, lupus erythematosus and, less frequently, mumps, measles, diphtheria, and others; and maternal injury, infection, or systemic illness that affects the developing brain of the fetus during pregnancy.

All people inherit varying degrees of susceptibility to seizures. The genetic factor is assumed to be greater when no specific cause can be identified. Mutations in several genes have been linked to some types of epilepsy. Several genes that code for protein subunits of voltage-gated and ligand-gated ion channels have been associated with forms of generalized epilepsy and infantile seizure syndromes [4]. One interesting finding in animals is that repeated low-level electrical stimulation (kindling) to some brain sites can lead to permanent increases in seizure susceptibility. Certain chemicals can also induce seizures. One mechanism proposed for this is called excitotoxicity.

Epilepsies are classified in five ways: their etiology; semiology, observable manifestations of the seizures; location in the brain where the seizures originate; identifiable medical syndromes; and the event that triggers the seizures, such as flashing lights. This classification is based on observation (clinical and EEG) rather than underlying pathophysiology or anatomy. In 1989, the International League Against Epilepsy proposed a classification scheme for epilepsies and epileptic syndromes. It is broadly described as a two-axis scheme having the cause on one axis and the extent of localization within the brain on the other.

There are many different epilepsy syndromes, each presenting with its own unique combination of seizure type, typical age of onset, EEG findings, treatment, and prognosis. *Temporal lobe epilepsy* is the most common epilepsy of adults. In most cases, the epileptogenic region is found in the mesial temporal structures (e.g., the hippocampus, amygdala, and parahippocampal gyrus). Seizures begin in late childhood or adolescence. There is an association with febrile seizures in childhood, and some studies have shown herpes simplex virus (HSV) DNA in these regions, suggesting this epilepsy has an infectious etiology. Most of these patients have complex partial seizures sometimes preceded by an aura, and some temporal lobe epilepsy patients also suffer from secondary generalized tonic-clonic seizures. *Absence epilepsy* is the most common childhood epilepsy and affects children between the ages of 4 and 12 years of age. These patients have recurrent absence seizures that can occur hundreds of times a day. On their EEG, one finds the stereotypical generalized 3-Hz spike and wave discharges.

The first line of epilepsy treatment is anticonvulsant medication. In some cases the implantation of a vagus nerve stimulator or a special ketogenic diet can be helpful. Neurosurgical operations for epilepsy can be palliative, reducing the frequency or severity of seizures; however, in some patients, an operation can be curative. Although antiepileptic drug treatment is the standard therapy for epilepsy, one third of all patients remain unresponsive to currently available medication. There is gen-

eral agreement that, despite pharmacological and surgical advances in the treatment of epilepsy, seizures cannot be controlled in many patients and there is a need for new therapeutic approaches [5–7]. Of those unresponsive to anticonvulsant medication, 7% to 8% may profit from epilepsy surgery. However, about 25% of people with epilepsy will continue to experience seizures even with the best available treatment [8]. Unfortunately for those responsive to medication, many antiepileptic medicines have significant side effects that have a negative impact on quality of life. Some side effects can be of particular concern for women, children, and the elderly.

For these reasons, the need for more effective treatments for pharmacoresistant epilepsy was among the driving force behind a White House–initiated Curing Epilepsy: Focus on the Future (Cure) Conference held in March 2000 that emphasized specific research directions and benchmarks for the development of effective and safe treatment for people with epilepsy. There is growing awareness that the development of new therapies has slowed, and to move toward new and more effective therapies, novel approaches to therapy discovery are needed [9]. A growing body of research indicates that controlling seizures may be possible by employing a seizure prediction, closed-loop treatment strategy. If it were possible to predict seizures with high sensitivity and specificity, even seconds before their onset, therapeutic possibilities would change dramatically [10]. One might envision a simple warning system capable of decreasing both the risk of injury and the feeling of helplessness that results from seemingly unpredictable seizures. Most people with epilepsy seize without warning. Their seizures can have dangerous or fatal consequences especially if they come at a bad time and lead to an accident. In the brain, identifiable electrical changes precede the clinical onset of a seizure by tens of seconds, and these changes can be recorded in an EEG.

The early detection of a seizure has many potential benefits. Advanced warning would allow patients to take action to minimize their risk of injury and, possibly in the near future, initiate some form of intervention. An automatic detection system could be made to trigger pharmacological intervention in the form of fast-acting drugs or electrical stimulation. For patients, this would be a significant breakthrough because they would not be dependent on daily anticonvulsant treatment. Seizure prediction techniques could conceivably be coupled with treatment strategies aimed at interrupting the process before a seizure begins. Treatment would then only occur when needed, that is, on demand and in advance of an impending seizure. Side effects from treatment with antiepileptic drugs, such as sedation and clouded thinking, could be reduced by on-demand release of a short-acting drug or electrical stimulation during the preictal state.

Paired with other suitable interventions, such applications could reduce morbidity and mortality as well as greatly improve the quality of life for people with epilepsy. In addition, identifying a preictal state would greatly contribute to our understanding of the pathophysiological mechanisms that generate seizures. We discuss the most available seizure detection and prediction algorithms as well as their potential use and limitations in later sections in this chapter. First, however, we review the dynamic aspects of epilepsy and the most widely used approach to detect and predict epileptic seizures.

## 6.2 Epilepsy as a Dynamic Disease

The EEG is a complex signal. Its statistical properties depend on both time and space [11]. Characteristics of the EEG, such as the existence of limit cycles (alpha activity, ictal activity), instances of bursting behavior (during light sleep), jump phenomena (hysteresis), amplitude-dependent frequency behavior (the smaller the amplitude the higher the EEG frequency), and existence of frequency harmonics (e.g., under photic driving conditions), are among the long catalog of properties typical of nonlinear systems. The presence of nonlinearities in EEGs recorded from an epileptogenic brain further supports the concept that the epileptogenic brain is a nonlinear system. By applying techniques from nonlinear dynamics, several researchers have provided evidence that the EEG of the epileptic brain is a nonlinear signal with deterministic and perhaps chaotic properties [12–14].

The EEG can be conceptualized as a series of numerical values (voltages) over time and space (gathered from multiple electrodes). Such a series is called a *multivariate time series*. The standard methods for time-series analysis (e.g., power analysis, linear orthogonal transforms, and parametric linear modeling) not only fail to detect the critical features of a time series generated by an autonomous (no external input) nonlinear system, but may falsely suggest that most of the series is random noise [15]. In the case of a multidimensional, nonlinear system such as the EEG generators, we do not know, or cannot measure, all of the relevant variables. This problem can be overcome mathematically. For a dynamic system to exist, its variables must be related over time. Thus, by analyzing a single variable (e.g., voltage) over time, we can obtain information about the important dynamic features of the whole system. By analyzing more than one variable over time, we can follow the dynamics of the interactions of different parts of the system under investigation. Neuronal networks can generate a variety of activities, some of which are characterized by rhythmic or quasirhythmic signals. These activities are reflected in the corresponding local EEG field potential. An essential feature of these networks is that variables of the network have both a strong nonlinear range and complex interactions. Therefore, they belong to a general class of nonlinear systems with complex dynamics. Characteristics of the dynamics depend strongly on small changes in the control parameters and/or the initial conditions. Thus, real neuronal networks behave like nonlinear complex systems and can display changes between states such as small-amplitude, quasirandom fluctuations and large-amplitude, rhythmic oscillations. Such dynamic state transitions are observed in the brain during the transition between interictal and epileptic seizure states.

One of the unique properties of the brain as a system is its relatively high degree of plasticity. It can display adaptive responses that are essential to implementing higher functions such as memory and learning. As a consequence, control parameters are essentially plastic, which implies that they can change over time depending on previous conditions. In spite of this plasticity, it is necessary for the system to stay within a stable working range in order for it to maintain a stable operating point. In the case of the patient with epilepsy, the most essential difference between a normal and an epileptic network can be conceptualized as a decrease in the distance between operating and bifurcation points.

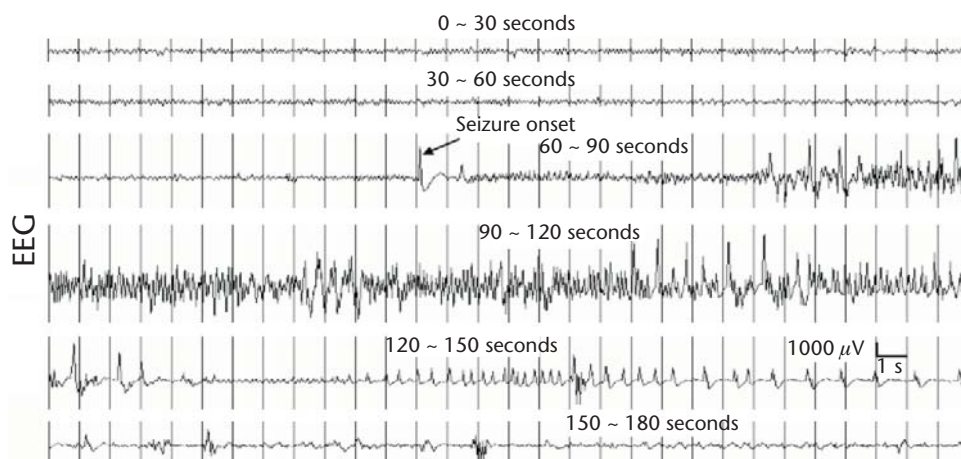
In considering epilepsies as dynamic diseases of brain systems, Lopes da Silva and colleagues proposed two scenarios of how a seizure could evolve [11]. The first is that a seizure could be caused by a sudden and abrupt state transition, in which case it would not be preceded by detectable dynamic changes in the EEG. Such a scenario would be conceivable for the initiation of seizures in primary generalized epilepsy. Alternatively, this transition could be a gradual change or a cascade of changes in dynamics, which could in theory be detected and even anticipated.

In the sections that follow, we use these basic concepts of brain dynamics and review the state-of-the-art seizure detection and seizure prediction methodologies and give examples using real data from human and rat epileptic time series.

### 6.3 Seizure Detection and Prediction

The majority of the current state-of-the-art techniques used to detect or predict an epileptic seizure involve linearly or nonlinearly transforming the signal using one of several mathematical black boxes, and subsequently trying to predict or detect the seizure based off the output of the black box. These black boxes include some purely mathematical transformations, such as the Fourier transform, or some class of machine learning techniques, such as artificial neural networks, or some combination of the two. In this section, we review some of the techniques for detection and prediction of seizures that have been reported in the literature.

Many techniques have been used in an attempt to detect epileptic seizures in the EEG. Historically, a visual confirmation was used to detect seizures. The onset and duration of a seizure could be identified on the EEG by a qualified technician. Figure 6.1 is an example of a typical spontaneous seizure in a laboratory animal model. Recently much research has been put into trying to predict or detect a seizure based off the EEG. The majority of these techniques use some kind of time-series analysis method to detect seizures offline. Time-series analysis of an EEG in general falls under one of the following two groups:



**Figure 6.1** Three minutes of EEG (demonstrated by six sequential 30-second segments) data recorded from the left hippocampus, showing a sample seizure from an epileptic rat.

1. *Univariate time-series analyses* are time-series analyses that consist of a single observation recorded sequentially over equal time increments. Some examples of univariate time series are the stock price of Microsoft, daily fluctuations in humidity levels, and single-channel EEG recordings. Time is an implicit variable in the time series. Information on the start time and the sampling rate of the data collection can allow one to visualize the univariate time series graphically as a function of time over the entire duration of data recording. The information contained in the amplitude value of the recorded EEG signal sampled in the form of a discrete time series  $x(t) = x(t_i) = x(i\Delta t)$ , ( $i = 1, 2, \dots, N$  and  $\Delta t$  is the sampling interval) can also be encoded through the amplitude and the phase of the subset of harmonic oscillations over a range of different frequencies. Time-frequency methods specify the map that translates between these representations.
2. *Multivariate time-series analyses* are time-series analyses that consist of more than one observation recorded sequentially in time. Multivariate time-series analysis is used when one wants to understand the interaction between the different components of the system under consideration. Examples include records of stock prices and dividends, concentration of atmospheric CO and global temperature, and multichannel EEG recordings. Time again is an implicit variable.

In the following sections some of the most commonly used measures for EEG time-series analysis will be discussed. First, a description of the linear and nonlinear univariate measures that operate on single-channel recordings of EEG data is given. Then some of the most commonly utilized multivariate measures that operate on more than a single channel of EEG data are described.

The techniques discussed next were chosen because they are representative of the different approaches used in seizure detection. Time–frequency analysis, nonlinear dynamics, signal correlation (synchronization), and signal energy are very broad domains and could be examined in a number of ways. Here we review a subset of techniques, examine each, and discuss the principles behind them.

## 6.4 Univariate Time-Series Analysis

### 6.4.1 Short-Term Fourier Transform

One of the more widely used techniques for detecting or predicting an epileptic seizure is based on calculating the power spectrum of one or more channels of the EEG. The core hypothesis, stated informally, is that the EEG signal, when partitioned into its component periodic (sine/cosine) waves, has a signature that varies between the ictal and the interictal states. To detect this signature, one takes the Fourier transform of the signal and finds the frequencies that are most prominent (in amplitude) in the signal. It has been shown that there is a relationship between the power spectrum of the EEG signal and ictal activity [16]. Although there appears to be a correlation between the power spectrum and ictal activity, the power spectrum is not used as a stand-alone detector of a seizure. In general, it is coupled with some other time-series prediction technique or machine learning to detect a seizure.



The Fourier transform is a generalization of the Fourier series. It breaks up any time-varying signal into its frequency components of varying magnitude and is defined in (6.1).

$$F(k) = \int_{-\infty}^{\infty} f(t)e^{-2\pi ikx} dx \quad (6.1)$$

Due to Euler's formula, this can also be written as shown in (6.2) for any complex function  $f(t)$  where  $k$  is the  $k$ th harmonic frequency:

$$F(k) = \int_{-\infty}^{\infty} f(t) \cos(-2\pi kx) dx + \int_{-\infty}^{\infty} f(t) i \sin(-2\pi kx) dx \quad (6.2)$$

We can represent any time-varying signal as a summation of sine and cosine waves of varying magnitudes and frequencies [17]. The Fourier transform is represented with the power spectrum. The power spectrum has a value for each harmonic frequency, which indicates how strong that frequency is in the given signal. The magnitude of this value is calculated by taking the modulus of the complex number that is calculated from the Fourier transform for a given frequency ( $|F(k)|$ ).

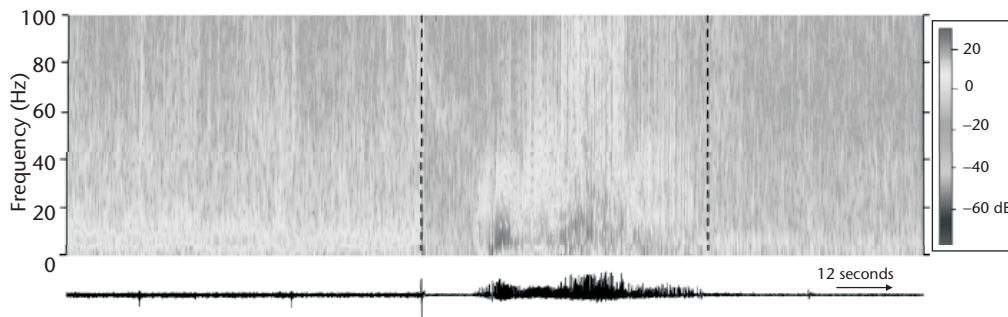
Stationarity is an issue that needs to be considered when using the Fourier transform. A stationary signal is one that is constant in its statistical parameters over time, and is assumed by the Fourier transform to be present. A signal that is made up of different frequencies at different times will yield the same transform as a signal that is made up of those same frequencies for the entire time period considered. As an example, consider two functions  $f_1$  and  $f_2$  over the domain  $0 = t = T$ , for any two frequencies  $\omega_1$  and  $\omega_2$  shown in (6.3) and (6.4):

$$f_1(t) = \sin(2\pi\omega_1 t) + \cos(2\pi\omega_2 t) \text{ if } 0 \leq t < T \quad (6.3)$$

and

$$f_2(t) = \begin{cases} \sin(2\pi\omega_1 t) & \text{if } 0 \leq t < T/2 \\ \cos(2\pi\omega_2 t) & \text{if } T/2 \leq t < T \end{cases} \quad (6.4)$$

When using the short-term Fourier transform, the assumption is made that the signal is stationary for some small period of time,  $T_s$ . The Fourier transform is then calculated for segments of the signal of length  $T_s$ . The short-term Fourier transform at time  $t$  gives the Fourier transform calculated over the segment of the signal lasting from  $(t - T_s)$  to  $t$ . The length of  $T_s$  determines the resolution of the analysis. There is a trade-off between time and frequency resolution. A short  $T_s$  yields better time resolution, but it limits the frequency resolution. The opposite of this is also true; a long  $T_s$  increases frequency resolution while decreasing the time resolution of the output. Wavelet analysis overcomes this limitation, and offers a tool that can maintain both time and frequency resolution. An example of Fourier transform calculated prior to, during, and following an epileptic seizure is given in Figure 6.2.



**Figure 6.2** Time-frequency spectrum plot for 180-second epoch of seizure. Black dotted lines mark the onset time and the offset times of the seizure.

#### 6.4.2 Discrete Wavelet Transforms

*Wavelets* are another closely related method used to predict epileptic seizures. Wavelet transforms follow the principle of superposition, just like Fourier transforms, and assume EEG signals are composed of various elements from a set of parameterized basis functions. Rather than being limited to sine and cosine wave functions, however, as in a Fourier transform, wavelets have to meet other mathematical criteria, which allow the basis functions to be far more general than those for simple sine/cosine waves. Wavelets make it substantially easier to approximate choppy signals with sharp spikes, as compared to the Fourier transform. The reason for this is that sine (and cosine) waves have infinite support (i.e., stretch out to infinity in time), which makes it difficult to approximate a spike. Wavelets are allowed to have finite support, so a spike in the EEG signal can be easily estimated by changing the magnitude of the component basis functions.

The discrete wavelet transform is similar to the Fourier transform in that it will break up any time-varying signal into smaller uniform functions, known as the basis functions. The basis functions are created by scaling and translating a single function of a certain form. This function is known as the mother wavelet. In the case of the Fourier transform, the basis functions used are sine and cosine waves of varying frequency and magnitude. Note that a cosine wave is just a sine wave translated by  $\pi/2$  radians, so the mother wavelet in the case of the Fourier transform could be considered to be the sine wave. However, for a wavelet transform the basis functions are more general. The only requirements for a family of functions to be a basis is that the functions are both complete and orthonormal under the inner product. Consider the family of functions  $\Psi = \{\psi_{ij} | -\infty < i, j < \infty\}$  where each  $i$  value specifies a different scale and each  $j$  value specifies a different translation based off of some mother wavelet function. Note that  $\Psi$  is considered to be complete if any continuous function  $f$ , defined over the real line  $x$ , can be defined by some combination of the functions in  $\Psi$  as shown in (6.5) [17]:

$$f(x) = \sum_{i,j=-\infty}^{\infty} c_{ij} \psi_{ij}(x) \quad (6.5)$$

In order for a family of functions to be orthonormal under the inner product, they must meet two criteria. It must be the case that for any  $i, j, l$ , and  $m$  where  $i \neq l$



and  $j \neq m$  that  $\langle y_{ij}, y_{lm} \rangle \geq 0$  and  $\langle y_{ij}, y_{ij} \rangle \geq 1$ , where  $\langle f, g \rangle$  is the inner product and is defined as in (6.6) and  $f(x)^*$  is the complex conjugate of  $f(x)$ :

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x)^* g(x) dx \quad (6.6)$$

The wavelet basis is very similar to the Fourier basis, with the exception that the wavelet basis does not have to be infinite. In a wavelet transform the basis functions can be defined over a certain window and then be zero everywhere else. As long as the family of functions defined by scaling and translating the mother wavelet is orthonormally complete, that family of functions can be used as the basis. With the Fourier transform, the basis is made up of sine and cosine waves that are defined over all values of  $x$  where  $-\infty < x < \infty$ .

One of the simplest wavelets is the *Haar wavelet* (Daubechies 2 wavelet). In a manner similar to the Fourier series, any continuous function  $f(x)$  defined on  $[0, 1]$  can be represented using the expansion shown in (6.7). The  $h_{j,k}(x)$  term is known as the Haar wavelet function and is defined as shown in (6.8);  $p_{j,k}(x)$  is known as the Haar scaling function and is defined in (6.9) [17]:

$$f(x) = \sum_{j=J}^{\infty} \sum_{k=0}^{2^j-1} \langle f, h_{j,k} \rangle h_{j,k}(x) + \sum_{k=0}^{2^J-1} \langle f, p_{J,k} \rangle p_{J,k}(x) \quad (6.7)$$

$$h_{j,k}(x) = \begin{cases} 2^{j/2} & \text{if } 0 \leq 2^j x - k < 1/2 \\ -2^{j/2} & \text{if } 1/2 \leq 2^j x - k < 1 \\ 0 & \text{otherwise} \end{cases} \quad (6.8)$$

$$p_{j,k}(x) = \begin{cases} 2^{j/2} & \text{if } 0 \leq 2^j x - k < 1 \\ 0 & \text{otherwise} \end{cases} \quad (6.9)$$

The combination of the Haar scaling function at the largest scale, along with the Haar wavelet functions, creates a set of functions that provides an orthonormal basis for functions in  $\mathbb{R}^2$ .

Wavelets and short-term Fourier transforms also serve as the foundation for other measures. Methods such as the spectral entropy method calculate some feature based on the power spectrum. Entropy was first used in physics as a thermodynamic quantity describing the amount of disorder in a system. Shannon extended its application to information theory in the late 1940s to calculate the entropy for a given probability distribution [18]. The entropy measure that Shannon developed can be expressed as follows:

$$H = -\sum p_k \log p_k \quad (6.10)$$

Entropy is a measure of how much information there is to learn from a random event occurring. Events that are unlikely to occur yield more information than events that are very probable. For spectral entropy, the power spectrum is considered to be a probability distribution. This insinuates that the random events would be that the signal was made up of a sine or cosine wave of a given frequency. The

spectral entropy allows us to calculate the amount of information there is to be gained from learning the frequencies that make up the signal. When the Fourier transform is used, nonstationary signals need to be accounted for. To do this, the short-term Fourier transform is used to calculate the power spectrum over small parts of the signal rather than the entire signal itself. The spectral entropy is an indicator of the number of frequencies that make up a signal. A signal made up of many different frequencies (white noise, for example) would have a uniform distribution and therefore yield high spectral entropy, whereas a signal made up of a single frequency would yield low spectral entropy.

In practice, wavelets have been applied to electrocorticogram (ECoG) signals in an effort to try to predict seizures. In one report, the authors first partitioned the ECoG signal into seizure and nonseizure components using a wavelet-based filter. This filter was not specifically predictive of seizures. It flagged any increase in power or shift in frequency, whether this change in the signal was caused by a seizure, an interictal epileptiform discharge, or merely normal activity. After the filter decomposed the signal down into its components, it was passed through a second filter that tried to isolate the seizures from the rest of the events. By decomposing the ECoG signal into components and passing it through the second step of isolating the seizures, the authors were able to detect all seizures with an average of 2.8 false positives per hour [19]. Unfortunately, this technique did not allow them to predict (as opposed to detect) seizures.

### 6.4.3 Statistical Moments

When a cumulative distribution function for a random variable cannot be determined, it is possible to describe an approximation to the distribution of this variable using moments and functions of moments [20]. Statistical moments relate information about the distribution of the amplitude of a given signal. In probability theory, the  $k$ th moment is defined as in (6.11) where  $E[x]$  is the expected value of  $x$ :

$$\mu_k = E[x^k] = \int x^k p(x) \quad (6.11)$$

The first statistical moment is the mean of the distribution being considered.

In general, the statistical moments are taken about the mean. This is also known as the  $k$ th central moment and is defined by (6.12) where  $\mu$  is the mean of the dataset considered [20]:

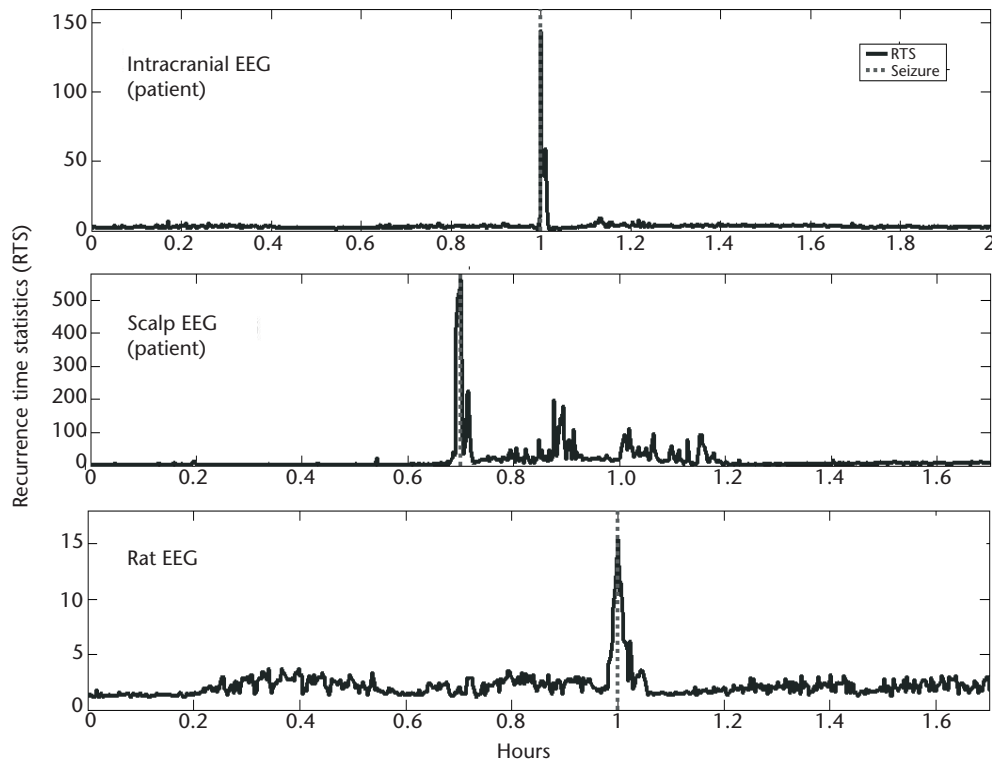
$$\mu_k = E[x - \mu]^k = \int (x - \mu)^k p(x) \quad (6.12)$$

The second moment about the mean would give the variance. The third and fourth moments about the mean would produce the skew and kurtosis, respectively. The skew of a distribution indicates the amount of asymmetry in that distribution, whereas the kurtosis shows the degree of peakedness of that distribution. The absolute value of the skewness  $|\mu_3|$  was used for seizure prediction in a review by Mormann et al. [14]. The paper showed that skewness was not able to significantly predict a seizure by detecting the state change from interictal to preictal. Although

unable to predict seizures, statistical moments may prove valuable as seizure detectors in recordings with large amplitude seizures.

#### 6.4.4 Recurrence Time Statistics

The recurrence time statistic (RTS),  $T_1$ , is a characteristic of trajectories in an abstract dynamic system. Stated informally, it is a measure of how often a given trajectory of the dynamic system visits a certain neighborhood in the phase space.  $T_1$  has been calculated for some ECoG data in an effort to detect seizures, with significant success. With two different patients and a total of 79 hours of data, researchers were able to detect 97% of the seizures with only an average of 0.29 false negatives per hour [21]. They did not, however, indicate any attempts to predict seizures. Results from our preliminary studies on human EEG signals showed that the RTS exhibited significant change during the ictal period that is distinguishable from the background interictal period (Figure 6.3). In addition, through the observations over multichannel RTS features, the spatial pattern from channel to channel can also be traced. Existence of these spatiotemporal patterns of RTS suggests that it is possible to utilize RTS to develop an automated seizure-warning algorithm.



**Figure 6.3** Studies on human EEG signals show that the recurrence time statistics exhibit changes during the ictal period that is distinguishable from the background interictal period.

### 6.4.5 Lyapunov Exponent

During the past decade, several studies have demonstrated experimental evidence that temporal lobe epileptic seizures are preceded by changes in dynamic properties of the EEG signal. A number of nonlinear time-series analysis tools have yielded promising results in terms of their ability to reveal preictal dynamic changes essential for actual seizure anticipation.

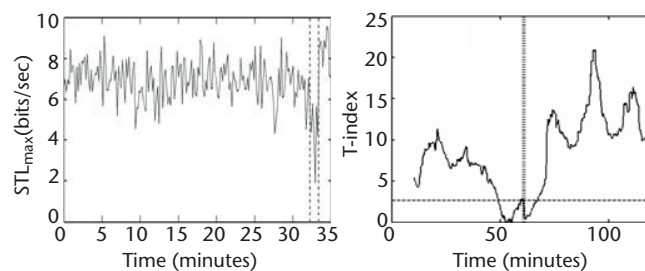
It has been shown that patients go through a preictal transition approximately 0.5 to 1 hour before a seizure occurs, and this preictal state can be characterized by the *Lyapunov exponent* [12, 22–29]. Stated informally, the Lyapunov exponent measures how fast nearby trajectories in a dynamic system diverge. The noted approach therefore treats the epileptic brain as a dynamic system [30–32]. It considers a seizure as a transition from a chaotic state (where trajectories are sensitive to initial conditions) to an ordered state (where trajectories are insensitive to initial conditions) in the dynamic system. The Lyapunov exponent is a nonlinear measure of the average rate of divergence/convergence of two neighboring trajectories in a dynamic system dependent on the sensitivity of initial conditions. It has been successfully used to identify preictal changes in EEG data [22–24]. Generally, Lyapunov exponents can be estimated from the equation of motion describing the time evolution of a given dynamic system. However, in the absence of the equation of motion describing the trajectory of the dynamic system, Lyapunov exponents are determined from observed scalar time-series data,  $x(t_n) = x(n \delta t)$ , where  $\delta t$  is the sampling rate for the data acquisition. In this situation, the goal is to generate a higher dimensional vector embedding of the scalar data  $x(t)$  that defines the state space of the multivariate brain dynamics from which the scalar EEG data is derived. Heuristically, this is done by constructing a higher dimensional vector  $\mathbf{x}_i$  from the data segment  $x(t)$  of given duration  $T$ , as shown in (6.13) with  $\tau$  defining the embedding delay used to construct a higher dimensional vector  $\mathbf{x}$  from  $x(t)$  with  $d$  as the selected dimension of the embedding space and  $t_i$  being the time instance within the period  $[T - (d - 1)\tau]$ :

$$\mathbf{x}_i = [x(t_i), x(t_i - \tau), \dots, x(t_i - (d - 1)\tau)] \quad (6.13)$$

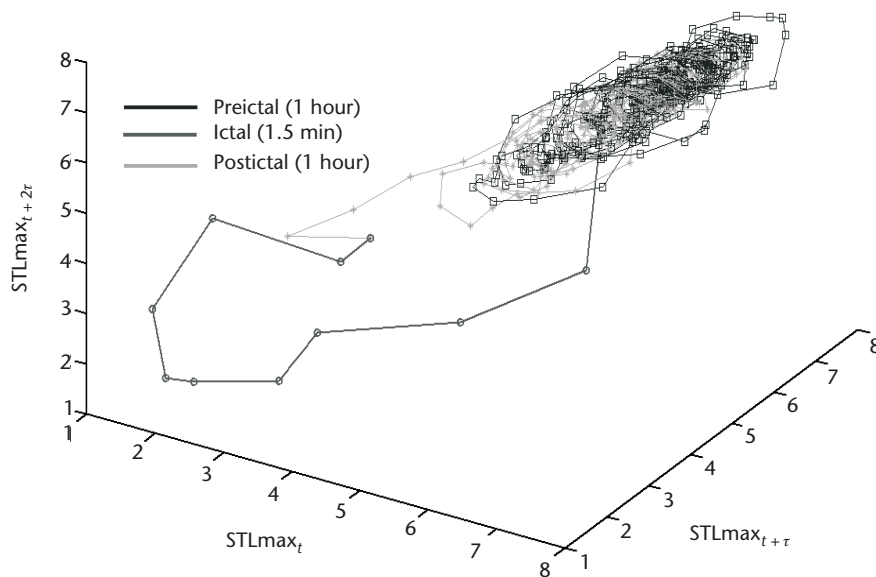
The geometrical theorem of [33] tells us that for an appropriate choice of  $d > d_{\min}$ ,  $\mathbf{x}_i$  provides a faithful representation of the phase space for the dynamic systems from which the scalar time series was derived. A suitable practical choice for  $d$ , the embedding dimension, can be derived from the “false nearest neighbor” algorithm. In addition, a suitable prescription for selecting the embedding delay,  $\tau$ , is also given in Abarbanel [34]. From  $\mathbf{x}_i$  a most stable short-term estimation of the largest Lyapunov exponent can be performed that is referred to as the *short-term largest Lyapunov exponent* ( $STL_{\max}$ ) [24]. The estimation  $L$  of  $STL_{\max}$  is obtained using (6.14) where  $\delta x_{ij}(0) = x(t_i) - x(t_j)$  is the displacement vector, defined at time points  $t_i$  and  $t_j$  and  $\delta x_{ij}(\Delta t) = x(t_i + \Delta t) - x(t_j + \Delta t)$  is the same vector after time  $\Delta t$ , and where  $N$  is the total number of local  $STL_{\max}$  that will be estimated within the time period  $T$  of the data segment, where  $T = N\Delta t + (d - 1)\tau$ :

$$L = \frac{1}{N\Delta t} \sum_{i=1}^N \log_2 \left| \frac{\delta x_{ij}(\Delta t)}{\delta x_{ij}(0)} \right| \quad (6.14)$$

A decrease in the Lyapunov exponent indicates this transition to a more ordered state (Figure 6.4). The assumptions underlying this methodology have been experimentally observed in the  $STL_{max}$  time-series data from human patients and rodents (Figure 6.5). This characterization by the Lyapunov exponent has, however, been successful only for EEG data recorded from particular areas in the neocortex and hippocampus and has been unsuccessful for other areas. Unfortunately, these areas can vary from seizure to seizure even in the same patient. The method is therefore



**Figure 6.4** Sample  $STL_{max}$  profile for a 35-minute epoch including a grade 5 seizure from an epileptic rat. Seizure onset and offset are indicated by dashed vertical lines. Note the drop in the  $STL_{max}$  value during the seizure period. (b) T-index profiles calculated from  $STL_{max}$  values of a pair of electrodes from rat A. The electrode pair includes a right hippocampus electrode and a left frontal electrode. Vertical dotted lines represent seizure onset and offset. The horizontal dashed line represents the critical entrainment threshold. Note a decline in the T-index value several minutes before seizure occurrence.



**Figure 6.5** Phase portrait of  $STL_{max}$  of a spontaneous rodent epileptic seizure (grade 5).

very sensitive to the electrode sites chosen. However, when the correct sites were chosen, the preictal transition was seen in more than 91% of the seizures. On average, this led to a prediction rate of 80.77% and an average warning time of 63 minutes [28]. Sadly, this method has been plagued by problems related to finding the critical electrode sites because their predictive capacity changes from seizure to seizure.

## 6.5 Multivariate Measures

Multivariate measures take more than one channel of EEG into account simultaneously. This is used to consider the interactions between the channels and how they correlate rather than looking at channels individually. This is useful if there is some interaction (e.g., synchronization) between different regions of the brain leading up to a seizure. Of the techniques discussed in the following sections, the simple synchronization measure and the lag synchronization measure fall under a subset of the multivariate measures, known as bivariate measures. Bivariate measures only consider two channels at a time and define how those two channels correlate. The remaining metrics take every EEG channel into account simultaneously. They do this by using a dimensionality reduction technique called principal component analysis (PCA). PCA takes a dataset in a multidimensional space, finds the most prominent dimensions in that dataset, and linearly transforms the original dataset to a lower dimensional space using the most prominent dimensions from the original dataset. PCA is used as a seizure detection technique itself [35]. It is also used as a tool to extract the most important dimensions from a data matrix containing pairwise correlation information for all EEG channels, as is the case with the correlation structure.

### 6.5.1 Simple Synchronization Measure

Several studies have shown that areas of the brain synchronize with one other during certain events. During seizures abnormally large amounts of highly synchronous activity are seen, and it has been suggested this activity may begin hours before the initiation of a seizure.

One multivariate method that has been used to calculate the synchronization between two EEG channels is a technique suggested by Quiroga et al. [36]. It first defines certain “events” for a pair of signals. Once the events have been defined in the signals, this method then counts the number of times the events in the two signals occur within a specified amount of time ( $\tau$ ) of each other [36]. It then divides this count by a normalizing term equivalent to the maximum number of events that could be synchronized in the signals.

For two discrete EEG channels  $x_i$  and  $y_j$ ,  $i = 1, \dots, N$ , where  $N$  is the number of points making up the EEG signal for the segment considered, event times are defined to be  $t_i^x$  and  $t_j^y$  ( $i = 1, \dots, m_x$ ;  $j = 1, \dots, m_y$ ). An event can be defined to be anything; however, events should be chosen so that the events appear simultaneously across the signals when they are synchronized. Quiroga et al. [36] define an event to be a local maximum over a range of  $K$  values. In other words, the  $i$ th point in signal  $x$



would be an event if  $x_i > x_{i \pm k}$ ,  $k = 1, \dots, K$ . The term  $\tau$  represents the time within which events from  $x$  and  $y$  must occur in order to be considered synchronized, and it must be less than half of the minimum interevent distance; otherwise, a single event in one signal could be considered to be synchronized with two different events in the other signal.

Finally, the number of events in  $x$  that appear “shortly” (within  $\tau$ ) after an event in  $y$  is counted as shown in (6.15) when  $J_{ij}^\tau$  is defined as in (6.16):

$$c^\tau(x|y) = \sum_{i=1}^{m_x} \sum_{j=1}^{m_y} J_{ij}^\tau \quad (6.15)$$

$$J_{ij}^\tau = \begin{cases} 1 & \text{if } 0 < t_i^x - t_j^y \\ 1/2 & \text{if } t_i^x = t_j^y \\ 0 & \text{else} \end{cases} \quad (6.16)$$

Similarly, the number of events in  $y$  that appear shortly after an event in  $x$  can also be defined in an analogous way. This would be denoted  $c^\tau(y|x)$ . With these two values, the synchronization measure  $Q_\tau$  can be calculated. This measure is shown in (6.17):

$$Q_\tau = \frac{c^\tau(x|y) + c^\tau(y|x)}{\sqrt{m_x m_y}} \quad (6.17)$$

The metric is normalized so that  $0 \leq Q_\tau \leq 1$  and  $Q_\tau$  is 1 if and only if  $x$  and  $y$  are fully synchronized (i.e., always have corresponding events within  $\tau$ ).

### 6.5.2 Lag Synchronization

When two different systems are identical with the exception of a shift by some time lag  $\tau$ , they are said to be lag synchronized [37]. This characteristic was tested by Mormann et al. [38] when applied to EEG channels in the interictal and preictal stage. To calculate the similarity of two signals they used a normalized cross-correlation function (6.18) as follows:

$$C(s_a, s_b)(\tau) = \frac{\text{corr}(s_a, s_b)(\tau)}{\sqrt{\text{corr}(s_a, s_a)(0) \cdot \text{corr}(s_b, s_b)(\tau)}} \quad (6.18)$$

where  $\text{corr}(s_a, s_b)(\tau)$  represents the linear cross-correlation function between the two time series  $s_a(t)$  and  $s_b(t)$  computed at lag time  $\tau$  as defined here:

$$\text{corr}(s_a, s_b)(\tau) = \int_{-\infty}^{\infty} s_a(t + \tau) s_b(t) dt \quad (6.19)$$

The normalized cross-correlation function yields a value between 0 and 1, which indicates how similar the two signals ( $s_a$  and  $s_b$ ) are. If the normalized cross-correlation function produces a value close to 1 for a given  $t$ , then the signals

are considered to be lag synchronized by a phase of  $\tau$ . Hence the final feature used to calculate the lag synchronization is the largest normalized cross correlation over all values of  $\tau$ , as shown in (6.20). A  $C_{\max}$  value of 1 indicates totally synchronized signals within some time lag  $\tau$  and unsynchronized signals produce a value very close to 0.

$$C_{\max} = \max_{\tau} \{C(s_a, s_b)(\tau)\} \quad (6.20)$$

## 6.6 Principal Component Analysis

Principal component analysis attempts to solve the problem of excessive dimensionality by combining features to reduce the overall dimensionality. By using linear transformations, it projects a high dimensional dataset onto a lower dimensional space so that the information in the original dataset is preserved in an optimal manner when using the least squared distance metric. An outline of the derivation of PCA is given here. The reader should refer to Duda et al. [39] for a more detailed mathematical derivation.

Given a  $d$ -dimensional dataset of size  $n$  ( $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ ), we first consider the problem of finding a vector  $\mathbf{x}_0$  to represent all of the vectors in the dataset. This comes down to the problem of finding the vector  $\mathbf{x}_0$ , which is closest to every point in the dataset. We can find this vector by minimizing the sum of the squared distances between  $\mathbf{x}_0$  and all of the points in the dataset. In other words, we would like to find the value of  $\mathbf{x}_0$  that minimizes the criterion function  $J_0$  shown in (6.21):

$$J_0(\mathbf{x}_0) = \sum_{k=1}^n \|\mathbf{x}_0 - \mathbf{x}_k\|^2 \quad (6.21)$$

It can be shown that the value of  $\mathbf{x}_0$  that minimizes  $J_0$  is the sample mean ( $1/N \sum \mathbf{x}_i$ ) of the dataset [39]. The sample mean has zero dimensionality and therefore does not give any information about the spread of the data, because it is a single point. To represent this information, the dataset would need to be projected onto a space with some dimensionality. To project the original dataset onto a one-dimensional space, we need to project it onto a line in the original space that runs through the sample mean. The data points in the new space can then be defined by  $\mathbf{x} = \mathbf{m} + a\mathbf{e}$ . Here,  $\mathbf{e}$  is the unit vector in the direction of the line and  $a$  is a scalar, which represents the distance from  $\mathbf{m}$  to  $\mathbf{x}$ . A second criterion function  $J_1$  can now be defined that calculates the sum of the squared distances between the points in the original dataset and the projected points on the line:

$$J_1(a_1, \dots, a_n, \mathbf{e}) = \sum_{k=1}^n \|(\mathbf{m} + a_k \mathbf{e}) - \mathbf{x}_k\|^2 \quad (6.22)$$

Taking into consideration that  $\|\mathbf{e}\| = 1$ , the value of  $a_k$  that minimizes  $J_1$  is found to be  $a_k = \mathbf{e}^t(\mathbf{x}_k - \mathbf{m})$ . To find the best direction  $\mathbf{e}$  for the line, this value of  $a_k$  is substituted back into (6.22) to get (6.23). Then  $J_1$  from (6.23) can be minimized with respect to  $\mathbf{e}$  to find the direction of the line. It turns out that the vector that minimizes

$J_1$  is one that satisfies the equation  $\mathbf{S}\mathbf{e} = \lambda\mathbf{e}$ , for some scalar value  $\lambda$ , where  $\mathbf{S}$  is the scatter matrix of the original dataset as defined in (6.24).

$$J_1(\mathbf{e}) = \sum_{k=1}^n a_k^2 - 2 \sum_{k=1}^n a_k^2 + \sum_{k=1}^n \|\mathbf{x}_k - \mathbf{m}\|^2 \quad (6.23)$$

$$\mathbf{S} = \sum_{k=1}^n (\mathbf{x}_k - \mathbf{m})(\mathbf{x}_k - \mathbf{m})^t \quad (6.24)$$

Because  $\mathbf{e}$  must satisfy  $\mathbf{S}\mathbf{e} = \lambda\mathbf{e}$ , it is easy to realize that  $\mathbf{e}$  must be an eigenvector of the scatter matrix  $\mathbf{S}$ . In addition to  $\mathbf{e}$  being an eigenvector of  $\mathbf{S}$ , Duda et al. [39] also showed that the eigenvector that yields the best representation of the original dataset is the one that corresponds to the largest eigenvalue. By projecting the data onto the eigenvectors of the scatter matrix that correspond to the  $d'$  highest eigenvalues, the original dataset can be projected down to a space of size  $d'$  dimensionality.

## 6.7 Correlation Structure

One method of seizure analysis is to consider the correlation over all of the recorded EEG channels. To do this, a correlation is defined over the given channels. To define the correlation matrix, a segment of the EEG signal is considered for a given window of a specified time. The EEG signal is then channel-wise normalized within this window. Given  $m$  channels, the correlation matrix  $\mathbf{C}$  is defined as in (6.25), where  $w_l$  specifies the length of the given window ( $w$ ) and  $EEG_i$  is the  $i$ th channel. The value of  $EEG_i$  has also been normalized to have zero mean and unit variance [6]. The  $C_{ij}$  term will yield a value of 0 when  $EEG_i$  and  $EEG_j$  are uncorrelated, a value of 1 when they are perfectly correlated, and a value of  $-1$  when they are anticorrelated. Note also that the correlation matrix is symmetrical since  $C_{ij} = C_{ji}$ . In addition,  $C_{ii} = 1$  for all values of  $i$  because any signal will be perfectly correlated with itself. It follows that the trace of the matrix ( $\sum C_{ii}$ ) will always equal the number of channels ( $m$ ).

$$C_{ij} = \frac{1}{w_l} \sum EEG_i(t) \cdot EEG_j(t) \quad (6.25)$$

To simplify the representation of the correlation matrix, the eigenvalues of the matrix are calculated. The eigenvalues reveal which dimensions of the original matrix have the highest correlation. When the eigenvalues ( $\lambda_1, \lambda_2, \dots, \lambda_m$ ) are sorted so that  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{\max}$ , they can then be used to produce a spectrum of the correlation matrix  $\mathbf{C}$  [40]. This spectrum is sorted by intensity of correlation. The spectrum is then used to track how the dynamics of all  $m$  EEG channels are affected when a seizure occurs.

## 6.8 Multidimensional Probability Evolution

Another nonlinear technique that has been used for seizure detection is based on a multidimensional probability evolution (MDPE) function. Using the probability density function, changes in the nature of the trajectory of the EEG signal, as it evolves, can be detected. To accomplish the task of detection, the technique tracks how often various parts of the state space are visited when the EEG is in the nonictal state. Using these statistics, anomalies in the dynamics of the system can then be detected, which usually implies the occurrence of a seizure. In one report, when MDPE was applied to test data, it was able to detect all of the seizures that occurred in the data [41]. However, there was no mention of the number of false positives, false negatives, or if the authors had tried to predict seizures at all.

## 6.9 Self-Organizing Map

The techniques just described are all based on particular mathematical transformations of the EEG signal. In contrast, a machine learning–based technique that has been used to detect seizures is the self-organizing map (SOM). The SOM is a particular kind of an artificial neural network that uses unsupervised learning to classify data; that is, it does not require training samples that are labeled with the class information (in the case of seizure detection, this would correspond to labeling the EEG signal as an ictal/interictal event); it is merely provided the data and the network learns on its own. Described informally, the SOM groups inputs that have “similar” attributes by assigning them to close by neurons in the network. This is achieved by incrementally rewarding the activation function of those artificial neurons in the network (and their neighbors) that favor a particular input data point. Competition arises because different input data points have to jockey for position on the network. One reported result transformed the EEG signal using a FFT, and subsequently used the FFT vector as input to a SOM. With the help of some additional stipulations on the amplitudes and frequencies, the SOM was able to detect 90% of the seizures with an average of 0.71 false positives per hour [42]. However, the report did not attempt to apply the technique to predicting seizures, which would most definitely have produced worse results.

## 6.10 Support Vector Machine

A more advanced machine learning technique that has been used for seizure detection is a support vector machine (SVM). As opposed to an SOM, an SVM is a reinforcement learning technique—it requires data that is labeled with the class information. A support vector machine is a classifier that partitions the feature space (or the kernel space in the case of a kernel SVM) into two classes using a hyperplane. Each sample is represented as a point in the feature space (or the kernel space, as the case may be) and is assigned a class depending on which side of the hyperplane it lies. The classifier that is yielded by the SVM learning algorithm is the optimal hyperplane that minimizes the expected risk of misclassifying unseen samples. Ker-

nel SVMs have been applied to EEG data after removing noise and other artifacts from the raw signals in the various channels. In one report, the author was able to detect 97% of the seizures using an online detection method that used a kernel SVM. Of the seizures that were detected, the author reported that he was able to predict 40% of the ictal events by an average of 48 seconds before the onset of the seizure [43].

## 6.11 Phase Correlation

Methods of measuring phase synchrony include methods based on spectral coherence. These methods incorporate both amplitude and phase information, detection of maximal values after filtering. For weakly coupled nonlinear equations, phases are locked, but the amplitudes vary chaotically and are mostly uncorrelated. To characterize the strength of synchronization, Tass [44] proposed two indices, one based on Shannon entropy and one based on conditional probability. This approach aims to quantify the degree of deviation of the relative phase distribution from a uniform phase distribution.

All of the techniques that have been described thus far approach the problem of detecting and predicting seizures from a traditional time-series prediction perspective. In all such cases, the EEG signal is viewed like any other signal that has predictive content embedded in it. The goal, therefore, is to transform the signal using various mathematical techniques so as to draw out this predictive content. The fact that an EEG signal is generated in a particular biological context, and is representative of a particular physical aspect of the system, does not play a significant role in these techniques.

## 6.12 Seizure Detection and Prediction

Seizure anticipation (or warning) can be classified into two broad categories: (1) *early seizure detection* in which the goal is to use EEG data to identify seizure onset, which typically occurs a few seconds in advance of the observed behavioral changes or during the period of early clinical manifestation of focal motor changes or loss of patient awareness, and (2) *seizure prediction* in which the aim is to detect preictal changes in the EEG signal that typically occur minutes to hours in advance of an impending epileptic seizure.

In seizure detection, since the aim of these algorithms is to causally identify an ictal state, the statistical robustness of early seizure detection algorithms is very high [45, 46]. The practical utility of these schemes in the development of an online seizure abatement strategy depends critically on the few seconds of time between the detection of an EEG seizure and its actual manifestation in patients in terms of behavioral changes. Recently Talathi et al. [47] conducted a review of a number of nonparametric early seizure detection algorithms to determine the critical role of the EEG acquisition methodology in improving the overall performance of these algorithms in terms of their ability to detect seizure onset early enough to provide a suitable time to react and intervene to abate seizures.

In seizure prediction, the effectiveness of seizure prediction techniques tends to be lower in terms of statistical robustness. This is because the time horizon of these methods ranges from minutes to hours in advance of an impending seizure and because the preictal state is not a well-defined state across multiple seizures and across different patients. Some studies have shown evidence of a preictal period that could be used to predict the onset of an epileptic seizure with high statistical robustness [13, 48]. However, many of these studies use a posteriori knowledge or do not use out-of-sample training [14]. This leads to a model that is “overfit” for the data being used. When this same model is applied to other data, the accuracy of the technique typically decreases dramatically.

A number of algorithms have been developed solely for seizure detection and not for seizure prediction. The goal in this case is to identify seizures from EEG signals offline. Technicians spend many hours going through days of recorded EEG activity in an effort to identify all seizures that occurred during the recording. A technique that could automate this screening process would save a great amount of time and money. Because the purpose is to identify every seizure, any part of the EEG data may be used. Particularly a causal estimation of algorithmic measures can be used to determine the time of seizure occurrence. Algorithms designed for this purpose typically have better statistical performance and can only be used as an offline tool to assist in the identification of EEG seizures in long records of EEG data.

### 6.13 Performance of Seizure Detection/Prediction Schemes

With so many seizure detection and prediction methods available, there needs to be a way to compare them so that the “best” method can be used. Many statistics that evaluate how well a method does are available. In seizure detection, the technique is supposed to discriminate EEG signals in the ictal (seizure) state from EEG signals in the interictal (nonseizure) state. In seizure prediction, the technique is supposed to discriminate EEG signals in the preictal (before the seizure) state from EEG signals in the interictal (nonseizure) state. The classification an algorithm gives to a particular segment of EEG for either seizure detection or prediction can be placed into one of four categories:

- *True positive (TP)*: A technique correctly classifies an ictal segment (preictal for prediction) of an EEG as being in the ictal state (preictal for prediction).
- *True negative (TN)*: A technique correctly classifies an interictal segment of an EEG as being in the interictal state.
- *False positive (FP)*: A technique incorrectly classifies an interictal segment of an EEG as being in the ictal state (preictal for prediction).
- *False negative (FN)*: A technique incorrectly classifies an ictal segment (preictal for prediction) of an EEG as being in the interictal state.

Next we discuss how these classifications can be used to create metrics for evaluating how well a seizure prediction/detection technique does. In addition, we also discuss the use of a posteriori information. A posteriori information is used by cer-



tain algorithms to improve their accuracy. However, in most cases, this information is not available when using the technique in an online manner so it cannot be generalized to online use.

### 6.13.1 Optimality Index

From these four totals (TP, TN, FP, FN) we can calculate two statistics that give a large amount of information regarding the success of a given technique. The first statistic is the sensitivity ( $S$ ), which is defined in (6.26). In detection this indicates the probability of detecting an existent seizure and is defined by the ratio of the number of detected seizures to the number of total seizures. In prediction this indicates the probability of predicting an existent seizure and is defined by the ratio of the number of predicted seizures to the number of total seizures.

$$S = \frac{TP}{TP + FN} \quad (6.26)$$

In addition to the sensitivity, the specificity ( $K$ ) is also used and is defined in (6.27). This indicates the probability of not incorrectly detecting/predicting a seizure and is defined by the ratio of the number of interictal segments correctly identified in comparison to the number of interictal segments.

$$K = \frac{TN}{TN + FP} \quad (6.27)$$

A third metric used to measure the quality of a given algorithm is the predictability. This indicates how far in advance of a seizure the seizure can be predicted or how long after the onset of the seizure it can be detected. In other words, the predictability ( $\Delta T$ ) is defined by  $\Delta T = T_a - T_e$  where  $T_a$  is the time at which the given algorithm detects the seizure and  $T_e$  is the time at which the onset of the seizure actually occurs according to the EEG.

Note that either of these metrics alone is not a sufficient measure of quality for a seizure detection/prediction technique. Consider a detection/prediction algorithm that always said the signal was in the ictal or preictal state, respectively. Such a method would produce a sensitivity of 1 and a specificity of 0. On the other hand, an algorithm that always said the signal was in the interictal state would produce a sensitivity of 0 and a specificity of 1. The ideal algorithm would produce a value of 1 for each. To accommodate this, Talathi et al. [47] defined the optimality index ( $O$ ), a single measure of goodness, which takes all three of these metrics into account. It is defined in (6.28), where  $D^*$  is the mean seizure duration of the seizures in the dataset:

$$O = \frac{S + K}{2} - \frac{\Delta T}{D^*} \quad (6.28)$$

### 6.13.2 Specificity Rate

The specificity rate is another metric used to assess the performance of a seizure prediction/detection algorithm [49]. It is calculated by taking the number of false predictions or detections divided by the length of time of the recorded data (FP/T). It gives an estimate of the number of times that the algorithm under consideration would produce a false prediction or detection in a unit time (usually an hour).

Morman et al. [49] also point out that the prediction horizon is important when considering the specificity rate of prediction algorithms. The prediction horizon is the amount of time before the seizure for which the given algorithm is trying to predict it. The reason is false positives are more costly as the prediction horizon increases. A false positive for an algorithm with a larger prediction horizon causes the patient to spend more time expecting a seizure that will not occur. This is in opposition to an algorithm with a smaller prediction horizon. Less time is spent expecting a seizure that will not occur when a false positive is given. To correct this, they suggest using a technique that reports the portion of time from the interictal period during which a patient is not in the state of falsely awaiting a seizure [49].

Another issue that should be considered when assessing a particular seizure detection/prediction technique is whether or not a posteriori information is used by the technique in question. A posteriori information is information that can be used to improve an algorithm's accuracy, but is specific to the dataset (EEG signal) at hand. When the algorithm is applied to other datasets where this information is not known, the accuracy of the algorithm can drop dramatically. In-sample optimization is one example of a posteriori information used in some algorithms [14, 49]. With in-sample optimization, the same EEG signal that is used to test the given technique is also used to train the technique. When training a given algorithm, certain parameters are adjusted in order to come up with a general method that can distinguish two classes. When training the technique, the algorithm is optimized to classify the training data. Therefore, when the same data that is used to test a technique is used to train the technique, the technique is optimized ("overfit") for the testing data. Although this produces promising results as far as accuracy, these results are not representative of what would be produced when the algorithm is applied to nontraining, that is, out-of-sample, data.

Another piece of a posteriori information that is used in some algorithms is optimal channel selection. When testing, other algorithms are given the channel of the EEG that produces the best results. It has been shown that out of the available EEG channels, not every channel provides information that can be used to predict or detect a seizure [47, 49]. Other channels provide information that would produce false positives. So when an optimal channel is provided to a given algorithm, the results produced from this technique again will be biased. Therefore, the algorithm does not usually generalize well to the online case when the optimal channel is not known.

## 6.14 Closed-Loop Seizure Prevention Systems

The majority of patients with epilepsy are treated with chronic medication that attempts to balance cortical inhibition and excitation to prevent a seizure from

occurring. However, anticonvulsant drugs only control seizures for about two-thirds of patients with epilepsy [50]. Electrical stimulation is an alternative treatment that has been used [51]. In most cases, open-loop stimulation is used. This type of treatment delivers electrical stimulus to the brain without any neurological feedback from the system. The stimulation is delivered on a preset schedule for predetermined lengths of time (Figure 6.6).

Electrically stimulating the brain on a preset schedule raises questions about the long-term effects of such a treatment. Constant stimulation of the neurons could cause long-term damage or totally alter the neuronal architecture. Because of this, recent research has been aimed at closed-loop and semi-closed-loop prevention systems. Both of these systems take neurological feedback into consideration when delivering the electrical stimulation. In semi-closed-loop prevention systems, the stimulus is supplied only when a seizure has been predicted or detected by some algorithm (Figure 6.7). The goal is to reduce the severity of or totally stop the oncoming seizure. In closed loop stimulation the neurological feedback is used to create an optimal stimulation pattern that is used to reduce seizure severity.

In general, an online seizure detection algorithm is used rather than a prediction algorithm. Although a technique that could predict a seizure beforehand would be ideal, in practice, prediction algorithms leave much to be desired as far as statistical accuracy goes when compared to seizure detection algorithms. As the prediction horizon increases, the correlation between channels tends to decrease. Therefore, the chance of accurately predicting a seizure decreases as well. However, the downside of using an online detection algorithm is that it does not always detect the seizure in enough time to give the closed-loop seizure prevention system sufficient warning to prevent the seizure from occurring.

Finally, factors concerning the collection of the EEG data also play a significant role in the success of seizure detection algorithms [47]. Parameters such as the location of EEG electrodes, the type of the electrode, and the sampling rate of the electrodes can play a vital role in the success of a given online detection algorithm. By increasing the sampling rate, the detection technique is supplied with more data points for a given time period. This gives the detector more chances to pick up on any patterns that would be indicative of a seizure

## 6.15 Conclusion

Epilepsy is a dynamic disease, characterized by numerous types of seizures and presentations. This has led to a rich set of electrographical records to analyze. To

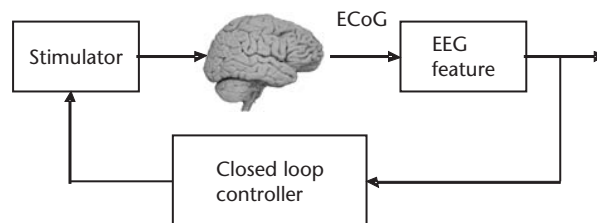
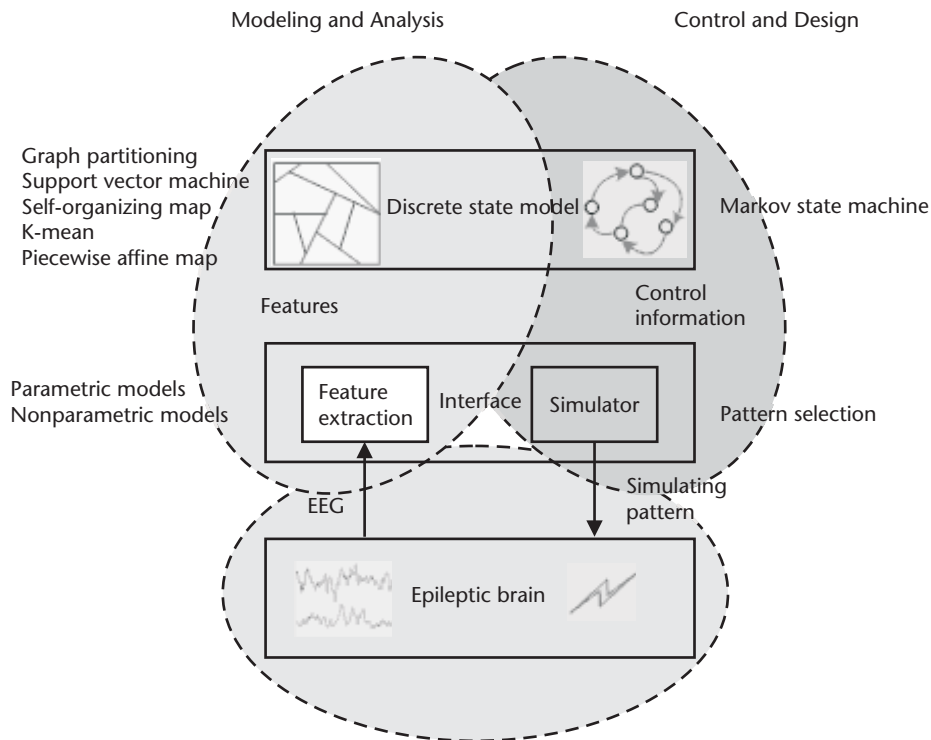


Figure 6.6 Schematic diagram for seizure control.



**Figure 6.7** A hybrid system that is composed of four parts of modeling phases: modeling, analysis, control, and design.

understand these signals, investigators have started to employ various signal processing techniques. Researchers have a wide assortment of both univariate and multivariate tools at their disposal. Even with these tools, the richness of the datasets has meant that these techniques have been met with limited success in predicting seizures. To date, there has been limited amount of research into comparing techniques on the same datasets. Oftentimes the initial success of a measure has been difficult to repeat because the first set of trials was the victim of overtraining. No measure has been able to reliably and repeatedly predict seizures with a high level of specificity and sensitivity.

While the line between seizure prediction, early detection, and detection can sometimes blur, it is important to note they do comprise three different questions. While unable to predict a seizure, many of these measures can detect a seizure. Seizures often present themselves as electrical storms in the brain, which are easily detectable, by eye, on an EEG trace. Seizure prediction seeks to tease out minute changes in the EEG signal. Thus far the tools that are able to detect one of these minor fluctuations often fall short when trying to replicate their success in slightly altered conditions. Coupled with the proper type of intervention (e.g., chemical stimulation or directed pharmacological delivery) early detection algorithms could usher in a new era of epilepsy treatment. The techniques presented in this chapter need to be continually studied and refined. They should be tested on standard datasets in order for their results to be accurately compared. Additionally, they need

to be tested on out-of-sample datasets to determine their effectiveness in a clinical setting.

## References

- [1] Blume, W., et al., "Glossary of Descriptive Terminology for Ictal Semiology: Report of the ILAE Task Force on Classification and Terminology," *Epilepsia*, Vol. 42, No. 9, 2001, pp. 1212–1218.
- [2] Fisher, R., et al., "Epileptic Seizures and Epilepsy: Definitions Proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE)," *Epilepsia*, Vol. 46, No. 4, 2005, pp. 470–472.
- [3] World Health Organization, *Epilepsy: Aetiology, Epidemiology, and Prognosis*, 2001.
- [4] Meisler, M. H., and J. A. Kearney, "Sodium Channel Mutations in Epilepsy and Other Neurological Disorders," *J. Clin. Invest.*, Vol. 115, No. 8, 2005, pp. 2010–2017.
- [5] Hauser, W. A., and D. C. Hesdorffer, *Causes and Consequences*, New York: Demos Medical Publishing, 1990.
- [6] Jacobs, M. P., et al., "Future Directions for Epilepsy Research," *Neurology*, Vol. 257, 2001, pp. 1536–1542.
- [7] Theodore, W. H., and R. Fisher, "Brain Stimulation for Epilepsy," *Lancet Neurol.*, Vol. 3, No. 6, 2004, p. 332.
- [8] Annegers, J. F., "United States Perspective on Definitions and Classifications," *Epilepsia*, Vol. 38, Suppl. 11, 1997, S9–S12.
- [9] Stables, J. P., et al., "Models for Epilepsy and Epileptogenesis," *Report from the NIH Workshop*, Bethesda, MD, *Epilepsia*, Vol. 43, No. 11, 2002, pp. 1410–1420.
- [10] Elger, C. E., "ARTICLE TITLE??" *Curr. Opin. Neurol.*, Vol. 14, No. 2, April 2001, pp. 185–186.
- [11] Lopes da Silva, F. H., "EEG Analysis: Theory and Practice; Computer-Assisted EEG Diagnosis: Pattern Recognition Techniques," in *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, E. Niedermeyer and F. H. Lopes da Silva, (eds.), pp. 871–897. Baltimore, MD: Williams & Wilkins, 1987.
- [12] Iasemidis, L. D., "On the Dynamics of the Human Brain in Temporal Lobe Epilepsy," University of Michigan, Ann Arbor, 1991.
- [13] Le Van Quyen, M., et al., "Anticipating Epileptic Seizure in Real Time by a Nonlinear Analysis of Similarity Between EEG Recordings," *Neuroreport*, Vol. 10, 1999, pp. 2149–2155.
- [14] Mormann, F., et al., "On the Predictability of Epileptic Seizures," *Clin. Neurophysiol.*, Vol. 116, No. 3, 2005, pp. 569–587.
- [15] Oppenheim, A. V., "Signal Processing in the Context of Chaotic Signals," *IEEE Int. Conf. ASSP*, 1992.
- [16] Blanco, S., "Applying Time-Frequency Analysis to Seizure EEG Activity: A Method to Help to Identify the Source of Epileptic Seizures," *IEEE Eng. Med. Biol. Mag.*, Vol. 16, 1997, pp. 64–71.
- [17] Walnut, D. F., *An Introduction to Wavelet Analysis*, Boston, MA: Birkhauser, 2002.
- [18] Shannon, C. E., "A Mathematical Theory of Communication," *Bell Syst. Tech. J.*, Vol. 27, 1948, pp. 379–423.
- [19] Osorio, I., "Real-Time Automated Detection and Quantitative Analysis of Seizures and Short-Term Prediction of Clinical Onset," *Epilepsia*, Vol. 39, No. 6, 1998, pp. 615–627.
- [20] Wilks, S. S., *Mathematical Statistics*, New York: Wiley, 1962.
- [21] Liu, H., "Epileptic Seizure Detection from ECoG Using Recurrence Time Statistics," *Proceedings of the 26th Annual International Conference of the IEEE EMBS*, 2004, pp. 29–32.

- [22] Iasemidis, L. D., et al., "Adaptive Epileptic Seizure Prediction System," *IEEE Trans. on Biomed. Eng.*, Vol. 50, 2003, pp. 616–627.
- [23] Iasemidis, L. D., et al., "Dynamic Resetting of the Human Brain at Epileptic Seizures: Application of Nonlinear Dynamics and Global Optimization Techniques," *IEEE Trans. on Biomed. Eng.*, Vol. 51, 2004, pp. 493–506.
- [24] Iasemidis, L. D., et al., "Quadratic Binary Programming and Dynamic Systems Approach to Determine the Predictability of Epileptic Seizures," *J. Combinat. Optim.*, Vol. 5, 2001, pp. 9–26.
- [25] Iasemidis, L. D., and J. C. Sackellares, "The Temporal Evolution of the Largest Lyapunov Exponent on the Human Epileptic Cortex," in *Measuring Chaos in the Human Brain*, D. W. Duke and W. S. Pritchard, (eds.), Singapore: World Scientific, 1991, pp. 49–82.
- [26] Iasemidis, L. D., and J. C. Sackellares, "Long Time Scale Temporo-Spatial Patterns of Entrainment of Preictal Electroencephalographic Data in Human Temporal Lobe Epilepsy," *Epilepsia*, Vol. 31, No. 5, 1990, p. 621.
- [27] Iasemidis, L. D., "Time Dependencies in the Occurrences of Epileptic Seizures," *Epilepsy Res.*, Vol. 17, No. 1, 1994, pp. 81–94.
- [28] Pardalos, P. M., "Seizure Warning Algorithm Based on Optimization and Nonlinear Dynamics," *Mathemat. Program.*, Vol. 101, No. 2, 2004, pp. 365–385.
- [29] Sackellares, J. C., "Epileptic Seizures as Neural Resetting Mechanisms," *Epilepsia*, Vol. 38, Suppl. 3, 1997, p. 189.
- [30] Degan, H., A. Holden, and L. F. Olsen, *Chaos in Biological Systems*, New York: Plenum, 1987.
- [31] Marcus, M., S. M. Aller, and G. Nicolis, *From Chemical to Biological Organization*, New York: Springer-Verlag, 1988.
- [32] Sackellares J. C., et al., "Epilepsy—When Chaos Fails," in *Chaos in Brain?*, K. Lehnertz, et al., (eds.), Singapore: World Scientific, 2000, pp. 112–133.
- [33] Takens, F., *Detecting Strange Attractors in Turbulence of Dynamical Systems and Turbulence*, New York: Springer-Verlag, 1981.
- [34] Abarbanel, H. D. I., *Analysis of Observed Chaotic Data*, New York: Springer-Verlag, 1996.
- [35] Milton, J., and P. Jung, *Epilepsy as a Dynamic Disease*, New York: Springer, 2003.
- [36] Quiroga, R. Q., T. Kreuz, and P. Grassberger, "Event Synchronization: A Simple and Fast Method to Measure Synchronicity and Time Delay Patterns," *Phys. Rev. E*, Vol. 66, No. 041904, 2002.
- [37] Rosenblum, M. G., A. S. Pikovsky, and J. Kurths, "From Phase to Lag Synchronization in Coupled Chaotic Oscillators," *Phys. Rev. Lett.*, Vol. 78, No. 22, 1997, pp. 4193–4196.
- [38] Mormann, F., et al., "Automated Detection of a Preseizure State Based on a Decrease in Synchronization in Intracranial Electroencephalogram Recordings from Epilepsy Patients," *Phys. Rev. E*, Vol. 67, No. 2, 2003.
- [39] Duda, R. O., P. E. Hart, and D. G. Stork, *Pattern Classification*, New York: Wiley-Interscience, 1997, pp. 114–117.
- [40] Schindler, K., et al., "Assessing Seizure Dynamics by Analysing the Correlation Structure of Multichannel Intracranial EEG," *Brain*, Vol. 130, No. 1, 2007, p. 65.
- [41] McSharry, P. E., "Linear and Non-Linear Methods for Automatic Seizure Detection in Scalp Electroencephalogram Recordings," *Med. Biol. Eng. Comput.*, Vol. 40, No. 4, 2002, pp. 447–461.
- [42] Gabor, A. J., "Automated Seizure Detection Using a Self-Organizing Neural Network," *Electroencephalog. Clin. Neurophysiol.*, Vol. 99, 1996, pp. 257–266.
- [43] Gardner, A. B., "A Novelty Detection Approach to Seizure Analysis from Intracranial EEG," Georgia Institute of Technology, Atlanta, 2004.
- [44] Tass, P., *Phase Resetting in Medicine and Biology: Stochastic Modeling and Data Analysis*, New York: Springer-Verlag, 1999.



- [45] Osorio, I., et al., "Automated Seizure Abatement in Humans Using Electrical Stimulation," *Ann. Neurol.*, Vol. 57, 2005, pp. 258–268.
- [46] Saab, M. E., and J. Gotman, "A System to Detect the Onset of Epileptic Seizures in Scalp EEG," *Clin. Neurophysiol.*, Vol. 116, 2005, pp. 427–442.
- [47] Talathi, S. S., et al., "Non-Parametric Early Seizure Detection in an Animal Model of Temporal Lobe Epilepsy," *J. Neural Eng.*, Vol. 5, 2008, pp. 1–14.
- [48] Martinerie, J., C. Adam, and M. Le Van Quyen, "Epileptic Seizures Can Be Anticipated by Non-Linear Analysis," *Nature Med.*, Vol. 4, 1998, pp. 1173–1176.
- [49] Mormann, F., C. Elger, and K. Lehnertz, "Seizure Anticipation: From Algorithms to Clinical Practice," *Curr. Opin. Neurol.*, Vol. 19, 2006, pp. 187–193.
- [50] Li, Y., and D. J. Mogul, "Electrical Control of Epileptic Seizures," *J. Clin. Neurophysiol.*, Vol. 24, No. 2, 2007, p. 197.
- [51] Colpan, M. E., et al., "Proportional Feedback Stimulation for Seizure Control in Rats," *Epilepsia*, Vol. 48, No. 8, 2007, pp. 1594–1603.

### Selected Bibliography

Mane, R., D. Rand, and L. S. Young, *Dynamical Systems and Turbulence*, New York: Springer-Verlag, 1981.